

*Pacific
Journal of
Mathematics*

Volume 342 No. 2

June 2026

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

EDITORS

Don Blasius (Managing Editor)
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
blasius@math.ucla.edu

Matthias Aschenbrenner
Fakultät für Mathematik
Universität Wien
Vienna, Austria
matthias.aschenbrenner@univie.ac.at

Vyjayanthi Chari
Department of Mathematics
University of California
Riverside, CA 92521-0135
chari@math.ucr.edu

Atsushi Ichino
Department of Mathematics
Kyoto University
Kyoto 606-8502, Japan
atsushi.ichino@gmail.com

Kefeng Liu
School of Sciences
Chongqing University of Technology
Chongqing 400054, China
liu@math.ucla.edu

Sucharit Sarkar
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
sucharit@math.ucla.edu

Dimitri Shlyakhtenko
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
shlyakht@ipam.ucla.edu

Ruixiang Zhang
Department of Mathematics
University of California
Berkeley, CA 94720-3840
ruixiang@berkeley.edu

PRODUCTION

Silvio Levy, Scientific Editor, production@msp.org


See inside back cover or msp.org/pjm for submission instructions.

The subscription price for 2026 is US \$710/year for the electronic version, and \$965/year for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004.

PJM peer review and production are managed by EditFLOW[®] from Mathematical Sciences Publishers.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2026 Mathematical Sciences Publishers

ON MULTILINEAR MAXIMAL OPERATORS ALONG HOMOGENEOUS CURVES

LARS BECKER AND BEN KRAUSE

Suppose that

$$\vec{\gamma}(t) := (\gamma_1(t), \dots, \gamma_n(t)) = (a_1 t^{d_1}, \dots, a_n t^{d_n}), \quad 1 \leq d_1 < \dots < d_n \in \mathbb{Z}, \quad a_i \neq 0$$

is a homogeneous polynomial curve. We prove that whenever $p_1, \dots, p_n > 1$ and $1/p = \sum_{j=1}^n 1/p_j \leq 1$, there exists an absolute constant $0 < C = C_{p_1, \dots, p_n} < \infty$ such that

$$\left\| \sup_{r>0} \frac{1}{r} \int_0^r \prod_{i=1}^n |f_i(x - \gamma_i(t))| dt \right\|_{L^p(\mathbb{R})} \leq C \cdot \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}.$$

Our main tool is a smoothing estimate, adapted from work of Kosz, Mirek, Peluse, Wan, and Wright.

1. Introduction

The study of multilinear maximal functions dates back to celebrated work of Lacey [11], who proved the following theorem.

Theorem 1.1. *Suppose that $p_1, p_2 > 1$, and that $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p} < \frac{3}{2}$. Then there exists an absolute constant $0 < C_{p_1, p_2} < \infty$ such that*

$$\begin{aligned} \|B^{\vec{\gamma}}(f_1, f_2)\|_{L^p(\mathbb{R})} &:= \left\| \sup_{r>0} \frac{1}{r} \int_0^r |f(x - \gamma_1(t))g(x - \gamma_2(t))| dt \right\|_{L^p(\mathbb{R})} \\ &\leq C_{p_1, p_2} \|f_1\|_{L^{p_1}(\mathbb{R})} \|f_2\|_{L^{p_2}(\mathbb{R})}, \end{aligned}$$

whenever $\gamma_i(t) = a_i t$ and $a_1 \neq a_2$ are nonzero.

The key property of this operator is its modulation invariance, which necessitated an approach using time-frequency analysis, building off ideas of Lacey and Thiele in their work on the bilinear Hilbert transform [12; 13]; this method was later adapted to handle multilinear extensions, see [2].

On the other hand, when the modulation invariance embedded in $\vec{\gamma}$ is eliminated, different techniques can be used. This was first explored in the singular integral context in [14; 16], with subsequent work of [15] establishing the following; see also [5].

MSC2020: 42A99.

Keywords: multilinear maximal operators, Littlewood–Paley theory, smoothing inequalities.

Theorem 1.2. *Suppose that $\gamma_1(t) = t$, $\gamma_2(t) = P(t)$, where $P(t)$ is a polynomial of degree d which vanishes to degree ≥ 2 at the origin. Then whenever $p_1, p_2 > 1$ and $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p} < \frac{d}{d-1}$, there exists an absolute constant $0 < C_{p_1, p_2} < \infty$ such that*

$$\|B^{\vec{\gamma}}(f_1, f_2)\|_{L^p(\mathbb{R})} \leq C_{p_1, p_2} \|f_1\|_{L^{p_1}(\mathbb{R})} \|f_2\|_{L^{p_2}(\mathbb{R})}.$$

The key ingredient in establishing Theorem 1.2 was a *Sobolev estimate*, a representative case of which is stated below.

Proposition 1.3 (special case). *Suppose that $\gamma_1(t) = t$, $\gamma_2(t) = t^2$. There exist absolute constants $0 < c < C < \infty$ such that if \hat{f}_i vanishes outside $\{|\xi| \leq 2^{l+ik}\}$ for some i , then*

$$\left\| \int_0^1 f_1(x - \gamma_1(2^{-k}t)) f_2(x - \gamma_2(2^{-k}t)) dt \right\|_{L^1(\mathbb{R})} \leq C 2^{-cl} \|f_1\|_{L^2(\mathbb{R})} \|f_2\|_{L^2(\mathbb{R})}.$$

In other words, the only obstruction to the estimate

$$\left\| \int_0^1 f_1(x - \gamma_1(2^{-k}t)) f_2(x - \gamma_2(2^{-k}t)) dt \right\|_{L^1(\mathbb{R})} \ll \|f_1\|_{L^2(\mathbb{R})} \|f_2\|_{L^2(\mathbb{R})},$$

where \ll informally stands for “much smaller than”, arises from zero-frequency considerations; note that the modulation invariance from Theorem 1.1 precludes such an argument. Aside from their utility in studying the operators $\{B^{\vec{\gamma}}\}$, Sobolev estimates have found a wide use in problems in Euclidean Ramsey theory, dating back to the work of Bourgain [1], with more recent contributions found in [3; 4; 6; 9; 10] among others. The current state of the art for operators of the form $\{B^{\vec{\gamma}}\}$ is essentially due to Hu and Lie [7], who addressed trilinear formulations

$$\vec{\gamma} = (P_1(t), P_2(t), P_3(t))$$

provided P_i are distinct degree polynomials which vanish at different rates at 0; see Observation 1.2(i) of [7] and [5, Remark 2]. their key input was a trilinear analogue of Proposition 1.3. We state below a slightly stronger version of their estimate [5, Theorem 3.1]. Their estimate has an additional dependence on k , which we drop here as justified by Proposition 1.6 below.

Proposition 1.4 (special case). *Suppose that $\gamma_i(t) = t^i$, $1 \leq i \leq 3$. There exist absolute constants $0 < c < C < \infty$ such that if \hat{f}_i vanishes outside $\{|\xi| \leq 2^{l+ik}\}$ for some i , then*

$$\left\| \int_0^1 \prod_{i=1}^3 f_i(x - \gamma_i(2^{-k}t)) dt \right\|_{L^1(\mathbb{R})} \leq C 2^{-cl} \prod_{i=1}^3 \|f_i\|_{L^3(\mathbb{R})}.$$

The goal of this paper is to address multilinear analogues of $B^{\vec{\gamma}}$ under the simplifying assumption that our curves are homogeneous polynomials.

Specifically, we will be concerned with polynomial curves

$$\vec{\gamma}(t) := (\gamma_1(t), \dots, \gamma_n(t)) = (a_1 t^{d_1}, \dots, a_n t^{d_n}), \quad 1 \leq d_1 < \dots < d_n \in \mathbb{Z}, \quad a_i \neq 0.$$

Our main result is this:

Theorem 1.5. *Suppose that $p_1, \dots, p_n > 1$, and that*

$$\frac{1}{p} = \frac{1}{p_1} + \dots + \frac{1}{p_n} \leq 1.$$

Then there exists an absolute constant $0 < C_{p_1, \dots, p_n; \vec{\gamma}} < \infty$ such that

$$\left\| \sup_{r>0} \frac{1}{r} \int_0^r \prod_{i=1}^n |f_i(x - \gamma_i(t))| dt \right\|_{L^p(\mathbb{R})} \leq C_{p_1, \dots, p_n; \vec{\gamma}} \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}.$$

As might be expected, the key input in proving Theorem 1.5 is the following multilinear Sobolev estimate:

Proposition 1.6. *Suppose that $\gamma_i(t) = a_i t^{d_i}$, $1 \leq i \leq n$ with integer exponents $1 \leq d_1 < d_2 < \dots < d_n \in \mathbb{Z}$, $a_i \neq 0$. There exist absolute constants $0 < c < C < \infty$ such that if \hat{f}_i vanishes outside $\{|\xi| \leq 2^{l+d_i k}\}$ for some i , then*

$$\left\| \int_0^1 \prod_{i=1}^n f_i(x - \gamma_i(2^{-k}t)) dt \right\|_{L^1(\mathbb{R})} \leq C 2^{-cl} \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}.$$

To establish Proposition 1.6, we adapt a recent result of Kosz, Mirek, Peluse, Wan, and Wright [8]; with Proposition 1.6 in hand, Theorem 1.5 readily presents. While it is reasonable to expect that an analogue of Proposition 1.6, and thus Theorem 1.5, should hold for more general distinct-degree polynomial curves γ_i which vanish to distinct degrees at the origin, the argument to deduce Proposition 1.6 from the results of [8] crucially relies on homogeneity, and does not readily adapt to the more general setting.

Notation. Throughout, we let φ denote various mean-one Schwartz functions, normalized in some sufficiently large seminorm. The precise choice of φ might differ from line to line. Similarly, we use ψ to denote a similar function, but with

$$(1.7) \quad \mathbf{1}_{|\xi| \approx 1} \leq \hat{\psi} \leq \mathbf{1}_{|\xi| \approx 1}$$

and such that

$$(1.8) \quad \sum_l \psi(\xi/2^l) = \mathbf{1}_{\xi \neq 0}.$$

We use the notation

$$\phi_k(x) := 2^k \phi(2^k x)$$

to denote L^1 -normalized dilations, and let

$$(1.9) \quad B_k^{\vec{\gamma}}(f_1, \dots, f_n)(x) := B_k(f_1, \dots, f_n)(x) := \int_0^1 \prod_{i=1}^n f_i(x - \gamma_i(2^{-k}t)) dt.$$

Below, we will regard $0 \neq a_1, \dots, a_n = O(1)$ as arbitrary but fixed, and will abbreviate

$$(1.10) \quad D := D(\vec{\gamma}) := d_1 + \dots + d_n.$$

Asymptotic notation. We will make use of the modified Vinogradov notation. We use $X \lesssim Y$, or $Y \gtrsim X$, to denote the estimate $X \leq CY$ for an absolute constant C . We use $X \approx Y$ as shorthand for $Y \lesssim X \lesssim Y$. We also make use of big-O notation: we let $O(Y)$ denote a quantity that is $\lesssim Y$. We let $f(t) := o_{t \rightarrow a}(X(t))$ denote a quantity such that $|f(t)|/X(t) \rightarrow 0$ as $t \rightarrow a$.

If we need C to depend on a parameter, we shall indicate this by subscripts, thus for instance $X \lesssim_p Y$ denotes the estimate $X \leq C_p Y$ for some C_p depending on p . We analogously define $O_p(Y)$.

2. Sobolev estimates

The main goal of this section is to prove the following Sobolev estimate.

Proposition 2.1. *There exists an absolute $c > 0$ such that the following single scale estimate holds whenever $s_i \geq 0$:*

$$\|B_k(\psi_{kd_1+s_1} * f_1, \dots, \psi_{kd_n+s_n} * f_n)\|_{L^1(\mathbb{R})} \lesssim 2^{-cs} \prod_{i=1}^n \|f_i\|_{L^n(\mathbb{R})},$$

where $s := \max\{s_i\}$.

The proof of Proposition 1.3 will be accomplished via a projection argument, anchored by [8, Theorem 6.1] in the case where $\gamma_i(t) = a_i t^{d_i}$, and $\mathbb{K} = \mathbb{R}$, which dictates that the operator

$$A_{-k}(F_1, \dots, F_n)(x_1, \dots, x_n) := \frac{1}{2^k} \int_0^{2^k} \prod_{i=1}^n F_i(x_1, \dots, x_i - a_i t^{d_i}, \dots, x_n) dt$$

satisfies nontrivial norm estimates whenever some \hat{F}_i vanishes in $|\xi_i| \leq 2^{-kd_i} \delta^{-1}$.

Lemma 2.2 (special case of Theorem 6.1 of [8]). *In the above setting, suppose that some \hat{F}_i vanishes in $|\xi_i| \leq 2^{-kd_i} \delta^{-1}$. Then there exists some absolute $c > 0$ such that*

$$(2.3) \quad \|A_{-k}(F_1, \dots, F_n)\|_{L^1(\mathbb{R}^n)} \lesssim_n (\delta^c + 2^{-kc}) \prod_{i=1}^n \|F_i\|_{L^n(\mathbb{R}^n)},$$

provided $k \geq 0$.

We now remove the dependence on k on the right side of (2.3), and address the case where $k \leq 0$. Specifically, we prove that for all k

$$(2.4) \quad \|A_{-k}(F_1, \dots, F_n)\|_{L^1(\mathbb{R}^n)} \lesssim_n \delta^c \prod_{i=1}^n \|F_i\|_{L^n(\mathbb{R}^n)}$$

whenever some \hat{F}_i vanishes in $|\xi_i| \leq 2^{-kd_i}\delta^{-1}$; for concreteness, suppose that this index is j .

To do so, introduce the operator

$$D_\lambda F(x_1, \dots, x_n) := F(\lambda^{d_1}x_1, \dots, \lambda^{d_n}x_n)$$

and choose $2^{k_0} \gg \delta^{-1}$. If we define

$$G_i(x_1, \dots, x_n) := D_{2^{k_0}} F_i(x_1, \dots, x_n),$$

then \hat{G}_i vanishes on $|\xi_j| \leq 2^{-k_0 d_j} \delta^{-1}$, so

$$\|A_{-k_0}(G_1, \dots, G_n)\|_{L^1(\mathbb{R}^n)} \lesssim \delta^c \prod_{i=1}^n \|G_i\|_{L^n(\mathbb{R}^n)}.$$

But now

$$\begin{aligned} & D_{2^{k_0-k}}(A_{-k_0}(G_1, \dots, G_n))(x_1, \dots, x_n) \\ &= 2^{-k_0} \int_0^{2^{k_0}} \prod_{i=1}^n G_i(2^{d_1(k_0-k)}x_1, \dots, 2^{d_i(k_0-k)}x_i + a_i t^{d_i}, \dots, 2^{d_n(k_0-k)}x_n) dt \\ &= 2^{-k} \int_0^{2^k} \prod_{i=1}^n G_i(2^{d_1(k_0-k)}x_1, \dots, 2^{d_i(k_0-k)}x_i + a_i 2^{d_i(k_0-k)}t^{d_i}, \dots, 2^{d_n(k_0-k)}x_n) dt \\ &= A_{-k}(D_{2^{k_0-k}}G_1, \dots, D_{2^{k_0-k}}G_n)(x) \\ &= A_{-k}(F_1, \dots, F_n)(x), \end{aligned}$$

so the result follows from changing variables.

Proof of Proposition 2.1. Set $g_i := \psi_{kd_i+s_i} * f_i$, let $\vec{1} := n^{-1/2}(1, \dots, 1) \in \mathbb{R}^n$, and let $\varphi : \mathbb{R}^n \rightarrow \mathbb{C}$ be a bump function with compactly supported Fourier transform that is constant along $\mathbb{R}\vec{1}$. For each

$$2^{-100D|k|-100s} \gg \epsilon > 0$$

sufficiently small (see (1.10)), define

$$F_i(x_1, \dots, x_n) := \epsilon^{1-1/n} \varphi(\epsilon x) g_i(x \cdot \vec{1})$$

so that $\|F_i\|_{L^n(\mathbb{R}^n)} \approx \|g_i\|_{L^n(\mathbb{R})}$ and \hat{F}_i is supported in an $O(\epsilon)$ neighborhood of

$$\{\xi \vec{1} : \hat{g}_i(\xi) \neq 0\};$$

in particular, for some j , \hat{F}_j vanishes when $|\xi| \lesssim 2^{kd_j+s}$. So,

$$\|A_k(F_1, \dots, F_n)\|_{L^1(\mathbb{R}^n)} \lesssim 2^{-cs} \prod_{i=1}^n \|g_i\|_{L^n(\mathbb{R})}.$$

On the other hand,

$$\begin{aligned} & \|A_k(F_1, \dots, F_n)(x_1, \dots, x_n)\|_{L^1(\mathbb{R}^n)} \\ &= \left\| 2^k \int_0^{2^{-k}} \left(\prod_{i=1}^n g_i(x \cdot \vec{1} - a_i t^{d_i}) \right) \cdot \left(\epsilon^{n-1} \prod_{i=1}^n \varphi(\epsilon x - \epsilon a_i t^{d_i} \vec{e}_i) \right) dt \right\|_{L^1(\mathbb{R}^n)} \\ &= \int_{(\mathbb{R}\vec{1})^\perp} \int_{\mathbb{R}} \left| 2^k \int_0^{2^{-k}} \left(\prod_{i=1}^n g_i(y - a_i t^{d_i}) \right) \left(\epsilon^{n-1} \prod_{i=1}^n \varphi(\epsilon z - \epsilon a_i t^{d_i} \vec{e}_i) \right) dt \right| dy dz \end{aligned}$$

using a change of variables (so in particular, z_1, \dots, z_{n-1} form an orthogonal basis for $(\mathbb{R}\vec{1})^\perp$); above, \vec{e}_i is the i -th coordinate vector.

By Taylor expansion,

$$\begin{aligned} & \int_{(\mathbb{R}\vec{1})^\perp} \int_{\mathbb{R}} \left| 2^k \int_0^{2^{-k}} \left(\prod_{i=1}^n g_i(y - a_i t^{d_i}) \right) \left(\epsilon^{n-1} \prod_{i=1}^n \varphi(\epsilon z - \epsilon a_i t^{d_i} \vec{e}_i) \right) dt \right| dy dz \\ &= \|\varphi^n\|_{L^1((\mathbb{R}\vec{1})^\perp)} \cdot \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \left| 2^k \int_0^{2^{-k}} \prod_{i=1}^n g_i(y - a_i t^{d_i}) dt \right| dy \\ &\quad + O(2^{O(k)} \epsilon) \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} 2^k \int_0^{2^{-k}} \prod_{i=1}^n |g_i(y - a_i t^{d_i})| dt dy \right) \epsilon^{n-1} (1 + \epsilon|z|)^{-100} dz \\ &= \|\varphi^n\|_{L^1((\mathbb{R}\vec{1})^\perp)} \|B_k(g_1, \dots, g_n)(x)\|_{L^1(\mathbb{R})} + O_k\left(\epsilon \prod_{i=1}^n \|g_i\|_{L^n(\mathbb{R})}\right). \end{aligned}$$

The result follows by sending $\epsilon \downarrow 0$. □

The following proposition will be used to complement Proposition 1.3 via interpolation; before we can prove it, we state a lemma.

Proposition 2.5. *Suppose $n \geq 2$, $p_1, \dots, p_n > 1$ and $\frac{1}{p} = \sum_i \frac{1}{p_i} \leq 1$. Then*

$$\left\| \sup_k |B_k(\psi_{kd_1+s_1} * f, \dots, \psi_{kd_n+s_n} * f)| \right\|_{L^p(\mathbb{R})} \lesssim s^n \prod_{i=1}^n \|f\|_{L^{p_i}(\mathbb{R})},$$

where $s := \max\{s_i\}$.

Lemma 2.6. *Suppose that φ is a Schwartz function. For $t \in \mathbb{R}$, consider the maximal function*

$$M^t f(x) := \sup_j |\varphi_j * f|(x - 2^{-j}t).$$

There exists an absolute constant $0 < C < \infty$ (independent of t) such that

$$\|M^t f\|_{L^{1,\infty}(\mathbb{R})} \leq C \log(2 + |t|) \|f\|_{L^1(\mathbb{R})}.$$

Proof. Since the maximal function is trivially bounded on $L^\infty(\mathbb{R})$ we may use vector-valued Calderón–Zygmund theory; see [17, §1], for example. In particular,

it suffices to show that

$$\int_{|x| \geq 10|y|} \sum_j |\varphi_j(x - 2^{-j}t - y) - \varphi_j(x - 2^{-j}t)| \lesssim \log |t|.$$

By dilation invariance we can normalize $|y| \approx 1$. When $2^j \leq 100^{-1}$ we use the mean-value theorem; when $100^{-1} \leq 2^j \lesssim |t|$ we use a single-scale estimate; and when $2^j \gg |t|$ we use the decay of φ . \square

Proof of Proposition 2.5. We have the pointwise bound

$$\begin{aligned} & |B_k(\psi_{kd_1+s_1} * f_1, \dots, \psi_{kd_n+s_n} * f_n)(x)| \\ &= \left| \int_0^1 \left(\int_{\mathbb{R}^n} \prod_{i=1}^n \psi_{kd_i+s_i}(x - a_i 2^{-kd_i} t^{d_i} - u_i) f_i(u_i) du \right) dt \right| \\ &\leq \int_0^1 \prod_{i=1}^n M^{2^{s_i} a_i t^{d_i}} f_i(x) dt, \end{aligned}$$

where M^t is the maximal function from Lemma 2.6 with φ chosen to be ψ . So

$$\begin{aligned} \left\| \sup_k |B_k(\psi_{kd_1+s_1} * f_1, \dots, \psi_{kd_n+s_n} * f_n)| \right\|_{L^p(\mathbb{R})} &\lesssim \int_0^1 \left\| \prod_{i=1}^n M^{2^{s_i} a_i t^{d_i}} f_i \right\|_{L^p(\mathbb{R})} dt \\ &\lesssim s^n \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}, \end{aligned}$$

as desired. \square

3. The proof of Theorem 1.5

We will prove that

$$(3.1) \quad \left\| \sup_k |B_k(f_1, \dots, f_n)| \right\|_{L^p(\mathbb{R})} \lesssim \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}.$$

Theorem 1.5 follows from (3.1), because averaging operators are positive and therefore the lacunary supremum in (3.1) dominates the full supremum in Theorem 1.5 up to a factor of two.

The proof is by induction. Thus, we will assume that for all $m < n$

$$\left\| \sup_k |B_k(f_1, \dots, f_m)| \right\|_{L^p(\mathbb{R})} \lesssim \prod_{i=1}^m \|f_i\|_{L^{p_i}(\mathbb{R})},$$

whenever $\vec{\gamma}$ is a homogeneous polynomial curve, $p_1, \dots, p_m > 1$, and $\frac{1}{p_1} + \dots + \frac{1}{p_m} = \frac{1}{p} \leq 1$, with the $n = 1$ case following from Hardy–Littlewood and convexity.

By Taylor expansion, for each k , we have the decomposition

$$(3.2) \quad \begin{aligned} & B_k(f_1, \dots, f_n) \\ &= \sum_{j \geq 0} \frac{1}{j!} ((\partial^j \varphi)_{kd_n} * f_n) B_{k, \neq n}^{(j)}(f_1, \dots, f_{n-1}) + B_k(f_1, \dots, f_{n-1}, f_n - \varphi_{kd_n} * f_n), \end{aligned}$$

where

$$B_{k,\neq n}^{(j)}(f_1, \dots, f_{n-1})(x) := 2^k \int_0^{2^{-k}n-1} \prod_{i=1}^{n-1} f_i(x - a_i t^{d_i}) (-a_n 2^{kd_n} t^{d_n})^j dt,$$

and $\partial^j \varphi$ satisfies all of the same Schwartz normalizations as φ up to factors of C^j , $C = O(1)$.

Thus, by iterating (3.2), induction and convexity, it suffices to prove that

$$\left\| \sup_k |B_k(f_1 - \varphi_{kd_1} * f_1, \dots, f_n - \varphi_{kd_n} * f_n)| \right\|_{L^p(\mathbb{R})} \lesssim \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}$$

in the above range of p, p_i .

Define

$$B_{k,s}(f_1, \dots, f_n) := \sum_{\substack{s_i \geq 0 \\ \max\{s_i\} = s}} B_k(\psi_{kd_1+s_1} * f_1, \dots, \psi_{kd_n+s_n} * f_n).$$

Then, by Proposition 1.3, we have the bound

$$\begin{aligned} \left\| \sup_k |B_{k,s}(f_1, \dots, f_n)| \right\|_{L^1(\mathbb{R})} &\leq \sum_k \left\| B_{k,s}(f_1, \dots, f_n) \right\|_{L^1(\mathbb{R})} \\ &\leq 2^{-cs} \sum_{0 \leq s_i \leq s} \sum_k \prod_{i=1}^n \|\psi_{kd_i+s_i} * f_i\|_{L^n(\mathbb{R})} \\ &\leq 2^{-cs} \sum_{s_i \leq s} \prod_{i=1}^n \left(\sum_k \|\psi_{kd_i+s_i} * f_i\|_{L^n(\mathbb{R})}^n \right)^{1/n} \\ &= 2^{-cs} \sum_{s_i \leq s} \prod_{i=1}^n \left\| \left(\sum_k |\psi_{kd_i+s_i} * f_i|^n \right)^{1/n} \right\|_{L^n(\mathbb{R})} \\ &\leq 2^{-cs} \sum_{s_i \leq s} \prod_{i=1}^n \|Sf_i\|_{L^n(\mathbb{R})}, \end{aligned}$$

where

$$Sf := \left(\sum_k |\psi_k * f|^2 \right)^{1/2}$$

is the Littlewood–Paley square function, which is bounded on $L^p(\mathbb{R})$, $1 < p < \infty$. So

$$(3.3) \quad \sum_k \|B_{k,s}(f_1, \dots, f_n)\|_{L^1(\mathbb{R})} \lesssim_n s^n 2^{-cs} \prod_{i=1}^n \|f_i\|_{L^n(\mathbb{R})}.$$

We will interpolate this with Proposition 2.5, to see that whenever $p_1, \dots, p_n > 1$, $\frac{1}{p} = \sum_i \frac{1}{p_i} \leq 1$, there exists an absolute $c = c_{p_1, \dots, p_n; p} > 0$ such that

$$(3.4) \quad \left\| \sup_k |B_{k,s}(f_1, \dots, f_n)| \right\|_{L^p(\mathbb{R})} \lesssim s^n 2^{-cs} \prod_{i=1}^n \|f_i\|_{L^{p_i}(\mathbb{R})}.$$

A final sum over $s \geq 1$ completes the proof, assuming (3.4) holds.

We now give the details of the interpolation argument. It is enough to prove (3.4) for nonnegative functions f_i , and by monotone convergence we may assume that

they take only finitely many values. We normalize them so that $\|f_i\|_{L^{p_i}(\mathbb{R})} = 1$. By monotone convergence it is enough to prove (3.4) with the supremum extended only over finitely many values k , at which point we can choose a measurable function $k(x)$ assigning to every x the maximizer. Thus, it suffices to estimate

$$\|B_{k(x),s}(f_1, \dots, f_n)(x)\|_{L^p(\mathbb{R})} = \sup_{\|h\|_{L^{p'}(\mathbb{R})} \leq 1} \int_{\mathbb{R}} h(x)\gamma(x)B_{k(x),s}(f_1, \dots, f_n)(x) dx,$$

where $\gamma(x) \in \{z \in \mathbb{C} : |z| = 1\}$ denotes the argument of $B_{k(x)}(f_1, \dots, f_n)(x)$:

$$\gamma(x) = \begin{cases} \frac{B_{k(x),s}(f_1, \dots, f_n)(x)}{|B_{k(x),s}(f_1, \dots, f_n)(x)|} & \text{if } B_{k(x),s}(f_1, \dots, f_n)(x) \neq 0, \\ 1 & \text{if } B_{k(x),s}(f_1, \dots, f_n)(x) = 0. \end{cases}$$

Replacing h by $|h|$ makes this expression larger. Using also monotone convergence, we may assume that h is a nonnegative simple function.

Now pick exponents $q_1, \dots, q_n > 1, \frac{1}{q} = \sum_i \frac{1}{q_i} \leq 1$ such that there exists $\theta \in (0, 1)$ with

$$\frac{1}{p_i} = \frac{1-\theta}{n} + \frac{\theta}{q_i}, \quad i = 1, \dots, n.$$

Define for complex z with $\Re z \in [0, 1]$

$$f_i^{(z)}(x) = \mathbf{1}_{f_i(x) \neq 0} \exp\left(\left(z \frac{p_i}{q_i} + (1-z) \frac{p_i}{n}\right) \log f_i(x)\right)$$

and

$$h^{(z)}(x) = \mathbf{1}_{h(x) \neq 0} \exp\left(z \frac{p'}{q'} \log h(x)\right).$$

These functions are well defined since f_i and h are nonnegative. For fixed x the functions $f_i^{(z)}(x)$ and $h^{(z)}(x)$ are all bounded and analytic in the interior of the strip $\Re z \in [0, 1]$, because f_i and h are simple functions. The claim (3.4) now follows from applying the Hadamard three lines theorem to the function $F(z) = \int_{\mathbb{R}} h^{(z)}(x)\gamma(x)B_{k(x),s}(f_1^{(z)}, \dots, f_n^{(z)})(x) dx$, utilizing (3.3) when $\Re z = 0$ and Proposition 2.5 when $\Re z = 1$. This completes the proof.

References

- [1] J. Bourgain, “A nonlinear version of Roth’s theorem for sets of positive density in the real line”, *J. Analyse Math.* **50** (1988), 169–181. MR
- [2] C. Demeter, T. Tao, and C. Thiele, “Maximal multilinear operators”, *Trans. Amer. Math. Soc.* **360**:9 (2008), 4989–5042. MR
- [3] P. Durcik and J. Roos, “A new proof of an inequality of Bourgain”, *Math. Res. Lett.* **31**:4 (2024), 1047–1060. MR
- [4] P. Durcik, S. Guo, and J. Roos, “A polynomial Roth theorem on the real line”, *Trans. Amer. Math. Soc.* **371**:10 (2019), 6973–6993. MR

- [5] A. Gaitan and V. Lie, “The boundedness of the (sub)bilinear maximal function along “non-flat” smooth curves”, *J. Fourier Anal. Appl.* **26**:4 (2020), art. id. 69, 33 pp. MR
- [6] G.-D. Hong, “Polynomial Szemerédi for sets with large Hausdorff dimension on the torus”, preprint, 2025. arXiv 2507.14407
- [7] B. Hu and V. Lie, “On the curved trilinear Hilbert transform”, preprint, 2023. arXiv 2308.10706
- [8] D. Kosz, M. Mirek, S. Peluse, R. Wan, and J. Wright, “The multilinear circle method and a question of Bergelson”, preprint, 2024. arXiv 2411.09478
- [9] B. Krause, “A non-linear Roth theorem for fractals of sufficiently large dimension”, preprint, 2019. arXiv 1904.10562
- [10] B. Krause, “On polynomial progressions inside sets of large dimension”, preprint, 2025. arXiv 2508.04680
- [11] M. T. Lacey, “The bilinear maximal functions map into L^p for $2/3 < p \leq 1$ ”, *Ann. of Math. (2)* **151**:1 (2000), 35–57. MR
- [12] M. Lacey and C. Thiele, “ L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$ ”, *Ann. of Math. (2)* **146**:3 (1997), 693–724. MR
- [13] M. Lacey and C. Thiele, “On Calderón’s conjecture”, *Ann. of Math. (2)* **149**:2 (1999), 475–496. MR
- [14] X. Li, “Bilinear Hilbert transforms along curves, I: The monomial case”, *Anal. PDE* **6**:1 (2013), 197–220. MR
- [15] X. Li and L. Xiao, “Uniform estimates for bilinear Hilbert transforms and bilinear maximal functions associated to polynomials”, *Amer. J. Math.* **138**:4 (2016), 907–962. MR
- [16] V. Lie, “On the boundedness of the bilinear Hilbert transform along “non-flat” smooth curves”, *Amer. J. Math.* **137**:2 (2015), 313–363. MR
- [17] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR

Received August 14, 2025. Revised December 2, 2025.

LARS BECKER
 MATHEMATICAL INSTITUTE
 UNIVERSITY OF BONN
 BONN
 GERMANY

and

DEPARTMENT OF MATHEMATICS
 PRINCETON UNIVERSITY
 PRINCETON, NJ
 UNITED STATES
 lbecker@math.princeton.edu

BEN KRAUSE
 DEPARTMENT OF MATHEMATICS
 UNIVERSITY OF BRISTOL
 BRISTOL
 UNITED KINGDOM
 ben.krause@bristol.ac.uk

THE PLAIN SPHERE NUMBER OF A LINK

RYAN BLAIR, ALEXANDRA KJUCHUKOVA AND ELLA PFAFF

Let L be a link in S^3 . We consider a natural class of meridional presentations for $\pi_1(S^3 \setminus L)$ in which the relations are witnessed by certain embedded two-spheres easily observed in a fixed diagram of L . The Wirtinger presentation is a special case. We prove that the bridge number of L equals the smallest number of generators of $\pi_1(S^3 \setminus L)$ over all such presentations.

1. Introduction

We introduce a new definition of bridge number of a link. The definition arises in the study of the meridional rank conjecture (MRC), which asks whether the bridge number of a link L equals the smallest number of meridional generators of $\pi_1(S^3 \setminus L)$. The conjecture, posed by Cappell and Shaneson [17, Problem 1.11], has been established in a variety of cases [1; 2; 3; 6; 7; 8; 9; 10; 11; 14; 19]. The analogous statement is also shown to hold for some knotted spheres in S^4 [16].

In [5], it is shown that the bridge number equals the minimal number of meridional generators over all presentations that allow only iterated Wirtinger relations in a fixed diagram. Our main result is that equality persists after significantly generalizing the presentations considered.

Denote the *bridge number* and *meridional rank* of a link L by $\beta(L)$ and $\mu(L)$, respectively. In [5], the authors introduce a new invariant, the *Wirtinger number* of a link, denoted $\omega(L)$. The Wirtinger number is defined in terms of a “coloring game” played in a fixed diagram of L ; see Definition 2.1. Performing a coloring move at a crossing c reflects the fact that the Wirtinger meridians of the overstrand and one understrand at c generate the Wirtinger meridian of the second understrand at c . Hence, a valid coloring sequence for D demonstrates that the Wirtinger meridians of the initially colored strands, or seeds (Definition 2.4), generate $\pi_1(S^3 \setminus L)$.

By starting with a diagram in minimal bridge position and choosing the strands containing the local maxima as seeds, we easily see that $\beta(L) \geq \omega(L)$; moreover, by definition, $\omega(L) \geq \mu(L)$. In [5], it is shown that in fact $\beta(L) = \omega(L)$. We now prove that the bridge number equals the smallest number of meridional generators

MSC2020: 57K10, 57K31, 57M05.

Keywords: knot, link, bridge number, meridian, Wirtinger presentation, fundamental group, link exterior.

of $\pi_1(S^3 \setminus L)$ across a considerably more general (though still “visible” in a single diagram) class of meridional presentations.

The *plain sphere number* of a link L (Definition 2.2) is the smallest number of Wirtinger meridians that generate $\pi_1(S^3 \setminus L)$ using only relations witnessed by certain embedded two-spheres, which we now describe. Consider a link diagram D , and let γ denote an embedded circle in the plane of projection, such that γ avoids neighborhoods of crossings and meets the interior of strands of D transversely. Further assume that γ meets D in exactly n points and that, of those, precisely one is contained in a given strand, s , of D . Then, the Wirtinger meridians of the remaining strands that meet γ generate the Wirtinger meridian of s . To see this, cap off the simple closed curve γ with two disks, D_{\pm}^2 , whose interiors are disjoint from the plane of projection and such that D_+^2 (resp. D_-^2) is above (resp. below) the plane. We refer to $D_+^2 \cup_{\gamma} D_-^2$ as a *plain sphere* — it is indeed in plain sight — and we say that the sphere witnesses a relation in the group. That is, the product of (appropriately oriented) Wirtinger meridians of strands intersecting γ , taken in the order determined by γ , is trivial. The Wirtinger relation at any crossing c corresponds to a circle γ equal to the boundary of a small neighborhood of c . Clearly, $\beta(L) = \omega(L) \geq \rho(L) \geq \mu(L)$.

Theorem 1.1. *Let L be a link in S^3 . The plain sphere number of L equals the bridge number of L .*

The theorem is proved in Section 3. In Section 2 we give the formal definition of $\rho(L)$, recall the definition of $\omega(L)$, and establish some terminology. The short Section 4 provides an example contrasting the plain sphere and Wirtinger numbers of a diagram, and describes a procedure for computing $\rho(D)$ for a fixed diagram. In Section 5 we describe some plain spheres in words.

2. Preliminaries

Recall that if L is a link in \mathbb{R}^3 and $P : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is the standard projection map given by $P(x, y, z) = (x, y)$, then $P(L)$ is a *link projection* if $P|_L$ is a regular projection. Hence, a link projection is a finite four-valent graph in the plane, and we refer to the vertices of this graph as crossings. A *link diagram* is a knot projection together with labels at each crossing that indicate which strand goes over and which goes under. By standard convention, these labels take the form of deleting parts of the under-arc at every crossing, and thus we think of a link diagram as a disjoint union of closed arcs, or *strands*, in the plane, together with instructions for how to connect these strands to form a union of simple closed curves in \mathbb{R}^3 . At times we will refer to the knot projection $P(D)$ corresponding to a given knot diagram D , where $P(D)$ is obtained from D by forgetting the under- and over- information at crossings.

Let D be a diagram of a link L with n crossings. Denote by $s(D)$ the set of strands s_1, s_2, \dots, s_n and let $v(D)$ denote the set of crossings c_1, c_2, \dots, c_n . Two strands s_i and s_j of D are *adjacent* if s_i and s_j are the understrands of some crossing in D . The diagram of the unknot with a single crossing is the unique knot diagram up to planar isotopy for which there exists a strand s_i of D that is adjacent to itself. In all cases we consider, adjacent arcs are understood to be distinct.

Definition 2.1. We call D *Wirtinger k -colorable* if we have specified a set A of strands of D and a nested sequence of subsets $A = A_0 \subset A_1 \subset \dots \subset A_{|s(D)|-|A|} = s(D)$ such that the following hold:

- (1) $|A| = k$.
- (2) $A_{i+1} \setminus A_i = \{s_j\}$ for some strand s_j in D .
- (3) Whenever $A_{i+1} \setminus A_i = \{s_j\}$, s_j is adjacent to s_i at some crossing $c \in v(D)$, and $s_i \in A_i$.
- (4) The over-strand s_k at c is an element of A_i .

When D is k -colorable in the above sense, the Wirtinger meridians of the strands in A generate all Wirtinger meridians in the diagram, using only Wirtinger relations at crossings.

Definition 2.2. We call D *plain sphere k -colorable* if we have specified a set A of strands of D and a nested sequence of subsets $A = A_0 \subset A_1 \subset \dots \subset A_{|s(D)|-|A|} = s(D)$ such that the following hold:

- (1) $|A| = k$.
- (2) $A_{i+1} \setminus A_i = \{s_j\}$ for some strand s_j in D .
- (3) Whenever $A_{i+1} \setminus A_i = \{s_j\}$, there exists an embedded circle L_{i+1} in the plane of projection such that: L_{i+1} is transverse to the projection $P(D)$; L_{i+1} is disjoint from small neighborhoods of crossings in D ; $|L_{i+1} \cap s_j| = 1$; and the points of intersection between L_{i+1} and strands of D are all contained in A_{i+1} .

We refer to the circles L_i as *loops*.

Condition (3) precisely guarantees that the Wirtinger meridian of s_j is generated by the Wirtinger meridians of the strands of D that have nontrivial intersection with L_{i+1} , other than s_j itself. This is expressed by a relation in $\pi_1(S^3 \setminus L)$, witnessed by a plain sphere as defined in the introduction: an embedded two-sphere S_j^2 that intersects the plane of projection in the loop L_{i+1} . The existence of a valid coloring sequence as above shows that the Wirtinger meridians of the strands in A generate all Wirtinger meridians in the diagram, using only relations witnessed by plain two-spheres.

Remark 2.3. We could almost regard the spheres S_j^2 in the above description as simultaneously and disjointly embedded in S^3 . Indeed we will show later that the loops L_i can be assumed pairwise nonintersecting. And, each S_j^2 is the boundary union of two embedded disks, on opposite sides of the plane of projection, whose interiors do not meet. However, since the upper disk D_{j+}^2 of each S_j^2 contains the basepoint, the spheres do, in fact, intersect at a point.

Definition 2.4. When a diagram D can be colored by a valid sequence of Wirtinger coloring moves (resp. plain sphere coloring moves) as in Definition 2.1 (resp. Definition 2.2), the elements of A are called the *seed strands*, or simply *seeds*, for the coloring.

The minimum value of k such that D is Wirtinger k -colorable is the *Wirtinger number* of D , denoted $\omega(D)$. Similarly, the minimum value of k such that D is plain sphere k -colorable is the *plain sphere number* of D , denoted $\rho(D)$. We use $\omega(D)$ and $\rho(D)$ to define invariants of L .

Definition 2.5. Let $L \subset S^3$ be a link. The *Wirtinger number* of L , denoted $\omega(L)$, is the minimal value of $\omega(D)$ over all diagrams D of L . Similarly, the *plain sphere number* of L , denoted $\rho(L)$, is the minimal value of $\rho(D)$ over all diagrams D of L .

We define a number of auxiliary terms related to Definitions 2.1, 2.2 and 2.5. These terms will be used extensively in the lemmas and theorems that follow. If D is a Wirtinger k -colorable diagram with $A_i \setminus A_{i-1} = \{s_j\}$, we say that the strand s_j is colored at *stage i* by a *Wirtinger coloring move*. Similarly, if D is a plain sphere k -colorable diagram with $A_i \setminus A_{i-1} = \{s_j\}$, we say that the strand s_j is colored at *stage i* by a *plain sphere coloring move* or, for short, a *loop coloring move*. A *partially plain sphere colored* link diagram D is a nested collection of sets $A = A_0 \subset A_1 \subset \cdots \subset A_r$ that meet all of the requirements of Definition 2.2 with the exception that A_r may be a proper subset of $s(D)$. Additionally, if D is a plain sphere k -colorable diagram, then a *plain sphere coloring sequence* for D is an ordered set of loops $\mathcal{L} = (L_1, L_2, \dots, L_{|s(D)|-k})$ that are as in condition (3) of Definition 2.2. Finally, if L_i is a loop in a plain sphere coloring sequence for D that is isotopic (via an isotopy that is transverse to the link projection) to the boundary of a regular neighborhood of a crossing of D , then we call L_i a *Wirtinger loop*. Clearly, any plain sphere coloring move performed using a Wirtinger loop is also achievable via a Wirtinger coloring move. In other words, a Wirtinger coloring sequence is merely a special case of a plain sphere coloring sequence. This shows that $\rho(D) \leq \omega(D)$ for all link diagrams D and hence $\rho(L) \leq \omega(L)$ for all links L in S^3 .

An easy example of a plain sphere coloring move that is not a Wirtinger move can be found in any diagram that is not visually prime. Namely, the connected

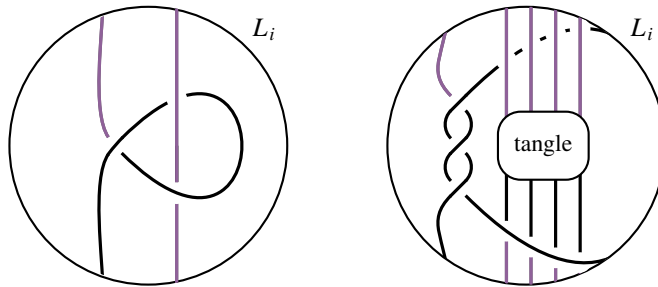


Figure 1. Possible plain sphere coloring moves at stage i of a coloring sequence which do not reduce to sequences of Wirtinger moves in the given tangle diagrams. On the left is the move used in Figure 6; on the right is a generalization. The oval contains an arbitrary n -strand tangle T_n , pictured in the case $n = 4$. The remaining tangle strand can be replaced by any one-strand tangle; as long as all but one of the points in $L_i \cap D$ are colored before stage i , the move will be valid.

sum sphere is a plain sphere intersecting two strands whose meridians cobound an annulus in the link complement. Two additional plain sphere coloring moves that are not Wirtinger moves are depicted in Figure 1. We refer to such loop moves as nontrivial loop moves. In Figure 9 of [4] there is an example of a minimal diagram D of a nonprime knot K such that $\rho(D) = \beta(D) = 5$ while $\omega(D)=6$. A prime knot diagram D with $\rho(D) < \omega(D)$ is given in Figure 6 (page 230). This exhibits a nontrivial plain sphere move that does not come from a connected sum sphere. For a family of plain sphere moves see Figure 1.

3. Proof

We prove Theorem 1.1 by showing that the plain sphere number of L equals the Wirtinger number of L . The result then follows from [5], where it is shown that the Wirtinger number and bridge number are equal. The equality $\rho(L) = \omega(L)$ relies on the following lemmas.

Lemma 3.1. *Let D be a link diagram. The plain sphere number of D can be realized by a collection of disjoint circles. That is, if $\rho(D) = n$, then there exists a set of n seed strands in D , together with an ordered set $\mathcal{L} = (L_1, \dots, L_{|s(D)|-n})$ of disjointly embedded circles, each transverse to D and disjoint from small neighborhoods of all crossings, such that \mathcal{L} defines a valid plain sphere coloring sequence for D .*

Proof. Let \mathcal{L} be an ordered collection of loops realizing the plain sphere number of D . That is, \mathcal{L} gives a coloring sequence for D starting from n seeds. After a small perturbation if needed, we may assume that the pairwise intersections of circles in \mathcal{L} are disjoint from D . We define the *complexity* of \mathcal{L} to be the (unsigned)

count of intersection points between pairs of circles in \mathcal{L} :

$$c(\mathcal{L}) := \sum_{\substack{L_i, L_j \in \mathcal{L} \\ i \neq j}} |L_i \cap L_j|.$$

Without loss of generality we may assume that \mathcal{L} has minimal complexity over all coloring sequences for D starting with n seeds. (Recall that $\rho(D) \leq \omega(D)$, and the Wirtinger number allows us to define a plain sphere coloring sequence where the loops are pairwise disjoint. Thus, it is clear that the complexity of a coloring sequence can be chosen to be zero after potentially increasing the number of seeds. What we show here is that \mathcal{L} can be chosen to realize $\rho(D)$ while also satisfying $c(\mathcal{L}) = 0$.)

If $c(\mathcal{L}) \neq 0$, we can find $L_i, L_r \in \mathcal{L}$ such that $i \neq r$ and $L_i \cap L_r \neq \emptyset$. Select L_i with the property that it is the last loop in the ordered set \mathcal{L} that has nonempty intersection with other loops in \mathcal{L} . That is, for all $k > i$, L_k is disjoint from the other loops in \mathcal{L} .

We will prove the statement by contradiction; we will show that as long as $c(\mathcal{L}) \neq 0$, the coloring sequence \mathcal{L} does not in fact minimize complexity over all plain sphere coloring sequences starting with n seeds. We will do this by producing a new plain sphere coloring sequence in which we replace L_i with another circle L_i^* , such that L_i^* defines a valid plain sphere coloring move at stage i and colors the same strand as L_i . (This implies that all subsequent moves $L_{i+1}, \dots, L_{|S(D)|-n}$ in \mathcal{L} can be performed without modification.) Lastly, we will show that $c((L_1, \dots, L_{i-1}, L_i^*, L_{i+1}, \dots, L_{|S(D)|-n})) < c(\mathcal{L})$.

In order to simplify the exposition, assume from now on that we have already carried out the coloring moves determined by L_1, L_2, \dots, L_{i-1} , and we are at stage i of the coloring process. That is, the coloring moves determined by L_1, \dots, L_{i-1} have been performed, and now L_i determines a valid plain sphere coloring move.

Denote by E_i the disk bounded by L_i in the plane of projection, and denote by \mathcal{A} the set of components of intersection between E_i and loops in \mathcal{L} :

$$\mathcal{A} := E_i \cap \left(\bigcup_{\substack{L_j \in \mathcal{L} \\ j \neq i}} L_j \right).$$

Note that \mathcal{A} is the union of properly embedded 1-manifolds (i.e. arcs and loops) in E_i . Consider the possibility that \mathcal{A} contains a loop L_j . Our assumptions about L_i imply that either L_j is disjoint from all other loops in \mathcal{L} or all points in $L_j \cap D$ have already been colored at stage i of the coloring sequence. With this in mind, we turn our attention to components of \mathcal{A} that are arcs. The boundaries of the arcs in \mathcal{A} cut $L_i = \partial E_i$ into a collection of arcs \mathcal{A}' , and moreover the arcs in \mathcal{A} cut E_i into a collection of disks \mathcal{D} (some of which may contain closed components of

\mathcal{A} in their interiors). Observe that each arc in \mathcal{A} is contained in a circle in \mathcal{L} that intersects L_i . Therefore, coloring moves determined by these circles have already been performed. This ensures that all points of intersection between D and arcs in \mathcal{A}' , with the exception of one point — the point being colored at stage i , using the loop L_i — are colored before stage i of the loop coloring sequence.

Since L_i is a circle in \mathcal{L} , we know that $L_i \cap D$ is a set of isolated points. Since L_i defines a coloring move, the intersection $L_i \cap D$ contains at least two points. At stage i of the coloring sequence, L_i determines a valid coloring move. Therefore, only one of the points in $L_i \cap D$ is uncolored. Denote this point by x and the arc in \mathcal{A}' containing x by a . Then a is contained on the boundary of one of the disks in \mathcal{D} , seen in Figure 2. Denote this disk by E_i^* and observe that $\partial E_i^* =: L_i^*$ is the union of a and other subarcs of \mathcal{A} . In particular, every point of intersection between D and L_i^* lies on an arc contained in either \mathcal{A} or \mathcal{A}' . It follows from the above that, at stage i of the sequence, x is the only point in $L_i^* \cap D$ that is not colored. Thus,

$$\mathcal{L}^* := (L_1, \dots, L_{i-1}, L_i^*, L_{i+1}, \dots, L_{|s(D)|-n})$$

is a valid plain sphere coloring sequence starting from n seeds. (To satisfy transversality conditions in the above, we use L_i^* to denote the boundary of a slightly smaller disk contained in the interior of E_i^* .)

Lastly, we check that $c(\mathcal{L}) < c(\mathcal{L}^*)$. By construction, L_i^* is the boundary of a slightly shrunken copy of the disk E_i^* . Therefore, L_i^* itself is disjoint from \mathcal{L} . By contrast, L_i has the property that $L_i \cap L_j \neq \emptyset$ for some $j \neq i$. Therefore, as claimed, replacing L_i by L_i^* has the effect of strictly decreasing the complexity of the plain sphere coloring sequence. \square

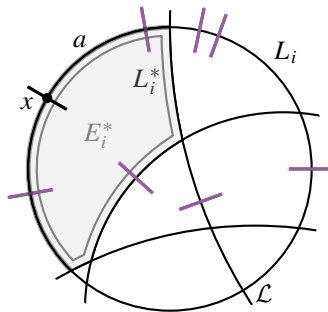


Figure 2. A region in the plane containing a link diagram at stage i of the coloring process. Subarcs of the diagram are represented by short line segments, and circles in \mathcal{L} that intersect L_i are represented by arcs. The loop L_i , which colors the strand containing the point x , can be replaced with the loop L_i^* , which is disjoint from all other loops in \mathcal{L} and represents a valid plain sphere coloring move at stage i for the same strand.

Lemma 3.2. *Let D be a diagram of a link L with $\rho(D) = n$. There exists a diagram D' of L such that the conclusion of Lemma 3.1 holds, and such that there exists a set of n seeds for D' and a valid coloring sequence \mathcal{L}' for D' , starting from those seeds, with the property that all innermost loops of \mathcal{L}' are Wirtinger loops.*

Proof. Since $\rho(D) = n$, there exists a coloring sequence \mathcal{L} for D starting from n seeds. Using Lemma 3.1, after possibly modifying \mathcal{L} , we can assume that the circles contained in \mathcal{L} are disjoint.

Let L_i be an innermost loop in \mathcal{L} . This means that L_i bounds a disk E_i whose interior is disjoint from \mathcal{L} . As before, we consider L_i at stage i of the coloring sequence. That is, among the points in the intersection $L_i \cap D$, only one point, denoted x , is not colored. Let s be the strand of D containing x and let $\partial(s) = \{a, b\}$. By the definition of a coloring move, L_i intersects s transversely in exactly one point. Therefore, a and b are contained in the two different components of $\mathbb{R}^2 \setminus L_i$. Without loss of generality, assume $a \in E_i, b \notin E_i$. Denote by c the crossing in D incident to a . See Figure 3. The key fact here is that $s \cap E_i$ is the *only arc* in $D \cap E_i$ that is not colored at stage i . This is a consequence of the following two observations: x is the only noncolored point in $\partial E_i \cap D$, so all other strands intersecting ∂E_i are colored; and L_i is innermost, so no strand entirely contained in E_i° can be colored via another circle in \mathcal{L} . Therefore, no strand in E_i° can change color after stage i . So, indeed, at stage i , of all components of $D \cap E_i$ only $s \cap E_i$ is not colored.

Case 1: Assume that s is not the overstrand at c . From the above discussion, we know that the overstrand and second understand at c are both colored before stage i . Therefore, we can color s via a Wirtinger move at the crossing c . In particular, we can replace the loop L_i in \mathcal{L} by a circle bounding a small neighborhood of c in the plane. This defines a valid coloring sequence \mathcal{L}' in which the innermost loop L_i was replaced by one that satisfies the conclusions of the lemma.

Case 2: Assume that s is the overstrand at c . We also know that c is the crossing where s terminates. Hence, s is both the overstrand and an understrand at c . In $P(D)$, the planar projection determined by D , s is the union of edges of $P(D)$

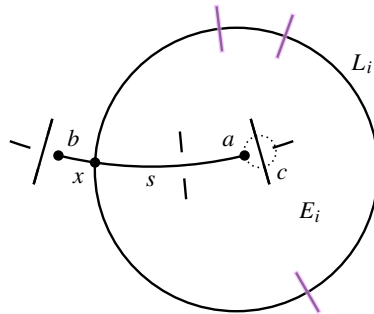


Figure 3. Exactly one uncolored strand s intersects the loop $L_i \in \mathcal{L}$ at stage i .

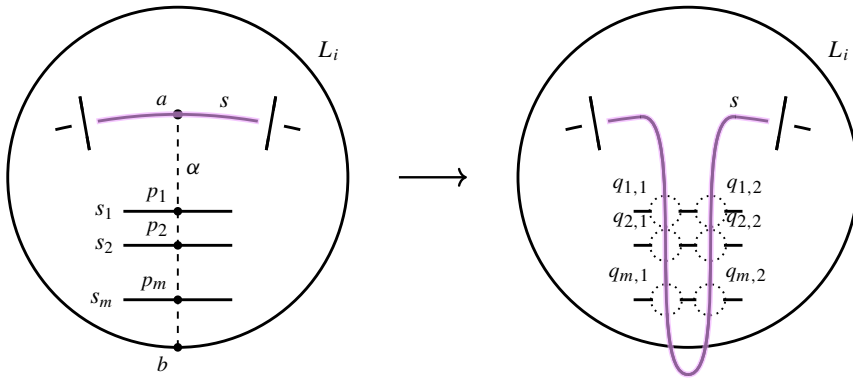


Figure 4. An isotopy to ensure no seed strand s is contained entirely in the interior of any disk E_i in the plane with $\partial(E_i) = L_i$. After a sequence of m Reidemeister II moves, $2m$ -many Wirtinger coloring loops are added in appropriate places to the sequence \mathcal{L} in order to obtain \mathcal{L}' .

and, since s is both the overstrand and understrand at c , we can conclude that the union of a subset of the edges in $P(D)$ constitutes an embedded loop γ in E_i . Moreover, γ bounds a disk $G \subset E_i$ such that $\gamma = \partial G$ is contained in the projection of the strand s . Hence, there is an isotopy of D along G creating a diagram D' in which crossing c has been resolved in the direction that preserves the number of components of the link. Note that the image of the strand s under this isotopy is contained in a strand of D' that intersects L_i in a second, necessarily colored, point. This eliminates the move determined by L_i from the coloring sequence.

One of the two cases will apply to each innermost loop in \mathcal{L} . Therefore, each innermost loop L_i may be removed after an isotopy contained in E_i or replaced by a Wirtinger loop. The result is a coloring sequence starting from n seeds in which all innermost loops represent Wirtinger moves. \square

Definition 3.3. Let D be a partially colored link diagram and \mathcal{L} an ordered set of disjoint circles transverse to D and representing valid coloring sequence moves on D . Let loop C be in \mathcal{L} and let G be the disk bounded in the plane by C . A loop C is *depth-two* if: there exists at least one loop in \mathcal{L} that is contained in the interior of G ; and all loops contained in the interior of G are innermost loops of \mathcal{L} . For $n > 2$, the definition is inductive in the natural way: a loop C is *depth- n* if disk G contains at least one depth- $(n-1)$ loop and no nested sequence of n or more loops.

Lemma 3.4. Let D be a diagram of a link L with $\rho(D) = n$. There exists a diagram D' of L and a valid coloring sequence \mathcal{L}' for D' starting from n seeds, such that all the conclusions of Lemma 3.2 are satisfied and, moreover, if L_i is any loop in \mathcal{L}' , then no seed for the coloring sequence is entirely contained in the interior of the disk E_i bounded by L_i in the plane of projection.

Proof. By Lemma 3.2 we can assume that $\rho(D) = n$ and that \mathcal{L} is a plain sphere coloring sequence starting with seeds s_1, s_2, \dots, s_n such that all innermost loops in \mathcal{L} are Wirtinger loops. If $L_i \in \mathcal{L}$ represents a Wirtinger move, then without loss of generality we may assume that L_i is the boundary of a small neighborhood of a crossing in D , and the desired conclusion automatically holds for L_i . In general, however, E_i may contain multiple crossings of D and potentially entire strands. Assume that a seed strand s is contained in E_i° . Define α to be an embedded arc in E_i , transverse to D , disjoint from small neighborhoods of crossings of D , and with the property that $\partial\alpha =: \{a, b\}$ has $a \in s^\circ$ and $b \in (L_i \setminus D)$. See Figure 4. We may use α to guide an isotopy of s in the plane of projection to produce a new strand such that the image of a after the isotopy lies outside of E_i , in a small neighborhood of b . The isotopy can be thought of as a sequence of Reidemeister II moves in which s passes over every strand it encounters along α .

Let D_1 be the diagram that results from the above isotopy. We claim that $\rho(D_1) = n$ and that D_1 admits a coloring sequence that is obtained from \mathcal{L} by inserting $2m$ additional loops representing Wirtinger moves, where $m := |\alpha \cap D|$. More specifically, let $\alpha \cap D =: \{p_1, \dots, p_m\}$. The additional Wirtinger loops will be interspersed across \mathcal{L} according to when the strands in D containing the p_j are colored. See Figure 4 for the isotopy of s and the new strands and crossings created.

Note that each point p_j corresponds to a Reidemeister II move performed while isotoping s along α . Thus, in a neighborhood of p_j , two new crossings—call them $q_{j,1}$ and $q_{j,2}$ —are created during the isotopy, and at both crossings s is the overstrand. Let s_j denote the strand of D containing p_j . (After the isotopy, s_j is subdivided into three strands by $q_{j,1}$ and $q_{j,2}$.) At some stage in the coloring sequence \mathcal{L} for D , the strand s_j is colored. This implies that, at the corresponding stage of the coloring sequence for D_1 , one of the understrands at either $q_{j,1}$ or $q_{j,2}$ is colored. But the overstrand at these crossings, s , is a seed, so it is colored as well. Thus, we can perform two consecutive Wirtinger moves at these crossings and color all three strands into which s_j has been subdivided. As a result, any further coloring moves given by \mathcal{L} whose validity relies on s_j being colored will be valid. Repeating this procedure for each $j \in \{1, \dots, m\}$ produces the desired coloring sequence for D_1 .

In sum, after an isotopy supported in a neighborhood of an arc, we ensured that the seed strand s is no longer contained in the interior of E_i . Moreover, we produced a coloring sequence, with the same number of seeds, for the diagram resulting from this isotopy. Repeating this procedure for every instance where a seed is contained entirely within the interior of one of the disks E_i , we arrive at the desired diagram D' and valid coloring sequence \mathcal{L}' from n seeds. \square

Before presenting the next lemma, we review some standard properties of tangles. An m -strand tangle is a collection of m disjoint arcs properly embedded in a 3-ball.

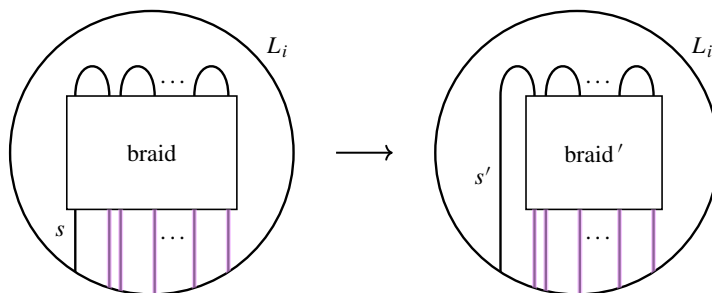


Figure 5. An isotopy at stage i of the coloring sequence of a rational m -strand tangle supported in disk E_i with $\partial(E_i) = L_i$. After isotopy, starting with $2m-1$ colored strands that intersect L_i , all strands that meet E_i can be colored using only Wirtinger moves.

An m -strand tangle is *rational* if all arcs of the tangle can be simultaneously isotoped into the boundary of the 3-ball.

Remark 3.5. Given a tangle R properly embedded in $D^2 \times [0, 1]$ such that $\partial R \subset (\partial D^2) \times [0, 1]$, R has a tangle diagram, analogous to a knot diagram, achieved by projecting R onto $D^2 \times \{0\}$. Let $x \in D^2$ denote the point with coordinates $(0, 1)$, where D^2 is identified with the unit disk in the plane. If R is a rational tangle embedded in $D^2 \times [0, 1]$ such that $\partial R \subset \partial D^2 \times [0, 1]$, then after planar isotopy, and possibly Reidemeister moves supported in the interior of $D^2 \times \{0\}$, R has a tangle diagram as in the left image in Figure 5. In brief, this is true since the disks of parallelism for the arcs in R can all be isotoped to be disjoint from the disk that is a regular neighborhood of $(D^2 \times \{1\}) \cup (D^2 \times \{0\}) \cup (\{x\} \times [0, 1])$ in $\partial(D^2 \times [0, 1])$. Isotoping the arcs of R along these disks of parallelism until they almost lie in $\partial(D^2 \times [0, 1])$ results in a diagram of R that lies in an annular neighborhood of $\partial D^2 \times \{0\}$ in $D^2 \times \{0\}$ such that the projection of R is disjoint from a neighborhood of $\{x\} \times \{0\}$ and each arc of the projection has exactly one maximum with respect to the radial height function on $D^2 \times \{0\}$ (where we think of height as increasing as we approach the center of the disk). After a planar isotopy, this diagram becomes the diagram of the left image in Figure 5.

Remark 3.6. Any tangle R properly embedded in $D^2 \times [0, 1]$ such that each arc of R has a unique local maximum (or each has a unique local minimum) with respect to projection on the $[0, 1]$ component of $D^2 \times [0, 1]$ is a rational tangle. In brief, this is true since if an arc α in R has a unique local maximum at height $a \in [0, 1]$, then there is a disk of parallelism between the subarc of α above $D^2 \times \{a - \epsilon\}$ and an arc in $D^2 \times \{a - \epsilon\}$. Since every arc in R contains a unique local maximum, this disk can be extended downward by moving both endpoints along the arc until it becomes a disk of parallelism between α and a subarc in $\partial(D^2 \times [0, 1])$. After generating a disk of parallelism in this way for each arc of R , we can ensure that

these disks are pairwise disjoint by an application of a standard innermost disk and outermost arc argument. Thus, R is rational.

Lemma 3.7. *Let D be a diagram of a link L with $\rho(D) = n$, and let $\mathcal{L} = (L_1, L_2, \dots, L_{|s(D)|-n})$ be a plain sphere coloring sequence for D , satisfying the conclusions of Lemma 3.4. Let L_i be a depth-two loop in \mathcal{L} and denote by E_i the disk in the plane of projection bounded by L_i . There exists an isotopy of D , supported in a small neighborhood of E_i , such that the resulting diagram D' has a coloring sequence \mathcal{L}' with n seeds and such that L_i is replaced by a collection of innermost loops.*

Proof. By assumption, \mathcal{L} is a coloring sequence for D starting from n seeds and satisfying the conclusions of Lemma 3.4. Since L_i is a depth-two loop, there exists at least one loop of \mathcal{L} contained in E_i and moreover all such loops are innermost. Since the coloring sequence satisfies the conclusions of Lemma 3.2, all loops contained in E_i are Wirtinger loops. Since the conclusions of Lemma 3.4 hold as well, we also have that no seed for the coloring sequence \mathcal{L} is entirely contained in E_i . Let s be the unique strand of D that intersects L_i and is colored at stage i in the coloring process.

Recall that $P : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ denotes the standard projection map and that $P(L)$ is the regular projection that gives rise to D . We can assume that $L \subset \mathbb{R}^2 \times [0, 1]$. If $L_i \cap D = 2m$, then, since no seed is entirely contained in E_i , $E_i \times [0, 1] = P^{-1}(E_i) \cap (\mathbb{R}^2 \times [0, 1])$ is a 3-ball containing a properly embedded m -strand tangle $T = L \cap E_i \times [0, 1]$. In particular, the lack of seeds in E_i implies T contains no simple closed curves. Double $E_i \cap D$ along ∂E_i to obtain a link diagram D^{dbl} of the link L^{dbl} where L^{dbl} is obtained by doubling T along $\partial E_i \times [0, 1]$. Since E_i does not entirely contain any seed strands and all loops contained in E_i are Wirtinger loops, the set of $2m$ strands of D^{dbl} that intersect ∂E_i , call them $\sigma_1, \dots, \sigma_{2m}$, are a collection of seed strands for a Wirtinger coloring of D^{dbl} . By Theorem 1.2 of [4], there is an isotopy of L^{dbl} achieved by modifying only the z -coordinate of points on L^{dbl} and preserving the projection $P(L^{dbl})$ at all times after which L^{dbl} has exactly $2m$ local maxima with respect to projection to the z -axis and each of these maxima project to exactly one of $\sigma_1 \cap \partial E_i, \dots, \sigma_{2m} \cap \partial E_i$. This isotopy restricts to a proper isotopy of T in $E_i \times [0, 1]$ after which each strand of T contains a unique local minimum with respect to the z -axis. Hence, by Remark 3.6, T must be a rational m -strand tangle. Since T is rational, then, by Remark 3.5, there is a sequence of Reidemeister moves supported in E_i that produces a new diagram D^* as in Figure 5, left (see previous page). That is, $T \cap E_i$ is obtained from a braid by taking the plat closure to one side. By replacing the point x in Remark 3.5 by a point in L_i close to $s \cap L_i$ in the clockwise direction, we can ensure that s is the leftmost strand entering the braid box from below, as in the left image in Figure 5.

Furthermore, it is well known that there is an isotopy of T fixing ∂T that pulls one arc out of the braid box, shown in Figure 5; see, for example, the Claim within the proof of Theorem 1 in [20]. After isotopy, call the resulting link diagram D' (Figure 5, right) and let s' be the strand containing the image of s . Notice that since all strands entering the braid box from below are colored at stage i , then all strands of D' that meet E_i can be colored via Wirtinger loops, even s' .

Thus, after an isotopy supported in a neighborhood of E_i , we can replace L_i by a collection of Wirtinger loops. In particular, we eliminated a depth-two loop from the coloring sequence. \square

Corollary 3.8. *If L is a link and D is a diagram of L with $\rho(D) = \rho(L) = n$, then there exists a diagram D' , equivalent to D , such that $\rho(D') = \omega(D') = n$.*

Proof. Let D be a link diagram and \mathcal{L} a plain sphere coloring sequence for D starting with n seeds. By repeated application of Lemma 3.7, we can eventually eliminate all depth-two loops. Eliminating a single loop as described in the lemma may not immediately reduce the number of depth-two loops as it could simultaneously produce a new depth-two loop from a previous depth-three loop; however, the procedure can be iterated as many times as needed.

After the necessary amount of persistence, we arrive at a diagram D' , equivalent to D , which admits a plain sphere coloring sequence \mathcal{L}' , also with n seeds, such that \mathcal{L}' contains no depth-two loops. Since D was chosen to be minimal with respect to plain sphere number, then $\rho(D') = n$. But this in fact implies that all loops in \mathcal{L}' are innermost. By Lemma 3.2, all moves in the coloring sequence \mathcal{L}' are Wirtinger moves. Therefore, $n \geq \omega(D')$. But we also have, by definition, that $\omega(D') \geq \rho(D') = n$. \square

Proof of Theorem 1.1. Let L be a link in S^3 with $\rho(L) = n$. We wish to show that $\omega(L) = n$ as well. Let D be a diagram realizing $\rho(L)$. This implies that D admits a plain sphere coloring sequence starting from n seeds. By Corollary 3.8, there exists a diagram D' of L that can be colored starting with n seeds and using only Wirtinger moves. Therefore, $n \geq \omega(D') \geq \omega(L) \geq \rho(L) = n$. It then follows from [5] that $\beta(L) = \omega(L) = n$. \square

4. Example and computations

4.1. Example. We compute the plain sphere number of a minimal diagram of the knot $K = 14n1527$, shown in Figure 6. The bridge number of K is 3 and equals the plain sphere number of the pictured diagram. The Wirtinger number of the diagram is 4. We thank Nathan Dunfield for detecting this example and sharing it with us.

4.2. Computing the plain sphere number of diagrams. Let D be a link diagram in S^2 , $P(D)$ be the projection corresponding to D . Recall that a region of D is the

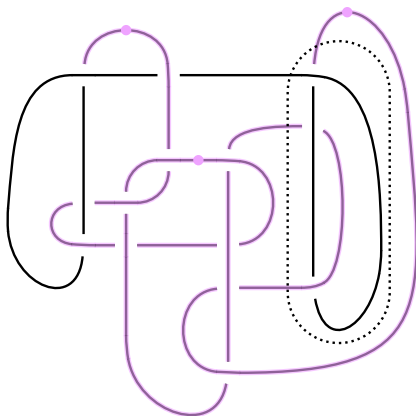


Figure 6. A partially colored minimal diagram D of the knot $K = 14n1527$ with $\beta(K) = \rho(D) = 3$ and $\omega(D) = 4$. Seed strands are marked with dots. Colored strands are purple. At this stage of the coloring sequence, only Wirtinger moves have been performed, and no further Wirtinger moves are possible. The dotted circle L_8 represents a loop coloring move. (There are several alternative loop coloring moves.) After coloring the black strand intersecting L_8 , the rest of the diagram can be colored by Wirtinger moves.

closure of a connected component of $S^2 \setminus P(D)$. Note that the boundary of a region is a cycle in $P(D)$ where $P(D)$ is viewed as a 4-valent graph. Let Γ_D denote the dual graph to D .

Definition 4.1. Let L be an embedded loop in S^2 that is transverse to D and is disjoint from neighborhoods of crossings of D . We say L is *tight* if for every region R of D such that $L \cap R \neq \emptyset$, $L \cap R$ is an arc properly embedded in R with endpoints on distinct edges of $P(D)$ in ∂R .

Note that every tight loop in S^2 is isotopic in S^2 to an embedded loop in Γ_D via an isotopy that is transverse to the edges of $P(D)$ and disjoint from the vertices of $P(D)$.

Lemma 4.2. *Let D be a link diagram. There exists a minimal plain sphere coloring sequence for D consisting entirely of tight loops.*

Proof. Given a plain sphere coloring sequence \mathcal{L} for a diagram D , define

$$\tau(\mathcal{L}) = \sum_{L \in \mathcal{L}} |L \cap D|.$$

Let \mathcal{L} be a minimal plain sphere coloring sequence such that the loops in \mathcal{L} are disjointly embedded. Assume that \mathcal{L} minimizes $\tau(\mathcal{L})$ over all such sequences. By Definition 2.2, for every $L_i \in \mathcal{L}$ and for every region R of D , if $L_i \cap R \neq \emptyset$, then $L_i \cap R$ is a collection of arcs.

Suppose toward a contradiction that there exists $L_i \in \mathcal{L}$ and a region R of D such that $L_i \cap R$ consists of two or more arcs. If L_i colors the strand s of D , by definition, $L_i \cap s$ is a single point. Let γ_1 and γ_2 be two distinct arcs of $L_i \cap R$ that are contained in the boundary of the same component of $R \setminus L_i$. Call the closure of this component C . Let α be an arc properly embedded in C with one endpoint in γ_1 and one endpoint in γ_2 . Let N be a rectangular I -fibered neighborhood of α in C such that $\partial N = \beta_1 \cup \beta_2 \cup \delta_1 \cup \delta_2$ where β_i is embedded in γ_i and $\delta_i^\circ \subset C^\circ$. Surger L_i along α in the standard way by removing β_1 and β_2 and gluing in δ_1 and δ_2 . This results in two loops in S^2 , L'_i and L''_i , both of which are transverse to D and disjoint from neighborhoods of crossings in D . Both L'_i and L''_i have nontrivial intersection with D , so

$$|L'_i \cap D| < |L_i \cap D| \quad \text{and} \quad |L''_i \cap D| < |L_i \cap D|.$$

Exactly one of L'_i and L''_i have nontrivial intersection with s , say L'_i does. All other points of intersection of D with L'_i and L''_i are contained in colored strands. Then $\mathcal{L}' = (L_1, \dots, L_{i-1}, L'_i, L_{i+1}, \dots, L_m)$ is a minimal plain sphere coloring sequence with $\tau(\mathcal{L}') < \tau(\mathcal{L})$, a contradiction. \square

Consequently, $\rho(D)$ can be realized by a sequence of loops represented by embedded cycles in Γ_D , an approach suggested by Nathan Dunfield. In particular, computing $\rho(D)$ for a fixed link diagram D is algorithmic. The forthcoming version 3.3 of SnapPy [13] will include a feature computing $\rho(D)$. A computation performed using this feature shows that over 600 diagrams in the knot table through 16 crossings have plain sphere number strictly less than Wirtinger number (Nathan Dunfield, private communication, 2025).

We conclude by pointing out that our plain spheres generalize the connected sum spheres that appear in the study of the visual primeness of links; see [18; 12; 15]. Of course, not all relators in the group of a link are necessarily witnessed by plain spheres in a given diagram D . This is readily seen to be the case even when D is a diagram of the unknot. Since any finite set of meridians of a link can be realized as Wirtinger meridians in some diagram, one can regard the meridional rank conjecture as positing equality between the bridge number and a certain elusive sphere number of links, allowing immersed spheres that more easily evade the eye.

5. Figures

*Caught in tumbleweed, in perpetual slumber
 They roll: round, squished, immersed, cucumber
 Elusive, blistering, fierce
 Faint! Diabolical! Spheres!
 Does each breed breed the bridge number?*

Acknowledgements

AK and EP are partially supported by NSF grant DMS 2204349. RB is partially supported by NSF grant DMS 2424734. We are particularly grateful to Nathan Dunfield for implementing a computation of the plain sphere number of a diagram in SnapPy and comparing the performance of the plain sphere and Wirtinger numbers on minimal diagrams through 16 crossings.

References

- [1] S. Baader and A. Kjuchukova, “Symmetric quotients of knot groups and a filtration of the Gordian graph”, *Math. Proc. Cambridge Philos. Soc.* **169**:1 (2020), 141–148. MR
- [2] S. Baader, R. Blair, and A. Kjuchukova, “Coxeter groups and meridional rank of links”, *Math. Ann.* **379**:3-4 (2021), 1533–1551. MR
- [3] S. Baader, R. Blair, A. Kjuchukova, and F. Misev, “The bridge number of arborescent links with many twigs”, *Algebr. Geom. Topol.* **23**:1 (2023), 75–85. MR
- [4] R. Blair, A. Kjuchukova, and M. Ozawa, “The incompatibility of crossing number and bridge number for knot diagrams”, *Discrete Math.* **342**:7 (2019), 1966–1978. MR
- [5] R. Blair, A. Kjuchukova, R. Velazquez, and P. Villanueva, “Wirtinger systems of generators of knot groups”, *Comm. Anal. Geom.* **28**:2 (2020), 243–262. MR
- [6] M. Boileau and H. Zieschang, “Nombre de ponts et générateurs méridiens des entrelacs de Montesinos”, *Comment. Math. Helv.* **60**:2 (1985), 270–279. MR
- [7] M. Boileau and B. Zimmermann, “The π -orbifold group of a link”, *Math. Z.* **200**:2 (1989), 187–208. MR
- [8] M. Boileau, E. Dutra, Y. Jang, and R. Weidmann, “Meridional rank of knots whose exterior is a graph manifold”, *Topology Appl.* **228** (2017), 458–485. MR
- [9] G. Burde, “Links covering knots with two bridges”, *Kobe J. Math.* **5**:2 (1988), 209–219. MR
- [10] C. R. Cornwell, “Knot contact homology and representations of knot groups”, *J. Topol.* **7**:4 (2014), 1221–1242. MR
- [11] C. R. Cornwell and D. R. Hemminger, “Augmentation rank of satellites with braid pattern”, *Comm. Anal. Geom.* **24**:5 (2016), 939–967. MR
- [12] P. R. Cromwell, “Positive braids are visually prime”, *Proc. London Math. Soc.* (3) **67**:2 (1993), 384–424. MR
- [13] M. Culler, N. M. Dunfield, M. Goerner, and J. R. Weeks, “SnapPy: a computer program for studying the geometry and topology of 3-manifolds”, software, 2010–2025, available at <http://snappy.computop.org>.
- [14] E. Dutra, “Meridional rank of Whitehead doubles”, preprint, 2022. arXiv 2212.13081
- [15] P. Feller, L. Lewark, and M. Orbegozo Rodriguez, “Homogeneous braids are visually prime”, *J. Topol.* **18**:3 (2025), art. id. e70040, 28 pp. MR
- [16] J. Joseph and P. Pongtanapaisan, “Meridional rank and bridge number of knotted 2-spheres”, *Canad. J. Math.* **77**:1 (2025), 282–299. MR
- [17] R. Kirby, “Problems in low-dimensional topology”, pp. 35–473 in *Geometric topology* (Athens, GA, 1993), edited by R. Kirby, AMS/IP Stud. Adv. Math. **2.2**, Amer. Math. Soc., 1997. MR
- [18] W. Menasco, “Closed incompressible surfaces in alternating knot and link complements”, *Topology* **23**:1 (1984), 37–44. MR

- [19] M. Rost and H. Zieschang, “Meridional generators and plat presentations of torus links”, *J. London Math. Soc. (2)* **35**:3 (1987), 551–562. MR
- [20] J. Schultens, “Additivity of bridge numbers of knots”, *Math. Proc. Cambridge Philos. Soc.* **135**:3 (2003), 539–544. MR

Received June 24, 2025. Revised February 20, 2026.

RYAN BLAIR
DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY
LONG BEACH, CA
UNITED STATES
ryan.blair@csulb.edu

ALEXANDRA KJUCHUKOVA
UNIVERSITY OF NOTRE DAME
NOTRE DAME, IN 46556
UNITED STATES
akjuchuk@nd.edu

ELLA PFAFF
UNIVERSITY OF NOTRE DAME
NOTRE DAME, IN 46556
UNITED STATES
epfaff@nd.edu

ON THE TEST PROPERTIES OF THE FROBENIUS ENDOMORPHISM

OLGUR CELIKBAS, ARASH SADEGHI AND YONGWEI YAO

We prove two theorems concerning the test properties of the Frobenius endomorphism over commutative Noetherian local rings of prime characteristic p . Our first theorem generalizes a result of Funk and Marley on the vanishing of Ext and Tor modules, while our second theorem generalizes one of our previous results on maximal Cohen–Macaulay tensor products. In these earlier results, we replace ${}^e R$ with a more general module ${}^e M$, where R is a Cohen–Macaulay ring, M is a Cohen–Macaulay R -module with full support, and ${}^e M$ is the module viewed as an R -module via the e -th iteration of the Frobenius endomorphism. We also provide examples and present applications of our results, yielding new characterizations of the regularity of local rings.

1. Introduction

Throughout the paper, all rings are assumed to be commutative and Noetherian. By (R, \mathfrak{m}, k) , we mean R is a local ring with a unique maximal ideal \mathfrak{m} and residue field k .

Let R be a ring of prime characteristic p , $F : R \rightarrow R$ be the Frobenius endomorphism, and let M be an R -module. Each iteration F^e of F defines a new R -module structure on M , denoted by ${}^e M$, whose scalar multiplication is given as follows: For $r \in R$ and $x \in {}^e M$, we have that $r \cdot x = r^{p^e} x$. We say that R is F -finite if, for some $e \geq 1$ (or equivalently, for all $e \geq 1$), the module ${}^e R$ is finitely generated over R ; see, for example, [24]. We denote by $F_R^e(-)$ the scalar extension along the e -th iteration $F_R^e : R \rightarrow R$ of F . Thus, if $\sum_i n_i \otimes s_i \in F_R^e(M)$, where $F_R^e(M) = M \otimes_R {}^e R$, $n_i \in M$ and $r_i \in R$, then $r \cdot (\sum_i n_i \otimes s_i) = \sum_i n_i \otimes (r s_i)$, with $r s_i$ being the product of r and s_i in R . Note that $F_R^e(M)$ is the S -module $M \otimes_R S$ obtained via the base change $F^e : R \rightarrow S = R$.

The module structure of ${}^e R$ (as an R -module) contains important information about the homological properties of the ring R . For example, a remarkable result

MSC2020: primary 13A35, 13C14, 13D07; secondary 13C10, 13C11.

Keywords: Cohen–Macaulay module, depth and torsion properties of tensor products of modules, Frobenius endomorphism, ring of prime characteristic, Ext and Tor.

of Kunz [17] shows that R is regular if and only if eR is a flat R -module for some (or equivalently, for all) $e \geq 1$; see also [4] and [25] for extensions of this result. Motivated by Kunz's result, the test properties of the Frobenius endomorphism have been extensively studied.

If N is a finitely generated R -module, it follows from the work of Herzog [12] and Peskine and Szpiro [23; 24] that $\text{pd}_R(N) < \infty$ if and only if $\text{Tor}_i^R({}^eR, N) = 0$ for infinitely many e and for all $i \geq 1$. Avramov and Miller [3] showed that, if R is a complete intersection, the vanishing of a single $\text{Tor}_n^R({}^eR, N)$ for some $e \geq 1$ and $n \geq 1$ suffices to conclude that $\text{pd}_R(N) < \infty$. Koh and Lee [16] developed ideas rooted in techniques of Burch [7], Herzog [12], and Hochster [13], and showed that eR detects the finiteness of N even when finitely many Tor modules vanish. Specifically, Koh and Lee proved that, given integers $e \gg 0$ and $t \geq 1$, if $\text{Tor}_i^R({}^eR, N) = 0$ for $i = t, \dots, t + \text{depth}(R)$, then $\text{pd}_R(N) < \infty$. They further showed that in the case where R is Cohen–Macaulay, the number of vanishing Tor modules can be reduced by one. We refer the reader to the expository work [21] of Miller for further details.

In this paper, we focus on the following result of Funk and Marley [11; 10], which examines the vanishing of $\text{Tor}_i^R({}^eR, N)$ for the case where R is Cohen–Macaulay and N is possibly an infinitely generated R -module.

1.1 (Funk and Marley [11, 3.1 and 3.2]). Let (R, \mathfrak{m}, k) be a d -dimensional Cohen–Macaulay local ring of prime characteristic p , with $d \geq 1$, and let N be an R -module. Given integers $e \gg 0$ and $t \geq 1$, the following hold:

- (i) If $\text{Tor}_i^R({}^eR, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{fd}_R(N) \leq d$.
- (ii) If R is F -finite and $\text{Ext}_R^i({}^eR, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N) \leq d$.

One of the main goals of this paper is to generalize the result of Funk and Marley stated in 1.1. In fact, we prove more and establish the following theorem:

Theorem 1.2. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Assume $\text{depth}(R) \geq 1$. Given integers $t \geq 1$ and $e \gg 0$, we have the following:*

- (i) *If $\text{Tor}_i^R({}^eM, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{fd}_R(N) \leq d$.*
- (ii) *If R is F -finite and $\text{Ext}_R^i({}^eM, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N) \leq d$.*
- (iii) *If N is finitely generated and $\text{Ext}_R^i(N, {}^eM) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{pd}_R(N) \leq t - 1$.*

Parts (i) and (ii) of Theorem 1.2 recover the Funk–Marley result for the case where $M = R$. Note that, in each part of Theorem 1.2, only d consecutive vanishings

of Ext or Tor modules are needed to conclude the homological property the module of N .

We first record several preliminary results in Section 2 and then prove the first two parts of Theorem 1.2 as Theorem 3.2 in Section 3. The third part of Theorem 1.2 is established as Theorem 4.3 in Section 4. Additionally, in Example 3.4, we show that the conclusion of Theorem 1.2 may fail if the ring R in question has zero depth.

Li [18] proved that, if (R, \mathfrak{m}, k) is a Cohen–Macaulay local ring, N is a finitely generated R -module with rank, and $F_R^e(N)$ is maximal Cohen–Macaulay for some $e \gg 0$, then N is free. In [8], the authors of the present paper replaced the rank hypothesis on N with the weaker assumption that N is generically free, and proved the following result:

1.3 (Celikbas, Sadeghi, Yao [8, 1.3]). Let (R, \mathfrak{m}, k) be a Cohen–Macaulay local ring of prime characteristic p , and let M and N be finitely generated R -modules. Assume N is generically free, that is, $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for all $\mathfrak{p} \in \text{Ass}(R)$. If $F_R^e(N)$ is maximal Cohen–Macaulay for some $e \gg 0$, then N is free.

In Section 4, as a byproduct of Theorem 1.2(iii), we generalize 1.3 and prove the following result; see Corollary 4.5. Note that $F_R^e(N) \otimes_R M$ is the S -module $(N \otimes_R S) \otimes_S M$, where $R \rightarrow S = R$ is the e -th iteration of the Frobenius endomorphism.

Theorem 1.4. *Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p , and let M and N be finitely generated R -modules. Assume the following conditions hold:*

- (i) M is Cohen–Macaulay and $\text{Supp}_R(M) = \text{Spec}(R)$.
- (ii) N is generically free, that is, $N_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$ -module for all $\mathfrak{p} \in \text{Ass}(R)$.

If $F_R^e(N) \otimes_R M$ is maximal Cohen–Macaulay for some $e \gg 0$, then N is free.

Examples 4.8 and 4.9 showcase the necessity of the hypotheses $\text{Supp}_R(M) = \text{Spec}(R)$ and N is generically free in Theorem 1.4. An immediate consequence of Theorem 1.4 over one-dimensional rings can be stated as follows:

Corollary 1.5. *Let (R, \mathfrak{m}, k) be a one-dimensional reduced local ring of prime characteristic p and let $0 \neq I$ be an ideal of R . If $F_R^e(N) \otimes_R I$ is torsion-free for some finitely generated R -module N and $e \gg 0$, then N is free.*

Theorem 1.2(iii), in addition to Theorem 1.4, has other applications, namely Corollaries 4.4, 4.6, and 4.7. Moreover, in Corollary 4.10, we obtain new characterizations of the regularity in terms of the vanishing of Ext and Tor.

2. Preliminaries

In this section, we record several preliminary results and observations that are necessary for our arguments in the subsequent sections. For the main results of this paper, one can skip this section and proceed to Sections 3 and 4.

2.1. Let (R, \mathfrak{m}, k) be a local ring and let M be an R -module.

- (i) We set $M^\vee = \text{Hom}_R(M, E_R(k))$, where $E_R(k)$ is the injective hull of k . Note that $\text{Hom}_R(-, E_R(k))$ is a faithful exact functor.
- (ii) Assume that M is finitely generated over R . Given $n \geq 1$, we denote by $\Omega_R^n M$ the n -th syzygy of M , namely, the image of the n -th differential map in a minimal free resolution of M . By convention, $\Omega_R^0 M = M$.
- (iii) If $M \neq 0$ is finitely generated over R and $\dim_R(M) = t$, we define the *Hilbert–Samuel multiplicity* of M as

$$e_R(M) = t! \lim_{n \rightarrow \infty} \frac{\text{length}_R(M/\mathfrak{m}^n M)}{n^t},$$

which is a positive integer; see, for example, [20, p. 107].

2.2. Let (R, \mathfrak{m}, k) be a local ring and let N and X be R -modules.

- (1) Assume $\text{Soc}_R(X) \not\subseteq \mathfrak{m}X$ and let $x \in \text{Soc}_R(X) - \mathfrak{m}X$. This implies that $Rx \cong k$ and the nonzero map $Rx \hookrightarrow X \twoheadrightarrow X/\mathfrak{m}X$, $x \mapsto \bar{x}$, splits. Therefore, $Rx \hookrightarrow X$ splits and thus k is a direct summand of X as an R -module.
- (2) Assume R has prime characteristic p and $X = {}^e M$ for some R -module $0 \neq M$ and $e \geq 0$.
 - (a) If $\text{Soc}_R(X) \not\subseteq \mathfrak{m}X$, that is, ${}^e(0 :_M \mathfrak{m}^{[p^e]}) \not\subseteq {}^e(\mathfrak{m}^{[p^e]}M)$, then part (1) shows that k is a direct summand of ${}^e M$ as an R -module.
 - (b) Assume M is finitely generated over R with $\text{depth}_R(M) = 0$. For all $e \gg 0$, we have ${}^e(0 :_M \mathfrak{m}^{[p^e]}) = \text{Soc}_R({}^e M) \supseteq {}^e(\text{Soc}_R(M)) \not\subseteq {}^e(\mathfrak{m}^{[p^e]}M)$. Hence, part (i) shows that k is a direct summand of ${}^e M$ for every $e \gg 0$.
 - (c) Assume (R, \mathfrak{m}, k) is Artinian. Then ${}^e(0 :_M \mathfrak{m}^{[p^e]}) \not\subseteq {}^e(\mathfrak{m}^{[p^e]}M)$ for all $e \gg 0$. By part (i), k is a direct summand of ${}^e M$ for all $e \gg 0$. Thus, given $i \geq 1$, if $\text{Ext}_R^i({}^e M, N) = 0$ for some $e \gg 0$, then $\text{Ext}_R^i(k, N) = 0$ and hence N is injective; see [6, 3.1.12] and also [9, 2.0.10].

2.3. Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p . There exists a local flat ring homomorphism $(R, \mathfrak{m}, k) \rightarrow (S, \mathfrak{n}, \ell)$ such that S is F-finite, $\mathfrak{m}S = \mathfrak{n}$ and $|\ell| = \infty$; hence S is faithfully flat over R , $\dim(S) = \dim(R)$ and $e(R) = e(S)$; it also follows that R is Cohen–Macaulay (respectively, regular) if and only if S is Cohen–Macaulay (respectively, regular). (For example, with $\hat{R} \cong k[[x_1, \dots, x_m]]/I$, we can pick $S = \bar{k}[[x_1, \dots, x_m]]/I\bar{k}[[x_1, \dots, x_m]]$, where \bar{k} is the algebraic closure of k .) In the case where S is such an extension of R and M is a finitely generated R -module, it follows that M is a Cohen–Macaulay R -module with $\text{Supp}_R(M) = \text{Spec}(R)$ if and only if $M \otimes_R S$ is a Cohen–Macaulay S -module with $\text{Supp}_S(M \otimes_R S) = \text{Spec}(S)$; also, M is free over R if and only if $M \otimes_R S$ is free over S .

2.4. Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p and let $M \neq 0$ be a finitely generated R -module. We set:

$$\begin{aligned} \text{crs}_p(M) &= \min\{e \geq 0 \mid (0 :_{M/(\underline{x})M} \mathfrak{m}^{[p^e]}) \not\subseteq \mathfrak{m}^{[p^e]}(M/(\underline{x})M) \\ &\hspace{15em} \text{for an } M\text{-regular sequence } \underline{x}\}, \\ \text{drs}_p(M) &= \min\{e \geq 0 \mid \mathfrak{m}^{[p^e]}M \subseteq (\underline{x})M \text{ for a system of parameters } \underline{x} \text{ on } M\}. \end{aligned}$$

It seems unknown whether or not $\sup\{\text{crs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Supp}_R(M)\}$ is finite in general.

2.5. Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p and let $M \neq 0$ be a finitely generated Cohen–Macaulay R -module. It follows that $\text{crs}_p(M) \leq \text{drs}_p(M)$. If $|k| = \infty$, then $\text{drs}_p(M) \leq \lceil \log_p e_R(M) \rceil$; see, for example, [8, 2.4]. Therefore, we have

$$\begin{aligned} \sup \{ \text{crs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Supp}_R(M) \} &\leq \begin{cases} \max \{ \lceil \log_p e_R(M) \rceil, \text{crs}_p(M) \}, \\ \lceil \log_p e_R(M) \rceil & \text{if } |k| = \infty, \end{cases} \\ \sup \{ \text{drs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Supp}_R(M) \} &\leq \begin{cases} \max \{ \lceil \log_p e_R(M) \rceil, \text{drs}_p(M) \}, \\ \lceil \log_p e_R(M) \rceil & \text{if } |k| = \infty. \end{cases} \end{aligned}$$

2.6. Let $R \rightarrow S$ be a ring homomorphism, M be an S -module, and let X be an R -module.

- (i) $\text{Hom}_R(M, X)$ has an S -module structure defined as follows: For $s \in S$ and $\alpha \in \text{Hom}_R(M, X)$, we set $s \cdot \alpha \in \text{Hom}_R(M, X)$ as $(s \cdot \alpha)(m) = \alpha(sm)$ for all $m \in M$.
- (ii) $\text{Hom}_R(X, M)$ has an S -module structure defined as follows: For $s \in S$ and $\alpha \in \text{Hom}_R(X, M)$, we set $s \cdot \alpha \in \text{Hom}_R(X, M)$ as $(s \cdot \alpha)(x) = s\alpha(x)$ for all $x \in X$.
- (iii) $M \otimes_R X$ has an S -module structure as follows: For $s \in S$ and $\sum_i m_i \otimes x_i \in M \otimes_R X$, with $m_i \in M$ and $x_i \in X$, we define

$$s \cdot \left(\sum_i m_i \otimes x_i \right) = \sum_i (sm_i) \otimes x_i \in M \otimes_R X.$$

The following observation is used in several proofs in the sequel. Note that, over a Noetherian ring, every finitely generated module is finitely presented.

2.7. Let $f : R \rightarrow S$ be a ring homomorphism, A be an R -module, and let B be an S -module. Assume E is an injective S -module. Given $n \geq 0$, the following hold:

- (i) $\text{Hom}_S(\text{Tor}_n^R(A, B), E) \cong \text{Ext}_R^n(A, \text{Hom}_S(B, E))$; see [26, 10.63].
- (ii) If A is finitely presented over R , then

$$\text{Hom}_S(\text{Ext}_R^n(A, B), E) \cong \text{Tor}_n^R(A, \text{Hom}_S(B, E)).$$

(iii) If (R, \mathfrak{m}, k) is local, $R = S$, and $f = \text{id}$, then parts (i) and (ii) imply that

$$\text{fd}_R(B) = \text{id}_R(B^\vee) \quad \text{and} \quad \text{id}_R(B) = \text{fd}_R(B^\vee).$$

2.8 (Auslander–Buchsbaum [2]). Let (R, \mathfrak{m}, k) be a d -dimensional local ring and let N be an R -module. If $\text{id}_R(N) < \infty$, then $\text{id}_R(N) \leq d$. Thus, if $\text{fd}_R(N) < \infty$, then $\text{fd}_R(N) \leq d$; see 2.7(iii).

The next observation is used in the proofs of Lemma 3.8 and Proposition 3.11.

2.9. Let (R, \mathfrak{m}, k) be a d -dimensional local ring and let N be an R -module. Consider a minimal injective resolution of N

$$\mathbf{I} = (0 \rightarrow I^0 \xrightarrow{h_1} I^1 \xrightarrow{h_2} \dots \xrightarrow{h_n} I^n \xrightarrow{h_{n+1}} I^{n+1} \rightarrow \dots),$$

where

$$I^j = \bigoplus_{\mathfrak{p} \in \text{Spec}(R)} E_R(R/\mathfrak{p})^{\oplus \mu_j(\mathfrak{p}, N)} \quad \text{and} \quad \mu_j(\mathfrak{p}, N) = \text{rank}_{k(\mathfrak{p})}(\text{Ext}_{R_{\mathfrak{p}}}^j(k(\mathfrak{p}), N))$$

for $j \geq 0$. Here, $\mu_j(\mathfrak{p}, N)$ is not necessarily finite. Note that $\text{Hom}_{R_{\mathfrak{p}}}(k(\mathfrak{p}), (h_j)_{\mathfrak{p}}) = 0$ for $j \geq 0$. If $\mathfrak{q} \in \text{Spec}(R)$ and we localize \mathbf{I} at \mathfrak{q} , then the resulting $\mathbf{I}_{\mathfrak{q}}$ is a minimal injective resolution of $N_{\mathfrak{q}}$ over $R_{\mathfrak{q}}$ and $I_{\mathfrak{q}}^j = \bigoplus_{\mathfrak{p} \subseteq \mathfrak{q}} E_R(R/\mathfrak{p})^{\oplus \mu_j(\mathfrak{p}, N)}$ for all $j \geq 0$; see, for example, [5] and [15, 3.15 and Appendix 20–24].

Assume $\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < \infty$ for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$. Then $\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) \leq \dim(R_{\mathfrak{p}}) \leq d - 1$ for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$; see 2.8. This implies that $I^i = E_R(k)^{\oplus \mu_i(\mathfrak{m}, N)}$ for all $i \geq d$. Consequently, \mathbf{I} has the form

$$\mathbf{I} = (0 \rightarrow I^0 \rightarrow \dots \rightarrow I^{d-1} \rightarrow E_R(k)^{\oplus \mu_d(\mathfrak{m}, N)} \xrightarrow{h_{d+1}} E_R(k)^{\oplus \mu_{d+1}(\mathfrak{m}, N)} \rightarrow \dots).$$

2.10. Let (R, \mathfrak{m}, k) be a local ring, I be an ideal of R such that $\sqrt{I} = \sqrt{\text{Ann}_R(M)}$, M be a finitely generated R -module, and let N be an R -module. Then:

(i) $\text{Hom}_R(M, N) = 0 \iff \text{grade}_R(I, N) \geq 1 \iff \text{Hom}_R(R/I, N) = 0$; see [6, 1.2.3 or 1.2.10(e)].

(ii) It follows from part (i) and 2.7(i) that

$$\begin{aligned} M \otimes_R N = 0 &\iff \text{Hom}_R(M, N^\vee) = 0 \iff \text{Hom}_R(R/I, N^\vee) = 0 \\ &\iff (R/I \otimes_R N)^\vee = 0 \iff R/I \otimes_R N = 0 \iff N = IN. \end{aligned}$$

(iii) If $\text{Supp}_R(M) = \text{Spec}(R)$ (or equivalently, $I = 0$), it follows from parts (i) and (ii) that

$$\text{Hom}_R(M, N) = 0 \iff M \otimes_R N = 0 \iff N = 0.$$

(iv) $\text{Ass}_R(\text{Hom}_R(M, N)) = \text{Supp}_R(M) \cap \text{Ass}_R(N)$; see, for example, [6, 1.2.28].

Remark 2.11. Assume I , M , and N are as in 2.10, but R is (Noetherian as always) not necessarily local. Then the implications considered in 2.10 still hold since they can all be verified locally. More precisely, it follows that

$$\text{Hom}_R(M, N) = 0 \iff \text{Hom}_R(R/I, N) = 0 \quad \text{and} \quad M \otimes_R N = 0 \iff N = IN.$$

In particular, if $\text{Supp}_R(M) = \text{Spec}(R)$, we have

$$\text{Hom}_R(M, N) = 0 \iff N = 0 \iff M \otimes_R N = 0.$$

These implications establish the fact $\text{Ass}_R(\text{Hom}_R(M, N)) = \text{Supp}_R(M) \cap \text{Ass}_R(N)$, namely the equality stated in 2.10(iv), still holds.

2.12. Let (R, \mathfrak{m}, k) be a local ring and let M be a finitely generated R -module. Consider a minimal free representation $P_1 \xrightarrow{\partial_1} P_0 \rightarrow M \rightarrow 0$. The *transpose* $\text{Tr}_R M$ of M is defined as the cokernel of the R -dual map $\partial_1^* = \text{Hom}_R(\partial_1, R)$. We refer the reader to [1] for the details of the following:

- (i) There is an exact sequence of R -modules $0 \rightarrow M^* \rightarrow P_0^* \rightarrow P_1^* \rightarrow \text{Tr}_R M \rightarrow 0$.
- (ii) It follows that, up to isomorphism, $\text{Tr}_R M$ is uniquely determined.
- (iii) It follows that $\text{Tr}_R(\text{Tr}_R M) \cong M$.
- (iv) M is free if and only if $\text{Tr}_R M$ is free.

The following result from [8] is necessary for our proof of Theorem 1.4.

2.13 [8, 2.2]. Let (R, \mathfrak{m}, k) be a local ring and let M and N be finitely generated R -modules. Assume $n \geq 1$ is an integer. Assume further the following conditions hold:

- (i) $N_{\mathfrak{p}}$ is free for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$.
- (ii) $\text{depth}_R(M \otimes_R N) \geq n$.
- (iii) $\text{depth}_R(M) \geq n - 1$.

Then $\text{Ext}_R^i(\text{Tr}_R N, M) = 0$ for all $i = 1, \dots, n$.

3. A generalization of a result of Funk and Marley

The main results of this section are captured in Theorem 3.2 below, stated earlier as parts (i) and (ii) of Theorem 1.2. They generalize this result already quoted in 1.1:

3.1 (Funk and Marley [11, 3.1 and 3.2]). Let (R, \mathfrak{m}, k) be a d -dimensional Cohen–Macaulay local ring of prime characteristic p , with $d \geq 1$, and let N be an R -module. Given integers $t \geq 1$ and $e \gg 0$, the following hold:

- (i) If $\text{Tor}_i^R({}^e R, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{fd}_R(N) \leq d$.
- (ii) If R is F-finite and $\text{Ext}_R^i({}^e R, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N) \leq d$.

The original statement of 3.1 in [11] includes the case where $d = 0$; in fact this case follows from the techniques of Koh and Lee used in the proof of [16, 2.6]; see also [11, 2.8] and [21, 2.2.8]. Note also that [11, 3.1 and 3.2], namely 3.1, is stated in terms of complexes of R -modules, but its proof naturally reduces to the case of modules. For this reason, we consider only modules when generalizing 3.1 in Theorems 3.2 and 4.3, which can also be extended to the complex case in a similar manner, as explained in the proof of [11, 3.1].

Theorem 3.2. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, the following hold:*

- (i) *If $\text{Tor}_i^R({}^eM, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{fd}_R(N) \leq d$.*
- (ii) *If R is F-finite and $\text{Ext}_R^i({}^eM, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N) \leq d$.*

Remark 3.3. Let R be a ring of prime characteristic p , M be a finitely generated R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Let $e \geq 0$.

- (i) Assume R is F-finite. Then $\text{Hom}_R({}^eM, N) = 0 \implies N = 0$; see 2.10(iii).
- (ii) Without the F-finite assumption, we have ${}^eM \otimes N = 0 \implies N = 0$, as locally this reduces to the F-finite case and then follows from 2.10(iii).

Thus, Theorem 3.2 still holds when $t = 0$. This relies on $\text{Supp}_R({}^eM) = \text{Spec}(R)$ and does not depend on the choice of e or M being Cohen–Macaulay.

Theorem 3.2 generalizes 3.1 in the case where $\text{depth}(R) \geq 1$. For a generalization of 3.1 in the case where $\text{depth}(R) = 0$ (that is, the $d = 0$ case of 3.1), see Proposition 3.10. Before presenting our proof of Theorem 3.2 at the end of the section, we would like to discuss the sharpness of the result and list some corollaries of the theorem. We will also prove Propositions 3.10 and 3.11, which the proof of Theorem 3.2 relies on.

The following example shows that the positive depth assumption on the ring is necessary for Theorem 3.2.

Example 3.4. Let $R = \mathbb{F}_p[[x, y]]/(x^2, xy)$ and let $M = R/(x)$. Then R is an F-finite ring with $\text{depth}(R) = 0$ and $\dim(R) = 1$, and M is a Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$. Moreover, ${}^eM \cong M^{\oplus p^e}$ for all $e \geq 0$.

Let $N = R/(y)$. Then $\text{Tor}_1^R(M, N) = 0$ so that $\text{Tor}_1^R({}^eM, N) = 0$ for all $e \geq 0$. If $\text{pd}_R(N) < \infty$, then N is free since $\text{depth}(R) = 0$. Hence $\text{pd}_R(N) = \text{fd}_R(N) = \infty$. This shows that the positive depth assumption is needed for Theorem 3.2(i).

Next let $N = M$. Then $\text{Ext}_R^1(M, N) = 0$ so that $\text{Ext}_R^1({}^eM, M) = 0$ for all $e \geq 0$. If $\text{id}_R(N) < \infty$, then $\text{id}_R(N) = \text{depth}(R) = 0$, that is, N is injective. Hence, $\text{id}_R(N) = \infty$ (one can also conclude that $\text{id}_R(N) = \infty$ since R is not Cohen–Macaulay). Thus the positive depth assumption is needed for Theorem 3.2(ii).

We give several corollaries of Theorem 3.2. The next corollary covers the particular case where $S = \hat{R}$, the \mathfrak{m} -adic completion of (R, \mathfrak{m}) .

Corollary 3.5. *Let $f : (R, \mathfrak{m}, k) \rightarrow (S, \mathfrak{n}, \ell)$ be a flat local ring homomorphism, where R is a d -dimensional local ring of prime characteristic p and $\mathfrak{m}S = \mathfrak{n}$, M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be a finitely generated S -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, if $\text{Tor}_i^R({}^eM, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{pd}_S(N) \leq d$.*

Proof. Observe $\text{Tor}_i^R(k, N) \cong \text{Tor}_i^S(\ell, N)$ for all $i \geq 0$. By Theorem 3.2, we have that $\text{fd}_R(N) \leq d$. Set $r = \text{fd}_R(N)$. It follows that $\text{Tor}_r^R(k, N) \neq 0 = \text{Tor}_{r+1}^R(k, N)$. Therefore, $\text{pd}_S(N) = r \leq d$. □

Corollary 3.6. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, if $\text{Ext}_R^i({}^eM, N^\vee) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N^\vee) \leq d$.*

Proof. The vanishing of $\text{Ext}_R^i({}^eM, N^\vee)$ yields the vanishing of $\text{Tor}_i^R({}^eM, N)$; see 2.7(i). Thus $\text{id}_R(N^\vee) = \text{fd}_R(N) \leq d$ by Theorems 3.2(i) and 2.7(iii). □

Corollary 3.7. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, assume at least one of the following conditions holds:*

- (i) $\text{Ext}_R^i(N, ({}^eM)^\vee) = 0$ for all $i = t, \dots, t + d - 1$.
- (ii) $\text{Ext}_R^i(N, {}^e(M^\vee)) = 0$ for all $i = t, \dots, t + d - 1$.

Then $\text{fd}_R(N) \leq d$.

Proof. The vanishing of $\text{Ext}_R^i(N, ({}^eM)^\vee)$ yields the vanishing of $\text{Tor}_i^R(N, {}^eM)$; see 2.7(i). Thus, case (i) follows from Theorem 3.2(i).

Similarly, the vanishing of $\text{Ext}_R^i(N, {}^e(M^\vee))$ yields the vanishing of $\text{Tor}_i^R(N, {}^eM)$ by 2.7(i), where $R \rightarrow S$ is the Frobenius F^e with $(S, \mathfrak{n}, \ell) = (R, \mathfrak{m}, k)$, $A = N$ over R , $B = M$ over S , and $E = E_S(\ell)$ is the injective hull of ℓ over S . Therefore, case (ii) also follows from Theorem 3.2(i). □

Next, we prepare some auxiliary results for the proof of Theorem 3.2. To begin with, we present Lemma 3.8 and Corollary 3.9 which are akin to [11, 4.5 and 4.6].

Lemma 3.8. *Let (R, \mathfrak{m}, k) be a d -dimensional F -finite local ring of prime characteristic p , M be a finitely generated R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Set $\delta = \max\{\text{depth}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(R)\}$ and let $t \geq 1$ be an integer. If*

$$\sup\{e \mid \text{Ext}_R^i({}^eM, N) = 0 \text{ for all } i = t, \dots, t + \delta\} \geq \sup\{\text{crs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(R)\},$$

then $\text{id}_R(N) \leq \min\{d, t + \delta - 1\}$. Therefore, if $\text{Ext}_R^i({}^eM, N) = 0$ for all $i = t, \dots, t + \delta$ and for infinitely many e , then $\text{id}_R(N) \leq \min\{d, t + \delta - 1\}$.

Proof. By 2.8, it suffices to show $\text{id}_R(N) \leq t + \delta - 1$, and we proceed by induction on d . The case where $d = 0$ follows from 2.2(2)(iii). Hence, we assume $d \geq 1$.

As R is F-finite, eM is a finitely presented R -module. Thus, for all $\mathfrak{p} \in \text{Spec}(R)$, we have the isomorphisms

$$\text{Ext}_R^i({}^eM, N)_{\mathfrak{p}} \cong \text{Ext}_{R_{\mathfrak{p}}}^i({}^eM_{\mathfrak{p}}, N_{\mathfrak{p}}).$$

So, for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$, the induction hypothesis dictates $\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < t + \delta$. By 2.9, N has a minimal injective resolution of the form

$$I = (0 \rightarrow I^0 \rightarrow \dots \rightarrow I^{t+\delta-1} \rightarrow E_R(k)^{\oplus \mu_{t+\delta}} \rightarrow E_R(k)^{\oplus \mu_{t+\delta+1}} \rightarrow \dots),$$

in which $\mu_i = \mu_i(\mathfrak{m}, N)$ for $i \geq t + \delta$. Set $\text{depth}_R(M) = v$. By assumption, there exists $e \geq \text{crs}_{\mathfrak{p}}(M)$ such that $\text{Ext}_R^i({}^eM, N) = 0$ for $i = t + \delta - v, \dots, t + \delta$. Since $e \geq \text{crs}_{\mathfrak{p}}(M)$, 2.4 yields a maximal M -regular sequence $\underline{x} = \{x_1, \dots, x_v\}$ such that

$$(0 :_{M/(\underline{x})M} \mathfrak{m}^{[p^e]}) \not\subseteq \mathfrak{m}^{[p^e]}(M/(\underline{x})M).$$

Thus, k is a direct summand of ${}^e(M/(\underline{x})M)$; see 2.2. As x_1 is M -regular, there is a short exact sequence $0 \rightarrow {}^eM \rightarrow {}^eM \rightarrow {}^e(M/x_1M) \rightarrow 0$. This, together with $\text{Ext}_R^i({}^eM, N) = 0$ for $i = t + \delta - v, \dots, t + \delta$, implies that

$$\text{Ext}_R^i({}^e(M/x_1M), N) = 0 \quad \text{for all } i = t + \delta - v + 1, \dots, t + \delta.$$

As $\underline{x} = \{x_1, \dots, x_v\}$ is M -regular, inductively we get $\text{Ext}_R^{t+\delta}({}^e(M/(\underline{x})M), N) = 0$. Therefore, $\text{Ext}_R^{t+\delta}(k, N) = 0$ since k is a direct summand of ${}^e(M/(\underline{x})M)$. In view of 2.9, we deduce

$$\mu_{t+\delta} = \text{rank}_k(\text{Ext}_R^{t+\delta}(k, N)) = 0.$$

Therefore, $I^{t+\delta} = E_R(k)^{\oplus \mu_{t+\delta}} = 0$, which concludes that $\text{id}_R(N) \leq t + \delta - 1$. \square

Corollary 3.9. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Let S be a ring extension of R as in 2.3. Set*

$$\delta = \max\{\text{depth}_{S_{\mathfrak{p}}}(M \otimes_R S_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(S)\}$$

and let $t \geq 1$ be an integer. If

$$\sup\{e \mid \text{Tor}_i^R({}^eM, N) = 0 \text{ for all } i = t, \dots, t + \delta\} \geq \sup\{\text{crs}_{\mathfrak{p}}(M \otimes_R S_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(S)\},$$

then $\text{fd}_R(N) \leq d$. In particular, if $\text{Tor}_i^R({}^eM, N) = 0$ for all $i = t, \dots, t + \delta$ and for infinitely many e , then $\text{fd}_R(N) \leq d$.

Proof. It suffices to show $\text{fd}_S(N \otimes_R S) \leq \dim(S)$. Thus, given the assumption, we may assume that $R = S$ is already F-finite. By 2.7(i), the vanishing of $\text{Tor}_i^R({}^eM, N)$ yields vanishing of $\text{Ext}_R^i({}^eM, N^\vee)$. Now the claim follows from Lemma 3.8; see 2.7(iii). \square

As an application of Lemma 3.8 and Corollary 3.9, we obtain the following proposition that generalizes a result of Takahashi and Yoshino [27, 5.3].

Proposition 3.10. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Then, given $t \geq 1$ and $e \gg 0$, the following hold:*

- (i) *If $\text{Tor}_i^R({}^eM, N) = 0$ for all $i = t, \dots, t + d$, then $\text{fd}_R(N) \leq d$.*
- (ii) *If R is F-finite and $\text{Ext}_R^i({}^eM, N) = 0$ for all $i = t, \dots, t + d$, then $\text{id}_R(N) \leq d$.*

Proof. For (i), we may assume that R is F-finite with $|k| = \infty$; see 2.3. In this case, we have $\sup\{\text{crs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(R)\} \leq \lceil \log_p e_R(M) \rceil < \infty$. The rest follows from Corollary 3.9.

For (ii), without the assumption that $|k| = \infty$, we have

$$\sup\{\text{crs}_p(M_{\mathfrak{p}}) \mid \mathfrak{p} \in \text{Spec}(R)\} \leq \max\{\lceil \log_p e_R(M) \rceil, \text{crs}_p(M)\} < \infty.$$

The rest follows from Lemma 3.8. \square

Our proof of Proposition 3.11 is inspired by some of the techniques employed by Funk and Marley in the proof of [11, 3.2]. To distinguish various module structures in the proof, we present 3.11 in the context of a general ring homomorphism $f : R \rightarrow S$. Subsequently, in the proof of Theorem 3.2, we apply Proposition 3.11 to $F^e : R \rightarrow R$, the iterated Frobenius endomorphism. Recall that, over a local ring (R, \mathfrak{m}, k) , we set $(-)^{\vee} = \text{Hom}_R(-, E_R(k))$.

Proposition 3.11. *Let $f : (R, \mathfrak{m}, k) \rightarrow (S, \mathfrak{n}, \ell)$ be a module-finite local homomorphism of local rings, with $d = \dim(R) \geq 1$, M be an S -module, and let N be an R -module. Assume the following:*

- (i) *$\text{Ext}_R^i(M, N) = 0$ for all $i = t, \dots, t + d - 1$ for some $t \geq 1$.*
- (ii) *M is a finitely generated S -module such that $\text{Supp}_R(M) = \text{Spec}(R)$.*
- (iii) *There exists $\underline{x} = \{x_1, \dots, x_d\} \subseteq \mathfrak{n}$ such that \underline{x} is M -regular and $\mathfrak{m}M \subseteq (\underline{x})M$.*
- (iv) *$\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < \infty$ for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$.*
- (v) *$\text{fd}_{R_{\mathfrak{p}}}((N^{\vee})_{\mathfrak{p}}) < \infty$ for all $\mathfrak{p} \in \text{Ass}(R)$.*

Then $\text{id}_R(N) \leq d$.

Proof. As $\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < \infty$ for each $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$, 2.9 dictates that N has a minimal injective resolution of the form

$$\mathbf{I} = (0 \rightarrow I^0 \xrightarrow{h_1} \dots \rightarrow I^{d-1} \xrightarrow{h_d} E_R(k)^{\oplus \mu_d(\mathfrak{m}, N)} \rightarrow \dots \xrightarrow{h_{t+d}} E_R(k)^{\oplus \mu_{t+d}(\mathfrak{m}, N)} \rightarrow \dots).$$

We apply $\text{Hom}_R(M, -)$ to \mathbf{I} and use our assumption (i) to obtain an exact sequence

$$(3.11.1) \quad M_{t-1} \xrightarrow{g_t} \dots \xrightarrow{g_{t+d-2}} M_{t+d-2} \xrightarrow{g_{t+d-1}} M_{t+d-1} \xrightarrow{g_{t+d}} M_{t+d} \xrightarrow{\rho} C \rightarrow 0,$$

where $M_i = \text{Hom}_R(M, I^i)$, $g_i = \text{Hom}_R(M, h_i)$ and $C = \text{coker}(g_{t+d})$. Note that each M_i is an S -module; see 2.6(i). Hence, C is an S -module as well.

Claim 1. *The induced sequence $\overline{M_{t+d-1}} \xrightarrow{\overline{g_{t+d}}} \overline{M_{t+d}} \xrightarrow{\overline{\rho}} \overline{C} \rightarrow 0$, in which*

$$\overline{M}_i = \text{Hom}_R(M/(\underline{\mathbf{x}})M, I^i), \quad \overline{g}_{t+d} = \text{Hom}_R(M/(\underline{\mathbf{x}})M, h_{t+d})$$

and $\overline{C} = \text{Hom}_S(S/(\underline{\mathbf{x}}), C)$ is exact.

Proof. Since I^i is injective over R and x_1 is regular on M , an application of $\text{Hom}_R(-, I^i)$ to the exact sequence $0 \rightarrow M \xrightarrow{x_1} M \rightarrow M/x_1M \rightarrow 0$ induces a short exact sequence

$$(\Lambda_i) \quad 0 \longleftarrow M_i \xleftarrow{x_1} M_i \longleftarrow \text{Hom}_R(M/x_1M, I^i) \longleftarrow 0.$$

We combine the short exact sequences (Λ_i) , for all $i = t - 1, \dots, t + d$, with the exact sequence (3.11.1), and obtain the exact sequence of S -modules

$$\text{Hom}_R(M/x_1M, I^t) \rightarrow \dots \rightarrow \text{Hom}_R(M/x_1M, I^{t+d}) \rightarrow \text{Hom}_S(S/x_1S, C) \rightarrow 0.$$

Inductively, as $\underline{\mathbf{x}} = \{x_1, \dots, x_d\}$ is an M -regular sequence, we realize the following exact sequence of S -modules that is naturally induced from (3.11.1), as claimed:

$$\text{Hom}_R(M/(\underline{\mathbf{x}})M, I^{t+d-1}) \xrightarrow{\overline{g}_{t+d}} \text{Hom}_R(M/(\underline{\mathbf{x}})M, I^{t+d}) \xrightarrow{\overline{\rho}} \text{Hom}_S(S/(\underline{\mathbf{x}})S, C) \rightarrow 0.$$

Claim 2. $\ker(\rho) = 0$.

Proof. The assumption $\mathfrak{m}M \subseteq (\underline{\mathbf{x}})M$ says that, as an R -module, $M/(\underline{\mathbf{x}})M$ is a direct sum of copies of k . Since \mathbf{I} is a minimal injective resolution of N , we see that

$$\overline{g}_{t+d} = \text{Hom}_R(M/(\underline{\mathbf{x}})M, h_{t+d}) = 0.$$

In view of the exact sequence in Claim 1, we get $\ker(\overline{\rho}) = 0$. Moreover, $\overline{\rho}$ can be identified as $\text{Hom}_S(S/(\underline{\mathbf{x}}), \rho)$ up to the natural isomorphism $\text{Hom}_S(S/(\underline{\mathbf{x}}), M_i) \cong \text{Hom}_R(M/(\underline{\mathbf{x}})M, I^i)$. Hence $(0 :_{\ker(\rho)}(\underline{\mathbf{x}})) \cong \ker(\overline{\rho}) = 0$. Given that M is finitely generated over R , we have

$$\text{Ass}_R(\ker(\rho)) \subseteq \text{Ass}_R(M_{t+d}) \subseteq \text{Ass}_R(E_R(k)^{\oplus \mu_{t+d}(\mathfrak{m}, N)}) \subseteq \{\mathfrak{m}\}.$$

Thus $\text{Ass}_S(\ker(\rho)) \subseteq \{\mathfrak{n}\}$, as \mathfrak{n} is the only prime ideal of S lying over \mathfrak{m} . So, if $\ker(\rho) \neq 0$, then $\mathfrak{n} \in \text{Ass}_S(\ker(\rho))$ and hence $0 \neq (0 :_{\ker(\rho)} \mathfrak{n}) \subseteq (0 :_{\ker(\rho)} (\underline{\mathbf{x}}))$, which contradicts the conclusion $(0 :_{\ker(\rho)} (\underline{\mathbf{x}})) = 0$ above. This completes the proof of the claim.

Now that we know $\ker(\rho) = 0$, the exact sequence (3.11.1) forces $g_{t+d} = 0$, which gives rise to an exact sequence

$$\text{Hom}_R(M, I^{t+d-2}) \xrightarrow{g_{t+d-1}} \text{Hom}_R(M, I^{t+d-1}) \longrightarrow 0.$$

Since M is finitely presented over R , we apply $(-)^{\vee}$ to the exact sequence above and obtain the following exact sequence in light of 2.7(ii):

$$(3.11.2) \quad M \otimes_R (I^{t+d-2})^{\vee} \xleftarrow{1 \otimes h_{t+d-1}^{\vee}} M \otimes_R (I^{t+d-1})^{\vee} \longleftarrow 0.$$

Next, let us return to the injective resolution I of N . Consider the exact sequence

$$0 \rightarrow N \rightarrow I^0 \xrightarrow{h_1} \dots \rightarrow I^{t+d-2} \xrightarrow{h_{t+d-1}} I^{t+d-1} \xrightarrow{\theta} D \rightarrow 0,$$

where D is the cokernel of the map h_{t+d-1} . Applying $(-)^{\vee}$ to the exact sequence above, we get an exact sequence

$$(3.11.3) \quad 0 \leftarrow N^{\vee} \leftarrow (I^0)^{\vee} \leftarrow \dots \leftarrow (I^{t+d-2})^{\vee} \xleftarrow{h_{t+d-1}^{\vee}} (I^{t+d-1})^{\vee} \xleftarrow{\theta^{\vee}} D^{\vee} \leftarrow 0$$

with each $(I^i)^{\vee}$ flat over R ; see 2.7(iii). It follows that

$$\text{Ass}_R(D^{\vee}) \subseteq \text{Ass}_R((I^{t+d-1})^{\vee}) \subseteq \text{Ass}(R).$$

Claim 3. $(D^{\vee})_{\mathfrak{p}} = 0$ for all $\mathfrak{p} \in \text{Ass}(R)$.

Proof. Fix any $\mathfrak{p} \in \text{Ass}(R)$. From assumption (v) and 2.8, we see that

$$\text{fd}_{R_{\mathfrak{p}}}((N^{\vee})_{\mathfrak{p}}) \leq \dim(R_{\mathfrak{p}}) \leq d.$$

Thus, (3.11.3) localized at \mathfrak{p} gives rise to a flat resolution of $N_{\mathfrak{p}}$ over $R_{\mathfrak{p}}$, which can be used to compute $\text{Tor}_i^{R_{\mathfrak{p}}}(-, (N^{\vee})_{\mathfrak{p}})$. By 2.7(ii), we have

$$\text{Tor}_{t+d-1}^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (N^{\vee})_{\mathfrak{p}}) = (\text{Tor}_{t+d-1}^R(M, N^{\vee}))_{\mathfrak{p}} \cong (\text{Ext}_R^{t+d-1}(M, N^{\vee}))_{\mathfrak{p}} = 0.$$

Moreover, we have $\text{Tor}_{t+d}^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (N^{\vee})_{\mathfrak{p}}) = 0$ since $t+d > \text{fd}_{R_{\mathfrak{p}}}((N^{\vee})_{\mathfrak{p}})$. Next, we apply $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} -$ to (3.11.3) localized at \mathfrak{p} . The vanishing of $\text{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (N^{\vee})_{\mathfrak{p}})$, for $i = t+d-1$ and $i = t+d$, forces the exact sequence

$$M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} ((I^{t+d-2})^{\vee})_{\mathfrak{p}} \xleftarrow{1 \otimes (h_{t+d-1}^{\vee})_{\mathfrak{p}}} M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} ((I^{t+d-1})^{\vee})_{\mathfrak{p}} \xleftarrow{1 \otimes (\theta^{\vee})_{\mathfrak{p}}} M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} (D^{\vee})_{\mathfrak{p}} \leftarrow 0.$$

Comparing this with (3.11.2) localized at \mathfrak{p} , we see $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} (D^{\vee})_{\mathfrak{p}} = 0$. Hence $(D^{\vee})_{\mathfrak{p}} = 0$ since $\text{Supp}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = \text{Spec}(R_{\mathfrak{p}})$; see 2.10. This completes the proof of the claim. \square

Overall, we have $\text{Ass}_R(D^{\vee}) \subseteq \text{Ass}(R)$ and $(D^{\vee})_{\mathfrak{p}} = 0$ for all $\mathfrak{p} \in \text{Ass}(R)$. This forces $D^{\vee} = 0$, which implies $D = 0$. Therefore, $\text{id}_R(N) < t + d$, so $\text{id}_R(N) \leq d$ by 2.8. \square

We are now ready to prove Theorem 3.2, a consequence of Propositions 3.10 and 3.11. Recall that $F : R \rightarrow R$ denotes the Frobenius endomorphism.

Proof of Theorem 3.2. It suffices to prove part (i) for the case where R is F-finite with $|k| = \infty$; see 2.3. Thus, we can obtain part (i) from part (ii) via duality; see 2.7(iii).

To prove part (ii), note that eM is a finitely generated (and hence a finitely presented) R -module with $\text{Supp}_R({}^eM) = \text{Supp}_R(M) = \text{Spec}(R)$. Since e is sufficiently large, there exists an M -regular sequence $\underline{x} = \{x_1, \dots, x_d\} \subseteq \mathfrak{m}$ such that $\mathfrak{m}({}^eM) \subseteq {}^e(\underline{x})M$; see 2.4. In light of 2.7(ii), the vanishing of $\text{Ext}_R^i({}^eM, N)$ implies the vanishing of $\text{Tor}_i^R({}^eM, N^{\vee})$ for all $i = t, \dots, t + d - 1$. Therefore, $\text{id}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < \infty$ and $\text{fd}_{R_{\mathfrak{p}}}((N^{\vee})_{\mathfrak{p}}) < \infty$ for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$; see Proposition 3.10. Note that $\text{Ass}(R) \subseteq \text{Spec}(R) - \{\mathfrak{m}\}$ since $\text{depth}(R) \geq 1$. Now we apply Proposition 3.11, with $R = S$ and $f = F^e$, and deduce that $\text{id}_R(N) \leq d$. \square

Remark 3.12. We conclude this section by pointing out some lower bounds for the integer e that ensure the validity of the proofs of certain previously stated results.

- (a) It is enough to assume $e \geq \lceil \log_p e_R(M) \rceil$ in part (i) of both Theorem 3.2 and Proposition 3.10.
- (b) When $|k| = \infty$, it is enough to assume $e \geq \lceil \log_p e_R(M) \rceil$ in part (ii) of both Theorem 3.2 and Proposition 3.10. When $|k| < \infty$, it is enough to assume $e \geq \max \{ \lceil \log_p e_R(M) \rceil, \text{drs}_p(M) \}$ in Theorem 3.2. When $|k| < \infty$, it is enough to assume $e \geq \max \{ \lceil \log_p e_R(M) \rceil, \text{crs}_p(M) \}$ in Proposition 3.10.
- (c) It is enough to assume $e \geq \lceil \log_p e_R(M) \rceil$ in Corollaries 3.5, 3.6, and 3.7.

It is proved in [14, 2.17] that, if R is an excellent ring and M is a finitely generated R -module, then the set $\sup \{ e_{R_{\mathfrak{m}}}(M_{\mathfrak{m}}) : \mathfrak{m} \in \text{Max}(R) \}$ is finite; see [14, 2.17]. We use this fact and state a global version of Theorem 3.2.

Theorem 3.13. *Let R be a d -dimensional ring of prime characteristic p , with $d \geq 1$, M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be an R -module. Assume $\text{depth}(R_{\mathfrak{m}}) \geq \min \{ 1, \dim(R_{\mathfrak{m}}) \}$ for each maximal ideal \mathfrak{m} of R .*

- (i) Assume R is excellent, $s = \sup \{e_{R_m}(M_m) : m \in \text{Max}(R)\}$, and let e be an integer such that $e \geq \lceil \log_p s \rceil$. Given $t \geq 1$, if $\text{Tor}_i^R({}^e M, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{fd}_R(N) \leq d$.
- (ii) Assume R is F -finite and the residue field of R_m is infinite for each maximal ideal m of R . Let e be an integer such that

$$e \geq \sup \{ \lceil \log_p e_{R_m}(M_m) \rceil, \text{drs}_p(M_m) : m \in \text{Max}(R) \}.$$

Given $t \geq 1$, if $\text{Ext}_R^i({}^e M, N) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{id}_R(N) \leq d$.

Proof. Note that $s < \infty$ due to [14, 2.17]. In proving the first part, we have that

$$\text{Tor}_i^{R_m}({}^e M_m, N_m) = 0 \quad \text{for all } i = t, \dots, t + d - 1.$$

If $\dim(R_m) = 0$, then the residue field of R_m is a direct summand of ${}^e M_m$, and hence $\text{fd}_{R_m}(N_m) \leq \dim(R_m) \leq d$; see 2.2(2)(i). On the other hand, if $\dim(R_m) \geq 1$, then part (i) of Theorem 3.2 yields $\text{fd}_{R_m}(N_m) \leq d$; see Remark 3.12(a). This implies that $\text{fd}_R(N) \leq d$ as flat dimension can be computed locally. We can prove the second part similarly since $\sup \{ \lceil \log_p e_{R_m}(M_m) \rceil, \text{drs}_p(M_m) : m \in \text{Max}(R) \}$ is finite under our setting; see Remark 3.12(a). □

4. On the homological properties of the Frobenius endomorphism

The aim of this section is to prove Theorem 1.2(iii) and to establish Theorem 1.4. Let us point out that Theorem 1.4 generalizes [8, 1.3]; see 1.3.

The layout of this section is as follows: We begin by preparing some auxiliary results. Then we establish Theorem 1.2(iii) in Theorem 4.3. Finally, making use of 2.13 and Theorem 4.3, we produce a proof of Theorem 1.4. Along the way, we also discuss the sharpness of our results; see Examples 4.8 and 4.9.

The first auxiliary result, namely Proposition 4.1, is akin to [16, 2.6] and [22, Theorem A]. Even though it suffices to apply the proposition to the identity map $1_R : R \rightarrow R$ in the sequel, we present it more generally in terms of a ring homomorphism $f : R \rightarrow S$.

Proposition 4.1. *Let $f : (R, m, k) \rightarrow (S, n, \ell)$ be a local homomorphism of local rings of prime characteristic p , $M \neq 0$ be a finitely generated R -module, and let N be a finitely generated S -module. Set $d = \dim(R)$ and $v = \text{depth}_R(M)$. Given $t \geq 1$ and $e \geq \text{crr}_p(M)$, if $\text{Ext}_R^i(N, {}^e M) = 0$ for all $i = t, \dots, t + v$, then $\text{fd}_R(N) \leq t - 1$, and thus $\text{pd}_R(N) \leq t - 1 + d$.*

Proof. By our choice of e , there exists a maximal M -regular sequence $\underline{x} = \{x_1, \dots, x_v\}$ such that k is a direct summand of ${}^e(M/(\underline{x})M)$ over R ; see 2.2 and 2.4.

As x_1 is M -regular, there is a short exact sequence $0 \rightarrow {}^e M \rightarrow {}^e M \rightarrow {}^e(M/x_1 M) \rightarrow 0$. This, together with $\text{Ext}_R^i(N, {}^e M) = 0$ for $i = t, \dots, t + v$, implies that

$$\text{Ext}_R^i(N, {}^e(M/x_1M)) = 0 \quad \text{for all } i = t, \dots, t + v - 1.$$

Inductively, as $\underline{x} = \{x_1, \dots, x_v\}$ is M -regular, we get $\text{Ext}_R^t(N, {}^e(M/(\underline{x})M)) = 0$. Since k is a direct summand of ${}^e(M/(\underline{x})M)$, we see that $\text{Ext}_R^t(N, k) = 0$, which implies $\text{fd}_R(N) \leq t - 1$; see [22, 2.1]. Note that every flat R -module has projective dimension at most d ; see [28, 4.2.8]. Therefore, we conclude that $\text{pd}_R(N) \leq t - 1 + d$. \square

As in Propositions 3.11 and 4.1, we present Proposition 4.2 in the context of a general ring homomorphism $f : R \rightarrow S$, allowing us to distinguish various module structures in the proof. When Proposition 4.2 is applied in the proof of Theorem 4.3, the homomorphism will be $F^e : R \rightarrow R$, the e -th iteration of the Frobenius endomorphism.

Proposition 4.2. *Let (R, \mathfrak{m}, k) be a local ring, $f : R \rightarrow S$ be a ring homomorphism, N be a finitely generated R -module, and let M be a finitely generated S -module. Assume the following hold:*

- (i) $\text{Ext}_R^i(N, M) = 0$ for all $i = t, \dots, t + d - 1$ for some $d \geq 1$ and $t \geq 1$.
- (ii) *There exists $\underline{x} = \{x_1, \dots, x_d\} \subseteq \text{Jac}(S)$ such that \underline{x} is M -regular and $\mathfrak{m}M \subseteq (\underline{x})M$.*

Then $\text{Hom}_R(\Omega_R^t(N), M) = 0$.

Proof. Consider a minimal free resolution of N over R :

$$F = (\cdots \rightarrow F_{t+d} \xrightarrow{h_{t+d}} F_{t+d-1} \xrightarrow{h_{t+d-1}} \cdots \xrightarrow{h_{t+1}} F_t \xrightarrow{h_t} F_{t-1} \rightarrow \cdots \rightarrow 0).$$

Applying $\text{Hom}_R(-, M)$ to F and using assumption (i), we get an induced exact sequence

$$(4.2.1) \quad M_{t+d} \xleftarrow{g_{t+d}} M_{t+d-1} \xleftarrow{g_{t+d-1}} \cdots \xleftarrow{g_{t+1}} M_t \xleftarrow{g_t} M_{t-1} \xleftarrow{l} G \leftarrow 0,$$

in which $M_i = \text{Hom}_R(F_i, M)$, $g_i = \text{Hom}_R(h_i, M)$, and $G = \ker(g_t)$. All M_i , and hence G , are S -modules; see 2.6(ii). In fact, each M_i is isomorphic to a finite direct sum of M . Thus all M_i are finitely generated S -modules.

As F_t is free over R and x_1 is regular on M , an application of $\text{Hom}_R(F_t, -)$ to the exact sequence $0 \rightarrow M \xrightarrow{x_1} M \rightarrow M/x_1M \rightarrow 0$ induces a short exact sequence

$$(\Gamma_i) \quad 0 \longrightarrow M_i \xrightarrow{x_1} M_i \longrightarrow \text{Hom}_R(F_t, M/x_1M) \longrightarrow 0.$$

We combine the exact sequences (Γ_i) , for all $i = t - 1, \dots, t + d$, with the exact sequence (4.2.1) and obtain the following exact sequence of S -modules:

$$\text{Hom}_R(F_{t+d-1}, M/x_1M) \leftarrow \cdots \leftarrow \text{Hom}_R(F_{t-1}, M/x_1M) \leftarrow G/x_1G \leftarrow 0.$$

Inductively, as $\underline{x} = \{x_1, \dots, x_d\}$ is an M -regular sequence, we realize the following exact sequence of S -modules that is naturally induced from (4.2.1):

$$(4.2.2) \quad \overline{M}_t \xleftarrow{\overline{g}_t} \overline{M}_{t-1} \xleftarrow{i} \overline{G} \longleftarrow 0,$$

in which $\overline{M}_i = \text{Hom}_R(F_i, M/(\underline{x})M)$, $\overline{g}_t = \text{Hom}_R(h_t, M/(\underline{x})M)$ and $\overline{G} = G/(\underline{x})G$. Up to isomorphism, we may write

$$\overline{M}_i = M_i/(\underline{x})M_i = M_i \otimes_S S/(\underline{x})S, \quad \overline{G} = G \otimes_S S/(\underline{x})S$$

and thus $i = \iota \otimes 1_{S/(\underline{x})S}$.

The assumption $\mathfrak{m}M \subseteq (\underline{x})M$ implies that $M/(\underline{x})M$ is annihilated by \mathfrak{m} . Since F is a minimal free resolution of N , we conclude $\overline{g}_t = \text{Hom}_R(h_t, M/(\underline{x})M) = 0$. Thus, the exactness of (4.2.2) forces $\text{im}(i) = \ker(\overline{g}_t) = \overline{M}_{t-1}$, meaning that $\text{im}(i) + (\underline{x})M_{t-1} = M_{t-1}$. As M_{t-1} is finitely generated over S and $(\underline{x})S \subseteq \text{Jac}(S)$, we obtain $\text{im}(i) = M_{t-1}$, thanks to Nakayama's lemma. This forces $\ker(g_{t+1}) = \text{im}(g_t) = 0$ due to the exactness of (4.2.1). Finally, as $\Omega_R^t(N) \cong \text{coker}(h_{t+1})$, we see that

$$\text{Hom}_R(\Omega_R^t(N), M) \cong \ker(\text{Hom}_R(h_{t+1}, M)) = \ker(g_{t+1}) = 0. \quad \square$$

Equipped with Propositions 4.1 and 4.2, we are now ready to prove the result stated in Theorem 1.2(iii).

Theorem 4.3. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be a finitely generated R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, if $\text{Ext}_R^i(N, {}^eM) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{pd}_R(N) \leq \min\{t - 1, \text{depth}(R)\}$.*

Proof. Since e is sufficiently large, there exists $\underline{x} = \{x_1, \dots, x_d\} \subseteq \mathfrak{m}$ such that \underline{x} is M -regular and $\mathfrak{m}({}^eM) \subseteq {}^e(\underline{x})M$; see 2.4. Upon an application of Proposition 4.2 to the map $F^e : R \rightarrow R$, we see that the vanishing of $\text{Ext}_R^i(N, {}^eM)$ for $i = t, \dots, t + d - 1$ implies $\text{Hom}_R(\Omega_R^t(N), {}^eM) = 0$.

To prove the claim by contradiction, suppose that $\Omega_R^t(N) \neq 0$ and select \mathfrak{p} in $\text{Ass}_R(\Omega_R^t(N)) \subseteq \text{Ass}(R)$; then $\mathfrak{p} \neq \mathfrak{m}$ since $\text{depth}(R) \geq 1$. It follows from Proposition 4.1 that $\text{pd}_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) < \infty$. Hence, $N_{\mathfrak{p}}$ is free by the Auslander–Buchsbaum formula. So, $\Omega_{R_{\mathfrak{p}}}^t(N)_{\mathfrak{p}} \neq 0$ is free over $R_{\mathfrak{p}}$. Also, since $\text{Supp}_R({}^eM) = \text{Supp}_R(M) = \text{Spec}(R)$, we see that $({}^eM)_{\mathfrak{p}} \neq 0$. Hence $\text{Hom}_{R_{\mathfrak{p}}}(\Omega_{R_{\mathfrak{p}}}^t(N)_{\mathfrak{p}}, ({}^eM)_{\mathfrak{p}}) \neq 0$, which contradicts the conclusion $\text{Hom}_R(\Omega_R^t(N), {}^eM) = 0$. Thus $\Omega_R^t(N) = 0$, so $\text{pd}_R(N) \leq t - 1$. This proves that $\text{pd}_R(N) \leq \min\{t - 1, \text{depth}(R)\}$. \square

We now record several corollaries of Theorem 4.3. It is worth noting that the positive depth assumption in the theorem is necessary; see Example 3.4.

Corollary 4.4. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be a finitely generated R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1$ and $e \gg 0$, assume that at least one of the following conditions holds:*

- (i) $\text{Tor}_i^R(N, ({}^eM)^\vee) = 0$ for all $i = t, \dots, t + d - 1$.
- (ii) $\text{Tor}_i^R(N, {}^e(M^\vee)) = 0$ for all $i = t, \dots, t + d - 1$.

Then $\text{pd}_R(N) \leq \min\{t - 1, \text{depth}(R)\}$.

Proof. The vanishing of $\text{Tor}_i^R(N, ({}^eM)^\vee)$ yields the vanishing of $\text{Ext}_R^i(N, {}^eM)$; see 2.7(ii). Thus, case (i) follows from Theorem 4.3. Similarly, the vanishing of $\text{Tor}_i^R(N, {}^e(M^\vee))$ yields the vanishing of $\text{Ext}_R^i(N, {}^eM)$ by 2.7(ii) where $R \rightarrow S$ is the Frobenius F^e with $(S, \mathfrak{n}, \ell) = (R, \mathfrak{m}, k)$, $A = N$ over R , $B = M$ over S , and $E = E_S(\ell)$ is the injective hull of ℓ over S . Therefore, case (ii) also follows from Theorem 4.3. □

Next, we provide a proof of Theorem 1.4, restated here as Corollary 4.5 for convenience. Before doing so, we present a few facts needed for the argument.

Recall that $F_R^e(-)$ denotes the scalar extension along $F^e : R \rightarrow R$, the e -th iteration of the Frobenius endomorphism. Given R -modules X and Y , it follows that

$$X \otimes_R {}^eY \cong {}^e(F_R^e(X) \otimes_R Y) \quad \text{and} \quad F_{R_p}^e(X_p) \cong (F_R^e(X))_p \quad \text{for all } p \in \text{Spec}(R).$$

Moreover, we have $\text{Supp}_R(X) = \text{Supp}_R(F_R^e(X))$. Although the finitely generated case suffices for our argument in Corollary 4.5, we briefly discuss the general situation where X may not be finitely generated. To establish the equality $\text{Supp}_R(X) = \text{Supp}_R(F_R^e(X))$, it is enough to assume R is local. If $X \neq 0$ is finitely generated, then $F_R^e(X) \neq 0$ via a surjection $X \twoheadrightarrow k$. In the general case, pick a flat local ring homomorphism $\varphi : R \rightarrow A$, where A is F-finite; such a ring map always exists by 2.3. Setting $S = R$ and $B = A$, and letting $F_A^e : A \rightarrow B$ denote the e -th iteration of the Frobenius of A , we see that the composition

$$\varphi \circ F_R^e : R \xrightarrow{F_R^e} S \xrightarrow{\varphi} B$$

agrees with $F_A^e \circ \varphi : R \xrightarrow{\varphi} A \xrightarrow{F_A^e} B$. Thus

$$(X \otimes_R S) = 0 \implies (X \otimes_R S) \otimes_S B = 0 \implies (X \otimes_R A) \otimes_A B = 0 \implies X \otimes_R A = 0 \implies X = 0.$$

Here, the second implication follows since $\varphi \circ F_R^e = F_A^e \circ \varphi$, the third implication follows from 2.10(iii) as A is F-finite, while the fourth one holds because φ is faithfully flat. This establishes what we want to show. We also refer the reader to Marley’s result in [19, 2.1] for an alternative proof of $\text{Supp}_R(X) = \text{Supp}_R(F_R^e(X))$.

Corollary 4.5. *Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p , and let M and N be finitely generated R -modules. Assume the following conditions hold:*

- (i) M is Cohen–Macaulay and $\text{Supp}_R(M) = \text{Spec}(R)$.
- (ii) N is generically free, that is, $N_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$ -module for all $\mathfrak{p} \in \text{Ass}(R)$.

If $F_R^e(N) \otimes_R M$ is a maximal Cohen–Macaulay R -module for some $e \gg 0$, then N is free.

Proof. Without loss of generality we may assume R is F-finite (and also $|k| = \infty$); see 2.3. Set $d = \dim(R)$ and proceed by induction on d .

As R is F-finite, the assumption that $F_R^e(N) \otimes_R M$ is maximal Cohen–Macaulay can be interpreted as that the module $N \otimes_R {}^eM$ is maximal Cohen–Macaulay. If $\text{depth}(R) = 0$, then N is free since we assume it is generically free. Hence, we assume $\text{depth}(R) \geq 1$. Note that, by the induction hypothesis, we may assume $N_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$ -module for all $\mathfrak{p} \in \text{Spec}(R) - \{\mathfrak{m}\}$. Note also that $\text{depth}_R({}^eM) = d = \text{depth}_R(N \otimes_R {}^eM)$. Therefore, we use 2.13 for the case where $n = d$ and conclude that $\text{Ext}_R^i(\text{Tr}_R N, {}^eM) = 0$ for all $i = 1, \dots, d$. Now, Theorem 4.3 implies that $\text{Tr}_R N$ is free. Consequently, N is free; see 2.12(iv). \square

Corollary 4.6. *Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p , and let M and N be finitely generated R -modules. Assume the following conditions hold:*

- (i) M is Cohen–Macaulay.
- (ii) M has constant rank on $\text{Min}(R)$ and N is generically free.

If $F_R^e(N) \otimes_R M$ is maximal Cohen–Macaulay for some $e \gg 0$, then N is free.

Proof. Suppose $\text{Min}(R) \not\subseteq \text{Supp}_R(M)$. Then $\text{Supp}_R(M) \cap \text{Min}(R) = \emptyset$ since M has constant rank on $\text{Min}(R)$. This implies that $\dim_R(M) < \dim(R)$. However, this is not possible as $\dim(R) = \dim_R(F_R^e(N) \otimes_R M) \leq \dim_R(M)$. So, $\text{Min}(R) \subseteq \text{Supp}_R(M)$ and hence $\text{Supp}_R(M) = \text{Spec}(R)$. Now, the claim follows from Corollary 4.5. \square

Corollary 4.7. *Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module, and let N be a finitely generated R -module. Assume:*

- (i) For each $\mathfrak{p} \in \text{Min}(R)$, there is an integer $r_{\mathfrak{p}} \geq 1$ such that $M_{\mathfrak{p}} \cong R_{\mathfrak{p}}^{\oplus r_{\mathfrak{p}}}$.
- (ii) N is generically free and $\dim_R(N) = \dim(R)$.

If $F_R^e(N) \otimes_R M$ is Cohen–Macaulay for some $e \gg 0$, then N is free.

Proof. It follows that $\dim_R(F_R^e(N) \otimes_R M) = \dim_R(F_R^e(N)) = \dim_R(N) = \dim(R)$ because $\text{Supp}_R(M) = \text{Spec}(R)$. Now, the claim follows from Corollary 4.5. \square

We illustrate the sharpness of Theorem 1.4 with some examples. Specifically, Example 4.8 demonstrates that the theorem’s conclusion may fail if the module M in question does not have full support. Similarly, Example 4.9 highlights the necessity of the assumption that M is generically free for the theorem to hold.

Example 4.8. Let $R = \mathbb{F}_p[[x, y]]/(xy)$, $M = R/(x)$, and $N = M$. Note that M is maximal Cohen–Macaulay, and $N_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$ -module for all $\mathfrak{p} \in \text{Ass}(R)$. It follows that $N \otimes_R^e M$ is maximal Cohen–Macaulay for all $e \geq 0$ because $N \otimes_R^e M \cong N \otimes_R M^{\oplus p^e} \cong M^{\oplus p^e}$. However, N is not free. Here $\text{Spec}(R) \ni (y) \notin \text{Supp}_R(M)$. This example also shows that the constant rank assumption in Corollary 4.6 is necessary. As a separate observation, $\text{Ext}_R^1(M, M) = 0$. Thus, the full support assumption is necessary in Theorem 4.3.

Example 4.9. Let $R = \mathbb{F}_p[[x, y]]/(x^2, xy)$, $M = R/(x)$, and $N = M$. Note that M is Cohen–Macaulay, $\text{Supp}_R(M) = \text{Spec}(R)$, and $N_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$ -module for all $\mathfrak{p} \in \text{Min}(R)$. It follows that $N \otimes_R^e M$ is maximal Cohen–Macaulay for all $e \geq 0$ since $N \otimes_R^e M \cong M^{\oplus p^e}$. However, N is not free. Here we have that $\text{Ass}(R) \ni \mathfrak{m} \notin \text{Min}(R)$.

It is proved in [4, 1.1] that, if there is a finitely generated module $N \neq 0$ over a local ring R of prime characteristic p such that $\text{fd}_R({}^rN) < \infty$ or $\text{id}_R({}^rN) < \infty$ for some $r \geq 1$, then R is regular. We use this fact and obtain the following consequences of Theorem 1.2; cf. [29, 6.8 and 6.10].

Corollary 4.10. *Let (R, \mathfrak{m}, k) be a d -dimensional local ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let $N \neq 0$ be a finitely generated R -module. Assume $\text{depth}(R) \geq 1$. Given $t \geq 1, r \geq 1$ and $e \gg 0$, we further assume that at least one of the following conditions holds:*

- (i) $\text{Tor}_i^R({}^eM, {}^rN) = 0$ for all $i = t, \dots, t + d - 1$.
- (ii) R is F -finite and $\text{Ext}_R^i({}^eM, {}^rN) = 0$ for all $i = t, \dots, t + d - 1$.
- (iii) rN is finitely generated over R and $\text{Ext}_R^i({}^rN, {}^eM) = 0$ for all $i = t, \dots, t + d - 1$.

Then R is regular.

Proof. In view of Theorem 1.2, the claims follow from [4, 1.1]. □

Remark 4.11. We point out some lower bounds for the integer e that ensure the validity of the proofs of certain previously stated results.

- (a) It is enough to assume $e \geq \lceil \log_p e_R(M) \rceil$ in Corollaries 1.5 and 4.5.
- (b) It is enough to assume $e \geq \lceil \log_p e_R(M) \rceil$ in both Theorem 4.3 and Corollary 4.4. This is because we may use 2.3 and assume $|k| = \infty$ in the proofs of these results.

- (c) It is enough to assume $e \geq \max \{ \lceil \log_p e_R(M) \rceil, \text{drs}_p(M) \}$ in Corollary 4.10 in general. If $|k| = \infty$, however, it suffices to assume $e \geq \lceil \log_p e_R(M) \rceil$ in Corollary 4.10.

We do not know whether the Cohen–Macaulay assumption on M is necessary in Theorem 1.4. This raises the following question:

Question 4.12. Let (R, \mathfrak{m}, k) be a local ring of prime characteristic p and let M and N be finitely generated R -modules. Assume $\text{Supp}_R(M) = \text{Spec}(R)$ and N is generically free. If $F_R^e(N) \otimes_R M$ is maximal Cohen–Macaulay for some $e \gg 0$, then must N be free? What if R is Cohen–Macaulay, or $N = M$?

We conclude this section with a global version of Theorem 4.3; we skip its proof as it is similar to that of Theorem 3.13.

Theorem 4.13. *Let R be a d -dimensional excellent ring of prime characteristic p , M be a finitely generated Cohen–Macaulay R -module such that $\text{Supp}_R(M) = \text{Spec}(R)$, and let N be a finitely generated R -module. Assume $\text{depth}(R_{\mathfrak{m}}) \geq \min\{1, \dim(R_{\mathfrak{m}})\}$ for each maximal ideal \mathfrak{m} of R . Set*

$$s = \sup \{ e_{R_{\mathfrak{m}}}(M_{\mathfrak{m}}) : \mathfrak{m} \in \text{Max}(R) \},$$

and let e be an integer such that $e \geq \lceil \log_p s \rceil$. Given $t \geq 1$, if $\text{Ext}_R^i(N, {}^e M) = 0$ for all $i = t, \dots, t + d - 1$, then $\text{pd}_R(N) \leq \min\{t - 1, \text{depth}(R)\}$.

Acknowledgements

Part of this work was completed during Celikbas’s visit to the Department of Mathematics and Statistics at Georgia State University in February 2018, and during the visit of Celikbas, Sadeghi, and Yao to the Nesin Mathematics Village in İzmir, Türkiye, in May 2019.

The authors thank these two institutions for their hospitality. They also thank Mohsen Asgarzadeh for his comments on a previous version of the manuscript. Finally, the authors thank the referee for a careful reading of the manuscript and for helpful comments.

References

- [1] M. Auslander and M. Bridger, *Stable module theory*, Memoirs of the American Mathematical Society **94**, American Mathematical Society, 1969. MR
- [2] M. Auslander and D. A. Buchsbaum, “Homological dimension in noetherian rings, II”, *Trans. Amer. Math. Soc.* **88** (1958), 194–206. MR
- [3] L. L. Avramov and C. Miller, “Frobenius powers of complete intersections”, *Math. Res. Lett.* **8**:1-2 (2001), 225–232. MR
- [4] L. L. Avramov, M. Hochster, S. B. Iyengar, and Y. Yao, “Homological invariants of modules over contracting endomorphisms”, *Math. Ann.* **353**:2 (2012), 275–291. MR

- [5] H. Bass, “On the ubiquity of Gorenstein rings”, *Math. Z.* **82** (1963), 8–28. MR
- [6] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1993. MR
- [7] L. Burch, “On ideals of finite homological dimension in local rings”, *Proc. Cambridge Philos. Soc.* **64** (1968), 941–948. MR
- [8] O. Celikbas, A. Sadeghi, and Y. Yao, “Tensoring with the Frobenius endomorphism”, *Homology Homotopy Appl.* **20**:1 (2018), 251–257. MR
- [9] T. H. Funk, *Frobenius and homological dimensions of complexes*, PhD thesis, University of Nebraska – Lincoln, 2021, available at <https://digitalcommons.unl.edu/mathstudent/107>.
- [10] T. Funk and T. Marley, “Correction to: Frobenius and homological dimensions of complexes”, *Collect. Math.* **71**:2 (2020), 299–300. MR
- [11] T. Funk and T. Marley, “Frobenius and homological dimensions of complexes”, *Collect. Math.* **71**:2 (2020), 287–297. MR
- [12] J. Herzog, “Ringe der Charakteristik p und Frobeniusfunktoren”, *Math. Z.* **140** (1974), 67–78. MR
- [13] M. Hochster, “Cyclic purity versus purity in excellent Noetherian rings”, *Trans. Amer. Math. Soc.* **231**:2 (1977), 463–488. MR
- [14] M. Hochster and Y. Yao, “Frobenius splitting, strong F-regularity, and small Cohen–Macaulay modules”, *Trans. Amer. Math. Soc.* **376**:9 (2023), 6729–6765. MR
- [15] S. B. Iyengar, G. J. Leuschke, A. Leykin, C. Miller, E. Miller, A. K. Singh, and U. Walther, *Twenty-four hours of local cohomology*, Graduate Studies in Mathematics **87**, American Mathematical Society, 2007. MR
- [16] J. Koh and K. Lee, “Some restrictions on the maps in minimal resolutions”, *J. Algebra* **202**:2 (1998), 671–689. MR
- [17] E. Kunz, “Characterizations of regular local rings of characteristic p ”, *Amer. J. Math.* **91** (1969), 772–784. MR
- [18] J. Li, “Frobenius criteria of freeness and Gorensteinness”, *Arch. Math. (Basel)* **98**:6 (2012), 499–506. MR
- [19] T. Marley, “The Frobenius functor and injective modules”, *Proc. Amer. Math. Soc.* **142**:6 (2014), 1911–1923. MR
- [20] H. Matsumura, *Commutative ring theory*, 2nd ed., Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, 1989. MR
- [21] C. Miller, “The Frobenius endomorphism and homological dimensions”, pp. 207–234 in *Commutative algebra* (Grenoble and Lyon, 2001), Contemp. Math. **331**, Amer. Math. Soc., 2003. MR
- [22] S. Nasseh, M. Tousi, and S. Yassemi, “Characterization of modules of finite projective dimension via Frobenius functors”, *Manuscripta Math.* **130**:4 (2009), 425–431. MR
- [23] C. Peskine and L. Szpiro, “Sur la topologie des sous-schémas fermés d’un schéma localement noethérien, définis comme support d’un faisceau cohérent localement de dimension projective finie”, *C. R. Acad. Sci. Paris Sér. A-B* **269** (1969), A49–A51. MR
- [24] C. Peskine and L. Szpiro, “Dimension projective finie et cohomologie locale. Applications à la démonstration de conjectures de M. Auslander, H. Bass et A. Grothendieck”, *Inst. Hautes Études Sci. Publ. Math.* **42** (1973), 47–119. MR
- [25] A. G. Rodicio, “On a result of Avramov”, *Manuscripta Math.* **62**:2 (1988), 181–185. MR

- [26] J. J. Rotman, *An introduction to homological algebra*, 2nd ed., Springer, 2009. MR
- [27] R. Takahashi and Y. Yoshino, “Characterizing Cohen-Macaulay local rings by Frobenius maps”, *Proc. Amer. Math. Soc.* **132**:11 (2004), 3177–3187. MR
- [28] J. Xu, *Flat covers of modules*, Lecture Notes in Mathematics **1634**, Springer, 1996. MR
- [29] M. R. Zargar, O. Celikbas, M. Gheibi, and A. Sadeghi, “Homological dimensions of rigid modules”, *Kyoto J. Math.* **58**:3 (2018), 639–669. MR

Received March 13, 2025. Revised January 31, 2026.

OLGUR CELIKBAS
SCHOOL OF MATHEMATICAL AND DATA SCIENCES
WEST VIRGINIA UNIVERSITY
MORGANTOWN, WV
UNITED STATES
olgur.celikbas@math.wvu.edu

ARASH SADEGHI
DOKHANIAT 49179-66686
GORGAN
IRAN
sadeghiarash61@gmail.com

YONGWEI YAO
DEPARTMENT OF MATHEMATICS AND STATISTICS
GEORGIA STATE UNIVERSITY
ATLANTA, GA
UNITED STATES
yyao@gsu.edu

SYMPLECTIC AUTOMORPHISMS OF A SURFACE WITH GENUS TWO FIBRATION AND THEIR ACTION ON CH_0

JIABIN DU AND WENFEI LIU

Let S be a complex smooth projective surface with a genus two fibration, and $\mathrm{Aut}_s(S)$ the group of symplectic automorphisms, fixing every holomorphic 2-forms (if any) on S . Based on the work of Jin-Xing Cai, we show that, if $\chi(\mathcal{O}_S) \geq 5$, then $|\mathrm{Aut}_s(S)| \leq 2$. Then we verify, under some conditions, that $\mathrm{Aut}_s(S)$ acts trivially on the Albanese kernel $\mathrm{CH}_0(S)_{\mathrm{alb}}$ of the 0-th Chow group, which is predicted by a conjecture of Bloch and Beilinson. As a consequence, if an automorphism $\sigma \in \mathrm{Aut}(S)$ acts trivially on $H^{i,0}(S)$ for $0 \leq i \leq 2$, then it also acts trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$.

1. Introduction	259
2. Preliminaries	263
2.1. Fibered surfaces and their automorphisms	263
2.2. The induced action on the Albanese variety and the 0-th Chow group	264
2.3. Useful criteria for a symplectic automorphism to act trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$	266
3. Symplectic automorphisms of surfaces with a genus two fibration	267
Acknowledgements	272
References	272

1. Introduction

Throughout the paper, we work over the complex numbers field \mathbb{C} .

Let X be a n -dimensional connected smooth projective variety over \mathbb{C} . When studying the automorphism group $\mathrm{Aut}(X)$, it is natural to look at the induced action of $\mathrm{Aut}(X)$ on the cohomology groups that are naturally attached to X , such as $H^*(X, R)$ with R an abelian group or $H^*(X, \Omega_X^p)$ for $0 \leq p \leq \dim X$, where Ω_X^p is the coherent sheaf of holomorphic p -forms on X . On the other hand, there is also an induced action of $\mathrm{Aut}(X)$ on the Chow groups $\mathrm{CH}_*(X)$, which is a more refined invariant than the cohomology groups. A deep conjecture of Bloch and Beilinson predicts that, roughly speaking, there is a natural decreasing filtration

MSC2020: 14J50, 14J29, 14C15.

Keywords: surface of general type, fibration of genus two, symplectic automorphism, Chow group, Bloch–Beilinson conjecture.

$F^\bullet \text{CH}_*(X)_\mathbb{Q}$ on the $\text{CH}_*(X)_\mathbb{Q} := \text{CH}_*(X) \otimes_{\mathbb{Z}} \mathbb{Q}$ such that the induced action of $\text{Aut}(X)$ on the graded pieces of $F^\bullet \text{CH}_*(X)_\mathbb{Q}$ is determined by the action of $\text{Aut}(X)$ on the cohomology groups $H^*(X, \mathbb{Q})$. We refer to [Voi03, Section 11.2.2] for a precise statement of the Bloch–Beilinson conjecture.

For a smooth projective surface S , it is clear that $\text{CH}_2(S) \cong \mathbb{Z}$ and $\text{CH}_1(S) \cong \text{Pic}(S)$. Thus the focus is on the still mysterious 0-th Chow group $\text{CH}_0(S)$. There is a natural filtration on $\text{CH}_0(S)$

$$(1-1) \quad 0 \subset \text{CH}_0(S)_{\text{alb}} \subset \text{CH}_0(S)_{\text{hom}} \subset \text{CH}_0(S)$$

where $\text{CH}_0(S)_{\text{hom}}$ is the kernel of the degree map $\text{deg} : \text{CH}_0(S) \rightarrow \mathbb{Z}$, and $\text{CH}_0(S)_{\text{alb}}$ is the kernel of the Albanese map $\text{alb}_S : \text{CH}_0(S)_{\text{hom}} \rightarrow \text{Alb}(S)$. A famous result of Mumford [Mum68] says, if the geometric genus $p_g(S) > 0$, then there is no uniform integer $d > 0$ such that any $\alpha \in \text{CH}_0(S)_{\text{hom}}$ can be written as $\alpha = \alpha^+ - \alpha^- \in \text{CH}_0(S)$ with both of α^+ and α^- effective and $\text{deg } \alpha^+ = \text{deg } \alpha^- = d$. It is thus legitimate to call $\text{CH}_0(S)_{\text{hom}}$ and also the Albanese kernel $\text{CH}_0(S)_{\text{alb}}$ infinite-dimensional.

The filtration (1-1) is supposed to be the filtration in the aforementioned Bloch–Beilinson conjecture, and as a consequence, we have:

Conjecture 1.1 [Voi03, Conjecture 11.19]. *Let S and T be smooth projective surfaces, Γ a cycle of codimension 2 in $S \times T$ such that the map $[\Gamma]^* : H^{2,0}(T) \rightarrow H^{2,0}(S)$ vanishes. Then the map $\Gamma_* : \text{CH}_0(S)_{\text{alb}} \rightarrow \text{CH}_0(T)_{\text{alb}}$ vanishes.*

By taking $T = S$ with $p_g(S) = 0$, and $\Gamma = \Delta_S \subset S \times S$ the diagonal in Conjecture 1.1, we recover Bloch’s initial conjecture.

Conjecture 1.2 ([Blo75]; see also [Voi03, Conjecture 11.2]). *Let S be a smooth projective surface. If $p_g(S) = 0$ then $\text{CH}_0(S)_{\text{alb}} = 0$.*

Bloch’s conjecture has been verified in various special cases (see [DL23, page 444] for a discussion), but is widely open in general.

We are interested in the induced action of automorphisms on $\text{CH}_0(S)$. So, take $T = S$, and $\Gamma := \Gamma_\sigma - \Delta_S$, where Γ_σ is the graph of an automorphism σ in a subgroup $G \subset \text{Aut}(S)$, and we obtain from Conjecture 1.1 the following:

Conjecture 1.3 (cf. [Voi12, Conjecture 1.2]). *Let S be a smooth projective surface, and G a group of automorphisms of S acting trivially on $H^{2,0}(S)$. Then G acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*

Define the group $\text{Aut}_s(S)$ of *symplectic automorphisms*

$$\text{Aut}_s(S) := \{\sigma \in \text{Aut}(S) \mid \sigma \text{ induces the trivial action on } H^{2,0}(S) = H^0(S, K_S)\}.$$

and the group $\text{Aut}_{\mathcal{O}}(S)$ of *\mathcal{O} -cohomologically trivial automorphisms*

$$\text{Aut}_{\mathcal{O}}(S) := \{\sigma \in \text{Aut}(S) \mid \sigma \text{ induces the trivial action on } H^i(S, \mathcal{O}_S) \text{ for any } i\}.$$

By Serre duality, we have $H^0(S, K_S) \cong H^2(S, \mathcal{O}_S)^\vee$, and hence $\text{Aut}_{\mathcal{O}}(S) \subset \text{Aut}_s(S)$; the two groups coincide if the irregularity $q(S) := \dim H^1(S, \mathcal{O}_S) = 0$. The automorphisms in $\text{Aut}_s(S)$ are those fixing a general holomorphic 2-form (if any) on S , and hence called symplectic.

Conjecture 1.3 amounts to saying that $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$.

For surfaces with $p_g(S) > 0$, Conjecture 1.3 is known to hold in certain cases:

- S is an abelian surface and $G = \text{Aut}_s(S)$ [BKL76, Theorems A.1 and A.7; Paw19, Corollary 1.5];
- S is a K3 surface, and $G \subset \text{Aut}_s(S)$ is finite [Voi12; Huy12];
- $S = K(A)$ is the Kummer K3 surface associated to an abelian surface A , and G is generated by $\sigma \in \text{Aut}_s(S)$ that lifts to a group automorphism $\tilde{\sigma}$ of A [Paw19, Theorem 1.10];
- S is a K3 surface with an elliptic fibration $f : S \rightarrow B$, and G preserves the fibration structure f [DL23, Theorem 1.5];
- S is a K3 surface with either the Picard number $\rho(S) \geq 3$ or S admitting a Jacobian fibration, and $G = \text{Aut}_s(S)$ [LYZ24, Theorem 1.3];
- S has Kodaira dimension one and $G = \text{Aut}_{\mathcal{O}}(S)$ [DL23, Theorem 5.10];
- S has Kodaira dimension one with $(p_g, q) \notin \{(1, 1), (2, 2)\}$, and $G = \text{Aut}_s(S)$ [DL23, Theorem 1.3];
- S has a finite-dimensional Chow motive in the sense of Kimura [Kim05] (this is the case if S has an isotrivial fibration), and $G \subset \text{Aut}_s(S)$ is finite [DL23, Lemma 5.9].

In this paper, we investigate Conjecture 1.3 for surfaces of general type with a genus two fibration. Recall that a fibration of genus g on a smooth projective surface S means a morphism $f : S \rightarrow B$ onto a smooth projective curve B with connected fibers of genus g .

Theorem 1.4. *Let S be a surface of general type with a genus two fibration $f : S \rightarrow B$ and $\chi(\mathcal{O}_S) \geq 5$. Suppose that $\text{Aut}_s(S)$ is nontrivial.*

- (i) $|\text{Aut}_s(S)| = 2$.
- (ii) *If the canonical map $\phi_{K_S} : S \dashrightarrow \mathbb{P}^{p_g(S)-1}$ is composed with a pencil, then $\text{Aut}_s(S)$ preserves every fiber of f , and acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*
- (iii) *If the canonical map $\phi_{K_S} : S \dashrightarrow \mathbb{P}^{p_g(S)-1}$ is generically finite onto its image $T := \phi_{K_S}(S)$, then $\text{Aut}_{\mathcal{O}}(S)$ is trivial, and $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$ unless*
 - (a) $q(S) = g(B) \geq \chi(\mathcal{O}_S) - 3$,

- (b) *a smooth model of the quotient surface $S/\langle\sigma\tau\rangle$ is of general type with $p_g = q = 0$, where σ is the generator of $\text{Aut}_s(S)$ and τ is the hyperelliptic fibration of f , and*
- (c) *f is not isotrivial.*

The bound $|\text{Aut}_s(S)| \leq 2$ of Theorem 1.4(i) is implicit in Jin-Xing Cai's work [Cai06a; Cai06b] on the automorphism group H of genus two fibrations acting trivially on $H^2(S, \mathbb{Q})$. Specifically, by Hodge decomposition, one has $H \subset \text{Aut}_s(S)$, and his proof that $|H| \leq 2$ uses only this fact.

Cai also constructed various fibered surfaces of genus two with an involution acting trivially on $H^2(S, \mathbb{Q})$ in [Cai06a; Cai06b; Cai07]. Only one series of them, namely [Cai06b, Example 3.5], satisfies all of the conditions (a)–(c) of Theorem 1.4, and we do not know whether or not $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$ for S in this example.

Theorem 1.4 follows from Propositions 3.4 and 3.7 and Corollary 2.4. Let us explain the ideas of the proofs. Following [Cai06a; Cai06b], the proof of Theorem 1.4(i) is based on the results of Xiao on genus two fibrations [Xiao85]. Specifically, let $G \subset \text{Aut}(S)$ be the subgroup generated by $\text{Aut}_s(S)$ together with the hyperelliptic involution τ . Then the canonical map ϕ_{K_S} of S factors through the quotient map $S \rightarrow S/G$ (Lemma 3.1), and the explicit bounds for ϕ_{K_S} of Xiao (Theorem 3.3) give the bound $|G| \leq 4$ and hence $|\text{Aut}_s(S)| \leq 2$.

Once the bound on $|\text{Aut}_s(S)|$ is established, we have

$$G = \{\text{id}_S, \sigma, \tau, \sigma\tau\} \cong (\mathbb{Z}/2\mathbb{Z})^2.$$

We can then decompose $\text{CH}_0(S)_{\text{alb}, \mathbb{Q}}$ into eigenspaces with respect to the G -action, and see that $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$ as soon as $\text{CH}_0(S/\langle\sigma\tau\rangle)_{\text{alb}} = 0$ (Lemma 2.6). Since $p_g(S/\langle\sigma\tau\rangle) = 0$, the Albanese kernel $\text{CH}_0(S/\langle\sigma\tau\rangle)_{\text{alb}}$ vanishes if (a smooth model of) the quotient surface $S/\langle\sigma\tau\rangle$ is not of general type by [BKL76]. The bulk of our arguments is devoted to verify the latter condition, once the surface $f : S \rightarrow B$ is not isotrivial but violates one of the conditions (a) and (b) in Theorem 1.4(iii). The case of isotrivial fibrations has been settled once and for all by applying Kimura's finite-dimensionality of such surfaces (Corollary 2.4).

Assuming Conjecture 1.2 holds, then $\text{CH}_0(S/\langle\sigma\tau\rangle)_{\text{alb}} = 0$, and we are done again (Proposition 3.8).

Concerning the condition $\chi(\mathcal{O}_S) \geq 5$ in the theorem, we remark that minimal surfaces of general type with $\chi(\mathcal{O}_S) < 5$ form (only) a bounded family. This condition is used mainly to ensure that the canonical map of S is well-behaved and that S has at most one genus two fibration on it.

Finally, as a consequence of Theorem 1.4, Conjecture 1.3 holds for the subgroup $\text{Aut}_{\mathcal{O}}(S) \subset \text{Aut}_s(S)$ for surfaces of general type with a genus two fibration, whose invariants are not so small.

Corollary 1.5. *Let S be a surface of general type with a genus two fibration $f : S \rightarrow B$ and $\chi(\mathcal{O}_S) \geq 5$. Then $\text{Aut}_{\mathcal{O}}(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*

As pointed out by a referee, since there is an $\text{Aut}_{\mathcal{O}}(S)$ -fixed point $x \in S$, $\text{Aut}_{\mathcal{O}}(S)$ even acts trivially on the whole $\text{CH}_0(S)$, provided that $q(S) = 0$.

Outline. We recall in Section 2 some relevant notions and facts such as fibration-preserving automorphisms, the induced action on the Albanese variety and the 0-th Chow group, as well as useful criteria for the triviality of the induced action on the Albanese kernel. In Section 3, we focus on genus two fibrations, first using the canonical map to bound the symplectic automorphism group $\text{Aut}_s(S)$, and then verifying the triviality of its induced action on the Albanese kernel as stated in Theorem 1.4.

Notation and conventions. Let S be smooth projective surface over \mathbb{C} .

- For a coherent sheaf \mathcal{F} on S , we denote $h^i(S, \mathcal{F}) := \dim_{\mathbb{C}} H^i(S, \mathcal{F})$, and $\mathcal{F}^{\vee} := \text{Hom}_{\mathcal{O}_S}(\mathcal{F}, \mathcal{O}_S)$, the dual of \mathcal{F} .
- For $0 \leq i \leq 2$, Ω_S^i denotes the sheaf of i -forms on S .
- ω_S denotes the canonical sheaf of S , which can be identified with Ω_S^2 , and K_S denotes a canonical divisor of S .
- The following numerical invariants are attached to S :
 - the *geometric genus* $p_g(S) := h^0(S, \omega_S)$;
 - the *irregularity* $q(S) := h^1(S, \mathcal{O}_S) = h^0(S, \Omega_S^1)$;
 - the *holomorphic Euler characteristic* $\chi(\mathcal{O}_S) = 1 - q(S) + p_g(S)$.

For a singular projective surface T , say with rational singularities, its geometric genus $p_g(T)$, irregularity $q(T)$ are defined as the corresponding invariants of its smooth model.

For a finite group G and an element $\sigma \in G$, their orders are denoted by $|G|$ and $|\sigma|$ respectively. An action of G on a set X is called *trivial* if $\sigma(x) = x$ for any $\sigma \in G$ and any $x \in X$.

2. Preliminaries

2.1. Fibered surfaces and their automorphisms. Let S be a normal projective surface, and $f : S \rightarrow B$ a fibration onto a smooth projective curve, that is, f is a surjective morphism with connected fibers. The fibration is called

- *of genus g* if its general fiber has genus g ;
- *hyperelliptic* (resp. *elliptic*, resp. *a \mathbb{P}^1 -fibration*) if its smooth fibers are hyperelliptic (resp. elliptic, resp. \mathbb{P}^1);
- *isotrivial* if its smooth fibers are mutually isomorphic.

We call $q_f := q(S) - g(B)$ the *relative irregularity* of f .

The subgroup $\text{Aut}_f(S)$ of *fibration-preserving automorphisms* is defined as follows:

$$\text{Aut}_f(S) := \{\sigma \in \text{Aut}(S) \mid \sigma \text{ maps fibers of } f \text{ to fibers}\}$$

There is an induced action of $\text{Aut}_f(S)$ on B , that is, a homomorphism $r : \text{Aut}_f(S) \rightarrow \text{Aut}(B)$, such that for any $\sigma \in \text{Aut}_f(S)$ there is an induced automorphism $\sigma_B = r(\sigma) \in \text{Aut}(B)$ such that the following diagram is commutative:

$$\begin{array}{ccc} S & \xrightarrow{\sigma} & S \\ f \downarrow & & \downarrow f \\ B & \xrightarrow{\sigma_B} & B \end{array}$$

The elements of $\text{Aut}_B(S) := \ker r$ are called *fiber-preserving automorphisms*.

Lemma 2.1. *Let $f : S \rightarrow B$ be a non-isotrivial fibration of genus $g \geq 2$. Then $\text{Aut}_f(S)$ is a finite group.*

Proof. Consider the moduli map associated to f

$$\mu : B \rightarrow \overline{\mathcal{M}}_g,$$

where $\overline{\mathcal{M}}_g$ is the (compact) moduli space of stable curves of genus g . Since f is not isotrivial, μ is generically finite, and the image $\text{Aut}_f(S)|_B$ of $r : \text{Aut}_f(S) \rightarrow \text{Aut}(B)$ has order at most $\deg \mu$. It follows that

(2-1)

$$|\text{Aut}_f(S)| = |\text{Aut}_B(S)| \cdot |\text{Aut}_f(S)|_B \leq |\text{Aut}(F)| \cdot |\text{Aut}_f(S)|_B \leq |\text{Aut}(F)| \cdot \deg \mu$$

where F is a general fiber of f . Since $g(F) = g \geq 2$, $\text{Aut}(F)$ is a finite group, and hence $\text{Aut}_f(S)$ is also a finite group by (2-1). □

2.2. The induced action on the Albanese variety and the 0-th Chow group. The Chow group of a normal projective variety X is the group of rational equivalence classes of algebraic cycles on X ; we refer to [Voi03, Chapter 9] for the basic properties of Chow groups.

Let X be a smooth projective variety. Fixing a base point $x_0 \in X$, we may define a group homomorphism

$$\text{alb} : \text{CH}_0(X)_{\text{hom}} \rightarrow \text{Alb}(X) = H^0(X, \Omega_X^1)^\vee / H_1(X, \mathbb{Z}), \quad \sum_i n_i [x_i] \mapsto \sum_i n_i \left[\int_{x_0}^{x_i} \right],$$

where $\int_{x_0}^{x_i} : H^0(X, \Omega_X^1) \rightarrow \mathbb{C}$ maps a holomorphic 1-form η to the integral $\int_{x_0}^{x_i} \eta$, determined up to $\int_\gamma \eta$ for a closed loop γ on X . The homomorphism alb does not depend on the choice of the base point x_0 .

Any automorphism group $G \subset \text{Aut}(X)$ has an induced action on $\text{CH}_0(X)$ and $\text{Alb}(X)$ as follows: For $\sigma \in G$, $\sum_i n_i [x_i] \in \text{CH}_0(X)$, and $[\int_{x_0}^x] \in \text{Alb}(X)$,

$$\sigma_*\left(\sum_i n_i [x_i]\right) = \sum_i n_i [\sigma(x_i)], \quad \sigma_*\left([\int_{x_0}^x]\right) = \left[\int_{\sigma(x_0)}^{\sigma(x)}\right] = \left[\int_{x_0}^{\sigma(x)}\right] - \left[\int_{x_0}^{\sigma(x_0)}\right].$$

Note that G acts by group automorphisms on $\text{CH}_0(X)$ and $\text{Alb}(X)$, and the homomorphism alb is G -equivariant. The G -action extends in an obvious way to $\text{CH}_0(X)_{\mathbb{Q}} := \text{CH}_0(X) \otimes_{\mathbb{Z}} \mathbb{Q}$ and $\text{Alb}(X)_{\mathbb{Q}} := \text{Alb}(X) \otimes_{\mathbb{Z}} \mathbb{Q}$.

Let Z be a normal projective birational model of X , with at most rational singularities. Then we have natural identifications (see [JLZ23, Section 2.1] for a discussion for the Albanese variety of a singular variety):

$$\text{Alb}(Z) \cong \text{Alb}(X), \quad \text{CH}_0(Z) \cong \text{CH}_0(X).$$

The following basic but important fact about the G -actions on $\text{Alb}(X)_{\mathbb{Q}}$ and $\text{CH}_0(X)_{\mathbb{Q}}$ should be well-known, but we write a proof of it for completeness (cf. [Lat21, Lemma 1.6]).

Lemma 2.2. *Let X be a smooth projective variety and $G \subset \text{Aut}(X)$ a finite subgroup. Let Y be a normal projective birational model with rational singularities of the quotient variety X/G . Then*

$$\text{Alb}(X)_{\mathbb{Q}}^G \cong \text{Alb}(Y)_{\mathbb{Q}}, \quad \text{CH}_0(X)_{\text{alb}, \mathbb{Q}}^G \cong \text{CH}_0(Y)_{\text{alb}, \mathbb{Q}}, \quad \text{CH}_0(X)_{\text{hom}, \mathbb{Q}}^G \cong \text{CH}_0(Y)_{\text{hom}, \mathbb{Q}}.$$

Proof. Since X/G and Y have at most rational singularities, there are natural identifications

$$\begin{aligned} \text{Alb}(X/G) &= \text{Alb}(Y), \\ \text{CH}_0(X/G)_{\text{hom}, \mathbb{Q}} &\cong \text{CH}_0(Y)_{\text{hom}, \mathbb{Q}}, \\ \text{CH}_0(X/G)_{\text{alb}, \mathbb{Q}} &\cong \text{CH}_0(Y)_{\text{alb}, \mathbb{Q}}. \end{aligned}$$

Thus we may assume that $Y = X/G$.

Recall that

$$\text{Alb}(X)_{\mathbb{Q}} = \frac{H^0(X, \Omega_X^1)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}}{H_1(X, \mathbb{Q})},$$

and hence we have natural isomorphisms of \mathbb{Q} -vector spaces

$$(2-2) \quad \text{Alb}(X)_{\mathbb{Q}}^G \cong \frac{(H^0(X, \Omega_X^1)^G)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}}{H_1(X, \mathbb{Q})^G} \cong \frac{H^0(Y, \Omega_Y^1)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}}{H_1(Y, \mathbb{Q})} \cong \text{Alb}(X/G)_{\mathbb{Q}},$$

where Y is a smooth projective model of X/G .

By the universal property of the Albanese morphism and by the flatness of \mathbb{Q} as a \mathbb{Z} -module, the quotient map $\pi : X \rightarrow X/G$ induces the following commutative

diagram with exact rows:

$$\begin{array}{ccccccc}
 0 & \longrightarrow & \mathrm{CH}_0(X)_{\mathrm{alb}, \mathbb{Q}} & \longrightarrow & \mathrm{CH}_0(X, \mathbb{Q})_{\mathrm{hom}, \mathbb{Q}} & \longrightarrow & \mathrm{Alb}(X)_{\mathbb{Q}} \longrightarrow 0 \\
 & & \downarrow \pi_* & & \downarrow \pi_* & & \downarrow \pi_* \\
 0 & \longrightarrow & \mathrm{CH}_0(X/G)_{\mathrm{alb}, \mathbb{Q}} & \longrightarrow & \mathrm{CH}_0(X/G)_{\mathrm{hom}, \mathbb{Q}} & \longrightarrow & \mathrm{Alb}(X/G)_{\mathbb{Q}} \longrightarrow 0
 \end{array}$$

Restricting to the G -invariant parts of the \mathbb{Q} -vector spaces in the first row, we obtain the following commutative diagram with exact rows:

$$\begin{array}{ccccccc}
 0 & \longrightarrow & \mathrm{CH}_0(X)_{\mathrm{alb}, \mathbb{Q}}^G & \longrightarrow & \mathrm{CH}_0(X, \mathbb{Q})_{\mathrm{hom}, \mathbb{Q}}^G & \longrightarrow & \mathrm{Alb}(X)_{\mathbb{Q}}^G \longrightarrow 0 \\
 & & \downarrow \alpha & & \downarrow \beta & & \downarrow \gamma \\
 0 & \longrightarrow & \mathrm{CH}_0(X/G)_{\mathrm{alb}, \mathbb{Q}} & \longrightarrow & \mathrm{CH}_0(X/G)_{\mathrm{hom}, \mathbb{Q}} & \longrightarrow & \mathrm{Alb}(X/G)_{\mathbb{Q}} \longrightarrow 0
 \end{array}$$

and we have seen in (2-2) that γ is an isomorphism.

For a point $x \in X$, let \bar{x} be its image in X/G . Any 0-cycle on X/G is rationally equivalent to one with support outside the branch locus \mathcal{B} of π [Voi12, Fact 3.3]. For $[z] \in \mathrm{CH}_0(X/G)_{\mathrm{hom}, \mathbb{Q}}$, we may thus assume that $z = \sum_i n_i \bar{x}_i$ with $\bar{x}_i \notin \mathcal{B}$ for any i , and one sees that β has an inverse

$$\beta^{-1}([z]) = \frac{1}{|G|} \sum_i \sum_{g \in G} [g(x_i)].$$

Therefore, β is an isomorphism.

By the five lemma, α is also an isomorphism. □

2.3. Useful criteria for a symplectic automorphism to act trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$.

Lemma 2.3 (cf. [DL23, Lemma 5.9]). *Let S be a smooth projective surface. Assume that the Chow motive $\mathfrak{h}(S)$ of S is finite-dimensional in the sense of Kimura [Kim05] and O’Sullivan. Then any symplectic automorphism of finite order of S acts trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$.*

In particular:

Corollary 2.4. *Let $f : S \rightarrow B$ be an isotrivially fibered surface. Then any finite order symplectic automorphism of S acts trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$.*

Proof. Since f is an isotrivial fibration, S is birational to $(\tilde{B} \times F)/G$, where \tilde{B} is a smooth projective curve dominating B and F is a general fiber of f . Therefore, S is dominated by $\tilde{B} \times F$ and hence has finite-dimensional Chow motive by [Kim05]. Now apply Lemma 2.3. □

Corollary 2.5. *Let $f : S \rightarrow B$ be a fibered surface of genus g such that $q_f = g$. Then any finite order symplectic automorphism of S acts trivially on $\mathrm{CH}_0(S)_{\mathrm{alb}}$.*

Proof. Since $g = q_f$, the fibration f is isotrivial by [Bea82], and hence the assertion follows from Corollary 2.4. \square

If a surface S has several commuting involutions, then we have the following criterion for the triviality of the $\text{Aut}_s(S)$ -action on the Albanese kernel.

Lemma 2.6. *Let S be a smooth projective surface. Let $G = \langle \sigma, \tau \rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$ be a subgroup of $\text{Aut}(S)$. Suppose that $\text{CH}_0(S/\langle \tau \rangle)_{\text{alb}} = 0$. Then σ acts trivially on $\text{CH}_0(S)_{\text{alb}}$ if and only if $\text{CH}_0(S/\langle \sigma\tau \rangle)_{\text{alb}} = 0$.*

Proof. Since $\text{CH}_0(S)_{\text{alb}}$ has no torsion [Roj80], we may work on the Albanese kernels with \mathbb{Q} -coefficients. Since $\text{CH}_0(S/\langle \tau \rangle)_{\text{alb}} = 0$, τ acts as -1 on $\text{CH}_0(S)_{\text{alb}, \mathbb{Q}}$. Thus σ acts as identity on $\text{CH}_0(S)_{\text{alb}, \mathbb{Q}}$ if and only if $\sigma\tau$ acts as -1 on $\text{CH}_0(S)_{\text{alb}, \mathbb{Q}}$, which is equivalent to the vanishing of the $(\sigma\tau)$ -invariant part $\text{CH}_0(S)_{\text{alb}, \mathbb{Q}}^{\sigma\tau} \cong \text{CH}_0(S/\langle \sigma\tau \rangle)_{\text{alb}, \mathbb{Q}}$. \square

If a surface with vanishing geometric genus is not of general type, then it has trivial Albanese kernel [BKL76]. Thus we may obtain from Lemma 2.6 the following statement.

Corollary 2.7 [DL23, Lemma 5.1]. *Let S be a smooth projective surface. Let $G = \langle \sigma, \tau \rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$ be a subgroup of $\text{Aut}(S)$ such that (the smooth models of) the quotient surfaces $S/\langle \tau \rangle$ and $S/\langle \sigma\tau \rangle$ are not of general type and have vanishing geometric genus. Then σ acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*

3. Symplectic automorphisms of surfaces with a genus two fibration

For a smooth projective surface S with $p_g(S) > 0$, one may use the canonical system $|K_S|$ to define a rational map $\phi_{K_S} : S \dashrightarrow \mathbb{P}^{p_g(S)-1}$, called the *canonical map* of S . This map is a main tool in our study of symplectic automorphisms of surfaces with a genus two fibration, due to the following observation.

Lemma 3.1. *Let S be a smooth projective surface of general type with a hyperelliptic fibration $f : S \rightarrow B$, and $G \subset \text{Aut}(S)$ the subgroup generated by the hyperelliptic involution $\tau \in \text{Aut}_B(S)$ together with the symplectic automorphism group $\text{Aut}_s(S)$. Suppose that $p_g := p_g(S) > 0$, so that the canonical map $\phi_{K_S} : S \dashrightarrow \mathbb{P}^{p_g-1}$ is well-defined. Then ϕ_{K_S} factors through the quotient map $\pi : S \rightarrow S/G$, that is, there is a rational map $\varphi : S/G \dashrightarrow \mathbb{P}^{p_g-1}$ such that $\phi_{K_S} = \varphi \circ \pi$.*

Proof. Since τ acts as -1 on $H^0(S, K_S)$ while $\text{Aut}_s(S)$ acts trivially on $H^0(S, K_S)$, ϕ_{K_S} is a G -equivariant map with G acting trivially on the target. The assertion of the lemma follows. \square

Before recalling the fundamental results of Xiao on the canonical map of a surface of general type with a genus two fibration, we need to introduce some notation first.

Notation 3.2. Let S be a surface of general type, and $f : S \rightarrow B$ a genus two fibration. We abbreviate the Euler characteristic $\chi(\mathcal{O}_S)$, the geometric genus $p_g(S)$, the irregularity $q(S)$, the genus $g(B)$ of the base curve B by χ , p_g , q , b . Let $M \subset f_*\omega_S$ be an invertible subsheaf with maximal degree and $e := 2 \deg M - \deg f_*\omega_S$. Then we have $\deg f_*\omega_S = \chi + 3(b - 1)$ (see the proof of [Xiao85, théorème 2.1]) and hence

$$(3-1) \quad \deg M = \frac{1}{2}(e + \deg f_*\omega_S) = \frac{1}{2}(e + \chi + 3(b - 1)).$$

As in Lemma 3.1, we will denote by τ the hyperelliptic involution of f , and let $G \subset \text{Aut}(S)$ be the subgroup generated by $\text{Aut}_s(S)$ and τ . Note that τ acts as -1 on $H^0(S, K_S)$ and hence $\tau \notin \text{Aut}_s(S)$ as soon as $p_g(S) > 0$.

Theorem 3.3 [Xiao85, pages 71–73, corollaire 1 and proposition 5.2]. *Let S be a smooth projective surface of general type with a genus two fibration $f : S \rightarrow B$, as in Notation 3.2.*

- (i) *If $p_g \geq 3$ and the image of the canonical map ϕ_{K_S} is a curve, then ϕ_{K_S} factors through f , and there are three possibilities for the numerical invariants:*
 - (a) $q = b = 0$ and $e = p_g$;
 - (b) $b = 0, q = 1$ and $e = p_g + 1$;
 - (c) $q = b = 1$ and $e = p_g$.
- (ii) *If $\chi(\mathcal{O}_S) \geq 4$ and the image of the canonical map is a surface, then $\deg \phi_{K_S} \in \{2, 4\}$. Moreover, if $\deg \phi_{K_S} = 4$, then*
 - (a) $p_g(S) \leq 2b + 2$;
 - (b) $T := \phi_{K_S}(S)$ is either a rational surface or a cone over an elliptic curve.

Now we discuss $\text{Aut}_s(S)$ for a fibered surface $f : S \rightarrow B$ of genus two according to the behavior of the canonical map, given by Theorem 3.3.

Proposition 3.4. *Let S be a smooth projective surface of general type with a genus two fibration $f : S \rightarrow B$ such that $p_g(S) \geq 3$. Suppose that the image of the canonical map ϕ_{K_S} is a curve.*

- (i) $\text{Aut}_s(S) \subset \text{Aut}_B(S)$, that is, each symplectic automorphism of S is fiber-preserving.
- (ii) $\text{Aut}_s(S)$ has order at most 2, and it acts trivially on $\text{CH}_0(S)_{\text{alb}}$.

Proof. By Theorem 3.3(i), ϕ_{K_S} factors through f . Therefore, the fibration f is canonical, and every automorphism of S preserves it. In other words, we have $\text{Aut}(S) = \text{Aut}_f(S)$.

By (3-1), we have

$$\deg M = \begin{cases} p_g - 1 & \text{in cases (a) and (b) of Theorem 3.3(i),} \\ p_g & \text{in case (c) of Theorem 3.3(i).} \end{cases}$$

In all cases, $\text{deg } M \geq 2b + 1$ and hence M is a very ample invertible sheaf on B . By counting dimensions, we have $H^0(B, M) \cong H^0(B, f_*\omega_S) \cong H^0(S, \omega_S)$, and hence the canonical map ϕ_{K_S} factors as

$$\phi_{K_S} : S \xrightarrow{f} B \xrightarrow{\phi_M} \mathbb{P}^{p_g-1}$$

where ϕ_M is the embedding defined by the complete linear system $|M|$.

The morphisms ϕ_{K_S} , f , and ϕ_M are all G -equivariant. Since $\text{Aut}_s(S)$ acts trivially on \mathbb{P}^{p_g-1} and B embeds into \mathbb{P}^{p_g-1} , the induced action of $\text{Aut}_s(S)$ on B is trivial, and hence $\text{Aut}_s(S) \subset \text{Aut}_B(S)$. This proves (i).

(ii) follows from (i) and the next result. □

Lemma 3.5. *Let S be a smooth projective surface, and $f : S \rightarrow B$ a fibration of genus two. If $p_g(S) > 0$, then $\text{Aut}_B(S) \cap \text{Aut}_s(S)$ has order at most 2, and it acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*

Proof. Let F be a general fiber of f . Then $\text{Aut}_B(S)$ injects into $\text{Aut}(F)$, which is finite. Suppose that $\sigma \in \text{Aut}_B(S) \cap \text{Aut}_s(S) \setminus \{\text{id}_S\}$. Then

$$p_g(S/\langle\sigma\rangle) = p_g(S) > 0.$$

This implies that $g(F/\langle\sigma\rangle) = 1$, since it is smaller than 2 but cannot be 0, due to the positivity of p_g . By the Riemann–Hurwitz formula, we have

$$(3-2) \quad 2g(F) - 2 = |\sigma| \left(2g(F/\langle\sigma\rangle) - 2 + \sum_i \left(1 - \frac{1}{r_i} \right) \right).$$

Since the automorphism group $\langle\sigma\rangle$ generated by σ is abelian, the quotient map $F \rightarrow F/\langle\sigma\rangle$ has at least two branch points and hence $\sum_i \left(1 - \frac{1}{r_i} \right) \geq 1$. Then the equality (3-2) implies that $|\sigma| \leq 2$.

The hyperelliptic involution $\tau \in \text{Aut}_B(S)$ commutes with σ , and they generate a Klein group $G := \langle\sigma, \tau\rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$. Since the induced fibrations $S/\langle\sigma\tau\rangle \rightarrow B$ and $S/\langle\tau\rangle \rightarrow B$ have genus 1 and 0 respectively, so both surfaces are not of general type, and they have vanishing p_g , we may conclude by Corollary 2.7 that σ acts trivially on $\text{CH}_0(S)_{\text{alb}}$. □

Lemma 3.6. *Let S be a smooth projective surface of general type with a genus two fibration $f : S \rightarrow B$. If $\chi(\mathcal{O}_S) \geq 5$, then $\text{Aut}(S) = \text{Aut}_f(S)$, that is, every automorphism of S preserves the fibration f .*

Proof. This is because f is the unique genus two fibration on S by [Xiao85, proposition 6.4 and théorème 6.5], and hence preserved by any automorphism of S . □

Proposition 3.7. *Let S be a smooth projective surface of general type with a genus two fibration $f : S \rightarrow B$, as in Notation 3.2. Suppose that $\chi(\mathcal{O}_S) \geq 5$, that the canonical map ϕ_{K_S} is generically finite onto its image $T := \phi_{K_S}(S)$, and that $\text{Aut}_s(S)$ is nontrivial.*

- (i) $\text{Aut}_s(S)$ has order 2, and preserves the fibration f .
- (ii) $S \dashrightarrow T$ is birationally a $(\mathbb{Z}/2\mathbb{Z})^2$ -cover, i.e., the induced extension $K(S)/K(T)$ of function fields is Galois with group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$.
- (iii) $b \geq 2$, and the induced action of $\text{Aut}_s(S)$ on B is not trivial.
- (iv) $\text{Aut}_{\mathcal{O}}(S)$ is trivial.
- (v) If $q_f > 0$, then $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$.

Proof. By Lemma 3.1, ϕ_{K_S} factors through the quotient map $S \rightarrow S/G$, where $G \subset \text{Aut}(S)$ is the subgroup generated by $\text{Aut}_s(S)$ and the hyperelliptic involution τ , which does not lie in $\text{Aut}_s(S)$. Therefore,

$$2|\text{Aut}_s(S)| \leq |G| \leq \deg \phi_{K_S} \in \{2, 4\}$$

where we used the bound on $\deg \phi_{K_S}$ from Theorem 3.3(ii). Since $\text{Aut}_s(S)$ is nontrivial by assumption, we infer that

$$\deg \phi_{K_S} = 4, \quad |\text{Aut}_s(S)| = 2, \quad G \cong (\mathbb{Z}/2\mathbb{Z})^2.$$

By Lemma 3.6, $\text{Aut}_s(S)$ preserves the fibration f . Thus (i) is proved.

For (ii), note that ϕ_{K_S} has degree 4 and factors through the quotient map $S \rightarrow S/\langle \tau, \sigma \rangle$, which also has degree 4. Thus the induced map $S/\langle \tau, \sigma \rangle \dashrightarrow T$ is birational, and (ii) follows.

(iii) By Theorem 3.3(ii), we have

$$(3-3) \quad p_g \leq 2b + 2.$$

It follows that

$$b \geq p_g - b - 2 \geq p_g - q - 2 = \chi - 3 \geq 2$$

where the last inequality is because $\chi \geq 5$ by assumption.

By Theorem 3.3(ii), T is either a rational surface or a cone over an elliptic curve. This implies that the smooth models of T (and of S/G) have irregularity at most 1. Thus, the dimension of $H^0(S, \Omega_S^1)^G$ is at most 1. On the other hand, since $S/\langle \tau \rangle$ is a \mathbb{P}^1 -fibration over B , we have

$$\dim H^0(S, \Omega_S^1)^\tau = b \geq 2.$$

It follows that $H^0(S, \Omega_S^1)^G \subsetneq H^0(S, \Omega_S^1)^\tau$, and hence σ does not act trivially on $H^0(S, \Omega_S^1) \cong H^1(S, \mathcal{O}_S)^\vee$. Therefore, the automorphism $\sigma_B \in \text{Aut}(B)$ induced by σ is not the identity. This finishes the proof of (iii).

(iv) We have $\sigma \notin \text{Aut}_{\mathcal{O}}(S)$ by (iii), and hence $\text{Aut}_{\mathcal{O}}(S)$ is trivial.

(v) Note that $q_f \leq 2$ in any case. If $q_f = 2$, then the assertion follows directly from Corollary 2.5.

Thus we may assume that $q_f = 1$, and there is a one-form

$$\eta_0 \in H^0(S, \Omega_S^1) \setminus f^*H^0(B, \Omega_B),$$

which we may assume to be an eigenvector with eigenvalue λ under the action of σ . Then

$$H^0(S, \Omega_S^1) = f^*H^0(\Omega_B^1) \oplus \mathbb{C} \cdot \eta_0,$$

both summands of which are invariant under the action of $G = \langle \sigma, \tau \rangle$. Define a \mathbb{C} -linear map

$$(3-4) \quad \wedge \eta_0 : H^0(B, \Omega_B^1) \rightarrow H^0(S, K_S), \quad \eta \mapsto f^*\eta \wedge \eta_0,$$

which is easily seen to be injective and G -equivariant. Since σ acts trivially on $H^0(S, K_S)$ and hence also trivially on the subspace $f^*H^0(B, \Omega_B^1) \wedge \eta_0$, it follows that σ acts as the scalar multiplication by λ^{-1} on the whole $H^0(B, \Omega_B^1)$. Thus the canonical map $\phi_{K_B} : B \rightarrow \mathbb{P}^{b-1}$ factors through the quotient map $B \rightarrow B/\langle \sigma_B \rangle$. Since $b \geq 2$, we infer that B is a hyperelliptic curve and σ_B is its hyperelliptic involution, acting as -1 on $H^0(B, \Omega_B^1)$. It follows that $\lambda = -1$, and the eigenvalues of the action are as in the following table:

	$f^*H^0(\Omega_B^1)$	$\mathbb{C} \cdot \eta_0$
σ	-1	-1
τ	1	-1

It follows that

$$p_g(S/\langle \sigma \tau \rangle) = 0, \quad q(S/\langle \sigma \tau \rangle) = 1$$

and hence $S/\langle \sigma \tau \rangle$ is not of general type [Bea96, Chapter VI]. By Corollary 2.7, σ acts trivially on $\text{CH}_0(S)_{\text{alb}}$. □

Finally, we make an observation that Conjecture 1.2 (for surfaces with $p_g = 0$) implies Conjecture 1.1 for surfaces of general type with a genus two fibration and $\chi(\mathcal{O}_S) \geq 5$.

Proposition 3.8. *Let S be a smooth projective surface of general type with a genus two fibration and $\chi(\mathcal{O}_S) \geq 5$. Assume that Conjecture 1.2 holds. Then $\text{Aut}_s(S)$ acts trivially on $\text{CH}_0(S)_{\text{alb}}$.*

Proof. We may assume that $\text{Aut}_s(S)$ is nontrivial. By Theorem 1.4, we may assume that the canonical map ϕ_{K_S} is generically finite, and $\text{Aut}_s(S)$ has order two. Let σ be the generator of $\text{Aut}_s(S)$, and $\tau \in \text{Aut}_B(S)$ the hyperelliptic involution. Then $\langle \sigma, \tau \rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$, and since $S/\langle \tau \rangle \rightarrow B$ is a \mathbb{P}^1 -fibration, one has $\text{CH}_0(S/\langle \tau \rangle)_{\text{alb}} = 0$. Note also that $p_g(S/\langle \sigma\tau \rangle) = 0$.

Assuming Conjecture 1.2 holds, we have $\text{CH}_0(S/\langle \sigma\tau \rangle)_{\text{alb}} = 0$, and hence σ acts trivially on $\text{CH}_0(S)_{\text{alb}}$ by Lemma 2.6. \square

Acknowledgements

We thank Professor Jin-Xing Cai for communications on some questions related to the paper, and the referees for their careful reading and useful suggestions. Wenfei Liu was supported by the NSFC (no. 12571046).

References

- [Bea82] A. Beauville, “Appendice: L’inégalité $p_g \geq 2q - 4$ pour les surfaces de type général”, *Bull. Soc. Math. France* **110**:3 (1982), 343–346. In O. Debarre, “Inégalités numériques pour les surfaces de type général”, pp. 319–346 in the same issue. MR
- [Bea96] A. Beauville, *Complex algebraic surfaces*, 2nd ed., London Mathematical Society Student Texts **34**, Cambridge University Press, 1996. MR
- [BKL76] S. Bloch, A. Kas, and D. Lieberman, “Zero cycles on surfaces with $p_g = 0$ ”, *Compositio Math.* **33**:2 (1976), 135–145. MR
- [Blo75] S. Bloch, “ K_2 of Artinian Q -algebras, with application to algebraic cycles”, *Comm. Algebra* **3** (1975), 405–428. MR
- [Cai06a] J.-X. Cai, “Automorphisms of fiber surfaces of genus 2, inducing the identity in cohomology”, *Trans. Amer. Math. Soc.* **358**:3 (2006), 1187–1201. MR
- [Cai06b] J.-X. Cai, “Automorphisms of fiber surfaces of genus 2, inducing the identity in cohomology, II”, *Internat. J. Math.* **17**:2 (2006), 183–193. MR
- [Cai07] J.-X. Cai, “Classification of fiber surfaces of genus 2 with automorphisms acting trivially in cohomology”, *Pacific J. Math.* **232**:1 (2007), 43–59. MR
- [DL23] J. Du and W. Liu, “On symplectic automorphisms of elliptic surfaces acting on CH_0 ”, *Sci. China Math.* **66**:3 (2023), 443–456. MR
- [Huy12] D. Huybrechts, “Symplectic automorphisms of K3 surfaces of arbitrary finite order”, *Math. Res. Lett.* **19**:4 (2012), 947–951. MR
- [JLZ23] Z. Jiang, W. Liu, and H. Zhao, “On numerically trivial automorphisms of threefolds of general type”, *Math. Res. Lett.* **30**:6 (2023), 1751–1785. MR
- [Kim05] S.-I. Kimura, “Chow groups are finite dimensional, in some sense”, *Math. Ann.* **331**:1 (2005), 173–201. MR
- [Lat21] R. Laterveer, “Bloch’s conjecture for some numerical Campedelli surfaces”, *Asian J. Math.* **25**:1 (2021), 49–64. MR
- [LYZ24] Z. Li, X. Yu, and R. Zhang, “Bloch’s conjecture for (anti-)autoequivalences on K3 surfaces”, preprint, 2024. arXiv 2305.10078v2

- [Mum68] D. Mumford, “Rational equivalence of 0-cycles on surfaces”, *J. Math. Kyoto Univ.* **9** (1968), 195–204. MR
- [Paw19] R. R. Pawar, “Action of correspondences on filtrations on cohomology and 0-cycles of abelian varieties”, *Math. Z.* **292**:1-2 (2019), 655–675. MR
- [Roj80] A. A. Rojzman, “The torsion of the group of 0-cycles modulo rational equivalence”, *Ann. of Math. (2)* **111**:3 (1980), 553–569. MR
- [Voi03] C. Voisin, *Hodge theory and complex algebraic geometry, II*, Cambridge Studies in Advanced Mathematics **77**, Cambridge University Press, 2003. MR
- [Voi12] C. Voisin, “Symplectic involutions of $K3$ surfaces act trivially on CH_0 ”, *Doc. Math.* **17** (2012), 851–860. MR
- [Xiao85] G. Xiao, *Surfaces fibrées en courbes de genre deux*, Lecture Notes in Mathematics **1137**, Springer, 1985. MR

Received January 20, 2025. Revised May 25, 2025.

JIABIN DU
SHANGHAI INSTITUTE FOR MATHEMATICS AND INTERDISCIPLINARY SCIENCES (SIMIS)
SHANGHAI
CHINA

and

RESEARCH INSTITUTE OF INTELLIGENT COMPLEX SYSTEMS
FUDAN UNIVERSITY
SHANGHAI
CHINA

jjabin.du@simis.cn

WENFEI LIU
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
XIAMEN, FUJIAN
CHINA

wliu@xmu.edu.cn

NON-BRAID-POSITIVE HYPERBOLIC L -SPACE KNOTS

KEISUKE HIMENO

An L -space knot is a knot that admits a positive Dehn surgery yielding an L -space. Many known hyperbolic L -space knots are braid positive, meaning they can be represented as the closure of a positive braid. Recently, Baker and Kegel showed that the hyperbolic L -space knot $o9_30634$ from Dunfield's census is not braid-positive, and they constructed infinitely many candidates for hyperbolic L -space knots that may not be braid-positive. However, it remains unproven whether their examples are genuinely non-braid-positive. In this paper, we construct infinitely many hyperbolic L -space knots that are not braid-positive, and are distinct from those considered by Baker and Kegel.

1. Introduction

An L -space is a rational homology 3–sphere whose (hat version) Heegaard Floer homology has rank equal to the order of its first homology. A knot is called an L -space knot if it admits a positive Dehn surgery yielding an L -space. L -space knots were originally motivated by the study of knots admitting lens space surgeries [14], and they continue to be an active subject of research.

In this paper, we consider the braid positivity of L -space knots. A knot or link is said to be *braid-positive* if it can be expressed as the closure of a positive braid. Many known L -space knots are braid-positive; for example, positive torus knots are L -space knots, and then they are braid-positive. On the other hand, it is known that not all L -space knots are braid-positive; for example, the $(2, 3)$ -cable of the right-handed trefoil is not braid-positive (see Example 1 of [1]). However, the existence of non-braid-positive hyperbolic L -space knots remained an open problem for some time (see Problem 31.2 of [5], Question 2 of [1]).

Recently, Baker and Kegel examined Dunfield's list of 632 hyperbolic L -space knots. They showed that all but one of these knots are braid-positive, and that the exceptional knot, $o9_30634$, is not braid-positive [2]. Furthermore, they constructed infinitely many candidates for non-braid-positive hyperbolic L -space knots, although they could not prove that any of their examples are definitely non-braid-positive.

MSC2020: 57K10, 57K18.

Keywords: L -space knot, braid-positive.

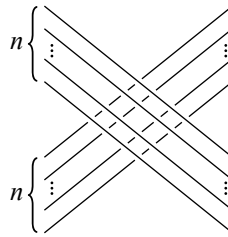


Figure 1. The braid X_n .

The purpose of this paper is to give infinitely many hyperbolic L -space knots that are not braid-positive. While all candidates by Baker and Keigel are represented as closures of 4-braids, we construct the infinite family of such knots by increasing the number of strands in the braid. For each $n \geq 2$, we define a knot K_n as the closure of a $2n$ -braid, as follows. Let $[\varepsilon_1 i_1, \varepsilon_2 i_2, \dots, \varepsilon_m i_m]$ denote the braid $\sigma_{i_1}^{\varepsilon_1} \sigma_{i_2}^{\varepsilon_2} \cdots \sigma_{i_m}^{\varepsilon_m}$, where σ_i is the i -th standard generator of the degree $2n$ braid group and $\varepsilon_j = \pm 1$. Define the braid

$$X_n = [n, n - 1, n + 1, n - 2, n, n + 2, \dots, 1, 3, \dots, 2n - 1, 2n - 2, \dots, 4, 2, \dots, n + 1, n - 1, n],$$

illustrated in Figure 1. Let K_n ($n \geq 1$) be a knot represented by the closure of $2n$ -braid

$$\beta_n = X_n^3 \cdot [-1, -2, \dots, -(n - 1), n, n - 1, n - 2, \dots, 2, 1, 1, 2, 3, \dots, n],$$

shown in Figure 2. Note that K_2 coincides with the knot $o9_30634$, and K_1 is the $(2, 5)$ -torus knot. Although the K_n for odd n do not appear in the statement of Theorem 1.1, we use them in the proof.

Theorem 1.1. *When n is even, K_n is a non-braid-positive hyperbolic L -space knot.*

Since the knots $\{K_n\}$ are mutually distinct (see Lemma 2.8), we have the following corollary.

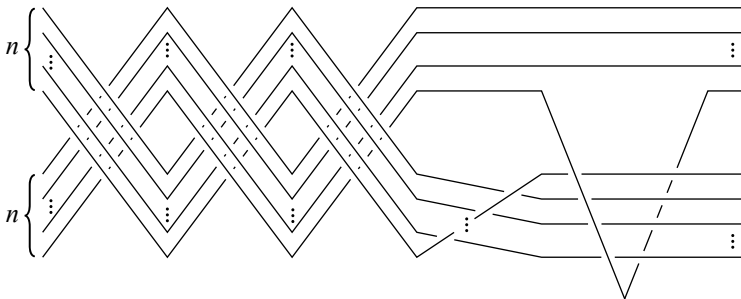


Figure 2. The knot K_n is the closure of this braid.

Corollary 1.2. *There exist infinitely many non-braid-positive hyperbolic L -space knots.*

When n is odd, we also expect K_n to be a non-braid positive hyperbolic L -space knot. However, in this case, the criterion we use for braid positivity failed to detect it.

2. Non-braid positivity

In this section, we prove that K_n is non-braid-positive when n is even. Our proof is based on Ito’s criterion using the HOMFLY polynomial [8].

2.1. HOMFLY polynomial and its zeroth coefficient polynomial. For an oriented link L , the HOMFLY polynomial $P_L(v, z)$ is a two-variable Laurent polynomial defined by the skein relation

$$v^{-1}P_{L_+}(v, z) - vP_{L_-}(v, z) = zP_{L_0}(v, z),$$

together with $P_U(v, z) = 1$, where U is the unknot. Here, the links (or diagrams) L_+ , L_- and L_0 coincide outside a small 3-ball, and inside the 3-ball, they differ as in Figure 3. (Throughout, all braids are assumed to be oriented from left to right.)

In [8], Ito provided a criterion for the braid positivity of a link using the HOMFLY polynomial. We state the result in the case of knots. Let

$$\tilde{P}_K(\alpha, z) = (-\alpha)^{-g(K)} P_K(v, z)|_{-v^2=\alpha},$$

where $g(K)$ denotes the genus of a knot K .

Theorem 2.1 [8, Theorem 2]. *If K is a braid-positive knot, then $\tilde{P}_K(\alpha, z)$ is positive, that is, all nonzero coefficients of $\tilde{P}_K(\alpha, z)$ are positive integers.*

The HOMFLY polynomial $P_L(v, z)$ can be expressed in the form

$$P_L(v, z) = (v^{-1}z)^{-\#L+1} \sum_{i=0}^{\#L-1} p_L^i(v)z^{2i},$$

where $\#L$ is the number of components of the link L . The polynomial $p_L^i(v)$ is called the i -th coefficient (HOMFLY) polynomial of L . In this paper, we focus on the zeroth coefficient polynomial $p_L^0(v)$. It is known that this polynomial satisfies several important properties; see, for example [9, Section 2] and [19, Section 2].

The zeroth coefficient polynomial $p_L^0(v)$ satisfies the skein relation

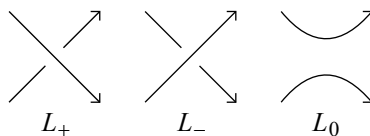


Figure 3. The small 3-ball in the diagrams of L_+ , L_- and L_0 .

$$(2-1) \quad v^{-2} p_{L_+}^0(v) - p_{L_-}^0(v) = \begin{cases} p_{L_0}^0(v) & (\delta = 0), \\ 0 & (\delta = 1), \end{cases}$$

where $\delta = \frac{1}{2}(\#L_+ - \#L_0 + 1) \in \{0, 1\}$. In particular, $\delta = 0$ if L_+ is a knot. Furthermore, for a two-component link $L = k_1 \cup k_2$, we have

$$(2-2) \quad p_L^0(v) = (v^{-2} - 1)v^{2 \cdot \text{lk}(k_1, k_2)} p_{k_1}^0(v) p_{k_2}^0(v),$$

where $\text{lk}(k_1, k_2)$ is the linking number of k_1 and k_2 . In particular, when L_+ is a knot (hence L_- is also a knot and L_0 is a two-component link $k_1 \cup k_2$), the skein relation (2-1) can be rewritten as

$$(2-3) \quad p_{L_-}^0(v) = v^{-2} p_{L_+}^0(v) + (1 - v^{-2})v^{2 \cdot \text{lk}(k_1, k_2)} p_{k_1}^0(v) p_{k_2}^0(v).$$

2.2. Degree of p_L^0 for a braid-positive link L . For a braid β , let $L(\beta)$ denote the closure of β . A positive braid β is a *minimal positive braid* if the number of strands of β is minimum among all the positive braid representative of $L(\beta)$. The following proposition is useful when applying the skein relation to a braid-positive link.

Proposition 2.2 [21, Lemma 2]. *Let L be a braid-positive link that is not an unlink. Then there exists a positive braid β such that L is the closure of a positive braid of the form $\sigma_i^2 \beta$ for some i . Moreover, such a positive braid representative $\sigma_i^2 \beta$ of L can be taken so that it is a minimal positive braid.*

Lemma 2.3. *Let β be a positive n -braid, and let e be the number of crossings of β . Then the degree of the zeroth coefficient polynomial satisfies*

$$\deg p_{L(\beta)}^0(v) \leq n + e - \#L(\beta).$$

Proof. We prove this lemma by induction on e . For an unlink U , we have $p_U^0(v) = (v^{-2} - 1)^{\#U-1}$, so $\deg p_U^0(v) = 0$.

Assume that $L(\beta)$ is not an unlink. By Proposition 2.2, we can write $L(\beta) = L(\sigma_i^2 \beta')$ for some i , where β' is a positive n -braid with $e - 2$ crossings. (If necessary, we may increase the number of crossings to realize β' as an n -braid.)

Applying the skein relation (2-1), we get

$$p_{L(\beta)}^0(v) = p_{L(\sigma_i^2 \beta')}^0(v) = v^2 p_{L(\beta')}^0(v) + \begin{cases} v^2 p_{L(\sigma_i \beta')}^0(v) & (\delta = 0), \\ 0 & (\delta = 1). \end{cases}$$

Note that $\#L(\beta') = \#L(\beta)$, and $\#L(\sigma_i \beta') = \#L(\beta) + 1$ if $\delta = 0$. By the assumption of induction, we have

$$\deg v^2 p_{L(\beta')}^0(v) \leq 2 + (n + e - 2 - \#L(\beta')) = n + e - \#L(\beta),$$

and if $\delta = 0$,

$$\deg v^2 p_{L(\sigma_i \beta')}^0(v) \leq 2 + (n + e - 1 - \#L(\sigma_i \beta')) = n + e - \#L(\beta).$$

The claim follows. □

Definition 2.4. A positive n -braid β is said to be *sharp* if

$$\deg p_L^0(\beta)(v) = n + e - \#L(\beta),$$

where e is the number of crossings of β .

We now state several lemmas concerning the notion of sharpness.

Lemma 2.5. *If positive braid $\sigma_i^2\beta$ is sharp, then at least one of β and $\sigma_i\beta$ is sharp.*

Proof. The claim follows immediately from the proof of Lemma 2.3. □

Lemma 2.6. *If a positive braid β is sharp, then it is a minimal positive braid.*

Proof. Let β be a positive n -braid that is not minimal, and let β' be a minimal positive braid representative of $L(\beta)$ with n' strands ($n' < n$). The number of crossings of β is $n - \chi(L(\beta))$ and that of β' is $n' - \chi(L(\beta))$ [18]. Here $\chi(L)$ is the maximal Euler characteristic among all compact, connected, oriented surfaces whose boundary is a link L .

By Lemma 2.3,

$$\deg p_{L(\beta)}^0(v) \leq n' + n' - \chi(L(\beta)) - \#L(\beta) < n + n - \chi(L(\beta)) - \#L(\beta).$$

This implies that β is not sharp. □

Lemma 2.7. *Let $L = k_1 \cup k_2$ be a two-component link represented as the closure of a positive braid β . Suppose that each component knot k_1 and k_2 can also be represented as the closure of positive braids β_1 and β_2 , respectively. If β is sharp, then both β_1 and β_2 are sharp.*

Proof. Let n, n_1, n_2 be the numbers of strands, and e, e_1, e_2 be the numbers of crossings of β, β_1 and β_2 , respectively. Then,

$$n = n_1 + n_2, \quad e = e_1 + e_2 + \text{lk}(k_1, k_2).$$

By (2-2) and Lemma 2.3, we have

$$\deg p_L^0 \leq 2 \cdot \text{lk}(k_1, k_2) + (n_1 + e_1 - 1) + (n_2 + e_2 - 1) = n + e - 2.$$

Equality holds if and only if β_1 and β_2 are sharp. Since β is sharp, equality holds, and the result follows. □

2.3. Non-braid positivity for K_n . We are ready to prove that K_n is not braid-positive when n is even.

Lemma 2.8. *For $n \geq 2$, the top term of $p_{K_n}^0(v)$ is $(-1)^n v^{3n^2+3n}$.*

Proof. We prove this by induction on n . Using Sage [17], one can confirm that the top term of $p_{K_2}^0(v)$ is v^{18} .

Throughout the following, we identify each braid with its closure diagram in the skein tree.

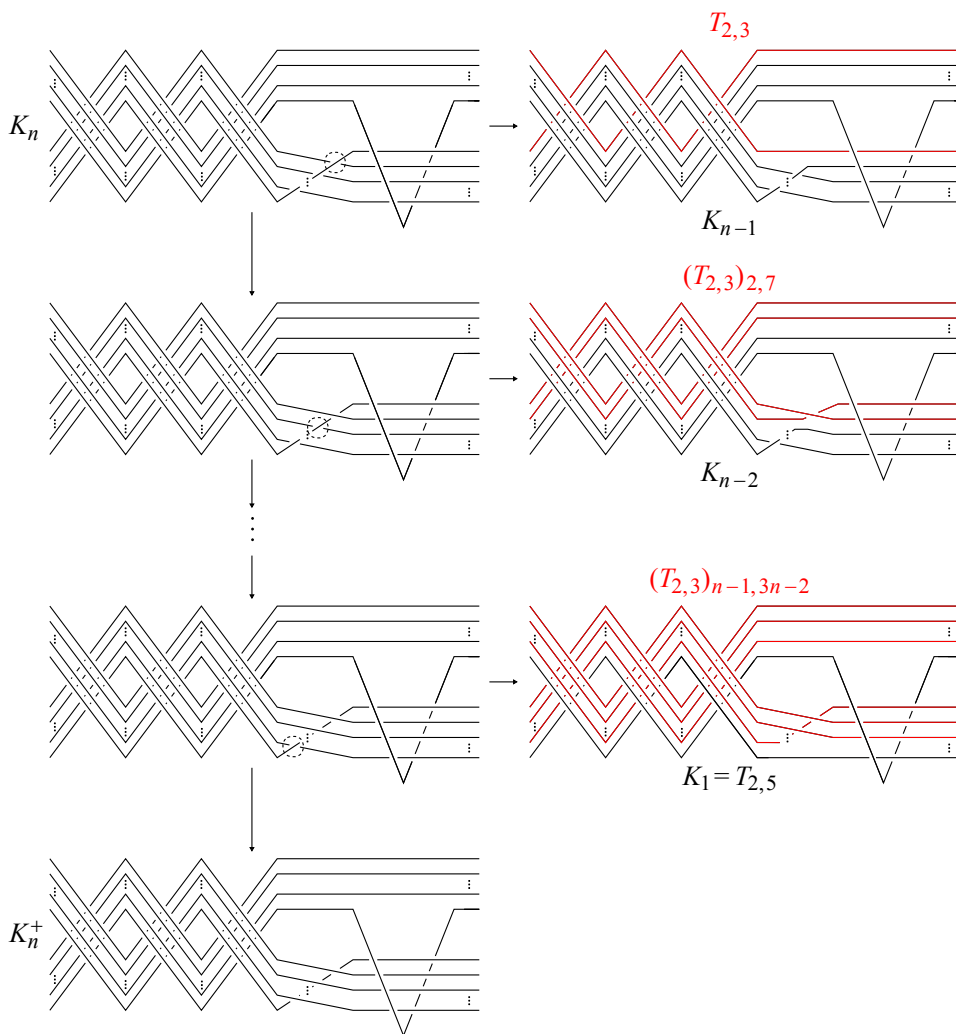


Figure 4. The skein tree for the knot K_n . Consider the crossing change and the smoothing at the crossing indicated by the dashed circle.

Assume that $n \geq 3$. From the skein tree shown in Figure 4 and equation (2-3), we obtain

$$\begin{aligned}
 p_{K_n}^0(v) &= (1 - v^{-2})v^{2(3(n-1)\cdot 1+1)} p_{T_{2,3}}^0(v) p_{K_{n-1}}^0(v) \\
 &\quad + (1 - v^{-2})v^{-2} \cdot v^{2(3(n-2)\cdot 2+2)} p_{(T_{2,3})_{2,7}}^0(v) p_{K_{n-2}}^0(v) + \dots \\
 &\quad + (1 - v^{-2})v^{-2(n-2)} \cdot v^{2(3\cdot 1\cdot (n-1)+n-1)} p_{(T_{2,3})_{n-1, 3n-2}}^0(v) p_{K_1}^0(v) \\
 &\quad + v^{-2(n-1)} p_{K_n^+}^0(v) \\
 &= \sum_{k=1}^{n-1} v^{-2(k-1)} (1 - v^{-2})v^{2(3(n-k)\cdot k+k)} p_{(T_{2,3})_{k, 3k+1}}^0(v) p_{K_{n-k}}^0(v) + v^{-2(n-1)} p_{K_n^+}^0(v),
 \end{aligned}$$

where $(T_{2,3})_{k,3k-1}$ denotes the $(k, 3k-1)$ -cable of the right-handed trefoil and K_n^+ is the braid-positive knot shown at the bottom in Figure 4.

By the induction assumption and the fact $\deg p_{K_1}^0(v) = \deg p_{T_{2,5}}^0(v) = 6$, we have $\deg p_{K_i}^0 = 3i^2 + 3i$ for $1 \leq i \leq n-1$. Since $(T_{2,3})_{k,3k+1}$ is represented by the closure of the positive braid $X_k^3 \cdot [1, 2, \dots, k-1]$, Lemma 2.3 gives

$$\begin{aligned} &\deg(v^{-2(k-1)}(1-v^{-2})v^{2(3(n-k)-k+k)}p_{(T_{2,3})_{k,3k+1}}^0(v)p_{K_{n-k}}^0(v)) \\ &\leq -2(k-1) + 2(3(n-k)k+k) + (2k+3k^2+(k-1)-1) + 3(n-k)^2 + 3(n-k) \\ &= 3n^2 + 3n. \end{aligned}$$

Equality holds if and only if the positive braid $X_k^3 \cdot [1, 2, \dots, k-1]$ is sharp.

Claim 2.9. *For $k \geq 2$, the positive braid $X_k^3 \cdot [1, 2, \dots, k-1]$ is not sharp.*

Proof. We consider the skein tree

$$\begin{array}{ccccccc} D_0 & \rightarrow & D_1 & \rightarrow & \dots & \rightarrow & D_i & \rightarrow & D_{i+1} & \rightarrow & \dots & \rightarrow & D_k, \\ \downarrow & & \downarrow & & & & \downarrow & & \downarrow & & & & \\ D_0^s & & D_1^s & & & & D_i^s & & D_{i+1}^s & & & & \end{array}$$

where D_0 is the braid $X_k^3 \cdot [1, 2, \dots, k-1]$ and each D_i ($i = 1, \dots, k$) and D_i^s ($i = 0, \dots, k-1$) are as illustrated in Figures 5 and 6. The right arrows correspond to crossing changes, and the down arrows correspond to smoothings. Note that all links in the skein tree are braid-positive links.

D_k contains exactly one occurrence of the generator σ_{2k-1} , and is hence a nonminimal positive braid. Each D_i^s ($0 \leq i \leq k-1$) is also nonminimal, as shown in Figures 7 and 8. Thus, by Lemmas 2.5 and 2.6, the claim follows. \square

On the other hand, K_n^+ is represented as the closure of the positive braid

$$X_n^3 \cdot [1, 2, \dots, n-1, n, n-1, n-2, \dots, 2, 1, 1, 2, 3, \dots, n].$$

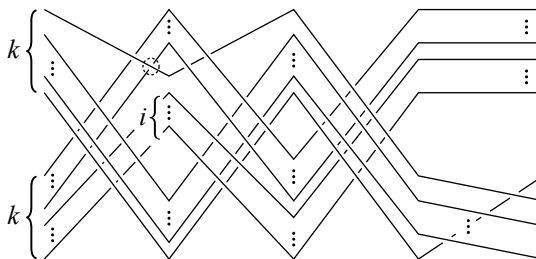


Figure 5. The diagram D_i ($0 \leq i \leq k$). Performing a crossing change at the crossing indicated by the dashed circle, followed by a Reidemeister II move, results in the diagram D_{i+1} . Alternatively, smoothing the crossing produces the diagram D_i^s .

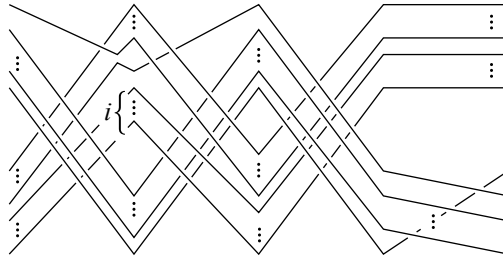


Figure 6. The diagram D_i^s ($0 \leq i \leq k - 1$).

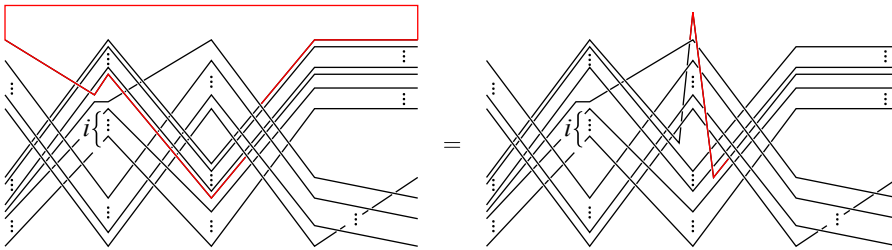


Figure 7. The positive braid D_i^s ($0 \leq i \leq k - 2$) is not minimal.

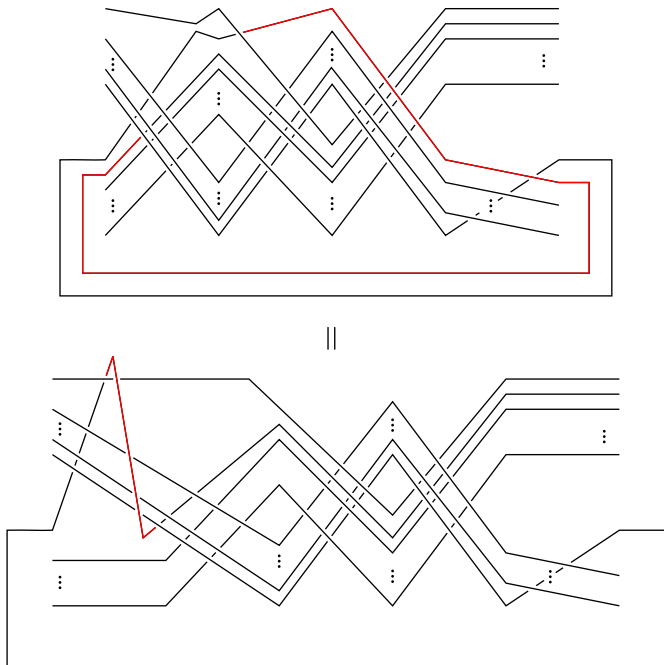


Figure 8. The positive braid D_{k-1}^s is not minimal.

Lemma 2.3 gives

$$\deg v^{-2(n-1)} p_{K_n^+}^0(v) \leq -2(n-1) + 2n + (3n^2 + n - 1 + 2n) - 1 = 3n^2 + 3n.$$

Equality holds if and only if the positive braid

$$X_n^3 \cdot [1, 2, \dots, n-1, n, n-1, n-2, \dots, 2, 1, 1, 2, 3, \dots, n]$$

is sharp again.

Claim 2.10. For $n \geq 3$, the positive braid

$$X_n^3 \cdot [1, 2, \dots, n-1, n, n-1, n-2, \dots, 2, 1, 1, 2, 3, \dots, n]$$

is not sharp.

Proof. We consider the skein tree

$$\begin{array}{ccccccccccc} E_0 & \rightarrow & E_1 & \rightarrow & \dots & \rightarrow & E_i & \rightarrow & E_{i+1} & \rightarrow & \dots & \rightarrow & E_n, \\ \downarrow & & \downarrow & & & & \downarrow & & \downarrow & & & & \\ E_0^s & & E_1^s & & & & E_i^s & & E_{i+1}^s & & & & \end{array}$$

where E_0 is the braid described in the claim, and E_n is $X_n^3 \cdot [1, 2, \dots, n-1]$, which is not sharp by Claim 2.9. The diagrams E_i and E_i^s are illustrated in Figures 9 and 10.

As shown in Figure 10, the diagram E_i^s consists of two components k_1 and k_2 , represented by the positive braids γ_i^1 and γ_i^2 respectively. Figure 11 shows that γ_i^1 ($1 \leq i \leq n-1$) is not minimal, and hence nonsharp by Lemma 2.6. Then, Lemma 2.7 implies that E_i^s is not sharp for $1 \leq i \leq n-1$.

For $i = 0$, we observe that γ_0^1 is a 1-braid (thus sharp), so we analyze γ_0^2 . We perform a crossing change and smoothing at the crossing marked in Figure 12. The resulting diagrams (Figures 13 and 14) demonstrate that γ_0^2 is not sharp, and hence E_0^s is not sharp. □

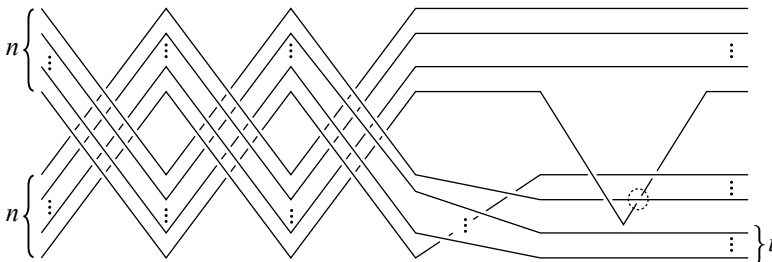


Figure 9. The diagram E_i ($0 \leq i \leq n$). Performing a crossing change at the crossing indicated by the dashed circle, results in the diagram E_{i+1} . Alternatively, smoothing the crossing produces the diagram E_i^s .

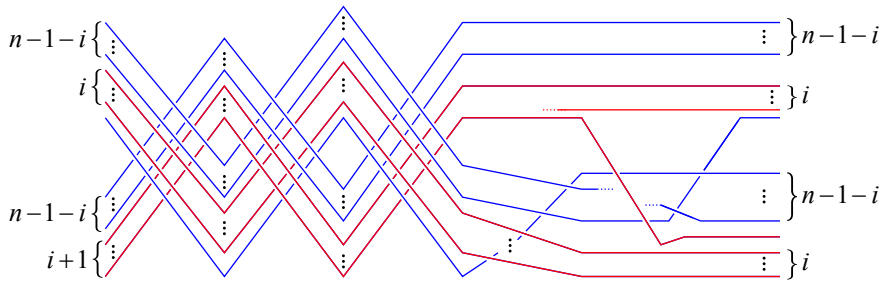


Figure 10. The diagram E_i^s ($0 \leq i \leq n-1$) represents the two-component link $k_1 \cup k_2$. k_1 is represented by the positive braid γ_i^1 (red), and k_2 by γ_i^2 (blue).

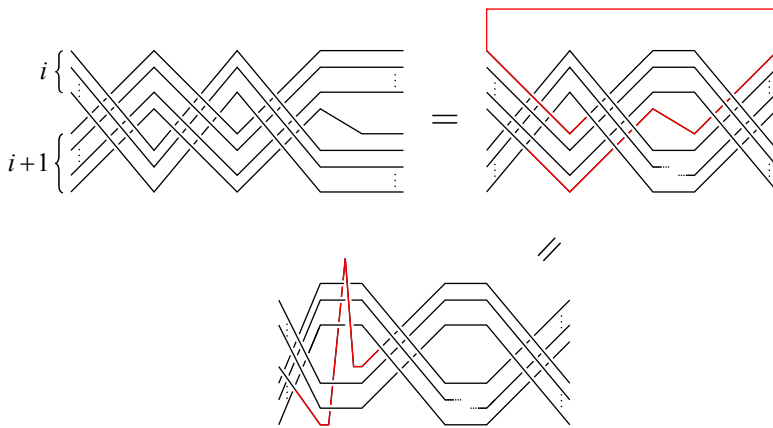


Figure 11. The positive braid γ_i^1 ($1 \leq i \leq n-1$) is not minimal. The first deformation is a conjugation that moves the X-shaped part on the left side of the braid to the right.

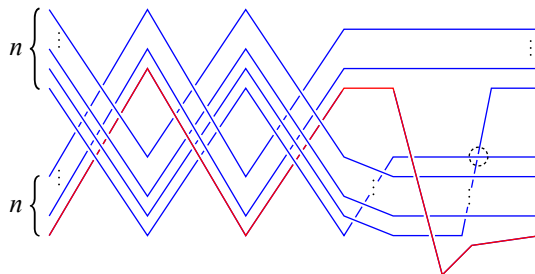


Figure 12. The diagram E_0^s . The red line represents γ_0^1 , and the blue line represents γ_0^2 . The crossing change at the crossing indicated by the dashed circle changes γ_0^2 into the diagram as in Figure 13. Alternatively, the smoothing yields the diagram as in Figure 14.

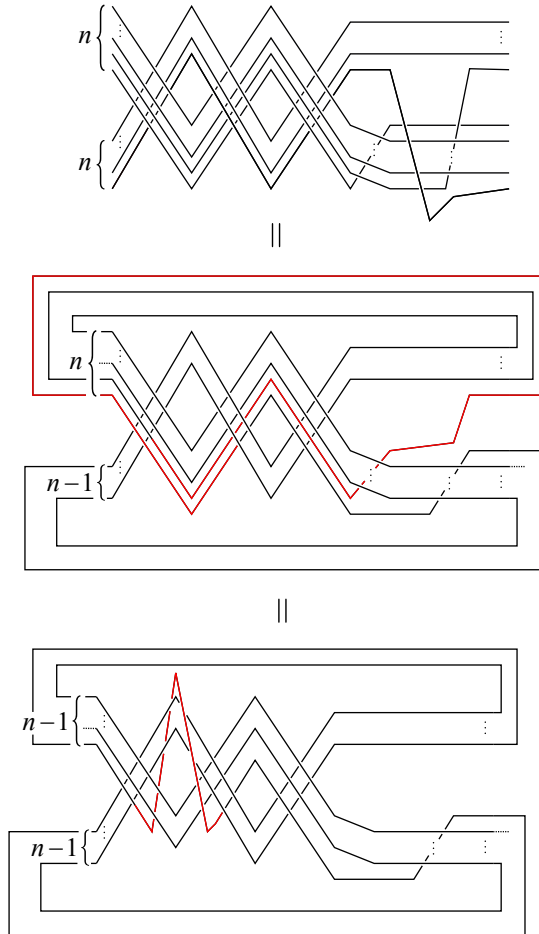


Figure 13. The diagram obtained by the crossing change of γ_0^2 . As indicated, the corresponding positive braid is not minimal, hence not sharp.

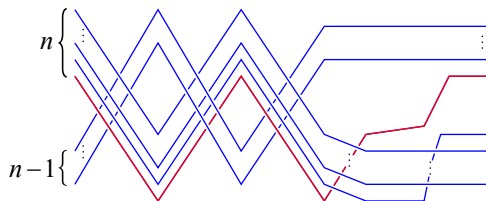


Figure 14. The diagram obtained by the smoothing of γ_0^2 . The blue line is the positive braid $X_{n-1}^3[1, 2, \dots, n-2]$, which is not sharp by Claim 2.9.

By Claims 2.9 and 2.10, the top term of $p_{K_n}^0(v)$ arises from the term

$$(1 - v^{-2})v^{2(3(n-1)\cdot 1+1)}p_{T_{2,3}}^0(v)p_{K_{n-1}}^0(v),$$

which evaluates to $v^{2(3(n-1)\cdot 1+1)}(-v^4)(-1)^{n-1}v^{3(n-1)^2+3(n-1)} = (-1)^n v^{3n^2+3n}$. The proof of Lemma 2.8 is complete. \square

Lemma 2.11. *When n is even, the genus of K_n is given by*

$$g(K_n) = \frac{3}{2}n^2 - \frac{1}{2}n + 1.$$

Proof. By [13] and Proposition 3.1 below, K_n is an L -space knot, and hence fibered. A quasi-positive Seifert surface for K_n can be constructed by attaching $3n^2 + n + 1$ bands to $2n$ disks, where one of the bands corresponds to the braid

$$[-1, -2, \dots, -(n-1), n, n-1, \dots, 2, 1]$$

and the remaining bands correspond to the other positive crossings. By [16], the surface is incompressible and therefore serves as the fiber surface of K_n . A direct calculation then yields the genus $g(K_n) = \frac{3}{2}n^2 - \frac{1}{2}n + 1$. \square

Proposition 2.12. *If n is even, then K_n is not braid-positive.*

Proof. Set $n = 2k$. By Lemmas 2.8 and 2.11, the top term of $(-\alpha)^{-g(K_{2k})} p_{K_{2k}}^0(v)|_{-v^2=\alpha}$ is

$$(-\alpha)^{-(6k^2-k+1)}(-1)^{2k}(-\alpha)^{6k^2+3k} = (-\alpha)^{4k-1} = -\alpha^{4k-1}.$$

Since this is the term of z^0 in $\tilde{P}_{K_n}(\alpha, z)$, $\tilde{P}_{K_n}(\alpha, z)$ is a non-positive polynomial. Therefore, Theorem 2.1 implies that K_n is not braid-positive. \square

3. L -space surgery

In this section, we prove that K_n admits a Dehn surgery yielding an L -space if n is even. Throughout, we set $n = 2k$. We apply the Montesinos trick [12]: for a strongly invertible link L in S^3 , the manifold obtained by Dehn surgery on L is the double branched cover of a link ℓ arising from a tangle replacement on the axis. See [2; 20] for details.

Figure 15 illustrates a strongly invertible position of the link $K \cup C_1 \cup C_2 \cup C_3$, where performing (-1) -surgery on C_1 and 1-surgery on both C_2 and C_3 transforms K into K_{2k} . Note that r -surgery on K corresponds to $(8k^2 + r)$ -surgery on K_{2k} .

Proposition 3.1. *$(12k^2 + 2k)$ -surgery on K_{2k} yields an L -space.*

Proof. When $k = 1$, the 14-surgery on $K_2 = o9_30634$ yields an L -space [2]. Assume that $k \geq 2$. Consider the quotient of the link in Figure 15 under the involution around the axis, as shown in Figure 16. Note that the surgery coefficient on K is $4k^2 + 2k$, and the writhe of K in the diagram is $4k^2 + 2k + 1$. Hence, $(4k^2 + 2k)$ -surgery on K corresponds to a tangle replacement by the (-1) -tangle.

Figure 17 illustrates a deformation of the quotient. Performing the indicated

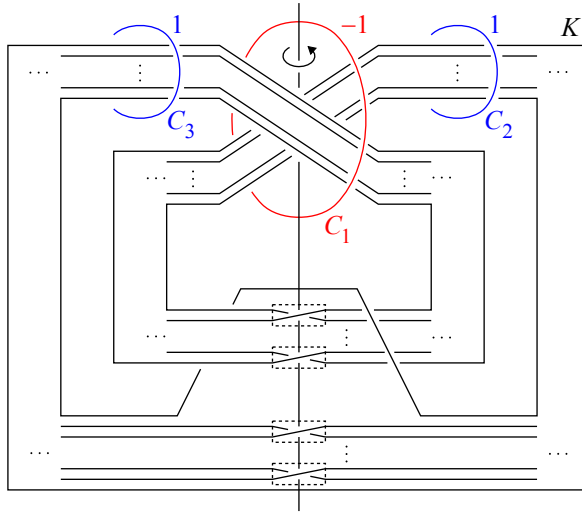


Figure 15. A strongly invertible position of the link $K \cup C_1 \cup C_2 \cup C_3$. The dashed boxes total $2k - 1$: k in the upper part and $k - 1$ in the lower.

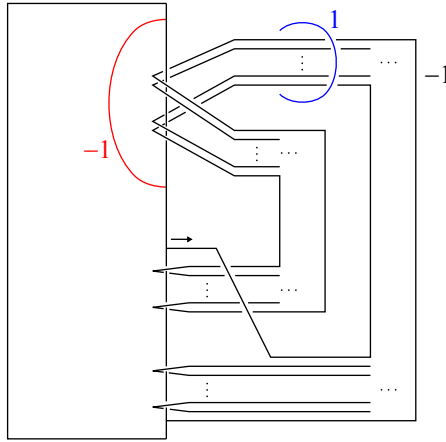


Figure 16. The quotient by the involution around the axis.

surgeries and tangle replacements on the middle right portion of the figure yields the link diagram in the bottom portion, which we denote by ℓ .

Claim 3.2. *The double branched cover of the link ℓ is an L -space.*

Proof. Using the Goeritz matrix derived from a checkerboard coloring of the diagram of ℓ (Figure 17, bottom), we compute $\det \ell = 12k^2 + 2k$. By smoothing the $2k - 1$ crossings indicated by the dashed box in that figure, as shown in Figure 18, we obtain the links (or knots) ℓ_∞^i ($i = 1, \dots, 2k - 1$) and a knot ℓ_0 .

A computation gives $\det \ell_\infty^i = 12k^2 + 2k - (6k + 1)i$ and $\det \ell_0 = 6k + 1$. Hence,

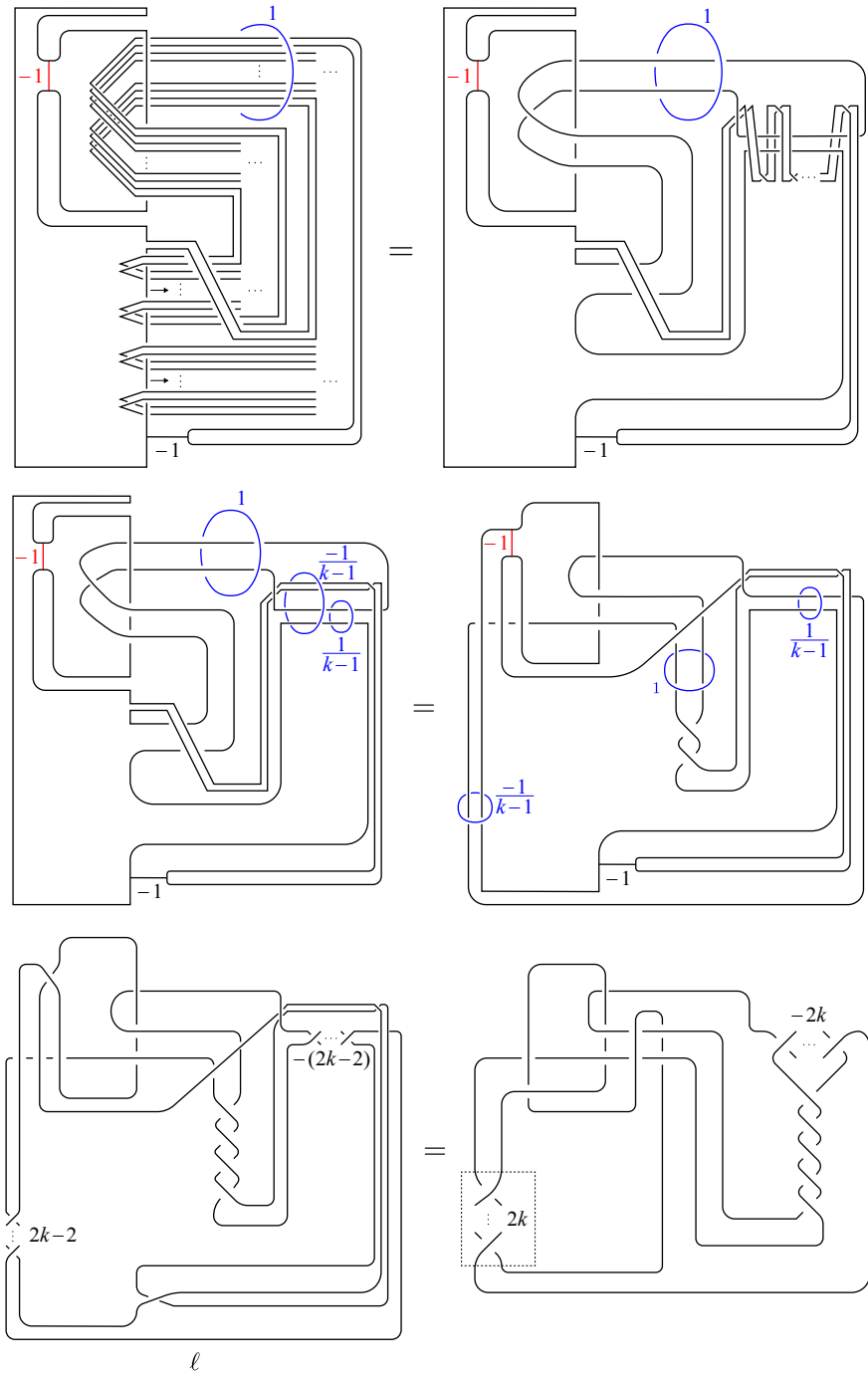


Figure 17. Three deformation steps for the quotient from Figure 16. In the bottom part, integers indicate the number of half-twists: right-handed if positive, left-handed otherwise.

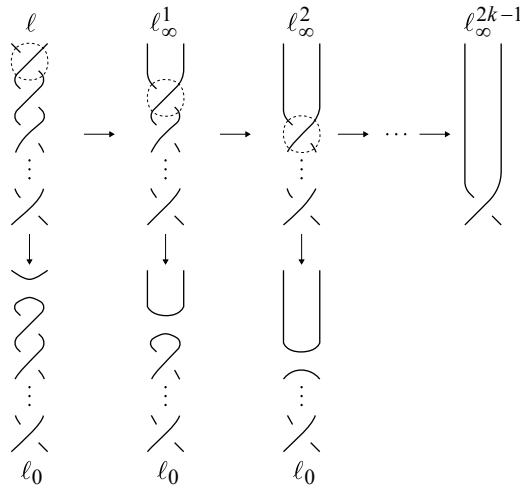


Figure 18. Smoothing the crossings indicated by the dashed circle.

$$\det l_\infty^i = \det l_\infty^{i+1} + \det l_0 \quad \text{for } i = 1, \dots, 2k-2, \quad \text{and} \quad \det l = \det l_\infty^1 + \det l_0.$$

By [14, Proposition 2.1] and [15, Proposition 2.1], it suffices to show that the double branched covers of l_0 and l_∞^{2k-1} are both L -spaces.

As shown in Figure 19, the knot l_0 is the Montesinos knot $M\left(-\frac{2}{3}, \frac{1}{2}, \frac{2k}{6k-1}\right)$, which is not quasi-alternating by [6]. Its double branched cover is the Seifert fibered

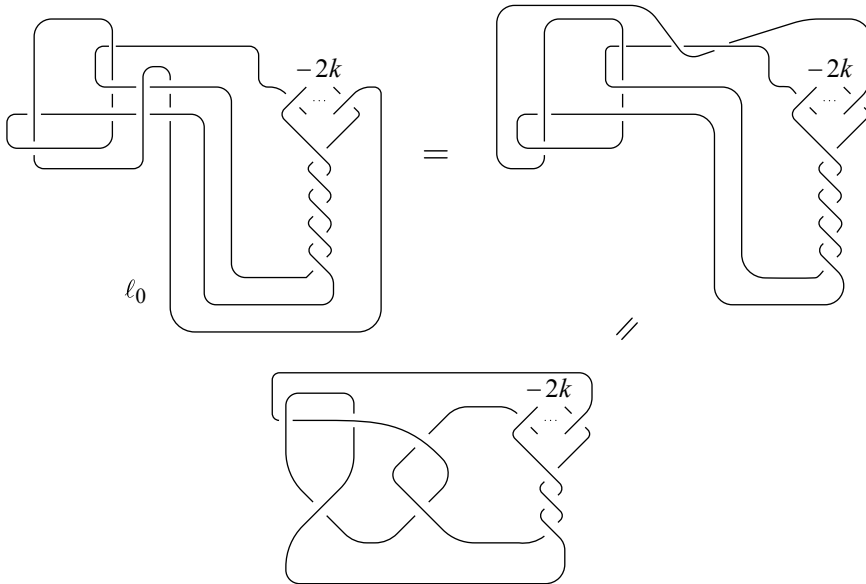


Figure 19. l_0 is a Montesinos knot.

space

$$M(0; -\frac{2}{3}, \frac{1}{2}, \frac{2k}{6k-1}) = M(-1; \frac{1}{2}, \frac{2k}{6k-1}, \frac{1}{3}),$$

following the convention of [11], consistent with [2; 20].

According to [10; 11], such a Seifert fibered space $M(-1; r_1, r_2, r_3)$ with $1 \geq r_1 \geq r_2 \geq r_3 \geq 0$ is an L -space if and only if there are no relatively prime integers $m > a > 0$ satisfying $mr_1 < a < m(1 - r_2)$ and $mr_3 < 1$. For $r_1 = \frac{1}{2}$, $r_2 = \frac{2k}{6k-1}$, and $r_3 = \frac{1}{3}$, the condition $mr_3 < 1$ with $m > a > 0$ gives $a = 1$, $m = 2$, but this violates $mr_1 < a$. Thus, the double branched cover of ℓ_0 is an L -space.

Similarly, Figure 20 shows that ℓ_∞^{2k-1} is the Montesinos knot $M(\frac{2}{5}, -\frac{1}{2}, \frac{2k}{14k-1})$,

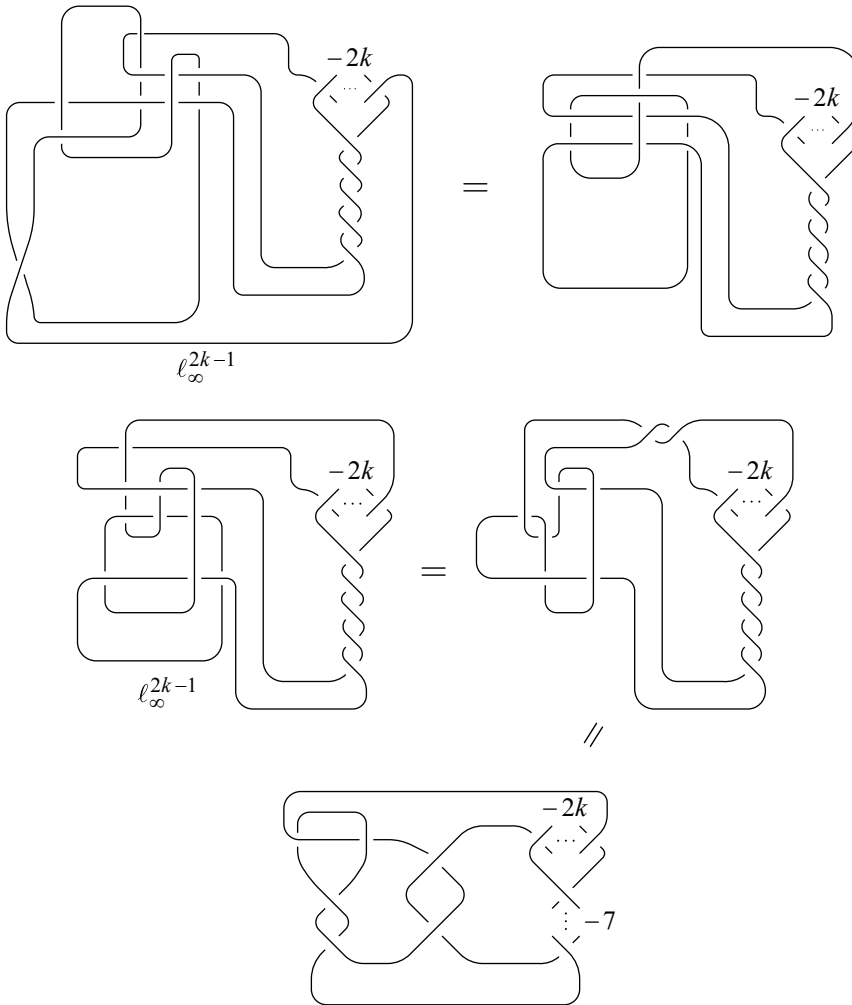


Figure 20. A deformation of ℓ_∞^{2k-1} , showing that it is a Montesinos knot.

and its double branched cover is

$$M\left(0; \frac{2}{5}, -\frac{1}{2}, \frac{2k}{14k-1}\right) = M\left(-1; \frac{1}{2}, \frac{2}{5}, \frac{2k}{14k-1}\right).$$

Set $r_1 = \frac{1}{2}$, $r_2 = \frac{2}{5}$, and $r_3 = \frac{2k}{14k-1}$. The condition $mr_3 < 1$ implies $m < 7 - \frac{1}{2k}$, so $m = 2, 3, \dots, 6$. But for these values, there exists no integer a satisfying $\frac{1}{2}m < a < \frac{3}{5}m$. Therefore, the double branched cover of ℓ_∞^{2k-1} is an L -space. \square

By the Montesinos trick and Claim 3.2, the proof of Proposition 3.1 is complete. \square

4. Hyperbolicity

To complete the proof of Theorem 1.1, we show that K_n is hyperbolic. To this end, we prove that the braid β_n representing K_n , regarded as an element of the mapping class group of a punctured 2-disk, is pseudo-Anosov in the sense of the Nielsen–Thurston classification. We apply the criterion of Bestvina and Handel [3], see also [4].

Let D be a 2-disk, and let $P = \{p_1, \dots, p_k\}$ be a set of k punctures. Take k small circles c_i ($i = 1, \dots, k$), each centered at p_i , such that the interior of c_i contains no other punctures. Choose a finite graph G embedded in D such that it is homotopy equivalent to $D \setminus P$, and contains $C = \{c_1, \dots, c_k\}$ as a subgraph. We allow G to have loops, but assume that it has no vertices of valence 1 or 2. Let $V(G)$ and $E(G)$ denote the sets of vertices and edges of G , respectively. A *walk* τ in G is a finite sequence alternating between vertices and edges $(v_1, e_1, v_2, e_2, \dots, v_\ell, e_\ell, v_{\ell+1})$, where $v_1, \dots, v_{\ell+1} \in V(G)$ and $e_1, \dots, e_\ell \in E(G)$ such that the endpoints of e_i are v_i and v_{i+1} . There is no confusion in denoting this walk by $\tau = e_1 e_2 \cdots e_\ell$. Let $W(G)$ be the set of all walks in G .

A homeomorphism f on D that preserves the set of punctures P induces a *graph map*

$$g: (V(G), W(G)) \rightarrow (V(G), W(G)),$$

which preserves the set C set-wise. (The definition of a graph map involves the notion of a *fibred surface* associated with G , but we omit those details.)

Definition 4.1. A graph map g is said to be *efficient* if there are no integers $m \geq 1$ and edges $e \in E(G)$ such that

$$g^m(e) = \cdots e_i e_i \cdots$$

for some $e_i \in E(G)$. If such integer $m \geq 1$ and an edge $e \in E(G)$ exist, then we say that $g^m(e)$ has a *back track*.

Let $\text{Vec}(G)$ be the real vector space spanned by $E(G)$. Each walk in G determines an element of $\text{Vec}(G)$ by mapping to the linear combination in which the coefficient

of each e_i is its multiplicity in the walk. For a graph map g , we define its transition matrix T_g as the linear transformation

$$\mathcal{T}_g : \text{Vec}(G) \rightarrow \text{Vec}(G)$$

that maps each edge e to $g(e)$, regarded as an element of $\text{Vec}(G)$.

Let C^{pre} be the set of edges $e \in E(G)$ such that $g^m(e)$ is contained in C for some $m \geq 1$, and define the set of *real edges* by

$$E^{\text{re}}(G) = E(G) \setminus (C \cup C^{\text{pre}}).$$

By the definition of g and C^{pre} , the transition matrix \mathcal{T}_g has the block form

$$\mathcal{T}_g = \begin{pmatrix} C & \mathcal{A} & \mathcal{B} \\ 0 & C^{\text{pre}} & \mathcal{D} \\ 0 & 0 & \mathcal{T}_g^{\text{re}} \end{pmatrix},$$

where C , C^{pre} and $\mathcal{T}_g^{\text{re}}$ are the transition matrices associated with C , C^{pre} and $E^{\text{re}}(G)$.

Recall that a nonnegative square matrix M is called *irreducible* if, for any indices (i, j) , there exists $m \geq 1$ such that the (i, j) -entry of M^m is positive. By the Perron–Frobenius theorem, such a matrix M has a real, positive eigenvalue greater than the absolute values of all other eigenvalues. This is called the *Perron–Frobenius eigenvalue* and is denoted by $\lambda(M)$ (it coincides with the spectral radius of M).

Theorem 4.2 [3]. *Let f be a homeomorphism on D that preserves the set of punctures P , and g be an induced graph map for f . Suppose that*

- g is efficient, and
- the transition matrix $\mathcal{T}_g^{\text{re}}$ with respect to the real edges is irreducible with $\lambda(\mathcal{T}_g^{\text{re}}) > 1$.

Then the mapping class of f (up to isotopy) is pseudo-Anosov with dilatation equal to $\lambda(\mathcal{T}_g^{\text{re}})$.

Lemma 4.3. *The mapping class corresponding to the braid*

$$\beta_n = X_n^3 \cdot [-1, -2, \dots, -(n-1), n, n-1, n-2, \dots, 2, 1, 1, 2, 3, \dots, n] \quad (n \geq 2)$$

is pseudo-Anosov.

Proof. When $n = 2$, the braid β_2 is pseudo-Anosov (confirmed using Sage [17]). Assume that $n \geq 3$.

Consider the braid

$$\beta'_n = [n, n-1, n-2, \dots, 2, 1, 1, 2, 3, \dots, n] \cdot X_n^3 \cdot [-1, -2, \dots, -(n-1)],$$

which is conjugate to β_n .

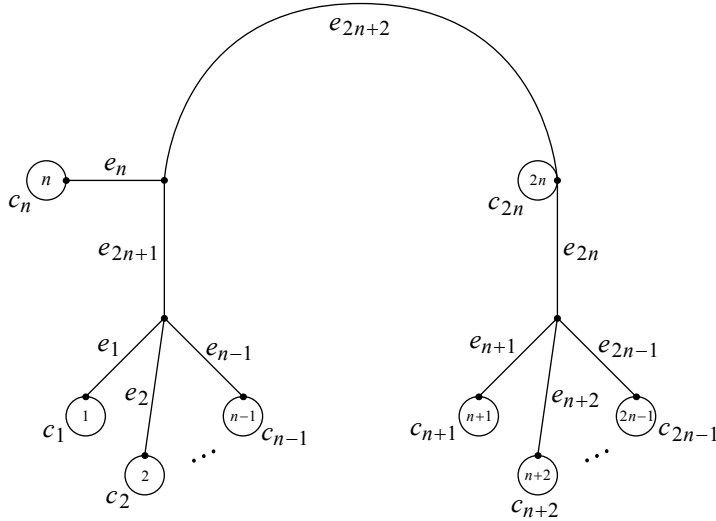


Figure 21. The graph G embedded in $D \setminus \{p_1, \dots, p_{2n}\}$. Each integer inside a circle c_i represents a puncture.

Let G be a graph embedded in $D \setminus \{p_1, \dots, p_{2n}\}$, as shown in Figure 21. (Note that this graph was chosen so that Theorem 4.2 can be applied.) Figures 22 and 23 illustrate a deformation of G induced by the braid β'_n . The induced graph map g acts on the real edges as follows:

$$\begin{aligned}
 g(e_i) &= e_{n+1+i} \text{ for } i = 1, \dots, n - 1, \\
 g(e_{n+i}) &= e_i \text{ for } i = 1, \dots, n - 1, \\
 g(e_n) &= e_{2n+1}e_{n-1}c_{n-1} \cdots e_{2n}c_{2n}e_{2n}e_{n+1}, \\
 g(e_{2n}) &= e_{2n+1}, \\
 g(e_{2n+1}) &= e_{n+1}c_{n+1}e_{n+1} \cdots e_{2n}c_{n-1}e_{2n+2}, \text{ and} \\
 g(e_{2n+2}) &= e_n c_n e_n \cdots e_{n-1} e_{2n+1} e_n.
 \end{aligned}$$

Note that e_1, \dots, e_{2n+2} are real edges, and $g(e)$ has no back tracks for any real edge e .

Suppose that $g^k(e)$ has a back track for some $e \in E^{\text{re}}(G)$ and $k \geq 1$. Then it must occur at e_n , that is, $g^k(e)$ contains a subwalk of the form

$$g^k(e) = \cdots e_{2n+1}e_n e_{2n+1} \cdots .$$

This implies that $g^{k-1}(e)$ contains a subwalk of the form

$$g^{k-1}(e) = \cdots e_{2n+2}e_{2n} \cdots \quad \text{or} \quad \cdots e_{2n}e_{2n+2} \cdots ,$$

which does not occur by Figure 23. Hence, g is efficient.

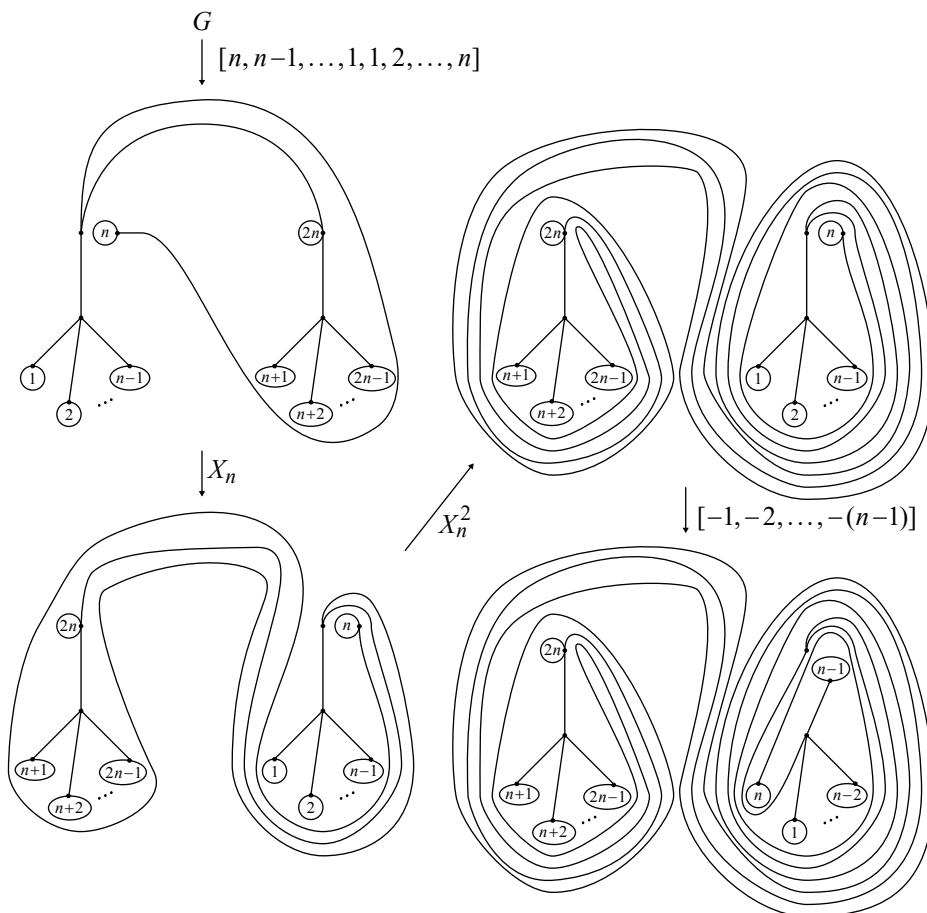


Figure 22. The deformation of G induced by the braid β'_n .

Since $g(e_n)$ passes through all real edges of G , and for any real edge e , there exists $k \geq 1$ such that $g^k(e)$ passes through e_n , it follows that some power $(\mathcal{T}_g^{\text{re}})^m$ of the transition matrix has all entries that are positive integers. We can choose m such that the trace of $(\mathcal{T}_g^{\text{re}})^m$ is greater than $\#E^{\text{re}}(G)$. Therefore, $(\mathcal{T}_g^{\text{re}})$ is irreducible and satisfies $\lambda(\mathcal{T}_g^{\text{re}}) > 1$. By Theorem 4.2, β'_n is pseudo-Anosov, and hence so is β_n . \square

Proposition 4.4. K_n ($n \geq 2$) is a hyperbolic knot.

Proof. Note that K_n can be obtained by the closure of the braid $X_n \beta_n X_n^{-1}$, and $X_n \beta_n X_n^{-1}$ is also pseudo-Anosov by Lemma 4.3.

By Figure 24, the Dehornoy floor of $X_n \beta_n X_n^{-1}$ is greater than 1 (see [7] for definition of the Dehornoy floor). Therefore, by Theorem 1.3 of [7], the closure of β_n , namely K_n , is a hyperbolic knot. \square

Proof of Theorem 1.1. By Propositions 2.12, 3.1 and 4.4, we have the conclusion. \square

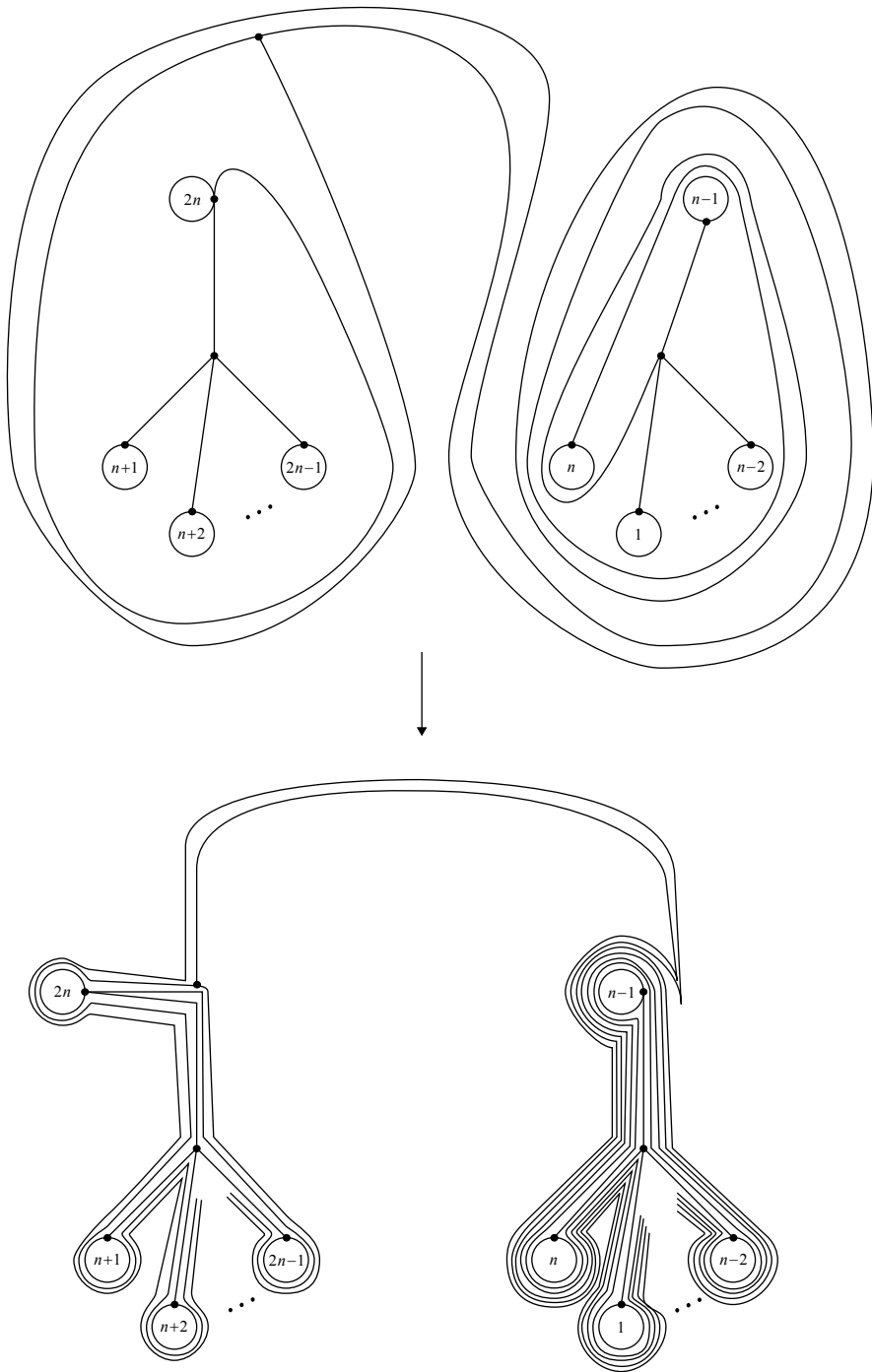


Figure 23. Continued from Figure 22. Here, the deformation from the right bottom in Figure 22 to the bottom of this figure is by an isotopy.

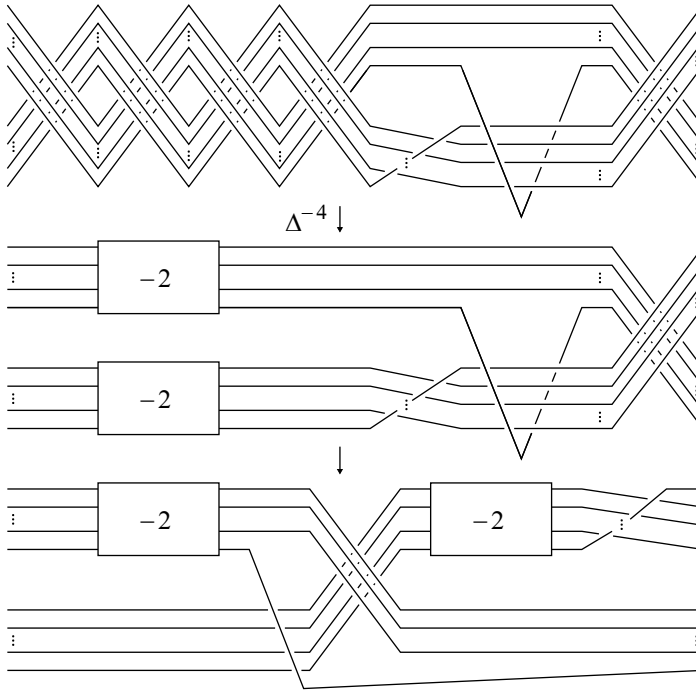


Figure 24. Top: the braid $X_n \beta_n X_n^{-1}$. Middle: the braid obtained by rewriting the braid $\Delta^{-4} X_n \beta_n X_n^{-1}$, where Δ is the Garside fundamental $2n$ -braid. Here the box with -2 indicates two left-handed full-twists. Bottom: the braid obtained by further modifying the middle braid. Since this braid is σ_1 -positive, we have $\Delta^4 <_D X_n \beta_n X_n^{-1}$ where $<_D$ is the Dehornoy order.

Acknowledgments

The author thanks Masakazu Teragaito for his thoughtful guidance and helpful discussions about this work. Additional thanks go to Tetsuya Ito, who gave helpful comments for the HOMFLY polynomial. The author was supported by JST SPRING, grant number JPMJSP2132.

References

- [1] C. Anderson, K. L. Baker, X. Gao, M. Kegel, K. Le, K. Miller, S. Onaran, G. Sangston, S. Tripp, A. Wood, and A. Wright, “L-space knots with tunnel number > 1 by experiment”, *Exp. Math.* **32**:4 (2023), 600–614. MR
- [2] K. L. Baker and M. Kegel, “Census L-space knots are braid positive, except for one that is not”, *Algebr. Geom. Topol.* **24**:1 (2024), 569–586. MR
- [3] M. Bestvina and M. Handel, “Train-tracks for surface homeomorphisms”, *Topology* **34**:1 (1995), 109–140. MR

- [4] E. Hironaka and E. Kin, “A family of pseudo-Anosov braids with small dilatation”, *Algebr. Geom. Topol.* **6** (2006), 699–738. MR
- [5] J. Hom, R. Lipschitz, and D. Ruberman, “Thirty yeas of Floer theory for 3–manifolds”, 2017, available at <https://www.birs.ca/cmo-workshops/2017/17w5011/report17w5011.pdf>.
- [6] A. Issa, “The classification of quasi-alternating Montesinos links”, *Proc. Amer. Math. Soc.* **146**:9 (2018), 4047–4057. MR
- [7] T. Ito, “Braid ordering and the geometry of closed braid”, *Geom. Topol.* **15**:1 (2011), 473–498. MR
- [8] T. Ito, “A note on HOMFLY polynomial of positive braid links”, *Internat. J. Math.* **33**:4 (2022), art. id. 2250031, 18 pp. MR
- [9] T. Ito, “An obstruction of Gordian distance one and cosmetic crossings for genus one knots”, *New York J. Math.* **28** (2022), 175–181. MR
- [10] P. Lisca and G. Matić, “Transverse contact structures on Seifert 3-manifolds”, *Algebr. Geom. Topol.* **4** (2004), 1125–1144. MR
- [11] P. Lisca and A. I. Stipsicz, “Ozsváth–Szabó invariants and tight contact 3-manifolds, III”, *J. Symplectic Geom.* **5**:4 (2007), 357–384. MR
- [12] J. M. Montesinos, “Surgery on links and double branched covers of S^3 ”, pp. 227–259 in *Knots, groups, and 3-manifolds: papers dedicated to the memory of R. H. Fox*, Ann. of Math. Stud. **84**, Princeton Univ. Press, 1975. MR
- [13] Y. Ni, “Knot Floer homology detects fibred knots”, *Invent. Math.* **170**:3 (2007), 577–608. MR
- [14] P. Ozsváth and Z. Szabó, “On knot Floer homology and lens space surgeries”, *Topology* **44**:6 (2005), 1281–1300. MR
- [15] P. Ozsváth and Z. Szabó, “On the Heegaard Floer homology of branched double-covers”, *Adv. Math.* **194**:1 (2005), 1–33. MR
- [16] L. Rudolph, “Quasipositivity and new knot invariants”, *Rev. Mat. Univ. Complut. Madrid* **2**:1 (1989), 85–109. MR
- [17] *SageMath, the Sage Mathematics Software System*, 2024, available at <http://www.sagemath.org>. Version 10.3.
- [18] J. R. Stallings, “Constructions of fibred knots and links”, pp. 55–60 in *Algebraic and geometric topology* (Stanford, CA, 1976), Proc. Sympos. Pure Math. **32** part 2, Amer. Math. Soc., 1978. MR
- [19] H. Takioka, “ $2n$ -moves and the Γ -polynomial for knots”, *Kyungpook Math. J.* **64**:3 (2024), 511–518. MR
- [20] M. Teragaito, “Hyperbolic L -space knots and their formal semigroups”, *Internat. J. Math.* **33**:12 (2022), art. id. 2250080, 20 pp. MR
- [21] J. M. Van Buskirk, “Positive knots have positive Conway polynomials”, pp. 146–159 in *Knot theory and manifolds* (Vancouver, BC, 1983), Lecture Notes in Math. **1144**, Springer, 1985. MR

Received July 7, 2025. Revised November 5, 2025.

KEISUKE HIMENO
 FACULTY OF EDUCATION
 YAMAGUCHI UNIVERSITY
 YAMAGUCHI
 JAPAN
 himekei@yamaguchi-u.ac.jp

THE RANKIN–SELBERG INTEGRAL ON \mathbf{GSp}_2 FOR SQUARE-FREE LEVELS

SEIJI KUGA AND MASAO TSUZUKI

We explicitly compute the Rankin–Selberg type integral introduced by Piatetski-Shapiro over adèles for vector-valued Siegel cusp forms of square-free levels $\Gamma_0(N)$. On the way, for particular test functions in the Bessel models of irreducible admissible representations, exact evaluations of the local zeta integrals are given.

1. Introduction

Lately, the Bessel periods on the symplectic similitude group $G := \mathbf{GSp}_2$ of rank 2 have been gaining in importance in the arithmetic of Siegel modular forms. The notion of Bessel period and the associated local models of admissible representations of G were originally introduced by Novodvorski and Piatetski-Shapiro in their work [27; 26; 29], where a modern treatment of Andrianov’s integral representation [1; 2] of the spinor L -function for Siegel modular forms was developed. Related to this, we should also mention independent work by Sugano [42], which handles vector-valued Hilbert–Siegel cusp forms on arithmetic groups defined by maximal orders of indefinite division algebras. The Bessel period also plays an essential role in a formulation of Böcherer’s conjecture posed by Liu [22] in a style of the refined Gan–Gross–Prasad conjecture. This version of Böcherer’s conjecture has been completely solved by Furusawa and Morimoto [9; 10].

In this paper, we compute the Rankin–Selberg type integral à la Piatetski-Shapiro for holomorphic vector-valued Siegel cusp forms of square-free levels explicitly, using results of Pitale and Schmidt [32] on local new vectors on $G(\mathbb{Q}_p)$ ($p < \infty$); since we rely on the local-global method, our result is immediately applied to nonholomorphic Siegel cusp forms of discrete series type (cf. [20; 23]).

We explain the background of our investigation for scalar-valued Siegel modular forms of weight l . Let E be an imaginary quadratic field of discriminant $D < 0$; depending on E , we define a closed \mathbb{Q} -subgroup $G^\# \hookrightarrow G$ as in Section 2.1 and form the Eisenstein series $E(s, \Lambda, \mu)$ ($s \in \mathbb{C}$) on $G^\#(\mathbb{A})$ attached to an ideal class

MSC2020: primary 11F46; secondary 11F67.

Keywords: Siegel modular forms, Rankin–Selberg integral.

character Λ of E and a finite order character $\mu : \mathbb{A}^\times / \mathbb{Q}^\times \rightarrow \mathbb{C}^1$ (see Section 3.1). Our focus is an asymptotic formula as $N + l \rightarrow \infty$ of the average

$$(1-1) \quad \frac{1}{[\mathbf{Sp}_2(\mathbb{Z}) : \Gamma_0(N)]} \sum_{\varphi \in \mathcal{B}(l, N)} \langle E(s, \Lambda, \mu) \mid \varphi \rangle B^{E, \Lambda}(\varphi),$$

where $B^{E, \Lambda}(\varphi)$ is the Bessel period of φ in an orthonormal basis $\mathcal{B}(l, N)$ of Siegel cusp forms of weight l and square-free level $\Gamma_0(N)$, and $\langle E(s, \Lambda, \mu) \mid \varphi \rangle$ is the L^2 inner product of $E(s, \Lambda, \mu)$ and φ restricted to $G^\#(\mathbb{A})$. We suppose the ramifications of E/\mathbb{Q} , μ , π are mutually disjoint. Actually, for application [19], we consider a more general average than (1-1) by requiring $\mathcal{B}(l, N)$ to consist of joint eigenforms by Hecke operators and by putting the eigenvalue of such a Hecke operator at each φ in the formation of the sum (1-1). With such generality, we obtain a completely explicit expression of (1-1) in terms of

- the spinor L -function $L(s, \pi, \mu)$ (see Section 4.6) attached to an irreducible cuspidal representation $\pi \cong \bigotimes_p \pi_p$ twisted by μ ,
- an average of Fourier coefficients of a newform φ_π^0 in π , denoted by $R(\varphi_\pi^0, E, \Lambda)$ (see (4-12)) originally introduced in [18],
- an average of the local Bessel periods (5-10) of π_p , or its computed form (5-12).

Precise statements, including the vector-valued case, are given in Proposition 5.1 and Theorem 5.2. Owing to the explicit nature of our result, we are able to show that, asymptotically as the prime level N grows to infinity with the weight l being fixed, the contribution to (1-1) of the old forms, the Yoshida lifts and the Saito–Kurokawa lifts are 0 (see Theorem 5.4).

These results are necessary in our forthcoming paper [19], where we shall work out the geometric side of a relative trace formula whose spectral side is essentially (1-1). We should remark that an average similar to our computed average (5-11) without the factors $\hat{L}(s + 1/2, \pi, \mu)$ and $\hat{f}_S(\pi)$ is studied in [7, §3]. When $E = \mathbb{Q}(i)$, our average (5-11) is almost the same as the one in [5, §8] except the appearance of the local periods $t(\pi, \mu)$.

Outline of paper. In Section 2, we introduce basic objects related to algebraic subgroups of the symplectic group G and Haar measures on their p -adic, real and adelic points. In Section 3.1, the Rankin–Selberg type integral and the global Bessel function for Siegel cusp forms are recalled; after this preparation, the basic identity (Lemma 3.2) is stated, which computes the pairing $\langle E(s, \Lambda, \mu), \varphi \rangle$ in (1-1) in terms of the Mellin transform of the global Bessel function attached to φ [29, Theorem 5.2]. We review the proof to determine a constant depending on Haar measures exactly. In Section 4, we recall the definition of the quantity $R(\varphi, E, \Lambda)$ for an ideal class character Λ , which was originally considered in [18]; we describe its behavior under

the Galois conjugation $\Lambda \mapsto \Lambda^\dagger$ (Lemma 4.6). The local multiplicity one theorem for Bessel models of π_p 's allows us to split $B^{E,\Lambda}(\varphi)$ for $\varphi \in \pi$ corresponding to a pure tensor in $\bigotimes_p \pi_p$ to an Euler product of local Bessel functions for π_p up to a constant. To define our newform φ_π^0 and identify this constant with $R(\varphi_\pi^0, E, \Lambda)$ (see Lemma 4.8), we fix pairs of a local Bessel functional ℓ_p on π_p and a local new vector ξ_p of π_p such that $\ell_p(\xi_p) = 1$; in Section 4.4, we explain that the main result of [32] ensures the existence of such $\{(\ell_p, \xi_p)\}_p$ when π corresponds to a new form on $\Gamma_0(N)$. In Section 6, we compute the local zeta integrals by using results of [32]. The unramified computations are known (see Section 6.2). Other cases, as well as the explicit determination of the local periods of prime level (see Section 6.8), seem to be new, so that some details are given in Sections 6.4, 6.5 and 6.6. In Section 4.6, the functional equation of the L -function is deduced from the basic identity by using Lemma 4.6. Although a “nice” functional equation of the L -function itself is known in a broader generality [39, Lemma 1.2] owing to Arthur’s classification of the discrete spectrum, we include this section for the sake of completeness and to provide a proof free from Arthur’s result, confirming that the local L -factor and the local ε -factor defined by [29] coincides with the ones proposed in [36, Tables 2 and 3]. In Section 5, we prove Proposition 5.1 and Theorem 5.2; by invoking a result by Furusawa and Morimoto [9, Theorem 8.1], we obtain an explicit expression of (1-1) solely in terms of L -functions (Theorem 5.3).

2. Preliminaries

Let V denotes the space of symmetric matrices of degree 2; we consider V as a \mathbb{Z} -scheme by defining

$$V(R) = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \mid a, b, c \in R \right\}$$

for any commutative ring R . The group $\mathbf{GL}_2(R)$ acts on this space from the right as

$$(2-1) \quad V(R) \times \mathbf{GL}_2(R) \ni (T, h) \mapsto Ts(h) := \frac{1}{\det h} {}^t h T h \in V(R).$$

Let G denote the symplectic similitude group

$$G := \left\{ g \in \mathbf{GL}_4 \mid {}^t g \begin{bmatrix} O & 1_2 \\ -1_2 & O \end{bmatrix} g = \nu(g) \begin{bmatrix} O & 1_2 \\ -1_2 & O \end{bmatrix} (\exists \nu(g) \in \mathbf{GL}_1) \right\}$$

(usually denoted by \mathbf{GSp}_2), where $\nu : G \rightarrow \mathbf{GL}_1$ is the similitude character. The center of G is denoted by Z , which coincides with the set of all the invertible scalar matrices of degree 4. Let $P = MN$ be the Siegel parabolic subgroup of G , where

$$\begin{aligned} M &:= \left\{ m(A, c) := \begin{bmatrix} A & O \\ O & {}^t A^{-1} c \end{bmatrix} \mid A \in \mathbf{GL}_2, c \in \mathbf{GL}_1 \right\}, \\ N &:= \left\{ n(X) := \begin{bmatrix} 1_2 & X \\ O & 1_2 \end{bmatrix} \mid X \in V \right\}. \end{aligned}$$

Note that $v(m(A, c)) = c$ and $v(n(X)) = 1$. Let

$$T := \left\{ m\left(\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, c\right) \mid a, b, c \in \mathbf{GL}_1 \right\},$$

which is a maximal torus of G . The Weyl group $W := N_G(T)/T$ of G is generated by the images of

$$s_1 := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad s_2 := \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

2.1. Embedding of groups. Let $E = \mathbb{Q}(\sqrt{D}) \subset \mathbb{C}$ be an imaginary quadratic field of discriminant $D < 0$, so that $\sqrt{D} = i|D|^{1/2}$ with i being the imaginary unit. Let \mathcal{O}_E denote the ring of integers of E . Choose a \mathbb{Z} -basis $\{1, \theta\}$ of \mathcal{O}_E such that

$$(2-2) \quad \theta - \bar{\theta} = -\sqrt{D},$$

where $\tau \mapsto \bar{\tau}$ denotes the nontrivial automorphism of E/\mathbb{Q} .

The differential ideal $\mathfrak{d}_{E/\mathbb{Q}}$, defined by $\mathfrak{d}_{E/\mathbb{Q}}^{-1} := \{\tau \in E \mid \text{tr}_{E/\mathbb{Q}}(\tau \mathcal{O}_E) \subset \mathbb{Z}\}$, is $\sqrt{D}\mathcal{O}_E$; thus a symplectic \mathbb{Z} -bilinear form $\langle \cdot, \cdot \rangle : \mathcal{O}_E^2 \times \mathcal{O}_E^2 \rightarrow \mathbb{Z}$ is defined by

$$\langle x, y \rangle := \text{tr}_{E/\mathbb{Q}} \left(\frac{-1}{\sqrt{D}} \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \right), \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathcal{O}_E^2.$$

Note that the following vectors of \mathcal{O}_E^2 form a symplectic \mathbb{Z} -basis, i.e., $\langle v_i^+, v_j^- \rangle = \delta_{ij}$ and $\langle v_i^+, v_j^+ \rangle = \langle v_i^-, v_j^- \rangle = 0$ for all i, j :

$$v_1^+ := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2^+ := \begin{bmatrix} \theta \\ 0 \end{bmatrix}, \quad v_1^- := \begin{bmatrix} 0 \\ -\theta \end{bmatrix}, \quad v_2^- := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For any ring R , set

$$G^\#(R) := \{h \in \mathbf{GL}_2(\mathcal{O}_E \otimes_{\mathbb{Z}} R) \mid \det(h) \in R^\times\}.$$

For $h \in G^\#(R)$, viewed as an R -linear automorphism of $(\mathcal{O}_E \otimes_{\mathbb{Z}} R)^2$, let $\iota_\theta(h)$ denote the 4×4 matrix representing h in the R -basis $\{v_1^+, v_2^+, v_1^-, v_2^-\}$ of $(\mathcal{O}_E \otimes_{\mathbb{Z}} R)^2$, i.e.,

$$(2-3) \quad [h(v_1^+), h(v_2^+), h(v_1^-), h(v_2^-)] = [v_1^+, v_2^+, v_1^-, v_2^-] \iota_\theta(h).$$

Then $\iota_\theta(h) \in G(R)$ and $v(\iota_\theta(h)) = \det h$. Thus, we have an embedding

$$\iota_\theta : G^\#(R) \rightarrow G(R).$$

Let $B^\#$ be the Borel subgroup of $G^\#$ such that $B^\#(R)$ coincides with the set of all points

$$(2-4) \quad \begin{bmatrix} a\tau & \beta \\ 0 & \tau^{-1} \end{bmatrix}, \quad \tau \in (E \otimes_{\mathbb{Q}} R)^\times, a \in R^\times, \beta \in R.$$

Let $N^\#$ denote the unipotent radical of $B^\#$, i.e., $N^\#(R)$ is the set of points (2-4) with $\tau = a = 1$. Set $Z^\# := \{a1_2 \mid a \in \mathbf{GL}_1\}$; then $Z^\#$ is a subgroup of the center of $G^\#$ of index 2.

Let $I_\theta : E \rightarrow \mathbf{M}_2(\mathbb{Q})$ denote the matrix representation of the regular representation of the \mathbb{Q} -algebra E with respect to the basis $\{1, \theta\}$, i.e.,

$$(2-5) \quad [\tau, \tau\theta] = [1, \theta] I_\theta(\tau), \quad \tau \in E,$$

or, explicitly,

$$(2-6) \quad I_\theta(a + b\theta) = \begin{bmatrix} a & -bN_{E/\mathbb{Q}}(\theta) \\ b & a + b \operatorname{tr}_{E/\mathbb{Q}}(\theta) \end{bmatrix}, \quad a, b \in \mathbb{Q}.$$

For $\beta = b_2 + b_3\theta \in E_R$ with $b_2, b_3 \in R$, define an element X_β of $V(R)$ as

$$(2-7) \quad X_\beta := \begin{bmatrix} b_1 & b_2 \\ b_2 & b_3 \end{bmatrix} \quad \text{with } b_1 := -b_2 \operatorname{tr}_{E/\mathbb{Q}}(\theta) - b_3 N_{E/\mathbb{Q}}(\theta).$$

We have

$$(2-8) \quad \iota_\theta \left(\begin{bmatrix} \tau & 0 \\ 0 & a\tau^{-1} \end{bmatrix} \right) = m(I_\theta(\tau), a), \quad a \in R^\times, \tau \in E_R^\times,$$

$$(2-9) \quad \iota_\theta \left(\begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \right) = n(X_\beta), \quad \beta \in E_R.$$

By formulas (2-6) and (2-8), we have $\iota_\theta(a1_2) = a1_4$ for $(a \in R^\times)$, which implies

$$(2-10) \quad \iota_\theta(Z^\#) = Z.$$

We then have $N_{E/\mathbb{Q}}(x + \theta y) = [x, y] T \begin{bmatrix} x \\ y \end{bmatrix}$ ($x, y \in R$) with

$$(2-11) \quad T_\theta := \begin{bmatrix} 1 & 2^{-1} \operatorname{tr}_{E/\mathbb{Q}}(\theta) \\ 2^{-1} \operatorname{tr}_{E/\mathbb{Q}}(\theta) & N_{E/\mathbb{Q}}(\theta) \end{bmatrix}.$$

Let $V^{T_\theta}(\mathbb{Q})$ be the orthogonal of T_θ in $V(\mathbb{Q})$ with respect to the nondegenerate quadratic form $\operatorname{tr}(XY)$ on $V(\mathbb{Q})$, i.e.,

$$V^{T_\theta}(\mathbb{Q}) := \{X \in V(\mathbb{Q}) \mid \operatorname{tr}(T_\theta X) = 0\}.$$

Note that $\operatorname{tr}(T_\theta^2) = 1 + (\operatorname{tr}_{E/\mathbb{Q}}(\theta)/2)^2 + N_{E/\mathbb{Q}}(\theta)^2 > 0$, $\det(T_\theta) = -D/4 > 0$ by (2-2). We have $V^{T_\theta}(\mathbb{Q}) = \{X_\beta \mid \beta \in E\}$, and

$$(2-12) \quad V(\mathbb{Q}) = \mathbb{Q}T_\theta \oplus V^{T_\theta}(\mathbb{Q}), \quad \iota_\theta(N^\#) = n(V^{T_\theta})$$

by (2-9). The group M acts on the space of rational homomomorphisms $\operatorname{Hom}(N, \mathbb{G}_a)$ by the rule $\operatorname{Ad}^*(m)\chi(n) = \chi(m^{-1}nm)$ for $m \in M, n \in N$ and $\chi \in \operatorname{Hom}(N, \mathbb{G}_a)$. For $T \in V$, let M_T denote the stabilizer of $\chi_T : n(X) \mapsto \operatorname{tr}(TX)$, and M_T° the identity component with respect to the Zariski topology. Then, for any \mathbb{Q} -algebra R ,

$$M_{T_\theta}^\circ(R) = \{m(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau)) \mid \tau \in E_R\}.$$

Define a bilinear form on $V(R)$ as $(X, Y) := -\text{tr}(XY^\dagger)$ and let $\mathbf{SO}(V(R))$ be the special orthogonal group of (\cdot, \cdot) .

2.2. Open compact subgroups at finite places. Let p be a prime number. Set $E_p := E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and $\mathcal{O}_{E,p} := \mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Then, $\mathbf{K}_p^\# := \mathbf{G}^\#(\mathbb{Z}_p)$, which coincides with the stabilizer in $\mathbf{G}^\#(\mathbb{Q}_p)$ of the $\mathcal{O}_{E,p}$ -lattice $\mathcal{O}_{E,p}^2 \subset E_p^2$, is a maximal compact subgroup of $\mathbf{G}^\#(\mathbb{Q}_p)$, and $\mathbf{K}_p := \mathbf{G}(\mathbb{Z}_p)$ is the standard maximal compact subgroup of $\mathbf{G}(\mathbb{Q}_p)$. For a nonzero ideal $\mathfrak{n} \subset \mathbb{Z}_p$, set

$$\mathbf{K}_0(\mathfrak{n}) := \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{G}(\mathbb{Z}_p) \mid C \in \mathfrak{n} \right\}.$$

For a nonzero ideal $\mathfrak{a} \subset \mathcal{O}_{E,p}$, set

$$\mathbf{K}_0^\#(\mathfrak{a}) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{G}^\#(\mathbb{Z}_p) \mid c \in \mathfrak{a} \right\}.$$

Thus, $\mathbf{K}_0(\mathbb{Z}_p) = \mathbf{K}_p$ and $\mathbf{K}_0^\#(\mathcal{O}_{E,p}) = \mathbf{K}_p^\#$.

Lemma 2.1. *For a nonzero ideal $\mathfrak{n} \subset \mathbb{Z}_p$, we have $\mathbf{K}_0^\#(\mathfrak{n}\mathcal{O}_{E,p}) = \iota_\theta^{-1}(\mathbf{K}_0(\mathfrak{n}))$.*

Proof. Indeed, both $\iota_\theta^{-1}(\mathbf{K}_p)$ and $\mathbf{K}_p^\#$ coincides with the stabilizer of $\mathcal{O}_{E,p}^2 = \langle v_1^+, v_2^+, v_1^-, v_2^- \rangle_{\mathbb{Z}_p}$ in $\mathbf{G}^\#(\mathbb{Q}_p)$. Hence $\iota_\theta^{-1}(\mathbf{K}_p) = \mathbf{K}_p^\#$. In the remaining part of the proof, we suppose $\mathfrak{n} \subset p\mathbb{Z}_p$. Then for $g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{K}_p$, we have $g \in \mathbf{K}_0(\mathfrak{n})$ if and only if $g(v_1^+), g(v_2^+) \in \langle v_1^+, v_2^+ \rangle_{\mathbb{Z}_p} + \mathfrak{n}\langle v_1^-, v_2^- \rangle_{\mathbb{Z}_p}$. For $g = \iota_\theta(h)$ with $h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{K}_p^\#$, this last condition becomes $\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} a\theta \\ c\theta \end{bmatrix} \in \begin{bmatrix} \mathcal{O}_{E,p} \\ \mathfrak{n}\mathcal{O}_{E,p} \end{bmatrix}$, or, equivalently, $c \in \mathfrak{n}\mathcal{O}_{E,p}$. \square

For $N \in \mathbb{Z}_{>0}$, define open compact subgroups $\mathbf{K}_0(N) \subset \mathbf{G}(\mathbb{A}_f)$ and $\mathbf{K}^\#(N\mathcal{O}_E) \subset \mathbf{G}^\#(\mathbb{A}_f)$ by

$$\mathbf{K}_0(N) := \prod_{p < \infty} \mathbf{K}_0(N\mathbb{Z}_p), \quad \mathbf{K}_0^\#(N\mathcal{O}_E) := \prod_{p < \infty} \mathbf{K}_0^\#(N\mathcal{O}_{E,p}).$$

2.3. Maximal compact subgroup at the archimedean place. The identity connected component of $\mathbf{G}(\mathbb{R})$ is $\mathbf{G}(\mathbb{R})^0 = \{g \in \mathbf{G}(\mathbb{R}) \mid \nu(g) > 0\}$. Set $\mathbf{K}_\infty := \mathbf{G}(\mathbb{R})^0 \cap \mathbf{O}(4)$, which is a maximal compact subgroup of $\mathbf{G}(\mathbb{R})^0$ given as

$$\mathbf{K}_\infty = \left\{ \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \mid A, B \in \mathbf{M}_2(\mathbb{R}), A + iB \in \mathbf{U}(2) \right\}.$$

Note that $\mathbf{K}_\infty \subset \mathbf{Sp}_2(\mathbb{R})$. The action of $\mathbf{G}(\mathbb{R})^0$ on the Siegel upper-half space $\mathfrak{h}_2 := \{Z \in \mathbf{M}_2(\mathbb{C}) \mid {}^t Z = Z, \text{Im}(Z) \gg 0\}$, denoted by $(g, Z) \mapsto g\langle Z \rangle$, is defined by the usual formula, i.e.,

$$g\langle Z \rangle := (AZ + B)(CZ + D)^{-1}, \quad g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{G}(\mathbb{R})^0, \quad Z \in \mathfrak{h}_2.$$

The stabilizer of $i1_2 \in \mathfrak{h}_2$ in $\mathbf{G}(\mathbb{R})^0$ coincides with $Z(\mathbb{R})\mathbf{K}_\infty$ and the map $\phi : \mathbf{G}(\mathbb{R})^0 \ni g \mapsto g\langle i1_2 \rangle \in \mathfrak{h}_2$ induces a diffeomorphism $\mathbf{G}(\mathbb{R})^0/Z(\mathbb{R})\mathbf{K}_\infty \cong \mathfrak{h}_2$. Let \mathfrak{p} denote the subspace of $\mathfrak{g} := \text{Lie}(\mathbf{G}(\mathbb{R})^0)$ that is mapped bijectively onto the tangent space of \mathfrak{h}_2 at $i1_2$ under the tangent map of ϕ at 1_4 . Let $\mathfrak{p}^+ (\subset \mathfrak{p}_\mathbb{C})$ be

the space of holomorphic tangent vectors and \mathfrak{p}^- the antiholomorphic ones. Set $\gamma_\infty := \mathfrak{m}(-1_2, -1) = \begin{bmatrix} -1_2 & 0 \\ 0 & 1_2 \end{bmatrix} \in \mathrm{G}(\mathbb{R})$; then $\mathbf{K}_\infty \cup \mathbf{K}_\infty \gamma_\infty$ is a maximal compact subgroup of $\mathrm{G}(\mathbb{R})$. The four vectors

$$u_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad u_2 := \begin{bmatrix} i \\ 0 \end{bmatrix}, \quad u_3 := \begin{bmatrix} 0 \\ -i \end{bmatrix}, \quad u_4 := \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

in $E_{\mathbb{R}}^2$ form an \mathbb{R} -basis such that

$$(2-13) \quad [v_1^+, v_2^+, v_1^-, v_2^-] = [u_1, u_2, u_3, u_4] (b_{\mathbb{R}}^\theta)^{-1}$$

with

$$(2-14) \quad b_{\mathbb{R}}^\theta := \mathfrak{m}(A_\theta, 2^{-1} \sqrt{|D|})^{-1}, \quad A_\theta := \begin{bmatrix} 1 & 2^{-1} \mathrm{tr}_{E/\mathbb{Q}}(\theta) \\ 0 & -2^{-1} \sqrt{|D|} \end{bmatrix} \in \mathbf{GL}_2(\mathbb{R}).$$

Note that $b_{\mathbb{R}}^\theta \in \mathrm{G}(\mathbb{R})^0$ because $v(b_{\mathbb{R}}^\theta)^{-1} = 2^{-1} \sqrt{|D|} > 0$. Set

$$\mathbf{K}_\infty^\# := \mathbf{U}(2) \cap \mathrm{G}^\#(\mathbb{R}) = \{k^\# \in \mathbf{U}(2) \mid \det(k^\#) = \pm 1\},$$

which is a maximal compact subgroup of $\mathrm{G}^\#(\mathbb{R})$. The identity connected component $(\mathbf{K}_\infty^\#)^0$ of $\mathbf{K}_\infty^\#$ is $\mathbf{SU}(2)$, and $\mathbf{K}_\infty^\# = (\mathbf{K}_\infty^\#)^0 \cup (\mathbf{K}_\infty^\#)^0 \delta_\infty$ with $\delta_\infty := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. A general element of $\mathbf{SU}(2)$ is written in the form

$$(2-15) \quad h = \begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix} \quad \text{with } a = a' + ia'', b = b' + ib'' \in \mathbb{C}, |a|^2 + |b|^2 = 1.$$

For such an h , a computation reveals the relation

$$(2-16) \quad [hu_1, hu_2, hu_3, hu_4] = [u_1, u_2, u_3, u_4] \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \\ \text{with } A := \begin{bmatrix} a' & -a'' \\ a'' & a' \end{bmatrix}, B := \begin{bmatrix} b'' & b' \\ b' & -b'' \end{bmatrix}.$$

Lemma 2.2. *We have $\iota_\theta^{-1}(\mathbf{K}_\infty) = (\mathbf{K}_\infty^\#)^0$. For $k_\infty^\# \in (\mathbf{K}_\infty^\#)^0$ as in (2-15), defining $A, B \in \mathbf{M}_2(\mathbb{R})$ as in (2-16), we have*

$$(2-17) \quad \iota_\theta(k_\infty^\#) = b_{\mathbb{R}}^\theta \begin{bmatrix} A & B \\ -B & A \end{bmatrix} (b_{\mathbb{R}}^\theta)^{-1}.$$

Proof. Equation (2-17) follows directly from (2-4), (2-16) and (2-13). From

$$[hu_1, hu_2, hu_3, hu_4] = [u_1, u_2, u_3, u_4](-\gamma)$$

and $-\gamma \notin \mathbf{K}_\infty$, the assertion follows. □

2.4. Haar measures. For locally compact unimodular topological groups H relevant to us, we fix Haar measures η_H on H in the following manner. Let \mathbb{A} be the adèle ring of \mathbb{Q} and $\psi : \mathbb{A}/\mathbb{Q} \rightarrow \mathbb{C}^1$ the basic character.

2.4.1. On the additive group \mathbb{C} , define $d\eta_{\mathbb{C}}(\tau) := dx dy = 2^{-1} |d\tau \wedge d\bar{\tau}|$ for $\tau = x + iy$ ($x, y \in \mathbb{R}$), where $dx dy$ is the Lebesgue measure on \mathbb{R}^2 .

The Haar measure $\eta_{\mathbb{C}^\times}$ on \mathbb{C}^\times (resp. $\eta_{\mathbb{R}^\times}$ on \mathbb{R}^\times) is defined by $d\eta_{\mathbb{C}^\times}(\tau) = |\tau|_{\mathbb{C}}^{-1} d\tau$ for $\tau = x + iy \in \mathbb{C}$ (resp. $|x|_{\mathbb{R}}^{-1} dx$). For each $p < \infty$, E_p^\times (resp. \mathbb{Q}_p^\times) is endowed with the Haar measure $\eta_{E_p^\times}$ such that $\eta_{E_p^\times}(\mathcal{O}_{E,p}^\times) = 1$ (resp. $\eta_{\mathbb{Q}_p^\times}(\mathbb{Z}_p^\times) = 1$). Then, viewing \mathbb{A}_E^\times as the restricted product of E_p^\times , we define $\eta_{\mathbb{A}_E^\times} = \prod_{p \leq \infty} \eta_{E_p^\times}$. Similarly, we set $\eta_{\mathbb{A}^\times} = \prod_{p \leq \infty} \eta_{\mathbb{Q}_p^\times}$.

2.4.2. For a matrix $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, set $X^\dagger := \begin{bmatrix} -d & -b \\ -c & -a \end{bmatrix}$, so that $XX^\dagger = (\det X) 1_2$, and $X \mapsto X^\dagger$ induces an involution of the \mathbb{Q} -vector space $V(\mathbb{Q})$, which is viewed as a quadratic space with the symmetric bilinear form $\text{tr}(XY^\dagger)$. We endow $V(\mathbb{A})$ with the self-dual Haar measure with respect to the self-duality defined by the bicharacter $(X, Y) \mapsto \psi(\text{tr}(XY^\dagger))$. Let $\eta_{\mathbb{A}}$ (resp. $\eta_{\mathbb{A}_E}$) be the self-dual Haar measure on \mathbb{A} (resp. \mathbb{A}_E) with respect to the bicharacter $(x, y) \mapsto \psi(xy)$ (resp. $(\alpha, \beta) \mapsto \psi(\text{tr}_{E/\mathbb{Q}}(\alpha\bar{\beta}))$). A computation shows the relation $-\text{tr}(X_\alpha X_\beta^\dagger) = \text{tr}_{E/\mathbb{Q}}(\alpha\bar{\beta})$ for $\alpha, \beta \in \mathbb{A}_E$, where $X_\alpha, X_\beta \in V^{T_\theta}(\mathbb{A})$ are defined by (2-7). Thus, by the isomorphism $\mathbb{A}_E \ni \beta \mapsto X_\beta \in V^{T_\theta}(\mathbb{A})$, the Haar measure $\eta_{\mathbb{A}_E}$ is transferred to a Haar measure on $N^\#(\mathbb{A})$. Since $\text{tr}(T_\theta T_\theta^\dagger) = -D/2 \neq 0$, we have the orthogonal direct sum decomposition $V(\mathbb{Q}) = \mathbb{Q}T_\theta^\dagger \oplus V^{T_\theta}(\mathbb{Q})$. By this and (2-12), write $X \in V(\mathbb{A})$ as $X = xT_\theta^\dagger + X_\beta = yT_\theta + X_\alpha$ ($x, y \in \mathbb{A}$, $\beta, \alpha \in \mathbb{A}_E$). By the definition of the Haar measures, we have $d\eta_{V(\mathbb{A})}(X) = d\eta_{\mathbb{A}}(x) \otimes d\eta_{\mathbb{A}_E}(\beta)$. Since the change of variables $(x, \beta) \rightarrow (y, \alpha)$ is given as

$$y = \frac{-D}{2} x, \quad \alpha = \beta + 2^{-1} \text{tr}_{E/\mathbb{Q}}(\theta)(x + y) + (N_{E/\mathbb{Q}}(\theta)x - y) \theta$$

and since $|-D/2|_{\mathbb{A}} = 1$, we get $d\eta_{V(\mathbb{A})}(X) = d\eta_{\mathbb{A}}(y) \otimes d\eta_{\mathbb{A}_E}(\alpha)$. Through the identification $V(\mathbb{A}) \ni X \mapsto n(X) \in N(\mathbb{A})$ and $\beta \in \mathbb{A}_E \mapsto \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \in N^\#(\mathbb{A})$, the groups $N(\mathbb{A})$ and $N^\#(\mathbb{A})$ acquire Haar measures, so that the integration formula

$$(2-18) \quad \int_{N(\mathbb{A})} f(n) dn = \int_{\mathbb{A}} \int_{N^\#(\mathbb{A})} f(n(xT_\theta) \iota_\theta(n^\#)) dx dn^\#$$

holds for any $f \in L^1(N(\mathbb{A}))$. Note that $\text{vol}(N(\mathbb{Q}) \backslash N(\mathbb{A})) = \text{vol}(N^\#(\mathbb{Q}) \backslash N^\#(\mathbb{A})) = 1$.

2.4.3. Let $p \leq \infty$. Any element $g^\# \in G^\#(\mathbb{Q}_p)$ is written as

$$(2-19) \quad g^\# = \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tau & 0 \\ 0 & \bar{\tau}a \end{bmatrix} k^\#$$

with $a \in \mathbb{Q}_p^\times$, $\tau \in E_p^\times$, $\beta \in E_p$ and $k^\# \in \mathbf{K}_p^\#$. Then, our Haar measure on $G^\#(\mathbb{Q}_p)$ is symbolically defined as

$$(2-20) \quad d\eta_{G^\#(\mathbb{Q}_p)}(g_p^\#) = |a|_p^2 \cdot d\eta_{E_p}(\beta) d\eta_{\mathbb{Q}_p^\times}(a) d\eta_{E_p^\times}(\tau) d\eta_{\mathbf{K}_p^\#}(k^\#),$$

where $d\eta_{\mathbf{K}_p^\#}(k^\#)$ is the Haar measure on $\mathbf{K}_p^\#$ with volume 1. For $p < \infty$, we have $\int_{\mathbf{K}_p^\#} \eta_{G^\#(\mathbb{Q}_p)} = 1$. On the adèle group $G^\#(\mathbb{A})$, we use the product measure of $\eta_{G^\#(\mathbb{Q}_p)}$

($p \leq \infty$). Then, from (2-20), we get

$$(2-21) \int_{\mathbf{G}^\#(\mathbb{A})} f(g^\#) dg^\# = \frac{\sqrt{|D|}}{2} \int_{\mathbf{N}^\#(\mathbb{A})} \int_{\mathbb{A}^\times} \int_{\mathbb{A}_E^\times} \int_{\mathbf{K}^\#} f\left(n^\# \begin{bmatrix} \tau & 0 \\ 0 & a\bar{\tau} \end{bmatrix} k^\#\right) |a|_{\mathbb{A}}^2 dn^\# d\eta_{\mathbb{A}^\times}(a) d\eta_{\mathbb{A}_E^\times}(\tau) dk^\#,$$

where $dk^\#$ is the normalized Haar measure on $\mathbf{K}^\# = \prod_{p \leq \infty} \mathbf{K}_p^\#$. Note that the measure on $\mathbf{N}^\#(\mathbb{A}) \cong \mathbb{A}_E$ coincides with the $\eta_{\mathbb{A}_E}$ defined above, which equals $(\sqrt{|D|}/2) \times \prod_p \eta_{E_p}$ ($p \leq \infty$) due to the formulas $(\prod_p \eta_{E_p})(\mathbb{A}_E/E) = |D|^{1/2}/2$ [47, Chapter V, §4 Proposition 7] and $\eta_{\mathbb{A}_E}(\mathbb{A}_E/E) = 1$.

2.4.4. We fix $\eta_{\mathbf{K}_\infty}$ so that $\mathrm{vol}(\mathbf{K}_\infty) = 1$. Then we normalize $\eta_{\mathbf{Sp}_2(\mathbb{R})}$ in such a way that the quotient $\eta_{\mathbf{Sp}_2(\mathbb{R})}/\eta_{\mathbf{K}_\infty}$ corresponds to the measure $(\det Z)^{-3} dX dY$ on $\mathfrak{h}_2 \cong \mathbf{Sp}_2(\mathbb{R})/\mathbf{K}_\infty$. Via $\mathbf{G}(\mathbb{R})^0 = \mathbf{Z}(\mathbb{R})^0 \mathbf{Sp}_2(\mathbb{R}) \cong \mathbb{R}_{>0} \times \mathbf{Sp}_2(\mathbb{R})$, $\eta_{\mathbf{G}(\mathbb{R})}$ is defined by demanding that its restriction to $\mathbf{G}(\mathbb{R})^0$ is $\eta_{\mathbb{R}^\times} \otimes \eta_{\mathbf{Sp}_2(\mathbb{R})}$. For $p < \infty$, we fix $\eta_{\mathbf{G}(\mathbb{Q}_p)}$ by demanding $\int_{\mathbf{K}_p} \eta_{\mathbf{G}(\mathbb{Q}_p)} = 1$. Then, $\eta_{\mathbf{G}(\mathbb{A})}$ is defined to be the restricted product of $\eta_{\mathbf{G}(\mathbb{Q}_p)}$ ($p \leq \infty$).

2.5. Idele class characters. Let Λ denote a character of the finite group

$$\mathrm{Cl}(E) := \mathbb{A}_E^\times / E^\times E_{\mathbb{R}}^\times \widehat{\mathcal{O}}_E^\times.$$

As is well known, this group is isomorphic to the ideal class group of E . We regard Λ as an idele class character of E^\times of finite order. Since $\mathbb{A}^\times = \mathbb{Q}^\times (\mathbb{R}_{>0}) \widehat{\mathbb{Z}}^\times$ is contained in the subgroup $E^\times E_{\mathbb{R}}^\times \widehat{\mathcal{O}}_E^\times$, we have

$$(2-22) \quad \Lambda|_{\mathbb{A}^\times} = \mathbf{1}.$$

Let v be a place of E and Λ_v the v -component of Λ , i.e., $\Lambda = \bigotimes_v \Lambda_v$. We say that v is inert in E/\mathbb{Q} , splits in E/\mathbb{Q} or ramifies in E/\mathbb{Q} if $E_v := E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is an unramified field extension of \mathbb{Q}_p , is isomorphic to $\mathbb{Q}_p \oplus \mathbb{Q}_p$, or is a ramified field extension of \mathbb{Q}_p , respectively. Since Λ is supposed to be trivial on $E_{\mathbb{R}}^\times$, we have $\Lambda_\infty = \mathbf{1}$; moreover,

$$(2-23) \quad \text{If } v \text{ is inert in } E/\mathbb{Q}, \text{ then } \Lambda_v = \mathbf{1}.$$

Indeed, since Λ is supposed to be trivial on $\widehat{\mathcal{O}}_E^\times$, the restriction $\Lambda_v|_{\mathcal{O}_{E,v}^\times}$ is trivial. If v is inert and p denote the residue characteristic of E_v , then $p \in \mathbb{Q}_p$ is a prime element of the local field E_v . We have $\Lambda_v(p) = 1$ due to (2-22). Hence, Λ_v is trivial on $p^{\mathbb{Z}} \mathcal{O}_{E,v}^\times = E_p^\times$.

Define the Galois conjugate Λ^\dagger of Λ by setting

$$(2-24) \quad \Lambda^\dagger(\tau) := \Lambda(\bar{\tau}), \quad \tau \in \mathbb{A}_E^\times.$$

We have $\Lambda(\tau)\Lambda^\dagger(\tau) = \Lambda(\tau\bar{\tau}) = 1$ for $\tau \in \mathbb{A}_E^\times$ by (2-22). Hence

$$(2-25) \quad \Lambda^\dagger = \Lambda^{-1} = \bar{\Lambda},$$

where $\bar{\Lambda}$ is the complex conjugate of Λ .

For an idele class character ξ of E^\times , let $\hat{L}(s, \xi)$ be the completed Hecke L -function of ξ , and $L_p(s, \xi)$ its local p -factor for $p \leq \infty$. Then, $L_p(s, \xi^\dagger) = L_p(s, \xi)$ for any $p < \infty$. Indeed, if p is not ramified in E/\mathbb{Q} , the equality is trivial. Suppose that E/\mathbb{Q} is ramified; if ξ_p is a ramified character of E_p^\times , then both L -factors are 1. If ξ_p is unramified, then for any prime element ϖ of E_p , we have $(\xi^\dagger)_p(\varpi) = \xi_p(\bar{\varpi}) = \xi_p(\varpi)$, which implies the equality between local p -factors.

For any character $\mu : \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_{>0} \rightarrow \mathbb{C}^1$, define $\mu_E := \mu \circ N_{E/\mathbb{Q}} : \mathbb{A}_E^\times/E^\times \rightarrow \mathbb{C}^1$. Then, μ_E is Galois invariant, and $\mu_E|_{E_\infty^\times} = \mathbf{1}$. Hence, $(\Lambda\mu_E)^\dagger = \Lambda^\dagger\mu_E^\dagger = \Lambda^{-1}\mu_E$, and

$$(2-26) \quad \hat{L}(s, \Lambda^{-1}\mu_E) = \hat{L}(s, \Lambda^\dagger\mu_E) = \hat{L}(s, \Lambda\mu_E) \quad (\text{Re}(s) > 1).$$

Note that $\hat{L}(s, \Lambda\mu_E)$ is holomorphic except for possible simple poles at $s = 1, 0$, which occurs if and only if both Λ and μ are trivial.

3. Eisenstein series and Rankin–Selberg integral

Let $\Lambda = \bigotimes_{p \leq \infty} \Lambda_p \in \widehat{\text{Cl}}(E)$ and $\mu = \bigotimes_{p \leq \infty} \mu_p \in \widehat{\mathbb{A}^\times/\mathbb{Q}^\times(\mathbb{R}_{>0})^\times}$.

3.1. Eisenstein series. For details, we refer to [14, §19]; the theory on $\mathbf{GL}_2(\mathbb{A}_E)$ developed there carries over into the group $G^\#(\mathbb{A})$ with minor modifications. For a finite-dimensional vector space V over a local field, let $\mathcal{S}(V)$ be the space of all Schwartz–Bruhat functions on V . For $p \leq \infty$, $\phi \in \mathcal{S}(E_p^2)$ and $s \in \mathbb{C}$ with $\text{Re}(s) > 0$, we define a function $f_\phi^{(s, \Lambda_p, \mu_p)} : G^\#(\mathbb{Q}_p) \rightarrow \mathbb{C}$ by

$$(3-1) \quad f_\phi^{(s, \Lambda_p, \mu_p)}(g^\#) = \mu_p(\det g^\#) |\det g^\#|_p^{s+1} \int_{E_p^\times} \phi\left(\begin{bmatrix} 0 & 1 \\ 0 & \tau \end{bmatrix} g^\#\right) \Lambda_p \mu_{E,p}(\tau) |\tau \bar{\tau}|_p^{s+1} d\eta_{E_p^\times}(\tau).$$

When $p = \infty$, we assume that ϕ is $\mathbf{K}_\infty^\#$ -finite. By local Tate theory, the function $s \mapsto f_\phi^{(s, \Lambda_p, \mu_p)}(g^\#)$ is continued meromorphically to \mathbb{C} in such a way that

$$(3-2) \quad f_\phi^{(s, \Lambda_p, \mu_p)}(g^\#) := L_p(s+1, \Lambda_p \mu_{E,p})^{-1} \times f_\phi^{(s, \Lambda_p, \mu_p)}(g^\#)$$

is holomorphic on \mathbb{C} . Then, $f_\phi^{(s, \Lambda_p, \mu_p)}$ belongs to the space $\mathcal{V}^\#(s, \Lambda_p, \mu_p)$ consisting of all smooth functions f on $G^\#(\mathbb{Q}_p)$ satisfying

$$f\left(\begin{bmatrix} \tau & \beta \\ 0 & a\bar{\tau} \end{bmatrix} g^\#\right) = \Lambda_p(\tau)^{-1} \mu_p(a)^{-1} |a|_p^{-(s+1)} f(g^\#)$$

for any $\begin{bmatrix} \tau & \beta \\ 0 & a\bar{\tau} \end{bmatrix} \in B^\#(\mathbb{Q}_p)$ and $g^\# \in G^\#(\mathbb{Q}_p)$.

The Fourier transform $\widehat{\phi}$ of $\phi \in \mathcal{S}(E_p^2)$ is defined by

$$(3-3) \quad \widehat{\phi}(x, y) := \int_{E_p^2} \phi(u, v) \psi(\langle [\begin{smallmatrix} x \\ y \end{smallmatrix}], [\begin{smallmatrix} u \\ v \end{smallmatrix}] \rangle) d\eta_{E_p}(u) d\eta_{E_p}(v).$$

Set $w_0 := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathrm{G}^\#(\mathbb{Q}_p)$. The standard intertwining operator

$$M(s) : \mathcal{V}^\#(s, \Lambda_p, \mu_p) \rightarrow \mathcal{V}^\#(-s, \Lambda_p^{-1}, \mu_p^{-1})$$

is defined on $\mathrm{Re}(s) > 0$ as the absolutely convergent integral

$$M(s)f(g^\#) = \int_{E_p} f(w_0 \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} g^\#) d\eta_{E_p}(\beta), \quad g^\# \in \mathrm{G}^\#(\mathbb{Q}_p).$$

The effect of $M(s)$ on the section $f_\phi^{(s, \Lambda_p, \mu_p)}$ is described by the Fourier transform $\widehat{\phi}$ as

$$(3-4) \quad f_{\widehat{\phi}}^{(-s, \Lambda_p^{-1}, \mu_p^{-1})}(g^\#) = c_p \Lambda_p (\sqrt{|D|})^{-1} |D|^{-s+1/2} \varepsilon_p(s, \Lambda \mu_E, \psi_{E_p}) \frac{L_p(1-s, \Lambda^{-1} \mu_E^{-1})}{L_p(s, \Lambda \mu_E)} M(s) f_\phi^{(s, \Lambda_p, \mu_p)}(g^\#)$$

with $c_p = 1$ if $p < \infty$ and $c_\infty = \sqrt{|D|}/2$ if $p = \infty$, where $\varepsilon(s, \cdot, \psi_{E_p})$ denote Tate’s local epsilon factor defined by the character $\psi_{E_p} := \psi_p \circ \mathrm{tr}_{E_p/\mathbb{Q}_p}$ of E_p and the associated self-dual measure on E_p as usual. Let $\mathcal{S}(\mathbb{A}_E^2)$ be the space of all the Schwartz–Bruhat functions on \mathbb{A}_E^2 . For a decomposable element $\phi = \bigotimes_{p \leq \infty} \phi_p$ in $\mathcal{S}(\mathbb{A}_E^2)$, we define

$$f_\phi^{(s, \Lambda, \mu)}(g^\#) = \prod_{p \leq \infty} f_{\phi_p}^{(s, \Lambda_p, \mu_p)}(g_p^\#), \quad g^\# = (g_p^\#)_p \in \mathrm{G}^\#(\mathbb{A}).$$

Note that $f_\phi^{(s, \Lambda, \mu)}$ is left- $\mathrm{B}^\#(\mathbb{Q})$ -invariant and right- $\mathbf{K}^\#$ -finite. The Eisenstein series attached to $f_\phi^{(s, \Lambda, \mu)}$ is defined by the absolutely convergent series

$$(3-4) \quad E(\phi, s, \Lambda, \mu; g^\#) := \widehat{L}(s+1, \Lambda \mu_E) \sum_{\delta \in \mathrm{B}^\#(\mathbb{Q}) \setminus \mathrm{G}^\#(\mathbb{Q})} f_\phi^{(s, \Lambda, \mu)}(\delta g^\#), \quad g^\# \in \mathrm{G}^\#(\mathbb{A}),$$

for $\mathrm{Re}(s) > 1$. The Fourier transform $\widehat{\phi}$ of $\phi \in \mathcal{S}(\mathbb{A}_E^2)$ is defined by a formula similar to (3-3) with respect to the measure $\eta_{\mathbb{A}_E} \otimes \eta_{\mathbb{A}_E}$. The following properties of the Eisenstein series are standard.

Proposition 3.1. *Let $\phi \in \mathcal{S}(\mathbb{A}_E^2)$, $s \in \mathbb{C}$, $\Lambda \in \widehat{\mathrm{Cl}(E)}$, $\mu \in \widehat{\mathbb{A}^\times/\mathbb{Q}^\times \mathbb{R}_{>0}}$ and $g^\# \in \mathrm{G}^\#(\mathbb{A})$.*

- (i) *The map $s \mapsto E(\phi, s, \Lambda, \mu; g^\#)$ ($\mathrm{Re}(s) > 1$) has a meromorphic continuation to \mathbb{C} , holomorphic in s unless $\Lambda \mu_E \neq \mathbf{1}$, in which case it has possible simple poles only at $s = 1, -1$. For a regular point $s \in \mathbb{C}$, the function $g^\# \mapsto E(\phi, s, \Lambda, \mu; g^\#)$ is an automorphic form on $\mathrm{G}^\#(\mathbb{Q})\mathrm{Z}^\#(\mathbb{A}) \setminus \mathrm{G}^\#(\mathbb{A})$.*

- (ii) We have the functional equation $E(\widehat{\phi}, -s, \Lambda^{-1}, \mu^{-1}; g^\#) = E(\phi, s, \Lambda, \mu; g^\#)$.
- (iii) For a relatively compact subset $\mathcal{N} \subset \mathbb{C}$ on which $s \mapsto E(\phi, s, \Lambda, \mu; g^\#)$ is regular and for a compact set $\mathcal{U} \subset G^\#(\mathbb{A})$, there exist constants $C > 0$ and $N > 0$ such that

$$|E(\phi, s, \Lambda, \mu; \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} g^\#)| \leq C |a|^N \quad (a \in \mathbb{R}_{>1}, g^\# \in \mathcal{U}, s \in \mathcal{N}).$$

3.2. Rankin–Selberg integral and the basic identity. For any cusp form φ on $Z(\mathbb{A})G(\mathbb{Q}) \backslash G(\mathbb{A})$, $\Lambda \in \text{Cl}(E)$ and $s \in \mathbb{C}$, the Rankin–Selberg integral is defined by

$$(3-5) \quad \langle E(\phi, s, \Lambda, \mu), \varphi \rangle := \int_{Z^\#(\mathbb{A})G^\#(\mathbb{Q}) \backslash G^\#(\mathbb{A})} E(\phi, s, \Lambda, \mu; g^\#) \varphi(\iota_\theta(g^\#)) dg^\#.$$

By Proposition 3.1(iii) and the Fourier expansion (4-5), it is straightforward to show the absolute convergence of the integral for $s \in \mathbb{C}$ where the Eisenstein series is regular. For $T \in V(\mathbb{Q})$, define a character $\psi_T : N(\mathbb{A}) \rightarrow \mathbb{C}^\times$ by

$$\psi_T(n(X)) = \psi(\text{tr}(TX)), \quad X \in V(\mathbb{A}).$$

The (T_θ, Λ) -Bessel period of a cusp form φ on $Z(\mathbb{A})G(\mathbb{Q}) \backslash G(\mathbb{A})$ is defined by the integral¹

$$(3-6) \quad B^{T_\theta, \Lambda}(\varphi; g) := \int_{\mathbb{A}_E^\times / E^\times \mathbb{A}^\times} \Lambda(\tau)^{-1} \int_{N(\mathbb{Q}) \backslash N(\mathbb{A})} \psi_{T_\theta}(n)^{-1} \varphi(m(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau))ng) dn d^\times \tau$$

for $g \in G(\mathbb{A})$, where $d^\times \tau$ denote the quotient measure on $\mathbb{A}_E^\times / \mathbb{A}^\times$ (see Section 2.4.1).

Formula (3-7) is the basic identity due to Piatetski-Shapiro [29]. To determine the constant exactly under our normalization of Haar measures, we reproduce the proof briefly.

Lemma 3.2. *Let $\varphi : Z(\mathbb{A}) \backslash G(\mathbb{Q})G(\mathbb{A}) \rightarrow \mathbb{C}$ be a cusp form. For $\text{Re}(s) > 1$, we have*

$$(3-7) \quad \langle E(\phi, s, \Lambda, \mu), \varphi \rangle = \frac{\sqrt{|D|}}{2} L(s+1, \Lambda \mu_E) \int_{\mathbb{A}^\times} \int_{\mathbf{K}_\infty^\# \mathbf{K}_0^\#(\mathcal{O}_E)} (f_\phi^{(s, \Lambda, \mu)}(k^\#) \mu(a) |a|_\mathbb{A}^{s-1} \times B^{T_\theta, \Lambda}(\varphi; m(a1_2, a)\iota_\theta(k^\#))) d^\times a dk^\#.$$

Proof. By substituting (3-4) into (3-5),

$$\begin{aligned} & L(s+1, \Lambda \mu_E)^{-1} \times \langle E(\phi, s, \Lambda, \mu), \varphi \rangle \\ &= \int_{Z^\#(\mathbb{A})G^\#(\mathbb{Q}) \backslash G^\#(\mathbb{A})} \sum_{\gamma \in B^\#(\mathbb{Q}) \backslash G^\#(\mathbb{Q})} f_\phi^{(s, \Lambda, \mu)}(\gamma g^\#) \varphi(\iota_\theta(g^\#)) dg^\# \end{aligned}$$

¹Note that $m(I_\theta(a), N_{E/\mathbb{Q}}(a)) = a1_4$ for $a \in \mathbb{A}^\times$. Then, due to φ being $Z(\mathbb{A})$ -invariant, the integral $B^{T_\theta, \Lambda}(\varphi; g)$ is 0 if $\Lambda|\mathbb{A}^\times \neq 1$. This trivial vanishing does not happen if $\Lambda|\mathbb{A}^\times = 1$.

$$\begin{aligned}
 &= \int_{Z^\#(\mathbb{A})\mathbb{B}^\#(\mathbb{Q})\backslash G^\#(\mathbb{A})} f_\phi^{(s,\Lambda,\mu)}(g^\#) \varphi(\iota_\theta(g^\#)) dg^\# \\
 &= \frac{\sqrt{|D|}}{2} \int_{\mathbb{A}^\times/\mathbb{Q}^\times} \int_{\mathbb{A}_E^\times/\mathbb{A}^\times E^\times} \int_{N^\#(\mathbb{Q})\backslash N^\#(\mathbb{A})} \int_{\mathbf{K}_0^\#(\varrho_E)\mathbf{K}_\infty^\#} f_\phi^{(s,\Lambda,\mu)}(k_\mathfrak{f}^\# k_\infty^\#) |a|_{\mathbb{A}}^{-(s+1)} \\
 &\quad \times \Lambda(\tau)^{-1} \mu(a)^{-1} \varphi(\iota_\theta(n^\#)) m(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau)) \begin{bmatrix} a^{-1} & 1_2 & 0 \\ 0 & & 1_2 \end{bmatrix} \iota_\theta(k_\mathfrak{f}^\# k_\infty^\#) \\
 &\quad \times |a|_{\mathbb{A}}^2 d^\times a d^\times \tau dn^\# dk_\mathfrak{f}^\# dk_\infty^\#,
 \end{aligned}$$

where the last equality is proved using (2-21). By Lemma 3.3, the last expression becomes

$$\begin{aligned}
 &\frac{\sqrt{|D|}}{2} \int_{\mathbb{A}^\times/\mathbb{Q}^\times} \int_{\mathbb{A}_E^\times/\mathbb{A}^\times E^\times} \int_{N(\mathbb{Q})\backslash N(\mathbb{A})} \int_{\mathbf{K}_0^\#(\varrho_E)\mathbf{K}_\infty^\#} f_\phi^{(s,\Lambda,\mu)}(k_\mathfrak{f}^\# k_\infty^\#) |a|_{\mathbb{A}}^{-(s+1)} \\
 &\quad \times \Lambda(\tau)^{-1} \mu(a)^{-1} \sum_{\alpha \in \mathbb{Q}^\times} \varphi(n m(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau)) \begin{bmatrix} \alpha a^{-1} & 1_2 & 0 \\ 0 & & 1_2 \end{bmatrix} \iota_\theta(k_\mathfrak{f}^\#)) \\
 &\quad \times \psi_{T_\theta}(n)^{-1} |a|_{\mathbb{A}}^2 d^\times a d^\times \tau dn dk_\mathfrak{f}^\#.
 \end{aligned}$$

The α -summation and the a -integral over $\mathbb{A}^\times/\mathbb{Q}^\times$ are combined to yield an integral over \mathbb{A}^\times . The change of variables $a \mapsto a^{-1}$ and (3-6) then give the desired formula. □

Lemma 3.3. For $g \in G(\mathbb{A})$,

$$\int_{N^\#(\mathbb{Q})\backslash N^\#(\mathbb{A})} \varphi(\iota_\theta(n^\#)g) dn^\# = \sum_{\alpha \in \mathbb{Q}^\times} \int_{N(\mathbb{Q})\backslash N(\mathbb{A})} \varphi(n \begin{bmatrix} \alpha 1_2 & 0 \\ 0 & & 1_2 \end{bmatrix} g) \psi_{T_\theta}(n)^{-1} dn.$$

Proof. Fix $g \in G(\mathbb{A})$ and define a function ϕ on \mathbb{A} by

$$\phi(x) := \int_{N^\#(\mathbb{Q})\backslash N^\#(\mathbb{A})} \varphi(n(xT_\theta)\iota_\theta(n^\#)g) dn^\#, \quad x \in \mathbb{A}.$$

Since ϕ is a \mathbb{Q} -periodic smooth function on \mathbb{A} , it can be expanded in a Fourier series, which is absolutely and normally convergent:

$$(3-8) \quad \phi(x) = \sum_{\alpha \in \mathbb{Q}} \int_{\mathbb{A}/\mathbb{Q}} \phi(y) \psi(\alpha y)^{-1} dy \times \psi(\alpha x), \quad x \in \mathbb{A}.$$

By (2-12) and (2-18),

$$\begin{aligned}
 &\int_{\mathbb{A}/\mathbb{Q}} \phi(y) \psi(\alpha y)^{-1} dy \\
 &= \int_{\mathbb{A}/\mathbb{Q}} \int_{V^{T_\theta}(\mathbb{Q})\backslash V^{T_\theta}(\mathbb{A})} \varphi(n(yT_\theta + Z)g)^{-1} \psi(\mathrm{tr}(\alpha T_\theta(yT_\theta^\vee + Z))) dy dZ \\
 &= \int_{V(\mathbb{Q})\backslash V(\mathbb{A})} \varphi(n(X)g) \psi_{T_\theta}(n(\alpha X))^{-1} dX,
 \end{aligned}$$

which for $\alpha = 0$ is zero due to the cuspidality of φ . Thus, by setting $x = 0$ in (3-8),

$$\begin{aligned}\phi(0) &= \sum_{\alpha \in \mathbb{Q}^\times} \int_{V(\mathbb{Q}) \backslash V(\mathbb{A})} \varphi(n(X)g) \psi_{T_\theta}(n(\alpha X))^{-1} dX \\ &= \sum_{\alpha \in \mathbb{Q}^\times} \int_{N(\mathbb{Q}) \backslash N(\mathbb{A})} \varphi(n \begin{bmatrix} \alpha^{1/2} & 0 \\ 0 & 1/2 \end{bmatrix} g) \psi_{T_\theta}(n)^{-1} dn.\end{aligned}$$

The last equality is obtained by the change of variables $X \mapsto \alpha^{-1}X$ and by the automorphy of φ together with the relation $n(\alpha^{-1}X) = \begin{bmatrix} \alpha^{-1/2} & 0 \\ 0 & 1/2 \end{bmatrix} n(X) \begin{bmatrix} \alpha^{1/2} & 0 \\ 0 & 1/2 \end{bmatrix}$. \square

4. Automorphic forms and Fourier coefficients

Let $(\mathbb{Z}^2)_{\text{dom}} := \{\lambda = (l_1, l_2) \in \mathbb{Z}^2 \mid l_1 \geq l_2\}$. Let $\lambda = (l_1, l_2) \in (\mathbb{Z}^2)_{\text{dom}}$ and ϱ be the representation of $\mathbf{GL}_2(\mathbb{C})$ on the space V_ϱ of homogeneous polynomials in X, Y of degree $l_1 - l_2$ defined by $\varrho(h)f(X, Y) = (\det h)^{l_2} f(aX + cY, bX + dY)$ for $h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{GL}_2(\mathbb{C})$ and $f \in V_\varrho$. As is well-known, any irreducible rational representation of $\mathbf{GL}_2(\mathbb{C})$, up to equivalence, is obtained this way. The space V_ϱ carries an $\mathbf{SU}(2)$ -invariant hermitian inner product $(\cdot | \cdot)_\varrho$ given by [9, (8.2.6)]. Recall $\mathbf{K}_\infty = \{k_\infty(u) := \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \mid u = A + iB \in \mathbf{U}(2)\}$. For $N \in \mathbb{Z}_{>0}$, let $S_\varrho(\mathbf{K}_0(N))$ denote the space of smooth functions $\varphi : \mathbf{G}(\mathbb{A}) \rightarrow V_\varrho$ that satisfy the following conditions:

- (i) $\varphi(z\gamma g) = \varphi(g)$ for $(z, \gamma, g) \in \mathbf{Z}(\mathbb{A}) \times \mathbf{G}(\mathbb{Q}) \times \mathbf{G}(\mathbb{A})$.
- (ii) $\varphi(gk_{\mathfrak{f}}k_\infty(u)) = \varrho(\bar{u})^{-1}\varphi(g)$ for $k_{\mathfrak{f}} \in \mathbf{K}_0(N)$ and $k_\infty(u) \in \mathbf{K}_\infty$.
- (iii) $R(X)\varphi = 0$ for all $X \in \mathfrak{p}^-$.
- (iv) φ is bounded on $\mathbf{G}(\mathbb{A})$.

For $T \in \mathbf{V}(\mathbb{R})$ with $\det(T) \neq 0$, define the function $\mathbf{B}_\varrho^T : \mathbf{G}(\mathbb{R})^0 \rightarrow \text{End}_{\mathbb{C}}(V_\varrho)$ by

$$(4-1) \quad \mathbf{B}_\varrho^T(g) := v(g)^{(l_1+l_2)/2} \varrho(Ci + D)^{-1} \exp(2\pi i \text{tr}(T g(i1_2))),$$

$$g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{G}(\mathbb{R})^0.$$

This function satisfies the conditions

$$(4-2) \quad \mathbf{B}_\varrho^T(m(A, c)g) = c^{(l_1+l_2)/2} \mathbf{B}_\varrho^T(g) \circ \varrho({}^t A^{-1} c)^{-1},$$

$$m(A, c) \in M_T^\circ(\mathbb{R}), g \in \mathbf{G}(\mathbb{R})^0,$$

$$(4-3) \quad \mathbf{B}_\varrho^T(n g k_\infty(u)) = \psi_T(n) \varrho(\bar{u})^{-1} \circ \mathbf{B}_\varrho^T(g), \quad (n, g, u) \in \mathbf{N}(\mathbb{R}) \times \mathbf{G}(\mathbb{R})^0 \times \mathbf{U}(2),$$

$$(4-4) \quad R(\mathfrak{p}^-) \mathbf{B}_\varrho^T = 0.$$

The function $B_{\varrho, v}^T : g \mapsto \mathbf{B}_\varrho^T(g)(v)$ with $v \in V_\varrho - (0)$ is bounded on $\mathbf{G}(\mathbb{R})^0$ if and only if $T \in \mathbf{V}(\mathbb{R})^+$, where $\mathbf{V}(\mathbb{R})^+$ denotes the set of positive definite elements in

$V(\mathbb{R})$. Set

$$\gamma := m(-1_2, -1) \in M(\mathbb{Q}).$$

Note that $\nu(\gamma) = -1$. Set $V(\mathbb{Q})^+ := V(\mathbb{Q}) \cap V(\mathbb{R})^+$.

Lemma 4.1. *Let $\varphi \in S_\varrho(\mathbf{K}_0(N))$. There is a unique family of vectors $a_\varphi(T; g_{\mathbf{f}}) \in V_\varrho$ ($T \in V(\mathbb{Q})$, $g_{\mathbf{f}} \in G(\mathbb{A}_{\mathbf{f}})$) such that*

$$(4-5) \quad \varphi(g_\infty g_{\mathbf{f}}) = \sum_{T \in V(\mathbb{Q})} \mathbf{B}_\varrho^T(g_\infty)(a_\varphi(T; g_{\mathbf{f}})), \quad g_\infty \in G(\mathbb{R})^\circ, g_{\mathbf{f}} \in G(\mathbb{A}_{\mathbf{f}}).$$

If $T \notin V(\mathbb{Q})^+$, then $a_\varphi(T; g_{\mathbf{f}}) = 0$ for all $g_{\mathbf{f}} \in G(\mathbb{A}_{\mathbf{f}})$. For $g_{\mathbf{f}} \in G(\mathbb{A}_{\mathbf{f}})$,

$$(4-6) \quad \int_{\mathbf{N}(\mathbb{A})/\mathbf{N}(\mathbb{Q})} \varphi(n g_\infty g_{\mathbf{f}}) \psi_T(n)^{-1} dn = \begin{cases} \mathbf{B}_\varrho^T(g_\infty)(a_\varphi(T; g_{\mathbf{f}})) & (g_\infty \in G(\mathbb{R})^\circ), \\ 0 & (g_\infty \in G(\mathbb{R}) - G(\mathbb{R})^\circ). \end{cases}$$

Proof. Fix $g_{\mathbf{f}} \in G(\mathbb{A}_{\mathbf{f}})$ and examine the integral, say $W(g_\infty)$, on the left side of (4-6). The function W on $G(\mathbb{R})^0$ satisfies conditions (4-3) and (4-4). Hence, there is a corresponding function $F : \mathfrak{h}_2 \rightarrow V_\varrho$ determined by the relation $F(Z) = \nu(g)^{-(l_1+l_2)/2} \varrho(Ci + D) W(g_\infty)$, with $g_\infty = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in G(\mathbb{R})^0$, and such that $Z = (Ai + B)(Ci + D)^{-1}$. Since $F(Z)$ satisfies $F(Z + X) = \psi_T(X) F(Z)$ ($X \in V(\mathbb{R})$), it is of the form $F(Z) = F_1(Y) \exp(2\pi i \operatorname{tr}(TX))$ with a V_ϱ -valued C^∞ -function $F_1(Y)$. Since $F(Z)$ is holomorphic, the Cauchy–Riemann equations yield $(d/dy_{ij}) F_1(Y) = -2\pi t_{ij} F_1(Y)$, which is uniquely solved as

$$F_1(Y) = a_\varphi(T; g_{\mathbf{f}}) \exp(-2\pi \operatorname{tr}(TY))$$

with a vector $a_\varphi(T; g_{\mathbf{f}}) \in V_\varrho$, or equivalently $W(g_\infty) = \mathbf{B}_\varrho^T(g_\infty)(a_\varphi(T; g_{\mathbf{f}}))$. Then, the formula in (4-5) is a consequence of the Fourier expansion of a $\mathbf{N}(\mathbb{Q})$ -periodic function on $\mathbf{N}(\mathbb{A})$. Since φ is bounded on $G(\mathbb{A})$, the function $W(g_\infty)$ should be bounded on $G(\mathbb{R})^0$; since $g_\infty \mapsto \mathbf{B}_\varrho^T(g_\infty)(v)$, with $v \in V_\varrho - \{0\}$, is unbounded if $T \notin V(\mathbb{R})^+$, we have $a_\varphi(T; g_{\mathbf{f}}) = 0$ for all $T \in V(\mathbb{Q}) - V(\mathbb{R})^+$. It remains to show the second case in (4-6). Let $g_\infty \notin G(\mathbb{R})^\circ$. Decompose $\gamma \in M(\mathbb{Q})$ as $\gamma_\infty \gamma_{\mathbf{f}}$ ($\gamma_\infty \in M(\mathbb{R})$, $\gamma_{\mathbf{f}} \in M(\mathbb{A}_{\mathbf{f}})$) in $M(\mathbb{A})$. By the left $G(\mathbb{Q})$ -invariance of φ ,

$$\varphi(n(X) g_\infty g_{\mathbf{f}}) = \varphi(\gamma n(X) g_\infty g_{\mathbf{f}}) = \varphi(n(-X) \gamma_\infty g_\infty \gamma_{\mathbf{f}} g_{\mathbf{f}}).$$

Integrate this in X over $V(\mathbb{Q}) \setminus V(\mathbb{A})$ and make a change of variables $n \rightarrow n^{-1}$ on the way; we have

$$\int_{\mathbf{N}(\mathbb{Q}) \setminus \mathbf{N}(\mathbb{A})} \varphi(n g_\infty g_{\mathbf{f}}) \psi_T(n)^{-1} dn = \int_{\mathbf{N}(\mathbb{Q}) \setminus \mathbf{N}(\mathbb{A})} \varphi(n \gamma_\infty g_\infty \gamma_{\mathbf{f}} g_{\mathbf{f}}) \psi_{-T}(n)^{-1} dn.$$

Since $\gamma_\infty g_\infty \in G(\mathbb{R})^\circ$, the last integral becomes $\mathbf{B}_\varrho^{-T}(\gamma_\infty g_\infty)(a_\varphi(-T; \gamma_{\mathbf{f}} g_{\mathbf{f}}))$ by the first case. Since $-T \notin V(\mathbb{R})^+$, we have $a(-T; \gamma_{\mathbf{f}} g_{\mathbf{f}}) = 0$. □

The vectors $a_\varphi(T; \mathfrak{g}_f)$ in V_ϱ are referred to as the adelic Fourier coefficients of φ . As is well-known, there is a linear bijective correspondence between $\varphi \in S_\varrho(\mathbf{K}_0(N))$ and the classical V_ϱ -valued Siegel cusp forms $f(Z)$ on

$$\Gamma_0(N) := \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{Sp}_2(\mathbb{Z}) \mid C \equiv 0 \pmod{N} \right\}$$

determined by the relation

$$f(g_\infty \langle i \ 1_2 \rangle) = \nu(g_\infty)^{-(l_1+l_2)/2} \varrho(Ci + D) \varphi(g_\infty), \quad g_\infty = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbf{G}(\mathbb{R})^0.$$

By the modular transformation law $f((AZ+B)(CZ+D)^{-1}) = \varrho(CZ+D)^{-1} f(Z)$ for $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \Gamma_0(N)$, since $-1_4 \in \Gamma_0(N)$, we have $f(Z) \equiv 0$; hence $S_\varrho(\mathbf{K}_0(N)) = (0)$, unless $l_1 \equiv l_2 \pmod{2}$. (For details, see [36, §3.2] and [4, §4].)

Since $\mathbf{N}(\mathbb{A}) \cap \mathbf{K}_0(N) = \{n(X) \mid X \in \mathbf{V}(\widehat{\mathbb{Z}})\}$, by (4-6), we have $a_\varphi(T; 1) = 0$ unless $T \in \mathcal{Q}$, where

$$\mathcal{Q} := \left\{ \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \mid a, b, c \in \mathbb{Z} \right\}$$

is the dual lattice of $\mathbf{V}(\mathbb{Z})$. Then, (4-5) with $\mathfrak{g}_f = 1_4$ reduces to

$$f(Z) = \sum_{T \in \mathcal{Q}^+} a_\varphi(T; 1_4) \exp(2\pi i \operatorname{tr}(TZ)), \quad Z \in \mathfrak{h}_2,$$

where $\mathcal{Q}^+ := \mathcal{Q} \cap \mathbf{V}(\mathbb{Q})^+$. This means that $a_\varphi(T) := a_\varphi(T; 1_4)$ ($T \in \mathcal{Q}^+$) is the Fourier coefficient of $f(Z)$ in the classical sense. Set

$$\delta := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in \mathbf{GL}_2(\mathbb{Q}).$$

The map $T \mapsto -Ts(\delta)$ preserves the set \mathcal{Q}^+ . If we identify \mathcal{Q}^+ with the set of positive definite integral binary forms $aX^2 + bXY + cY^2$, this operation corresponds to a sign change in b .

Lemma 4.2. *The set of adelic Fourier coefficients of $\varphi \in S_\varrho(\mathbf{K}_0(N))$ has the following properties:*

(4-7)

$$a_\varphi(T; \mathfrak{m}(h_f, \det h_f) \mathfrak{g}_f \kappa) = a_\varphi(Ts(h); \mathfrak{g}_f), \quad h \in \mathbf{GL}_2(\mathbb{Q}), \mathfrak{g}_f \in \mathbf{G}(\mathbb{A}_f), \kappa \in \mathbf{K}_0(N).$$

(4-8)
$$a_\varphi(-Ts(\delta); 1) = \varrho(\delta) a_\varphi(T; 1), \quad T \in \mathbf{V}(\mathbb{Q}).$$

Proof. The relation in (4-7) follows from (4-6) by a simple change of variables. Let us show (4-8). We have

(4-9)
$$\mathfrak{n}(-Xs(\delta)) = \kappa \mathfrak{n}(X) \kappa^{-1}, \quad X \in \mathbf{V}(\mathbb{A})$$

with $\kappa := \mathfrak{m}(\delta, 1) \in \mathbf{G}(\mathbb{Q})$. Write $\kappa = \kappa_f \kappa_\infty \in \mathbf{G}(\mathbb{A}_f) \mathbf{G}(\mathbb{R})$; then, $\kappa_f \in \mathbf{K}_0(N)$ and

$\kappa_\infty \in \mathbf{K}_\infty$. Using (4-9) and the left $G(\mathbb{Q})$ -invariance of φ , we have

$$\begin{aligned} \int_{\mathbf{N}(\mathbb{Q}) \backslash \mathbf{N}(\mathbb{A})} \varphi(n) \psi_{-T\mathfrak{s}(\delta)}(n)^{-1} \, dn &= \int_{\mathbf{V}(\mathbb{Q}) \backslash \mathbf{V}(\mathbb{A})} \varphi(n(-\mathfrak{s}(\delta)X)) \psi(\mathrm{tr}(TX))^{-1} \, dX \\ &= \int_{\mathbf{V}(\mathbb{Q}) \backslash \mathbf{V}(\mathbb{A})} \varphi(\kappa n(X)\kappa^{-1}) \psi(\mathrm{tr}(TX))^{-1} \, dX \\ &= \varrho(\delta) \int_{\mathbf{N}(\mathbb{Q}) \backslash \mathbf{N}(\mathbb{A})} \varphi(n) \psi_T(n)^{-1} \, dn. \end{aligned}$$

From this and (4-6),

$$\mathbf{B}_\varrho^{-T\mathfrak{s}(\delta)}(1_4)(a_\varphi(-T\mathfrak{s}(\delta); 1)) = \varrho(\delta) \mathbf{B}_\varrho^T(1_4)(a_\varphi(T; 1)).$$

By (4-1), this leads to $\mathbf{B}_\varrho^{-T\mathfrak{s}(\delta)}(1_4) = \mathbf{B}_\varrho^T(1_4) = e^{-2\pi \mathrm{tr}(T)} \mathrm{Id}_{V_\varrho}$. □

Define

$$(4-10) \quad \mathcal{Q}_{\mathrm{prim}}^+(D) := \left\{ T = \begin{bmatrix} a & 2^{-1}b \\ 2^{-1}b & c \end{bmatrix} \in \mathcal{Q} \mid a > 0, -\det T = D/4 \right\} :$$

D is a fundamental discriminant, so $T \in \mathcal{Q}_{\mathrm{prim}}^+(D)$ is primitive, i.e., $\mathrm{gcd}(a, b, c) = 1$; $\mathcal{Q}_{\mathrm{prim}}^+(D)$ is a subset of $\mathbf{V}(\mathbb{Q})$, which is stable under the action of the unimodular group $\mathbf{SL}_2(\mathbb{Z})$ induced by (2-1). The matrix T_θ defined by (2-11) belongs to $\mathcal{Q}_{\mathrm{prim}}^+(D)$. The $\mathbf{SL}_2(\mathbb{Z})$ -orbit of $T \in \mathcal{Q}_{\mathrm{prim}}^+(D)$ is denoted by $[T]$.

Lemma 4.3. *For each $u \in \mathbb{A}_E^\times$, let $I_\theta(u)$ decompose in $\mathbf{GL}_2(\mathbb{A})$ as*

$$(4-11) \quad I_\theta(u) = \gamma_u h_u \kappa_u, \quad \gamma_u \in \mathbf{GL}_2(\mathbb{Q}), h_u \in \mathbf{GL}_2(\mathbb{R})^\circ, \kappa_u = (\kappa_{u,p})_{p < \infty} \in \mathbf{GL}_2(\widehat{\mathbb{Z}})$$

and set $T_\theta(u) := T_\theta \mathfrak{s}(\gamma_u)$. Then:

- (i) $T_\theta(u) \in \mathcal{Q}_{\mathrm{prim}}^+(D)$.
- (ii) *The $\mathbf{SL}_2(\mathbb{Z})$ -equivalence class of $T_\theta(u)$ does not depend on the decomposition (4-11). If u, u' belongs to the same $E^\times E_\infty^\times \widehat{\theta}_E^\times$ -coset, then $T_\theta(u)$ and $T_\theta(u')$ are $\mathbf{SL}_2(\mathbb{Z})$ -equivalent.*
- (iii) *The map $[u] \mapsto [T_\theta(u)]$ from $\mathrm{Cl}(E) = \mathbb{A}_E^\times / E^\times E_\infty^\times \widehat{\theta}_E^\times$ to $\mathcal{Q}_{\mathrm{prim}}^+(D) / \mathbf{SL}_2(\mathbb{Z})$ is a bijection.*

Proof. This is an adelic reformulation of the classically well-known correspondence between the ideal classes and the $\mathbf{SL}_2(\mathbb{Z})$ -equivalence classes of integral binary quadratic forms. The proof is straightforward. □

Since $\mathfrak{m}(\mathbf{GL}_2(\widehat{\mathbb{Z}}), 1) \subset \mathbf{K}_0(N)$, the relation in (4-7) shows that the function $T \mapsto a_\varphi(T) := a_\varphi(T; 1_4)$ on $\mathbf{V}(\mathbb{Z})^+$ is $\mathbf{SL}_2(\mathbb{Z})$ -invariant. Hence the following sum of vectors in V_ϱ is well-defined:

$$(4-12) \quad \mathbf{R}(\varphi, E, \Lambda) := \sum_{[u] \in \mathrm{Cl}(E)} a_\varphi([T_\theta(u)])^{\mathrm{SO}} \Lambda(u)^{-1},$$

where $v^{SO} := \int_0^\pi \varrho\left(\begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}\right) v \, d\alpha$ for $v \in V_\varrho$. Recall the integral in (3-6) and the function in (4-1).

Lemma 4.4. *Let $\varphi \in S_\varrho(\mathbf{K}_0(N))$ and $\Lambda \in \widehat{\text{Cl}}(E)_{\text{prim}}$.*

(i) *For $g_\infty \in \mathbf{G}(\mathbb{R})^0$,*

$$B^{T_\theta, \Lambda}(\varphi; g_\infty) = w_D^{-1} \mathbf{B}_\varrho^{T_\theta}(g_\infty)(\mathbf{R}(\varphi, E, \Lambda))$$

where w_D is the number of units in \mathfrak{o}_E .

(ii) *Set $\sigma := \begin{bmatrix} 1 & \text{tr}_{E/\mathbb{Q}}(\theta) \\ 0 & -1 \end{bmatrix}$. Then*

$$(4-13) \quad \varrho(r)(B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta)) = B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta), \quad r \in \mathbf{SO}(2),$$

$$(4-14) \quad \varrho(\delta)(B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta)) = B^{-T_\theta \sigma(\theta), \Lambda}(\varphi; b_{\mathbb{R}}^\theta).$$

Proof. (i) The set $\mathbb{A}_E^\times / \mathbb{A}^\times E^\times E_{\mathbb{R}}^\times$ decomposes into a disjoint union of \mathfrak{o}_E^\times -orbits $[u]\mathfrak{o}_E^\times$ for different classes $[u] \in \text{Cl}(E)$. Let μ denote the quotient measure on $\mathbb{A}_E^\times / \mathbb{A}^\times E^\times E_{\mathbb{R}}^\times$. For $\tau \in \mathbb{A}_E^\times$, set $\mathfrak{m}_\theta(\tau) := \mathfrak{m}(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau))$. Then, by (3-6) and (4-6), $B^{T_\theta, \Lambda}(\varphi; g_\infty)$ equals

$$\begin{aligned} \mathbf{B}_\varrho^{T_\theta}(g_\infty) & \left(\int_{\mathbb{A}_E^\times / \mathbb{A}^\times E^\times E_{\mathbb{R}}^\times} \int_{E_\infty^\times / \mathbb{R}^\times} N_{E_\infty/\mathbb{R}}(\tau_\infty)^{(l_1+l_2)/2} \varrho(t I_\theta(\tau_\infty)^\dagger)^{-1} a_\varphi(T_\theta; \mathfrak{m}_\theta(u)) \right. \\ & \quad \left. \times \Lambda(u)^{-1} d^\times \tau_\infty d\mu(u) \right) \\ & = \mathbf{B}_\varrho^{T_\theta}(g_\infty) \left(\sum_{[u] \in \text{Cl}(E)} \int_{[u]\mathfrak{o}_E^\times} a_\varphi(T_\theta; \mathfrak{m}_\theta(u))^{\text{SO}} \Lambda(u)^{-1} d\mu(u) \right), \end{aligned}$$

The quotient measure $\eta_{E_\infty^\times/\mathbb{R}^\times}$ is identified with $d\alpha$ under the identification

$$E_\infty^\times / \mathbb{R}^\times = \left\{ \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \mid \alpha \in [0, \pi) \right\}.$$

Since $\mathfrak{m}_\theta(\widehat{\mathcal{O}}_E^\times) \subset \mathbf{K}_0(N)$, by (4-6), the formula in parenthesis becomes

$$\sum_{[u] \in \text{Cl}(E)} \mu([u]\mathfrak{o}_E^\times) a_\varphi(T_\theta; \mathfrak{m}_\theta(u)) \Lambda(u)^{-1}.$$

Let $I_\theta(u)$ decompose as in (4-11). Then, by (4-7) and Lemma 4.3, we have

$$a_\varphi(T_\theta; \mathfrak{m}_\theta(u)) = a_\varphi(T_\theta \mathfrak{s}(\gamma_u); 1) = a_\varphi(T_\theta(u)).$$

It remains to compute $\mu([u]\mathfrak{o}_E^\times)$. Since any fiber of the natural surjection from \mathfrak{o}_E^\times onto $[u]\mathfrak{o}_E^\times$ has a simply transitive action of the group \mathcal{O}_E^\times , we have $\mu([u]\mathfrak{o}_E^\times) = 1/\#\mathfrak{o}_E^\times = 1/w_D$.

(ii) We have $I_\theta(e^{i\alpha}) = A_\theta^{-1} \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} A_\theta$, or equivalently

$$\mathfrak{m}(I_\theta(e^{i\alpha}), 1) = b_{\mathbb{R}}^\theta \text{diag}\left(\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}\right) (b_{\mathbb{R}}^\theta)^{-1}$$

for $\alpha \in \mathbb{R}$. Due to this and $\Lambda_\infty = 1$, formula (4-13) is proved by a change of variables $\tau \mapsto \tau u^{-1}$ in the integral (3-6). Since $b_{\mathbb{R}}^\theta m(\delta, 1)(b_{\mathbb{R}}^\theta)^{-1} = m(\sigma, 1)$, formula (4-14) is proved in the same way as in the proof of Lemma 4.2. \square

Let $\varphi \in S_\varrho(\mathbf{K}_0(N)) - (0)$; then $l_1 \equiv l_2 \pmod{2}$ as noted above. Therefore, $\dim_{\mathbb{C}}(V_\varrho^{\mathrm{SO}(2)}) = 1$; indeed, $V_\varrho^{\mathrm{SO}(2)} = \mathbb{C}v_\varrho^0$ with $v_\varrho^0 := (X^2 + Y^2)^{(l_1-l_2)/2}$. Set

$$v_\varrho^\theta := \left(\frac{-2}{\sqrt{|D|}}\right)^{\frac{l_1+l_2}{2}} \varrho(t A_\theta) v_\varrho^0 = \left(\frac{-2}{\sqrt{|D|}}\right)^{\frac{l_1-l_2}{2}} (X^2 + \mathrm{tr}_{E/\mathbb{Q}}(\theta)XY + N_{E/\mathbb{Q}}(\theta)Y^2)^{\frac{l_1-l_2}{2}}.$$

Note that $\mathbf{B}_\varrho^{T_\theta}(b_{\mathbb{R}}^\theta) = \left(\frac{-2}{\sqrt{|D|}}\right)^{-(l_1+l_2)/2} e^{-2\pi\sqrt{|D|}} \varrho(t A_\theta)^{-1}$, so that $\mathbf{B}_\varrho^{T_\theta}(b_{\mathbb{R}}^\theta)(v_\varrho^\theta) = e^{-2\pi\sqrt{|D|}} v_\varrho^0$. Thus, as a corollary to Lemma 4.4, the vector $\mathbf{R}(\varphi, E, \Lambda) \in V_\varrho$ is nonzero if and only if $B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta) \neq 0$, in which case there exists a unique scalar $R(\varphi, E, \Lambda)$ such that $\mathbf{R}(\varphi, E, \Lambda) = R(\varphi, E, \Lambda)v_\varrho^\theta$, or equivalently

$$(4-15) \quad B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta) = \pi w_D^{-1} e^{-2\pi\sqrt{|D|}} R(\varphi, E, \Lambda)v_\varrho^0.$$

Note that $(v_\varrho^0)^{\mathrm{SO}} = \pi v_\varrho^0$. Define a C^∞ -function $B_\varrho^{T_\theta} : \mathrm{G}(\mathbb{R})^0 \rightarrow V_\varrho$ as

$$(4-16) \quad B_\varrho^{T_\theta}(g_\infty) := \mathbf{B}_\varrho^{T_\theta}(g_\infty)(v_\varrho^\theta), \quad g_\infty \in \mathrm{G}(\mathbb{R})^0.$$

4.1. Sign condition. The Galois group $\mathrm{Gal}(E/\mathbb{Q})$ acts on $\mathrm{Cl}(E)$ naturally, hence on the orbit space $\mathcal{Q}_{\mathrm{prim}}^+(D)/\mathbf{SL}_2(\mathbb{Z})$ through the bijection in Lemma 4.3(iii). The following lemma describes the conjugate action on $\mathcal{Q}_{\mathrm{prim}}^+(D)/\mathbf{SL}_2(\mathbb{Z})$ explicitly. Recall the element $\sigma \in \mathbf{GL}_2(\mathbb{Z})$ defined in Lemma 4.4(ii).

Lemma 4.5. *For $u \in \mathrm{Cl}(E)$, the element $T_\theta(\bar{u})$ is $\mathbf{SL}_2(\mathbb{Z})$ -equivalent to $-T_\theta(u)s(\delta)$. In other words, the conjugate action on $\mathcal{Q}_{\mathrm{prim}}^+(D)/\mathbf{SL}_2(\mathbb{Z})$ is induced by the map $T \mapsto -Ts(\delta)$.*

Proof. The defining formula of I_θ in (2-5) yields

$$I_\theta(a + \theta b) = \begin{bmatrix} 1 & \theta \\ 1 & \bar{\theta} \end{bmatrix}^{-1} \begin{bmatrix} a+\theta b & 0 \\ 0 & a+\bar{\theta}b \end{bmatrix} \begin{bmatrix} 1 & \theta \\ 1 & \bar{\theta} \end{bmatrix}, \quad a, b \in \mathbb{Q}.$$

Set $t := \mathrm{tr}_{E/\mathbb{Q}}(\theta)$. Then, since $\bar{\theta} = t - \theta$, we have $a + b\bar{\theta} = a' + b'\theta$ with $a' = a + tb \in \mathbb{Q}$ and $b' = -b \in \mathbb{Q}$. Then, a computation reveals

$$(4-17) \quad I_\theta(a + b\bar{\theta}) = \sigma I_\theta(a + b\theta)\sigma^{-1}.$$

By the decomposition of $I_\theta(u)$ for $u \in \mathbb{A}_E^\times$ in (4-11), we may take $\gamma_{\bar{u}} = \sigma\gamma_u\sigma^{-1}$. Thus, $T_\theta(\bar{u}) = T_\theta s(\gamma_{\bar{u}}) = T_\theta s(\sigma\gamma_u\sigma^{-1})$. A computation shows $T_\theta s(\sigma) = -T_\theta$. Hence,

$$T_\theta(\bar{u}) = -T_\theta s(\gamma_u\sigma^{-1}) = -T_\theta(u)s(\sigma^{-1}) = -T_\theta(u)s(\delta)s\left(\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\right).$$

The last matrix is $\mathbf{SL}_2(\mathbb{Z})$ -equivalent to $-T_\theta(u)s(\delta)$. This completes the proof. \square

For $\Lambda \in \widehat{\mathrm{Cl}(E)}$, recall from (2-24) its conjugate $\Lambda^\dagger \in \widehat{\mathrm{Cl}(E)}$.

Lemma 4.6. *Let $\varphi \in S_\varrho(\mathbf{K}_0(N))$ and $\Lambda \in \widehat{\text{Cl}}(E)$ and suppose $B^{T_\theta, \Lambda}(\varphi; b_{\mathbb{R}}^\theta) \neq 0$. Then,*

$$R(\varphi, E, \Lambda^\dagger) = (-1)^{l_2} R(\varphi, E, \Lambda).$$

Proof. This is proved by Lemma 4.5 and (4-15) and by $\varrho(\delta)v_\varrho^0 = (-1)^{l_2}v_\varrho^0$. Note that $\delta^{-1}\sigma = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \in \mathbf{SL}_2(\mathbb{Z})$, so that $a_\varphi(-T_\theta s(\sigma), 1) = a_\varphi(-T_\theta s(\delta), 1)$. \square

4.2. Automorphic representations. Let $\Pi_{\text{cusp}}(\mathbb{Z} \backslash \mathbb{G})$ denote the set of all those irreducible cuspidal representations of $\mathbb{G}(\mathbb{A})$ with trivial central characters, i.e., those irreducible $(\mathfrak{g}, \mathbf{K}_\infty) \times \mathbb{G}(\mathbb{A}_f)$ -submodules (π, V_π) of the space of cusp forms on $\mathbb{Z}(\mathbb{A})\mathbb{G}(\mathbb{Q}) \backslash \mathbb{G}(\mathbb{A})$. We endow V_π with the restriction of the L^2 inner product

$$(\varphi \mid \varphi_1)_{L^2} := \int_{\mathbb{Z}(\mathbb{A})\mathbb{G}(\mathbb{Q}) \backslash \mathbb{G}(\mathbb{A})} \varphi(g) \overline{\varphi_1(g)} dg, \quad \varphi, \varphi_1 \in V_\pi.$$

For $\pi \in \Pi_{\text{cusp}}(\mathbb{Z} \backslash \mathbb{G})$, we fix its restricted tensor decomposition $\pi \cong \bigotimes_{p \leq \infty} \pi_p$ with (π_p, V_{π_p}) being an irreducible admissible unitarizable representation of $\mathbb{G}(\mathbb{Q}_p)$ with trivial central character such that a \mathbf{K}_p -invariant vector $\xi_p^0 \in V_{\pi_p} - (0)$ is preassigned for almost all $p < \infty$. An element $\lambda \in (\mathbb{Z}^2)_{\text{dom}}$ is the highest weight of the minimal \mathbf{K}_∞ -type of a holomorphic discrete series representation (HDS for short) of $\mathbf{Sp}_2(\mathbb{R})$ if and only if $l_2 > 2$, in which case the Harish-Chandra parameter of the HDS is $(l_1 - 1, l_2 - 2)$. For such λ and $N \in \mathbb{Z}_{>0}$, let $\Pi_{\text{cusp}}(\lambda, N)$ be a subset of $\pi \in \Pi_{\text{cusp}}(\mathbb{Z} \backslash \mathbb{G})$ having the following properties:

- (i) As a $(\mathfrak{g}, \mathbf{K}_\infty)$ -module $\pi_\infty \cong D_{l_1-1, l_2-2} \oplus D_{-l_2+2, -l_1+1}$, where D_{m_1, m_2} is the discrete series representation of $\mathbb{G}(\mathbb{R})^0$ of Harish-Chandra parameter (m_1, m_2) with central character $\mathbb{Z}(\mathbb{R}) \cong \mathbb{R}^\times \ni z \mapsto \text{sgn}(z)^{m_1+m_2+1}$.
- (ii) $V_\pi^{\mathbf{K}_0(N)} \neq (0)$.
- (iii) Let (ϱ, V_ϱ) be the irreducible rational representation of highest weight λ as before. The space $S_\varrho(\mathbf{K}_0(N))$ is an orthogonal direct sum of spaces $V_\pi^{\mathbf{K}_0(N)}[\varrho] := \{\varphi \in S_\varrho(\mathbf{K}_0(N)) \mid \varphi_v \in V_\varrho (\forall v \in V_\varrho)\}$ for $\pi \in \Pi_{\text{cusp}}(\lambda, N)$, where $\varphi_v(g) := (\varphi(g) \mid v)_\varrho$ for $\varphi \in S_\varrho(\mathbf{K}_0(N))$ and $v \in V_\varrho$.

Although there may be many choices of $\Pi_{\text{cusp}}(\lambda, N)$, we fix one of them once and for all. For $\pi \in \Pi_{\text{cusp}}(\lambda, N)$, we fix a \mathbf{K}_∞ -intertwining map $V_\varrho \oplus \overline{V}_\varrho \hookrightarrow D_{l_1-1, l_2-2} \oplus D_{-l_2+2, -l_1+1} (\cong \pi_\infty)$ once and for all, where $\overline{V}_\varrho := \overline{\mathbb{C}} \otimes_{\mathbb{C}} V_\varrho$ is the complex conjugate of V_ϱ .

4.3. Basic assumptions. From now on, we fix a triple $(\lambda = (l_1, l_2), N, M) \in (\mathbb{Z}^2)_{\text{dom}} \times \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ satisfying the following conditions:

- (A-i) N is square-free.
- (A-ii) $l_1 \equiv l_2 \pmod{2}$, and $l_2 \in \mathbb{Z}_{\geq 3}$ so that $(l_1 - 1, l_2 - 2)$ is the Harish-Chandra parameter of an HDS.

(A-iii) N and M are coprime.

(A-iv) M is odd.

(A-v) All prime divisors of NM are inert in E/\mathbb{Q} .

We fix characters $\Lambda \in \widehat{\mathrm{Cl}}(E)$ and $\mu = \bigotimes_{p \leq \infty} \mu_p \in \widehat{\mathbb{A}^\times/\mathbb{Q}^\times \mathbb{R}_{>0}}$ with $M := \mathrm{cond}(\mu)$ as well. As usual, the character μ induces a primitive Dirichlet character $\tilde{\mu} : (\mathbb{Z}/M\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$ in such a way that $\mu(u) = \tilde{\mu}(a)^{-1}$ for all $u \in \widehat{\mathbb{Z}}^\times$ and $a \in \mathbb{Z}$ with $u - a \in M\widehat{\mathbb{Z}}$; thus $\mu_p(p) = \tilde{\mu}(p)$ for primes $p \nmid M$, and

$$\prod_{p|M} W_{\mathbb{Q}_p}(\mu_p, \psi_p) = M^{-1/2} G(\tilde{\mu}),$$

with $W_{\mathbb{Q}_p}(\mu_p, \psi_p)$ as in Definition 6.2 and $G(\tilde{\mu}) := \sum_{a \in (\mathbb{Z}/M\mathbb{Z})^\times} \tilde{\mu}(a) e^{2\pi i a/M}$, the Gauss sum of $\tilde{\mu}$. Let $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ denote the set of all those $\pi \in \Pi_{\mathrm{cusp}}(\lambda, N)$ that satisfies the condition

(A-vi) π admits a global (T_θ, Λ) -Bessel model, i.e., there exists $\varphi \in V_\pi$ such that $B^{T_\theta, \Lambda}(\varphi; g) \neq 0$ for some $g \in \mathrm{G}(\mathbb{A})$.

We remark that the conditions listed above (when $\mu = \mathbf{1}$ and $l_1 = l_2$) are also imposed in the most part of [7]. Let $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\mathrm{SK}}$ denote the set of $\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ that is the Saito–Kurokawa lift [28] of a cuspidal automorphic representation of $\mathbf{PGL}_2(\mathbb{A})$, which is locally described by [37] (see also [38]). From [37, §4], the set $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\mathrm{SK}} = \emptyset$ unless $l_1 = l_2$, i.e., V_ϱ is one dimensional. We note that all the representations $\pi \cong \bigotimes_p \pi_p$ in $\Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N) \setminus \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{SK}}$ are non-CAP (cf. [41, 17, §3.5, 30, Corollary 4.5]); then, by [48], π_p is tempered for all $p \nmid N$. By invoking [3] (see also [39]), any $\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N) \setminus \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{SK}}$ is either a Yoshida lift, i.e., there exists a pair of irreducible cuspidal automorphic representations (σ_1, σ_2) of $\mathbf{GL}_2(\mathbb{A})$ such that $L(s, \pi) = L(s, \sigma_1)L(s, \sigma_2)$, or a “general type”, that is, there exists an irreducible cuspidal automorphic representation Π of $\mathbf{GL}_4(\mathbb{A})$ such that $L(s, \Pi, \wedge^2)$ has a pole at $s = 1$ and $L(s, \pi) = L(s, \Pi)$. Let $\Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{Y}}$ (resp. $\Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{G}}$) be the set of $\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)$ that is a Yoshida lift (resp. a general type). Then,

$$(4-18) \quad \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N) = \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{G}} \cup \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{Y}} \cup \Pi_{\mathrm{cusp}}^{(T_\theta, \chi)}(l, N)^{\mathrm{SK}}$$

(disjoint union).

On the other hand, the set $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ can be separated into the subsets consisting of all newforms (i.e., $N_\pi = N$) and all oldforms (i.e., $N_\pi \neq N$) denoted by $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\mathrm{new}}$ and $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\mathrm{old}}$ respectively. For $\bullet \in \{\mathrm{G}, \mathrm{Y}, \mathrm{SK}\}$ and $\ast \in \{\mathrm{new}, \mathrm{old}\}$, we set

$$\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\bullet, \ast} = \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\bullet} \cap \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\ast}.$$

Let N_π be the product of primes $p < \infty$ such that π_p is not spherical. By condition (ii) in Section 4.2, we have $N_\pi \mid N$. Due to (A-i), the representations π_p for $p \mid N_\pi$ have nonzero $\mathbf{K}_0(p\mathbb{Z}_p)$ -fixed vectors, which, in turn, implies that π_p is Iwahori spherical so that their isomorphism classes are listed in [36, Table 3]. (For extended tables, we refer to [34, Appendix].) Note that in [36] a different symplectic form is used to define the symplectic group; up to the adjustment for this difference, the group P_1 in [36] is our $\mathbf{K}_0(p\mathbb{Z}_p)$. In [34], all irreducible admissible representations of $G(\mathbb{Q}_p)$ that admit local Bessel models are classified, and the result is conveniently summarized in [32, Table 2]. Here is a summary of what is available for our π (as in Section 4.3):

- π_∞ as a representation of $G(\mathbb{R})^0$ is a direct sum of an HDS and its complex conjugate; thus, π is CAP if and only if it is a Saito–Kurokawa lift from a cuspidal representation of $\mathbf{PGL}_2(\mathbb{A})$, which happens only if $l_1 = l_2$.
- Suppose $p \nmid N_\pi$. Then, the local representation π_p is of type I and tempered if π is non-CAP, and is of type IIb when π is CAP.
- Suppose $p \mid N_\pi$. Then, the local representation π_p is either of type IIIa, in which case $\dim_{\mathbb{C}} V_{\pi_p}^{\mathbf{K}_0(p\mathbb{Z}_p)} = 2$, or of type VIb, in which case $\dim_{\mathbb{C}} V_{\pi_p}^{\mathbf{K}_0(p\mathbb{Z}_p)} = 1$; when π is CAP, π_p has to be of type VIb.

4.4. Bessel models. Let $p < \infty$. For any irreducible admissible representation (π_p, V_{π_p}) of $G(\mathbb{Q}_p)$, let $(V_{\pi_p}^*)^{T_\theta, \Lambda_p}$ denote the space of all \mathbb{C} -linear forms $\ell : V_{\pi_p} \rightarrow \mathbb{C}$ that satisfy

$$\ell(\pi_p(m(I_\theta(\tau), N_{E/\mathbb{Q}}(\tau))n)\xi) = \Lambda_p(\tau)\psi_{T_\theta}(n)\ell(\xi), \quad \xi \in V_{\pi_p}, \tau \in E_p^\times, n \in N(\mathbb{Q}_p).$$

It is known that $\dim_{\mathbb{C}}(V_{\pi_p}^*)^{T_\theta, \Lambda_p} \leq 1$ [26; 33]. We say that π_p has a local (T_θ, Λ_p) -Bessel model if $(V_{\pi_p}^*)^{T_\theta, \Lambda_p} \neq (0)$; when this is the case, the space of functions of the form $g \mapsto \ell(\pi_p(g)\xi)$ with $\xi \in V_{\pi_p}$ is independent of $\ell \in (V_{\pi_p}^*)^{T_\theta, \Lambda_p} - (0)$; this space is denoted by $\mathcal{B}(T_\theta, \Lambda_p)[\pi_p]$ and is called the local (T_θ, Λ_p) -Bessel model of π_p .

When π_p is spherical, it is known that π_p has a local (T_θ, Λ_p) -Bessel model, and the space $\mathcal{B}(T_\theta, \Lambda_p)[\pi_p]$ contains a unique \mathbf{K}_p -invariant function B_{π_p} such that $B_{\pi_p}^0(1_4) = 1$ [42, Theorem 2-I; 6]. Let ξ_p^0 be the nonzero \mathbf{K}_p -fixed vector in V_{π_p} ; then there exists a unique element $\ell_{\pi_p}^0 \in (V_{\pi_p}^*)^{T_\theta, \Lambda_p} - (0)$ such that $\ell_{\pi_p}^0(\pi_p(g)\xi_p^0) = B_{\pi_p}^0(g)$ for all $g \in G(\mathbb{Q}_p)$. The pair $(\ell_{\pi_p}^0, \xi_p^0)$ is referred to as the unramified (T_θ, Λ_p) -Bessel datum for π_p .

Let $\pi \cong \bigotimes_{p \leq \infty} \pi_p \in \Pi_{\text{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ (see Section 4.3).

Definition 4.7. A system $\{(\ell_p, \xi_p)\}_{p < \infty}$ with $\ell_p \in (V_{\pi_p}^*)^{T_\theta, \Lambda_p} - (0)$ and $\xi_p \in V_{\pi_p}^{\mathbf{K}_0(N_\pi\mathbb{Z}_p)}$, is called a (T_θ, Λ) -Bessel data for π if $(\ell_p, \xi_p) = (\ell_{\pi_p}^0, \xi_{\pi_p}^0)$ for $p \nmid N_\pi$ and $\ell_p(\xi_p) = 1$ for all $p < \infty$.

By [32, Theorem 2.8.2 and 9.3], a (T_θ, Λ) -Bessel data exists for our π . Once a (T_θ, Λ) -Bessel data $\{(\ell_p, \xi_p)\}_{p < \infty}$ is fixed, one can define $\varphi_{\pi, \varrho}^0 \in V_\pi^{\mathbf{K}_0(N_\pi)}[\varrho]$ to be the V_ϱ -valued cusp form such that for any $v \in V_\varrho$, the function $(\varphi_{\pi, \varrho}^0(g) \mid v)_\varrho$ in $V_\pi \cong \bigotimes_{p \leq \infty} V_{\pi_p}$ corresponds to the pure tensor $\bar{v} \otimes (\bigotimes_{p < \infty} \xi_p)$, where $\bar{v} = 1 \otimes v \in \bar{V}_\varrho \hookrightarrow V_{\pi_\infty}$. A particular choice of $\{(\ell_p, \xi_p)\}$ will be made in Section 6.8 so that $\varphi_{\pi, \varrho}^0$ corresponds to a newform on the arithmetic quotient $\Gamma_0(N) \backslash \mathfrak{h}_2$ in the sense of [7, §3.2].

Lemma 4.8. *Let $\{(\ell_p, \xi_p)\}_{p < \infty}$ be a (T_θ, Λ) -Bessel data for $\pi \in \Pi_{\mathrm{cusp}}(\lambda, N)$. Let $\phi = \bigotimes_p \phi_p \in \mathcal{S}(\mathbb{A}_E^2)$ be a decomposable element. Then, for any $\varphi \in V_\pi^{\mathbf{K}_0(N)}[\varrho]$ such that, for any $v \in V_\varrho$, φ_v corresponds to the pure tensor $v \otimes (\bigotimes_{p < \infty} v_p) \in \bigotimes_{p \leq \infty} V_{\pi_p}$, we have*

$$(4-19) \quad B^{T_\theta, \Lambda}(\varphi; g) = B^{T_\theta, \Lambda}(\varphi_{\pi, \varrho}^0; g_\infty) \prod_{p < \infty} \ell_p(\pi_p(g_p)v_p), \quad g = (g_p)_p \in G(\mathbb{A}).$$

For $\mathrm{Re}(s) > 1$, $v \in V_\varrho$ and $b_\mathfrak{f} = (b_p)_{p < \infty} \in G(\mathbb{A}_\mathfrak{f})$,

$$(4-20) \quad \langle E(\phi, s, \Lambda, \mu), R(b_\mathfrak{f} b_\mathbb{R}^\theta \bar{\varphi}_v) \rangle = \frac{\sqrt{|D|}}{2} Z_v^{(\infty)}(\phi_\infty, \varphi_\pi^0; s, \mu_\infty, \Lambda) \prod_{p \leq \infty} Z_p(\phi_p, \bar{B}_{v_p}; s, \mu_p; b_p),$$

where

$$\begin{aligned} Z_v^{(\infty)}(\phi_\infty, \varphi_\pi^0; s, \mu_\infty, \Lambda) &:= \int_{\mathbf{K}_\infty^\#} \int_{\mathbb{R}^\times} f_{\phi_\infty}^{(s, 1, \mu_\infty)}(k_\infty^\#) \mu_\infty(a) |a|_{\mathbb{R}}^{s-1} \\ &\quad \times (v \mid B^{T_\theta, \Lambda}(\varphi_{\pi, \varrho}^0; m(a)1_2, a)\iota_\theta(k_\infty^\#) b_\mathbb{R}^\theta)_\varrho \, d^\times a \, dk_\infty^\#, \\ Z_p(\phi_p, \bar{B}_{v_p}; s, \mu_p; b_p) &:= \int_{\mathbf{K}_p^\#} \int_{\mathbb{Q}_p^\times} f_{\phi_p}^{(s, \Lambda_p, \mu_p)}(k_p^\#) \mu_p(a) |a|_p^{s-1} \\ &\quad \times \bar{B}_{v_p}(m(a_p)1_2, a_p)\iota_\theta(k_p^\#) b_p \, d^\times a_p \, dk_p^\#, \end{aligned}$$

with $B_{v_p}(g_p) := \ell_p(\pi_p(g_p)v_p)$ for $g_p \in G(\mathbb{Q}_p)$.

Proof. Fix $g_\infty \in G(\mathbb{R})^0$ and $v \in V_\varrho$ and regard $\varphi \mapsto B^{T_\theta, \Lambda}(\varphi_v; g_\infty)$ as a linear functional on $\bigotimes_{p < \infty} V_{\pi_p}$ by the natural inclusion

$$\bigotimes_{p < \infty} V_{\pi_p} \hookrightarrow v \otimes \left(\bigotimes_{p < \infty} V_{\pi_p} \right) \hookrightarrow V_\pi.$$

Then, by the local multiplicity-one theorem for Bessel functionals on $G(\mathbb{Q}_p)$ recalled above, there exists $C_v(g_\infty) \in \mathbb{C}$ such that

$$(4-21) \quad (v \mid B^{T_\theta, \Lambda}(\varphi; g_\infty))_\varrho = C_v(g_\infty) \prod_{p < \infty} \overline{\ell_p(v_p)}$$

for φ corresponding to $\bigotimes_{p < \infty} v_p$. To determine $C_v(g_\infty)$, set $v_p = \xi_p$ for all $p < \infty$; then by $\ell_p(\xi_p) = 1$, we get $(v \mid B^{T_\theta, \Lambda}(\overline{\varphi_{\pi, \varrho}^0}; g_\infty))_\varrho = C_v(g_\infty)$. Now we apply (3-7) with T_θ, Λ and φ replaced by $-T_\theta, \Lambda^{-1}$ and $R(b_\mathfrak{f})\varphi$. □

4.5. Gamma factor and global Bessel period. Recall the point $b_{\mathbb{R}}^{\theta}$ in (2-14) and the vector $v_{\varrho}^{\theta} \in V_{\varrho}$ in (4-16). Since $v(b_{\mathbb{R}}^{\theta}) > 0$, the point $m(a1_2, a)\iota_{\theta}(k_{\infty}^{\#})b_{\mathbb{R}}^{\theta}$ with $a \in \mathbb{R}^{\times}$ and $k_{\infty}^{\#} \in \mathbf{K}_{\infty}^{\#}$ belongs to $G(\mathbb{R})^0$ if and only if $\text{sgn}(a) \text{sgn}(k_{\infty}^{\#}) = +1$, where $\text{sgn}(k_{\infty}^{\#}) := \text{sgn}(v(\iota_{\theta}(k_{\infty}^{\#})))$. Thus, by (3-6) and (4-6), we have

$$B^{T_{\theta}, \Lambda}(\varphi_{\pi}^0; m(a1_2, a)\iota_{\theta}(k_{\infty}^{\#})b_{\mathbb{R}}^{\theta}) = 0$$

unless $\text{sgn}(a) \text{sgn}(k_{\infty}^{\#}) = +1$. By Lemma 4.4 and (4-16),

$$B^{T_{\theta}, \Lambda}(\varphi_{\pi}^0; g_{\infty}) = \pi w_D^{-1} R(\varphi_{\pi}^0, E, \Lambda) B_{\varrho}^{T_{\theta}}(g_{\infty}), \quad g_{\infty} \in G(\mathbb{R})^0.$$

Substituting this, we have that $Z_v^{(\infty)}(\phi_{\infty}, \varphi_{\pi}^0; s, \Lambda)$ equals

$$(4-22) \quad 2\pi w_D^{-1} \overline{R(\varphi_{\pi}^0, E, \Lambda)} \int_{(\mathbf{K}_{\infty}^{\#})^0} \int_0^{\infty} |a|_{\mathbb{R}}^{s-1} f_{\phi_{\infty}}^{(s, 1, \mu_{\infty})}(k_{\infty}^{\#})(v | B_{\varrho}^{T_{\theta}}(m(a1_2, a)\iota_{\theta}(k_{\infty}^{\#})b_{\mathbb{R}}^{\theta}))_{\varrho} d^{\times} a dk_{\infty}^{\#}.$$

Now we specify ϕ_{∞} . Recall that the highest weight of ϱ is $\lambda = (l_1, l_2)$ and that $l_1 - l_2 \in 2\mathbb{Z}_{>0}$. Set $d := l_1 - l_2$ and define

$$f_{\varrho}(u) := (d + 1) \frac{(\varrho(C\bar{u}C^{-1})v_{\varrho}^0 | v_{\varrho}^0)_{\varrho}}{(v_{\varrho}^0 | v_{\varrho}^0)_{\varrho}}, \quad u \in (\mathbf{K}_{\infty}^{\#})^0,$$

with $C := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$. If $u = \begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix} \in (\mathbf{K}_{\infty}^{\#})^0 = \mathbf{SU}(2)$ with $a = a' + ia''$, $b = b' + ib''$ and $A, B \in \mathbf{Mat}_2(\mathbb{R})$ defined as in (2-16), so that $(b_{\mathbb{R}}^{\theta})^{-1} \iota_{\theta}(u) b_{\mathbb{R}}^{\theta} = k_{\infty}(A + iB)$ (Lemma 2.2), then a computation reveals $C\bar{u}C^{-1} = A - iB$. Thus, the automorphism $u \mapsto C\bar{u}C^{-1}$ of $\mathbf{SU}(2)$ brings the subgroup

$$\mathbf{B}^{\#}(\mathbb{R}) \cap (\mathbf{K}_{\infty}^{\#})^0 = \left\{ \begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} \mid a \in \mathbb{C}^{\times} \right\}$$

to the subgroup $\mathbf{SO}(2)$, which fixes v_{ϱ}^0 . Hence, f_{ϱ} is left- $\mathbf{B}^{\#}(\mathbb{R}) \cap (\mathbf{K}_{\infty}^{\#})^0$ -invariant. Thus, since $G^{\#}(\mathbb{R}) = \mathbf{B}^{\#}(\mathbb{R})(\mathbf{K}_{\infty}^{\#})^0$, there exists a unique element $f_{\varrho}^{(s)}$ of $\mathcal{V}^{\#}(s, 1, \mu_{\infty})$ such that $f_{\varrho}^{(s)} | \mathbf{K}_{\infty}^{\#} = f_{\varrho}$. Define $\phi_{\infty} \in \mathcal{S}(\mathbb{C}^2)$ by

$$(4-23) \quad \phi_{\infty}(x, y) := (d + 1) |D|^{-d/4} \sum_{j=0}^{d/2} \binom{d/2}{j}^2 (x\bar{x})^j (-y\bar{y})^{d/2-j} \frac{2}{\pi} \exp\left(-\frac{2\pi}{\sqrt{|D|}}(x\bar{x} + y\bar{y})\right).$$

Then, noting that $\varrho(C)^{-1}v_{\varrho}^0 = (XY)^{d/2}$ and using formulas (8.2.4), (8.2.5) and (8.2.6) of [9], we easily confirm the relation

$$\phi_{\infty}([0, \bar{\tau}]u) = f_{\varrho}(u) |D|^{-d/4} |\tau|_{\mathbb{C}}^{d/2} \exp\left(-\frac{2\pi}{\sqrt{|D|}}|\tau|_{\mathbb{C}}\right), \quad \tau \in \mathbb{C}^{\times}, u \in \mathbf{SU}(2).$$

Then, by computing the integral in (3-1) for $g_{\infty} \in \mathbf{SU}(2)$, we get the relation

$$(4-24) \quad f_{\phi_{\infty}}^{(s, 1, \mu_{\infty})} = |D|^{\frac{s+1}{2}} (-1)^{\frac{l_1-l_2}{2}} \Gamma_{\mathbb{C}}\left(s + \frac{l_1-l_2}{2} + 1\right) f_{\varrho}^{(s)}.$$

For $u \in (\mathbf{K}_\infty^\#)^0$ and A, B as above, by (4-1), (4-16) and Lemma 2.2, we have

$$\begin{aligned} & (v \mid B_\varrho^{T_\theta} (m(a1_2, a) \iota_\theta(u) b_\mathbb{R}^\theta))_\varrho \\ &= (-1)^{\frac{l_1+l_2}{2}} a^{\frac{l_1+l_2}{2}} \exp(-2\sqrt{|D|}\pi a) (v \mid \varrho(C\bar{u}C^{-1})^{-1} v_\varrho^0)_\varrho. \end{aligned}$$

By using the orthogonality of matrix coefficients on $\mathrm{SU}(2)$, the integral in (4-22) is computed as

$$\begin{aligned} & |D|^{\frac{s}{2}} (-1)^{l_1} \Gamma_{\mathbb{C}}(s + \frac{l_1-l_2}{2} + 1) (v \mid v_\varrho^0)_\varrho \int_0^\infty a^{s+\frac{l_1+l_2}{2}-1} \exp(-2\sqrt{|D|}\pi a) d^\times a \\ &= \frac{1}{2} |D|^{\frac{1}{2}(1-\frac{l_1+l_2}{2})} (-1)^{l_1} (v \mid v_\varrho^0)_\varrho \Gamma_{\mathbb{C}}(s + \frac{l_1-l_2}{2} + 1) \Gamma_{\mathbb{C}}(s + \frac{l_1+l_2}{2} - 1) \end{aligned}$$

for $\mathrm{Re}(s) + \frac{l_1+l_2}{2} > 1$. Recall $L(s, \pi_\infty) = \Gamma_{\mathbb{C}}(s + \frac{l_1-l_2}{2} + \frac{1}{2}) \Gamma_{\mathbb{C}}(s + \frac{l_1+l_2}{2} - \frac{3}{2})$. Thus,

$$(4-25) \quad \begin{aligned} Z_v^{(\infty)}(\phi_\infty, \varphi_\pi^0; s, \mu_\infty, \Lambda) \\ = \pi w_D^{-1} \overline{R(\varphi_\pi^0, E, \Lambda)} (v \mid v_\varrho^0)_\varrho (-1)^{l_1} |D|^{\frac{1}{2}(1-\frac{l_1+l_2}{2})} L(s + \frac{1}{2}, \pi_\infty). \end{aligned}$$

The formula in [46, 7.23 Lemma] yields

$$(4-26) \quad M(s) f_\varrho^{(s)} = \frac{\pi}{s-d/2} \prod_{j=1}^{d/2} \frac{s-d/2+j-1}{s+d/2-j} f_\varrho^{(-s)}.$$

Combining this with (3-4) and (4-24), we easily deduce

$$(4-27) \quad f_{\hat{\phi}_\infty}^{(-s, 1, \mu_\infty)} = \frac{|D|}{4} (-1)^{\frac{l_1-l_2}{2}} f_{\phi_\infty}^{(-s, 1, \mu_\infty)}.$$

4.6. The spinor L -function and its functional equation. Let $\pi \cong \otimes_v \pi_v$ be a cuspidal automorphic representation of $\mathrm{G}(\mathbb{A})$ with the trivial central character: then, $\bar{\pi}_p \cong \pi_p^\vee \cong \pi_p$ for all $p < \infty$ by [43, Proposition 2.3]. The twist π_μ of π by an idele class character $\mu : \mathbb{A}^\times/\mathbb{Q}^\times \rightarrow \mathbb{C}^1$ is defined on the space V_π of π as $\pi_\mu(g) := \pi(g) \cdot (\mu \circ \nu)(g)$ for $g \in \mathrm{G}(\mathbb{A})$, so that the central character of π_μ is μ^2 . Let $\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ and μ be as in Section 4.3, so that $\lambda = (l_1, l_2) \in \mathbb{Z}^2$, $l_1 \equiv l_2 \pmod{2}$, $l_1 \geq l_2 > 2$ and the ramification loci of π and μ are disjoint. We define the spinor L -function of π twisted by μ as the Euler product

$$L(s, \pi, \mu) := \prod_{p < \infty} L(s, (\pi_\mu)_p), \quad \mathrm{Re}(s) > \frac{5}{2},$$

with $L(s, (\pi_\mu)_p)$ the local L -factor listed in [34, Table A.8]. Define

$$\hat{L}(s, \pi, \mu) := \Gamma_{\mathbb{C}}(s + \frac{l_1-l_2}{2} + \frac{1}{2}) \Gamma_{\mathbb{C}}(s + \frac{l_1+l_2}{2} - \frac{3}{2}) L(s, \pi, \mu).$$

Using Proposition 3.1 and Lemmas 4.8 and 4.6, combining (4-25), (4-27) and the computations of the local zeta integrals for cases 1, 4, 5, and 6 in table (6-8), we

obtain a meromorphic continuation of $L(s, \pi, \mu)$ to \mathbb{C} as well as the functional equation

$$\hat{L}(s, \pi, \mu) = \varepsilon(s, \pi, \mu) \hat{L}(1 - s, \pi, \bar{\mu})$$

with

$$\varepsilon(s, \pi, \mu) := (-1)^{l_2} \tilde{\mu}(N_\pi^2) \left(\frac{G(\bar{\mu})}{\sqrt{M}} \right)^4 (M^4 N_\pi^2)^{1/2-s},$$

as expected. Moreover $\hat{L}(s, \pi, \mu)$ is holomorphic except for possible simple poles at $s = 3/2, -1/2$, which does not occur when μ is nontrivial. By (4-26), $M(s)f_\rho^{(s)}$ has a simple zero at $s = 1$ if $l_1 > l_2$, which in turn implies the holomorphy at $s = 1$ of the global intertwining operator applied to the section $f_\phi^{(s)}$ as well as the Eisenstein series $E(\phi, s, \Lambda, \mu)$ for ϕ_∞ as above. Hence, by (4-20) with an appropriate ϕ , $\hat{L}(s, \pi, \mu)$ is holomorphic at $s = 3/2$ when $l_1 > l_2$. If μ is real-valued, then

$$G(\bar{\mu})/\sqrt{M} \in \{1, i\} \quad \text{and} \quad \varepsilon(1/2, \pi, \mu) = (-1)^{l_2},$$

so that $L(1/2, \pi, \mu) = 0$ unless l_2 is even.

5. Spectral average of Rankin–Selberg integrals

The space $S_\rho(\mathbf{K}_0(N))$ is endowed with the Hermitian inner product associated to the norm $\int_{Z(\mathbb{A})G(\mathbb{Q}) \backslash G(\mathbb{A})} (\varphi(g) | \varphi(g))_\rho dg$ ($\varphi \in S_\rho(\mathbf{K}_0(N))$). Let $\mathcal{H}(G(\mathbb{Q}_p) // \mathbf{K}_p)$ be the Hecke algebra for $(G(\mathbb{Q}_p), \mathbf{K}_p)$ for $p < \infty$. For any finite set S of primes p that is prime to N , define

$$\mathcal{H}_S := \bigotimes_{p \in S} \mathcal{H}(G(\mathbb{Q}_p) // \mathbf{K}_p).$$

The \mathbb{C} -algebra \mathcal{H}_S acts on the finite-dimensional Hilbert space $S_\rho(\mathbf{K}_0(N))$ normally by

$$[R(f_S)\varphi](g) = \int_{G(\mathbb{Q}_S)} \varphi(gx_S) f_S(x_S) dx_S, \quad g \in G(\mathbb{A}), f_S \in \mathcal{H}_S, \varphi \in S_\rho(\mathbf{K}_0(N)),$$

where $G(\mathbb{Q}_S) := \prod_{p \in S} G(\mathbb{Q}_p)$. Let μ and M be as in Section 4.3. We define the Schwartz–Bruhat function $\phi \in \mathcal{S}(\mathbb{A}_E^2)$ associated with μ as

$$\phi = \prod_{p \leq \infty} \phi_p, \quad \phi_p(x, y) = \begin{cases} \mathbb{1}_{\sigma_{E_p}}(x) \mathbb{1}_{\sigma_{E_p}}(y) & (p < \infty, p \nmid M), \\ \mathbb{1}_{p^e \sigma_{E_p}}(x) \mathbb{1}_{1+p^e \sigma_{E_p}}(y) & (p < \infty, p^e \parallel M, e \geq 1), \end{cases}$$

with ϕ_∞ as in (4-23). In this section, we investigate the averages

$$(5-1) \quad \mathbb{I}^{(s)}(\lambda, N, f_S) := \frac{1}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \sum_{\varphi \in \mathcal{B}(\lambda, N)} \langle E(\phi, s, \Lambda, \mu), R(b)R(f_S)\overline{\varphi_{\mathbf{v}_\rho^0}} \rangle B^{T_\theta, \Lambda}(\varphi_{\mathbf{v}_\rho^0}; b_\mathbb{R}^\theta),$$

where $\mathbf{v}_\rho^0 \in V_\rho^{\mathbf{SO}(2)}$ is the vector from Section 4 and $b = (b_p)_{p \leq \infty} \in G(\mathbb{A})$ is defined

by

$$(5-2) \quad b_p = \begin{cases} 1_4 & (p < \infty, p \nmid NM), \\ \eta_p & (p < \infty, p \mid N), \\ b_p^M := \begin{bmatrix} p^e & 1_2 & T_\theta^\dagger \\ 0 & & 1_2 \end{bmatrix} & (p < \infty, p^e \parallel M, e \geq 1), \\ b_{\mathbb{R}}^\theta & (p = \infty). \end{cases}$$

Here

$$\eta_p := \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & p & 0 & 0 \\ -p & 0 & 0 & 0 \end{bmatrix} \in \mathrm{G}(\mathbb{Q}_p)$$

is the Atkin–Lehner element, $\mathbf{K}_f := \mathbf{K}_0(1)$, and $\mathcal{B}(\lambda, N) = \bigcup_{\pi \in \Pi_{\mathrm{cusp}}(\lambda, N)} \mathcal{B}_\pi(\lambda, N)$ is an orthonormal basis of $S_\varrho(\mathbf{K}_0(N))$ with $\mathcal{B}_\pi(\lambda, N)$ an orthonormal basis of $V_\pi^{\mathbf{K}_0(N)}[\varrho]$. The sum is independent of the choice of an orthonormal basis and can be written as $\mathbb{I}^{(s)}(\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N), f_S)$, where for any subset $X \subset \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ we define

$$(5-3) \quad \mathbb{I}^{(s)}(X, f_S) = \frac{1}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \sum_{\pi \in X} \hat{f}_S(\pi_S) \times \sum_{\varphi \in \mathcal{B}_\pi(\lambda, N)} \langle E(\phi, s, \Lambda, \mu), R(b)\overline{\varphi_{V_\varrho^0}} \rangle B^{T_\theta, \Lambda}(\varphi_{V_\varrho^0}; b_{\mathbb{R}}^\theta),$$

where $\hat{f}_S(\pi_S)$ is the spherical Fourier transforms of f_S at $\pi_S := \bigotimes_{p \in S} \pi_p$, which is defined as the eigenvalue of the operator $\pi_S(f_S)$ on the $\mathbf{K}_S = \prod_{p \in S} \mathbf{K}_p$ -fixed vectors $\pi_S^{\mathbf{K}_S} \cong \mathbb{C}$. For $\bullet \in \{\mathrm{T}, \mathrm{G}, \mathrm{Y}, \mathrm{SK}\}$ and $*$ $\in \{\mathrm{new}, \mathrm{old}\}$, we define $\mathbb{I}^{(s)}(\lambda, N, f_S)^\bullet$, $\mathbb{I}^{(s)}(\lambda, N, f_S)^*$, and $\mathbb{I}^{(s)}(\lambda, N, f_S)^{\bullet,*}$ to be $\mathbb{I}^{(s)}(X, f_S)$ with $X = \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^\bullet$, $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^*$ and $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{\bullet,*}$, respectively. Then, due to (4-18), the average (5-1) has the expression

$$(5-4) \quad \mathbb{I}^{(s)}(\lambda, N, f_S) = \mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{G}, \mathrm{new}} + \mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{Y}, \mathrm{new}} + \mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{SK}, \mathrm{new}} + \mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{old}}.$$

5.1. A construction of orthonormal basis. Let $\pi \cong \bigotimes_{p \leq \infty} \pi_p$ be an element of the set $\Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)$ (see Section 4.3). We fix a (T_θ, Λ) -Bessel data $\{(\ell_p, \xi_p)\}_{p < \infty}$ of π (see Section 4.4) once and for all. Set $\varphi_\pi^0(g) := (\varphi_{\pi, \varrho}^0(g) \mid \mathbf{v}_\varrho^0)_\varrho$, which is an element of $V_\pi^{\mathbf{K}_0(N)}$. Since π is a unitary representation, all of its factors π_p are unitarizable. For each $p < \infty$, we can uniquely fix a $\mathrm{G}(\mathbb{Q}_p)$ -invariant inner product $(\cdot \mid \cdot)_p$ on V_{π_p} by demanding $(\xi_p \mid \xi_p)_p = 1$; for $p = \infty$, we fix a $\mathrm{G}(\mathbb{R})$ -invariant inner product so that its pullback to $V_\varrho \hookrightarrow \pi_\infty$ (cf. property (i) in Section 4.2) coincides with the inner product $(\cdot \mid \cdot)_\varrho$ of V_ϱ . Let φ_π^0 be the global new form

attached to $\{\xi_p\}_{p<\infty}$ (see Section 4.4). Then,

$$(5-5) \quad \frac{(\varphi \mid \varphi)_{L^2}}{(\varphi_\pi^0 \mid \varphi_\pi^0)_{L^2}} = (v \mid v)_\varrho \prod_{p<\infty} (v_p \mid v_p)_p$$

for any $\varphi \in V_\pi$ that corresponds to $v \otimes \left(\bigotimes_{p<\infty} v_p \right)$ with $v \in V_\varrho$. As below, for $p < \infty$, fix an orthonormal basis

$$\mathcal{B}(\pi_p, \mathbf{K}_0(N\mathbb{Z}_p))$$

of $V_{\pi_p}^{\mathbf{K}_0(N\mathbb{Z}_p)}$ in such a way that

$$\mathcal{B}(\pi_p, \mathbf{K}_0(N\mathbb{Z}_p)) = \{\xi_{\pi_p}^0\} \quad \text{if } p \nmid N.$$

Then, a pure tensor of the form

$$(5-6) \quad (\varphi_\pi^0 \mid \varphi_\pi^0)_{L^2}^{-1/2} \times \bigotimes_{p<\infty} v_p, \quad v_p \in \mathcal{B}(\pi_p, \mathbf{K}_0(N\mathbb{Z}_p)),$$

yields an element $\varphi \in V_\pi^{\mathbf{K}_0(N)}[l]$ such that for $v \in V_\varrho$ the element $\varphi_v \in V_\pi$ corresponds to v tensored with (5-6). The set of functions φ obtained in this way from (5-6) will be denoted by $\mathcal{B}_\pi(\lambda, N)$. If $\pi \in \Pi_{\text{cusp}}(\lambda, N)$ does not satisfy condition (A-iv) in Section 4.3, then we fix arbitrary orthonormal basis $\mathcal{B}_\pi(\lambda, N)$ of the space $V_\pi^{\mathbf{K}_0(N)}[\lambda]$. Let $\mathcal{B}(\lambda, N)$ be the union of the sets $\mathcal{B}_\pi(\lambda, N)$ for $\pi \in \Pi_{\text{cusp}}(\lambda, N)$; then, by (ii) in Section 4.2, the set $\mathcal{B}(\lambda, N)$ is an orthonormal basis of $S_\varrho(\mathbf{K}_0(N))$.

5.2. Computation of the average. Let φ correspond to a pure tensor as in (5-6). Then, by (4-20) and the computations recalled in Section 4.5 and Section 6.2,

$$(5-7) \quad \langle E(\phi, s, \Lambda, \mu), R(b)\overline{\varphi_{v_\varrho^0}} \rangle \\ = (\varphi_\pi^0 \mid \varphi_\pi^0)_{L^2}^{-1/2} \pi w_D^{-1} \overline{R(\varphi_\pi^0, E, \Lambda)} \hat{L}(s + \frac{1}{2}, \pi, \mu) \\ \times (v_\varrho^0 \mid v_\varrho^0)_\varrho (-1)^{l_2} 2^{-2} |D|^{\frac{1}{2}(4 - \frac{l_1 + l_2}{2})} \prod_{p \mid N} Z_p^*(\phi_p, \bar{B}_{v_p}; s, \mu_p; b_p),$$

where

$$Z_p^*(\phi_p, \bar{B}_{v_p}; s, \mu_p; b_p) := L(s + \frac{1}{2}, \pi_p, \mu_p)^{-1} Z_p(\phi_p, \bar{B}_{v_p}; s, \mu_p; b_p)$$

is the normalized local zeta integral. Moreover, by (4-21) and Lemma 4.4,

$$(5-8) \quad B^{T_\theta, \Lambda}(\varphi_{v_\varrho^0}; b_\mathbb{R}^\theta) = \\ (\varphi_\pi^0 \mid \varphi_\pi^0)_{L^2}^{-1/2} \pi w_D^{-1} R(\varphi_\pi^0, E, \Lambda) (B_\varrho^{T_\theta}(b_\mathbb{R}^\theta) \mid v_\varrho^0)_\varrho \prod_{p \mid N} \ell_p(v_p).$$

Note that

$$B_\varrho^{T_\theta}(b_\mathbb{R}^\theta) = (-1)^{\frac{l_1 + l_2}{2}} \exp(-2\pi \sqrt{|D|}) v_\varrho^0$$

by (4-1) and (2-14). From (5-3), (5-7) and (5-8), we get:

Proposition 5.1. *Let $f_S \in \mathcal{H}_S$. Then, $\mathbb{I}^{(s)}(\lambda, N, f_S)$ equals*

$$(5-9) \quad \frac{2^{-2}\pi^2 |D|^{\frac{1}{2}(3-\frac{l_1+l_2}{2})} e^{-2\pi\sqrt{|D|}}}{w_D^2[\mathbf{K}_f : \mathbf{K}_0(N)]} M^{s-6} \zeta_M(1)\zeta_M(4)\tilde{\mu}(2D)G(\tilde{\mu}) \\ \times (v_\varrho^0 | v_\varrho^0)_\varrho^2 \sum_{\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)} \hat{f}_S(\pi_S) \frac{|R(\varphi_\pi^0, E, \Lambda)|^2}{(\varphi_\pi^0 | \varphi_\pi^0)_{L^2}} \hat{L}(s + \frac{1}{2}, \pi, \mu) \\ \times \prod_{p < \infty} \mathbb{I}_{\pi_p, \mathbf{K}_0(N\mathbb{Z}_p)}^{(s)}(\xi_p, \phi_p, \Lambda_p, \mu_p; b_p),$$

where ξ_p for $p < \infty$ is from the fixed (T_θ, Λ) -Bessel data $\{(\ell_p, \xi_p)\}_{p < \infty}$ of π , and

$$(5-10) \quad \mathbb{I}_{\pi_p, \mathbf{K}_0(N\mathbb{Z}_p)}^{(s)}(\xi_p, \phi_p, \Lambda_p, \mu_p; b_p) := \sum_{v \in \mathcal{B}(\pi_p, \mathbf{K}_0(N\mathbb{Z}_p))} Z_p^*(\phi_p, \bar{B}_{v_p}; s, \mu_p, b_p) \ell_p(v).$$

We also use the notation $\zeta_M(s)$ to denote $\prod_{p|M} (1 - p^{-s})^{-1}$.

Note that if π_p with $p \nmid N$ is unramified and $(\ell_{\pi_p}^0, \xi_{\pi_p}^0)$ is the unramified Bessel datum of π_p , then (5-10) is 1. For other cases, we compute (5-10) in Section 6.8 completely. Substituting them, we obtain:

Theorem 5.2. *Let $\lambda = (l_1, l_2)$, $N, \mu, \tilde{\mu}$ and M be as in Section 4.3. Let S be a finite set of prime numbers relatively prime to DMN . Then, for any $f_S = \bigotimes_{p \in S} f_p \in \mathcal{H}_S$, the value $\mathbb{I}^{(s)}(\lambda, N, f_S)$ equals*

$$(5-11) \quad \frac{2^{-2}\pi^2 |D|^{\frac{1}{2}(3-\frac{l_1+l_2}{2})} e^{-2\pi\sqrt{|D|}}}{w_D^2[\mathbf{K}_f : \mathbf{K}_0(N)]} \\ \times M^{s-6} \zeta_M(1)\zeta_M(4)\tilde{\mu}(2D)G(\tilde{\mu})N^{s-1}\tilde{\mu}(N) \prod_{p|N} (1 + p^{-2})^{-1} \\ \times (v_\varrho^0 | v_\varrho^0)_\varrho^2 \sum_{\pi \in \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)} \hat{f}_S(\pi_S) \frac{|R(\varphi_\pi^0, E, \Lambda)|^2}{(\varphi_\pi^0 | \varphi_\pi^0)_{L^2}} \hat{L}(s + \frac{1}{2}, \pi, \mu) \mathfrak{t}^{(s)}(\pi, \mu),$$

with $\mathfrak{t}^{(s)}(\pi, \mu) = \prod_{p|N} \mathfrak{t}^{(s)}(\pi_p, \mu_p)$ and $\mathfrak{t}^{(s)}(\pi_p, \mu_p)$ defined as

$$(5-12) \quad \begin{cases} 1 & \text{if } p \mid N_\pi \text{ and } \pi_p \text{ is of type VIb,} \\ 2 & \text{if } p \mid N_\pi \text{ and } \pi_p \text{ is of type IIIa,} \\ 2(p-1)p^{-5}L(1, \pi_p, \mathrm{Std}) \\ & \times (1 - \frac{\mu_p(p)p^{-s}}{p+1} \mathrm{tr}(p^{-1}T_{1,0} + \eta_p|\pi_p^{\mathbf{K}_0(p\mathbb{Z}_p)}) + \mu_p^2(p)p^{-2s}) \\ & \text{if } p \mid \frac{N}{N_\pi}. \end{cases}$$

To refine (5-11) by considering $\mathbb{I}^{(s)}(\lambda, N, f_S)^{\bullet, \text{new}}$ ($\bullet \in \{G, Y, SK\}$), we use an explicit formula of the quantity $|R(\varphi_\pi^0, E, \Lambda)|^2 / (\varphi_\pi^0 | \varphi_\pi^0)_{L^2}$ first proved by Dickson, Pitale, Saha, and Schmidt [7, Theorems 1.13 and 3.13] when $l_1 = l_2$ under the refined GGP conjecture for Bessel periods posed by Liu [22]. Thanks to a theorem by Furusawa and Morimoto [10, Theorem 8.1], the formula in [7, Theorems 1.13 and 3.13] is extended to vector-valued forms unconditionally so that it can be applied to (5-11).

Theorem 5.3. *Let the notation and assumptions be as in Theorem 5.2.*

(i) $\mathbb{I}^{(s)}(\lambda, N, f_S)^{G, \text{new}}$ equals

$$\frac{2^{\#S(N)+2l_1-6} \pi^2 |D|^{\frac{l_1+l_2+2}{4}} e^{-2\pi\sqrt{|D|}}}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \times M^{s-6} \zeta_M(1) \zeta_M(4) \tilde{\mu}(2D) G(\tilde{\mu}) N^{s-1} \tilde{\mu}^{-1}(N) \prod_{p \in S(N)} (1 + p^{-1})$$

$$\times \sum_{\pi \in \Pi_{\text{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{G, \text{new}}} \hat{f}_S(\pi_S) \frac{\hat{L}(\frac{1}{2}, \pi \times AI(\Lambda^{-1}))}{\hat{L}(1, \pi; \text{Ad})} \hat{L}(s + \frac{1}{2}, \pi, \mu),$$

where $S(N)$ is the set of all the prime divisors of N , and $AI(\Lambda^{-1})$ is the automorphic induction to $\mathbf{GL}_2(\mathbb{A})$ from the character Λ^{-1} of $\mathbb{A}_E^\times / E^\times$.

(ii) If N has an even number of prime divisors, then $\mathbb{I}^{(s)}(\lambda, N, f_S)^{Y, \text{new}} = 0$. If N has an odd number of prime divisors, then $\mathbb{I}^{(s)}(\lambda, N, f_S)^{Y, \text{new}}$ equals

$$\frac{2^{\#S(N)+2l_1-7} \pi^2 |D|^{\frac{l_1+l_2+2}{4}} e^{-2\pi\sqrt{|D|}}}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \times M^{s-6} \zeta_M(1) \zeta_M(4) \tilde{\mu}(2D) G(\tilde{\mu}) N^{s-1} \tilde{\mu}^{-1}(N) \prod_{p \in S(N)} (1 + p^{-1})$$

$$\times \sum_{\substack{\pi_1 \in \Pi_{\mathbf{PGL}_2, \text{cusp}}(l_1+l_2-2, N)^{\text{new}} \\ \pi_2 \in \Pi_{\mathbf{PGL}_2, \text{cusp}}(l_1-l_2+2, N)^{\text{new}}}} \left(\hat{f}_S(Y(\pi_1, \pi_2)_S) \right.$$

$$\left. \times \frac{\hat{L}(\frac{1}{2}, \pi_1 \times AI(\Lambda^{-1})) \hat{L}(\frac{1}{2}, \pi_2 \times AI(\Lambda^{-1})) \hat{L}(s + \frac{1}{2}, \pi_1 \times \mu) \hat{L}(s + \frac{1}{2}, \pi_2 \times \mu)}{\hat{L}(1, \pi_1; \text{Ad}) \hat{L}(1, \pi_2; \text{Ad}) \hat{L}(1, \pi_1 \times \pi_2)} \right),$$

where $\Pi_{\mathbf{PGL}_2, \text{cusp}}(k, N)^{\text{new}}$ is the set of all irreducible cuspidal representations of $\mathbf{PGL}_2(\mathbb{A})$ associated to holomorphic newforms of weight k and level N and $Y(\pi_1, \pi_2) \in \Pi_{\text{cusp}}(\lambda, N)^{\text{new}}$ denotes the Yoshida lift of π_1 and π_2 .

(iii) If $\Lambda \neq \mathbb{1}$ or $l_1 > l_2$, then $\mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{SK}, \mathrm{new}} = 0$. If $\Lambda = \mathbb{1}$ and $l_1 = l_2 (= l)$, then $\mathbb{I}^{(s)}((l, l), N, f_S)^{\mathrm{SK}, \mathrm{new}}$ equals

$$\frac{3 \cdot 2^{2l-3} \pi^2 |D|^{(l+2)/2} e^{-2\pi\sqrt{|D|}}}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \cdot \frac{s}{4\pi} \\ \times M^{s-6} \zeta_M(1) \zeta_M(4) \tilde{\mu}(2D) G(\tilde{\mu}) N^{s-1} \tilde{\mu}^{-1}(N) \prod_{p \in S(N)} (1+p^{-1}) \cdot \hat{L}(s+1, \mu) \hat{L}(s, \mu) \\ \times \sum_{\pi_0 \in \Pi_{\mathrm{PGL}_2, \mathrm{cusp}}^{(T_\theta, 1)}(2l-2, N)^{\mathrm{new}}} \hat{f}_S(\mathrm{SK}(\pi_0)_S) \frac{\hat{L}(\frac{1}{2}, \pi_0 \times \chi_D) \hat{L}(1, \chi_D)^2}{\hat{L}(\frac{3}{2}, \pi_0) \hat{L}(1, \pi_0; \mathrm{Ad})} \hat{L}(s + \frac{1}{2}, \pi_0 \times \mu),$$

where $\Pi_{\mathrm{PGL}_2, \mathrm{cusp}}^{(T_\theta, 1)}(2l-2, N)^{\mathrm{new}}$ is the set of all $\pi_0 \in \Pi_{\mathrm{PGL}_2, \mathrm{cusp}}(2l-2, N)^{\mathrm{new}}$ such that the Saito–Kurokawa lift $\mathrm{SK}(\pi_0)$ of π_0 has the $(T_\theta, \mathbb{1})$ -Bessel model, and χ_D is the Kronecker character of modulo D .

Proof. The equalities in (i) and (ii) are a direct corollary to Theorem 5.2 and [9, Theorem 8.1] (for the scalar case we refer to [7, Theorems 1.13 and 3.14]); note that the polynomial $Q_{S, \varrho}$ (with $S = T_\theta$) in [9, (8.2.16)] equals

$$(-1)^{\frac{l_1-l_2}{2}} \left(\frac{2}{\sqrt{|D|}}\right)^{\frac{l_1+l_2}{2}} v_\varrho^\theta,$$

our $R(\varphi, E, \Lambda)$ equals the quantity $w_D \left(\frac{\sqrt{|D|}}{2}\right)^{-\frac{l_1+l_2}{2}} \pi \mathcal{B}_\Lambda(\varphi; E) / (Q_{S, \varrho}, Q_{S, \varrho})_{l_1-l_2}$ defined by [9, (8.2.17)], and $(Q_{S, \varrho}, Q_{S, \varrho})_{l_1-l_2} = \left(\frac{|D|}{4}\right)^{-(l_1+l_2)/2} (v_\varrho^0, v_\varrho^0)_{l_1-l_2} = \left(\frac{|D|}{4}\right)^{-(l_1+l_2)/2} (v_\varrho^0 | v_\varrho^0)_\varrho$. We also note the formula

$$\hat{L}(s, \pi, \mu) = \hat{L}(s, \pi_1 \times \mu) \hat{L}(s, \pi_2 \times \mu),$$

where $\pi = Y(\pi_1, \pi_2)$. This is obtained immediately by comparing the local factors. It is noted in [7, p. 296] that each local representation π_p of $\pi \cong \bigotimes_p \pi_p \in \Pi_{\mathrm{cusp}}^{(T_\theta, \Lambda)}(\lambda, N)^{Y, \mathrm{new}}$ for $p \mid N$ is of type VIb. Finally, we check statement (iii). As quoted in [7, Theorem 3.11], the (T_θ, Λ) -Bessel periods of the Saito–Kurokawa lifts are zero unless Λ is trivial due to Qiu. By comparing the local L -functions [38, Theorem 5.2(ii)] (cf. [28, Theorem 3.1]), we have

$$(5-13) \quad \hat{L}(s, \pi, \mu) = \frac{1}{4\pi} (s-1/2) \hat{L}(s, \pi_0 \times \mu) \hat{L}(s+1/2, \mu) \hat{L}(s-1/2, \mu),$$

where $\pi = \mathrm{SK}(\pi_0)$. By [38] and the definition, the set $\Pi_{\mathrm{cusp}}^{(T_\theta, 1)}((l, l), N)^{\mathrm{SK}}$ corresponds bijectively to the set $\Pi_{\mathrm{PGL}_2, \mathrm{cusp}}^{(T_\theta, 1)}(2l-2, N)$ via the Saito–Kurokawa lifting if $l \geq 3$. Note that the condition $l \geq 3$ is ensured by assumption (A-ii) in Section 4.3. □

By (5-4), $\mathbb{I}^{(s)}(\lambda, N, f_S) - \mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{G}, \mathrm{new}}$ is the sum of $\mathbb{I}^{(s)}(\lambda, N, f_S)^{Y, \mathrm{new}}$, $\mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{SK}, \mathrm{new}}$ and $\mathbb{I}^{(s)}(\lambda, N, f_S)^{\mathrm{old}}$. Our main focus is their values at $s = 0$.

For those, we have the following upper bounds in the level aspect. For $f_S \in \mathcal{H}_S$, let $\|f_S\|_1 := \int_{G(\mathbb{Q}_S)} |f_S(g_S)| dg_S$ denote its L^1 -norm.

Theorem 5.4. *Let $l, \mu, \tilde{\mu}, M, S$ be as in Theorem 5.2.*

(i) *Let λ be fixed as before and N an inert prime in E/\mathbb{Q} . Then, there exist constants $C_{D,\lambda,\mu}, C'_{D,\lambda,\mu} > 0$ depending only on D, l, μ , and f_S such that*

$$(5-14) \quad |\mathbb{I}^{(0)}(\lambda, N, f_S)^{\text{old}}| < C_{D,\lambda,\mu} \|f_S\|_1 N^{-8},$$

$$(5-15) \quad |\mathbb{I}^{(0)}(\lambda, N, f_S)^{\text{Y,new}}| < C'_{D,\lambda,\mu} \|f_S\|_1 N^{-\frac{3}{2}}.$$

(ii) *Let N be 1 or an inert prime in E/\mathbb{Q} . When $\mu \neq \mathbf{1}$, $\mathbb{I}^{(0)}(l, N, f_S)^{\text{SK,new}} = 0$. When $\mu = \mathbf{1}$, there exists a constant $C > 0$ such that*

$$(5-16) \quad |\mathbb{I}^{(0)}((l, l), N, f_S)^{\text{SK,new}}| < C |D|^{\frac{l}{2}} l^{\frac{1}{2}} \|f_S\|_1 N^{-\frac{3}{2}}.$$

Proof. We have the inequality $|\hat{f}_S(\pi_S)| \leq \|f_S\|_1$ for all irreducible unitary representation π_S of $G(\mathbb{Q}_S)$. Because N is a prime, one can check that the summation range of $\mathbb{I}^{(0)}(\lambda, N, f_S)^{\text{old}}$ is $\Pi_{\text{cusp}}^{(T_\theta, \chi)}(\lambda, N)^{\text{old}} = \Pi_{\text{cusp}}^{(T_\theta, \chi)}(\lambda, 1)$, which is independent of N . From the value of $t(\pi_p, \mu_p)$ in (5-12) for $p = N$, combined with the temperedness of π_N due to [48] and the matrix of $T_{1,0}$ in [31, Table 3], we get $t(\pi_N, \mu_N) = O(N^{-4})$. This and the equation $[\mathbf{K}_f : \mathbf{K}_0(N)] = N^3(1 + N^{-2})(1 + N^{-1})$ for prime N yield the bound (5-14).

Next we treat the average for Yoshida lifts. For $\pi_1 \in \Pi_{\mathbf{PGL}_2, \text{cusp}}(l_1 + l_2 - 2, N)^{\text{new}}$ and $\pi_2 \in \Pi_{\mathbf{PGL}_2, \text{cusp}}(l_1 - l_2 + 2, N)^{\text{new}}$, we need the lower bound

$$L(1, \pi_1 \times \pi_2) \gg_{l_1, l_2} \exp(-C \sqrt{\log N})$$

uniform in N . This is a special case of Lemma 5.5, because π_1 and π_2 are everywhere tempered by Deligne’s estimate and by the fact that the local p -components of π_i for $p \mid N$ are the (twisted) Steinberg representation, which is tempered. Note that $\pi_1 \not\cong \pi_2$ due to the weight condition. Now, we deduce the inequality (5-15) bounding the sum from above by a product of the average considered in [8, Theorem 1.1]; to do this, we invoke the subconvexity bound

$$(5-17) \quad L\left(\frac{1}{2}, \pi\right) = O(C(\pi)^{\frac{1}{4}-\delta}) \quad (\exists \delta > 0)$$

for automorphic cuspidal representations π of $\mathbf{GL}_2(\mathbb{A})$ [24] and the nonnegativity of $L\left(\frac{1}{2}, \pi_i \times \mathcal{AI}(\Lambda^{-1})\right) = L\left(\frac{1}{2}, \text{BC}_{E/\mathbb{Q}}(\pi_i) \otimes \Lambda^{-1}\right)$ for $i = 1, 2$ due to [15].

Let us prove (5-16); by Theorem 5.3(ii), we may assume $\lambda = (l, l)$. Suppose $\mu \neq \mathbf{1}$; then, by Theorem 5.3(ii), $\mathbb{I}^{(s)}((l, l), N, f_S)^{\text{SK,new}} \times s^{-1}$ is entire, hence $\mathbb{I}^{(0)}((l, l), N, f_S)^{\text{SK,new}} = 0$ because $\hat{L}(s, \mu)$, as well as $\hat{L}(s, \pi_0 \times \chi_D)$, is entire. In the rest of the proof, we assume $\mu = \mathbf{1}$. Then, $s \hat{L}(s, \mu) \hat{L}(s + 1, \mu)$ has a simple pole at $s = 0$. By [37, Theorem 3.1 and Table 2], $\text{SK}(\pi_0)$ with $\pi_0 \in \Pi_{\text{cusp}, \mathbf{PGL}_2}(2l - 2, N)$ has the global (T_θ, Λ) -Bessel model only if the sign of the functional equation of

π_0 is -1 so that $\hat{L}(s + 1/2, \pi_0)$ has a zero at $s = 0$. Hence, Theorem 5.3(ii) gives us the following majorant of $\mathbb{1}^{(0)}((l, l), N, f_S)^{\mathrm{SK}, \mathrm{new}}$:

$$\frac{2^3 \pi^{\frac{5}{2}} (l-1)^{-2}}{[\mathbf{K}_f : \mathbf{K}_0(N)]} \cdot \frac{\Gamma(l) \|f_S\|_1}{\Gamma(l - \frac{1}{2})} |D|^{\frac{l}{2}} \sum_{\pi_0 \in \Pi_{\mathrm{PGL}_2, \mathrm{cusp}}^{(T_\theta, 1)}(2l-2, N)} \frac{L(\frac{1}{2}, \pi_0 \times \chi_D) |L'(\frac{1}{2}, \pi_0)|}{L(1, \pi_0; \mathrm{Ad})}.$$

To estimate this, we invoke a subconvexity bound $L(1/2, \pi_0 \times \chi_D) = O((l^2 N)^{\frac{1}{4}-\delta})$ for some $\delta > 0$ from (5-17) and a lower bound

$$L(1, \pi_0; \mathrm{Ad}) > c_0 \exp(-c_1 \sqrt{\log(1 + l^2 N)})$$

for some constant $c_0, c_1 > 0$ which is known by [12, Theorem 0.1] (see also the remark after [21, Corollary 7]). The lower bound implies $L(1, \pi_0; \mathrm{Ad})^{-1} = O((l^2 N)^{2\delta})$. By a common argument, we derive a subconvexity bound for the central derivative $L'(1/2, \pi_0) = O((l^2 N)^{\frac{1}{4}-\delta})$ from the bound $L(1/2, \pi_0 | \cdot |_{\mathbb{A}}^{it}) = O((l^2 N(1 + |t|))^{1/4-\delta})$ that follows from (5-17). Finally, (5-16) follows using the uniform bound $\#\Pi_{\mathrm{PGL}_2, \mathrm{cusp}}(2l-2, N) = O(lN)$ and the asymptotic $\frac{\Gamma(l)}{\Gamma(l-\frac{1}{2})} \sim l^{\frac{1}{2}}$ ($l \rightarrow \infty$). □

Lemma 5.5. *Let π_1 and π_2 be irreducible cuspidal automorphic representations of PGL_n such that $C(\pi_1), C(\pi_2) \leq Q$ with $Q > 2$. We assume that $\pi_1 \not\cong \pi_2$ and both of them are self-dual and tempered everywhere. Then,*

$$L(1, \pi_1 \times \pi_2) \gg \exp(-C \sqrt{\log Q})$$

with an absolute constant $C > 0$.

Proof. We recall the argument indicated in [35] (attributed originally to [25]), which eliminates a possibility of Siegel zeros of the L -function $L(s, \pi_1 \times \pi_2)$. Fix $t \in \mathbb{R}$. Then, $L(s, \Pi \times \tilde{\Pi})$ with Π being the isobaric sum

$$\Pi := \pi_1 \boxplus (\pi_1 \times | \cdot |_{\mathbb{A}}^{it}) \boxplus (\pi_1 \times | \cdot |_{\mathbb{A}}^{-it}) \boxplus \pi_2 \boxplus (\pi_2 \times | \cdot |_{\mathbb{A}}^{it}) \boxplus (\pi_2 \times | \cdot |_{\mathbb{A}}^{-it})$$

is a Dirichlet series with nonnegative coefficients [13, Lemma a]. Moreover, by [16, Proposition 9.4] and the self-duality of π_1 and π_2 , one expresses $L(s, \Pi \times \tilde{\Pi})$ as

$$L(s, \Pi \times \tilde{\Pi}) = \prod_{1 \leq i, j \leq 2} \left[\frac{L(s, \pi_i \times \pi_j)^3 L(s-it, \pi_i \times \pi_j)^2 L(s+it, \pi_i \times \pi_j)^2}{\times L(s-2it, \pi_i \times \pi_j) L(s+2it, \pi_i \times \pi_j)} \right]^2,$$

which shows that $L(s, \Pi \times \tilde{\Pi})$ has a pole of order 6 at $s = 1$ (due to $\pi_1 \not\cong \pi_2$), and has a zero of order 8 at $s = \sigma$ if $L(\sigma + it, \pi_i \times \pi_j) = 0$ ($i, j = 1, 2$). Hence, by [11, p. 178, Lemma], one can show that $L(s, \pi_1 \times \pi_2)$ has no zeros on the interval $(1 - \frac{C_0}{\log M}, 1)$ for some constant $C_0 > 0$ with

$$M = (1 + |t|)^{24} C(\pi_2 \times \pi_2)^{18} C(\pi_1 \times \pi_1)^9 C(\pi_2 \times \pi_2)^2,$$

where $C(\pi_i \times \pi_j)$ is the analytic conductor of $L(s, \pi_i \times \pi_j)$. By [13, Lemma b], we have $C(\pi_i \times \pi_j) \ll Q^{2n}$. Thus, for an absolute constant $C'_0 > 0$, $L(s, \pi_1 \times \pi_2)$ is zero-free on the region $1 - \frac{C'_0}{\log Q(1+|\text{Im}(s)|)} < \text{Re}(s) < 1$. Now, we apply the argument of [21, Corollary 7] to the L -function $L(s, \pi_1 \times \pi_2)$. Since π_1, π_2 are assumed to be tempered everywhere, the optimal bound of the lambda function $\Lambda_{\mathcal{A}}(n)$ for $\mathcal{A} = \pi_1 \times \pi_2$ is available so that the automorphy of $\pi_1 \times \pi_2$ is not necessary, which simplifies the proof to get the lower bound of $L(1, \pi_1 \times \pi_2)$. \square

As a consequence of Theorem 5.4, we get

$$(5-18) \quad \mathbb{I}^{(0)}(\lambda, N, f_S)^{\text{G,new}} = \mathbb{I}^{(0)}(\lambda, N, f_S) + O_{\Lambda, \lambda, \mu}(\|f_S\|_1 N^{-\frac{3}{2}}).$$

6. Computation of local zeta integrals for p -adic fields

In this section, let F be a nonarchimedean local field of characteristic 0, \mathcal{O} the integer ring of F , \mathfrak{p} the maximal ideal of \mathcal{O} , ϖ a generator of \mathfrak{p} and $q = \#(\mathcal{O}/\mathfrak{p})$. Let $|\cdot|$ denote the normalized absolute value of F , i.e., $|\varpi| = q^{-1}$. Fix a nontrivial additive character $\psi : F \rightarrow \mathbb{C}^1$ with $\text{cond}(\psi) := \min\{n \in \mathbb{Z} ; \psi|_{\varpi^n \mathcal{O}} = 1\} = 0$.

We compute the local zeta integral à la Piatetski-Shapiro [29] for several representations, taking particular test functions; as a result, we determine the local L -factors and the local ε -factors in [29] to confirm that they coincides with the expected ones listed in [34, Tables A8 and A9]. As explained in Section 4.3, for a particular global application in mind, we only deal with representations of types I, IIb, IIIa and VIb (but allowing the central characters to be nontrivial when we are concerned with newvectors.)

6.1. Local zeta integral for Bessel models. We first review some generalities on local zeta integrals and then recall results from [32] on explicit formulas of Bessel functions for Iwahori spherical representations of G , which are possible local components of $\pi \cong \otimes_p \pi_p \in \Pi_{\text{cusp}}^{(T_\theta, \Delta)}(\lambda, N)$.

Let $K = G(\mathcal{O})$ be a standard maximal compact subgroup of $G(F)$ and

$$K_0(\mathfrak{p}^e) := \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in K \mid C \equiv 0 \pmod{\mathfrak{p}^e} \right\}$$

be the congruence subgroup of level \mathfrak{p}^e .

For a symmetric matrix $T = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \in \text{Sym}_2(F)$ such that $d := b^2 - 4ac \neq 0$, set $\xi = \begin{bmatrix} b/2 & c \\ -a & -b/2 \end{bmatrix}$. Let $L := F \oplus F\xi$ be the two-dimensional F -algebra, \mathcal{O}_L be its integer ring, and $\mathfrak{p}_L := \mathfrak{p}\mathcal{O}_L$. Define the additive character ψ_L on L by $\psi_L := \psi \circ \text{tr}_{L/F}$. The maps

$$\begin{aligned} L \ni x + y\xi &\mapsto x + \frac{\sqrt{d}}{2}y \in F(\sqrt{d}) && (d \notin (F^\times)^2), \\ L \ni x + y\xi &\mapsto \left(x + \frac{\sqrt{d}}{2}y, x - \frac{\sqrt{d}}{2}y\right) \in F \oplus F && (d \in (F^\times)^2), \end{aligned}$$

are isomorphisms of F -algebras. We assume that

- (i) $a \in \mathcal{O}^\times$ and $b, c \in \mathcal{O}$.
- (ii) If $d \notin (F^\times)^2$, then $\sqrt{d}\mathcal{O}_L$ is the different ideal of L/F .
- (iii) If $d \in (F^\times)^2$, then $d \in \mathcal{O}^\times$.

We can check that

$$L^\times = \{ \tau \in \mathbf{GL}_2(F) \mid {}^t \tau T \tau = \det \tau \cdot T \}.$$

The map $L \ni \tau \mapsto \tau^\dagger \in L$ is the nontrivial F -automorphism on L . We set $\theta_0 = b/2 - \xi \in \mathcal{O}_L$, then $\{a, \theta_0\}$ is an \mathcal{O} -basis of \mathcal{O}_L . Let $\langle \cdot, \cdot \rangle$ denote the symplectic form on L^2 over F defined by

$$\langle x, y \rangle := \mathrm{tr}_{L/F}(-2\xi)^{-1}(x_1 y_2 - x_2 y_1), \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in L^2.$$

We set

$$\begin{aligned} G^\#(F) &:= \{ g \in \mathbf{GL}_2(L) \mid \det g \in F^\times \}, \\ B^\#(F) &:= \left\{ \begin{bmatrix} \tau & \beta \\ 0 & a\tau^\dagger \end{bmatrix} \mid \tau \in L^\times, \beta \in L, a \in F^\times \right\}, \\ K^\# &:= G^\#(F) \cap \mathbf{GL}_2(\mathcal{O}_L), \\ K_0^\#(\mathfrak{p}^e) &:= G^\#(F) \cap \begin{bmatrix} \mathcal{O}_L & \mathcal{O}_L \\ \mathfrak{p}_L^e & \mathcal{O}_L \end{bmatrix}. \end{aligned}$$

Since $\langle g^\#x, g^\#y \rangle = \det g^\# \langle x, y \rangle$ for any $x, y \in L^2$ and $g^\# \in G(F)$, we get a natural embedding $G^\#(F) \ni g \mapsto \iota(g^\#) \in G(F)$. More precisely, $\iota(g^\#)$ is the representation matrix of the action $L^2 \ni x \mapsto g^\#x \in L^2$ with respect to an F -basis $\left\{ \begin{bmatrix} a \\ 0 \end{bmatrix}, \begin{bmatrix} \theta_0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -a^{-1}\theta_0^\dagger \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$. A computation yields

$$(6-1) \quad \iota\left(\begin{bmatrix} \tau & 0 \\ 0 & a\tau^\dagger \end{bmatrix}\right) = m(\tau, a \det \tau),$$

$$\iota\left(\begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1_2 & X\beta \\ 0 & 1_2 \end{bmatrix}, \quad \beta = \beta_2 a + \beta_3 \theta_0 \in L, \quad X\beta := \begin{bmatrix} -a^{-1}b\beta_2 - a^{-1}c\beta_3 & \beta_2 \\ \beta_2 & \beta_3 \end{bmatrix},$$

$$(6-2) \quad \iota\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & -a^{-2}b & a^{-1} \\ 0 & 0 & a^{-1} & 0 \\ 0 & a & 0 & 0 \\ a & b & 0 & 0 \end{bmatrix}.$$

We note that $K^\# = \iota^{-1}(K)$ and $K_0^\#(\mathfrak{p}^e) = \iota^{-1}(K_0(\mathfrak{p}^e))$. We call

$$R := \left\{ \iota\left(\begin{bmatrix} \tau & 0 \\ 0 & \tau^\dagger \end{bmatrix}\right)n \mid \tau \in L^\times, n \in \mathbf{N}(F) \right\}$$

the Bessel subgroup of $G(F)$ (with respect to T). For a character $\Lambda : L^\times \rightarrow \mathbb{C}^1$, we can check that the map

$$R \ni \iota\left(\begin{bmatrix} \tau & 0 \\ 0 & \tau^\dagger \end{bmatrix}\right)n(X) \mapsto \Lambda(\tau)\psi(\mathrm{tr}(TX)) \in \mathbb{C}^1$$

defines a character on R . We denote this character by $\Lambda \otimes \psi_T$.

Let (π, V_π) be an irreducible admissible representation of $G(F)$. Let $(V_\pi^*)^{T, \Lambda}$ denote the space of all \mathbb{C} -linear forms $\ell : V_\pi \rightarrow \mathbb{C}$ that satisfies

$$\ell(\pi(r)\xi) = \Lambda \otimes \psi_T(r) \ell(\xi), \quad \xi \in V_\pi, r \in R.$$

Then, $\dim_{\mathbb{C}}(V_{\pi}^*)^{T,\Lambda} \leq 1$ [26; 33]. We say that π has a local (T, Λ) -Bessel model if $(V_{\pi}^*)^{T,\Lambda} \neq (0)$. In this case, we can define the local (T, Λ) -Bessel model of π as in Section 4.4. Moreover, when π is spherical and ψ_T and Λ are unramified, we define the unramified (T, Λ) -Bessel datum $(\ell_{\pi}^0, \xi_{\pi}^0) \in (V_{\pi}^*)^{T,\Lambda} \times V_{\pi}^K$ for π as in Section 4.4, so that

$$B_{\pi}^0(g) = \ell_{\pi}^0(\pi(g)\xi_{\pi}^0), \quad g \in G(\mathbb{Q}_p),$$

is the unramified Bessel function such that $B_{\pi}^0(1_4) = 1$. Recall that all irreducible admissible representations of $G(F)$ that admit local Bessel models are classified in [34], and the result is conveniently summarized in [32, Table 2]. For our cases, as we mentioned in Section 4.4, we may assume that π is of type I, IIb, IIIa, or VIb. We quote some of its properties from [32, Table 2]:

type	π	$\dim V_{\pi}^K$	$\dim V_{\pi}^{K_0(\mathfrak{p})}$	cent. char.
I	$\chi \times \chi' \rtimes \sigma$	1	4	$\chi\chi'\sigma^2$
IIb	$\chi 1_{\mathbf{GL}_2} \rtimes \sigma$	1	3	$\chi^2\sigma^2$
IIIa	$\chi \rtimes \sigma_{\text{St}_{\mathbf{GL}_2}}$	0	2	$\chi\sigma^2$
VIb	$\tau(T, \cdot ^{-\frac{1}{2}}\sigma)$	0	1	σ^2

Here, χ, χ' , and σ are unramified quasicharacters on F^{\times} such that representations are irreducible and admit (T, Λ) -Bessel models. For later use we set $\alpha = \chi(\varpi)$, $\beta = \chi'(\varpi)$, and $\gamma = \sigma(\varpi)$. According to [34, Table A2], for type IIIa and type VIb, the corresponding representations are unitarizable if and only if the inducing quasicharacters are unitary, whereas, for type I, the unitarity of the inducing quasicharacters is equivalent to the corresponding representations being unitarizable and tempered. In what follows, having a global application in mind, we suppose the unitarity of all the inducing characters, i.e., $|\alpha| = |\beta| = |\gamma| = 1$.

Similarly to (3-1), for a Schwartz–Bruhat function $\phi \in \mathcal{S}(L^2)$, a character μ on F^{\times} , and $s \in \mathbb{C}$ with $\text{Re}(s) > 1$, we define a function on $G^{\#}(F)$ by

$$f_{\phi}^{(s,\Lambda,\mu)}(g^{\#}) = \mu(\det g^{\#})|\det g^{\#}|^{s+1} \int_{L^{\times}} \phi\left(\begin{bmatrix} 0 & 1 \\ \tau & 0 \end{bmatrix} g^{\#}\right) \Lambda \mu_L(\tau) |\tau \tau^{\dagger}|^{s+1} d^{\times} \tau,$$

where $\mu_L(\tau) := \mu(\tau \tau^{\dagger})$, $\tau \in L^{\times}$ and $d^{\times} \tau$ is the normalized Haar measure on L^{\times} such that $\text{vol}(\mathcal{O}_L) = 1$. For an irreducible smooth admissible representation π of $Z(F) \backslash G(F)$, admitting (T, Λ) -Bessel model and $B \in \mathcal{B}(T, \Lambda)[\pi]$, define the zeta integral by

$$Z(\phi, B, s, \mu; g) = \int_{F^{\times}} \int_{K^{\#}} B(m(a1_2, a)\iota(k^{\#})g)\mu(a)|a|^{s-1} f_{\phi}^{(s,\Lambda,\mu)}(k^{\#})dk^{\#}d^{\times} a$$

for $s \in \mathbb{C}$ with $\text{Re}(s) > 1$. It is shown that $Z(\phi, B, \mu; g)$ is a rational function in q^{-s} [40, Lemma 5.3.1]. Then, the local L -function $L(s, \pi, \mu)$ is defined to be the

unique function such that $L(s, \pi, \mu)^{-1}$ is a polynomial in q^{-s} with constant term 1 and such that $Z(\phi, B, s, \mu, 1)/L(s, \pi, \mu) \in \mathbb{C}[q^s, q^{-s}]$ for all ϕ and B ([29]; see also [40, Proposition 5.3.3]). For later use, we need the following:

Lemma 6.1. *Suppose Λ is unramified. Let $B \in \mathcal{B}(T, \Lambda)^{K_0(\mathfrak{p})}$ and $\phi = \mathbb{1}_{\mathfrak{o}_L \oplus \mathfrak{o}_L}$. Then, for $\mathrm{Re}(s) \gg 0$,*

$$(6-5) \quad Z(\phi, B, s, \mu; \eta) = \frac{L(s+1, \Lambda\mu_L)}{q^2+1} \left\{ \sum_{l=0}^{\infty} \eta B(h(l, 0)) \mu(\varpi)^l q^{-l(s-1)} + q^{s+1} \mu(\varpi)^{-1} \sum_{l=0}^{\infty} B(h(l, 0)) \mu(\varpi)^l q^{-l(s-1)} \right\},$$

where ηB denotes the right-translation of B by $\eta := \begin{bmatrix} & & & -1 \\ & & & \varpi \\ & & & \varpi \\ -\varpi & & & \end{bmatrix}$.

Proof. It is easy to check the equality $f_{\phi}^{(s, \Lambda, \mu)}(k^{\#}) = L(s+1, \Lambda\mu_L)$ for all $k^{\#} \in K^{\#}$. By considering the left $K_0^{\#}(\mathfrak{p})$ -coset decomposition of $K^{\#}$,

$$(6-6) \quad K^{\#} = K_0^{\#}(\mathfrak{p}) \sqcup \left(\bigsqcup_{\xi \in \mathfrak{o}_L/\mathfrak{p}_L} \begin{bmatrix} 1 & \xi \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} K_0^{\#}(\mathfrak{p}) \right),$$

the zeta integral $Z(\phi, B, s, \mu; \eta)$ becomes $[K^{\#} : K_0^{\#}(\mathfrak{p})]^{-1} L(s+1, \Lambda\mu_L)$ times

$$\int_{F^{\times}} \left(\eta B(m(a1_2, a)) + \sum_{\xi \in \mathfrak{o}_L/\mathfrak{p}_L} \eta B(m(a1_2, a) \iota \left(\begin{bmatrix} 1 & \xi \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \right) \mu(a) |a|^{s-1} d^{\times} a.$$

For $a \in \varpi^l \mathfrak{o}^{\times}$ ($l \in \mathbb{Z}$), by the left R -equivariance and the right $K_0(\mathfrak{p})$ -invariance of B , we have

$$B(m(a1_2, a) \iota \left(\begin{bmatrix} 1 & \xi \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right)) = \psi(a \mathrm{tr}(T_{\theta} X_{\xi})) B(h(l+1, 0))$$

due to the relation $m(a1_2, a) \iota \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \eta \in h(l+1, 0) K_0(\mathfrak{p})$, which is easily confirmed by (6-2). Note that $\mathrm{tr}(T_{\theta} X_{\xi}) = 0$. Hence, the ξ -sum becomes $q^2 \times B(h(l+1, 0))$ because $\#(\mathfrak{o}_L/\mathfrak{p}_L) = q^2$. From this and the vanishing of $B(h(l, 0))$ for $l < 0$ — see (6-10) — we get (6-5). \square

By [29, Proposition 3.2], there exists an entire function $\varepsilon(\pi, s, \mu, \psi)$ such that the local functional equation

$$(6-7) \quad \varepsilon(\pi, s+1/2, \mu, \psi) \frac{Z(\phi, B, s, \mu; g)}{L(s+1/2, \pi, \mu)} = \frac{Z(\hat{\phi}, B^{\dagger}, -s, \mu^{-1}; g)}{L(-s+1/2, \hat{\pi}, \mu^{-1})}$$

holds. Here, we have defined $B^{\dagger} \in \mathcal{B}(T, \Lambda^{\dagger})[\pi]$ by setting

$$B^{\dagger}(g) := B \left(\begin{bmatrix} 1 & a^{-1}b & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -a^{-1}b & 1 \end{bmatrix} g \right) \quad (g \in \mathrm{G}(F)),$$

we have denoted by Λ^{\dagger} the Galois conjugate of Λ defined by $\Lambda^{\dagger}(\tau) := \Lambda(\tau^{\dagger})$ for

$\tau \in L^\times$, and $\hat{\phi}$ is the Fourier transform of $\phi \in \mathcal{S}(L^2)$, defined as

$$\hat{\phi}(x, y) = \int_{L^2} \phi(u, v) \psi_L\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right]\right) du dv$$

du and dv being the Haar measures on L such that $\text{vol}(\mathcal{O}_L) = 1$.

We shall get explicit evaluations of the zeta integrals $Z(\phi, B, s, \mu; g)$ for unramified Λ and particular choices of ϕ, B and g as shown in the next table, which allow us to determine $\varepsilon(s, \pi, \mu, \psi)$:

	type	ϕ	B	μ	L/F	g
(6-8)	case 1	I or IIb	$\mathbb{1}_{\mathcal{O}_L \oplus \mathcal{O}_L}$	K -fixed	unramified	$- \quad 1_4$
	case 2	I or IIb	$\mathbb{1}_{\mathfrak{p}_L^e \oplus 1 + \mathfrak{p}_L^e}$	K -fixed	ramified	inert b
	case 3	I or IIb	$\widehat{\mathbb{1}_{\mathfrak{p}_L^e \oplus 1 + \mathfrak{p}_L^e}}$	K -fixed	ramified	inert b
	case 4	I or IIb	$\mathbb{1}_{\mathcal{O}_L \oplus \mathcal{O}_L}$	$K_0(\mathfrak{p})$ -fixed	unramified	inert η
	case 5	IIIa	$\mathbb{1}_{\mathcal{O}_L \oplus \mathcal{O}_L}$	$K_0(\mathfrak{p})$ -fixed	unramified	inert η
	case 6	VIb	$\mathbb{1}_{\mathcal{O}_L \oplus \mathcal{O}_L}$	$K_0(\mathfrak{p})$ -fixed	unramified	inert η

Here, the integer e is the conductor of μ defined to be the minimum nonnegative integer n satisfying $\mu|_{(1+\mathfrak{p}^n) \cap \mathcal{O}^\times} = 1$, $b \in G(F)$ is given by

$$(6-9) \quad b := \begin{bmatrix} \varpi^e & 1_2 & T^\dagger \\ 0 & & 1_2 \end{bmatrix},$$

and L/F is said to be inert if it is an unramified field extension. For $l, m \in \mathbb{Z}$, set

$$h(\ell, m) := \begin{bmatrix} \varpi^{\ell+2m} & & & \\ & \varpi^{\ell+m} & & \\ & & 1 & \\ & & & \varpi^m \end{bmatrix} \in G(F).$$

As in [32], for a smooth representation (π, V_π) of $G(F)$, we define an operator $T_{\ell, m} \in \text{End}(V_\pi)$ by

$$T_{\ell, m} v := \text{vol}(K_0(\mathfrak{p}))^{-1} \int_{K_0(\mathfrak{p})h(\ell, m)K_0(\mathfrak{p})} \pi(g)v \, dg, \quad v \in V_\pi,$$

with dg the Haar measure on $G(F)$ such that $\text{vol}(K) = 1$, and the Atkin–Lehner involution by

$$\eta v := \pi(\eta)v, \quad \eta := \begin{bmatrix} & & & -1 \\ & & 1 & \\ & & & \varpi \\ -\varpi & & & \end{bmatrix}, \quad v \in V_\pi.$$

If π is irreducible and has the (T, Λ) -Bessel models, then Lemma 4.4(i) of [32] gives for any $B \in \mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$ the support conditions

$$(6-10) \quad B(h(\ell, m)) = B(h(\ell, m)s_2) = 0 \quad \text{if } \ell < 0, ,$$

$$(6-11) \quad B(h(\ell, m)s_1s_2) = B(h(\ell, m)s_2s_1s_2) = 0 \quad \text{if } \ell < -1.$$

6.2. Unramified computation (case 1). When π is spherical, π is of type I or IIb in our cases. Recall that there exists a unique element $B_\pi^0 \in \mathcal{B}(T, \Lambda)[\pi]^K$ such that $B_\pi^0(1_4) = 1$. Then, the following formula for an unramified μ in its greatest generality is due to Sugano [42, Theorem 2.1]:

$$Z(\phi, B_\pi^0, s, \mu; 1_4) = L(s + 1/2, \pi, \mu), \quad \mathrm{Re}(s) \gg 0.$$

In [20, Proposition 5.9], a proof of this formula is given when $L = F \oplus F$ and π is of type I based on the explicit formula of B_π^0 due to [6].

6.3. Unramified computation (cases 2 and 3). Now we assume that $e > 0$ and that $T \in \mathrm{Sym}^2(F)$ satisfies the following conditions:

- L/F is an unramified field extension.
- $T \in \mathbf{GL}_2(\mathcal{O})$, $-\frac{d}{2} = \mathrm{tr}(TT^\dagger) \in \mathcal{O}^\times$.

Then ϖ remains a prime element in \mathfrak{p}_L . Let b be as in (6-9), and set $\phi = \mathbb{1}_{\mathfrak{p}_L^e \oplus 1 + \mathfrak{p}_L^e} \in \mathcal{S}(L^2)$. Our goal in this subsection is to calculate the zeta integrals $Z(\phi, B_\pi^0, s, \mu; b)$ and $Z(\hat{\phi}, B_\pi^0, s, \mu; b)$ explicitly. We note that $Z(\phi, B_\pi^0, s, \mu; 1_4)$ must be zero because the a -integral in (6-4) vanishes.

Definition 6.2 (root number). Let ψ and μ be as above, we define the Gauss sum $W_F(\mu, \psi)$ by

$$W_F(\mu, \psi) := q^{-\frac{e}{2}} \mu(\varpi)^{-e} \sum_{\alpha \in (\mathcal{O}/\mathfrak{p}^e)^\times} \psi(\varpi^{-e}\alpha) \mu(\alpha).$$

This sum has absolute value 1 and is independent of the choice of ϖ . The following lemma is more or less well-known [44, (2.18)].

Lemma 6.3. *Let ψ and μ be as above. For $n \in \mathbb{Z}$,*

$$\int_{\mathcal{O}^\times} \psi(\varpi^n a) \mu(a) d^\times a = \begin{cases} q^{-\frac{e}{2}+1} (q-1)^{-1} \mu(\varpi)^e W_F(\mu, \psi) & (n = -e), \\ 0 & (n \neq -e). \end{cases}$$

Lemma 6.4. *Let L/F be an unramified quadratic extension. For any character μ of F with $e := \mathrm{cond}(\mu) > 0$,*

$$W_L(\mu_L, \psi_L) = (-1)^e W_F(\mu, \psi)^2.$$

Proof. For a virtual representation V of the Weil group \mathfrak{W}_F of F , let $\varepsilon(s, V, \psi)$ denote the local epsilon factor à la Deligne–Langlands [45]. By [45, (3.4.8)], we have

$$(6-12) \quad \varepsilon(s, \mathrm{Ind}_{\mathfrak{W}_L}^{\mathfrak{W}_F} \rho, \psi) = \varepsilon(s, \rho, \psi_L)$$

for any ρ of degree 0. Any character μ of F^\times is viewed as a one-dimensional representation of \mathfrak{W}_F by $F^\times \cong \mathfrak{W}_F^{\mathrm{ab}}$. If $\mu_L := \mu \circ N_{L/F}$, then $\mathrm{Ind}_{\mathfrak{W}_L}^{\mathfrak{W}_F} \mu_L \cong \mu \oplus \mu \eta_{L/F}$. If μ is unramified, then $\varepsilon(s, \mu, \psi) = 1$; since L/F is unramified,

so is the character $\eta_{L/F}$. From these observations, we see that equality (6-12) holds true when $\rho = \mathbf{1}$. Thus, (6-12) is true when $\rho = \mu$ with $e > 0$. Since $\varepsilon(s, \mu\eta_{L/F}^j, \psi) = q^{-e(s-1/2)}W_F(\bar{\mu}\eta_{L/F}^j, \psi)$ for $j = 0, 1$ and $\varepsilon(s, \mu_L, \psi_L) = q^{-2e(s-1/2)}W_L(\bar{\mu}_L, \psi_L)$, we obtain $W_L(\mu_L, \psi_L) = W_F(\mu, \psi)W_F(\mu\eta_{L/F}, \psi)$. Since $e > 0$ and $\eta_{L/F}(\varpi) = -1$, we have

$$W_F(\mu\eta_{L/F}, \psi) = \eta_{L/F}(\varpi)^e W_F(\mu, \psi) = (-1)^e W_F(\mu, \psi). \quad \square$$

Lemma 6.5. For $k^\# = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in K^\#$, we have

$$f_\phi^{(s, \Lambda, \mu)}(k^\#) = \begin{cases} q^{-2e+2}(q^2-1)^{-1}\mu_L(d)^{-1}\mu(\det k^\#) & (c \in \mathfrak{p}_L^e), \\ 0 & (\text{otherwise}), \end{cases}$$

$$f_{\hat{\phi}}^{(s, \Lambda, \mu)}(k^\#) = \begin{cases} q^{e(2s-3)+2}(q^2-1)^{-1}\Lambda(\varpi)^{-e}\mu_L(c)^{-1}W_L(\mu_L, \psi_L) & (c \in \mathcal{O}_L^\times), \\ 0 & (c \notin \mathcal{O}_L^\times). \end{cases}$$

Proof. These equations immediately follow by Lemma 6.3. □

Lemma 6.6. For $u \in \mathcal{O}^\times$,

$$\sum_{\substack{\eta \in \mathcal{O}_L/\mathfrak{p}_L^e \\ u + N_{L/F}(\eta) \in \mathcal{O}^\times}} \mu(u + N_{L/F}(\eta)) = (-1)^e q^e \mu(u).$$

Proof. Because L/F is an unramified extension, the norm map $N_{L/F} : \mathcal{O}_L^\times \rightarrow \mathcal{O}^\times$ is surjective. Thus, there exists $v \in \mathcal{O}_L^\times$ such that $u = N_{L/F}(v)$. By replacing η by $v\eta$, without loss of generality, we may assume that $u = 1$.

For an integer i with $0 \leq i \leq e + 1$, we set

$$\overline{\mathcal{O}_{L,i}} = \begin{cases} \{\eta \in (\mathcal{O}_L/\mathfrak{p}_L^e)^\times \mid 1 + N_{L/F}(\eta) \not\equiv 0 \pmod{\mathfrak{p}}\} & (i = 0), \\ \mathfrak{p}_L^i/\mathfrak{p}_L^e & (0 < i \leq e), \\ \emptyset & (i = e + 1) \end{cases}$$

and

$$\overline{\mathcal{U}_i} = \begin{cases} \mathcal{O}^\times/(1 + \mathfrak{p}^e) & (i = 0), \\ (1 + \mathfrak{p}^i)/(1 + \mathfrak{p}^e) & (0 < i \leq e), \\ \emptyset & (i = e + 1). \end{cases}$$

When $0 \leq i < e/2$, we can check that

$$\overline{\mathcal{O}_{L,i}} \setminus \overline{\mathcal{O}_{L,i+1}} \ni \eta \pmod{\mathfrak{p}_L^e} \mapsto 1 + N_{L/F}(\eta) \pmod{\mathfrak{p}^e} \in \overline{\mathcal{U}_{2i}} \setminus \overline{\mathcal{U}_{2i+1}}$$

is a surjective map having the fiber of cardinality $q^{e-1}(q + 1)$ at each point.

When $e/2 \leq i \leq e$, it is immediate that

$$\mu(1 + N_{L/F}(\eta)) = 1, \quad \eta \in \overline{\mathcal{O}_{L,i}}.$$

Hence

$$\begin{aligned} & \sum_{\substack{\eta \in \mathcal{O}_L/\mathfrak{p}_L^e \\ 1+N_{L/F}(\eta) \in \mathcal{O}^\times}} \mu(1+N_{L/F}(\eta)) \\ &= \sum_{i=0}^e \sum_{\eta \in \overline{\mathcal{O}_{L,i}} \setminus \overline{\mathcal{O}_{L,i+1}}} \mu(1+N_{L/F}(\eta)) \\ &= q^{e-1}(q+1) \sum_{0 \leq i < e/2} \sum_{\varepsilon \in \overline{\mathcal{U}_{2i}} \setminus \overline{\mathcal{U}_{2i+1}}} \mu(\varepsilon) + \sum_{e/2 \leq i \leq e} \#(\overline{\mathcal{O}_{L,i}} \setminus \overline{\mathcal{O}_{L,i+1}}). \end{aligned}$$

Then the lemma follows from

$$\sum_{\varepsilon \in \overline{\mathcal{U}_{2i}} \setminus \overline{\mathcal{U}_{2i+1}}} \mu(\varepsilon) = \begin{cases} 0 & (0 \leq i < \frac{e-1}{2}) \\ -1 & (i = \frac{e-1}{2}) \end{cases} \quad \text{and} \quad \sum_{e/2 \leq i \leq e} \#(\overline{\mathcal{O}_{L,i}} \setminus \overline{\mathcal{O}_{L,i+1}}) = q^{2e-2\lceil e/2 \rceil},$$

where $\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$ is the ceiling function. □

Lemma 6.7. *Let $B \in \mathcal{B}(T, \Lambda)[\pi]^K$ and $\eta \in \mathcal{O}$ with $\frac{1}{4}a^6d + N_{L/F}(\eta) \in \varpi^j \mathcal{O}^\times$ ($0 \leq j \leq e$).*

(i) *If we set $Y_\eta = -a^2T^\dagger + X_\eta$, then we have*

$$\det Y_\eta = -\frac{1}{4}a^4d - a^{-2}N_{L/F}(\eta).$$

(ii) *There exist $\tau \in \mathcal{O}_L^\times$ and $A \in \mathbf{GL}_2(\mathcal{O})$ such that*

$$Y_\eta = \tau \begin{bmatrix} \varpi^j & 0 \\ 0 & 1 \end{bmatrix} A.$$

(iii) *For $n \in \mathbb{Z}$ and $a \in \mathcal{O}^\times$,*

$$\begin{aligned} & B(m(\varpi^n a 1_2, \varpi^n a) \iota \left(\begin{bmatrix} 0 & -1 \\ 1 & \eta^\dagger \end{bmatrix} \right) b) \\ &= \psi(-\varpi^n a \cdot \frac{1}{2}a^4d(\frac{1}{4}a^6d + N_{L/F}(\eta))^{-1}) \cdot B(h(e+n-2j, j)). \end{aligned}$$

Proof. (i) Noting that $\det T = -d/4$, $\det X_\eta = -a^{-2}N_{L/F}(\eta)$, and $\mathrm{tr}(TX_\eta) = \mathrm{tr}(X_\eta T^\dagger) = 0$, we obtain

$$\begin{aligned} \det Y_\eta &= \frac{1}{2} \mathrm{tr}(Y_\eta Y_\eta^\dagger) = \frac{1}{2} \mathrm{tr}(a^4 T T^\dagger + X_\eta^\dagger X_\eta) \\ &= a^4 \det T + \det(X_\eta) = -\frac{1}{4}a^4d - a^{-2}N_{L/F}(\eta). \end{aligned}$$

(ii) By (i) and the assumption that $\frac{1}{4}a^6d + N_{L/F}(\eta) \in \varpi^j \mathcal{O}^\times$, we have $Y_\eta \in \mathbf{GL}_2(F)$. The disjoint union $\mathbf{GL}_2(F) = \bigsqcup_{m \geq 0} L^\times \begin{bmatrix} \varpi^m & 0 \\ 0 & 1 \end{bmatrix} \mathbf{GL}_2(\mathcal{O})$ (see [42, Lemma 2–4]) implies that there exist $m \in \mathbb{Z}_{\geq 0}$, $\ell \in \mathbb{Z}$, $\tau \in \mathcal{O}_L^\times$, and $A \in \mathbf{GL}_2(\mathcal{O})$ such that

$$Y_\eta = (\varpi^\ell \tau) \begin{bmatrix} \varpi^m & 0 \\ 0 & 1 \end{bmatrix} A = \tau \begin{bmatrix} \varpi^{m+\ell} & 0 \\ 0 & \varpi^\ell \end{bmatrix} A.$$

We now prove that $m = j$ and $\ell = 0$. Because $\tau \in \mathbf{GL}_2(\mathcal{O})$, the Smith normal form of Y_η is $\begin{bmatrix} \varpi^{m+\ell} & 0 \\ 0 & \varpi^\ell \end{bmatrix}$. Hence, by the theory of the Smith normal form, we only have to show that the elements of Y_η are coprime. This follows from the identity $\mathrm{tr}(Y_\eta T) = \frac{1}{2}a^2d$ and the assumption $\frac{1}{2}d \in \mathcal{O}^\times$.

(iii) By using (6-1), we can check that

$$\iota\left(\begin{bmatrix} 1 & 0 \\ -\eta^\dagger & 1 \end{bmatrix}\right) = \begin{bmatrix} 1_2 & 0 \\ X_\eta^\dagger & 1_2 \end{bmatrix}, \quad \iota\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right) b \iota\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1_2 & 0 \\ -a^2 T & \varpi^e 1_2 \end{bmatrix}.$$

Hence we have

$$m(\varpi^n a 1_2, \varpi^n a) \iota\left(\begin{bmatrix} 0 & -1 \\ 1 & \eta^\dagger \end{bmatrix}\right) b = \begin{bmatrix} \varpi^n a 1_2 & 0 \\ Y_\eta^\dagger & \varpi^e 1_2 \end{bmatrix} \iota\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right).$$

By (ii), we may choose elements $\tau \in \mathcal{O}_L^\times$ and $A \in \mathbf{GL}_2(\mathcal{O})$ with $Y_\eta = \tau \begin{bmatrix} \varpi^j & 0 \\ 0 & 1 \end{bmatrix} A$. Then, a computation shows that

$$\begin{bmatrix} \varpi^n a 1_2 & 0 \\ Y_\eta^\dagger & \varpi^e 1_2 \end{bmatrix} = n(\varpi^n a (\det Y_\eta)^{-1} Y_\eta) \begin{bmatrix} \tau & 0 \\ 0 & \tau^\dagger \end{bmatrix} h(e + n - 2j, j) k,$$

where $k = \begin{bmatrix} A & 0 \\ 0 & {}^t A^\dagger \end{bmatrix} \begin{bmatrix} a \varpi^j (\det Y_\eta)^{-1} 1_2 & 0 \\ 0 & 1_2 \end{bmatrix} \begin{bmatrix} 0 & -1_2 \\ 1_2 & \varpi^e (\det Y_\eta)^{-1} Y_\eta \end{bmatrix} \in K$. Since Λ is unramified and B is K -invariant, we have the desired equality. \square

Proposition 6.8. *Suppose that π is of type I or IIb. Let $\mu : F^\times \rightarrow \mathbb{C}^1$ and $\Lambda : L^\times \rightarrow \mathbb{C}^1$ be characters such that $\text{cond}(\mu) = e > 0$ and Λ is unramified. Suppose that $T = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \in \text{Sym}^2(F) \cap \mathbf{GL}_2(\mathcal{O})$ satisfies that L/F is an unramified field extension and $2d \in \mathcal{O}^\times$. When $\phi = 1_{\mathfrak{p}_L^e} \oplus 1_{1+\mathfrak{p}_L^e}$ and $\text{Re}(s) > 1$, we have*

$$\begin{aligned} Z(\phi, B_\pi^0, s, \mu; b) &= q^{e(s-\frac{11}{2})+5} (q^4 - 1)^{-1} (q - 1)^{-1} \mu(-2^{-1}d)^{-1} W_F(\mu, \psi), \\ Z(\hat{\phi}, B_\pi^0, s, \mu; b) &= (-1)^e q^{e(3s-\frac{11}{2})+5} (q^4 - 1)^{-1} (q - 1)^{-1} \Lambda(\varpi)^{-e} \mu(-2^{-1}a^2) \\ &\quad \times W_L(\mu_L, \psi_L) W_F(\mu, \psi). \end{aligned}$$

Proof. Let $K^\#(\mathfrak{p}^e) = \{k^\# \in K^\# \mid k^\# \equiv 1_2 \pmod{\mathfrak{p}_L^e}\}$ be the principal congruence subgroup of level \mathfrak{p}_L^e . If $\phi \in \mathcal{S}(L^2)$ is right $K^\#(\mathfrak{p}^e)$ -invariant, the function

$$K^\# \ni k^\# \mapsto f_\phi^{(s, \Lambda, \mu)}(k^\#) \int_{F^\times} B_\pi^0(m(a 1_2, a) \iota(k^\#) b) \mu(a) |a|^{s-1} \in \mathbb{C}$$

is left- $B^\#(\mathcal{O})$ -invariant and right $K^\#(\mathfrak{p}^e)$ -invariant because $b^{-1} \iota(K^\#(\mathfrak{p}^e)) b \subset K$. Since $K^\#(\mathfrak{p}^e)$ is a normal subgroup of $K^\#$ and $B^\#(\mathcal{O}) K^\#(\mathfrak{p}^e) = K_0^\#(\mathfrak{p}^e)$, we have

$$\begin{aligned} (6-13) \quad Z(\phi, B_\pi^0, s, \mu; b) &= \sum_{[\gamma] \in B^\#(\mathcal{O}) \backslash K^\# / K^\#(\mathfrak{p}^e)} \text{vol}(B^\#(\mathcal{O}) \gamma K^\#(\mathfrak{p}^e)) f_\phi^{(s, \Lambda, \mu)}(\gamma) \Upsilon^{(s, \mu)}(\gamma) \\ &= [K^\# : K_0^\#(\mathfrak{p}^e)]^{-1} \sum_{[\gamma] \in K_0^\#(\mathfrak{p}^e) \backslash K^\#} f_\phi^{(s, \Lambda, \mu)}(\gamma) \Upsilon^{(s, \mu)}(\gamma) \end{aligned}$$

where we set $\Upsilon^{(s, \mu)}(\gamma) := \int_{F^\times} B_\pi^0(m(a 1_2, a) \iota(\gamma) b) \mu(a) |a|^{s-1} d^\times a$ to make room.

A complete set of representatives for $K^\#(\mathfrak{p}^e) \backslash K^\#$ is given by

$$(6-14) \quad \begin{bmatrix} 1 & 0 \\ \xi & 1 \end{bmatrix} \quad (\xi \in \mathfrak{p}_L / \mathfrak{p}_L^e), \quad \begin{bmatrix} 0 & -1 \\ 1 & \eta^\dagger \end{bmatrix} \quad (\eta \in \mathcal{O}_L / \mathfrak{p}_L^e).$$

Now we prove the first equation. By Lemma 6.5, when γ runs through the above elements, we have $f_\phi^{(s, \Lambda, \mu)}(1_2) = q^{-2e+2} (q^2 - 1)^{-1}$ and $f_\phi^{(s, \Lambda, \mu)}(\gamma) = 0$ if $\gamma \neq 1$.

By substituting this in (6-13), noting that $[K^\# : K_0^\#(\mathfrak{p}^e)] = q^{2e-2}(q^2 + 1)$, and Lemma 6.3, we have

$$\begin{aligned}
 Z(\phi, B_\pi^0, s, \mu; b) &= q^{-4e+4}(q^4 - 1)^{-1} \int_{F^\times} B_\pi^0 \left(\begin{bmatrix} a\varpi^e & 1_2 & aT^\dagger \\ & & 1_2 \end{bmatrix} \right) \mu(a) |a|^{s-1} d^\times a \\
 &= q^{-4e+4}(q^4 - 1)^{-1} \int_{F^\times} B_\pi^0 \left(\begin{bmatrix} a\varpi^e & 1_2 & 0 \\ & & 1_2 \end{bmatrix} \right) \psi(a \cdot \mathrm{tr}(T T^\dagger)) \mu(a) |a|^{s-1} d^\times a \\
 &= q^{-4e+4}(q^4 - 1)^{-1} \mu(-2^{-1}d)^{-1} \\
 &\quad \times \sum_{n \in \mathbb{Z}} B_\pi^0 \left(\begin{bmatrix} \varpi^{e+n} & 1_2 & 0 \\ & & 1_2 \end{bmatrix} \right) \mu(\varpi)^n q^{-n(s-1)} \int_{\mathcal{O}^\times} \psi(\varpi^n a) \mu(a) d^\times a \\
 &= q^{e(s-\frac{11}{2})+5} (q^4 - 1)^{-1} (q - 1)^{-1} \mu(-2^{-1}d)^{-1} W_F(\mu, \psi).
 \end{aligned}$$

Next, we prove the second equation. When we replace ϕ by $\hat{\phi}$ in (6-13) and consider the representative (6-14), we can ignore the contribution of ξ -terms in the summation from Lemma 6.5. Hence, by noting the K -invariance of B , we have

$$\begin{aligned}
 (6-15) \quad Z(\hat{\phi}, B_\pi^0, s, \mu; b) &= q^{e(2s-5)+4} (q^4 - 1)^{-1} \Lambda(\varpi)^{-e} W_L(\mu_L, \psi_L) \\
 &\quad \times \sum_{\eta \in \mathcal{O}_L/\mathfrak{p}_L^e} \int_{F^\times} B_\pi^0(m(a)1_2, a) \iota \left(\begin{bmatrix} 0 & -1 \\ 1 & \eta^\dagger \end{bmatrix} \right) b \mu(a) |a|^{s-1} d^\times a.
 \end{aligned}$$

By replacing η by $\eta + \varpi^e u$ for some $u \in \mathcal{O}_L$ if we need, we may assume that $a^6 d + N_{L/F}(\eta) \in \varpi^j \mathcal{O}_L^\times$ with $0 \leq j \leq e$. Then, Lemmas 6.3 and 6.7(iii) imply

$$\begin{aligned}
 &\int_{F^\times} B_\pi^0(m(a)1_2, a) \iota \left(\begin{bmatrix} 0 & -1 \\ 1 & \eta^\dagger \end{bmatrix} \right) b \mu(a) |a|^{s-1} d^\times a \\
 &= \sum_{n \in \mathbb{Z}} B_\pi^0(h(e+n-2j, j)) \mu(\varpi)^n q^{-n(s-1)} \\
 &\quad \times \int_{\mathcal{O}^\times} \psi(-\varpi^n a \cdot \frac{1}{2} a^4 d (\frac{1}{4} a^6 d + N_{L/F}(\eta))^{-1}) \mu(a) d^\times a \\
 &= \mu(-2a^{-4} d^{-1} \varpi^{-j} (\frac{1}{4} a^6 d + N_{L/F}(\eta))) \\
 &\quad \times \sum_{n \in \mathbb{Z}} B_\pi^0(h(e+n-2j, j)) \mu(\varpi)^n q^{-n(s-1)} \int_{\mathcal{O}^\times} \psi(\varpi^{n-j} a) \mu(a) d^\times a \\
 &= q^{-\frac{e}{2}+1+(e-j)(s-1)} (q-1)^{-1} \mu(-2a^{-4} d^{-1}) \mu(\frac{1}{4} a^6 d + N_{L/F}(\eta)) \\
 &\quad W_F(\mu, \psi) \cdot B_\pi^0(h(-j, j)).
 \end{aligned}$$

The last expression is equal to 0 unless $j = 0$ from (6-10). By substituting this into (6-15) and Lemma 6.6, we have that $Z(\phi, B_\pi^0, s, \mu; b)$ equals

$$\begin{aligned}
 &q^{e(3s-\frac{13}{2})+5} (q^4 - 1)^{-1} (q - 1)^{-1} \Lambda(\varpi)^{-e} \mu(-2a^{-4} d^{-1}) \\
 &\quad \times W_L(\mu_L, \psi_L) W_F(\mu, \psi) \sum_{\substack{\eta \in \mathcal{O}_L/\mathfrak{p}_L^e \\ a^6 d + N_{L/F}(\eta) \in \mathcal{O}^\times}} \mu(\frac{1}{4} a^6 d + N_{L/F}(\eta)) \\
 &= (-1)^e q^{e(3s-\frac{11}{2})+5} (q^4 - 1)^{-1} (q - 1)^{-1} \Lambda(\varpi)^{-e} \mu(-2^{-1} a^2) W_L(\mu_L, \psi_L) W_F(\mu, \psi).
 \end{aligned}$$

□

Corollary 6.9. *Let $\pi, T, \mu,$ and Λ be as in Proposition 6.8. Then*

$$\begin{aligned} \varepsilon(\pi, s, \mu, \psi) &= (-1)^e q^{4e(\frac{1}{2}-s)} \Lambda(\varpi)^{-e} \mu(-a^{-2}d) \overline{W_L(\mu_L, \psi_L) W_F(\mu, \psi)^2} \\ &= q^{4e(\frac{1}{2}-s)} \Lambda(\varpi)^{-e} \mu(-a^{-2}d) \overline{W_F(\mu, \psi)^4}. \end{aligned}$$

Proof. The first equality follows from Proposition 6.8 and (6-7); the second equality is due to Lemma 6.4. □

6.4. Computation for old forms of type I and IIb (case 4). Now we assume that π has the trivial central character. Then, π has the (T, Λ) -Bessel model if and only if $\Lambda = 1$. Recall that the dimension of $V_\pi^{K_0(\mathfrak{p})}$ is equal to 4 or 3 according as π is of type I or type IIb.

Proposition 6.10. *Let π be a smooth admissible irreducible representation of type I or type IIb with trivial central character. Suppose that $\mu : F^\times \rightarrow \mathbb{C}^1$ is an unramified character. When $\phi = \mathbb{1}_{\mathcal{O}_L \oplus \mathcal{O}_L}$ and $\text{Re}(s) > 1$, we have*

$$\begin{aligned} Z(\phi, B, s, \mu; \eta) &= \frac{1}{q^2+1} L(s+1/2, \pi, \mu) \\ &\times [\eta B + q^{-1} T_{1,0} B + (\mu(\varpi)^{-1} q^{s+1} + \mu(\varpi) q^{-s+1} - \text{tr}((q^{-1} T_{1,0} + \eta)|_{V_\pi^{K_0(\mathfrak{p})}})) B] (1_4) \end{aligned}$$

for $B \in \mathcal{B}(T, 1)[\pi]^{K_0(\mathfrak{p})}$.

Proof. We will give a proof of the case that π is of type I. In the case of type IIb, the proof is similar. Let $\{B_j^*\}_{1 \leq j \leq 4} \subset \mathcal{B}(T, 1)[\pi]^{K_0(\mathfrak{p})}$ be an eigenbasis of $T_{1,0}$ with eigenvalues $\{\lambda_j\}_{1 \leq j \leq 4}$. Then, we suffice to check the desired equation only for $B = B_j^*$ ($1 \leq j \leq 4$). Let $(\eta_{ij})_{1 \leq i, j \leq 4}$ be the representation matrix of $\eta|_{K_0(\mathfrak{p})}$ with respect to $\{B_j^*\}_{1 \leq j \leq 4}$. Then, (6-21) shows that

$$B_j^*(h(l+1, 0)) = \lambda_j q^{-3} B_j^*(h(l, 0)), \quad l \in \mathbb{Z}_{\geq 0}.$$

By applying Lemma 6.1 to $B = B_j^*$, we obtain

$$Z(\phi, B_j^*, s, \mu; \eta) = \frac{L(s+1, \mu_L)}{q^2+1} \left(\sum_{i=1}^4 \frac{\eta_{ij} B_i^*(1_4)}{1-\lambda_i \mu(\varpi) q^{-(s+2)}} + \frac{\mu(\varpi)^{-1} q^{s+1} B_j^*(1_4)}{1-\lambda_j \mu(\varpi) q^{-(s+2)}} \right).$$

Thus, we may put

$$Z(\phi, B_j^*, s, \mu; \eta) = \frac{1}{q^2+1} \prod_{i=1}^4 (1-\lambda_i q^{-2} X)^{-1} \frac{A_0 + A_1 X + A_2 X^2 + A_3 X^3 + A_4 X^4}{X(1-q^{-2} X^2)}$$

with $X = \mu(\varpi) q^{-s}$ for some $A_i \in \mathbb{C}$ ($0 \leq i \leq 4$).

By noting that $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{\alpha\beta\gamma q^{3/2}, \beta\gamma q^{3/2}, \alpha\gamma q^{3/2}, \gamma q^{3/2}\}$ from [34, Table A.8], we have

$$(6-16) \quad (q^2+1) \cdot \frac{Z(\phi, B_j^*, s, \mu; 1_4)}{L(s+1/2, \pi, \mu)} = \frac{A_0 + A_1 X + A_2 X^2 + A_3 X^3 + A_4 X^4}{X(1-q^{-2} X^2)}.$$

A direct computation shows that

$$(6-17) \quad A_0 = qB_j^*(1_4),$$

$$(6-18) \quad A_1 = \eta B_j^*(1_4) + q^{-1} \mathrm{tr}(T_{1,0}|_{V_\pi^{K_0(\mathfrak{p})}}) \cdot B_j^*(1_4) - q^{-1} \cdot T_{1,0} B_j^*(1_4).$$

The functional equation (6-7) and the relation $\varepsilon(\pi, s, \mu, \psi) = 1$ (see Theorem 4.4 in [29]) show that the right side of (6-16) is invariant under the transformation $X \rightarrow X^{-1}$. By comparing coefficients, we have

$$(6-19) \quad A_2 = q^{-2}(q^2 - 1)A_0, \quad A_3 = -q^{-2}A_1, \quad A_4 = -q^{-2}A_0.$$

Substituting (6-17), (6-18), and (6-19) into (6-16), the right side of (6-16) becomes

$$[\eta B_j^* + q^{-1} T_{1,0} B_j^* + (\mu(\varpi)^{-1} q^{s+1} + \mu(\varpi) q^{-s+1} - \mathrm{tr}(q^{-1} T_{1,0}|_{V_\pi^{K_0(\mathfrak{p})}})) B_j^*](1_4). \quad \square$$

6.5. Computation for newforms of type IIIa (case 5). In this subsection, we do not assume the triviality of the central character. By [32, §9.1], there exists a basis $\{B_1, B_2\}$ of $\mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$ such that

$$(6-20) \quad \begin{aligned} T_{1,0} B_1 &= \alpha \gamma q B_1, & T_{1,0} B_2 &= \gamma q B_2, \\ T_{0,1} B_1 &= \alpha \gamma^2 (\alpha q + 1) q B_1, & T_{0,1} B_2 &= \alpha \gamma^2 (\alpha^{-1} q + 1) q B_2, \\ \eta B_1 &= \alpha \gamma B_2, & \eta B_2 &= \gamma B_1. \end{aligned}$$

We now consider the values of B_1 and B_2 at the identity element. Lemma 9.1 of [32] ensures their nonvanishing and $\Lambda(\varpi) = \alpha \gamma^2$ by the condition of central character of π .

Lemma 6.11. *Let $T \in \mathrm{Sym}^2(F)$ such that L/F is an unramified field extension. Suppose that $\Lambda : L^\times \rightarrow \mathbb{C}^1$ is an unramified character. Then*

$$B_2(1_4) = \alpha^{-1} B_1(1_4) \neq 0.$$

Proof. For $B \in \mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$ and any nonnegative integer l , by Lemmas 5.1 and 5.3 in [32], we get the following relations:

$$(6-21) \quad T_{1,0} B(h(l, 0)) = q^3 B(h(l+1, 0)),$$

$$(6-22) \quad T_{1,0} B(h(l, 0)_{s_2}) = q^2(q-1)B(h(l+1, 0)) + q^2 B(h(l-1, 1)_{s_1 s_2}),$$

$$(6-23) \quad T_{1,0} B(h(l, 0)_{s_2 s_1 s_2}) = q^2(q-1)B(h(l+1, 0)) + \Lambda(\varpi) B(h(l-1, 0)_{s_2 s_1 s_2}) + (q^2 - 1) B(h(l-1, 1)_{s_1 s_2}),$$

$$(6-24) \quad T_{0,1} B(h(l, 0)) = q^3(q+1)B(h(l, 1)),$$

$$(6-25) \quad T_{0,1} B(h(l, 0)_{s_2}) = q^4 B(h(l, 1)_{s_1 s_2}) + q \Lambda(\varpi) B(h(l-2, 1)_{s_1 s_2}) + q^2(q-1)B(h(l, 1)) + \begin{cases} -q \Lambda(\varpi) B(1_4) & (l = 0), \\ q(q-1) \Lambda(\varpi) B(h(l, 0)) & (l \geq 1). \end{cases}$$

Set $B = B_1$. By putting $l = 0$ in the above relations, applying the relation (6-20), and using the equation that $B_2(h(l, 0)) = \gamma B_1(h(l-1, 0)s_2s_1s_2)$ for $l \geq 0$, we have

$$(6-26) \quad \alpha\gamma q B_1(s_2) = \alpha\gamma(q-1)B_1(1_4) + q^2 B_1(h(-1, 1)s_1s_2),$$

$$(6-27) \quad \Lambda(\varpi)\gamma^{-1}(q^{-1}-1)B_2(1_4) = \alpha\gamma(q-1)B_1(1_4) + (q^2-1)B_1(h(-1, 1)s_1s_2),$$

$$(6-28) \quad \Lambda(\varpi)(\alpha q + 1)q B_1(s_2) = q^4 B_1(h(0, 1)s_1s_2) + \frac{\Lambda(\varpi)(\alpha q^2 - \alpha q - q^2 - 1)}{q+1} B_1(1_4).$$

Similarly, by putting $l = 1$ in (6-23), we have

$$\Lambda(\varpi)q^{-3}(1-q)B_2(1_4) = \Lambda(\varpi)\alpha q^{-2}(q-1)B_1(1_4) + (q^2-1)B_1(h(0, 1)s_1s_2).$$

From this last equation, together with (6-26), (6-27), (6-28), we obtain $B_2(1_4) = \alpha^{-1} B_1(1_4)$. The nonvanishing of B_1 at 1_4 follows from Theorem 9.3 of [32]. \square

Proposition 6.12. *Suppose that π is of type IIIa. Let $T \in \text{Sym}^2(F)$ be such that L/F is an unramified field extension. Suppose that $\mu : F^\times \rightarrow \mathbb{C}^1$ and $\Lambda : L^\times \rightarrow \mathbb{C}^1$ are unramified characters. When $\phi = 1_{\sigma_L \oplus \sigma_L}$ and $\text{Re}(s) > 1$, we have*

$$(6-29) \quad Z(\phi, B, s, \mu; \eta) = \frac{\Lambda(\varpi)^{-1} \mu(\varpi)^{-1} q^{s+1}}{q^2 + 1} L(s + 1/2, \pi, \mu) \cdot B(1_4),$$

$$B \in \mathcal{B}(s, \Lambda)[\pi]^{K_0(\mathfrak{p})}.$$

Proof. By substituting $B = B_i$ in (6-5) for $i = 1, 2$, the statement is immediate for $B = B_i$ by (6-20), (6-21), and Lemma 6.11. Since B_1 and B_2 span $\mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$, we complete the proof. \square

Corollary 6.13. *Let π, T, μ , and Λ be as in Proposition 6.12, then*

$$\varepsilon(\pi, s, \mu, \psi) = \mu(\varpi)^2 q^{2(\frac{1}{2}-s)}.$$

6.6. Computation for newforms of type VIb (case 6). Suppose that π is of type VIb; we do not assume the triviality of the central character. By [32, Table 3], any element $B \in \mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$ satisfies

$$(6-30) \quad T_{1,0}B = \gamma q B, \quad \eta B = \gamma B.$$

By [32, Theorem 9.3], we have $B(1_4) \neq 0$ if and only if $B \neq 0$.

Proposition 6.14. *Suppose that π is of type VIb. Let $T \in \text{Sym}^2(F)$ be such that L/F is an unramified field extension. Suppose that $\mu : F^\times \rightarrow \mathbb{C}^1$ and $\Lambda : L^\times \rightarrow \mathbb{C}^1$ are unramified characters. When $\phi = 1_{\sigma_L \oplus \sigma_L}$ and $\text{Re}(s) > 1$, we have*

$$Z(\phi, B, s, \mu; \eta) = \frac{\Lambda(\varpi)^{-1} \mu(\varpi)^{-1} q^{s+1}}{q^2 + 1} L(s + 1/2, \pi, \mu) \cdot B(1_4),$$

$$B \in \mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}.$$

Proof. Similar to that of Proposition 6.12, using relations (6-21)–(6-25). \square

Corollary 6.15. *Let $\pi, T, \mu,$ and Λ be as in Proposition 6.14. Then*

$$\varepsilon(\pi, s, \mu, \psi) = \mu(\varpi)^2 q^{2(\frac{1}{2}-s)}.$$

6.7. The values of Bessel models at $\mathbf{1}_4$ for old forms. For quasicharacters χ, χ' , and σ on F^\times which are trivial on \mathcal{O}^\times , the representation $(\chi \times \chi' \rtimes \sigma, V)$ can be realized as the right regular representation of the space of all smooth functions f on $G(F)$ satisfying

$$f \left(\begin{bmatrix} a & * & * & * \\ 0 & b & * & * \\ 0 & 0 & ca^{-1} & 0 \\ 0 & 0 & * & cb^{-1} \end{bmatrix} g \right) = |a^2 b| |c|^{-\frac{3}{2}} \chi(a) \chi'(b) \sigma(c) f(g), \quad a, b, c \in F^\times, g \in G(F).$$

The space V has a unique K -invariant element $f_K \in V$ such that $f_K(\mathbf{1}_4) = 1$. Let

$$I = K \cap \begin{bmatrix} \mathcal{O} & \mathfrak{p} & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \mathcal{O} \\ \mathfrak{p} & \mathfrak{p} & \mathcal{O} & \mathcal{O} \\ \mathfrak{p} & \mathfrak{p} & \mathfrak{p} & \mathcal{O} \end{bmatrix}$$

be the Iwahori subgroup of K . For $w \in W$, let $f_w \in V^I$ be the unique element such that $f_w|_K = \mathbb{1}_{IwI}$.

By [31, Section 2], the space V^I is 8-dimensional space spanned by

$$f_{1_4}, \quad f_{s_1}, \quad f_{s_2}, \quad f_{s_2 s_1}, \quad f_{s_1 s_2 s_1}, \quad f_{s_1 s_2}, \quad f_{s_1 s_2 s_1 s_2}, \quad f_{s_2 s_1 s_2}.$$

A basis of the 4-dimensional space $V^{K_0(\mathfrak{p})}$ is given by

$$f_1^I := f_{1_4} + f_{s_1}, \quad f_2^I := f_{s_2} + f_{s_2 s_1}, \quad f_3^I := f_{s_1 s_2 s_1} + f_{s_1 s_2}, \quad f_4^I := f_{s_1 s_2 s_1 s_2} + f_{s_2 s_1 s_2}.$$

Recall that π is isomorphic to $\chi \times \chi' \rtimes \sigma$ or $\chi \mathbf{1}_{\mathrm{GL}_2} \rtimes \sigma$ for some characters χ, χ' , and σ on F^\times according as π is of type I or IIb. The representation $(\chi \mathbf{1}_{\mathrm{GL}_2} \rtimes \sigma, V)$ can be realized as a subrepresentation of $|\cdot|^{1/2} \chi \times |\cdot|^{-1/2} \chi \rtimes \sigma$, and a basis of 3-dimensional space $V^{K_0(\mathfrak{p})}$ is given by

$$f_1^{\mathrm{IIb}} := f_{1_4} + f_{s_1}, \quad f_2^{\mathrm{IIb}} := f_{s_2} + f_{s_2 s_1} + f_{s_1 s_2 s_1} + f_{s_1 s_2}, \quad f_3^{\mathrm{IIb}} := f_{s_1 s_2 s_1 s_2} + f_{s_2 s_1 s_2}.$$

The representation matrices of operators $T_{1,0}$ and η on $V_\pi^{K_0(\mathfrak{p})}$ are given in [31, Table 3] and [31, Lemma 2.1].

Now we assume that π has the trivial central character. Let (ℓ_π^0, f_K) be the Bessel data for π defined above. For $\bullet \in \{\mathrm{I}, \mathrm{IIb}\}$, define a basis $\{B_i^\bullet\}_{1 \leq i \leq \dim V_\pi^{K_0(\mathfrak{p})}}$ of $V_\pi^{K_0(\mathfrak{p})}$ by

$$B_i^\bullet(g) := \ell_\pi^0(\pi(g) f_i^\bullet), \quad g \in G(F).$$

We need the values of these functions at $g = \mathbf{1}_4$, which are given as follows.

Proposition 6.16. *Let π be a smooth admissible irreducible K -spherical representation of type I or IIb, and $\{B_i^\bullet\}$ be as above for $\bullet \in \{\mathrm{I}, \mathrm{IIb}\}$.*

(i) If π is of type I,

$$B_1^I(1_4) = \frac{\alpha\beta}{(q-\alpha)(q-\beta)}, \quad B_2^I(1_4) = \frac{-q\beta}{(q-\alpha)(q-\beta)},$$

$$B_3^I(1_4) = \frac{-q\alpha}{(q-\alpha)(q-\beta)}, \quad B_4^I(1_4) = \frac{q^2}{(q-\alpha)(q-\beta)}.$$

(ii) If π is of type IIb,

$$B_1^{\text{IIb}}(1_4) = \frac{\alpha^2}{(q^{1/2}-\alpha)(q^{3/2}-\alpha)}, \quad B_2^{\text{IIb}}(1_4) = \frac{-q^{1/2}(1+q)\alpha}{(q^{1/2}-\alpha)(q^{3/2}-\alpha)},$$

$$B_3^{\text{IIb}}(1_4) = \frac{q^2}{(q^{1/2}-\alpha)(q^{3/2}-\alpha)}.$$

Proof. Set $d := \dim V_\pi^{K_0(\mathfrak{p})}$. By the construction of the elements $\{f_i^\bullet\}_{1 \leq i \leq d}$, we have $B_\pi^0 = \sum_{i=1}^d B_i^\bullet$. For any $\ell \in \mathbb{Z}_{\geq 0}$, we let $T_{1,0}^\ell$ act on the both sides of this equation and then evaluate at $g = 1_4$ using (6-21); thus,

$$B_\pi^0(h(\ell, 0)) = q^{-3\ell} \sum_{i=1}^d (T_{1,0}^\ell B_i^\bullet)(1_4).$$

The value on the left-hand side is given in [6, Corollary 1.8]; by [31, Table 3], we have a system of linear equations among $B_i^\bullet(1_4)$, which is solved easily. \square

6.8. Local periods. For an irreducible admissible unitalizable $K_0(\mathfrak{p})$ -spherical representation (π, V_π) on $G(F)$ having the (T, Λ) -Bessel model, we fix a pair $(\ell^\pi, \xi^\pi) \in (V_\pi^*)^{T, \Lambda} \times V_\pi^{K_0(\mathfrak{p})}$ satisfying $\ell_\pi(\xi^\pi) = 1$. When π is K -spherical, we choose (ℓ^π, ξ^π) as the unramified Bessel datum (ℓ_0^π, ξ_0^π) . Then, there exists a unique $G(F)$ -invariant inner product $\langle \cdot, \cdot \rangle$ on V_π such that $\langle \xi^\pi | \xi^\pi \rangle = 1$. For $s \in \mathbb{C}$, $\phi \in \mathcal{S}(L^2)$, $g \in G(F)$, and a character μ on F^\times , we define the local period of π as

$$\mathbb{1}_{\pi, K_0(\mathfrak{p})}^{(s)}(\xi^\pi, \phi, \Lambda, \mu; g) = \sum_{B \in \mathcal{B}((V_\pi^*)^{T, \Lambda}, K_0(\mathfrak{p}))} Z^*(\phi, B, s, \mu; g) \overline{B(1_4)},$$

where

$$Z^*(\phi, B, s, \mu; g) := L(s + 1/2, \pi, \mu)^{-1} Z(\phi, B, s, \mu; g)$$

is the normalized zeta integral and $\mathcal{B}((V_\pi^*)^{T, \Lambda}, K_0(\mathfrak{p}))$ is an orthonormal basis of $\mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$ with respect to the inner product on $\mathcal{B}(T, \Lambda)[\pi]$ via the $G(F)$ -isomorphism

$$(6-31) \quad V_\pi \ni \xi \mapsto (g \mapsto \ell_\pi(\pi(g)\xi)) \in \mathcal{B}(T, \Lambda)[\pi].$$

6.8.1. Local periods for type I and IIb. When π is of type I or IIb, we recall that the representation space can be realized as in Section 6.7. We choose $\xi^\pi = \xi_0^\pi = f_K$. Then, since we are dealing with induced representations from unitary characters,

the standard inner product $\langle f|g \rangle := \int_K f(k)\overline{g(k)}dk$ ($f, g \in V_\pi$) becomes $G(F)$ -invariant and satisfies $\langle \xi_0^\pi | \xi_0^\pi \rangle = 1$.

For $\bullet \in \{I, \text{IIb}\}$, the set $\{\langle B_i^\bullet | B_i^\bullet \rangle^{-\frac{1}{2}} B_i^\bullet\}_{1 \leq i \leq \dim V_\pi^{K_0(\mathfrak{p})}}$ forms an orthonormal basis of $\mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$. By [36, Remark 2.1.3], the values $\langle B_i^\bullet | B_i^\bullet \rangle$ are given as follows:

$$\begin{aligned} \langle B_i^I | B_i^I \rangle &= q^{i-1}(q+1) \quad (1 \leq i \leq 4), \\ \langle B_1^{\text{IIb}} | B_1^{\text{IIb}} \rangle &= (q+1), \quad \langle B_2^{\text{IIb}} | B_2^{\text{IIb}} \rangle = q(q+1)^2, \quad \langle B_3^{\text{IIb}} | B_3^{\text{IIb}} \rangle = q^3(q+1). \end{aligned}$$

By Propositions 6.10 and 6.16, and the matrix representation of $T_{1,0}$ and η given in [31, Table 3] and [31, Lemma 2.1], we can find the local period by a direct calculation.²

Proposition 6.17. *Let π be a smooth admissible irreducible unitary representation of type I or type IIb with trivial central character. Suppose that Λ and μ are unramified. When $\phi = \mathbb{1}_{\sigma_L \oplus \sigma_L}$, we have*

$$\begin{aligned} \mathbb{I}_{\pi, K_0(\mathfrak{p})}^{(s)}(\xi_0^\pi, \phi, \Lambda, \mu; \eta) &= \frac{2(q-1)}{q^5(q^2+1)} L(1, \pi; \text{Std}) \\ &\quad \times \left(\mu(\varpi)^{-1} q^{s+1} + \mu(\varpi) q^{-s+1} - \frac{1}{q+1} \text{tr}(T_{1,0} + q\eta|_{V_\pi^{K_0(\mathfrak{p})}}) \right), \end{aligned}$$

where $L(s, \pi; \text{Std})$ is the standard L-function of π defined by

$$L(s, \pi; \text{Std}) = (1-\alpha q^{-s})^{-1} (1-\beta q^{-s})^{-1} (1-\alpha^{-1} q^{-s})^{-1} (1-\beta^{-1} q^{-s})^{-1} (1-q^{-s})^{-1}$$

or

$$L(s, \pi; \text{Std}) = (1-\alpha q^{-s+\frac{1}{2}})^{-1} (1-\alpha q^{-s-\frac{1}{2}})^{-1} (1-\alpha^{-1} q^{-s+\frac{1}{2}})^{-1} (1-\alpha^{-1} q^{-s-\frac{1}{2}})^{-1} (1-q^{-s})^{-1},$$

according as π is of type I or IIb.

6.8.2. Local period for type IIIa. Suppose that π is of type IIIa. Then, as remarked before, $|\alpha| = |\gamma| = 1$ and $\alpha \neq 1$. By Lemma 6.11, there exists a basis $\{B_i\}_{i=1,2}$ of $\mathcal{B}(T, \Lambda)[\pi]^{K_0(\mathfrak{p})}$, unique up to constant, satisfying the relations (6-20) and $B_2(1_4) = \alpha^{-1} B_1(1_4) \neq 0$. Thus, we can uniquely fix $\{B_i^{\text{IIIa}}\}_{i=1,2}$ by requiring $B_1^{\text{IIIa}}(1_4) = 1$ and the datum (ℓ_π, ξ^π) by $\ell_\pi(\xi^\pi) = B_1^{\text{IIIa}}$ to form (6-31) and induce the $G(F)$ -invariant inner product on $\mathcal{B}(T, \Lambda)[\pi]$ such that $\langle B_1^{\text{IIIa}} | B_1^{\text{IIIa}} \rangle = 1$. Since $\langle B_1 | B_2 \rangle$ vanishes by [7, Lemma 2.11], we then have $\langle B_2^{\text{IIIa}} | B_2^{\text{IIIa}} \rangle = 1$ and $\langle B_1^{\text{IIIa}} | B_2^{\text{IIIa}} \rangle = 0$ due to $|\alpha| = 1$. The following proposition is immediate from Proposition 6.12.

Proposition 6.18. *Let π be a smooth admissible irreducible unitary representation of type IIIa. Suppose that Λ and μ are unramified. When $\phi = \mathbb{1}_{\sigma_L \oplus \sigma_L}$,*

$$\mathbb{I}_{\pi, K_0(\mathfrak{p})}^{(s)}(\xi^\pi, \phi, \Lambda, \mu; \eta) = \frac{2\Lambda(\varpi)^{-1} \mu(\varpi)^{-1} q^{s+1}}{q^2 + 1}.$$

²We have used MATHEMATICA 13.

6.8.3. Local period for type VIb. Suppose that π is of type VIb. We recall that $\dim(V_\pi^{K_0(\mathfrak{p})}) = \dim(\mathcal{B}(T, \Lambda)^{K_0(\mathfrak{p})}) = 1$; thus, by [32, Theorem 9.3], there exists a unique element $B^{\text{VIb}} \in \mathcal{B}(T, \Lambda)^{K_0(\mathfrak{p})}$ satisfying $B^{\text{VIb}}(1_4) = 1$. We fix the datum (ℓ_π, ξ^π) by setting $\ell_\pi(\xi^\pi) = B^{\text{VIb}}$. The following result is immediate from Proposition 6.14.

Proposition 6.19. *Let π be a smooth admissible irreducible unitary representation of type VIb. Suppose that Λ and μ are unramified. When $\phi = \mathbb{1}_{\sigma_L \oplus \sigma_L}$,*

$$\mathbb{1}_{\pi, K_0(\mathfrak{p})}^{(s)}(\xi^\pi, \phi, \Lambda, \mu; \eta) = \frac{\Lambda(\varpi)^{-1} \mu(\varpi)^{-1} q^{s+1}}{q^2 + 1}.$$

Acknowledgments

The authors thank the referee for a careful reading and constructive suggestions.

Kuga was supported by JSPS KAKENHI Grant Number 23KJ1931. Tsuzuki was supported by JSPS KAKENHI Grant Number 20K03565.

References

- [1] A. N. Andrianov, “Dirichlet series with Euler product in the theory of Siegel modular forms of genus two”, *Trudy Mat. Inst. Steklov.* **112** (1971), 73–94, 386. In Russian; translated in *Proc. Steklov Inst. Math.* **112** (1971), 70–93. MR
- [2] A. N. Andrianov, “Euler products that correspond to Siegel’s modular forms of genus 2”, *Uspehi Mat. Nauk* **29:3(177)** (1974), 43–110. In Russian; translated in *Russ. Math. Surveys* **29:3**, 1974, 45–116. MR
- [3] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, AMS Colloquium Publications **61**, American Mathematical Society, 2013. MR
- [4] M. Asgari and R. Schmidt, “Siegel modular forms and representations”, *Manuscripta Math.* **104:2** (2001), 173–200. MR
- [5] V. Blomer, “Spectral summation formulae for $\text{GSp}(4)$ and moments of spinor L -functions”, *J. Eur. Math. Soc.* **21:6** (2019), 1751–1774. MR
- [6] D. Bump, S. Friedberg, and M. Furusawa, “Explicit formulas for the Waldspurger and Bessel models”, *Israel J. Math.* **102** (1997), 125–177. MR
- [7] M. Dickson, A. Pitale, A. Saha, and R. Schmidt, “Explicit refinements of Böcherer’s conjecture for Siegel modular forms of squarefree level”, *J. Math. Soc. Japan* **72:1** (2020), 251–301. MR
- [8] B. Feigon and D. Whitehouse, “Averages of central L -values of Hilbert modular forms with an application to subconvexity”, *Duke Math. J.* **149:2** (2009), 347–410. MR
- [9] M. Furusawa and K. Morimoto, “Refined global Gross–Prasad conjecture on special Bessel periods and Böcherer’s conjecture”, *J. Eur. Math. Soc.* **23:4** (2021), 1295–1331. MR
- [10] M. Furusawa and K. Morimoto, “On the Gross–Prasad conjecture with its refinement for $(\text{SO}(5), \text{SO}(2))$ and the generalized Böcherer conjecture”, *Compos. Math.* **160:9** (2024), 2115–2202. MR
- [11] D. Goldfeld, J. Hoffstein, and D. Lieman, “Appendix: An effective zero-free region”, *Ann. of Math. (2)* **140:1** (1994), 177–181. MR
- [12] J. Hoffstein and P. Lockhart, “Coefficients of Maass forms and the Siegel zero”, *Ann. of Math. (2)* **140:1** (1994), 161–176. MR

- [13] J. Hoffstein and D. Ramakrishnan, “Siegel zeros and cusp forms”, *Internat. Math. Res. Notices* **1995**:6 (1995), 279–308. MR
- [14] H. Jacquet, *Automorphic forms on $\mathrm{GL}(2)$, II*, Lecture Notes in Mathematics **278**, Springer, 1972. MR
- [15] H. Jacquet and N. Chen, “Positivity of quadratic base change L -functions”, *Bull. Soc. Math. France* **129**:1 (2001), 33–90. MR
- [16] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, “Rankin–Selberg convolutions”, *Amer. J. Math.* **105**:2 (1983), 367–464. MR
- [17] H. H. Kim, S. Wakatsuki, and T. Yamauchi, “An equidistribution theorem for holomorphic Siegel modular forms for GSp_4 and its applications”, *J. Inst. Math. Jussieu* **19**:2 (2020), 351–419. MR
- [18] E. Kowalski, A. Saha, and J. Tsimerman, “Local spectral equidistribution for Siegel modular forms and applications”, *Compos. Math.* **148**:2 (2012), 335–384. MR
- [19] S. Kuga and M. Tsuzuki, “An asymptotic formula of spectral average of central L -values on $\mathrm{GSp}(2)$ for square free levels”, preprint, 2024. arXiv 2410.03166
- [20] F. Lemma, “On higher regulators of Siegel threefolds, II: The connection to the special value”, *Compos. Math.* **153**:5 (2017), 889–946. MR
- [21] X. Li, “Upper bounds on L -functions at the edge of the critical strip”, *Int. Math. Res. Not.* **2010**:4 (2010), 727–755. MR
- [22] Y. Liu, “Refined global Gan–Gross–Prasad conjecture for Bessel periods”, *J. Reine Angew. Math.* **717** (2016), 133–194. MR
- [23] D. Loeffler, V. Pilloni, C. Skinner, and S. L. Zerbes, “Higher Hida theory and p -adic L -functions for GSp_4 ”, *Duke Math. J.* **170**:18 (2021), 4033–4121. MR
- [24] P. Michel and A. Venkatesh, “The subconvexity problem for GL_2 ”, *Publ. Math. Inst. Hautes Études Sci.* **111** (2010), 171–271. MR
- [25] C. J. Moreno, “Analytic proof of the strong multiplicity one theorem”, *Amer. J. Math.* **107**:1 (1985), 163–206. MR
- [26] M. E. Novodvorskii, “Uniqueness theorems for generalized Bessel models”, *Mat. Sb. (N.S.)* **90**(132) (1973), 275–287, 326. In Russian; translated in *Math. USSR Sbornik* **19**:2, 1973, 275–286. MR
- [27] M. E. Novodvorskii and I. I. Piateckii-Shapiro, “Generalized Bessel models for the symplectic group of rank 2”, *Mat. Sb. (N.S.)* **90**(132) (1973), 246–256. In Russian; translated in *Math. USSR Sbornik* **19**:2, 1973, 243–255. MR
- [28] I. I. Piatetski-Shapiro, “On the Saito–Kurokawa lifting”, *Invent. Math.* **71**:2 (1983), 309–338. MR
- [29] I. I. Piatetski-Shapiro, “ L -functions for GSp_4 ”, *Pacific J. Math.* (special issue) (1997), 259–275. MR
- [30] A. Pitale and R. Schmidt, “Ramanujan-type results for Siegel cusp forms of degree 2”, *J. Ramanujan Math. Soc.* **24**:1 (2009), 87–111. MR
- [31] A. Pitale and R. Schmidt, “Bessel models for $\mathrm{GSp}(4)$: Siegel vectors of square-free level”, preprint (long version of [32]), 2013, available at <https://sites.math.unt.edu/~schmidt/papers/BSI.pdf>.
- [32] A. Pitale and R. Schmidt, “Bessel models for $\mathrm{GSp}(4)$: Siegel vectors of square-free level”, *J. Number Theory* **136** (2014), 134–164. MR
- [33] D. Prasad and R. Takloo-Bighash, “Bessel models for $\mathrm{GSp}(4)$ ”, *J. Reine Angew. Math.* **655** (2011), 189–243. MR

- [34] B. Roberts and R. Schmidt, *Local newforms for $\mathrm{GSp}(4)$* , Lecture Notes in Mathematics **1918**, Springer, 2007. MR
- [35] P. Sarnak, “Nonvanishing of L -functions on $\mathfrak{R}(s) = 1$ ”, pp. 719–732 in *Contributions to automorphic forms, geometry, and number theory*, Johns Hopkins Univ. Press, Baltimore, MD, 2004. MR
- [36] R. Schmidt, “Iwahori-spherical representations of $\mathrm{GSp}(4)$ and Siegel modular forms of degree 2 with square-free level”, *J. Math. Soc. Japan* **57**:1 (2005), 259–293. MR
- [37] R. Schmidt, “The Saito–Kurokawa lifting and functoriality”, *Amer. J. Math.* **127**:1 (2005), 209–240. MR
- [38] R. Schmidt, “On classical Saito–Kurokawa liftings”, *J. Reine Angew. Math.* **604** (2007), 211–236. MR
- [39] R. Schmidt, “Packet structure and paramodular forms”, *Trans. Amer. Math. Soc.* **370**:5 (2018), 3085–3112. MR
- [40] R. Schmidt and L. Tran, “Zeta integrals for $\mathrm{GSp}(4)$ via Bessel models”, *Pacific J. Math.* **296**:2 (2018), 437–480. MR
- [41] D. Soudry, “The CAP representations of $\mathrm{GSp}(4, \mathbb{A})$ ”, *J. Reine Angew. Math.* **383** (1988), 87–108. MR
- [42] T. Sugano, “On holomorphic cusp forms on quaternion unitary groups of degree 2”, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**:3 (1985), 521–568. MR
- [43] R. Takloo-Bighash, “ L -functions for the p -adic group $\mathrm{GSp}(4)$ ”, *Amer. J. Math.* **122**:6 (2000), 1085–1120. MR
- [44] J. T. Tate, Jr., *Fourier analysis in number fields and Hecke’s zeta-functions*, Ph.D. thesis, Princeton University, 1950, available at <https://www.proquest.com/docview/304411725>. MR
- [45] J. Tate, “Number theoretic background”, pp. 3–26 in *Automorphic forms, representations and L -functions* (Corvallis, OR, 1977), vol. 2, Proc. Sympos. Pure Math. **33**(2), Amer. Math. Soc., 1979. MR
- [46] N. R. Wallach, “Representations of reductive Lie groups”, pp. 71–86 in *Automorphic forms, representations and L -functions* (Corvallis, OR, 1977), vol. 1, Proc. Sympos. Pure Math. **33**(1), Amer. Math. Soc., 1979. MR
- [47] A. Weil, *Basic number theory*, Grundle Math. Wissen. **144**, Springer, 1967. MR
- [48] R. Weissauer, *Endoscopy for $\mathrm{GSp}(4)$ and the cohomology of Siegel modular threefolds*, Lecture Notes in Mathematics **1968**, Springer, 2009. MR

Received July 21, 2024. Revised November 22, 2025.

SEIJI KUGA
FACULTY OF SCIENCE AND TECHNOLOGY
SOPHIA UNIVERSITY
TOKYO
JAPAN
s-kuga-2g7@sophia.ac.jp

MASAO TSUZUKI
FACULTY OF SCIENCE AND TECHNOLOGY
SOPHIA UNIVERSITY
TOKYO
JAPAN
m-tsuduk@sophia.ac.jp

SUM OF THE SQUARES OF THE p' -CHARACTER DEGREES

NGUYEN N. HUNG, J. MIQUEL MARTÍNEZ AND GABRIEL NAVARRO

We study the sum of the squares of the irreducible character degrees not divisible by some prime p of a finite group, and its relationship with the corresponding quantity in a p -Sylow normalizer. This led us to study a recent conjecture by E. Giannelli, which we prove for $p = 2$ and in some other cases.

1. Introduction

There are some strikingly simple and seemingly innocent statements that follow from the recently proven McKay conjecture [CS24] (or from the techniques developed to prove it) whose validity appears to resist any elementary justification. For instance, W. Feit was interested in proving that if P is an abelian Sylow p -subgroup of G , then $k(G) \geq k(N_G(P))$, where $k(G)$ is the number of conjugacy classes of the finite group G [F80]. Although this inequality now follows as a consequence of [CS24], no alternative proof is currently known. (Equality happens if and only if $P \trianglelefteq G$ by the Itô–Michler theorem.)

This current project — long awaited by the third author — was expected to provide another example of such phenomena. And yet, it does not. What was anticipated to be a theorem remains, for now, a conjecture. As usual, $\text{Irr}_{p'}(G)$ is the set of irreducible complex characters of the finite group G whose degree is not divisible by the prime p , and $P' = [P, P]$ is the derived subgroup of the group P .

Conjecture A. *Let G be a finite group and $P \in \text{Syl}_p(G)$. Then*

$$\sum_{\chi \in \text{Irr}_{p'}(G)} \chi(1)^2 \geq |N_G(P) : P'|,$$

with equality if and only if $N_G(P)$ has a normal complement in G .

N. N. Hung gratefully acknowledges the support of the AMS-Simons Research Enhancement Grant (AWD-000167 AMS). The work of J. M. Martínez and G. Navarro is supported by Grant PID2022-137612NB-I00 funded by MCIN/AEI/ 10.13039/501100011033 and ERDF “A way of making Europe.” We thank Gunter Malle for helpful conversations and for noticing gaps in the proofs of Proposition 3.2 and Lemma 4.8 in an earlier version. We thank the referees for their thorough reading and their many comments and suggestions, which helped improve the exposition of this paper. MSC2020: 20D20, 20C15.

Keywords: character degree, Sylow normalizer, McKay conjecture.

The McKay theorem establishes that there exists a bijection $f : \text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$. One way to prove Conjecture A, perhaps the most obvious, is to show that we can choose f to satisfy the additional condition that $f(\chi)(1) \leq \chi(1)$ for all $\chi \in \text{Irr}_{p'}(G)$. This turns out to be a conjecture by E. Giannelli [G25], and it does not appear to follow easily from the proof of the McKay conjecture. In Theorem 3.5 below we carry out the standard reduction of the McKay conjecture now incorporating Giannelli's strengthening, which is therefore reduced now to a question on simple groups. This question is verified in a number of important cases in Sections 4 and 5 below. In Theorem 5.2, we completely prove it for $p = 2$, and this establishes:

Theorem B. *Conjecture A holds for $p = 2$.*

Can Conjecture A be proven independently of Giannelli's strengthening of McKay? We do not know the answer to this question. Characters of p' -degree in a group tend to have much larger degree than those in the normalizer of a Sylow p -subgroup, but so far, no effective strategy has been developed to exploit this observation.

Theorem C below, whose proof is rather involved — but does not use the classification of finite simple groups — takes care of the equality in Conjecture A (assuming Giannelli's conjecture), and might have independent interest.

Theorem C. *Let G be a finite group, p a prime, and $P \in \text{Syl}_p(G)$. Then all irreducible characters of p' -degree of $N_G(P)$ extend to G if and only if there is $K \trianglelefteq G$ such that $KN_G(P) = G$ and $N_K(P) = 1$.*

This extends [I86, Theorem B], where it is required that *all* the irreducible characters of $N_G(P)$ extend to G . As we shall explain, in order to prove Theorem C, we shall need to establish a relative version of it with respect to a normal subgroup. We find it surprising that a character restriction type of theorem admits a relative version, since these versions regarding group structure occur very rarely.

There are some variations of Conjecture A which we do not attempt here. Among others, it seems reasonable to replace P' by the Frattini subgroup $\Phi(P)$ and the set of p' -degree irreducible characters by the so-called *almost p -rational* characters of p' -degree. Proving this, however, seems much more complicated.

Concerning Giannelli's conjecture, let us mention here that in general it is not always possible to find McKay bijections $f : \text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$ such that $f(\chi)$ is an irreducible constituent of $\chi_{N_G(P)}$, as shown by A_5 for $p = 2$, or $\text{GL}_2(3)$ for $p = 3$.

A final but important remark is in order: Why do we care about the sum of the squares of the irreducible character degrees not divisible by p ? First of all, this

number would be the dimension of any complex algebra affording the p' -degree irreducible characters. Also, of the famous list of problems by Richard Brauer [B63], Problem 2 asks when non-isomorphic finite groups G and H have isomorphic group algebras $\mathbb{C}G$ and $\mathbb{C}H$. If G has a normal p -complement, Isaacs proved that H has a normal p -complement too [N18, Theorem 7.8]. The question of whether the normality of a Sylow p -subgroup P of G is recognized by $\mathbb{C}G$ remains open. In [N04], it is suggested that, perhaps, this happens if and only if

$$\left(\sum_{\chi \in \text{Irr}_{p'}(G)} \chi(1)^2 \right)_{p'} = |G|_{p'},$$

where $n_{p'} = n/n_p$ with n_p denoting the largest power of the prime p dividing the integer n .

2. Proof of Theorem C

Our notation for ordinary characters follows [I06] and [N18]. In this section we prove Theorem C. In order to do that, we need to prove a relative version of it, that allows us to use induction. This makes the proof more complicated. We start with some preliminary results.

Lemma 2.1. *Let $N \trianglelefteq G$ and let $\theta \in \text{Irr}(N)$ be G -invariant. Then*

$$\ker \theta = \bigcap_{\chi \in \text{Irr}(G|\theta)} \ker \chi.$$

Proof. It is clear that $\ker \theta$ is contained in $\ker \chi$ for every $\chi \in \text{Irr}(G|\theta)$, since $\chi_N = e_\chi \theta$ for $\chi \in \text{Irr}(G|\theta)$. Suppose that $x \in \bigcap_{\chi \in \text{Irr}(G|\theta)} \ker \chi$. Then $\theta^G(x) = \theta^G(1) = |G : N|\theta(1) \neq 0$, and therefore $x \in N$. Since θ is G -invariant, we have $\theta^G(x) = |G : N|\theta(x)$, and thus $\theta(x) = \theta(1)$, as wanted. \square

Lemma 2.2. *Suppose that $K \trianglelefteq G$, $H \leq G$, $N = K \cap H$ and $G = KH$. Let $\gamma \in \text{Irr}(K)$ be G -invariant, and let $\theta \in \text{Irr}(N)$ be an irreducible H -invariant constituent of γ_N . Suppose that restriction defines a bijection $\text{Irr}(G|\gamma) \rightarrow \text{Irr}(H|\theta)$. Then $\gamma_N = \theta$.*

Proof. We have

$$\sum_{\chi \in \text{Irr}(G|\gamma)} (\chi(1)/\gamma(1))^2 = |G : K| \quad \text{and} \quad \sum_{\tau \in \text{Irr}(H|\theta)} (\tau(1)/\theta(1))^2 = |H : N|.$$

By hypothesis,

$$\begin{aligned} \sum_{\tau \in \text{Irr}(H|\theta)} (\tau(1)/\theta(1))^2 &= \sum_{\chi \in \text{Irr}(G|\gamma)} (\chi(1)/\theta(1))^2 \\ &= (\gamma(1)/\theta(1))^2 \sum_{\chi \in \text{Irr}(G|\gamma)} (\chi(1)/\gamma(1))^2. \end{aligned}$$

Since $|H : N| = |G : K|$, we deduce that $\gamma_N = \theta$. \square

Lemma 2.3. *Suppose that $G = KP$, where K is a normal p -complement and $P \in \text{Syl}_p(G)$. Let $\gamma \in \text{Irr}(G)$. If $P \subseteq \ker \gamma$, then $\gamma_{C_K(P)}$ is irreducible.*

Proof. Since $\ker \gamma \trianglelefteq G$, we have $[K, P]P \subseteq \ker \gamma$. Thus γ_K is an irreducible character of K with $[K, P]$ contained in its kernel. By coprime action (for instance, Lemma 4.28 of [I08]), we have $C_K(P)[K, P] = K$ and thus $\gamma_{C_K(P)}$ is irreducible. \square

Lemma 2.4. *Suppose that $N \trianglelefteq G$, $P \in \text{Syl}_p(G)$. Let $N \leq K_i \trianglelefteq G$ be complementing $NN_G(P)/N$ for $i = 1, 2$ in G/N . Then $K_1 = K_2$.*

Proof. By working in G/N , we may assume that $N = 1$. Then G/K_i has a normal Sylow p -subgroup. Let $L = K_1 \cap K_2$. Then G/L has a normal Sylow p -subgroup. Hence, $LN_G(P) = G$. Then $K_i = K_i \cap LN_G(P) = L(K_i \cap N_G(P)) = L$, as wanted. \square

In several occasions, we shall use Tate's theorem in the following form.

Theorem 2.5. *Suppose that G is a finite group, $P \in \text{Syl}_p(G)$, and $K \trianglelefteq G$. If $K \cap P = K \cap P'$, then K has a normal p -complement.*

Proof. We may assume that $G = KP$. In the notation of [I06, Theorem 6.31], we have $P \cap A^p(G) = A^p(P) = P'$, where $A^p(G)$ is the smallest normal subgroup L of G such that G/L is an abelian p -group. It then follows that G has a normal p -complement by [I06, Theorem 6.31]. Therefore, K has a normal p -complement. \square

Let $N \trianglelefteq G$ and let $\theta \in \text{Irr}(N)$ be G -invariant. We let

$$\text{Irr}_{p', \text{rel}}(G|\theta) := \{\chi \in \text{Irr}(G|\theta) \mid p \text{ does not divide } \chi(1)/\theta(1)\}.$$

The following easily implies Theorem C (when N is trivial).

Theorem 2.6. *Let $N \trianglelefteq G$ and let $\theta \in \text{Irr}(N)$ be G -invariant. Let $P \in \text{Syl}_p(G)$. Assume that θ extends to NP . The following statements are equivalent.*

- (i) *Every $\chi \in \text{Irr}_{p', \text{rel}}(NN_G(P)|\theta)$ extends to G .*
- (ii) *Every $\chi \in \text{Irr}(NN_G(P)|\theta)$ extends to G .*
- (iii) *There is a normal complement K/N to $N_G(P)N/N$ in G/N such that θ has an extension $\hat{\theta} \in \text{Irr}(K)$.*
- (iv) *There is a normal complement K/N to $N_G(P)N/N$ in G/N such that θ has a G -invariant extension $\hat{\theta} \in \text{Irr}(K)$.*

Proof. First we prove that (iv) and (iii) are equivalent. We only have to prove that (iii) implies (iv). Since $PN \cap K = N$, we have that K/N is a group of order not divisible by p . By Sylow theory, $N_G(PN) = N_G(P)N$, and therefore $C_{K/N}(P) = C_{K/N}(PN/N) \subseteq K/N \cap N_{G/N}(PN/N) = 1$. Let $\eta \in \text{Irr}(K)$ be an extension of θ , and let Δ be the set of extensions of θ to K . We have that

$\Delta = \{\lambda\eta \mid \lambda \in \text{Irr}(K/N) \text{ is linear}\}$, by Gallagher's theorem [I06, Corollary 6.17]. Therefore $|\Delta|$ is not divisible by p . Since P acts on Δ by conjugation, because θ is G -invariant, by counting we have that there is some P -invariant $\hat{\theta} \in \Delta$ extending θ . Since $C_{K/N}(P) = 1$, we have that $\hat{\theta}$ is the unique P -invariant character of K lying over θ (using [I06, Theorem 13.31 and Problem 13.10]). We shall use this argument several times). We claim that $\hat{\theta}$ is also G -invariant. It is enough to show that $\hat{\theta}$ is $N_G(P)$ -invariant. If $n \in N_G(P)$, then $\hat{\theta}^n$ is a P -invariant extension of $\theta^n = \theta$. By uniqueness, $\hat{\theta}^n = \hat{\theta}$, as wanted.

To prove that (iv) implies (ii), we apply [N18, Lemma 6.8(d)].

Since it is clear that (ii) implies (i), to complete the proof of the theorem, it is enough to show that (i) implies (iii). We argue this by induction on $|G : N|$. Recall that $NN_G(P)/N = N_{G/N}(PN/N)$. By using the theory of character triple isomorphisms (see [I06, Chapter 11]), we may assume that $N \subseteq Z(G)$. Hence, $N \subseteq N_G(P)$. In particular, θ is linear. Write $\theta = \theta_p\theta_{p'}$, where θ_p has order a power of p and $\theta_{p'}$ has order not divisible by p . Since θ extends to PN by hypothesis, therefore so does θ_p , which is a power of θ . Hence, θ_p extends to some linear character $\nu \in \text{Irr}(G)$ (by [I06, Theorem 6.26]). If $\chi \in \text{Irr}_{p',\text{rel}}(N_G(P)|\theta_{p'})$, then $\chi\nu_{N_G(P)} \in \text{Irr}_{p',\text{rel}}(N_G(P)|\theta)$. If $\eta \in \text{Irr}(G)$ extends $\chi\nu_{N_G(P)}$, then $\eta\nu^{-1}$ extends χ . Therefore, we may assume that θ is a linear character of p' -order. By modding out by $\ker \theta$, we may assume that N is a p' -group and that θ is faithful. Hence, our hypothesis is that every $\chi \in \text{Irr}_{p'}(N_G(P)|\theta)$ extends to G . We want to show that there is $K \trianglelefteq G$ complementing $N_G(P)/N$ in G , and that θ extends to K . For each $\tau \in \text{Irr}_{p'}(N_G(P)|\theta)$, fix $\tilde{\tau} \in \text{Irr}(G)$ an extension of τ to G .

If $P' \subseteq X \leq P$ and $X \trianglelefteq N_G(P)$, we claim that there exists $L \trianglelefteq G$ such that $L \cap N_G(P) = NX$. Let $\tilde{\theta} = \theta \times 1_X$. Notice that all $\text{Irr}(N_G(P)|\tilde{\theta})$ have p' -degree because $N_G(P)/X$ has an abelian normal Sylow p -subgroup P/X . Let

$$U = \bigcap_{\tau \in \text{Irr}(N_G(P)|\tilde{\theta})} \ker \tilde{\tau} \trianglelefteq G.$$

Then, using Lemma 2.1, we have $U \cap N_G(P) = \ker \tilde{\theta} = X$. Let $L = UN \trianglelefteq G$. Therefore $L \cap N_G(P) = NX$, as claimed. Notice that in this case, $X = P \cap L \in \text{Syl}_p(L)$.

By letting $X = P'$ in the claim in the previous paragraph of this proof, let $L \trianglelefteq G$ such that $L \cap N_G(P) = NP'$. Since $P \cap L = P'$, by Tate's theorem (Theorem 2.5), L has a normal p -complement K . Also, P' is a Sylow p -subgroup of L . Let $W = KN_G(P) = LN_G(P)$. Notice that $C_{K/N}(P) = 1$ since $N_K(P) = N$, and therefore over θ there is a unique P -invariant $\eta \in \text{Irr}(K)$ (again using [I06, Theorem 13.31 and Problem 13.10]). By uniqueness, notice that η is $N_G(P)$ -invariant. We have

$$|\text{Irr}_{p'}(W|\theta)| = |\text{Irr}_{p'}(N_G(P)|\theta)|$$

by the relative McKay conjecture for p -solvable groups [N18, Theorem 10.26]. (Notice that W is indeed p -solvable, so we have not used the solution of the general McKay conjecture.) We claim that $\text{Irr}_{p'}(W|\theta) = \{\tilde{\tau}_W \mid \tau \in \text{Irr}_{p'}(N_G(P)|\theta)\}$. Indeed, let $f : \text{Irr}_{p'}(N_G(P)|\theta) \rightarrow \text{Irr}_{p'}(W|\theta)$ given by $f(\tau) = \tilde{\tau}_W$. Since $(\tilde{\tau})_{N_G(P)} = \tau$, we do have $f(\tau) \in \text{Irr}_{p'}(W|\theta)$. Since f is clearly injective, it is necessarily bijective and the claim follows. Since η is the only P -invariant irreducible character of K over θ , it easily follows that $\text{Irr}_{p'}(G|\theta) = \text{Irr}_{p'}(G|\eta)$ and $\text{Irr}_{p'}(W|\theta) = \text{Irr}_{p'}(W|\eta)$. Indeed, if $\chi \in \text{Irr}_{p'}(G|\theta)$, then χ_K has some P -invariant irreducible constituent, using that χ has p' -degree. All irreducible constituents of χ_K lie over θ , so we deduce that this P -invariant constituent should be η . Notice then that every $\text{Irr}_{p'}(W|\eta)$ extends to G . We claim now that η is G -invariant. First notice that since η is P -invariant and K is a p' -group, η extends to KP (using Corollary 6.2 of [N18]). Hence, there is some p' -degree irreducible character of W over η . Therefore, there exists $\tau \in \text{Irr}_{p'}(N_G(P)|\theta)$ such that $\tilde{\tau}_W$ lies over η . Since η is invariant in W , it follows that $W \leq G_\eta$. If ϵ is the Clifford correspondent of $\tilde{\tau}$ over η , it follows that $\epsilon^G = \tilde{\tau}$. Since $\tilde{\tau}_{G_\eta}$ is irreducible, necessarily $G_\eta = G$. Hence, if $K > N$, we can apply induction to $\text{Irr}_{p'}(W|\eta)$ with respect to G .

Suppose first that $P' \trianglelefteq G$, and assume next that $P' > 1$. Working in $\bar{G} = G/P'$ and using induction, we conclude that there is $P' \leq R \trianglelefteq G$ such that $RN_G(P) = G$, and $R \cap N_G(P) = NP'$, and that $\theta \times 1_{p'}$ extends to $\gamma \in \text{Irr}(R)$. Since $P \cap R = P'$, we have that R has a normal p -complement S by Tate's theorem, and Sylow p -subgroup P' . This normal p -complement complements $N_G(P)/N$ in G/N . Since $\gamma_N = \theta$, we have that γ_S extends θ and we are done, in the case that $P' \trianglelefteq G$ and $P' > 1$. Assume now that P is abelian. Hence, all $\text{Irr}(N_G(P)|\theta) = \text{Irr}_{p'}(N_G(P)|\theta)$ extend to G , by hypothesis. By the claim in the fourth paragraph of this proof (letting $X = P$), let $Y \trianglelefteq G$ such that $Y \cap N_G(P) = NP$. Then $P \in \text{Syl}_p(Y)$ and $G = YN_G(P)$ by the Frattini argument. Now, $P \subseteq Z(N_Y(P))$, and by Burnside's p -complement theorem (Theorem 5.13 of [I08]), Y has a normal p -complement Q . Then $QN_G(P) = G$ and $Q \cap N_G(P) = N$. Notice that G is p -solvable. Since $C_{Q/N}(P) = 1$, let $\mu \in \text{Irr}(Q)$ be the unique P -invariant over θ . By uniqueness, μ is $N_G(P)$ -invariant, and therefore G -invariant. Also $\text{Irr}_{p'}(G|\theta) = \text{Irr}_{p'}(G|\mu) = \text{Irr}(G|\mu)$, using that μ extends to $QP = Y$, Gallagher [I06, Corollary 6.17], and the fact that P is abelian. By the relative version of the McKay conjecture (in p -solvable groups; [N18, Theorem 10.26]), we have

$$|\text{Irr}(G|\mu)| = |\text{Irr}_{p'}(G|\mu)| = |\text{Irr}_{p'}(G|\theta)| = |\text{Irr}_{p'}(N_G(P)|\theta)| = |\text{Irr}(N_G(P)|\theta)|.$$

By a previous argument, we see that $\text{Irr}(G|\mu) = \{\tilde{\tau} \mid \tau \in \text{Irr}(N_G(P)|\theta)\}$, and we see that restriction defines a bijection $\text{Irr}(G|\mu) \rightarrow \text{Irr}(N_G(P)|\theta)$. By Lemma 2.2, we have $\mu_N = \theta$, and we are done. Therefore we may assume that P' is not normal in G . Hence, using the notation of the fifth paragraph of this proof, we may assume that

$K > N$. (Otherwise, $L = NP' = N \times P'$ and necessarily $P' \trianglelefteq G$.) Every $\text{Irr}_{p'}(W|\eta)$ extends to G . By induction, there exists $K \leq E \trianglelefteq G$ complementing W/K and some $\rho \in \text{Irr}(E)$ such that $\rho_K = \eta$. Notice that E complements $N_G(P)/N$. In particular $C_{E/N}(P) = 1$, and ρ is the only P -invariant over θ . Hence $\text{Irr}_{p'}(G|\theta) = \text{Irr}_{p'}(G|\rho)$. By a previous argument, we know that ρ is $N_G(P)$ -invariant, and therefore G -invariant. We only need to show that $\rho_N = \theta$. Recall that G is p -solvable. Hence $|\text{Irr}_{p'}(G|\rho)| = |\text{Irr}_{p'}(G|\theta)| = |\text{Irr}_{p'}(N_G(P)|\theta)|$ (again by [N18, Theorem 10.26]), and therefore we deduce that

$$\text{Irr}_{p'}(G|\rho) = \{\tilde{\tau} \mid \tau \in \text{Irr}(N_G(P)|\theta)\}.$$

Using that E is a p' -group, let $\hat{\rho} \in \text{Irr}(EP)$ be an extension of ρ . Since $\hat{\rho}$ has p' -degree and $EP \trianglelefteq G$ (because $EN_G(P) = G$), it follows that $\hat{\rho}$ lies under some p' -irreducible character of G , call it χ . Hence $\hat{\rho}$ lies under some $\tilde{\tau} = \chi$ for some $\tau \in \text{Irr}_{p'}(N_G(P)|\theta)$. However $\tau_{P'}$ is a multiple of $1_{P'}$. Thus $\hat{\rho}_{P'}$ is a multiple of $1_{P'}$. Write $C = C_E(P')$. By Lemma 2.3 applied to EP' , we have that $\varphi = \hat{\rho}_C = \rho_C$ is irreducible. Hence, restriction defines a bijection $\text{Irr}(G|\rho) \rightarrow \text{Irr}(N_G(P')|\varphi)$, by [N18, Lemma 6.8(d)]. Notice that $\text{Irr}(N_G(P')|\varphi) = \text{Irr}(N_G(P')|\theta)$ since $C_{C/N}(P) = 1$. By the p -solvable case of the McKay conjecture [N18, Theorem 10.26] we know that $|\text{Irr}_{p'}(N_G(P')|\theta)| = |\text{Irr}_{p'}(N_G(P)|\theta)|$. Therefore

$$\text{Irr}_{p'}(N_G(P')|\theta) = \{\tilde{\tau}_{N_G(P')} \mid \tau \in \text{Irr}_{p'}(N_G(P)|\theta)\}.$$

Since $N_G(P') < G$, by induction, $N_G(P)/N$ has a normal complement in $N_G(P')/N$, which by Lemma 2.4 has to be C and that θ extends to C . By the first paragraph of this proof, θ has a P -invariant extension to C . Since $C_{C/N}(P) = 1$, then this P -invariant extension should be $\varphi = \rho_C$, and therefore ρ extends θ , as desired. \square

3. The McKay conjecture and inequality between character degrees

The following refinement of the McKay conjecture has been proposed in [G25].

Conjecture 3.1. *Let G be a finite group, p a prime and $P \in \text{Syl}_p(G)$. Then there is a bijection*

$$* : \text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$$

with

$$\chi^*(1) \leq \chi(1)$$

for all $\chi \in \text{Irr}_{p'}(G)$.

We observe now that if Conjecture 3.1 is true, then so is Conjecture A.

Proposition 3.2. *Let G be a finite group, let p be a prime and let $P \in \text{Syl}_p(G)$. If Conjecture 3.1 is true for G then*

$$\sum_{\chi \in \text{Irr}_{p'}(G)} \chi(1)^2 \geq |N_G(P) : P'|$$

with equality if and only if $N_G(P)$ has a normal complement in G .

Proof. Let $\Omega : \text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$ be the bijection from Conjecture 3.1. Then

$$\sum_{\chi \in \text{Irr}_{p'}(G)} \chi(1)^2 \geq \sum_{\chi \in \text{Irr}_{p'}(G)} \Omega(\chi)(1)^2 = \sum_{\psi \in \text{Irr}_{p'}(N_G(P))} \psi(1)^2 = |N_G(P) : P'|$$

and the inequality part follows. If

$$\sum_{\chi \in \text{Irr}_{p'}(G)} \chi(1)^2 = |N_G(P) : P'|$$

then for every $\chi \in \text{Irr}_{p'}(G)$ we have $\Omega(\chi)(1) = \chi(1)$. We claim that this implies that restriction is a bijection $\text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$. Write $H = N_G(P)$. Suppose that $a_1 = 1 < \dots < a_k$ are the degrees of $\text{Irr}_{p'}(G)$, and therefore, of $\text{Irr}_{p'}(H)$, occurring with multiplicities m_1, \dots, m_k , respectively. Let $\{\psi_1, \dots, \psi_t\}$ and $\{\chi_1, \dots, \chi_t\}$ be the irreducible characters in $\text{Irr}_{p'}(H)$ and $\text{Irr}_{p'}(G)$ of maximal degree a_k , where $t = m_k$. Now, $(\psi_i)^G$ contains an irreducible character of p' -degree of G of degree at least $\psi_i(1)$, using that $(\psi_i)^G(1)$ is not divisible by p . Necessarily, ψ_i contains some χ_j . Then $(\chi_j)_H = \psi_i$, and we can reorder so that $(\chi_i)_H = \psi_i$. Suppose that $\{\delta_1, \dots, \delta_s\}$ and $\{\eta_1, \dots, \eta_s\}$ are the irreducible characters in $\text{Irr}_{p'}(H)$ and $\text{Irr}_{p'}(G)$ of the next degree a_{k-1} , with $s = m_{k-1}$. Again $(\delta_i)^G$ contains a p' -irreducible character τ with degree at least a_{k-1} . If $\tau(1) = a_k$, then we know that τ_H is irreducible and then $\tau_H = \delta_i$. This is impossible. Hence, $(\delta_i)^G$ contains some η_j , which necessarily extends δ_i . Reordering, we may assume that $(\eta_i)_H = \delta_i$. We proceed like this until the claim is proven.

We may now apply Theorem 2.6 with $N = 1$ to conclude. □

The purpose of this section is to give a reduction of Conjecture 3.1 to simple groups. The following is a slight refinement of the inductive McKay condition, and will be the condition we impose on quasisimple groups in our reduction. We use the *central isomorphism* relation \geq_c from [N18, Definition 10.14], which generalizes the notion of a character triple isomorphism.

Conjecture 3.3. *Let S be a quasisimple group with cyclic center, $P \in \text{Syl}_p(S)$ and $A = \text{Aut}(S)_P$. Then there is an A -stable subgroup $N_S(P) \leq M < S$ and an A -equivariant bijection*

$$\Psi : \text{Irr}_{p'}(S) \rightarrow \text{Irr}_{p'}(M)$$

with

$$\Psi(\chi)(1) \leq \chi(1) \quad \text{and} \quad (S \rtimes A_\chi, S, \chi) \geq_c (M \rtimes A_\chi, M, \Psi(\chi))$$

for every $\chi \in \text{Irr}_{p'}(S)$.

Theorem 3.4. *Let $K \trianglelefteq G$ be perfect, and assume $K/\mathbf{Z}(K) \cong X^n$ where X is a finite simple group and $\mathbf{Z}(K)$ cyclic. Assume every quasisimple group S with $S/\mathbf{Z}(S) \cong X$ and $\mathbf{Z}(S)$ cyclic satisfies Conjecture 3.3. Let $Q \in \text{Syl}_p(G)$ and $R = Q \cap K$. Then there exists a $N_G(R)$ -stable subgroup $N_K(R) \leq H < K$ and a $N_G(R)$ -equivariant bijection*

$$\Psi : \text{Irr}_{p'}(K) \rightarrow \text{Irr}_{p'}(H)$$

such that

$$\Psi(\theta)(1) \leq \theta(1) \quad \text{and} \quad (G_\theta, K, \theta) \geq_c (HN_G(R)_\theta, H, \Psi(\theta))$$

for all $\theta \in \text{Irr}_{p'}(K)$.

Proof. K is a central product $S_1 \cdots S_n$ where $\mathbf{Z}(S_i) = \mathbf{Z}(K)$, each S_i is perfect and $S_i/\mathbf{Z}(S_i) \cong X$. Use the proof of [N18, Theorem 10.25] using the S_i 's instead of the universal covering group of X , and noticing that the construction of the bijection Ψ satisfies the degree inequality if one assumes Conjecture 3.3. Note that the subgroup H is constructed by taking the product of the subgroups M appearing in the statement of Conjecture 3.3. □

The next result follows the proof of [N18, Theorem 10.26]. We sketch it for the reader's convenience. Recall that we say a nonabelian finite simple group S is *involved* in G if there exist $H \trianglelefteq K \leq G$ with $K/H \cong S$.

Theorem 3.5. *Assume that, for every simple group X of order divisible by p involved in G , Conjecture 3.3 holds whenever S is a quasisimple group with $S/\mathbf{Z}(S) \cong X$ and $\mathbf{Z}(S)$ is cyclic. Let $Z \trianglelefteq G$, $P \in \text{Syl}_p(G)$, $\lambda \in \text{Irr}(Z)$ and assume λ is P -invariant. Then there is a bijection*

$$\Omega : \text{Irr}_{p'}(G|\lambda) \rightarrow \text{Irr}_{p'}(ZN_G(P)|\lambda)$$

with $\Omega(\chi)(1) \leq \chi(1)$ for all $\chi \in \text{Irr}_{p'}(G)$.

Proof. We argue by induction on $|G : Z|$. Since $|G : G_\lambda| \geq |ZN_G(P) : (ZN_G(P))_\lambda|$, by the Clifford correspondence we may assume λ is G -invariant. By character triple isomorphisms we may assume Z is central and cyclic and λ is linear and faithful. Let L/Z be a chief factor of G/Z . Let Δ be a G -transversal on the set of P -invariant characters in $\text{Irr}(L|\lambda)$ lying under some $\chi \in \text{Irr}_{p'}(G)$ and notice that [N18, Lemma 9.3] implies that Δ is also a $N_G(P)$ -transversal on the P -invariant characters in $\text{Irr}(L|\lambda)$ lying under some $\chi \in \text{Irr}_{p'}(LN_G(P))$. This implies that

$$\text{Irr}_{p'}(G|\lambda) = \bigsqcup_{\theta \in \Delta} \text{Irr}_{p'}(G|\theta)$$

and

$$\text{Irr}_{p'}(LN_G(P)|\lambda) = \bigsqcup_{\theta \in \Delta} \text{Irr}_{p'}(LN_G(P)|\theta).$$

By induction, for every $\mu \in \Delta$ there is a bijection

$$\Omega_\mu : \text{Irr}_{p'}(G|\mu) \rightarrow \text{Irr}_{p'}(LN_G(P)|\mu)$$

satisfying $\Omega_\mu(\chi)(1) \leq \chi(1)$ for every $\chi \in \text{Irr}_{p'}(G|\mu)$. We define

$$\Omega_1 : \text{Irr}_{p'}(G|\lambda) \rightarrow \text{Irr}_{p'}(LN_G(P)|\lambda)$$

by $\Omega_1(\chi) = \Omega_\mu(\chi)$ if $\chi \in \text{Irr}(G|\mu)$ and we have Ω_1 is a bijection satisfying $\Omega_1(\chi)(1) \leq \chi(1)$ for every $\chi \in \text{Irr}_{p'}(G|\lambda)$. If $LN_G(P) < G$ then by induction we have a bijection

$$\Omega_0 : \text{Irr}_{p'}(LN_G(P)|\lambda) \rightarrow \text{Irr}_{p'}(N_G(P)|\lambda)$$

also satisfying $\Omega_0(\psi)(1) \leq \psi(1)$ for every $\psi \in \text{Irr}_{p'}(G|\lambda)$. We take $\Omega := \Omega_0 \circ \Omega_1$ and this bijection satisfies the desired properties. Thus we may assume $LN_G(P) = G$. In particular, L/Z is not a p -group.

If L/Z is a p' -group then we write $Z = Z_p \times Z'_p$ where $Z_p \in \text{Syl}_p(Z)$ and $\mu = \lambda_{Z_p}$ and $\nu = \lambda_{Z'_p}$. Let $K \trianglelefteq L$ be a p -complement of L . If Δ is a G -transversal on the set of P -invariant characters of $\text{Irr}(K|\mu)$ and $*$: $\text{Irr}_P(K) \rightarrow \text{Irr}(C)$ is the P -Glauberman correspondence, where $C = C_K(P)$, then notice that this correspondence restricts to a bijection $*$: $\text{Irr}_P(K|\mu) \rightarrow \text{Irr}(C|\mu)$. We have

$$\text{Irr}_{p'}(G|\lambda) = \bigsqcup_{\theta \in \Delta} \text{Irr}_{p'}(G|\theta \times \nu)$$

and

$$\text{Irr}_{p'}(N_G(P)|\lambda) = \bigsqcup_{\theta \in \Delta} \text{Irr}_{p'}(N_G(P)|\theta^* \times \nu)$$

again applying [N18, Lemma 9.3]. By [T08, Theorem 6.5] and [T09, Theorem 7.12] the character triples (G_θ, K, θ) and $(N_G(P)_{\theta^*}, C, \theta^*)$ are isomorphic, which implies that there is a bijection

$$\Phi_\theta : \text{Irr}(G_\theta|\theta) \rightarrow \text{Irr}(N_G(P)_{\theta^*}|\theta^*)$$

preserving character degree ratios (and in particular, restricting to p' -degree characters). Since $\theta^*(1) \leq \theta(1)$, we have $\Phi_\theta(\psi)(1) \leq \psi(1)$ for all $\psi \in \text{Irr}_{p'}(G_\theta|\theta)$. Further, it follows from the definition of character triple isomorphism that Φ_θ restricts to a bijection

$$\Phi_\theta : \text{Irr}_{p'}(G_\theta|\theta \times \nu) \rightarrow \text{Irr}_{p'}(N_G(P)_{\theta^*}|\theta^* \times \nu)$$

with $\Phi_\theta(\psi)(1) \leq \psi(1)$ for all $\psi \in \text{Irr}_{p'}(G_\theta|\theta \times \nu)$. Since $G_\theta \cap N_G(P) = N_G(P)_{\theta^*}$, using the Clifford correspondence we find a bijection

$$\hat{\Phi}_\theta : \text{Irr}_{p'}(G|\theta \times \nu) \rightarrow \text{Irr}_{p'}(N_G(P)|\theta^* \times \nu)$$

with $\hat{\Phi}_\theta(\chi)(1) \leq \chi(1)$ for all $\chi \in \text{Irr}_{p'}(G_\theta|\theta \times \nu)$ and we are done by defining

$$\Omega : \text{Irr}_{p'}(G|\lambda) \rightarrow \text{Irr}_{p'}(N_G(P)|\lambda)$$

by $\Omega(\chi) := \hat{\Phi}_\theta(\chi)$ where $\theta \in \Delta$ lies below χ .

Therefore we may assume L/Z is semisimple. Let $K = L'$, $Z_1 = Z \cap K = \mathbf{Z}(K)$, and $\nu = \lambda_{Z_1}$. Then K is perfect and satisfies the hypotheses of Theorem 3.4. Let $R = P \cap K$. There exists a $N_G(R)$ -stable subgroup $N_K(R) \leq H < K$ and an $N_G(R)$ -equivariant bijection

$$\Psi : \text{Irr}_{p'}(K) \rightarrow \text{Irr}_{p'}(H)$$

satisfying $\Psi(\eta)(1) \leq \eta(1)$ and inducing central character triple isomorphisms. By the definition of \geq_c we see that Ψ restricts to a bijection

$$\Psi : \text{Irr}_{p'}(K|\nu) \rightarrow \text{Irr}_{p'}(H|\nu)$$

with the same properties. The character triple isomorphisms induce bijections

$$\Phi_\mu : \text{Irr}_{p'}(G_\mu|\mu) \rightarrow \text{Irr}_{p'}(HN_G(R)_{\Psi(\mu)}|\Psi(\mu))$$

which satisfy $\Phi_\mu(\chi)(1) \leq \chi(1)$, and send characters over the central product $\mu \cdot \lambda \in \text{Irr}_{p'}(L)$ to characters over $\Psi(\mu) \cdot \lambda \in \text{Irr}_{p'}(HZ)$. Again, $|G : G_\mu| \geq |HN_G(R) : MN_G(R)_{\Psi(\mu)}|$ so by the Clifford correspondence and the above remark we may find a bijection

$$\hat{\Phi}_\mu : \text{Irr}_{p'}(G|\mu \cdot \lambda) \rightarrow \text{Irr}_{p'}(HN_G(R)|\Psi(\mu) \cdot \lambda)$$

satisfying $\hat{\Phi}_\mu(\chi)(1) \leq \chi(1)$ for all $\chi \in \text{Irr}_{p'}(G|\mu \cdot \lambda)$. Now by taking transversals over the P -invariant characters in $\text{Irr}_{p'}(K|\nu)$ that lie under characters $\chi \in \text{Irr}_{p'}(G)$ and arguing as before we obtain a bijection

$$\Omega_0 : \text{Irr}_{p'}(G|\lambda) \rightarrow \text{Irr}_{p'}(HN_G(R)|\lambda)$$

satisfying $\Omega_0(\chi)(1) \leq \chi(1)$ for every $\chi \in \text{Irr}_{p'}(G|\lambda)$. Since $H < K$ and $LN_G(P) = G$ we have $N_G(P) \leq HN_G(R) < G$ so, by induction we have a bijection

$$\Omega_1 : \text{Irr}_{p'}(HN_G(R)|\lambda) \rightarrow \text{Irr}_{p'}(N_G(P)|\lambda)$$

satisfying $\Omega_1(\psi)(1) \leq \psi(1)$ for all $\psi \in \text{Irr}_{p'}(HN_G(R)|\lambda)$ and the result follows by taking $\Omega := \Omega_1 \circ \Omega_0$. □

Conjecture 3.1 can be recovered by setting $Z = 1$ in Theorem 3.5. By arguing as in Proposition 3.2 and using Theorem 3.5 we obtain a relative version of Conjecture A. If $N \trianglelefteq G$ and $\theta \in \text{Irr}(N)$ recall that we write

$$\text{Irr}_{p', \text{rel}}(G|\theta) = \{ \chi \in \text{Irr}(G|\theta) \mid p \text{ does not divide } \chi(1)/\theta(1) \}.$$

Proposition 3.6. *Let G be a finite group and $P \in \text{Syl}_p(G)$. Let $N \trianglelefteq G$, and let $\theta \in \text{Irr}(N)$ be G -invariant such that θ extends to NP . Assume Conjecture 3.3 holds for every covering group of every simple group of order divisible by p involved in G . Then*

$$\sum_{\chi \in \text{Irr}_{p', \text{rel}}(G|\theta)} (\chi(1)/\theta(1))^2 \geq \sum_{\tau \in \text{Irr}_{p', \text{rel}}(NN_G(P)|\theta)} (\tau(1)/\theta(1))^2,$$

with equality if and only if there is a normal complement K/N to $N_G(P)N/N$ in G/N such that θ has an extension $\hat{\theta} \in \text{Irr}(K)$.

4. Quasisimple groups

In this section and the next, we present evidence supporting Conjecture 3.3. We verify the conjecture for, among other cases, all quasisimple groups of exceptional Lie type with respect to all primes, as well as for all groups of Lie type defined in characteristic equal to the given prime p . These results, in particular, allow us to confirm the conjecture for all quasisimple groups when $p = 2$.

In the cases mentioned, we prove a slightly stronger version of Conjecture 3.3:

Conjecture 4.1. *Let S be a quasisimple group with cyclic center, $P \in \text{Syl}_p(S)$ and $A = \text{Aut}(S)_P$. Then there exists an A -equivariant bijection*

$$\Psi : \text{Irr}_{p'}(S) \rightarrow \text{Irr}_{p'}(N_S(P))$$

such that

$$\chi(1) \geq \Psi(\chi)(1) \quad \text{and} \quad (S \rtimes A_\chi, S, \chi) \geq_c (N_S(P) \rtimes A_\chi, N_S(P), \Psi(\chi))$$

for every $\chi \in \text{Irr}_{p'}(S)$.

The existence of an A -equivariant bijection Ψ from $\text{Irr}_{p'}(S)$ to $\text{Irr}_{p'}(N_S(P))$ satisfying the so-called *central isomorphism*

$$(S \rtimes A_\chi, S, \chi) \geq_c (N_S(P) \rtimes A_\chi, N_S(P), \Psi(\chi))$$

(of character triples) has been established in several papers and was completed in [CS24] by Cabanes and Späth. In fact, by [CS24, Theorem B], we now know that such bijection, which we shall refer to as a *McKay-good* bijection, exists for all finite groups. It is the extra *degree condition* $\Psi(\chi)(1) \leq \chi(1)$ that we need to consider in this paper.

Following the literature, we say that a quasisimple group S satisfies the *inductive McKay condition* if it satisfies Conjecture 3.3, possibly excluding the degree condition. Note that there are other equivalent formulations of the inductive McKay conditions; see, for example, [IMN07, §10]. In some cases, we will work with this version of the inductive condition. We also note that if Conjecture 4.1 holds for a quasisimple group S (with or without cyclic center), then it also holds for every

quotient of S by a central subgroup. For further discussion, we refer the reader to [CS24, Remark 2.9].

Notation 4.2. Let G be a finite group.

- (i) $d(G) := \min\{\chi(1) : \chi \in \text{Irr}(G), \chi(1) > 1\}$ is the smallest nontrivial degree of a (complex) irreducible character of G . By convention, $d(G) = 1$ if G is abelian.
- (ii) $m_p(G) := \min\{\chi(1) : \chi \in \text{Irr}_{p'}(G), \chi(1) > 1\}$ is the smallest nontrivial degree of an irreducible p' -character of G . By convention, $m_p(G) = 1$ if G has no non-linear p' -degree irreducible character.
- (iii) $b_p(G) := \max\{\chi(1) : \chi \in \text{Irr}_{p'}(G)\}$ is the largest degree of an irreducible p' -character of G .
- (iv) We will use $m(G)$ and $b(G)$, respectively, for $m_p(G)$ and $b_p(G)$, whenever p is implicitly known or the presence of p is not important.

Hypothesis 4.3. $m(S) \geq b(N_S(P))$, where S is a quasisimple group, p is a prime, and $P \in \text{Syl}_p(S)$.

Proposition 4.4. Fix a quasisimple group S and a prime p . If Hypothesis 4.3, or the stronger condition $d(S) \geq b(N_S(P))$, is true for S , then so is Conjecture 4.1 for S .

Proof. If $m(S) \geq b(N_S(P))$, then any bijection Ψ from $\text{Irr}_{p'}(S)$ to $\text{Irr}_{p'}(N_S(P))$ sending 1_S to $1_{N_S(P)}$ automatically satisfies the degree condition $\Psi(\chi)(1) \leq \chi(1)$ for every $\chi \in \text{Irr}_{p'}(S)$. □

Hypothesis 4.3, unfortunately, fails quite often. For instance, in general, it fails when S is a cover of an alternating or a simple classical group (in characteristic not equal to p). It also fails for certain groups of Lie type in characteristic p (see the proof of Proposition 4.11).

4.1. Groups of Lie type in characteristic p . The failure of Hypothesis 4.3 when S is a group of Lie type in characteristic p arises in the case of unitary groups. This case requires additional work. Our notation for finite simple groups (and related ones) follows [C85; Atl].

Low-degree irreducible representations of the special unitary groups $\text{SU}_n(q)$ ($q = p^f$ is a power of a prime p and $n \geq 3$, excluding $(n, q) = (3, 2)$) are studied in [TZ96, §4] and [LOST10, §6.1]. Among these are the so-called irreducible Weil characters, denoted by $\zeta_{n,q}^i$ for $0 \leq i \leq q$. The characters $\zeta_{n,q}^i$ with $i > 0$ have degree $(q^n - (-1)^n)/(q + 1)$, while $\zeta_{n,q}^0$ has degree $(q^n + q(-1)^n)/(q + 1)$. In fact, these $\zeta_{n,q}^i$ account for all nontrivial characters of $\text{SU}_n(q)$ with degree at most $d(\text{SU}_n(q)) + 1$. Hence, if p is the defining characteristic of the group, we have

$$m_p(\text{SU}_n(q)) = (q^n - (-1)^n)/(q + 1).$$

For our purposes, we need to construct a non-Weil character of the unitary groups that is invariant under the automorphism groups.

To do so, we first describe the Weil characters of $G := \mathrm{SU}_n(q)$ via the notion of Lusztig series and semisimple characters, as follows. The set $\mathrm{Irr}(G)$ of irreducible characters of G is partitioned into the Lusztig series $\mathcal{E}(G, s)$, where s runs over a complete set of representatives of conjugacy classes of semisimple elements of $G^* := \mathrm{PGU}_n(q)$ (see [GM20, Theorem 2.6.2]). Each $\mathcal{E}(G, s)$ contains one or more special members called semisimple characters whose degrees are equal to $|G^* : \mathbf{C}_{G^*}(s)|_{p'}$ (see [GM20, Definition 2.6.9]).

Centralizers of semisimple elements in classical groups are well known (see [C81; D18] for instance). We recall here the needed result for $\mathrm{GU}_n(q)$.

Lemma 4.5. *Let $G = \mathrm{GU}_n(q)$ and $s \in G$ be a semisimple element. For a monic polynomial $g(t) = t^d + a_{d-1}t^{d-1} + \dots + a_1t + a_0 \in \mathbb{F}_{q^2}(t)$ not equal to t , write $g^*(t) := t^d + (a_1/a_0)q t^{d-1} + \dots + (a_{d-1}/a_0)q t + (1/a_0)q$. Let $f(t)$ be the characteristic polynomial of s and assume that its decomposition into irreducible monic polynomials over \mathbb{F}_{q^2} is*

$$f(t) = \pm \prod_{i=1}^a f_i(t)^{n_i} \times \prod_{j=1}^b (g_j(t)g_j^*(t))^{m_j},$$

where $f_i = f_i^*$ and $g_j \neq g_j^*$ for every i and j and they are pairwise distinct. Let $d_i := \deg(f_i)$ and $e_j := \deg(g_j)$. Then

$$\mathbf{C}_G(s) \cong \prod_{i=1}^a \mathrm{GU}_{n_i}(q^{d_i}) \times \prod_{j=1}^b \mathrm{GL}_{m_j}(q^{2e_j}).$$

Remark 4.6. Irreducible factors of the characteristic polynomial $f(t)$ of a semisimple element in $\mathrm{GU}_n(q)$ are subject to certain restrictions. For example, a polynomial g is a factor if and only if g^* is also a factor. Additionally, any factor g satisfying $g = g^*$ must have odd degree.

Let δ be a generator of the multiplicative group $\mathbb{F}_{q^2}^\times$. For $1 \leq i \leq q$, consider the semisimple element $s_i \in G^*$ to be the image of the diagonal matrix

$$\tilde{s}_i := \mathrm{diag}(\delta^{i(q-1)}, 1^{n-1}) \in \tilde{G} := \mathrm{GU}_n(q)$$

under the natural projection $\pi : \tilde{G} \rightarrow G^*$. For general s , $\mathcal{E}(G, s)$ may contain more than one semisimple character, but we claim that, each $\mathcal{E}(G, s_i)$ contains a unique one.

Note that \tilde{G} is self-dual (in the sense of [C85, §4.3]) and we may identify \tilde{G} with its dual group. As the center of the ambient algebraic group of \tilde{G} is connected, the series $\mathcal{E}(\tilde{G}, \tilde{s}_i)$ contains a unique semisimple character [GM20, p. 171]. We

denote this character by ψ_i . Then $\psi_i(1) = |\tilde{G} : C_{\tilde{G}}(\tilde{s}_i)|_{p'}$. By Lemma 4.5,

$$C_{\tilde{G}}(\tilde{s}_i) \cong \text{GU}_{n-1}(q) \times \text{GU}_1(q).$$

The choice of eigenvalues of s_i 's implies that $C_{G^*}(s_i) = C_{\tilde{G}}(\tilde{s}_i)/\mathbf{Z}\tilde{G}$. (To see this, let $C := \pi^{-1}(C_{G^*}(s_i))$.) Then the mapping $\tau : C \rightarrow \mathbf{Z}\tilde{G}$ defined by $gs g^{-1} = \tau(g)s$ is a homomorphism with $\text{Ker } \tau = C_{\tilde{G}}(\tilde{s}_i)$. However, the fact that s and $gs g^{-1}$ have the same eigenvalues forces $\tau(g) = 1$ for every $g \in C$.) It follows that

$$|G^* : C_{G^*}(s_i)|_{p'} = |\tilde{G} : C_{\tilde{G}}(\tilde{s}_i)|_{p'} = \frac{q^n - (-1)^n}{q + 1},$$

so that semisimple characters in $\mathcal{E}(G, s_i)$ have the same degree as the (only) semisimple character ψ_i in $\mathcal{E}(\tilde{G}, \tilde{s}_i)$. By [GM20, Corollary 2.6.18], semisimple characters in $\mathcal{E}(G, s_i)$ are irreducible constituents of the restriction of ψ_i from \tilde{G} to G^* . We deduce that $\mathcal{E}(G, s_i)$ has a unique semisimple character, as claimed. This is the Weil character $\zeta_{n,q}^i$ mentioned above.

Lemma 4.7. *Assume the above notation. Then a Weil character $\zeta_{n,q}^i$ for some $1 \leq i \leq q$ is invariant under all field automorphisms of G if and only if $i(p - 1)$ is divisible by $q + 1$. In particular, if $p = 2$, then every $\zeta_{n,q}^i$ is moved by some automorphism of G .*

Proof. The character $\zeta_{n,q}^i$ is invariant under all field automorphisms of G if and only if $\delta^{i(q-1)} = \delta^{ip(q-1)}$, which means that $i(q - 1) \equiv ip(q - 1) \pmod{q^2 - 1}$, and the first statement follows. When $p = 2$, each $\zeta_{n,q}^i$ (for $1 \leq i \leq q$) is moved by some field automorphisms of G , as desired. \square

Lemma 4.8. *Let $G = \text{SU}_n(q)$ where $n \geq 3$ is odd, $(n, q + 1) = 1$, and $q = p^f$ for some prime p . Then G possesses an irreducible character χ of degree not divisible by p such that χ is $\text{Aut}(G)$ -invariant and $\chi(1) > q^{n-1}$.*

Proof. By the hypothesis, G is simple and self-dual. By [B09, Lemma 5], p' -degree irreducible characters of G are precisely the semisimple characters of G , which in turn can be labeled by conjugacy classes of semisimple elements of $G^* := \text{PGU}_n(q) \cong G$.

It is well known that $\text{Aut}(G)$ permutes the Lusztig series of G . In our situation, by identifying the automorphisms of G with the corresponding automorphisms of G^* under the natural isomorphism, every $\varphi \in \text{Aut}(G)$ maps $\mathcal{E}(G, s)$ to $\mathcal{E}(G, \varphi(s))$ (see [T18, Proposition 7.2]). Since each $\mathcal{E}(G, s)$ contains a unique semisimple character (using again [GM20, p. 171] and the fact that the ambient algebraic group of $\text{PGU}_n(q)$ has connected center), of degree $|G^* : C_{G^*}(s)|_{p'}$, it follows that a p' -degree irreducible characters of G is $\text{Aut}(G)$ -invariant if and only if the semisimple conjugacy class labeling it is $\text{Aut}(G^*)$ -invariant. Therefore, the result follows if we are able to produce a semisimple element $s \in G^*$ such that its G^* -conjugacy

class is $\text{Aut}(G^*)$ -invariant and $|G^* : C_{G^*}(s)|_{p'} > q^{n-1}$. Recall that $\text{Out}(G^*) \cong C_{2f}$ is the cyclic group of order $2f$ consisting of the field automorphisms of G^* .

Suppose first that q is odd. Let s be the image of a diagonal matrix $\tilde{s} \in \tilde{G} := \text{GU}_n(q)$ with spectrum $\text{Spec}(\tilde{s}) = \{1, \dots, 1, -1, -1\}$ under the projection $\pi : \tilde{G} \rightarrow G^*$ mentioned earlier. Note that n is odd. Similar arguments as above show that $C_{\tilde{G}}(\tilde{s})$ is the complete inverse image of $C_{G^*}(s)$. Moreover, $C_{\tilde{G}}(\tilde{s}) \cong \text{GU}_{n-2}(q) \times \text{GU}_2(q)$. It follows that

$$|G^* : C_{G^*}(s)|_{p'} = |\tilde{G} : C_{\tilde{G}}(\tilde{s})|_{p'} = \frac{(q^n + 1)(q^{n-1} - 1)}{(q^2 - 1)(q + 1)}.$$

Observe that, as \tilde{s} is invariant under all field automorphisms, s is $\text{Aut}(G^*)$ -invariant. The required character χ can be taken to be the (only) semisimple character in the Lusztig series associated to s .

Assume now that $q = 2^f$ with f even. We then take \tilde{s} to be a diagonal matrix with $\text{Spec}(\tilde{s}) = \{\delta^{(q^2-1)/3}, \delta^{2(q^2-1)/3}, 1^{n-2}\}$ instead, where δ , as before, is a generator of \mathbb{F}_q^\times . Again, by Lemma 4.5, we have $C_{\tilde{G}}(\tilde{s}) \cong \text{GU}_{n-2}(q) \times \text{GL}_1(q^2)$ and $C_{\tilde{G}}(\tilde{s})$ is the complete inverse image of $C_{G^*}(s)$. (This is clear when $n > 3$. When $n = 3$, the image of the homomorphism $\tau : \pi^{-1}(C_{G^*}(s)) \rightarrow \mathbf{Z}\tilde{G}$ defined by $gsg^{-1} = \tau(g)s$ is contained in $\{\text{Id}, \delta^{(q^2-1)/3}\text{Id}, \delta^{2(q^2-1)/3}\text{Id}\}$. However, both $\delta^{(q^2-1)/3}\text{Id}$ and $\delta^{2(q^2-1)/3}\text{Id}$ have order 3, which does not divide $|\mathbf{Z}\tilde{G}| = q + 1$. Therefore we still have $\tau(g) = \text{Id}$ for all $g \in \pi^{-1}(C_{G^*}(s))$.) We then have

$$|G^* : C_{G^*}(s)|_{p'} = \frac{(q^n + 1)(q^{n-1} - 1)}{(q^2 - 1)(q - 1)}.$$

It is straightforward to check that this is greater than q^{n-1} , and we conclude as in the previous case.

We solve the remaining case $q = 2^f$ with f odd by a somewhat different argument, working with the ambient algebraic group of G and its associated Frobenius map instead. So let \mathbf{G} be a simple algebraic group of adjoint type of type A , defined over an algebraically closed field of characteristic 2. Let F_f be the standard Frobenius map on \mathbf{G} , raising all matrix entries to the 2^f -th power and ρ the inverse transpose. Set $F := F_f \circ \rho$. Then $G = G^*$ is precisely the group of F -fixed points in \mathbf{G} . Let $F_0 := F_1 \circ \rho$. As f is odd, we have $F = F_f \circ \rho = F_0^f$. Note that F_1 induces a generator, say σ_{F_1} , for $\text{Out}(G) \cong C_{2f}$; on the other hand, F_0 induces $\sigma_{F_0} \in \text{Out}(G)$ of order f and $\langle \sigma_{F_0} \rangle = \langle (\sigma_{F_0})^2 \rangle = \langle (\sigma_{F_1})^2 \rangle$. By the work of Brunat [B09, Proposition 2], the number of σ_{F_0} -invariant irreducible characters of G of odd degree is equal to the number of semisimple characters of $\mathbf{G}^{F_0} = \text{PGU}_n(2) \cong \text{SU}_n(2)$. (Note that, by the assumptions, f is odd and $(n, 2^f + 1) = 1$. In particular, n is coprime to 3.) This number, in turn, is 2^{n-1} , by [C85, Corollary 8.3.6]. These 2^{n-1} characters are permuted by σ_{F_1} , forming orbits of size 1 or 2. The trivial character forms

its own orbit. Consequently, at least one of the nontrivial characters, say χ , is fixed by σ_{F_1} – that is, χ is $\text{Out}(G)$ -invariant. By Lemma 4.7, χ cannot be a Weil character. Moreover, under the assumptions on n, q , and f , we have $n \geq 5$. It then follows from [TZ96, Table V] that $\chi(1) \geq (q^n + 1)(q^{n-1} - q^2)/(q + 1)(q^2 - 1)$, and therefore $\chi(1) > q^{n-1}$. This concludes the proof. \square

Lemma 4.9. *Let $G = \mathbf{G}^F$ be the group of fixed points of a simple algebraic group \mathbf{G} of simply connected type over an algebraically closed field of characteristic p , under a Steinberg endomorphism F of \mathbf{G} , and assume that G/\mathbf{ZG} is simple. Let $P \in \text{Syl}_p(G)$. Then*

$$b(N_G(P)) \leq |T/\mathbf{ZG}|,$$

where $T := \mathbf{T}^F$ and \mathbf{T} is a maximally split F -stable maximal torus of \mathbf{G} .

Proof. Let \mathbf{B} be an F -stable Borel subgroup of \mathbf{G} containing \mathbf{T} . Let Φ be the root system of \mathbf{G} with respect to \mathbf{T} and \mathbf{B} . Let Φ^+ and $\Delta = \{\alpha_i : i \in I\}$ be the corresponding set of positive roots and simple roots, respectively. Also, let X_α denote the root subgroup associated to each $\alpha \in \Phi$.

Let $U := \prod_{\alpha \in \Phi^+} X_\alpha$, which is the unipotent radical of \mathbf{B} . According to [MT11, Corollary 24.11], we have $P := U^F \in \text{Syl}_p(G)$ and $\mathbf{B}^F = N_G(P)$. Furthermore,

$$N_G(P) = P \rtimes T,$$

where $T := \mathbf{T}^F$. At this point we have $b(N_G(P)) \leq |N_G(P)/P| = |T|$ but, in many cases, this upper bound for $b(N_G(P))$ is not sufficient. We need to know more details about the action of T on $\text{Irr}(P/P')$.

Let $U_c := \prod_{\alpha \in \Phi^+ \setminus \Delta} X_\alpha$. As explained in [C85, §2.9], U_c is normal in U and $U/U_c = \prod_{i \in I} X_{\alpha_i}$. The endomorphism F naturally acts on the roots, and thus on the root subgroups, given by $F(X_\alpha) = X_{F(\alpha)}$. As \mathbf{B} and \mathbf{T} are F -stable, F permutes the positive roots, as well as the simple roots. Both U and U_c are therefore F -stable. Let ρ denote the permutation on I induced from the action of F on the simple roots, and \mathcal{O} be the set of ρ -orbits on I . For each such an orbit J , let $X_J := \prod_{i \in J} X_{\alpha_i}$, which is an F -stable group. Further, let $X_J := X_J^F$. We then have

$$(4-1) \quad U^F/U_c^F = \prod_{J \in \mathcal{O}} X_J.$$

Note that each X_α is normalized by \mathbf{T} (see [C85, p. 18]). Thus U^F and U_c^F are both normalized by $T = \mathbf{T}^F$, and so T acts on the factor group U^F/U_c^F . In fact, each direct factor X_J with $J \in \mathcal{O}$ of U^F/U_c^F is normalized by T . Remark that $\mathbf{ZG} \leq T$ (see the proof of [MT11, Lemma 24.12]) and \mathbf{ZG} acts trivially on U^F . Therefore the maximal size of a T -orbit on U^F/U_c^F is at most $|T/\mathbf{ZG}|$.

Assume from now on that $G \notin \{\text{Sp}_{2n}(2), F_4(2), G_2(3)\}$. (For these exceptions, a Sylow 2-subgroup of G is self-normalizing and the desired inequality is immediate.)

Then, according to [B09, Lemma 5], U_c^F is the derived subgroup P' of $P = U^F$. As the actions of T on P/P' and $\text{Irr}(P/P')$ are isomorphic, $|T/\mathbf{Z}G|$ is also an upper bound for the sizes of the T -orbits on $\text{Irr}(P/P')$. Moreover, the restriction of every irreducible character of $P/P' \rtimes T$ to P/P' is multiplicity-free, as T is abelian and $|T|$ is coprime to p . The result now follows. \square

Lemma 4.10. *Assume the hypotheses of Lemma 4.9. Then $S = G/\mathbf{Z}G$ satisfies Hypothesis 4.3, unless $G = \text{SU}_n(q)$ with $n \geq 3$ odd and $(n, q + 1) = 1$.*

Proof. Clearly, $m(G) \leq m(S)$ and $b(N_G(Q)) \geq b(N_S(P))$ for $P \in \text{Syl}_p(S)$ and $Q \in \text{Syl}_p(G)$. Therefore, if Hypothesis 4.3 holds for G , then it also holds for S . Sylow p -subgroups and their normalizers of a finite reductive group in characteristic p are best described through the framework of Borel subgroups and their unipotent radicals in the ambient algebraic group. For convenience, we shall use the same P for a Sylow p -subgroup of G (in fact, $P \cong Q$, see [MT11, p. 214]).

Note that the group X_J in the proof of Lemma 4.9 is isomorphic to $\mathbb{F}_{q^{|J|}}$ and $|X_J| = q^{|J|}$, where q is the absolute value of all eigenvalues of F on the character group of an F -stable maximal torus of G . Furthermore, as noted in [C85, p. 74], we have

$$|T| = \prod_{J \in \mathcal{O}} (q^{|J|} - 1).$$

Lemma 4.9 therefore implies that

$$b(N_G(P)) \leq \frac{\prod_{J \in \mathcal{O}} (q^{|J|} - 1)}{|\mathbf{Z}G|}.$$

The ρ -action on the set of the simple roots, defined in the proof of Lemma 4.9, for all the relevant (G, F) is described in [C85, p. 37], allowing one to easily determine the sizes of ρ -orbits. Generally, $|J| \in \{1, 2, 3\}$ for all (G, F) and when G is untwisted, $|J| = 1$ for all J . On the other hand, the values of the smallest nontrivial p' -degree $m(G)$ can be read off from [TZ96] and [N10] when G is of classical type and those for exceptional types can be found in [Lü01].

Consider $G = \text{SL}_2(q)$ with q odd. Then $|T| = q - 1$ and so $b(N_G(P)) \leq (q - 1)/2$. On the other hand, $m(G) = (q - 1)/2$ and so Hypothesis 4.3 is satisfied. (In this case, in fact, $b(N_G(P)) = m(G) = (q - 1)/2$. Each element t of $T \cong \mathbb{F}_q^\times$ acts on $P \cong \mathbb{F}_q^+$ by mapping u to ut^2 , and so the stabilizer in T of any nontrivial (linear) character of P is precisely the order-2 subgroup of T . Therefore $N_G(P)$ has $q - 1$ linear characters and four characters of degree $(q - 1)/2$.) Similarly, for $G = \text{SL}_2(q)$ with q even, one has $m(G) = q - 1 = |T|$ and Hypothesis 4.3 is still valid.

Let $G = \text{SL}_n(q)$ with $n \geq 3$. Here $|T| = (q - 1)^{n-1}$ and $m(G) = (q^n - 1)/(q - 1)$, unless $(n, q) = (4, 3)$. Note that $m(\text{SL}_4(3)) = 26$. In any case, we have $|T| \leq m(G)$, and Hypothesis 4.3 is verified.

Let $G = \text{SU}_n(q)$ with $n \geq 3$. Then

$$|T| = \begin{cases} (q^2 - 1)^{(n-1)/2} & \text{if } n \text{ is odd,} \\ (q^2 - 1)^{(n-2)/2}(q - 1) & \text{if } n \text{ is even,} \end{cases}$$

and, as mentioned,

$$m(G) = (q^n - (-1)^n)/(q + 1).$$

We now can verify that Hypothesis 4.3 holds when n is even or n is odd and $|ZG| = (n, q + 1) > 1$.

For $G = \text{Sp}_{2n}(q)$ or $\text{Spin}_{2n+1}(q)$ with $n \geq 2$, we have $|T| = (q - 1)^n$ and $b(N_G(P)) \leq (q - 1)^n/(2, q - 1)$. One can verify from Table 1 that, $d(G)$, which serves as a lower bound for $m(G)$, is always at least $(q - 1)^n/(2, q - 1)$. Similarly, for $G = \text{Spin}_{2n}^\epsilon(q)$ with $\epsilon \in \{\pm\}$ and $n \geq 4$, we have $b(N_G(P)) \leq (q - 1)^{n-1}(q - \epsilon)/4$, while $m(G) \geq (q^n - 1)(q^{n-1} - 1)/2(q + 1)$, by [N10, Theorems 1.3 and 1.4]. Hypothesis 4.3 is again satisfied.

Let $G = {}^2B_2(q^2)$ or ${}^2G_2(q^2)$, where $q^2 = 2^{2n+1}$ or $q^2 = 3^{2n+1}$, respectively. In these cases, $|T| = q^2 - 1$, which is less than $m(G)$. (For type 2B_2 , $m(G)$ is at least $\sqrt{1/2}q(q^2 - 1)$, while for type 2G_2 , $m(G) = q^4 - q^2 + 1$.) Similarly, when $G = {}^2F_4(q^2)$, we have $|T| = (q^2 - 1)^2$ and $m(G) \geq \sqrt{1/2}q^9(q^2 - 1)$. In all these cases, the required inequality holds.

For the remaining exceptional-type groups, we always have $|T| \leq q^{rk(G)}$, where $rk(G)$ is the semisimple rank of G . However, the lower bound for the smallest nontrivial irreducible representation of G , as shown in Table 1, confirms that $m(G) > q^{rk(G)}$. □

Proposition 4.11. *Conjecture 4.1 is true when S is a quotient of a non-exceptional covering group of a simple group of Lie type in characteristic p .*

Proof. We assume that S is not the Tits group ${}^2F_4(2)'$, since Hypothesis 4.3 can be verified directly for this group using GAP 4.11.0 (<http://www.gap-system.org>). By Lemma 4.10 and Proposition 4.4, it remains to consider only the case where $G = \text{SU}_n(q)$ with $n \geq 3$ odd and $(n, q + 1) = 1$.

In this case, G can be viewed as a group of adjoint type and hence, by [C85, §2.9],

$$T \cong \prod_{J \in \mathcal{O}} T_J,$$

the direct product of cyclic groups $T_J \cong C_{q^2-1}$. (Here, every orbit J , defined in the proof of Lemma 4.9, has length 2.) Further, the action of T on $P/P' \cong \prod_{J \in \mathcal{O}} X_J$ (see (4-1)) is a “product” action, in the way that T_J acts trivially on $X_{J'}$ if $J \neq J'$ and transitively on $X_J \setminus \{1\}$. Therefore T has a (unique) regular orbit on P/P' , as well as on $\text{Irr}(P/P')$, implying that $b(N_G(P)) = |T|$ and $N_G(P)$ has a unique p' -degree irreducible character of degree $|T| = (q^2 - 1)^{(n-1)/2}$. We shall denote

this character by τ . Every other p' -degree irreducible character of $N_G(P)$ restricts trivially to at least one of X_J 's, and hence has degree at most $|T|/(q^2 - 1)$.

Note that, as $(n, q + 1) = 1$, G is simple and the group, say A , of outer automorphisms of G is cyclic and stabilizes (the unipotent subgroup) P . Moreover, a bijection between $\text{Irr}_{p'}(G)$ and $\text{Irr}_{p'}(N_G(P))$ is McKay-good if and only if it is A -equivariant. The existence of such a bijection was established in [B09; S12]. In other words, we know that, for every subgroup $B \leq A$, the numbers of B -fixed characters in $\text{Irr}_{p'}(G)$ and $\text{Irr}_{p'}(N_G(P))$ are the same. Note that the above-mentioned character τ of $N_G(P)$ is A -invariant (due to its uniqueness property). Let ξ be the character of G produced by Lemma 4.8; in particular, ξ is p' -degree and A -invariant. Now the numbers of B -fixed characters in $\text{Irr}_{p'}(G) \setminus \{\xi\}$ and $\text{Irr}_{p'}(N_G(P)) \setminus \{\tau\}$ are the same for every $B \leq A$. Using [I06, Lemma 13.23], one can construct an A -equivariant bijection from $\text{Irr}_{p'}(G) \setminus \{\xi\}$ to $\text{Irr}_{p'}(N_G(P)) \setminus \{\tau\}$, which can be extended to an A -equivariant bijection $\Psi : \text{Irr}_{p'}(G) \rightarrow \text{Irr}_{p'}(N_G(P))$ such that $\Psi(\xi) = \tau$.

We claim that $\chi(1) \geq \Psi(\chi)(1)$ for all $\chi \in \text{Irr}_{p'}(G)$. First, observe that

$$\xi(1) > q^{n-1} > (q^2 - 1)^{(n-1)/2} = \tau(1).$$

On the other hand, when $\chi \neq \xi$, we have $\Psi(\chi) \neq \tau$, and so $\Psi(\chi)(1) \leq |T|/(q^2 - 1) = (q^2 - 1)^{(n-3)/2}$, which implies that

$$\chi(1) \geq \frac{q^n - (-1)^n}{q + 1} \geq \Psi(\chi)(1),$$

and this finishes the proof. \square

4.2. Groups of Lie type in characteristic not equal to p .

Lemma 4.12. *Let $H \leq G$ be such that $|G : H|$ is not divisible by p . Then $b(H) \leq b(G)$.*

Proof. Let $\varphi \in \text{Irr}_{p'}(H)$ such that $\varphi(1) = b(H)$. By Frobenius reciprocity, φ is contained in χ_H for any irreducible constituent χ of φ^G . Since $\varphi^G(1) = \varphi(1)|G : H|$ is not divisible by p , at least one of those constituents has p' -degree. \square

Proposition 4.13. *Conjecture 4.1 is true when S is a non-exceptional covering group of a simple group of exceptional Lie type in characteristic different from p .*

Proof. Assume for now that S is not of Suzuki or Ree type. As before, S is a quotient of $G := \mathbf{G}^F$, where \mathbf{G} is a simple algebraic group of simply connected type and F is a Steinberg endomorphism on \mathbf{G} . Let q be the absolute value of all eigenvalues of F on the character group of an F -stable maximal torus of \mathbf{G} . We know that p does not divide q , and we may assume that p divides $|G|$.

We require some d -Harish-Chandra theory, particularly the concept of Sylow d -tori, which was first introduced by Broué and Malle in [BM92]. (For a detailed

G	conditions	$d(G)$
$SL_2(q)$	$q \geq 5$	$(q-1)/(2, q-1)$
$SL_n(q)$	$n \geq 3$ $(n, q) \notin \{(3, 2), (3, 4), (4, 2), (4, 3)\}$	$(q^n - 1)/(q - 1)$
$SU_n(q)$	$n \geq 3$ odd; $(n, q) \neq (3, 2)$ $n \geq 4$ even; $(n, q) \notin \{(4, 2), (4, 3)\}$	$(q^n - q)/(q + 1)$ $(q^n - 1)/(q + 1)$
$Sp_{2n}(q)$	$n \geq 2$; q odd $n \geq 2$; q even; $(n, q) \neq (2, 2)$	$(q^n - 1)/2$ $(q^n - 1)(q^n - q)/2(q + 1)$
$Spin_{2n+1}(q)$	$n \geq 3$; $q > 3$ odd $n \geq 3$; $q = 3$	$(q^{2n} - 1)/(q^2 - 1)$ $(q^n - 1)(q^n - q)/(q^2 - 1)$
$Spin_{2n}^+(q)$	$n \geq 4$; $q > 3$ $n \geq 4$; $q \in \{2, 3\}$; $(n, q) \neq (4, 2)$	$(q^n - 1)(q^{n-1} + q)/(q^2 - 1)$ $(q^n - 1)(q^{n-1} - 1)/(q^2 - 1)$
$Spin_{2n}^-(q)$	$n \geq 4$	$(q^n + 1)(q^{n-1} - q)/(q^2 - 1)$
${}^2B_2(q^2)$	$q^2 = 2^{2f+1} \geq 8$	$\sqrt{1/2}q(q^2 - 1)$
${}^2G_2(q^2)$	$q^2 = 3^{2f+1} \geq 27$	$q^4 - q^2 + 1$
${}^2F_4(q^2)$	$q^2 = 2^{2f+1} \geq 8$	$\sqrt{1/2}q^9(q^2 - 1)$
$G_2(q)$	$q \geq 3$	$\geq q^3 - 1$
${}^3D_4(q)$		$\geq q^3(q^2 - 1)$
$F_4(q)$		$\geq q^8 + q^4 + 1$
$E_6(q)_{sc}$		$\geq q^9(q^2 - 1)$
${}^2E_6(q)_{sc}$		$\geq q^9(q^2 - 1)$
$E_7(q)_{sc}$		$\geq q^{15}(q^2 - 1)$
$E_8(q)$		$\geq q^{27}(q^2 - 1)$

Table 1. Values or bounds for the minimal nontrivial degree of ordinary characters of finite reductive groups of simply connected type [TZ96; Lü01].

account, see also [GM20, §3.5].) Define e as the multiplicative order of q modulo p if $p > 2$ or if $p = 2$ and $q \equiv 1 \pmod{4}$. For $p = 2$ and $q \equiv -1 \pmod{4}$, let $e := 2$. As $p \mid |G|$, we know that $\Phi_e(q) \mid |G|$, where Φ_e is the e -th cyclotomic polynomial. Let S_e be a Sylow e -torus of G . It is, by definition, an F -stable torus of G whose order polynomial is the maximal power of Φ_e dividing the generic order of G (see [GM20, p. 259]).

Let $L_e := C_G(S_e)$, known as a (minimal) e -split Levi subgroup of G . It is F -stable (see [GM20, p. 258]) and we write $L_e := L_e^F$. Note that, by the conjugation property of Sylow d -tori, we have $N_G(S_e) = N_G(L_e)$. The quotient $W_G(L_e) := N_G(L_e)/L_e^F$ is referred to as the *relative Weyl group* of L_e in G .

According to [M07, Proposition 5.21], $N_G(S_e)$ contains a Sylow p -subgroup, say P , of G . In fact, by [M07, Theorem 7.8], $N_G(S_e)$ contains $N_G(P)$, unless

$p = 3$ and $G = G_2(q)$ with $q \equiv 2, 4, 5, \text{ or } 7 \pmod{9}$. For now, let us exclude this exception. It follows from Lemma 4.12 that

$$b(N_G(P)) \leq b(N_G(L_e)).$$

Suppose first that e is a regular number for (G, F) , which means that L_e is a maximal torus of G (see [GM20, p. 259]). Then $L_e := L_e^F$ is abelian, and thus $b(N_G(L_e)) \leq |N_G(L_e)/L_e|$, by [I06, Corollary 11.29]. In summary, if e is a regular number for (G, F) and G is not of type G_2 , then

$$b(N_G(P)) \leq |W_G(L_e)|.$$

Consider the case when the power of Φ_e in the order polynomial of G is precisely 1, or equivalently, where $W_G(L_e)$ is cyclic (see [GM20, Proposition 3.5.12]). As computed in [BM93, Table 8.1], with one exception at type E_8 and $e = 30$, we have $|W_G(L_e)| \leq 24$ for all relevant G and e . One can easily check that the minimal character degree $d(G)$ of G , displayed in Table 1, is always at least 26, and so we are done. For the exception, we have $|W_G(L_e)| = 30$, while $d(G) \geq q^{27}(q^2 - 1)$ and so Hypothesis 4.3 is satisfied. On the other hand, non-cyclic relative Weyl groups of minimal e -split Levi subgroups for exceptional types are available in [GM20, Table 3.2]. We have verified that $|W_G(L_e)|$ remains at most $d(G)$, except for the specific cases discussed in the next paragraph.

Consider $G = E_8(q)$, for instance. Then $|W_G(L_e)| \leq d(G)$ unless $e = 2, q = 2$, and $p = 3$. For the exception, $W_G(L_e)$ is isomorphic to the Weyl group $W(E_8) = C_2.GO_8^+(2)$ of E_8 , which has order 696729600, while the smallest nontrivial degree $d(E_8(2))$ of $E_8(2)$ is 545925250. However, since $b_3(N_G(L_e))$ divides $|W_G(L_e)|$, by [I06, Corollary 11.29], we have $b_3(N_G(L_e)) \leq 696729600_3 = 2867200 < d(G)$, and the result follows as before. Other exceptions occur for $(G, q, p, e) = (E_7(q)_{sc}, 2, 3, 2)$ or $({}^2E_6(2)_{sc}, 2, 3, 2)$, but the arguments are entirely similar.

Next we consider the case where e is not a regular number for (G, F) . This includes $e = 5$ for type E_6 ; $e = 10$ for 2E_6 ; $e \in \{4, 5, 8, 10, 12\}$ for E_7 ; and $e \in \{7, 9, 14, 18\}$ for E_8 . Note that L_e is no longer abelian. But, as $S_e := S_e^F$ is abelian, we shall use the bound

$$b(N_G(P)) \leq b(N_G(L_e)) \leq |N_G(L_e) : S_e| = |W_G(L_e)||L_e : S_e|$$

instead.

Let $G = E_7(q)_{sc}$ and $e = 4$. Then $|N_G(L_e)/L_e| = 96$ and L_e has type $\Phi_4^2 A_1(q)^3$ (see [GM20, Table 3.3]). Thus $|N_G(L_e) : S_e| \leq 96q^3(q^2 - 1)^3$, which is smaller than $d(E_7(q)_{sc})$, as desired. In the remaining cases, $W_G(L_e)$ is cyclic and its order together with the structure of L_e are again given in [BM93, Table 8.1]. Note that, in this case, the Sylow e -torus S_e has order $\Phi_e(q)$. It is now straightforward to check that $|W_G(L_e)||L_e : S_e| \leq d(G)$ for all relevant G and e .

For Suzuki and Ree groups, Broué and Malle introduced an adapted version of Φ_e , denoted by $\Phi_e^{(p)}$, which are cyclotomic polynomials over $\mathbb{Q}(\sqrt{2})$ or $\mathbb{Q}(\sqrt{3})$ (see [M07, §8]). For the Tits group, we can verify Hypothesis 4.3 directly using GAP; therefore, we assume that $S \neq {}^2F_4(2)'$. With this, there exists a Sylow $\Phi_e^{(p)}$ -torus S of G such that $N_G(S) \geq N_G(P)$ for some Sylow p -subgroup P of G , unless

- (i) $p = 2$ and $G = {}^2G_2(q)$, or
- (ii) $p = 3$ and $G = {}^2F_4(q)$ with $q \equiv 2, 5 \pmod{9}$.

(See [M07, Theorem 8.4].) The case when such a Sylow $\Phi_e^{(p)}$ -torus exists can be argued similarly as above. (Alternatively, one can use [IMN07, §16 and §17] and [A98] to achieve the result for these groups.) The case $p = 2$ and $G = {}^2G_2(q)$ follows from the proof of the McKay inductive conditions for the group (see [IMN07, §17]). When $p = 3$ and $G = {}^2F_4(q)$, according to [A98, (2B)], we have $|N_G(P)/P| \leq 48$, and the same reasoning applies.

Finally, the case $p = 3$ and $G = G_2(q)$ (that we previously excluded) follows from An's proof of the Alperin-McKay conjecture for $G_2(q)$. (See [A94, p. 190] where it was shown that $|N_G(P)/P|$ is bounded above by 16, which is smaller than $d(G_2(q))$.) \square

4.3. Exceptional covering groups. Here we deal with *exceptional covering groups* of finite simple groups. These include 3-fold and 6-fold covers of A_6 and A_7 , covers of sporadic simple groups (by convention), and certain covers of simple groups of Lie type with a non-generic Schur multiplier (see [MT11, Table 24.3]). Here, a (perfect central) cover S of a simple group of Lie type X is called exceptional if it is not a quotient of the finite reductive group of simply connected type covering X ; in particular, S is a proper cover of X .

Proposition 4.14. *Conjecture 4.1 is true when S is an exceptional covering group.*

Proof. The existence of a McKay-good bijection for exceptional covering groups was established by Malle [M08]. We note that, except the single case $S = 2 \cdot \text{PSL}_3(4)$, the group of outer automorphisms stabilizing $P \in \text{Syl}_p(S)$ is either trivial or cyclic of prime order, and so any bijection that respects the action of those outer automorphisms is McKay-good. Building on Malle's work, we show that in most cases, Hypothesis 4.3 holds, ensuring that any such bijection automatically satisfies the additional degree condition. In the remaining cases, it turns out that S has precisely one character of degree smaller than $b(N_S(P))$. In these instances, it suffices to identify a corresponding (p' -degree) character of $N_S(P)$ with smaller degree that is invariant under the action of the outer automorphisms.

Throughout we let $X := S/Z$ and $Q := PZ/Z \in \text{Syl}_p(X)$, where $Z := ZS$. Of course we may assume that $p \mid |S|$.

First suppose that $|S|_p = p$; in particular, p is odd. In this case, work of Dade [D66; D96] provides a natural bijection between the irreducible characters in any block B of S and those in its Brauer correspondent b , and this bijection satisfies the required degree condition. We may assume that B has full defect. When $|S|_p = p$, Dade's bijection is described in Lemmas 4.7–4.10 of [D96] and is known, in particular, to preserve decomposition numbers. Hence it suffices to verify the desired degree inequality for the corresponding bijection between the Brauer characters of B and b given in [D96, Lemma 4.7].

This bijection between Brauer characters is defined as follows. The Green correspondence sends the isomorphism classes of finite-dimensional non-projective indecomposable kS -modules M belonging to B bijectively onto the isomorphism classes of finite-dimensional non-projective indecomposable $kN_S(P)$ -modules \tilde{M} , where k is a suitable residue field of characteristic p (see [F82, III.5]). Moreover, this correspondence satisfies the degree inequality $\dim(M) \geq \dim(\tilde{M})$. If M is simple, the socle $S(\tilde{M})$ of \tilde{M} lies in a uniquely determined isomorphism class of simple $kN_S(P)$ -modules. Dade's bijection then sends the Brauer character afforded by M to that afforded by $S(\tilde{M})$, and therefore also satisfies the degree inequality. (We also note that in a recent preprint [Li25], Linckelmann proved that, when a defect group of B is cyclic in general, there exists a perfect isometry between $\mathbb{Z} \text{Irr}(B)$ and $\mathbb{Z} \text{Irr}(b)$ with the degree condition.) Furthermore, Koshitani and Späth [KS16] proved that Dade's bijection fulfills the inductive Alperin–McKay condition, and hence is McKay-good. We assume from now on that $p^2 \mid |S|$.

(I) Consider the sporadic groups. The structure of the quotient group $N_X(Q)/Q'$ is given in [W98]. When $X \in \{Fi'_{24}, B, M\}$, we have checked that $m(S) \geq |Z||N_X(Q)/Q| \geq |N_S(P)/P|$, thereby confirming that Hypothesis 4.3 holds. For the remaining sporadic groups, computations using GAP reveal that Hypothesis 4.3 fails in the following cases. We include here the relevant values of $m(S)$ and $b(N_S(P))$:

$$(S, p, b(N_S(P)), m(S)) \in \{(Co_2, 5, 24, 23), (Co_3, 3, 32, 23), (Co_3, 5, 24, 23), \\ (McL, 5, 24, 22), (3 \cdot McL, 5, 24, 22)\}.$$

When $S \in \{Co_2, Co_3\}$, as the outer automorphism group of S is trivial, any bijection (from $\text{Irr}_{p'}(S)$ to $\text{Irr}_{p'}(N_S(P))$) is McKay-good. In these cases, S has a unique irreducible p' -degree character, say ψ , of degree smaller than $b(N_S(P))$. Furthermore, we observe that $N_S(P)$ is not perfect. (Indeed, $N_S(P)/P' \cong 5^2 \rtimes (4 \cdot S_4)$ when $S = Co_2$, and $N_S(P)/P' \cong S_3 \times (3^2 : SD_{16})$ when $S = Co_3$.) Consequently, $N_S(P)$ has a nontrivial linear character, say τ . It then follows that any bijection sending ψ to τ satisfies the required degree condition.

Let $(S, p) \in \{(McL, 5), (3 \cdot McL, 5)\}$. In both cases, the outer automorphism group $\text{Out}(S) \cong C_2$ normalizes P . Moreover, S has exactly one irreducible character of degree less than $b(N_S(P)) = 24$, and this character is necessarily $\text{Out}(S)$ -invariant. Suppose first that $S = McL$. By [Atl], $N_S(P) = P \rtimes D$, where P is an extra-special group of order 5^3 and exponent 5, and $D = C_3 \rtimes C_8$. The group D acts Frobeniusly on $\text{Irr}(P/P') \cong C_5 \times C_5$, and hence $N_S(P)$ has a unique irreducible character of degree 24. This is the only $5'$ -degree character of $N_S(P)$ of degree greater than $m(S) = 22$. All other $5'$ -degree irreducible characters of $N_S(P)$ are trivial on P , and thus can be viewed as characters of D , with degrees 1 and 2. Now, the required bijection can be constructed in the way that this degree-24 character corresponds to (any) $\text{Out}(S)$ -invariant irreducible character of S of degree greater than 22. (The group McL indeed has several such characters.) When $S = 3 \cdot McL$, again $\text{Irr}_{5'}(N_S(P))$ contains a single member of degree 24. As in the case $S = McL$, this is the only $5'$ -degree character of $N_S(P)$ of degree greater than $m(S) = 22$, and the same reasoning applies.

(II) The simple groups of Lie type with a non-generic Schur multiplier are

$$\text{PSL}_2(4), \text{PSL}_2(9), \text{PSL}_3(2), \text{PSL}_3(4), \text{PSL}_4(2), \text{PSU}_4(2), \text{PSU}_4(3), \\ \text{PSU}_6(2), {}^2B_2(8), \text{PO}_7(3), \text{PSp}_6(2), \text{PO}_8^+(2), G_2(3), G_2(4), F_4(2), \text{ and } {}^2E_6(2).$$

(See [MT11, Table 24.3].)

Consider the (only) exceptional cover $S = 2 \cdot F_4(2)$ of $X = F_4(2)$. When $p = 2$, as all the faithful irreducible characters of S have even degree [Atl], the problem is reduced to the simple group $F_4(2)$, which was already solved in Proposition 4.11. When $p = 3$, the Sylow normalizer $N_X(Q)$ of X is contained in $\text{PSL}_4(3).2_2$ and lifts to the direct product $C_2 \times N_X(Q)$ in S , as described in [M08, §4]. From the proof of Proposition 4.11, the maximal $3'$ -degree of the Sylow 3-normalizer of $\text{SL}_4(3)$ is at most $(3 - 1)^3/2 = 4$, implying $b(N_S(P)) = b(N_X(Q)) \leq 8$. When $p = 5$, $N_X(Q)$ is an extension of C_5^2 with the complex reflection group G_8 , which has the structure $C_4.S_4$. In this case, $N_S(P) = Z \times N_X(Q)$, and elementary character theory yields $b(N_S(P)) \leq 24$. Similarly, for $p = 7$, $N_X(Q)$ is an extension of C_7^2 with the complex reflection group G_5 of structure $C_6.A_4$. Here, it can be shown that $b(N_S(P)) \leq 48$. As $m(S) = 52$ [Atl], Hypothesis 4.3 holds for all the relevant p .

Consider $X = {}^2E_6(2)$. This group has two exceptional covers $2 \cdot X$ and $6 \cdot X$ with cyclic center. When $p = 2$, as in the case $S = 2 \cdot F_4(2)$, the problem reduces to handling the covers X and $3 \cdot X$, which are actually non-exceptional covers of X , and we are done by Proposition 4.11. Consider $p = 3$. Then $N_X(Q)$ is contained in $\Omega_7(3)$ and is lifted to the direct product $C_2 \times N_X(Q)$ in $2 \cdot X$. The proof of Proposition 4.11 shows that the maximal $3'$ -degree of the Sylow 3-normalizer of $\Omega_7(3)$ is at most 4, and therefore $b(N_S(P)) = b(N_X(Q)) \leq 4$ for $S = 2 \cdot X$. For the

six-fold cover $S = 6 \cdot X$, as the $3'$ -degree irreducible characters of both S and $N_S(P)$ are trivial on ZS_3 , we still have $b(N_S(P)) = b(N_{S/ZS_3}(P/ZS_3)) \leq 4$. When $p = 5$, [M08, Table 2] shows that the largest $5'$ -degree of the Sylow normalizer of both $2 \cdot X$ and $6 \cdot X$ (and of X itself as well) is 48. For $p = 7$, as mentioned in [M08, §4], the Sylow 7-normalizer of X is contained in $F_4(2)$ and moreover, $F_4(2)$ is lifted to $2 \cdot F_4(2)$ in $2 \cdot X$ and to $3 \times 2 \cdot F_4(2)$ in $6 \cdot X$. It follows from the previous paragraph that $b(N_S(P)) \leq 48$, whether S is the double or 6-fold cover of X . On the other hand, the smallest nontrivial degree of $6 \cdot {}^2E_6(2)$ is 1938 (GAP). Hypothesis 4.3 again holds true in this case.

Finally, using GAP, we have checked that Hypothesis 4.3 also holds for all other exceptional covers of the simple groups of Lie type listed above, as well as for the 3-fold and 6-fold covers of A_6 and A_7 . The proof is complete. \square

A few remarks are in order. First, Propositions 4.11, 4.13, and 4.14 together imply that Conjecture 4.1 holds for every quasisimple group S of exceptional Lie type and for every prime p . Second, our arguments show that, in all cases considered, Hypothesis 4.3 holds for (S, p) , except possibly in the following situations: $S = PSU_n(q)$ with $n \geq 3$ odd, $(n, q + 1) = 1$, and q a power of p ; when S is sporadic and $|S|_p = p$; or when $(S, p) \in \{(Co_2, 5), (Co_3, 3), (Co_3, 5), (McL, 5), (3 \cdot McL, 5)\}$.

5. Odd-degree characters

In this section we confirm Conjectures 4.1, 3.3, and 3.1 for $p = 2$, thereby proving Theorem B.

Theorem 5.1. *Conjecture 4.1 is true for all quasisimple groups S and $p = 2$.*

Proof. By Propositions 4.14, 4.11, and 4.13, and noting that Sylow 2-subgroups of an alternating group or its double cover are self-normalizing [O76], we only need to consider classical groups over fields of odd characteristic. Furthermore, we may assume that S is a non-exceptional covering group of the simple group $X := S/ZS$. As before, we use G for the finite reductive group of simply connected type such that $G/ZG = X$, and so S is a certain quotient of G . By Proposition 4.4, the existence of a required bijection is guaranteed if we are able to show that

$$(5-1) \quad b(N_G(Q)) \leq m(G)$$

for $Q \in \text{Syl}_2(G)$.

Let $X = \text{PSL}_2(q)$ with $5 \leq q$ odd, and hence $G = \text{SL}_2(q)$. Assume first that $q \equiv \pm 3 \pmod{8}$. Then Q is the quaternion group of order 8 and $N_G(Q)$ is isomorphic to $\text{SL}_2(3)$ (see [IMN07, §15E]). We have $b(N_G(Q)) = 3$, while $m(G) = (q - 1)/2$ if $q \equiv 3 \pmod{8}$ and $m(G) = (q + 1)/2$ if $q \equiv -3 \pmod{8}$, and so (5-1) is satisfied. When $q \equiv \pm 1 \pmod{8}$, Q is self-normalizing and the inequality is trivial.

Let $X = \text{PSL}_n^\pm(q)$ with $n \geq 3$. Here, as usual, we use the superscript $+$ for linear groups, while $-$ for unitary groups. Then $G = \text{SL}_n^\pm(q)$. Let $\tilde{G} := \text{GL}_n^\pm(q)$, $R \in \text{Syl}_2(\tilde{G})$, and take $Q := R \cap G$. By [K05, Theorem 1], we have

$$N_{\tilde{G}}(Q) = N_{\tilde{G}}(R) = RC_{\tilde{G}}(R).$$

The structure of Sylow normalizers in \tilde{G} was determined by Carter and Fong in [CF64, Lemma 6 and Theorem 4], as follows

$$N_{\tilde{G}}(R) \cong R \times (C_{(q \mp 1)_{2^t}})^t,$$

where $(q \mp 1)_{2^t}$ is the odd part of $q \mp 1$ and t is the number of terms in the 2-adic expansion of n . It follows that

$$N_{\tilde{G}}(Q)/Q' = N_{\tilde{G}}(R)/Q' \cong (R/Q') \times (C_{(q \mp 1)_{2^t}})^t.$$

Let $\mathbf{b}(M)$ denote the largest degree of an irreducible character of a finite group M . We have

$$\mathbf{b}(N_{\tilde{G}}(Q)/Q') = \mathbf{b}(R/Q') \leq |R : Q| \leq q + 1.$$

Here, the inequality in the middle follows from [I06, Corollary 11.29] and the fact that Q/Q' is an abelian normal subgroup of R/Q' . The last inequality follows from $|R : Q| = |R : (R \cap G)| = |RG : G| \leq |\tilde{G} : G| \leq q + 1$. Since $N_G(Q)/Q'$ is a normal subgroup of $N_{\tilde{G}}(Q)/Q'$, we deduce that

$$\mathbf{b}(N_G(Q)) = \mathbf{b}(N_G(Q)/Q') \leq \mathbf{b}(N_{\tilde{G}}(Q)/Q') \leq q + 1.$$

The desired inequality $\mathbf{b}(N_G(Q)) \leq m(G)$ then follows immediately from the bound provided in Table 1.

Next, consider $X = \text{PSp}_{2n}(q)$ with $n \geq 2$ and q odd. Then $G = \text{Sp}_{2n}(q)$ with ZG being cyclic of order 2. The Sylow 2-subgroup Q is self-normalizing in G when $q \equiv \pm 1 \pmod{8}$; otherwise, $|N_G(Q)/Q| = 3^t$ where t is the number of terms in the 2-adic expansion of n (see [CF64, Theorem 4]). In the former case, Hypothesis 4.3 is trivial. In the latter, we have

$$\mathbf{b}(N_G(Q)) \leq |N_G(Q)/Q| = 3^t < (q^n - 1)/2 = d(G) \leq m(G),$$

and we are done again.

Lastly, consider $X = \Omega_{2n+1}(q)$ with $n \geq 3$ or $X = P\Omega_{2n}^\pm(q)$ with $n \geq 4$ (q again is odd). Here $G = \text{Spin}_{2n+1}(q)$ or $\text{Spin}_{2n}^\pm(q)$, respectively. According to [CF64, Theorem 5], Sylow 2-subgroups of G are self-normalizing, and we conclude as before. □

Theorem 5.2. *Conjecture 3.1 is true when $p = 2$.*

Proof. This follows from Theorems 5.1 and 3.5. □

As said in the introduction, Theorem B immediately follows from Theorem 5.2.

We conclude with some remarks. To complete the proof of Conjecture A and Conjecture 3.1 for all primes p , it remains to verify Conjecture 3.3 for quasisimple classical groups S in characteristic different from p , as well as for covers of alternating groups. This appears to be a nontrivial problem, since for these groups the normalizer $N_S(P)$ typically has many irreducible p' -characters whose degrees exceed the minimal p' -degree of S .

Conjecture 3.1 has now been established for symmetric groups by Giannelli [G25]. It may be possible to adapt the methods of [G25] to prove Conjecture 3.3 for alternating groups.

References

- [A94] J. B. An, “Alperin–McKay conjecture for the Chevalley groups $G_2(q)$ ”, *J. Algebra* **165**:1 (1994), 184–193. MR
- [A98] J. An, “The Alperin and Dade conjectures for Ree groups ${}^2F_4(q^2)$ in non-defining characteristics”, *J. Algebra* **203**:1 (1998), 30–49. MR
- [Atl] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups: maximal subgroups and ordinary characters for simple groups*, Oxford University Press, 1985. MR
- [B09] O. Brunat, “On the inductive McKay condition in the defining characteristic”, *Math. Z.* **263**:2 (2009), 411–424. MR
- [B63] R. Brauer, “Representations of finite groups”, pp. 133–175 in *Lectures on modern mathematics*, vol. 1, Wiley, New York, 1963. MR
- [BM92] M. Broué and G. Malle, “Théorèmes de Sylow génériques pour les groupes réductifs sur les corps finis”, *Math. Ann.* **292**:2 (1992), 241–262. MR
- [BM93] M. Broué and G. Malle, “Zyklotomische Heckealgebren”, pp. 119–189 in *Représentations unipotentes génériques et blocs des groupes réductifs finis*, Astérisque **212**, Soc. math. de France, 1993. MR
- [C81] R. W. Carter, “Centralizers of semisimple elements in the finite classical groups”, *Proc. London Math. Soc.* (3) **42**:1 (1981), 1–41. MR
- [C85] R. W. Carter, *Finite groups of Lie type: conjugacy classes and complex characters*, Wiley, New York, 1985. MR
- [CF64] R. Carter and P. Fong, “The Sylow 2-subgroups of the finite classical groups”, *J. Algebra* **1** (1964), 139–151. MR
- [CS24] M. Cabanes and B. Späth, “The McKay Conjecture on character degrees”, preprint, 2024. To appear in *Ann. of Math.* arXiv 2410.20392
- [D18] G. D. Franceschi, *Centralizers and conjugacy classes in finite classical groups*, D.Phil. thesis, The University of Auckland, 2018.
- [D66] E. C. Dade, “Blocks with cyclic defect groups”, *Ann. of Math.* (2) **84** (1966), 20–48. MR
- [D96] E. C. Dade, “Counting characters in blocks with cyclic defect groups, I”, *J. Algebra* **186**:3 (1996), 934–969. MR
- [F80] W. Feit, “Some consequences of the classification of finite simple groups”, pp. 175–181 in *The Santa Cruz Conference on Finite Groups* (Santa Cruz, CA, 1979), Proc. Sympos. Pure Math. **37**, Amer. Math. Soc., 1980. MR

- [F82] W. Feit, *The representation theory of finite groups*, North-Holland Mathematical Library **25**, North-Holland, Amsterdam, 1982. MR
- [G25] E. Giannelli, “McKay bijections and character degrees”, preprint, 2025. arXiv 2507.01730
- [GM20] M. Geck and G. Malle, *The character theory of finite groups of Lie type; a guided tour*, Cambridge Studies in Advanced Mathematics **187**, Cambridge University Press, 2020. MR
- [I06] I. M. Isaacs, *Character theory of finite groups*, Academic Press, New York, 1976. Reprinted AMS Chelsea, Providence, RI, 2006.
- [I08] I. M. Isaacs, *Finite group theory*, Graduate Studies in Mathematics **92**, American Mathematical Society, 2008. MR
- [I86] I. M. Isaacs, “Subgroups with the character restriction property”, *J. Algebra* **100**:2 (1986), 403–420. MR
- [IMN07] I. M. Isaacs, G. Malle, and G. Navarro, “A reduction theorem for the McKay conjecture”, *Invent. Math.* **170**:1 (2007), 33–101. MR
- [K05] A. S. Kondrat’ev, “Normalizers of Sylow 2-subgroups in finite simple groups”, *Mat. Zametki* **78**:3 (2005), 368–376. In Russian; translated in *Math. Notes* **78**:3–4 (2005), 338–346. MR
- [KS16] S. Koshitani and B. Späth, “The inductive Alperin–McKay and blockwise Alperin weight conditions for blocks with cyclic defect groups and odd primes”, *J. Group Theory* **19**:5 (2016), 777–813. MR
- [Li25] M. Linckelmann, “On character degrees of cyclic and Klein four defect blocks”, preprint, 2025. arXiv 2509.01394
- [LOST10] M. W. Liebeck, E. A. O’Brien, A. Shalev, and P. H. Tiep, “The Ore conjecture”, *J. Eur. Math. Soc.* **12**:4 (2010), 939–1008. MR
- [Lü01] F. Lübeck, “Smallest degrees of representations of exceptional groups of Lie type”, *Comm. Algebra* **29**:5 (2001), 2147–2169. MR
- [M07] G. Malle, “Height 0 characters of finite groups of Lie type”, *Represent. Theory* **11** (2007), 192–220. MR
- [M08] G. Malle, “The inductive McKay condition for simple groups not of Lie type”, *Comm. Algebra* **36**:2 (2008), 455–463. MR
- [MT11] G. Malle and D. Testerman, *Linear algebraic groups and finite groups of Lie type*, Cambridge Studies in Advanced Mathematics **133**, Cambridge University Press, 2011. MR
- [N04] G. Navarro, “Problems on characters and Sylow subgroups”, pp. 275–281 in *Finite groups 2003*, de Gruyter, Berlin, 2004. MR
- [N10] H. N. Nguyen, “Low-dimensional complex characters of the symplectic and orthogonal groups”, *Comm. Algebra* **38**:3 (2010), 1157–1197. MR
- [N18] G. Navarro, *Character theory and the McKay conjecture*, Cambridge Studies in Advanced Mathematics **175**, Cambridge University Press, 2018. MR
- [O76] J. r. B. Olsson, “McKay numbers and heights of characters”, *Math. Scand.* **38**:1 (1976), 25–42. MR
- [S12] B. Späth, “Inductive McKay condition in defining characteristic”, *Bull. Lond. Math. Soc.* **44**:3 (2012), 426–438. MR
- [T08] A. Turull, “Above the Glauberman correspondence”, *Adv. Math.* **217**:5 (2008), 2170–2205. MR
- [T09] A. Turull, “The Brauer–Clifford group”, *J. Algebra* **321**:12 (2009), 3620–3642. MR

- [T18] J. Taylor, “Action of automorphisms on irreducible characters of symplectic groups”, *J. Algebra* **505** (2018), 211–246. MR
- [TZ96] P. H. Tiep and A. E. Zalesskii, “Minimal characters of the finite classical groups”, *Comm. Algebra* **24**:6 (1996), 2093–2167. MR
- [W98] R. A. Wilson, “The McKay conjecture is true for the sporadic simple groups”, *J. Algebra* **207**:1 (1998), 294–305. MR

Received June 3, 2025. Revised March 5, 2026.

NGUYEN N. HUNG
DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF AKRON
AKRON, OH
UNITED STATES
hungnguyen@uakron.edu

J. MIQUEL MARTÍNEZ
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT DE VALÈNCIA
BURJASSOT, VALENCIA
SPAIN
josep.m.martinez@uv.es

GABRIEL NAVARRO
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT DE VALÈNCIA
BURJASSOT, VALENCIA
SPAIN
gabriel@uv.es

ON THE FIRST EIGENVALUE OF THE HODGE LAPLACIAN OF SUBMANIFOLDS

CHRISTOS-RAENT ONTI

We prove that equality in a sharp lower bound for the first p -eigenvalue of the Hodge Laplacian on closed submanifolds in space forms can occur only on topological spheres, assuming positivity.

1. Introduction

Let M^n be a closed, connected and oriented Riemannian manifold of dimension n . For each integer $1 \leq p \leq n-1$, the Hodge–Laplace operator (or the Hodge Laplacian) acting on p -forms is defined by

$$\Delta = d\delta + \delta d : \Omega^p(M^n) \rightarrow \Omega^p(M^n),$$

where d and δ are the differential and the co-differential operators, respectively. It is well known that the spectrum of the Hodge–Laplace operator is discrete and nonnegative, and that its kernel is isomorphic to the p -th de Rham cohomology group $H^p(M^n; \mathbb{R})$. If $\lambda_{1,p}(M^n)$ denotes its lowest eigenvalue, then

$$\lambda_{1,p}(M^n) = \inf_{\omega \in \Omega^p(M^n) \setminus \{0\}} \frac{\int_M (\|d\omega\|^2 + \|\delta\omega\|^2) dM}{\int_M \|\omega\|^2 dM}.$$

Since the above is invariant by the Poincaré duality induced by the Hodge $*$ -operator, we have $\lambda_{1,p}(M^n) = \lambda_{1,n-p}(M^n)$ and thus we may assume that $p \leq n/2$. Clearly, if $\lambda_{1,p}(M^n) > 0$, then $H^p(M^n; \mathbb{R}) = H^{n-p}(M^n; \mathbb{R}) = 0$.

The Hodge Laplacian satisfies for every p -form $\omega \in \Omega^p(M^n)$ the Bochner–Weitzenböck formula

$$(1) \quad \Delta\omega = \nabla^* \nabla \omega + \mathcal{B}^{[p]}\omega,$$

where $\nabla^* \nabla$ is the connection Laplacian and $\mathcal{B}^{[p]} : \Omega^p(M^n) \rightarrow \Omega^p(M^n)$ is a certain symmetric endomorphism on the bundle of p -forms, called the *Bochner–Weitzenböck operator*. Therefore, (1) implies that lower bounds on the Bochner–Weitzenböck operator lead naturally to lower bounds on the Hodge–Laplace operator.

MSC2020: 53C40, 53C42.

Keywords: Hodge Laplacian, first eigenvalue, isometric immersion.

In particular, from [6, Proposition 3] we get that

$$(2) \quad \text{if } \mathcal{B}^{[p]} \geq p(n-p)\Lambda \text{ for some } \Lambda > 0, \text{ then } \lambda_{1,p}(M^n) \geq p(n-p+1)\Lambda.$$

Let $f : M^n \rightarrow \tilde{M}^{n+m}$, $n \geq 3$, be an isometric immersion into a Riemannian manifold \tilde{M}^{n+m} of dimension $n+m$. The second fundamental form α_f is viewed as a section of the vector bundle $\text{Hom}(TM \times TM, N_fM)$, where N_fM is the normal bundle. For each unit normal vector field $\xi \in \Gamma(N_fM)$, the associated shape operator A_ξ is given by

$$\langle A_\xi X, Y \rangle = \langle \alpha_f(X, Y), \xi \rangle, \quad X, Y \in TM.$$

Recall that the traceless part of the second fundamental form is $\Phi = \alpha_f - \langle \cdot, \cdot \rangle \mathcal{H}$, where \mathcal{H} denotes the *mean curvature vector field* given by $\mathcal{H} = (\text{tr } \alpha_f)/n$, where tr means taking the trace. Finally, by H we denote the length of the *mean curvature*, that is, $H = \|\mathcal{H}\|$.

Proposition 1 (Onti and Vlachos [5, Proposition 16]). *If the curvature operator of \tilde{M}^{n+m} is bounded from below by a constant c , then the Bochner–Weitzenböck operator of M^n , for any $1 \leq p \leq \lfloor n/2 \rfloor$, satisfies pointwise the inequality*

$$(3) \quad \min_{\substack{\omega \in \Omega^p(M^n) \\ \|\omega\|=1}} \langle \mathcal{B}^{[p]} \omega, \omega \rangle \geq \frac{p(n-p)}{n} \left(n(H^2 + c) - \frac{n(n-2p)}{\sqrt{np(n-p)}} H \|\Phi\| - \|\Phi\|^2 \right).$$

If equality holds in (3) at a point $x \in M^n$, then:

- (i) *The shape operator $A_\xi(x)$ has at most two distinct eigenvalues with multiplicities p and $n-p$ for every unit vector $\xi \in N_fM(x)$. If in addition $p < n/2$ and the eigenvalue of multiplicity $n-p$ vanishes, then $A_\xi(x) = 0$.*
- (ii) *If $H(x) \neq 0$ and $p < n/2$, then $\text{Im } \alpha_f(x) = \text{span} \{ \mathcal{H}(x) \}$.*

Therefore, if

$$\kappa_p := \min_{x \in M^n} \left\{ (H^2 + c) - \frac{n-2p}{\sqrt{np(n-p)}} H \|\Phi\| - \frac{1}{n} \|\Phi\|^2 \right\}$$

for some $1 \leq p \leq \lfloor n/2 \rfloor$, then it follows from (2) and (3) that

$$(4) \quad \lambda_{1,p}(M^n) \geq p(n-p+1)\kappa_p.$$

Inequality (4) was first proved by Savo for hypersurfaces [6, Theorem 7], and subsequently extended by Cui and Sun to submanifolds of arbitrary codimension [3, Theorem 1.1]. Those authors also showed that the inequality is sharp by providing trivial examples attaining equality. However, no characterization was given of the submanifolds for which equality holds. The aim of this note is to shed light on the case of equality in (4) assuming $\lambda_{1,p}(M^n) > 0$, when $\tilde{M}^{n+m} = \mathbb{Q}_c^{n+m}$, where \mathbb{Q}_c^{n+m} denotes the complete simply connected space form of constant sectional curvature c .

In fact, we prove that in this case equality occurs only on topological spheres. For simplicity we assume that $c \in \{0, \pm 1\}$. Thus \mathbb{Q}_c^{n+m} is the Euclidean space \mathbb{R}^{n+m} ($c = 0$), the unit sphere \mathbb{S}^{n+m} ($c = 1$), or the hyperbolic space \mathbb{H}^{n+m} ($c = -1$).

Theorem. *Let $f : M^n \rightarrow \mathbb{Q}_c^{n+m}$, $n \geq 4$, be an isometric immersion of a closed, connected and oriented Riemannian manifold. If for some $1 \leq p \leq \lfloor n/2 \rfloor$ equality holds in (4) with $\lambda_{1,p}(M^n) > 0$, then M^n is homeomorphic to the sphere \mathbb{S}^n .*

2. Proof of the theorem

The idea of the proof is to show that M^n is a simply connected homology sphere over the integers and the proof will follow by the generalized Poincaré conjecture (Smale $n \geq 5$, Freedman $n = 4$).

Assume that for some $1 \leq p \leq \lfloor n/2 \rfloor$ equality holds in (4) with $\lambda_{1,p}(M^n) > 0$. Then Proposition 1 implies that the shape operator $A_\xi(x)$ at each point x will have at most two distinct eigenvalues of multiplicities p and $n-p$ for every unit vector $\xi \in N_f M(x)$. We claim that there exists a Morse function on M^n such that the index at each critical point is 0, p , $n-p$ or n . To this end, we distinguish two cases:

CASE $c \in \{0, 1\}$: Let $u \in \mathbb{R}^{n+m+c}$ be a vector such that the height function

$$\varphi : M^n \rightarrow \mathbb{R}, \quad \varphi(x) = \langle f_c(x), u \rangle$$

is a Morse function, where $f_c = f$ if $c = 0$, and $f_c = j \circ f$, where $j : \mathbb{S}^{n+m} \rightarrow \mathbb{R}^{n+m+1}$ denotes the standard inclusion, if $c = 1$. A direct computation gives that at a critical point x_0 of φ we have

$$u \in N_{f_c} M(x_0) \quad \text{and} \quad \text{Hess } \varphi(X, Y) = \langle \alpha_{f_c}(X, Y), u \rangle \quad \text{for all } X, Y \in T_{x_0} M.$$

Obviously, the second fundamental form of f_c has at most two distinct principal curvatures of multiplicities p and $n-p$ in every normal direction and the claim follows in this case.

CASE $c = -1$: We consider the function

$$\varphi : \mathbb{H}^{n+m} \rightarrow \mathbb{R}, \quad \varphi(x) = \frac{1}{2}r^2(x),$$

where $r(x)$ denotes the distance function issuing from some suitable choice of point $o \in \mathbb{H}^{n+m}$ to $x \in \mathbb{H}^{n+m}$. It is a standard fact that φ is smooth everywhere in \mathbb{H}^{n+m} . Let $\gamma(t)$ be a unit speed geodesic with $\gamma(0) = o$. Then, we have $\gamma'(t) = \text{grad } r(\gamma(t))$. For $X, Y \in \Gamma(T\mathbb{H})$ a direct computation gives

$$\langle X, \text{grad } \varphi \rangle = r \langle X, \text{grad } r \rangle$$

and

$$\text{Hess } \varphi(X, Y) = \langle X, \text{grad } r \rangle \langle Y, \text{grad } r \rangle + r \text{Hess } r(X, Y).$$

Consider $\tilde{\varphi} = \varphi \circ f : M^n \rightarrow (0, +\infty)$ and choose $o \in \mathbb{H}^{n+m}$ such that $\tilde{\varphi}$ is a Morse function on M^n (this is always possible). At a critical point $x_0 \in M^n$, we have $\text{grad } r(f(x_0)) \perp f_*(T_{x_0}M^n)$, that is the (unique unit speed) geodesic $\gamma(t)$ will hit $f(M^n)$ orthogonally, and

$$(5) \quad \text{Hess } \tilde{\varphi}(\tilde{X}, \tilde{Y}) = r(\text{Hess } r(f_*\tilde{X}, f_*\tilde{Y}) + \langle A_{\text{grad } r} \tilde{X}, \tilde{Y} \rangle), \quad \tilde{X}, \tilde{Y} \in T_{x_0}M^n.$$

Let $\gamma(\ell) = f(x_0)$ and consider a Jacobi field $J(t)$ along $\gamma(t)$ with $J(0) = 0$. It follows that

$$(6) \quad \text{Hess } r(J(\ell), J(\ell)) = \langle J'(\ell), J(\ell) \rangle = \frac{1}{2} \frac{d}{dt} \|J(t)\|^2|_{t=\ell}.$$

Recall that the Jacobi fields $J(t)$ in \mathbb{H}^{n+m} with $J'(0) \perp \gamma'(0)$ are given by

$$J(t) = \sinh t \cdot w(t),$$

where $w(t)$ is a parallel vector field along $\gamma(t)$ with $J'(0) = w$ and $\|w\| = 1$. Observe that $A_{\text{grad } r}(x_0)$ has at most two distinct eigenvalues, say λ and μ with multiplicities p and $n-p$, respectively. Consider an orthonormal basis $\{e_1, \dots, e_n\}$ of $T_{x_0}M^n$ such that

$$A_{\text{grad } r}(e_i) = \begin{cases} \lambda e_i & \text{for } 1 \leq i \leq p, \\ \mu e_i & \text{for } p+1 \leq i \leq n. \end{cases}$$

Let $w_i(t)$ be such that

$$w_i(\ell) = f_*(e_i), \quad 1 \leq i \leq n,$$

with corresponding Jacobi fields

$$J_i(t) = \sinh t \cdot w_i(t), \quad 1 \leq i \leq n.$$

Therefore, from (6) we obtain

$$\text{Hess } r(J_i(\ell), J_i(\ell)) = \frac{1}{2} \frac{d}{dt} \|J_i(t)\|^2|_{t=\ell} = \frac{1}{2} \sinh(2\ell), \quad \text{for all } 1 \leq i \leq n.$$

Hence (5) gives

$$\text{Hess } \tilde{\varphi}(e_i, e_i) = \begin{cases} \ell(\coth \ell + \lambda) & \text{for } 1 \leq i \leq p, \\ \ell(\coth \ell + \mu) & \text{for } p+1 \leq i \leq n, \end{cases}$$

and therefore it is now clear that $\text{Index } \tilde{\varphi}(x_0) \in \{0, p, n-p, n\}$, proving the claim.

Therefore, it follows from standard Morse theory (see [4, Theorem 3.5] or [2, Theorem 4.10]) that M^n has the homotopy type of a CW-complex with cells only in dimensions $0, p, n-p$ or n . Therefore

$$(7) \quad H_i(M^n; \mathbb{Z}) = 0 \quad \text{for all } i \neq 0, p, n-p, n.$$

Next, we claim that also

$$(8) \quad H_p(M^n; \mathbb{Z}) = H_{n-p}(M^n; \mathbb{Z}) = 0.$$

Indeed, our hypothesis implies $H^p(M^n; \mathbb{R}) = H^{n-p}(M^n; \mathbb{R}) = 0$. Hence

$$(9) \quad H_p(M^n; \mathbb{Z}) = \text{Tor}(H_p(M^n; \mathbb{Z})) \text{ and } H_{n-p}(M^n; \mathbb{Z}) = \text{Tor}(H_{n-p}(M^n; \mathbb{Z})).$$

By Poincaré duality, the universal coefficient theorem and (7), we have

$$\text{Tor}(H_p(M^n; \mathbb{Z})) \cong \text{Tor}(H_{n-p-1}(M^n; \mathbb{Z})) = 0$$

and

$$\text{Tor}(H_{n-p}(M^n; \mathbb{Z})) \cong \text{Tor}(H_{p-1}(M^n; \mathbb{Z})) = 0,$$

where \cong denotes the isomorphism of the corresponding groups. This, in combination with (9), proves (8). Hence M^n is a homology sphere over the integers.

Finally, we show that M^n is simply connected. If $p \neq 1$ this follows directly from [1, Proposition 4.5.7, p. 90], since in this case, φ has no critical points of index one. If $p = 1$, then since there are no 2-cells, it follows by the cellular approximation theorem that the inclusion of the 1-skeleton $X^{(1)} \hookrightarrow M^n$ induces isomorphism between the fundamental groups. Therefore, $\pi_1(M^n)$ is a free group on $\beta_1(M^n; \mathbb{Z}) = 0$ elements, and thus M^n is simply connected.

Therefore, M^n is a homotopy sphere and by the generalized Poincaré conjecture (Smale $n \geq 5$, Freedman $n = 4$), M^n is homeomorphic to \mathbb{S}^n , which concludes the proof of the theorem.

References

- [1] M. Audin and M. Damian, *Morse theory and Floer homology*, Springer, 2014.
- [2] J. Cheeger and D. G. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland Mathematical Library **9**, North-Holland, Amsterdam, 1975.
- [3] Q. Cui and L. Sun, “Optimal lower eigenvalue estimates for Hodge-Laplacian and applications”, *J. Differential Equations* **266**:12 (2019), 8320–8343.
- [4] J. Milnor, *Morse theory*, Annals of Mathematics Studies **51**, Princeton University Press, 1963.
- [5] C.-R. Onti and T. Vlachos, “Homology vanishing theorems for pinched submanifolds”, *J. Geom. Anal.* **32**:12 (2022), Paper No. 294, 33.
- [6] A. Savo, “The Bochner formula for isometric immersions”, *Pacific J. Math.* **272**:2 (2014), 395–422.

Received September 16, 2025. Revised November 21, 2025.

CHRISTOS-RAENT ONTI
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF CYPRUS
NICOSIA
CYPRUS
onti.christos-raent@ucy.ac.cy

HOWE DUALITY FOR THE DUAL PAIR $\mathrm{SL}_2(\mathbb{R}) \times F_{4,1}$: A PING PONG OF K -TYPES

GORDAN SAVIN

We prove Howe duality for an exceptional theta correspondence. To that end, we relate the K -types of corresponding representations by exploiting a pair of see-saw identities.

1. Introduction

Let \mathbb{O} be the algebra of Cayley octonions over the field of real numbers \mathbb{R} . Let J be the 27-dimensional space of 3×3 hermitian symmetric matrices with coefficients in \mathbb{O} . Let $N_J : J \rightarrow \mathbb{R}$ be the cubic form (the norm of J), essentially the determinant of 3×3 matrices. For every $e \in J$ such that $N_J(e) \neq 0$ there is a structure of exceptional Jordan algebra on J such that e is the identity of J . Let $G = \mathrm{Aut}(J, e)$ be the group of automorphisms of the resulting Jordan algebra, which is the same as the group of linear transformations of J preserving N_J and the point e . It is a simple Lie group of absolute type F_4 . See [6] for all of this. If we pick e to be

$$\begin{pmatrix} +1 & & \\ & +1 & \\ & & +1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -1 & & \\ & -1 & \\ & & +1 \end{pmatrix},$$

then G is compact for the first choice of e and of split rank one for the second [11]. The Jordan algebra, by way of the Koecher–Tits construction [8], gives rise to a simply connected group $G(J)$, of the exceptional type E_7 and split rank 3 over \mathbb{R} (the same group for both choices of e). The group $G(J)$ comes along with the dual pair (see [7])

$$\mathrm{SL}_2(\mathbb{R}) \times G \subset G(J).$$

These dual pairs are completely analogous to $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{O}(p, q)$ in $\mathrm{Sp}_{2n}(\mathbb{R})$ where $n = p + q$. Indeed, if we take J to be the space of $n \times n$ symmetric matrices with coefficients in \mathbb{R} , then orthogonal groups are stabilizers of generic points in J , and $G(J)$ is $\mathrm{Sp}_{2n}(\mathbb{R})$.

This work is supported by the Croatian Science Foundation under the project IP-2022-10-4615 and by a gift (no. 946504) from the Simons Foundation.

MSC2020: 22E46, 22E47.

Keywords: minimal representation, theta correspondence.

The group $G(J)$ has a minimal (holomorphic) representation Π that appears as a local component of a global representation [5]. In [3], Π was restricted to the dual pair $SL_2(\mathbb{R}) \times G$, with G compact, and the following decomposition was obtained:

$$\Pi = \bigoplus_{n \geq 0} \delta(2n + 12) \otimes E_n.$$

Here $\delta(2n + 12)$ is the holomorphic representation of the lowest weight $2n + 12$ and E_n is the irreducible representation of G of the highest weight $n\varpi_4$ where ϖ_4 is the fourth fundamental weight for F_4 . It is the highest weight of the 26-dimensional irreducible representation of G (the complement of the line through e in J).

Here we study the restriction of Π to the dual pair with G noncompact. Let K be the maximal compact subgroup of G . We emphasize that we do not work with continuous representations of noncompact groups, but with the corresponding (\mathfrak{g}, K) -modules, where \mathfrak{g} is the complex Lie algebra of G . Thus, if π is a (\mathfrak{g}, K) -module of finite length, we define

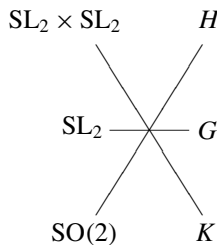
$$\Theta(\pi) = (\Pi \otimes \pi^\vee)_{\mathfrak{g}}.$$

Here π^\vee is the contragredient of π , and the subscript \mathfrak{g} is saying that we are taking co-invariants with respect to the action of \mathfrak{g} on the tensor product. If π is irreducible then $\Theta(\pi) \otimes \pi$ is the maximal π -isotypic quotient of Π (see [1]). In other words, the above definition generalizes the usual definition of the theta lift.

We can analogously define $\Theta(\sigma)$ for an $(\mathfrak{sl}_2, SO(2))$ -module σ of finite length. Observe that $\Theta(\pi)$ and $\Theta(\sigma)$ are naturally $(\mathfrak{sl}_2, SO(2))$ and (\mathfrak{g}, K) -modules, respectively. We shall show that $\Theta(\pi)$ and $\Theta(\sigma)$ are finite length modules, and that they have unique irreducible quotients, if π and σ are irreducible. The main input is the structure of lifts of types. More precisely, if τ is a K -type, then

$$\Theta(\tau) = (\Pi \otimes \tau^\vee)_K$$

is also an $(\mathfrak{sl}_2, SO(2))$ -module that we determine explicitly. Similarly, we have a description of the lift for $SO(2)$ -types. This is done in the last section using a strategy of Howe [4], involving the following see-saw diagram of dual pairs in $G(J)$:



Here H is a simply connected, hermitian symmetric group of absolute type E_6 . The centralizer of G sits diagonally in $SL_2 \times SL_2$. Thus $\Theta(\tau)$ is naturally an $(\mathfrak{sl}_2 + \mathfrak{sl}_2, SO(2) \times SO(2))$ -module. We compute it, and then restrict it to the diagonal \mathfrak{sl}_2 . A similar strategy is used for lifts of $SO(2)$ -types.

With the lift of types computed, we can play a game of ping pong with types: if $\sigma \otimes \pi$ is a quotient of Π and τ is a type of π then, by a see-saw identity, σ must have an $SO(2)$ -type determined by τ and vice versa. More details in the next section where main results are obtained. A similar strategy (and the name ping-pong) was used in [2] to establish Howe duality for exceptional p -adic dual pairs.

2. Main results

The correspondence with compact G establishes a correspondence of infinitesimal characters in the noncompact case. The reader can consult [10] for more details on this. Let us write down the correspondence. Using the standard realization of the F_4 root system, the infinitesimal character of E_n (the representation with the highest weight $n\varpi_4$) is

$$\frac{1}{2}(2n+11, 5, 3, 1).$$

On the other hand, the infinitesimal character of $\delta(2n + 12)$ is $2n + 11$, which we recognize as the first entry above. This means that if σ has infinitesimal character x , then $\Theta(\sigma)$ has infinitesimal character $\frac{1}{2}(x, 5, 3, 1)$. More generally, if σ is annihilated by an ideal in the center of $U(\mathfrak{sl}_2)$ of finite codimension, then $\Theta(\pi)$ is also annihilated by an ideal in the center of $U(\mathfrak{g})$ of finite codimension. Hence, for σ of finite length, in order to prove that $\Theta(\sigma)$ has finite length, it suffices to prove that it is admissible. The same goes for $\Theta(\pi)$.

The maximal compact subgroup of $SL_2(\mathbb{R})$ is $SO(2)$. Its irreducible representations are one-dimensional characters parameterized with integers n . Let (n) denote the corresponding one-dimensional representation. Since the center of $SL_2(\mathbb{R})$ is also the center of the simply connected $G(J)$, only even $n = 2m$ characters appear in Π .

The maximal compact subgroup of G is denoted by K . It is a simple group of type B_4 . The group K can be picked to be the intersection of G with the compact form of G , where the two groups are the stabilizers of the two choices for e , as in the introduction. Let \mathfrak{g} be the complex simple Lie algebra of G , and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ the corresponding Cartan decomposition. Here \mathfrak{p} is the 16-dimensional spin representation of K . Let μ be its highest weight, and let λ be the highest weight of the standard 9-dimensional irreducible representation of K . Let $\tau(m, n)$ be the irreducible representation of K of the highest weight

$$m\lambda + n\mu.$$

Applying the branching rule [9] to the representations E_n , we see that only these representations of K lie in Π . Let

$$\Theta(\tau(m, n)) = (\Pi \otimes \tau(m, n)^\vee)_K$$

be the lift of $\tau(m, n)$. It is naturally an \mathfrak{sl}_2 -module. We have, by Proposition 3.1,

$$\Theta(\tau(m, n)) \cong U(\mathfrak{sl}_2) \otimes_{U(\mathfrak{so}(2))} \otimes(2m+4).$$

A power of this identity is demonstrated by the following lemma:

Lemma 2.1. *Let σ be a finite length $(\mathfrak{sl}_2, \text{SO}(2))$ -module. Then*

$$\text{Hom}_K(\Theta(\sigma), \tau(m, n)) \cong \text{Hom}_{\mathfrak{sl}_2}(\Theta(\tau(m, n)), \sigma) \cong \text{Hom}_{\text{SO}(2)}(\otimes(2m+4), \sigma).$$

Proof. The first isomorphism is a see-saw identity, obtained by switching the order of taking \mathfrak{sl}_2 and K co-invariants. The second isomorphism follows from the Frobenius reciprocity, since $\Theta(\tau(m, n))$ is isomorphic to $U(\mathfrak{sl}_2) \otimes_{U(\mathfrak{so}(2))} \otimes(2m+4)$. \square

Now we have the following consequence, Santa Claus is coming to town:

Proposition 2.2. *Let σ be a finite length $(\mathfrak{sl}_2, \text{SO}(2))$ -module. Then $\Theta(\sigma) \neq 0$ if and only if σ has a type $2m+4$ for some $m \geq 0$. $\Theta(\sigma)$ has finite length. If σ is irreducible, $\Theta(\sigma)$ has multiplicity free K -types, consisting of all $\tau(m, n)$ such that $2m+4$ is a type of σ .*

Proof. This is all trivial from the lemma; only the finite length of $\Theta(\sigma)$ perhaps merits some explanation. It is a combination of admissibility (from the lemma) and the fact that $\Theta(\sigma)$ is annihilated by an ideal in $Z(\mathfrak{g})$ of finite codimension. \square

Now we go in the opposite direction. For a character $2m+4$ of $\text{SO}(2)$ consider $\Theta(2m+4)$. By Proposition 3.2, it is a quotient of

$$U(\mathfrak{g}) \otimes_{U(\mathfrak{k})} F_m$$

where $F_m = \mathbb{C}$ if $m \leq 0$, otherwise

$$F_m = \tau(0, 0) \oplus \tau(1, 0) \oplus \dots \oplus \tau(m, 0).$$

Lemma 2.3. *Let π be a finite length (\mathfrak{g}, K) -module. Then*

$$\text{Hom}_{\text{SO}(2)}(\Theta(\pi), \otimes(2m+4)) \cong \text{Hom}_{\mathfrak{g}}(\Theta(2m+4), \pi) \subseteq \text{Hom}_K(F_m, \pi).$$

Proof. The isomorphism is again a see-saw identity. The inclusion follows from the Frobenius reciprocity, since $\Theta(2m+4)$ is a quotient of $U(\mathfrak{g}) \otimes_{U(\mathfrak{k})} F_m$. \square

We now record an easy consequence.

Proposition 2.4. *Let π be a finite length (\mathfrak{g}, K) -module. Then $\Theta(\pi) \neq 0$ only if π contains a type $\tau(m, 0)$ for some m . $\Theta(\pi)$ is of finite length.*

We are now ready to state and prove the main result of this paper.

Theorem 2.5. *Let σ be an irreducible $(\mathfrak{sl}_2, SO(2))$ -module. Assume that σ contains the type $(2m+4)$, for $m \geq 0$, and no smaller types $2n+4$, with $n \geq 0$. Then $\Theta(\sigma)$ has a unique irreducible quotient. It contains the type $\tau(m, 0)$ with multiplicity one, and no types $\tau(n, 0)$ with $n < m$. Conversely, let π be an irreducible (\mathfrak{g}, K) -module containing the type $\tau(m, 0)$, and no smaller such types. If $\Theta(\pi)$ is nonzero, then $\Theta(\pi)$ has a unique irreducible quotient. It contains the type $(2m+4)$, and no smaller types $2n+4$, with $n \geq 0$.*

Proof. Assume π is a quotient of $\Theta(\sigma)$. We do not assume that π is irreducible. By Lemma 2.1, we have the sequence of inclusions

$$\begin{aligned} \text{Hom}_K(\pi, \tau(m, 0)) &\subseteq \text{Hom}_K(\Theta(\sigma), \tau(m, 0)) \cong \text{Hom}_{\mathfrak{sl}_2}(\Theta(\tau(m, 0)), \sigma) \\ &\cong \text{Hom}_{SO(2)}((2m+4), \sigma). \end{aligned}$$

We can run this sequence with any $2n+4$ in place of $2m+4$. If $n < m$, by the assumption, the last space is trivial, hence $\tau(n, 0)$ is not a type of π . We shall use this in a moment. Since π is a quotient of $\Theta(\sigma)$ and σ is irreducible, $\pi \otimes \sigma$ is a quotient of Π . But this implies that σ is a quotient of $\Theta(\pi)$, and by Lemma 2.3 we have a second sequence of inclusions (note that we are starting with the space of the same dimension as as the space we ended with in the first sequence):

$$\begin{aligned} \text{Hom}_{SO(2)}(\sigma, (2m+4)) &\subseteq \text{Hom}_{SO(2)}(\Theta(\pi), (2m+4)) \cong \text{Hom}_{\mathfrak{g}}(\Theta(2m+4), \pi) \\ &\subseteq \text{Hom}_K(F_m, \pi). \end{aligned}$$

Since π has no type $\tau(n, 0)$ with $n < m$, we ended with $\text{Hom}_K(\tau(m, 0), \pi)$, which has the same dimension as $\text{Hom}_K(\pi, \tau(m, 0))$, the starting space in the first sequence of inclusions. Thus all inclusions in the two sequences are isomorphisms, and all spaces are one-dimensional, since $\text{Hom}_{SO(2)}((2m+4), \sigma)$ is one-dimensional.

However, we did not assume that π is irreducible. If $\pi = \pi_1 \oplus \pi_2$ and if we run the above argument for each π_1 and π_2 , then we can write the chain

$$\begin{aligned} 1 + 1 &= \dim \text{Hom}_K(\pi_1, \tau(m, 0)) + \dim \text{Hom}_K(\pi_2, \tau(m, 0)) \\ &= \dim \text{Hom}_K(\pi, \tau(m, 0)) = 1, \end{aligned}$$

a contradiction. Thus $\Theta(\sigma)$ has a unique irreducible quotient. It contains $\tau(m, 0)$, with multiplicity one.

In the other direction, now π is irreducible and σ is a quotient of $\Theta(\pi)$, we start with the second sequence. That sequence ends with

$$\text{Hom}_K(F_m, \pi) \cong \text{Hom}_K(\tau(m, 0), \pi),$$

since π does not contain K -types $\tau(n, 0)$ with $n < m$. Next, we run the first sequence. The conclusion is that all spaces have the same dimension d , equal to

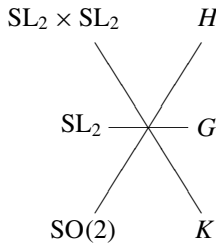
the multiplicity of $\tau(m, 0)$ in π which is not 0. Again, we did not assume that σ is irreducible. So, if $\sigma = \sigma_1 \oplus \sigma_2$ and we run the same argument for each σ_1 and σ_2 , since

$$\dim \text{Hom}_{\text{SO}(2)}(\sigma_1, (2m+4)) + \dim \text{Hom}_{\text{SO}(2)}(\sigma_2, (2m+4)) = \dim \text{Hom}_{\text{SO}(2)}(\sigma, (2m+4)),$$

we arrive at $d + d = d$, a contradiction. □

3. Computing lifts of types

In this section we verify the expressions for $\Theta(\tau(m, n))$ and $\Theta(2m+4)$ used in the proof of the main result. As indicated in the introduction, we use the following see-saw diagram in $G(J)$:



Here H is a simply connected, hermitian symmetric group of absolute type E_6 . Our SL_2 , the centralizer of G , sits diagonally in $\text{SL}_2 \times \text{SL}_2$, the centralizer of K . A word of caution here. If we pick a different SL_2 in $\text{SL}_2 \times \text{SL}_2$, the one consisting of all (g, hgh^{-1}) , where $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then G and H in the above are replaced by their compact forms. In other words, it is important how we identify groups isomorphic to $\text{SL}_2(\mathbb{R})$.

Let (e, h, f) be an \mathfrak{sl}_2 -triple such that $\mathbb{C} \cdot h$ is the Lie algebra of $\text{SO}(2)$. For an integer $n > 0$, let $\delta(n)$ be the irreducible lowest weight n module. Let v_n be a nonzero lowest weight vector. Let $\bar{\delta}(m)$ be the complex conjugate of $\delta(m)$. It is the irreducible highest weight $-m$ module. Observe that there is a natural map

$$U(\mathfrak{sl}_2) \otimes_{U(\mathfrak{so}(2))} (n - m) \rightarrow \delta(n) \otimes \bar{\delta}(m)$$

where $1 \in \mathbb{C} \cong (n - m)$ is mapped to $v_n \otimes v_m$. Since $\delta(n)$ is a free $\mathbb{C}[e]$ -module generated by v_n and $\bar{\delta}(m)$ is a free $\mathbb{C}[f]$ -module generated by v_m , the above map is easily checked to be an isomorphism.

Proposition 3.1. *Let $\tau(m, n)$ be the irreducible K -type as previously. Then*

$$\Theta(\tau(m, n)) \cong U(\mathfrak{sl}_2) \otimes_{U(\mathfrak{so}(2))} \otimes (2m+4).$$

Proof. Since the centralizer of K is $SL_2 \times SL_2$, $\Theta(\tau(m, n))$ is naturally an $SL_2 \times SL_2$ -module. By [13, Proposition 3.3.3] (careful with SL_2 's) we have

$$\Theta(\tau(m, n)) \cong \delta(2m + n + 8) \otimes \bar{\delta}(n + 4).$$

In view of the discussion above, and $(2m + n + 8) - (n + 4) = 2m + 4$, the proposition follows. \square

It remains to discuss $\Theta(2m + 4)$. Let L be a maximal compact subgroup of H . We can assume that $K \subset L$. Let \mathfrak{h} and \mathfrak{l} be the complex Lie algebras of H and L . Since H/L is a hermitian symmetric space, \mathfrak{l} is a Levi subalgebra such that $[\mathfrak{l}, \mathfrak{l}]$ is a simple Lie algebra of type D_5 . We have a Cartan decomposition

$$\mathfrak{h} = \bar{\mathfrak{u}} + \mathfrak{l} + \mathfrak{u}$$

such that $\mathfrak{q} = \mathfrak{l} + \mathfrak{u}$ is a parabolic subalgebra. If F is a finite-dimensional \mathfrak{l} -module, we can define a highest weight module

$$U(\mathfrak{h}) \otimes_{U(\mathfrak{q})} F \cong U(\bar{\mathfrak{u}}) \otimes F.$$

We now restrict this module to $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$. Recall that \mathfrak{p} is 16-dimensional spin-module. On the other hand, $\bar{\mathfrak{u}}$ and \mathfrak{u} are two 16-dimensional spin modules for $[\mathfrak{l}, \mathfrak{l}]$, the simple algebra of type D_5 . Hence \mathfrak{p} must embed diagonally into $\bar{\mathfrak{u}} + \mathfrak{u}$. Now it is not difficult to check that the natural map

$$U(\mathfrak{g}) \otimes_{U(\mathfrak{k})} F \rightarrow U(\mathfrak{h}) \otimes_{U(\mathfrak{q})} F$$

given by the identity on $1 \otimes F$ is an isomorphism. We are ready to prove the following:

Proposition 3.2. *For m integer, $\Theta(2m + 4)$ is a quotient of $U(\mathfrak{g}) \otimes_{U(\mathfrak{k})} F_m$ where $F_m = \mathbb{C}$ if $m \leq 0$; otherwise*

$$F_m = \tau(0, 0) \oplus \tau(1, 0) \oplus \dots \oplus \tau(m, 0).$$

Proof. $\Theta(2m + 4)$ is an (\mathfrak{h}, L) -module, determined in [12, Section 6]. It is a quotient of the Verma module $U(\mathfrak{h}) \otimes_{U(\mathfrak{q})} F_m$ where F_m is a one dimensional representation of L if $m \leq 0$. Otherwise F_m , restricted to $[L, L]$, is irreducible with the highest weight $(m, 0, 0, 0, 0)$. The restriction of this representation to K is the claimed sum. \square

References

- [1] P. Bakic, H. Y. Loke, and G. Savin, “Exceptional dual pair correspondences; case of real groups of split rank one”, preprint, 2025. arXiv 2508.01551
- [2] W. T. Gan and G. Savin, “Howe duality and dichotomy for exceptional theta correspondences”, *Invent. Math.* **232**:1 (2023), 1–78. MR

- [3] B. H. Gross and G. Savin, “Motives with Galois group of type G_2 : an exceptional theta-correspondence”, *Compositio Math.* **114**:2 (1998), 153–217. MR
- [4] R. Howe, “Transcending classical invariant theory”, *J. Amer. Math. Soc.* **2**:3 (1989), 535–552. MR
- [5] H. H. Kim, “Exceptional modular form of weight 4 on an exceptional domain contained in \mathbb{C}^{27} ”, *Rev. Mat. Iberoamericana* **9**:1 (1993), 139–200. MR
- [6] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, AMS Colloquium Publications **44**, American Mathematical Society, 1998. MR
- [7] T. Kobayashi and G. Savin, “Global uniqueness of small representations”, *Math. Z.* **281**:1-2 (2015), 215–239. MR
- [8] M. Koecher, “Imbedding of Jordan algebras into Lie algebras, I”, *Amer. J. Math.* **89** (1967), 787–816. MR
- [9] J. Lepowsky, “Multiplicity formulas for certain semisimple Lie groups”, *Bull. Amer. Math. Soc.* **77** (1971), 601–605. MR
- [10] J.-S. Li, “The correspondences of infinitesimal characters for reductive dual pairs in simple Lie groups”, *Duke Math. J.* **97**:2 (1999), 347–377. MR
- [11] A. L. Onishchik and E. B. Vinberg (editors), *Lie groups and Lie algebras III: structure of Lie groups and Lie algebras*, Encyclopaedia of Mathematical Sciences **41**, Springer, 1994.
- [12] P. Pandžić, A. Prlič, G. Savin, V. Souček, and V. Tuček, “On the classification of unitary highest weight modules in the exceptional cases”, *J. Algebra* **684** (2025), 524–562. MR
- [13] Y. Shan, “Exceptional theta correspondence $\mathbf{F}_4 \times \mathbf{PGL}_2$ for level one automorphic representations”, preprint, 2025. arXiv 2501.19101

Received December 4, 2025. Revised January 8, 2026.

GORDAN SAVIN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UT
UNITED STATES
gordan.savin@utah.edu

LIOUVILLE THEOREMS AND NEW GRADIENT ESTIMATES FOR POSITIVE SOLUTIONS TO $\Delta_p u + au^q = 0$ ON A COMPLETE MANIFOLD

YOUDE WANG AND LIQIN ZHANG

We use the Saloff-Coste Sobolev inequality and the Nash–Moser iteration method to study the local and global behaviors of positive solutions to the nonlinear elliptic equation $\Delta_p u + au^q = 0$ defined on a complete Riemannian manifold (M, g) with Ricci lower bound, where $p > 1$ is a constant and $\Delta_p u = \operatorname{div}(|\nabla u|^{p-2} \nabla u)$ is the usual p -Laplace operator. Under certain assumptions on a , p and q , we derive some gradient estimates and Liouville type theorems for positive solutions to the above equation. In particular, under certain assumptions on a , p and q we show whether or not the exact Cheng–Yau log-gradient estimates for the positive solutions to $\Delta_p u + au^q = 0$ on (M, g) with Ricci lower bound hold true is equivalent to whether or not the positive solutions to this equation fulfill Harnack inequality, and hence some new Cheng–Yau log-gradient estimates are established.

1. Introduction	396
2. Preliminaries	402
3. Gradient estimates	404
3.1. Estimate for the linearized operator of p -Laplace	404
3.2. Deducing the main integral inequality	412
3.3. L^{β_1} -bound of gradient in a geodesic ball with radius $3R/4$	417
3.4. Moser iteration for positive solutions of (1-1)	420
4. Proof of the main theorem and its consequences	422
Acknowledgements	425
References	425

Zhang is the corresponding author.

MSC2020: 35A01, 35B09, 35R01.

Keywords: gradient estimate, Nash–Moser iteration, Liouville type theorem.

1. Introduction

Gradient estimates are a fundamental technique in the study of partial differential equations on a Riemannian manifold. They can be used to deduce Liouville-type theorems [1; 2; 23; 12; 7; 19], to derive Harnack inequalities [23; 12], to infer local and global behavior of solutions, to study the geometry of manifolds [4; 20; 12; 11], and so on.

On the other hand, it is well-known that Liouville’s theorem has had a huge impact across many fields, such as complex analysis, partial differential equations, geometry, probability, discrete mathematics and complex and algebraic geometry. The impact of the Liouville theorem has been even larger as the starting point of many further developments. For details on the Liouville properties of harmonic functions and some related theory of function on a manifold we refer to an expository paper [5] written by T. H. Colding (see also [4]).

In this paper, we are concerned with the equation

$$(1-1) \quad \Delta_p u + au^q = 0$$

defined on a complete Riemannian manifold (M, g) equipped with a metric g , where $p > 1, a, q$ are constants and

$$\Delta_p u = \operatorname{div} (|\nabla u|^{p-2} \nabla u)$$

is the usual p -Laplace operator.

For simplicity, we will focus on equation (1-1) and try to establish some new gradient estimates on the positive solutions to this equation. Now we recall some relative results in the previous literature with the equation.

In the case that M is an Euclidean space, this equation was studied by Serrin and Zou in [15] and some Liouville theorems and universal estimates were established. Very recently, J. He, together with one of us (Wang) and G. Wei, [9] adopted a new way to employ Nash–Moser iteration to study the gradient estimates of this equation on a complete Riemannian manifold.

The new estimate

$$(1-2) \quad \frac{|\nabla u|^2}{u^2} + au^{q-1} \leq \frac{2n}{2-n \max\{0, q-1\}} \left(\frac{C_1^2(n-1)(1+\sqrt{\kappa}R) + C_2}{R^2} + 2\kappa + \frac{2nC_1^2}{(2+n \max\{0, q-1\})R^2} \right)$$

was obtained in [13] in the case $p = 2$ if the Ricci curvature of the domain manifold satisfies $\operatorname{Ric}_g \geq -(n-1)\kappa$ and $q < \frac{n+2}{n}$. Obviously, this is a stronger estimate than the logarithmic gradient estimate (also see [10]). Wang and Wei [19] also derived Cheng–Yau-type gradient estimates for positive solutions to $\Delta u + u^q = 0$ under the

assumption

$$q \in \left(-\infty, \frac{n+1}{n-1} + \frac{2}{\sqrt{n(n-1)}}\right).$$

Shortly afterward, the authors of [9] extended the Cheng–Yau estimate to the range

$$q \in \left(-\infty, \frac{n+3}{n-1}\right).$$

Recently, Z. Lu extended the estimate (1-2) in [13] to the range $q \in \left(-\infty, \frac{n+3}{n-1}\right)$.

The first goal of this paper is to give gradient estimates for positive solutions with positive lower bounds to (1-1), different from the exact log-gradient estimate.

As a second goal we try to answer two natural questions:

- *Is the value $\frac{n+3}{n-1}(p-1)$ above optimal for deriving the exact Cheng–Yau estimates for a positive C^1 solution to (1-1) on a complete manifold with Ricci curvature bounded below?*
- *Does the exact Cheng–Yau estimate hold true if u is a C^1 smooth positive solution to (1-1) that satisfies the standard Harnack inequality?*

Inspired by [8; 9; 21; 22], in the present paper we use the Nash–Moser iteration method to study the gradient estimate and the Liouville property of equation (1-1), defined on a complete Riemannian manifold.

Statement of main results. By a solution u of (1-1) in an (arbitrary) domain Ω we mean a positive solution $u \in C^1(\Omega) \cap C^3(\tilde{\Omega})$, where $\tilde{\Omega} = \{x \in \Omega \mid |\nabla u(x)| \neq 0\}$. Any solution of (1-1) satisfies $u \in C^{1,\alpha}(\Omega)$ for some $\alpha \in (0, 1)$ (see [6; 16; 17], for example). Moreover, u is in fact smooth in $\tilde{\Omega}$.

For brevity we define

$$h := \beta(p-1) \left[\frac{p-n}{(n-1)^2} \beta + \frac{2}{n-1} \right],$$

$$\phi_\beta := \begin{cases} \sup_{B(x_0,R)} u & \text{if } 0 < \beta < 2, \\ 1 & \text{if } \beta = 2, \\ \inf_{B(x_0,R)} u & \text{if } \beta > 2. \end{cases}$$

We suppose that β satisfies the condition

$$(1-3) \quad \beta \in \begin{cases} \left(0, \frac{2(n-1)}{n-p}\right) & \text{if } 1 < p < n, \\ (0, +\infty) & \text{if } p \geq n. \end{cases}$$

Now, we state our main results.

Theorem 1.1. *Let $p > 1$ and let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Assume u is*

a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$. If the constants a, p and q satisfy either

$$\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) - h^{1/2} < q < \frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) + h^{1/2} \quad (a \neq 0)$$

or

$$a \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) - q \right] \geq 0,$$

where β is a constant satisfying (1-3), then there exists a positive constant $C = C(n, p, q, \beta)$ such that

$$(1-4) \quad \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1 + \kappa R^2}{R^2} \phi_\beta^{2-\beta}.$$

If $\beta = 2$, we have

$$\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) + h^{1/2} = \frac{n+3}{n-1}(p-1) \quad \text{and} \quad \frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) - h^{1/2} = p-1,$$

so this case recovers the conclusion of Theorem 1.1 in [9]. At the same time, from (1-4) we can infer that

$$\sup_{B(x_0, R/2)} |\nabla u|^2 \leq C \frac{1 + \kappa R^2}{R^2} \phi_\beta^2 = C \frac{1 + \kappa R^2}{R^2} \sup_{B(x_0, R)} u^2,$$

if p and q satisfy the assumptions of Theorem 1.1 with $\beta \in (0, 2)$.

For convenience, we define

$$\Psi(I) := \sup_{\beta \in I} \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) + h^{1/2} \right] \quad \text{and} \quad \Gamma(I) := \inf_{\beta \in I} \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) - h^{1/2} \right].$$

If $2 \in I$, we obviously have

$$\Psi(I) \geq \frac{n+3}{n-1}(p-1) \quad \text{and} \quad \Gamma(I) \leq p-1.$$

So, we always have

$$\Psi \left(\left(0, \frac{2(n-1)}{n-p} \right) \right) \geq \frac{n+3}{n-1}(p-1),$$

since $p > 1$, and hence $2 \in \left(0, \frac{2(n-1)}{n-p} \right)$.

We then obtain the following consequences of Theorem 1.1:

Corollary 1.2. *Let $p > 1$ and let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Assume u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$. Assume also that the constants a, p and q satisfy one of the following five conditions:*

- $a > 0, p \geq n$ and $q \in \mathbb{R}$.
- $a > 0, 1 < p < n$ and $q < \Psi \left(\left(0, \frac{2(n-1)}{n-p} \right) \right)$.

- $a < 0, p \geq n$ and $q > \Gamma((0, +\infty))$.
- $a < 0, 1 < p < n$ and $q > \Gamma((0, \frac{2(n-1)}{n-p}))$.
- $a = 0, p > 1$.

Then there exist positive constants $C = C(n, p, q)$ and $\beta = \beta(n, p, q) \in (0, +\infty)$ such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1 + \kappa R^2}{R^2} \phi_\beta^{2-\beta}.$$

Note that in case 2 ($a > 0$ and $1 < p < n$), if there exists a point $\beta_0 \in (0, \frac{2(n-1)}{n-p})$ such that

$$\Psi\left(\left(0, \frac{2(n-1)}{n-p}\right)\right) = \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) + h^{1/2}\right] \Big|_{\beta=\beta_0},$$

then the condition $q < \Psi((0, \frac{2(n-1)}{n-p}))$ can be relaxed to $q \leq \Psi((0, \frac{2(n-1)}{n-p}))$. In the other four cases, we can obtain similar conclusions.

Further, if $1 < p < n$, it is easy to see that

$$\frac{n+1}{n-p}(p-1) \leq \Psi\left(\left(0, \frac{2(n-1)}{n-p}\right)\right).$$

Usually, this is a strict inequality; for instance, if we let $n = 3$ and $p = 2$, then

$$(1-5) \quad \frac{\beta}{2} \cdot \frac{n+1}{n-1}(p-1) + h^{1/2} = \beta + \sqrt{\beta(1-\beta/4)}$$

and $\beta \in (0, 4)$. Hence, we can check that (1-5) attains its maximum at an interior point $\beta_0 = 2 + 4/\sqrt{5} \in (0, 4)$. Therefore, we get

$$\Psi((0, 4)) = 2 + \sqrt{5} > 4 = \frac{3+1}{3-2} \cdot (2-1).$$

But we also have

$$\Psi((0, 4)) = 2 + \sqrt{5} > \frac{n+3}{n-1}(p-1) = 3.$$

This indicates that for $q \geq \frac{n+3}{n-1}(p-1)$ one also derives the gradient estimate.

Corollary 1.3. *Let $p > 1$ and let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Assume u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$. If the constants a, p and q satisfy*

$$a < 0 \quad \text{and} \quad q > \Gamma((0, 2]),$$

then there exist positive constants $C = C(n, p, q)$ and $\beta = \beta(n, p, q) \in (0, 2]$ such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1 + \kappa R^2}{R^2} \sup_{B(x_0, R)} u^{2-\beta}.$$

By using Theorem 1.1, we then reach the following conclusion.

Corollary 1.4. *Let $p > 1$ and let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Assume that u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$ and that*

$$u(x) \geq b > 0, \quad x \in B(x_0, R).$$

Assume also that the constants a, p and q satisfy one of the following conditions:

- $a > 0, p \geq n$ and $q \in \mathbb{R}$.
- $a > 0, 1 < p < n$ and $q < \Psi\left(\left[2, \frac{2(n-1)}{n-p}\right]\right)$.
- $a < 0, p \geq n$ and $q > \Gamma\left([2, +\infty)\right)$.
- $a < 0, 1 < p < n$ and $q > \Gamma\left(\left[2, \frac{2(n-1)}{n-p}\right]\right)$.
- $a = 0, p > 1$.

Then there exist positive constants $C = C(n, p, q, b)$ and $\beta = \beta(n, p, q) \in [2, +\infty)$, such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1 + \kappa R^2}{R^2}.$$

The next result states that whether or not the exact Cheng–Yau log-gradient estimates for the positive solutions to equation (1-1) hold true is equivalent to whether or not the positive solutions to (1-1) fulfill Harnack’s inequality.

Theorem 1.5. *Let $p > 1$ and (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Assume u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$ satisfying the Harnack inequality*

$$\sup_{B(x_0, R)} u \leq l \inf_{B(x_0, R)} u.$$

Assume also that the constants a, q and p satisfy one of the following conditions:

- $a > 0, p \geq n$ and $q \in \mathbb{R}$.
- $a > 0, 1 < p < n$ and $q < \Psi\left(\left(0, \frac{2(n-1)}{n-p}\right)\right)$.
- $a < 0, p \geq n$ and $q > \Gamma\left(\left(0, +\infty\right)\right)$.
- $a < 0, 1 < p < n$ and $q > \Gamma\left(\left(0, \frac{2(n-1)}{n-p}\right)\right)$.
- $a = 0, p > 1$.

Then there exists a positive constant $C = C(n, p, q, l)$ such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq C \frac{1 + \kappa R^2}{R^2}.$$

Corollary 1.6. *Let $p > 1$ and let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with nonnegative Ricci curvature. Assume u is a positive solution to equation (1-1) on any given geodesic ball $B(x_0, 2R) \subset M$, and that it satisfies the Harnack inequality*

$$\sup_{B(x_0, R)} u \leq l \inf_{B(x_0, R)} u,$$

where l is independent of u and R . Assume also that the constants a, q and p satisfy one of the following conditions:

- $a > 0, p \geq n$ and $q \in \mathbb{R}$.
- $a > 0, 1 < p < n$ and $q < \Psi\left(\left(0, \frac{2(n-1)}{n-p}\right)\right)$.
- $a < 0, p \geq n$ and $q > \Gamma\left(\left(0, +\infty\right)\right)$.
- $a < 0, 1 < p < n$ and $q > \Gamma\left(\left(0, \frac{2(n-1)}{n-p}\right)\right)$.
- $a = 0, p > 1$.

Then there exist a positive constant $C = C(n, p, q, l)$ such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2}.$$

Conversely, if the above log-gradient estimate holds true, then, for any given $B(x_0, R) \subset M$, there holds

$$\sup_{B(x_0, R)} u \leq l \inf_{B(x_0, R)} u,$$

where l is independent of u and R .

If we consider (1-1) in \mathbb{R}^n ($n \geq 3$), we can achieve the following conclusion.

Corollary 1.7. *Assume u is a positive solution to equation (1-1) on the ball $B(x_0, 2R) \subset \mathbb{R}^n$. Assume also that the constants a, q and p satisfy one of the following conditions:*

- $a > 0, 1 < p < n, p \neq q$ and $q \in \left(p-1, \frac{(p-1)n}{n-p}\right)$.
- $a > 0, p \geq n, q \neq p$ and $q \in (0, +\infty)$.
- $a \geq 1$ and $1 < p = q < n < p^2$.
- $a \geq 1$ and $p = q \geq n$.

Then there exist a positive constant $C = C(n, p, q, a)$ such that

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2}.$$

The corollary tells us that $\frac{n+3}{n-1}(p-1)$ is not an optimal bound for deriving an exact Cheng–Yau-type log-gradient estimate, since

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2}$$

in the case $a > 0$, $p \geq n$, $q \neq p$ and $q \in (0, +\infty)$.

The next result is a direct consequence of Corollary 1.7.

Theorem 1.8. *Assume u is a positive solution to equation (1-1) on \mathbb{R}^n and $n \geq 3$. Assume also that the constants a , p and q satisfy one of the following conditions:*

- $a > 0$, $1 < p < n$, $p \neq q$ and $q \in (p-1, \frac{(p-1)n}{n-p})$.
- $a > 0$, $p \geq n$, $p \neq q$ and $q \in (0, +\infty)$.
- $a \geq 1$ and $1 < p = q < n < p^2$.
- $a \geq 1$ and $p = q \geq n$.

then (1-1) admits no positive solution.

In the above conclusions, we always suppose that $\dim M = n \geq 3$. In fact, for the case $\dim M = 2$ we can also obtain similar conclusions. Since the proofs are similar to the case $\dim M \geq 3$, we will not give details.

Main ideas of proof and the organization of paper. In order to give the gradient estimates, we consider the linearized operator \mathcal{L}_p of the p -Laplace operator at a solution u , and let \mathcal{L}_p act on an auxiliary function given by

$$F(u) = \frac{|\nabla u|^2}{u^\beta}, \quad \beta > 0.$$

The use of such an auxiliary function is inspired by the gradient estimates established in [18] for another equation related to Ricci solitons. Then, we need to establish some suitable pointwise estimate of $\mathcal{L}_p(F)$ using the techniques of Cheng and Yau [3; 23], so that we can take a Nash–Moser iteration scheme to give the L^∞ -norm of $F(u)$. Saloff-Coste’s Sobolev inequalities play an important role in our arguments.

Outline. In Section 2, we recall some background and establish important lemmas, which will play a key role in the Nash–Moser iteration process. In Section 3, the main body of this paper, we prove the gradient estimates. In Section 4, we give the proofs of the main theorem and its corollaries.

2. Preliminaries

Throughout this paper, we let (M, g) be an n -dimensional Riemannian manifold ($n \geq 3$), and ∇ denotes the Levi-Civita connection corresponding to the metric g .

We denote the volume form on (M, g) by

$$d \text{ vol} = \sqrt{\det(g_{ij})} dx_1 \wedge \cdots \wedge dx_n,$$

where (x_1, \dots, x_n) is a local coordinate chart, and for simplicity we usually omit the volume form of integral over M .

Definition 2.1. We say that $u \in C^1(M) \cap W_{loc}^{1,p}(M)$ is a weak solution of (1-1) if for all $\psi \in W_0^{1,p}(M)$ we have

$$\int_M |\nabla u|^{p-2} \langle \nabla u, \nabla \psi \rangle = \int_M au^q \psi.$$

Next, we recall the Saloff-Coste Sobolev inequalities (see [14]), which shall play a key role in our proof of the main theorems.

Lemma 2.2 (Saloff-Coste [14]). *Let (M, g) be a complete manifold with $\text{Ric} \geq -(n-1)\kappa$. For $n > 2$, there exists a positive constant C_n , depending only on n , such that for all $B \subset M$ of radius R and volume V we have for $h_1 \in C_0^\infty(B)$*

$$\|h_1\|_{L^{2n/(n-2)}(B)}^2 \leq \exp\{C_n(1 + \sqrt{\kappa}R)\} V^{-2/n} R^2 \left(\int_B |\nabla h_1|^2 + R^{-2} h_1^2 \right).$$

For $n = 2$, the above inequality holds with n replaced by any fixed $n' > 2$.

Now we consider the linearization operator \mathcal{L}_p of p -Laplace operator:

$$(2-1) \quad \mathcal{L}_p(\psi) = \text{div}[f^{p/2-1} A(\nabla \psi)],$$

where $f = |\nabla u|^2$ and

$$(2-2) \quad A(\nabla \psi) = \nabla \psi + (p-2)f^{-1} \langle \nabla \psi, \nabla u \rangle \nabla u.$$

We first derive an useful expression of $\mathcal{L}_p(f)$.

Lemma 2.3. *The equality*

$$\mathcal{L}_p(f) = \left(\frac{p}{2} - 1\right) f^{p/2-2} |\nabla f|^2 + 2f^{p/2-1} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) + 2 \langle \nabla \Delta_p u, \nabla u \rangle$$

holds pointwise in $\{x : f(x) > 0\}$.

Proof. By the definition of A in (2-2), we have

$$(2-3) \quad A(\nabla f) = \nabla f + (p-2)f^{-1} \langle \nabla f, \nabla u \rangle \nabla u.$$

Combining (2-1) and (2-3), we obtain

$$(2-4) \quad \mathcal{L}_p(f) = \left(\frac{p}{2}-1\right) f^{p/2-2} |\nabla f|^2 + f^{p/2-1} \Delta f + (p-2) \left(\frac{p}{2}-2\right) f^{p/2-3} \langle \nabla f, \nabla u \rangle^2 \\ + (p-2) f^{p/2-2} \langle \nabla \langle \nabla f, \nabla u \rangle, \nabla u \rangle + (p-2) f^{p/2-2} \langle \nabla f, \nabla u \rangle \Delta u.$$

At the same time, by the definition of the p -Laplacian, we have

$$(2-5) \quad 2 \langle \nabla \Delta_p u, \nabla u \rangle \\ = (p-2) \left(\frac{p}{2}-2\right) f^{p/2-3} \langle \nabla f, \nabla u \rangle^2 + (p-2) f^{p/2-2} \langle \nabla \langle \nabla f, \nabla u \rangle, \nabla u \rangle \\ + (p-2) f^{p/2-2} \langle \nabla f, \nabla u \rangle \Delta u + 2 f^{p/2-1} \langle \nabla \Delta u, \nabla u \rangle.$$

We combine (2-4) and (2-5) to obtain

$$(2-6) \quad \mathcal{L}_p(f) = \\ \left(\frac{p}{2}-1\right) f^{p/2-2} |\nabla f|^2 + f^{p/2-1} \Delta f + 2 \langle \nabla \Delta_p u, \nabla u \rangle - 2 f^{p/2-1} \langle \nabla \Delta u, \nabla u \rangle.$$

From (2-6) and the Bochner formula

$$\frac{1}{2} \Delta f = |\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u) + \langle \nabla \Delta u, \nabla u \rangle$$

we get

$$\mathcal{L}_p(f) = \left(\frac{p}{2}-1\right) f^{p/2-2} |\nabla f|^2 + 2 f^{p/2-1} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) + 2 \langle \nabla \Delta_p u, \nabla u \rangle.$$

Thus, we finish the proof. \square

In the last section, we are going to use the following lemmas. We denote by B_R a ball with radius R and let Ω be a domain in \mathbb{R}^n .

Lemma 2.4 [15, Theorem 4.1(a)]. *Suppose $n > m$ and $s \in (m, \frac{m(n-1)}{n-m})$. Let ω be a nonnegative weak solution of the differential inequality*

$$(2-7) \quad \omega^{s-1} \leq -\Delta_m \omega \leq \Lambda \omega^{s-1} \quad \text{in } \Omega,$$

for some constant $\Lambda > 1$. Then there is a constant $C = C(n, m, s, \Lambda) > 0$ such that

$$(2-8) \quad \sup_{B_R} \omega(x) \leq C \inf_{B_R} \omega(x).$$

Lemma 2.5 [15, Theorem 4.3(a)]. *Let $n \leq m$. Assume the hypotheses of Lemma 2.4, except that the condition $s \in (m, \frac{m(n-1)}{n-m})$ is replaced by $s \in (1, +\infty)$, that is, $\frac{m(n-1)}{n-m} = +\infty$. Then (2-8) is valid with $C = C(n, m, s, \Lambda) > 0$.*

3. Gradient estimates

3.1. Estimate for the linearized operator of p -Laplace. First, we need to give the pointwise estimate of $\mathcal{L}_p(F)$, where

$$F = \frac{f}{u^\beta} \quad (\beta > 0)$$

and \mathcal{L}_p is the linearized operator of p -Laplacian at u .

Lemma 3.1. *The equality*

$$\begin{aligned} \mathcal{L}_p(F) = & u^{-\beta} \mathcal{L}_p(f) + \beta(\beta+1)(p-1)u^{-\beta-2}f^{p/2+1} \\ & - \beta\left(1 + \frac{p}{2}\right)(p-1)u^{-\beta-1}f^{p/2-1}\langle \nabla f, \nabla u \rangle - \beta(p-1)u^{-\beta-1}f^{p/2}\Delta u \end{aligned}$$

holds pointwise in $\{x : f(x) > 0\}$.

Proof. By the definition of A in (2-2), we have

$$(3-1) \quad A(\nabla F) = u^{-\beta}A(\nabla f) - \beta u^{-\beta-1}fA(\nabla u),$$

$$(3-2) \quad A(\nabla u) = (p-1)\nabla u,$$

$$(3-3) \quad A(\nabla f) = \nabla f + (p-2)f^{-1}\langle \nabla u, \nabla f \rangle \nabla u.$$

Combining (2-1) and (3-1), we obtain

$$(3-4) \quad \mathcal{L}_p(F) = \operatorname{div} [u^{-\beta}f^{p/2-1}A(\nabla f)] - \beta \operatorname{div} [u^{-\beta-1}f^{p/2}A(\nabla u)].$$

Direct computation shows that

$$(3-5) \quad \begin{aligned} \operatorname{div} [u^{-\beta}f^{p/2-1}A(\nabla f)] \\ = -\beta u^{-\beta-1}f^{p/2-1}\langle A(\nabla f), \nabla u \rangle + u^{-\beta} \operatorname{div} [f^{p/2-1}A(\nabla f)] \end{aligned}$$

and

$$(3-6) \quad \begin{aligned} \operatorname{div} [u^{-\beta-1}f^{p/2}A(\nabla u)] = & -(\beta+1)u^{-\beta-2}f^{p/2}\langle A(\nabla u), \nabla u \rangle \\ & + \frac{p}{2}u^{-\beta-1}f^{p/2-1}\langle A(\nabla u), \nabla f \rangle + u^{-\beta-1}f^{p/2} \operatorname{div} A(\nabla u). \end{aligned}$$

By substituting (2-1) and (3-3) into (3-5), we have

$$(3-7) \quad \operatorname{div} [u^{-\beta}f^{p/2-1}A(\nabla f)] = -\beta(p-1)u^{-\beta-1}f^{p/2-1}\langle \nabla f, \nabla u \rangle + u^{-\beta}\mathcal{L}_p(f).$$

Substituting (3-2) into (3-6) leads to

$$(3-8) \quad \begin{aligned} \operatorname{div} [u^{-\beta-1}f^{p/2}A(\nabla u)] = & -(\beta+1)(p-1)u^{-\beta-2}f^{p/2+1} \\ & + \frac{p}{2}(p-1)u^{-\beta-1}f^{p/2-1}\langle \nabla f, \nabla u \rangle + (p-1)u^{-\beta-1}f^{p/2}\Delta u. \end{aligned}$$

Now, we plug (3-7) and (3-8) into (3-4) to derive the required equality, and hence finish the proof of Lemma 3.1. \square

Lemma 3.2. *Let u be a positive solution of equation (1-1) in $\Omega \subset M$. Then*

$$(3-9) \quad \begin{aligned} \mathcal{L}_p(F) = & \left(\frac{p}{2} - 1\right)u^\beta f^{p/2-2}|\nabla F|^2 \\ & + 2a\left[\frac{\beta}{2}(p-1) - q\right]u^{q-\beta-1}f - p\beta u^{-1}f^{p/2-1}\langle \nabla F, \nabla u \rangle \\ & + 2u^{-\beta}f^{p/2-1}(|\nabla \nabla u|^2 + \operatorname{Ric}(\nabla u, \nabla u)) + \left[-\frac{1}{2}p\beta^2 + (p-1)\beta\right]u^{-\beta-2}f^{p/2+1} \end{aligned}$$

holds pointwise in $\{x \in \Omega : f(x) > 0\}$.

Proof. By summarizing Lemmas 2.3 and 3.1 we can achieve that

$$(3-10) \quad \mathcal{L}_p(F) = u^{-\beta} \times \left[\left(\frac{p}{2}-1\right) f^{p/2-2} |\nabla f|^2 + 2f^{p/2-1} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) + 2\langle \nabla \Delta_p u, \nabla u \rangle \right] \\ + \beta(\beta+1)(p-1)u^{-\beta-2} f^{p/2+1} \\ - \beta\left(1+\frac{p}{2}\right)(p-1)u^{-\beta-1} f^{p/2-1} \langle \nabla f, \nabla u \rangle - \beta(p-1)u^{-\beta-1} f^{p/2} \Delta u.$$

Since $F = f/u^\beta$, we can infer that

$$\nabla f = \beta u^{-1} f \nabla u + u^\beta \nabla F.$$

Hence, we have

$$(3-11) \quad \langle \nabla f, \nabla u \rangle = \beta u^{-1} f^2 + u^\beta \langle \nabla F, \nabla u \rangle,$$

$$(3-12) \quad |\nabla f|^2 = \beta^2 u^{-2} f^3 + 2\beta u^{\beta-1} f \langle \nabla F, \nabla u \rangle + u^{2\beta} |\nabla F|^2.$$

Substituting (1-1), (3-11) and (3-12) into (3-10), we obtain

$$(3-13) \quad \mathcal{L}_p(F) = \left(\frac{p}{2}-1\right) u^\beta f^{p/2-2} |\nabla F|^2 - 2a q u^{q-\beta-1} f \\ + 2u^{-\beta} f^{p/2-1} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) \\ + \left[-\frac{1}{2}(p^2-2p+2)\beta^2 + (p-1)\beta\right] u^{-\beta-2} f^{p/2+1} \\ - \frac{1}{2}(p^2-p+2)\beta u^{-1} f^{p/2-1} \langle \nabla F, \nabla u \rangle - \beta(p-1)u^{-\beta-1} f^{p/2} \Delta u.$$

From (1-1) and the equality

$$\Delta_p u = \text{div}(f^{p/2-1} \nabla u) = \left(\frac{p}{2}-1\right) f^{p/2-2} \langle \nabla f, \nabla u \rangle + f^{p/2-1} \Delta u,$$

it is easy to verify that

$$(3-14) \quad \Delta u = \left(1 - \frac{p}{2}\right) f^{-1} \langle \nabla f, \nabla u \rangle - a u^q f^{1-\frac{p}{2}}.$$

Substituting (3-11) into the above (3-14) yields

$$(3-15) \quad \Delta u = \left(1 - \frac{p}{2}\right) u^\beta f^{-1} \langle \nabla F, \nabla u \rangle + \left(1 - \frac{p}{2}\right) \beta u^{-1} f - a u^q f^{1-\frac{p}{2}}.$$

Hence, we substitute (3-15) into (3-13) to derive the required equality. Thus we complete the proof of Lemma 3.2. \square

Lemma 3.3. *Let $\alpha > 1$ and let u be a positive solution of equation (1-1) in $\Omega \subset M$. Then, the following holds pointwise in $\{x \in \Omega : f(x) > 0\}$:*

$$(3-16) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}_p(F^\alpha) \\ = \left(\alpha + \frac{p}{2} - 2\right) f^{p/2-1} |\nabla F|^2 + (\alpha-1)(p-2) f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 \\ + 2u^{-2\beta} f^{p/2} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) - \beta p u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle \\ + 2a \left[\frac{\beta}{2}(p-1) - q\right] u^{q-2\beta-1} f^2 + \left[-\frac{p}{2}\beta^2 + (p-1)\beta\right] u^{-2\beta-2} f^{p/2+2}$$

Proof. By the definition of A in (2-2), we have

$$(3-17) \quad A(\nabla(F^\alpha)) = \alpha F^{\alpha-1} A(\nabla F)$$

and

$$(3-18) \quad \langle A(\nabla F), \nabla F \rangle = |\nabla F|^2 + (p-2)f^{-1} \langle \nabla F, \nabla u \rangle^2.$$

Combining (2-1), (3-17) and (3-18), we obtain

$$(3-19) \quad \begin{aligned} \mathcal{L}_p(F^\alpha) &= \alpha \operatorname{div}[F^{\alpha-1} f^{p/2-1} A(\nabla F)] \\ &= \alpha(\alpha-1)F^{\alpha-2} f^{p/2-1} \langle A(\nabla F), \nabla F \rangle + \alpha F^{\alpha-1} \mathcal{L}_p(F) \\ &= \alpha(\alpha-1)F^{\alpha-2} f^{p/2-1} [|\nabla F|^2 + (p-2)f^{-1} \langle \nabla F, \nabla u \rangle^2] + \alpha F^{\alpha-1} \mathcal{L}_p(F). \end{aligned}$$

In view of (3-9) and (3-19), we can derive the required equality and complete the proof of Lemma 3.3. \square

Next, we need to consider the pointwise estimate of $\mathcal{L}_p(F^\alpha)$. We begin with two lemmas.

Lemma 3.4. *Let $a \neq 0$, $\alpha > 1$ and let u be a positive solution of (1-1) in $\Omega \subset M$ with $\operatorname{Ric} \geq -(n-1)\kappa$. Set*

$$(3-20) \quad H = \beta(p-1) \left[\frac{p-n}{2(n-1)} \beta + 1 \right] - \frac{\left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q \right]^2}{\frac{2}{n-1} - \frac{p-1}{(n-1)^2} \left[\alpha + \frac{p-n}{2(n-1)} \right]^{-1}}.$$

Then we have, pointwise in $\{x \in \Omega : f(x) > 0\}$,

$$(3-21) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + (p-1)\beta \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle + H u^{-2\beta-2} f^{p/2+2}.$$

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal frame of TM on a domain with $f \neq 0$ such that $e_1 = \nabla u / |\nabla u|$. We infer the equalities

$$(3-22) \quad 4 \sum_{i=1}^n u_{1i}^2 = f^{-1} |\nabla f|^2, \quad u_{11} = \frac{\langle \nabla f, \nabla u \rangle}{2f},$$

$$(3-23) \quad \Delta_p u = (p-1) f^{p/2-1} u_{11} + f^{p/2-1} \sum_{i=2}^n u_{ii}.$$

Substituting (3-11) and (3-12) into the two equations in (3-22) leads respectively to

$$(3-24) \quad 4 \sum_{i=1}^n u_{1i}^2 = u^{2\beta} f^{-1} |\nabla F|^2 + \beta^2 u^{-2} f^2 + 2\beta u^{\beta-1} \langle \nabla F, \nabla u \rangle$$

and

$$(3-25) \quad 2u_{11} = \beta u^{-1} f + u^\beta f^{-1} \langle \nabla F, \nabla u \rangle.$$

Combining (1-1) and (3-23), we obtain

$$(3-26) \quad \left(\sum_{i=2}^n u_{ii} \right)^2 = [(p-1)u_{11} + af^{1-p/2}u^q]^2.$$

By substituting (3-25) into (3-26), we have

$$(3-27) \quad \left(\sum_{i=2}^n u_{ii} \right)^2 = \left[\frac{p-1}{2}\beta u^{-1}f + \frac{p-1}{2}u^\beta f^{-1}\langle \nabla F, \nabla u \rangle + af^{1-\frac{p}{2}}u^q \right]^2.$$

By omitting some nonnegative terms in $|\nabla \nabla u|^2$ and using Cauchy's inequality, we arrive at

$$(3-28) \quad |\nabla \nabla u|^2 \geq \sum_{i=1}^n u_{1i}^2 + \sum_{i=2}^n u_{ii}^2 \geq \sum_{i=1}^n u_{1i}^2 + \frac{1}{n-1} \left(\sum_{i=2}^n u_{ii} \right)^2.$$

We plug (3-24) and (3-27) into (3-28) to obtain

$$(3-29) \quad |\nabla \nabla u|^2 \geq \frac{1}{4}u^{2\beta}f^{-1}|\nabla F|^2 + \frac{1}{4}\beta^2u^{-2}f^2 + \frac{1}{2}\beta u^{\beta-1}\langle \nabla F, \nabla u \rangle \\ + \frac{1}{n-1} \left[\frac{p-1}{2}\beta u^{-1}f + \frac{p-1}{2}u^\beta f^{-1}\langle \nabla F, \nabla u \rangle + af^{1-\frac{p}{2}}u^q \right]^2.$$

By expanding the last term of (3-29), we obtain

$$|\nabla \nabla u|^2 \geq \frac{1}{4}u^{2\beta}f^{-1}|\nabla F|^2 + \frac{(p-1)^2}{4(n-1)}u^{2\beta}f^{-2}\langle \nabla F, \nabla u \rangle^2 + \frac{\beta^2}{4} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-2}f^2 \\ + \frac{\beta}{2} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{\beta-1}\langle \nabla F, \nabla u \rangle + \frac{a^2}{n-1}u^{2q}f^{2-p} + a\frac{p-1}{n-1}\beta u^{q-1}f^{2-p/2} \\ + a\frac{p-1}{n-1}\beta u^{q-1}f^{2-p/2} + a\frac{p-1}{n-1}u^{q+\beta}f^{-\frac{p}{2}}\langle \nabla F, \nabla u \rangle.$$

Using this inequality and $\text{Ric} \geq -(n-1)\kappa$, we have

$$(3-30) \quad 2u^{-2\beta}f^{p/2}(|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) \\ \geq -2(n-1)\kappa u^{-2\beta}f^{p/2+1} + \frac{1}{2}f^{p/2-1}|\nabla F|^2 + \beta \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-\beta-1}f^{p/2}\langle \nabla F, \nabla u \rangle \\ + \frac{\beta^2}{2} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-2\beta-2}f^{p/2+2} + \frac{(p-1)^2}{2(n-1)}f^{p/2-2}\langle \nabla F, \nabla u \rangle^2 + \frac{2a^2}{n-1}u^{2q-2\beta}f^{2-p/2} \\ + a\frac{2(p-1)}{n-1}\beta u^{q-2\beta-1}f^2 + a\frac{2(p-1)}{n-1}u^{q-\beta}\langle \nabla F, \nabla u \rangle.$$

Substituting (3-30) into (3-16) yields

$$(3-31) \quad \frac{1}{\alpha}F^{2-\alpha}\mathcal{L}_p(F^\alpha) \\ \geq -2(n-1)\kappa u^{-2\beta}f^{p/2+1} + 2a \left[\frac{\beta}{2}\frac{n+1}{n-1}(p-1) - q \right] u^{q-2\beta-1}f^2 \\ + \left(\alpha + \frac{p}{2} - \frac{3}{2} \right) f^{p/2-1}|\nabla F|^2 + \left[(\alpha-1)(p-2) + \frac{(p-1)^2}{2(n-1)} \right] f^{p/2-2}\langle \nabla F, \nabla u \rangle^2 \\ + (p-1)\beta \left[\frac{p-n}{2(n-1)}\beta + 1 \right] u^{-2\beta-2}f^{p/2+2} + (p-1)\beta \frac{p-n}{n-1}u^{-\beta-1}f^{p/2}\langle \nabla F, \nabla u \rangle \\ + \frac{2a^2}{n-1}u^{2q-2\beta}f^{2-p/2} + a\frac{2(p-1)}{n-1}u^{q-\beta}\langle \nabla F, \nabla u \rangle.$$

By using $p > 1$, $\alpha > 1$ and

$$(3-32) \quad f^{p/2-1} |\nabla F|^2 \geq f^{p/2-2} \langle \nabla F, \nabla u \rangle^2,$$

we arrive at

$$(3-33) \quad \left(\alpha + \frac{p}{2} - \frac{3}{2}\right) f^{p/2-1} |\nabla F|^2 + \left[(\alpha - 1)(p - 2) + \frac{(p-1)^2}{2(n-1)}\right] f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 A \\ \geq (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right] f^{p/2-2} \langle \nabla F, \nabla u \rangle^2.$$

By substituting (3-33) into (3-31), we obtain

$$(3-34) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}_p(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q\right] u^{q-2\beta-1} f^2 \\ + (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right] f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 \\ + (p-1)\beta \left[\frac{p-n}{2(n-1)}\beta + 1\right] u^{-2\beta-2} f^{p/2+2} \\ + (p-1)\beta \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle + \frac{2a^2}{n-1} u^{2q-2\beta} f^{2-p/2} \\ + a \frac{2(p-1)}{n-1} u^{q-\beta} \langle \nabla F, \nabla u \rangle.$$

Noting that $\alpha > 1$ and using the inequality $a_1^2 - 2a_1a_2 \geq -a_2^2$, we infer

$$(3-35) \quad (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right] f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 + a \frac{2(p-1)}{n-1} u^{q-\beta} \langle \nabla F, \nabla u \rangle \\ \geq -\frac{a^2}{(n-1)^2} (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right]^{-1} u^{2q-2\beta} f^{2-p/2}.$$

We substitute (3-35) into (3-34) to obtain

$$(3-36) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q\right] u^{q-2\beta-1} f^2 \\ + (p-1)\beta \left[\frac{p-n}{2(n-1)}\beta + 1\right] u^{-2\beta-2} f^{p/2+2} \\ + (p-1)\beta \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle \\ + \left\{ \frac{2a^2}{n-1} - \frac{a^2}{(n-1)^2} (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right]^{-1} \right\} u^{2q-2\beta} f^{2-p/2}.$$

Using $\alpha > 1$ and the inequality $a_1^2 - 2a_1a_2 \geq -a_2^2$ again, we obtain

$$(3-37) \quad \left\{ \frac{2a^2}{n-1} - \frac{a^2}{(n-1)^2} (p-1) \left[\alpha + \frac{p-n}{2(n-1)}\right]^{-1} \right\} u^{2q-2\beta} f^{2-p/2} \\ + 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q\right] u^{q-2\beta-1} f^2 \\ \geq -\frac{\left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q\right]^2}{\frac{2}{n-1} - \frac{p-1}{(n-1)^2} \left[\alpha + \frac{p-n}{2(n-1)}\right]^{-1}} u^{-2\beta-2} f^{p/2+2}.$$

Now, we plug (3-37) into (3-36) to deduce the desired inequality and hence complete the proof of Lemma 3.4. \square

Lemma 3.5. *Let $\alpha > 1$ and let u be a positive solution of equation (1-1) in $\Omega \subset M$ with $\text{Ric} \geq -(n-1)\kappa$. Then we have, pointwise in $\{x \in \Omega : f(x) > 0\}$,*

$$(3-38) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}_p(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + \beta(p-1) \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle \\ + 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q \right] u^{q-2\beta-1} f^2 + \frac{\beta}{2} (p-1) \left(2 + \frac{p-n}{n-1} \beta \right) u^{-2\beta-2} f^{p/2+2}.$$

Proof. By using the inequality $(a_1 + a_2)^2 \geq a_1^2 + 2a_1a_2$, we have

$$(3-39) \quad \left[\frac{p-1}{2} \beta u^{-1} f + \frac{p-1}{2} u^\beta f^{-1} \langle \nabla F, \nabla u \rangle + a f^{1-\frac{p}{2}} u^q \right]^2 \\ \geq \frac{(p-1)^2}{4} \beta^2 u^{-2} f^2 + (p-1) \beta u^{-1} f \left(\frac{p-1}{2} u^\beta f^{-1} \langle \nabla F, \nabla u \rangle + a f^{1-\frac{p}{2}} u^q \right).$$

Substituting (3-39) into (3-29), we have

$$(3-40) \quad |\nabla \nabla u|^2 \geq \frac{1}{4} u^{2\beta} f^{-1} |\nabla F|^2 + \frac{\beta^2}{4} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-2} f^2 \\ + \frac{\beta}{2} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{\beta-1} \langle \nabla F, \nabla u \rangle + a \beta \frac{p-1}{n-1} u^{q-1} f^{2-p/2}.$$

By using (3-40) and the assumption $\text{Ric} \geq -(n-1)\kappa$, we have

$$(3-41) \quad 2u^{-2\beta} f^{p/2} (|\nabla \nabla u|^2 + \text{Ric}(\nabla u, \nabla u)) \\ \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + \frac{1}{2} f^{p/2-1} |\nabla F|^2 + \frac{\beta^2}{2} \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-2\beta-2} f^{p/2+2} \\ + \beta \left[1 + \frac{(p-1)^2}{n-1} \right] u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle + 2a \beta \frac{p-1}{n-1} u^{q-2\beta-1} f^2.$$

We plug (3-41) into (3-16) to derive

$$(3-42) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}_p(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} + \beta(p-1) \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle \\ + \left(\alpha + \frac{\beta}{2} - \frac{3}{2} \right) f^{p/2-1} |\nabla F|^2 + (\alpha-1)(p-2) f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 \\ + 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q \right] u^{q-2\beta-1} f^2 + \frac{\beta}{2} (p-1) \left(2 + \frac{p-n}{n-1} \beta \right) u^{-2\beta-2} f^{p/2+2}.$$

Noting that $p > 1$, $\alpha > 1$ and $f^{p/2-1} |\nabla F|^2 \geq f^{p/2-2} \langle \nabla F, \nabla u \rangle^2$, we arrive at

$$(3-43) \quad \left(\alpha + \frac{\beta}{2} - \frac{3}{2} \right) f^{p/2-1} |\nabla F|^2 + (\alpha-1)(p-2) f^{p/2-2} \langle \nabla F, \nabla u \rangle^2 \\ \geq \left(\alpha - \frac{1}{2} \right) (p-1) f^{p/2-2} \langle \nabla F, \nabla u \rangle^2.$$

In view of (3-42) and (3-43), we can derive the desired inequality and complete the proof of Lemma 3.5. \square

By using Lemmas 3.4 and 3.5, we can achieve the following pointwise estimate of $\mathcal{L}_p(F^\alpha)$.

Lemma 3.6. *Let u be a positive solution of equation (1-1) in $\Omega \subset M$ with $\text{Ric} \geq -(n-1)\kappa$. Set*

$$(3-44) \quad h = \beta(p-1) \left[\frac{p-n}{(n-1)^2} \beta + \frac{2}{n-1} \right]$$

and suppose that β satisfies

$$(3-45) \quad \beta \in \begin{cases} (0, \frac{2(n-1)}{n-p}) & \text{if } 1 < p < n, \\ (0, +\infty) & \text{if } p \geq n. \end{cases}$$

If the constants a , p and q satisfy either

$$\frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) - h^{1/2} < q < \frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) + h^{1/2} \quad (a \neq 0)$$

or

$$a \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) - q \right] \geq 0,$$

then there exists $\alpha > 1$ such that

$$(3-46) \quad \frac{1}{\alpha} \mathcal{L}(F^\alpha) \geq -2(n-1)\kappa u^{(p/2-1)\beta} F^{\alpha+p/2-1} \\ + Au^{(p/2)\beta-2} F^{\alpha+p/2} - Bu^{(\beta/2)(p-1)-1} F^{\alpha+(p-3)/2} |\nabla F|$$

pointwise in $\{x \in \Omega : f(x) > 0\}$, where A is a positive constant and

$$B = (p-1)\beta \frac{|n-p|}{n-1}.$$

Proof. Case 1: $\frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) - h^{1/2} < q < \frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) + h^{1/2}$, with $a \neq 0$.

This implies $\lim_{\alpha \rightarrow +\infty} H > 0$, where H is defined in (3-20). Thus, we can choose α_0 large enough that for any $\alpha > \alpha_0$,

$$(3-47) \quad H > 0.$$

Furthermore, we have

$$(3-48) \quad (p-1)\beta \frac{p-n}{n-1} u^{-\beta-1} f^{p/2} \langle \nabla F, \nabla u \rangle \geq -(p-1)\beta \frac{|n-p|}{n-1} u^{-\beta-1} f^{(p+1)/2} |\nabla F| \\ = -Bu^{-\beta-1} f^{(p+1)/2} |\nabla F|.$$

Combining (3-21) and (3-48), we can infer that

$$(3-49) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} \\ - Bu^{-\beta-1} f^{(p+1)/2} |\nabla F| + Hu^{-2\beta-2} f^{p/2+2} \quad (a \neq 0).$$

Combining (3-47) and (3-49) leads to (3-46).

Case 2: $a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q \right] \geq 0$.

In this case we have

$$(3-50) \quad 2a \left[\frac{\beta}{2} \frac{n+1}{n-1} (p-1) - q \right] u^{q-2\beta-1} f^2 \geq 0.$$

Combining (3-38), (3-48) and (3-50), we obtain

$$(3-51) \quad \frac{1}{\alpha} F^{2-\alpha} \mathcal{L}_p(F^\alpha) \geq -2(n-1)\kappa u^{-2\beta} f^{p/2+1} - Bu^{-\beta-1} f^{(p+1)/2} |\nabla F| \\ + \frac{\beta}{2} (p-1) \left(2 + \frac{p-n}{n-1} \beta \right) u^{-2\beta-2} f^{p/2+2}.$$

Since β satisfies (3-45), it is easy to see that

$$(3-52) \quad \frac{\beta}{2}(p-1)\left(2 + \frac{p-n}{n-1}\beta\right) > 0.$$

Combining the above, we complete the proof of Lemma 3.6. \square

3.2. Deducing the main integral inequality.

Lemma 3.7. *Let M be a complete manifold with $\text{Ric} \geq -(n-1)\kappa$ and let u be a positive solution of equation (1-1) in $B_{2R}(x_0) \subset M$. Then, there exist constants t large enough and $\mu_1 > 0$ such that*

$$\begin{aligned} \exp\{-C_n(1+\sqrt{\kappa}R)\} V^{2/n} R^{-2} &\|F^{\frac{p}{4}+\frac{\alpha-1}{2}+\frac{t}{2}}\eta\|_{L^{2n/(n-2)}(\Omega)}^2 + \mu t \phi_\beta^{\beta-2} \int_\Omega F^{t+\alpha+p/2} \eta^2 \\ &\leq ((n-1)\mu_1 t \kappa + R^{-2}) \int_\Omega F^{t+\alpha+p/2-1} \eta^2 + \mu_1 \int_\Omega F^{t+p/2+\alpha-1} |\nabla \eta|^2, \end{aligned}$$

where $\Omega = B_R(x_0)$, $\eta \in C_0^\infty(\Omega, \mathbb{R})$ is a nonnegative function and V is the volume of $B_R(x_0)$.

Proof. We choose a geodesic ball $\Omega = B_R(x_0) \subset M$ and a test function $\xi \cdot u^\lambda = F_\epsilon^t \eta^2 \cdot u^\lambda$, where $\eta \in C_0^\infty(\Omega, \mathbb{R})$ is nonnegative, $F_\epsilon = (F - \epsilon)^+$, $\epsilon > 0$, $t > 1$ and $\lambda \in \mathbb{R}$ are to be determined later. It follows from (2-1) that

$$\begin{aligned} &\frac{1}{\alpha} \int_\Omega \mathcal{L}_p(F^\alpha) \cdot \xi \cdot u^\lambda \\ &= - \int_\Omega \langle \nabla(\xi u^\lambda), f^{p/2-1} F^{\alpha-1} [\nabla F + (p-2)f^{-1} \langle \nabla u, \nabla F \rangle \nabla u] \rangle \\ &= - \int_\Omega f^{p/2-1} F^{\alpha-1} u^\lambda \langle \nabla F, \nabla \xi \rangle - \lambda \int_\Omega u^{\lambda-1} f^{p/2-1} F^{\alpha-1} \langle \nabla F, \nabla u \rangle \xi \\ &\quad - (p-2) \int_\Omega f^{p/2-2} F^{\alpha-1} u^\lambda \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \xi \rangle \\ &\quad - (p-2)\lambda \int_\Omega f^{p/2-1} F^{\alpha-1} u^{\lambda-1} \langle \nabla F, \nabla u \rangle \xi \\ &= - \int_\Omega f^{p/2-1} F^{\alpha-1} u^\lambda \langle \nabla F, \nabla \xi \rangle - (p-1)\lambda \int_\Omega f^{p/2-1} F^{\alpha-1} u^{\lambda-1} \langle \nabla F, \nabla u \rangle \xi \\ &\quad - (p-2) \int_\Omega f^{p/2-2} F^{\alpha-1} u^\lambda \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \xi \rangle. \end{aligned}$$

Since $\xi = F_\epsilon^t \eta^2$, we can achieve that

$$(3-53) \quad \begin{aligned} &\frac{1}{\alpha} \int_\Omega \mathcal{L}_p(F^\alpha) \cdot F_\epsilon^t \eta^2 \cdot u^\lambda \\ &= - \int_\Omega f^{p/2-1} F^{\alpha-1} u^\lambda \langle \nabla F, t F_\epsilon^{t-1} \eta^2 \nabla F + 2 F_\epsilon^t \eta \nabla \eta \rangle \\ &\quad - (p-1)\lambda \int_\Omega f^{p/2-1} F^{\alpha-1} u^{\lambda-1} \langle \nabla F, \nabla u \rangle F_\epsilon^t \eta^2 \\ &\quad - (p-2) \int_\Omega f^{p/2-2} F^{\alpha-1} u^\lambda \langle \nabla F, \nabla u \rangle \langle \nabla u, t F_\epsilon^{t-1} \eta^2 \nabla F + 2 F_\epsilon^t \eta \nabla \eta \rangle \end{aligned}$$

$$\begin{aligned}
&= -t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\
&\quad - 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla \eta \rangle \eta \\
&\quad - (p-1)\lambda \int_{\Omega} u^{(p/2-1)\beta+\lambda-1} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \eta^2 \\
&\quad - (p-2)t \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^{t-1} \langle \nabla F, \nabla u \rangle^2 \eta^2 \\
&\quad - 2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta.
\end{aligned}$$

Combining (3-46) and (3-53), we achieve

$$\begin{aligned}
&-t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\
&\quad - 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla \eta \rangle \eta \\
&\quad - (p-1)\lambda \int_{\Omega} u^{(p/2-1)\beta+\lambda-1} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \eta^2 \\
&\quad - (p-2)t \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^{t-1} \langle \nabla F, \nabla u \rangle^2 \eta^2 \\
&\quad - 2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta \\
&\geq -2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{\alpha+p/2-1} F_{\epsilon}^t \eta^2 \\
&\quad + A \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{\alpha+p/2} F_{\epsilon}^t \eta^2 - B \int_{\Omega} u^{\beta/2(p-1)-1+\lambda} F^{\alpha+(p-3)/2} |\nabla F| F_{\epsilon}^t \eta^2.
\end{aligned}$$

From this, we obtain

$$\begin{aligned}
(3-54) \quad &2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{\alpha+p/2-1} F_{\epsilon}^t \eta^2 \\
&\quad - 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla \eta \rangle \eta \\
&\quad - 2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta \\
&\quad + (B + (p-1)|\lambda|) \int_{\Omega} u^{(\beta/2)(p-1)-1+\lambda} F^{\alpha+u(p-3)/2} |\nabla F| F_{\epsilon}^t \eta^2 \\
&\geq t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\
&\quad + (p-2)t \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^{t-1} \langle \nabla F, \nabla u \rangle^2 \eta^2 \\
&\quad + A \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{\alpha+p/2} F_{\epsilon}^t \eta^2.
\end{aligned}$$

Set

$$\begin{aligned}
(3-55) \quad L_p = &t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\
&\quad + (p-2)t \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^{t-1} \langle \nabla F, \nabla u \rangle^2 \eta^2.
\end{aligned}$$

Then we need to consider two cases:

Case i: $p \geq 2$. Here we get from (3-34) that

$$(3-56) \quad L_p \geq t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2.$$

Case ii: $1 < p < 2$. Here, again from (3-34), we get

$$(3-57) \quad \begin{aligned} L_p &= t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\ &\quad - (2-p)t \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^{t-1} \langle \nabla F, \nabla u \rangle^2 \eta^2 \\ &\geq t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\ &\quad - (2-p)t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 \\ &= (p-1)t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2. \end{aligned}$$

Set

$$(3-58) \quad \theta(p) = \begin{cases} p-1 & \text{if } 1 < p < 2; \\ 1 & \text{if } p \geq 2. \end{cases}$$

Combining (3-55), (3-56), (3-57) and (3-58) yields

$$(3-59) \quad L_p \geq \theta(p)t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2.$$

By combining (3-54) and (3-59), we have

$$\begin{aligned} &2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{\alpha+p/2-1} F_{\epsilon}^t \eta^2 \\ &\quad - 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^t \langle \nabla F, \nabla \eta \rangle \eta \\ &\quad - 2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{p/2+\alpha-3} F_{\epsilon}^t \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta \\ &\quad + (B + (p-1)|\lambda|) \int_{\Omega} u^{(\beta/2)(p-1)-1+\lambda} F^{\alpha+(p-3)/2} |\nabla F| F_{\epsilon}^t \eta^2 \\ &\geq \theta(p)t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{p/2+\alpha-2} F_{\epsilon}^{t-1} |\nabla F|^2 \eta^2 + A \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{\alpha+p/2} F_{\epsilon}^t \eta^2. \end{aligned}$$

By letting $\epsilon \rightarrow 0^+$, we obtain

$$(3-60) \quad \begin{aligned} &2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+\alpha+p/2-1} \eta^2 \\ &\quad - 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-2} \langle \nabla F, \nabla \eta \rangle \eta \\ &\quad - 2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{t+p/2+\alpha-3} \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta \\ &\quad + (B + (p-1)|\lambda|) \int_{\Omega} u^{(\beta/2)(p-1)-1+\lambda} F^{t+\alpha+(p-3)/2} |\nabla F| \eta^2 \\ &\geq \theta(p)t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 + A \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{t+\alpha+p/2} \eta^2. \end{aligned}$$

By the absolute value inequality and the Cauchy inequality, we have

$$(3-61) \quad \begin{aligned} & (B+(p-1)|\lambda|) \int_{\Omega} u^{(\beta/2)(p-1)-1+\lambda} F^{t+\alpha+(p-3)/2} |\nabla F| \eta^2 \\ & \leq \frac{\theta(p)}{4} t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 \\ & \quad + \frac{1}{\theta(p)t} (B+(p-1)|\lambda|)^2 \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{t+\alpha+p/2} \eta^2, \end{aligned}$$

$$(3-62) \quad \begin{aligned} & -2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-2} \langle \nabla F, \nabla \eta \rangle \eta \\ & \leq 2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-2} |\nabla F| |\nabla \eta| \eta \\ & \leq \frac{\theta(p)}{4} t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 \\ & \quad + \frac{4}{\theta(p)t} \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-1} |\nabla \eta|^2, \end{aligned}$$

$$(3-63) \quad \begin{aligned} & -2(p-2) \int_{\Omega} u^{(p/2-2)\beta+\lambda} F^{t+p/2+\alpha-3} \langle \nabla F, \nabla u \rangle \langle \nabla u, \nabla \eta \rangle \eta \\ & \leq 2|p-2| \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-2} |\nabla F| |\nabla \eta| \eta \\ & \leq \frac{\theta(p)}{4} t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 \\ & \quad + \frac{4}{\theta(p)t} (p-2)^2 \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-1} |\nabla \eta|^2. \end{aligned}$$

Substituting (3-61), (3-62) and (3-63) into (3-60), we obtain

$$(3-64) \quad \begin{aligned} & \frac{\theta(p)}{4} t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 + A \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{t+\alpha+p/2} \eta^2 \\ & \leq 2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+\alpha+p/2-1} \eta^2 \\ & \quad + \frac{1}{\theta(p)t} (B+(p-1)|\lambda|)^2 \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{t+\alpha+p/2} \eta^2 \\ & \quad + \frac{4}{\theta(p)t} (1+(p-2)^2) \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-1} |\nabla \eta|^2. \end{aligned}$$

Now we choose t large enough that

$$(3-65) \quad \frac{1}{\theta(p)t} (B+(p-1)|\lambda|)^2 \leq \frac{A}{2}.$$

It follows from (3-64) and (3-65) that

$$\begin{aligned} & \frac{\theta(p)}{4} t \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 + \frac{A}{2} \int_{\Omega} u^{(p/2)\beta-2+\lambda} F^{t+\alpha+p/2} \eta^2 \\ & \leq 2(n-1)\kappa \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+\alpha+p/2-1} \eta^2 \\ & \quad + \frac{4}{\theta(p)t} (1+(p-2)^2) \int_{\Omega} u^{(p/2-1)\beta+\lambda} F^{t+p/2+\alpha-1} |\nabla \eta|^2. \end{aligned}$$

By letting $\lambda = \beta \left(1 - \frac{p}{2}\right)$, we obtain

$$(3-66) \quad \frac{\theta(p)t}{4} \int_{\Omega} F^{t+p/2+\alpha-3} |\nabla F|^2 \eta^2 + A/2 \int_{\Omega} u^{\beta-2} F^{t+\alpha+p/2} \eta^2 \\ \leq 2(n-1)\kappa \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 + \frac{4}{\theta(p)t} (1 + (p-2)^2) \int_{\Omega} F^{t+p/2+\alpha-1} |\nabla \eta|^2.$$

On the other hand, we have

$$(3-67) \quad |\nabla(F^{p/4+(\alpha-1+t)/2}\eta)|^2 \\ = \left|\frac{1}{4}(p+2t+2\alpha-2)F^{p/4+(t+\alpha-3)/2}\eta\nabla F + F^{p/4+(t-\alpha-1)/2}\nabla\eta\right|^2 \\ \leq \frac{1}{8}(p+2t+2\alpha-2)^2 F^{p/2+t+\alpha-3} |\nabla F|^2 \eta^2 + 2F^{p/2+t+\alpha-1} |\nabla \eta|^2.$$

Substituting (3-67) into (3-66) gives

$$(3-68) \quad \frac{2\theta(p)t}{(p+2t+2\alpha-2)^2} \int_{\Omega} |\nabla(F^{p/4+(\alpha-1+t)/2}\eta)|^2 + \frac{A}{2} \int_{\Omega} u^{\beta-2} F^{t+\alpha+p/2} \eta^2 \\ \leq 2(n-1)\kappa \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 \\ + \left\{ \frac{4}{\theta(p)t} (1 + (p-2)^2) + \frac{4\theta(p)t}{(p+2t+2\alpha-2)^2} \right\} \int_{\Omega} F^{t+p/2+\alpha-1} |\nabla \eta|^2.$$

At the same time, the Saloff-Coste Sobolev inequality implies

$$(3-69) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2}\eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ \leq \int_{\Omega} |\nabla(F^{p/4+(\alpha-1+t)/2}\eta)|^2 + R^{-2} \int_{\Omega} F^{p/2+\alpha+t-1} \eta^2.$$

Now, we substitute (3-69) into (3-68) to obtain

$$\frac{2\theta(p)t}{(p+2t+2\alpha-2)^2} \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2}\eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ \leq -\frac{A}{2} \int_{\Omega} u^{\beta-2} F^{t+\alpha+p/2} \eta^2 + \left[2(n-1)\kappa + \frac{2\theta(p)t}{(p+2t+2\alpha-2)^2 R^2}\right] \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 \\ + \left\{4/\theta(p)t(1 + (p-2)^2) + \frac{4\theta(p)t}{(p+2t+2\alpha-2)^2}\right\} \int_{\Omega} F^{t+\frac{p}{2}+\alpha-1} |\nabla \eta|^2.$$

We divide both sides of this inequality by $\frac{2\theta(p)t}{(p+2t+2\alpha-2)^2}$ to obtain

$$(3-70) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2}\eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ + \frac{A(p+2t+2\alpha-2)^2}{4\theta(p)t} \int_{\Omega} u^{\beta-2} F^{t+\alpha+p/2} \eta^2 \\ \leq \left[(n-1)\frac{(p+2t+2\alpha-2)^2}{\theta(p)t}\kappa + \frac{1}{R^2}\right] \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 \\ + \left\{\frac{2(p+2t+2\alpha-2)^2}{\theta^2(p)t^2}(1 + (p-2)^2) + 2\right\} \int_{\Omega} F^{t+p/2+\alpha-1} |\nabla \eta|^2.$$

Set

$$\mu_1 = \sup_{t \in [1, \infty)} \frac{2(p+2t+2\alpha-2)^2}{\theta^2(p)t^2} (1+(p-2)^2) + 2, \quad \mu = A \inf_{t \in [1, \infty)} \frac{(p+2t+2\alpha-2)^2}{4\theta(p)t^2}.$$

Then μ_1 and μ are both finite positive constants. Combining their definitions with (3-70) yields

$$(3-71) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2} \eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ + \mu t \int_{\Omega} u^{\beta-2} F^{t+\alpha+p/2} \eta^2 \\ \leq ((n-1)\mu_1 t \kappa + R^{-2}) \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 + \mu_1 \int_{\Omega} F^{t+p/2+\alpha-1} |\nabla \eta|^2.$$

Set

$$(3-72) \quad \phi_{\beta} = \begin{cases} \sup_{\Omega} u & \text{if } 0 < \beta < 2, \\ 1 & \text{if } \beta = 2, \\ \inf_{\Omega} u & \text{if } \beta > 2. \end{cases}$$

With this notation we obtain an inequality identical to (3-71), except that the summand on the second line is replaced by

$$(3-73) \quad \mu t \phi_{\beta}^{\beta-2} \int_{\Omega} F^{t+\alpha+p/2} \eta^2.$$

This is sufficient to prove the lemma. \square

3.3. L^{β_1} -bound of gradient in a geodesic ball with radius $3R/4$.

Lemma 3.8. *Let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Furthermore, suppose that a , q , p and β satisfy the conditions stated in Lemma 3.6. Let*

$$(3-74) \quad \beta_1 = \frac{n}{n-2} \cdot \frac{p+2t_0+2\alpha-2}{2}.$$

If u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$, then for t_0 large enough there exists $C = C(n, p, q, \beta) > 0$ such that

$$(3-75) \quad \|F\|_{L^{\beta_1}(B(x_0, 3R/4))} \leq CV^{1/\beta_1} (\kappa + R^{-2}) \phi^{2-\beta},$$

where V is the volume of the geodesic ball $B_R(x_0)$ and ϕ is defined in (3-72).

Proof. As observed at the end of the previous proof, we have

$$(3-76) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t_0)/2} \eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ + \mu t_0 \phi_{\beta}^{\beta-2} \int_{\Omega} F^{t_0+\alpha+p/2} \eta^2 \\ \leq ((n-1)\mu_1 t_0 \kappa + R^{-2}) \int_{\Omega} F^{t_0+\alpha+p/2-1} \eta^2 + \mu_1 \int_{\Omega} F^{t_0+p/2+\alpha-1} |\nabla \eta|^2,$$

where $t = t_0$ satisfies (3-65). Define

$$(3-77) \quad \Omega_1 := \left\{ x : F \geq \left(2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right) \phi^{2-\beta}, x \in \Omega \right\}.$$

Then we have

$$(3-78) \quad ((n-1)\mu_1 t_0 \kappa + R^{-2}) \int_{\Omega_1} F^{t_0+\alpha+p/2-1} \eta^2 \leq \frac{\mu}{2} t_0 \phi_\beta^{\beta-2} \int_{\Omega} F^{t_0+\alpha+p/2} \eta^2.$$

Set $\Omega_2 := \Omega \setminus \Omega_1 = \left\{ x : F < \left(2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right) \phi^{2-\beta}, x \in \Omega \right\}$. Then

$$(3-79) \quad \begin{aligned} & ((n-1)\mu_1 t_0 \kappa + R^{-2}) \int_{\Omega_2} F^{t_0+\alpha+p/2-1} \eta^2 \\ & \leq ((n-1)\mu_1 t_0 \kappa + R^{-2}) \left\{ \left(2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right) \phi^{2-\beta} \right\}^{t_0+\alpha+p/2-1} V, \end{aligned}$$

where V is the volume of $\Omega = B(x_0, R)$. Combining (3-78) and (3-79), we obtain

$$(3-80) \quad \begin{aligned} & ((n-1)\mu_1 t_0 \kappa + R^{-2}) \int_{\Omega} F^{t_0+\alpha+p/2-1} \eta^2 - \frac{\mu}{2} t_0 \phi_\beta^{\beta-2} \int_{\Omega} F^{t_0+\alpha+p/2} \eta^2 \\ & \leq ((n-1)\mu_1 t_0 \kappa + R^{-2}) \left[\left(2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right) \phi^{2-\beta} \right]^{t_0+\alpha+p/2-1} V. \end{aligned}$$

We set $\Omega_1 = B(x_0, 3R/4)$ and choose $\eta_1 \in C_0^\infty(\Omega)$ satisfying

$$0 \leq \eta_1 \leq 1, \quad \eta_1 \equiv 1 \quad \text{in } \Omega_1, \quad |\nabla \eta_1| \leq C/R,$$

and let

$$\eta = \eta_1^{t_0+p/2+\alpha}.$$

Then, we have

$$(3-81) \quad \begin{aligned} \mu_1 \int_{\Omega} F^{t_0+p/2+\alpha-1} |\nabla \eta|^2 &= \mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \int_{\Omega} F^{t_0+p/2+\alpha-1} \eta_1^{p+2\alpha-2+2t_0} |\nabla \eta_1|^2 \\ &\leq \mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \int_{\Omega} F^{t_0+p/2+\alpha-1} \eta_1^{p+2\alpha-2+2t_0}. \end{aligned}$$

By Hölder's inequality, (3-81) can be written as

$$(3-82) \quad \begin{aligned} & \mu_1 \int_{\Omega} F^{t_0+p/2+\alpha-1} |\nabla \eta|^2 \\ & \leq \mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \left(\int_{\Omega} F^{t_0+p/2+\alpha} \eta_1^{2t_0+p+2\alpha} \right)^{\frac{2t_0+p+2\alpha-2}{2t_0+p+2\alpha}} V^{\frac{2}{2t_0+p+2\alpha}} \\ & = \mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \left(\int_{\Omega} F^{t_0+p/2+\alpha} \eta^2 \right)^{\frac{2t_0+p+2\alpha-2}{2t_0+p+2\alpha}} V^{\frac{2}{2t_0+p+2\alpha}}. \end{aligned}$$

By using Young's inequality, we can write (3-82) as

$$(3-83) \quad \begin{aligned} & \mu_1 \int_{\Omega} F^{t_0+p/2+\alpha-1} |\nabla \eta|^2 \leq \frac{1}{2} \mu t_0 \phi^{\beta-2} \int_{\Omega} F^{t_0+p/2+\alpha} \eta^2 \\ & + \frac{2}{2t_0+p+2\alpha} \left[\mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \right]^{t_0+p/2+\alpha} \left[\frac{2(2t_0+p+2\alpha-2)}{(2t_0+p+2\alpha)\mu t_0} \phi^{2-\beta} \right]^{t_0+p/2+\alpha-1} V. \end{aligned}$$

Substituting (3-80) and (3-83) into (3-76) gives

$$(3-84) \quad \exp \left\{ -C_n(1 + \sqrt{\kappa}R) \right\} V^{2/n} R^{-2} \left\| F^{p/4+(\alpha-1+t_0)/2} \eta \right\|_{L^{2n/(n-2)}(\Omega)}^2 \\ \leq ((n-1)\mu_1 t_0 \kappa + R^{-2}) \left\{ \left[2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right] \phi^{2-\beta} \right\}^{t_0+\alpha+p/2-1} V \\ + \frac{2}{2t_0+p+2\alpha} \left[\mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \right]^{t_0+p/2+\alpha} \left[\frac{2(2t_0+p+2\alpha-2)}{(2t_0+p+2\alpha)\mu t_0} \phi^{2-\beta} \right]^{t_0+p/2+\alpha-1} V.$$

Taking $\frac{2}{2t_0+p+2\alpha-2}$ powers on both sides of (3-84), we obtain

$$\|F\|_{L^{\beta_1}(B(x_0, 3R/4))} \\ \leq \left\| F^{p/4+(\alpha-1+t_0)/2} \eta \right\|_{L^{2n/(n-2)}(\Omega)}^{4/2t_0+p+2\alpha-2} \\ \leq \exp \left\{ \frac{2C_n(1+\sqrt{\kappa}R)}{2t_0+p+2\alpha-2} \right\} V^{1/\beta_1} \\ \times \left\{ ((n-1)\mu_1 t_0 R^2 \kappa + 1) \left[2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right] \phi^{2-\beta} \right\}^{t_0+\alpha+p/2-1} \\ + \frac{2R^2}{2t_0+p+2\alpha} \left[\mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 \frac{C^2}{R^2} \right]^{t_0+p/2+\alpha} \\ \times \left[\frac{2(2t_0+p+2\alpha-2)}{(2t_0+p+2\alpha)\mu t_0} \phi^{2-\beta} \right]^{t_0+p/2+\alpha-1} \left\}^{\frac{2}{2t_0+p+2\alpha-2}},$$

where

$$\beta_1 = \frac{n}{n-2} \cdot \frac{p+2t_0+2\alpha-2}{2}.$$

Using the fact that $(a_1 + a_2)^{b_1} \leq 2^{b_1}(a_1^{b_1} + a_2^{b_1})$, valid for $a_i \geq 0$ and $b_1 > 0$, we infer from the preceding inequality that

$$(3-85) \quad \|F\|_{L^{\beta_1}(B(x_0, 3R/4))} \leq \exp \left\{ \frac{2C_n(1+\sqrt{\kappa}R)}{2t_0+p+2\alpha-2} \right\} V^{1/\beta_1} (I_1 + I_2),$$

where

$$I_2 := 2^{2/2t_0+p+2\alpha-2} \left(\frac{2}{2t_0+p+2\alpha} \right)^{\frac{2}{2t_0+p+2\alpha-2}} (\mu_1 \left(t_0 + \frac{p}{2} + \alpha \right)^2 C^2)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} \\ \frac{2(2t_0+p+2\alpha-2)}{(2t_0+p+2\alpha)\mu t_0} \phi^{2-\beta} R^{-2}$$

will be estimated later, and

$$(3-86) \quad I_1 := 2^{\frac{2}{2t_0+p+2\alpha-2}} ((n-1)\mu_1 t_0 R^2 \kappa + 1)^{\frac{2}{2t_0+p+2\alpha-2}} \left(2(n-1)\kappa \frac{\mu_1}{\mu} + \frac{2}{\mu t_0 R^2} \right) \phi^{2-\beta} \\ = 2^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} ((n-1)\mu_1 t_0 R^2 \kappa + 1)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} t_0^{-1} \mu^{-1} R^{-2} \phi^{2-\beta} \\ \leq 2^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} ((n-1)\mu_1 t_0 + 1)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} (1+R^2 \kappa)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} t_0^{-1} \mu^{-1} R^{-2} \phi^{2-\beta} \\ = 2^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} ((n-1)\mu_1 + t_0^{-1})^{\frac{2}{2t_0+p+2\alpha-2}} (1+R^2 \kappa)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} \mu^{-1} R^{-2} \phi^{2-\beta}.$$

Noticing that

$$\lim_{t_0 \rightarrow +\infty} 2^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} \left[(n-1)\mu_1 + t_0^{-1} \right]^{\frac{2}{2t_0+p+2\alpha-2}} = 2,$$

we can verify that

$$(3-87) \quad \sup_{t_0 \in [1, +\infty)} 2^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} \left[(n-1)\mu_1 + t_0^{-1} \right]^{\frac{2}{2t_0+p+2\alpha-2}} < +\infty.$$

Combining (3-86) and (3-87) leads to

$$(3-88) \quad I_1 \leq C_{I_1} (1 + R^2 \kappa)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} \mu^{-1} R^{-2} \phi^{2-\beta},$$

where C_{I_1} is a positive constant which depends only on n , p and q .

Similarly, we have

$$(3-89) \quad I_2 \leq C_{I_2} \mu^{-1} t_0 R^{-2} \phi^{2-\beta},$$

where C_{I_2} is a positive constant which depends only on n , p and q .

Substituting (3-88) and (3-89) into (3-85), we obtain

$$\begin{aligned} \|F\|_{L^{\beta_1}(B(x_0, 3R/4))} &\leq \exp \left\{ \frac{2C_n(1 + \sqrt{\kappa}R)}{2t_0 + p + 2\alpha - 2} \right\} \\ &\quad \times V^{1/\beta_1} \left[C_{I_1} (1 + \kappa R^2)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} + C_{I_2} t_0 \right] R^{-2} \mu^{-1} \phi^{2-\beta}. \end{aligned}$$

We can derive from this inequality that

$$(3-90) \quad \|F\|_{L^{\beta_1}(B(x_0, 3R/4))} \leq C_1 \exp \left\{ \frac{2C_n(1 + \sqrt{\kappa}R)}{2t_0 + p + 2\alpha - 2} \right\} V^{1/\beta_1} \left[(1 + \kappa R^2)^{\frac{2t_0+p+2\alpha}{2t_0+p+2\alpha-2}} + t_0 \right] R^{-2} \phi^{2-\beta},$$

where $C_1 := \max\{C_{I_1}, C_{I_2}\} \mu^{-1}$ is a positive constant depending only on n , β , p , q .

Now, let t_0 satisfy (3-65) and

$$(3-91) \quad (1 + \kappa R^2) \leq t_0 \leq C_0(1 + \kappa R^2),$$

where $C_0 = C_0(n, p, q, \beta)$ is a positive constant. Combining (3-90) and (3-91) gives

$$\|F\|_{L^{\beta_1}(B(x_0, 3R/4))} \leq C V^{1/\beta_1} (\kappa + R^{-2}) \phi^{2-\beta},$$

where C is a positive constant that depends only on n , β , p , q . This completes the proof of Lemma 3.8. \square

3.4. Moser iteration for positive solutions of (1-1).

Lemma 3.9. *Let (M, g) be an n -dimensional ($n \geq 3$) complete manifold with $\text{Ric} \geq -(n-1)\kappa$, where κ is a nonnegative constant. Suppose that a , q , p and β satisfy the same conditions as in Lemma 3.6. Let*

$$\beta_1 = \frac{n}{n-2} \cdot \frac{p+2t_0+2\alpha-2}{2}.$$

If u is a positive solution to equation (1-1) on the geodesic ball $B(x_0, 2R) \subset M$, then for t_0 large enough there exists $C = C(n, p, q) > 0$ such that

$$\|F\|_{L^\infty(B(x_0, R/2))} \leq CV^{-1/\beta_1} \|F\|_{L^{\beta_1}(\Omega_1)},$$

where V is the volume of the geodesic ball $B_R(x_0)$.

Proof. Recall the integral inequality (3-71) from the proof of Lemma 3.8 (page 417). By dropping the second nonnegative term in that inequality, we obtain

$$(3-92) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2} \eta\|_{L^{2n/(n-2)}(\Omega)}^2 \\ \leq ((n-1)\mu_1 t \kappa + R^{-2}) \int_{\Omega} F^{t+\alpha+p/2-1} \eta^2 + \mu_1 \int_{\Omega} F^{t+p/2+\alpha-1} |\nabla \eta|^2,$$

Then we set

$$r_m := \frac{R}{2} + \frac{R}{4^m} \quad \text{and} \quad \Omega_m := B(x_0, r_m),$$

and then choose $\eta_m \in C_0^\infty(\Omega_m)$ satisfying

$$0 \leq \eta_m \leq 1, \quad \eta_m \equiv 1 \quad \text{in} \quad B(x_0, r_{m+1}), \quad |\nabla \eta_m| \leq C \frac{4^m}{R}.$$

Replacing η by η_m in (3-92), we can easily verify that

$$(3-93) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} R^{-2} \|F^{p/4+(\alpha-1+t)/2}\|_{L^{2n/(n-2)}(\Omega_{m+1})}^2 \\ \leq ((n-1)\mu_1 t \kappa + R^{-2}) \int_{\Omega_m} F^{t+p/2+\alpha-1} + \mu_1 \frac{C^2 16^m}{R^2} \int_{\Omega_m} F^{t+p/2+\alpha-1}.$$

Next, we choose

$$\beta_1 = (t_0 + \frac{p}{2} + \alpha - 1) \frac{n}{n-2} \quad \text{and} \quad \beta_{m+1} = \frac{n\beta_m}{n-2},$$

and let $t = t_m$ such that

$$t_m + \frac{p}{2} + \alpha - 1 = \beta_m.$$

Then

$$(3-94) \quad \exp\{-C_n(1 + \sqrt{\kappa}R)\} V^{2/n} \left(\int_{\Omega_{m+1}} F^{\beta_{m+1}} \right)^{(n-2)/n} \\ \leq ((n-1)\mu_1 t_m R^2 \kappa + 1 + \mu_1 C^2 16^m) \int_{\Omega_m} F^{\beta_m}.$$

Taking $1/\beta_m$ powers on both sides of (3-94), we obtain

$$(3-95) \quad \|F\|_{L^{\beta_{m+1}}(\Omega_{m+1})} \\ \leq \exp\left\{ \frac{C_n(1 + \sqrt{\kappa}R)}{\beta_m} \right\} V^{-2/(n\beta_m)} ((n-1)\mu_1 t_m \kappa R^2 + 1 + \mu_1 C^2 16^m)^{1/\beta_m} \|F\|_{L^{\beta_m}(\Omega_m)}.$$

Keeping the definition of t_m in mind, from (3-95) we deduce that

$$\begin{aligned} \|F\|_{L^{\beta_{m+1}}(\Omega_{m+1})} &\leq \exp\left\{\frac{C_n(1+\sqrt{\kappa}R)}{\beta_m}\right\} V^{-2/n\beta_m} 16^{m/\beta_m} \\ &\quad \times \left((n-1)\mu_1\left(t_0 + \frac{p}{2} + \alpha - 1\right)\kappa R^2 + 1 + \mu_1 C^2\right)^{1/\beta_m} \|F\|_{L^{\beta_m}(\Omega_m)}. \end{aligned}$$

Noting that

$$\sum_{m=1}^{\infty} \frac{1}{\beta_m} = \frac{n}{2\beta_1} \quad \text{and} \quad \sum_{m=1}^{\infty} \frac{m}{\beta_m} = \frac{n^2}{4\beta_1},$$

we have

$$\begin{aligned} \|F\|_{L^\infty(B(x_0, R/2))} &\leq \exp\left\{\frac{nC_n(1+\sqrt{\kappa}R)}{2\beta_1}\right\} V^{-1/\beta_1} 16^{n^2/(4\beta_1)} \\ &\quad \times \left[(n-1)\mu_1\left(t_0 + \frac{p}{2} + \alpha - 1\right)\kappa R^2 + 1 + \mu_1 C^2\right]^{n/(2\beta_1)} \|F\|_{L^{\beta_1}(\Omega_1)} \\ &\leq C_3 \exp\left\{\frac{nC_n(1+\sqrt{\kappa}R)}{2t_0+p+2\alpha-2}\right\} V^{-1/\beta_1} (1 + \kappa R^2)^{n/(2\beta_1)} \|F\|_{L^{\beta_1}(\Omega_1)}, \end{aligned}$$

where $C_3 = C_3(n, p, q)$ is a positive constant. In view of (3-65) ($t = t_0$) and (3-91), it is not difficult to see that

$$\|F\|_{L^\infty(B(x_0, R/2))} \leq C V^{-\frac{1}{\beta_1}} \|F\|_{L^{\beta_1}(\Omega_1)},$$

where $C = C(n, p, q)$ is a positive constant. □

4. Proof of the main theorem and its consequences

Theorem 1.1 follows easily from Lemmas 3.8 and 3.9. Therefore, we omit its proof, but give those of Corollary 1.2, Theorem 1.5 and Corollary 1.7. We omit those of Corollary 1.6 (very easy) and those of Corollaries 1.3 and 1.4 (similar to that of Corollary 1.2).

Proof of Corollary 1.2. By using Theorem 1.1, we only need to confirm that the constants a, q and p satisfy either

$$\frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) - h^{1/2} < q < \frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) + h^{1/2} \quad (a \neq 0)$$

or

$$(4-1) \quad a \left[\frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1) - q \right] \geq 0.$$

Here, we only check case 1; the others are similar.

Case 1: $a > 0, p \geq n$ and $q \in \mathbb{R}$. Since $p \geq n$, we can see that $\beta \in (0, +\infty)$ by using (1-3). Furthermore, since $a > 0$, we can verify that (4-1) is equivalent to

$$(4-2) \quad q \leq \frac{\beta}{2} \cdot \frac{n+1}{n-1} (p-1).$$

Hence, for any fixed p ($p \geq n$), n and $q \in \mathbb{R}$, we can make (4-2) be true by letting β large enough. Therefore, we complete the proof of case 1. \square

Proof of Theorem 1.5. By using Corollary 1.2, we know that there exist positive constants $C = C(n, p, q)$ and $\beta = \beta(n, p, q) \in (0, \infty)$, such that

$$(4-3) \quad \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1+\kappa R^2}{R^2} \phi_\beta^{2-\beta}.$$

Since

$$(4-4) \quad \phi_\beta = \begin{cases} \sup_{B(x_0, R)} u & \text{if } 0 < \beta < 2, \\ 1 & \text{if } \beta = 2, \\ \inf_{B(x_0, R)} u & \text{if } \beta > 2, \end{cases}$$

we consider three cases:

Case 1: $\beta \in (0, 2)$. Combining (4-3) and (4-4), we have the estimate

$$(4-5) \quad \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \leq C \frac{1+\kappa R^2}{R^2} \sup_{B(x_0, R)} u^{2-\beta}.$$

Multiplying both sides of (4-5) by $\sup_{B(x_0, R)} u^{\beta-2}$ leads to

$$(4-6) \quad \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \sup_{B(x_0, R)} u^{\beta-2} \leq C \frac{1+\kappa R^2}{R^2} \sup_{B(x_0, R)} u^{2-\beta} \sup_{B(x_0, R)} u^{\beta-2}.$$

Since $\sup_{B(x_0, R)} u \leq l \inf_{B(x_0, R)} u$, we see that

$$(4-7) \quad \sup_{B(x_0, R)} u^{2-\beta} \sup_{B(x_0, R)} u^{\beta-2} = \left(\sup_{B(x_0, R)} u \right)^{2-\beta} \left(\inf_{B(x_0, R)} u \right)^{\beta-2} \leq l^{2-\beta}.$$

Furthermore,

$$(4-8) \quad \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \sup_{B(x_0, R/2)} u^{\beta-2} \leq \sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^\beta} \sup_{B(x_0, R)} u^{\beta-2}.$$

Now, substituting (4-7) and (4-8) into (4-6) leads to

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq C l^{2-\beta} \frac{1+\kappa R^2}{R^2}.$$

Thus we finish the proof of case 1.

Case 2: $\beta = 2$. Combining (4-3) and (4-4), we have the estimate

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq C \frac{1+\kappa R^2}{R^2}.$$

Therefore, we complete the proof of case 2.

Case 3: $\beta > 2$. The proof of case 3 is similar to case 1. Hence, we omit it. \square

Now, we turn to (1-1) in \mathbb{R}^n and give the proof of Corollary 1.7, which uses Lemmas 2.4 and 2.5.

Proof of Corollary 1.7. In order to use Lemmas 2.4 and 2.5, we need to consider separately the four cases from the statements. We give the proofs in cases 1 and 3 only; in cases 2 and 4 they are similarly to those of 1 and 3 respectively.

Case 1: $a > 0, 1 < p < n, p \neq q$ and $q \in (p-1, (p-1)n/(n-p))$. Define $u = a^{1/(p-q-1)}w$. Then w satisfies $\Delta_p w + w^q = 0$, thanks to (1-1). By using Lemma 2.4, we know that for any $1 < p < n$ and $q \in (p-1, \frac{(p-1)n}{n-p})$, the Harnack inequality

$$(4-9) \quad \sup_{B_R} w(x) \leq C \inf_{B_R} w(x)$$

holds true, where $C = C(n, p, q)$ is a positive constant. Combining Theorem 1.5 and (4-9), we deduce that for any $1 < p < n$ and $q \in (p-1, \frac{(p-1)n}{n-p})$, the estimate

$$(4-10) \quad \sup_{B(x_0, R/2)} \frac{|\nabla w|^2}{w^2} \leq \frac{C}{R^2}$$

holds true, where $C = C(n, p, q)$ is a positive constant. Substituting $w = a^{1/(q-p+1)}u$ into (4-10), we obtain

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2},$$

where $C = C(n, p, q)$ is a positive constant. This finishes the proof of case 1.

Case 3: $a \geq 1$ and $1 < p = q < n < p^2$. Since $p = q$ and $1 < p < n < p^2$, we know that $q \in (p-1, \frac{(p-1)n}{n-p})$. By using Lemma 2.4, we know that for any $1 < p = q < n < p^2$ the Harnack inequality

$$(4-11) \quad \sup_{B_R} u(x) \leq C \inf_{B_R} u(x)$$

holds true, where $C = C(n, p, q, a)$ is a positive constant. By using Theorem 1.5 and (4-11), we achieve the estimate

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2},$$

where $C = C(n, p, q, a)$ is a positive constant. This finishes the proof of case 3. \square

Proof of Theorem 1.8. By using Corollary 1.7, we know that if the constants a, q and p satisfy one of the four conditions listed in the theorem, then the estimate

$$\sup_{B(x_0, R/2)} \frac{|\nabla u|^2}{u^2} \leq \frac{C}{R^2}$$

holds true, where $C = C(n, p, q, a)$ is a positive constant. By letting $R \rightarrow +\infty$, we conclude that $\sup_{\mathbb{R}^n} |\nabla u| = 0$. Hence, u is a constant. But, no positive constant is a positive solution to (1-1). \square

Acknowledgements

The authors are supported by National Natural Science Foundation of China (grant no. 12431003).

References

- [1] M.-F. Bidaut-Véron, M. Garcia-Huidobro, and L. Véron, “A priori estimates for elliptic equations with reaction terms involving the function and its gradient”, *Math. Ann.* **378**:1-2 (2020), 13–56. MR
- [2] M.-F. Bidaut-Véron, M. Garcia-Huidobro, and L. Véron, “Radial solutions of scaling invariant nonlinear elliptic equations with mixed reaction terms”, *Discrete Contin. Dyn. Syst.* **40**:2 (2020), 933–982. MR
- [3] S. Y. Cheng and S. T. Yau, “Differential equations on Riemannian manifolds and their geometric applications”, *Comm. Pure Appl. Math.* **28**:3 (1975), 333–354. MR
- [4] T. H. Colding and W. P. Minicozzi, II, “Harmonic functions on manifolds”, *Ann. of Math. (2)* **146**:3 (1997), 725–747. MR
- [5] T. H. Colding and W. P. Minicozzi, II, “Liouville properties”, *ICCM Not.* **7**:1 (2019), 16–26. MR
- [6] E. DiBenedetto, “ $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations”, *Nonlinear Anal.* **7**:8 (1983), 827–850. MR
- [7] B. Gidas and J. Spruck, “Global and local behavior of positive solutions of nonlinear elliptic equations”, *Comm. Pure Appl. Math.* **34**:4 (1981), 525–598. MR
- [8] D. Han, J. He, and Y. Wang, “Gradient estimates for $\Delta_p u - |\nabla u|^q + b(x)|u|^{r-1}u = 0$ on a complete Riemannian manifold and Liouville type theorems”, preprint, 2023. arXiv 2309.03510
- [9] J. He, Y. Wang, and G. Wei, “Gradient estimate for solutions of the equation $\Delta_p v + av^q = 0$ on a complete Riemannian manifold”, *Math. Z.* **306**:3 (2024), art. id. 42, 19 pp. MR
- [10] P. Huang and Y. Wang, “Gradient estimates and Liouville theorems for Lichnerowicz equations”, *Pacific J. Math.* **317**:2 (2022), 363–386. MR
- [11] B. Kotschwar and L. Ni, “Local gradient estimates of p -harmonic functions, $1/H$ -flow, and an entropy formula”, *Ann. Sci. Éc. Norm. Supér. (4)* **42**:1 (2009), 1–36. MR
- [12] P. Li and S.-T. Yau, “On the parabolic kernel of the Schrödinger operator”, *Acta Math.* **156**:3-4 (1986), 153–201. MR
- [13] B. Peng, Y. Wang, and G. Wei, “Gradient estimates for $\Delta u + au^{p+1} = 0$ and Liouville theorems”, *Math. Theor. Appl.* **43** (2023), 32–43.
- [14] L. Saloff-Coste, “Uniformly elliptic operators on Riemannian manifolds”, *J. Differential Geom.* **36**:2 (1992), 417–450. MR
- [15] J. Serrin and H. Zou, “Cauchy–Liouville and universal boundedness theorems for quasilinear elliptic equations and inequalities”, *Acta Math.* **189**:1 (2002), 79–142. MR
- [16] P. Tolksdorf, “Regularity for a more general class of quasilinear elliptic equations”, *J. Differential Equations* **51**:1 (1984), 126–150. MR
- [17] K. Uhlenbeck, “Regularity for a class of non-linear elliptic systems”, *Acta Math.* **138**:3-4 (1977), 219–240. MR
- [18] J. Wang and Y. Wang, “Gradient estimates for $\Delta u + a(x)u \log u + b(x)u = 0$ and its parabolic counterpart under integral Ricci curvature bounds”, *Comm. Anal. Geom.* **32**:4 (2024), 923–975. MR

- [19] Y. Wang and G. Wei, “On the nonexistence of positive solution to $\Delta u + au^{p+1} = 0$ on Riemannian manifolds”, *J. Differential Equations* **362** (2023), 74–87. MR
- [20] X. Wang and L. Zhang, “Local gradient estimate for p -harmonic functions on Riemannian manifolds”, *Comm. Anal. Geom.* **19**:4 (2011), 759–771. MR
- [21] Y. Wang and L. Zhang, “Gradient estimates for $\Delta_p u + \Delta_q u + a|u|^{s-1}u + b|u|^{l-1}u = 0$ on a complete Riemannian manifold and applications”, *Internat. J. Math.* **36**:2 (2025), art. id. 2450070, 53 pp. MR
- [22] Y. Wang, J. Yang, and L. Zhang, “Liouville theorems and gradient estimates for positive solutions to $\Delta_p u + \Delta_q u + h(u) = 0$ on a complete manifold”, *Contemp. Math. [Internet]* **6**:3 (2025), 3135–83.
- [23] S. T. Yau, “Harmonic functions on complete Riemannian manifolds”, *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR

Received April 30, 2025. Revised January 24, 2026.

YOUDE WANG
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCES
GUANGZHOU UNIVERSITY
GUANGZHOU
CHINA

and

HUA LOO-KENG KEY LABORATORY MATHEMATICS
INSTITUTE OF MATHEMATICS
ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
BEIJING
CHINA

and

SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF CHINESE ACADEMY OF SCIENCES
BEIJING
CHINA

wyd@math.ac.cn

LIQIN ZHANG
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCES
GUANGZHOU UNIVERSITY
GUANGZHOU
CHINA

1061837643@qq.com

p -NUCLEARITY OF REDUCED GROUP L^p -OPERATOR ALGEBRAS

ZHEN WANG

Let $p \in (1, \infty)$. G. An, J.-J. Lee, and Z.-J. Ruan introduced p -nuclearity for L^p -operator algebras. They proved that the reduced group L^p -operator algebra $F_\lambda^p(G)$, where G is a discrete group, is p -nuclear and the p -pseudo-measure algebra $PM_p(G)$ is p -semidiscrete if G is amenable. In this paper, we show that the following are equivalent: (i) G is amenable; (ii) the reduced group L^p -operator algebra $F_\lambda^p(G)$ is p -nuclear; (iii) the p -pseudomeasure algebra $PM_p(G)$ is p -semidiscrete. This solves an open problem raised by N. C. Phillips concerning the p -nuclearity for reduced group L^p -operator algebras.

1. Introduction

For $p \in [1, \infty)$, we say that a Banach algebra A is an L^p -operator algebra if it is isometrically isomorphic to a norm-closed subalgebra of the algebra $\mathcal{B}(E)$ of all bounded linear operators on some L^p -space E . Clearly, L^p -operator algebras are a natural generalization of operator algebras on Hilbert spaces (and in particular C^* -algebras) by replacing Hilbert spaces with L^p -spaces.

The study of L^p -operator algebras traces back to C. Herz's influential work on harmonic analysis of group algebras in the 1970's [22; 23; 24]. For a locally compact group G , Herz introduced the Banach algebra $PF_p(G)$, defined as the operator norm closure of the image of the left regular representation $\lambda_p : L^1(G) \rightarrow \mathcal{B}(L^p(G))$. The Banach algebra $PF_p(G)$ is called p -pseudofunctions of G by C. Herz. This algebra is also called the reduced group L^p -operator algebra of G and it is denoted by $F_\lambda^p(G)$ in [18]. If $p = 2$, then $F_\lambda^2(G)$ is the reduced group C^* -algebra of G , which is usually denoted by $C_\lambda^*(G)$. We adopt the notation $F_\lambda^p(G)$ throughout the following of this paper.

Associated with $F_\lambda^p(G)$ there are two other natural algebras, the p -pseudo-measure algebra $PM_p(G)$ and the algebra of p -convolvers $CV_p(G)$. The p -pseudo-measure algebra $PM_p(G)$ is the weak* closure of $F_\lambda^p(G)$ in $\mathcal{B}(L^p(G))$. Let

$$\rho_p : G \rightarrow \mathcal{B}(l^p(G))$$

MSC2020: primary 47L10, 47L25; secondary 43A07, 43A15.

Keywords: reduced group L^p -operator algebra, p -pseudomeasure algebra, nuclearity, amenability.

be the right regular representation. The algebra of p -convolvers $CV_p(G)$ is the commutant of $\rho_p(G)$. If $p = 2$, then $PM_2(G)$ and $CV_2(G)$ are the group von Neumann algebra, which is denoted by $L(G)$. The reader is referred to [11; 12] for more research concerning them, especially on the problem of determining whether $PM_p(G) = CV_p(G)$.

Recently, interest in L^p -operator algebras has been renewed due to the work of N. C. Phillips. In the last ten years, Phillips introduced and studied L^p -operator algebras [31; 32; 33; 34; 35; 36; 37]. These studies encourage many authors to participate in the research of L^p -operator algebras. This includes the work on group L^p -operator algebras [18]; groupoid L^p -operator algebras [16]; L^p -operator crossed products [19; 42; 43] and the L^p -Toeplitz algebra [41]. Although most previous investigations have been very largely focused on various examples, some recent works were undertaken in a more abstract and systematic way [3; 6; 17]. Surprisingly, when $p \in [1, \infty) \setminus \{2\}$, the research on L^p -operator algebra has many wonderful results for rigidity problems [6; 7; 20]. The reader is referred to [15] for more historical comments and recent developments in L^p -operator algebras.

Nuclearity is an important property for C^* -algebras. This property was introduced by Takesaki [40]. A C^* -algebra A is called *nuclear* if for any C^* -algebra B there is a unique norm on the algebraic tensor product $A \otimes B$. By the remarkable work of Lance [28], Choi and Effros [5] and Kirchberg [26], the nuclearity is equivalent to *completely positive approximation property*, that is, there exist nets of contractive completely positive maps $\varphi_\alpha : A \rightarrow M_{n(\alpha)}$ and $\psi_\alpha : M_{n(\alpha)} \rightarrow A$ such that

$$\|\psi_\alpha \circ \varphi_\alpha(a) - a\| \rightarrow 0$$

for all $a \in A$. Also it is well known that the nuclearity is equivalent to the amenability for C^* -algebras [8; 21], which was originally introduced by B. E. Johnson [25] for Banach algebras.

The semidiscreteness of von Neumann algebras is close related to nuclearity of C^* -algebras. A von Neumann algebra M is *semidiscrete* if there exist nets of weak* continuous contractive completely positive maps $\varphi_\alpha : M \rightarrow M_{n(\alpha)}$ and $\psi_\alpha : M_{n(\alpha)} \rightarrow M$ such that

$$\langle \psi_\alpha \circ \varphi_\alpha(a) - a, f \rangle \rightarrow 0$$

for all $a \in M$ and $f \in M_*$, where M_* is the predual of M . The following theorem is a classical result concerning the nuclear reduced group C^* -algebras and semidiscrete group von Neumann algebras.

Theorem 1.1 [4, Theorem 2.6.8]. *Let G be a discrete group. The following statements are equivalent:*

- (i) G is amenable.

- (ii) *The reduced group C^* -algebra $C_\lambda^*(G)$ is nuclear.*
- (iii) *The group von Neumann algebra $L(G)$ is semidiscrete.*

In [1, Proposition 5.1(a)], G. An, J.-J. Lee and Z.-J. Ruan studied p -nuclearity of reduced group L^p -operator algebra $F_\lambda^p(G)$ and p -semidiscreteness of p -pseudo-measure algebra $PM_p(G)$. They proved the following proposition.

Proposition 1.2 [1, Proposition 5.1]. *Let $p > 1$ and let G be a discrete amenable group.*

- (i) *The reduced group L^p -operator algebra $F_\lambda^p(G)$ is p -nuclear.*
- (ii) *The p -pseudomeasure algebra $PM_p(G)$ is p -semidiscrete.*

Since $F_\lambda^1(G)$ is always 1-nuclear for all discrete group G (see [1, Theorem 6.4]), we only consider the following problem for $p \in (1, \infty)$.

Problem 1.3 [35, Problem 10.4]. *Let $p \in (1, \infty)$. If G is a discrete group and $F_\lambda^p(G)$ is p -nuclear, does it follow that G is amenable?*

In this paper, we solve N. C. Phillips’ problem by proving the following theorem.

Theorem 1.4. *Let $p \in (1, \infty)$ and let G be a discrete group. The following statements are equivalent:*

- (i) *G is amenable.*
- (ii) *$F_\lambda^p(G)$ is p -nuclear.*
- (iii) *$PM_p(G)$ is p -semidiscrete.*
- (iv) *There exists an isomorphism $\Phi : F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G) \rightarrow F_\lambda^p(G) \overset{\wedge}{\otimes} F_\lambda^p(G)$, where $\overset{\vee}{\otimes}$ and $\overset{\wedge}{\otimes}$ are the p -operator space injective and projective tensor products, respectively.*
- (v) *The canonical linear map $h : F_\lambda^p(G) \otimes F_\lambda^p(G) \rightarrow \mathcal{B}(l^p(G))$ given by*

$$h(\lambda_p(s) \otimes \lambda_p(t)) = \lambda_p(s)\rho_p(t)$$

is continuous with respect to the p -operator space injective tensor norm, where \otimes is the algebraic tensor product.

- (vi) *For any $f \in C_c(G)$, we have $\|\lambda_p(f)\| \geq \left| \sum_{t \in G} f(t) \right|$.*
- (vii) *For any finite subset $E \subset G$, we have $|E| = \left\| \sum_{t \in E} \lambda_p(t) \right\|$.*

Remark 1.5. Condition (vi) implies that the trivial representation 1_G extends to a representation of $F_\lambda^p(G)$. This is also equivalent to the amenability of G (see [13, Theorem 5.2]).

In the reduced group C^* -algebra case, one way to prove (ii) \Rightarrow (i) of Theorem 1.1 is based on the Arveson's extension theorem (see [4, Theorem 2.6.8]). In the proof of (iii) \Rightarrow (i) of Theorem 1.1, Arveson's extension theorem is also applied to construct unital completely positive map from $\mathcal{B}(l^2(G))$ to $L(G)$ that restricts to the identity on $L(G)$ (i.e., semidiscreteness implies injectivity). Arveson's extension theorem states that every contractive completely positive map $\varphi : B \rightarrow \mathcal{B}(H)$ can be extended to a contractive completely positive map $\bar{\varphi} : A \rightarrow \mathcal{B}(H)$, where A is a C^* -algebra, B is an operator subsystem of A , and H is a complex Hilbert space. However, J.-J. Lee gave an example that the Arveson–Wittstock–Hahn–Banach theorem does not hold for p -operator space [30], that is, there are p -operator spaces $V \subset W$, an SQ_p -space E , and a p -completely contractive map $\varphi : V \rightarrow \mathcal{B}(E)$ such that φ does not extend to a p -completely contractive map on W . The lack of a valid Arveson–Wittstock–Hahn–Banach theorem for p -operator spaces forces us to adopt alternative approaches. Our proof of Theorem 1.4 is inspired by the method of C. Anantharaman-Delaroche (see [2, Proposition 3.5]). Her proof is based on the functorial property [2, Proposition 2.6] of spatial and maximal tensor products and weak containment of unitary representations [2, Proposition 3.5]. Our proof relies on

- the functorial properties of p -operator spaces projective and injective tensor products (see Lemma 2.4, 2.5 and the proof of (ii) \Rightarrow (iv), (iii) \Rightarrow (v) in Theorem 1.4);
- the uniform convexity of $l^p(G)$ that is motivated by G. Pisier (see [38, Theorem 3.30] and the proof of (vii) \Rightarrow (i)).

The paper is organized as follows. In Section 2, we make some preparations for the proof of Theorem 1.4. In Section 3, we give a proof of Theorem 1.4.

2. Preliminaries

In this section, we recall some notation, definitions and lemmas for the proof of Theorem 1.4.

2.1. Reduced group L^p -operator algebras and p -pseudomeasure algebras. Let $p \in (1, \infty)$. For a discrete group G , we let $\lambda_p : G \rightarrow \mathcal{B}(l^p(G))$ denote the *left regular representation*, that is $\lambda_p(s)(\delta_t) = \delta_{st}$ for all $s, t \in G$, where $\{\delta_t\}_{t \in G}$ is the canonical basis of $l^p(G)$.

Definition 2.1. The reduced group L^p -operator algebra of G , denoted $F_\lambda^p(G)$, is the completion of $C_c(G)$ with respect to the norm $\|\lambda_p(f)\|$.

There are many equivalent definitions for amenable groups. We will use the following definition in the proof of the Theorem 1.4.

Definition 2.2 [9, Definition 11.2.3]. Let $p \in [1, \infty)$ and let G be a discrete group. The group G is amenable if there exists a net $f_\alpha \in l^p(G)$ such that $f_\alpha \geq 0$, $\|f_\alpha\|_p = 1$ and $\|\lambda_p(s)f_\alpha - f_\alpha\|_p \rightarrow 0$ for all $s \in G$.

For $p \in (1, \infty)$, and we denote by p' its conjugate exponent, which satisfies $\frac{1}{p} + \frac{1}{p'} = 1$. Let $\mathcal{N}(l^p(G)) = l^{p'}(G) \widehat{\otimes} l^p(G)$ denote the space of nuclear operators on $l^p(G)$, where $\widehat{\otimes}$ is the projective tensor product. Then $\mathcal{B}(l^p(G))$ is the dual space of $\mathcal{N}(l^p(G))$ by way of dual pairing $\langle T, \xi \otimes \eta \rangle = \langle \xi, T\eta \rangle$ where $\xi \in l^{p'}(G)$ and $\eta \in l^p(G)$. We say that a net (T_α) in $\mathcal{B}(l^p(G))$ converges weak* to an operator T in $\mathcal{B}(l^p(G))$ if $\langle T_\alpha, f \rangle \rightarrow \langle T, f \rangle$ for all $f \in l^{p'}(G) \widehat{\otimes} l^p(G)$.

Definition 2.3. The p -pseudomeasure algebra of G , denoted $PM_p(G)$, is the weak* closure of $F_\lambda^p(G)$ in $\mathcal{B}(l^p(G))$.

When $p = 2$, we have $PM_2(G) = L(G)$, where $L(G)$ is the group von Neumann algebra of G .

The p -pseudomeasure algebra has a predual $A_p(G)$, that is, $PM_p(G) = A_p(G)'$ [1]. The algebra $A_p(G)$ is called the Figà-Talamanca–Herz algebra and will be introduced next. Let $\Lambda_p : l^{p'}(G) \widehat{\otimes} l^p(G) \rightarrow C_0(G)$ be given by

$$\Lambda_p(\xi \otimes \eta)(s) = \langle \xi, \lambda_p(s)\eta \rangle$$

for all $s \in G, \eta \in l^p(G), \xi \in l^{p'}(G)$. Since $C_c(G)$ is dense in $l^p(G)$ and $l^{p'}(G)$, it follows that Λ_p maps into $C_0(G)$. Then $A_p(G)$ is defined to be the *coimage* of Λ_p , i.e., the space of $f \in C_0(G)$ for which there are $(\xi_n) \subset l^{p'}(G)$ and $(\eta_n) \subset l^p(G)$ such that

$$f(s) = \sum_{n=1}^{\infty} \xi_n * \check{\eta}_n(s) = \sum_{n=1}^{\infty} \langle \xi_n, \lambda_p(s)\eta_n \rangle$$

with norm

$$\|f\|_{A_p(G)} = \inf \left\{ \sum_{n=1}^{\infty} \|\xi_n\| \|\eta_n\| : f = \sum_{n=1}^{\infty} \xi_n * \check{\eta}_n \right\} < \infty,$$

where $\check{\eta}_n(s) = \eta_n(s^{-1})$ and $\xi_n * \check{\eta}_n(s) = \sum_{t \in G} \xi(t)\check{\eta}_n(t^{-1}s)$. It follows from [22] that $A_p(G)$ is a commutative Banach algebra with pointwise multiplication.

It follows from the definition that $A_p(G)$ can be identified with the quotient of nuclear space $\mathcal{N}(l^p(G))$. In fact, we have $A_p(G) = \mathcal{N}(l^p(G))/PM_p(G)_\perp$, where $PM_p(G)_\perp = \ker \Lambda_p$ and $PM_p(G)_\perp$ is called the pre-annihilator of $PM_p(G)$ in $\mathcal{N}(l^p(G))$. Therefore we have the isometric isomorphism $PM_p(G) = A_p(G)'$.

2.2. p -operator spaces. The notion of p -operator spaces is closely related to that of L^p -operator algebras. Let $p \in (1, \infty)$. For each positive integer n , let $l_n^p = L^p(\{1, 2, \dots, n\}, \nu)$, where ν is the counting measure on $\{1, 2, \dots, n\}$. We denote $M_n^p = \mathcal{B}(l_n^p)$. Let m be a positive integer, and we denote $M_{n,m}^p = \mathcal{B}(l_m^p, l_n^p)$. A p -operator space is defined to be a Banach space together with a matrix norm,

i.e., a norm $\|\cdot\|_n$ on each matrix space $M_n(V)$, which satisfies the following two conditions:

- (i) $\mathcal{D}_\infty : \|x \oplus y\|_{n+m} = \max\{\|x\|_n, \|y\|_m\}$ for $x \in M_n(V)$ and $y \in M_m(V)$.
- (ii) $\mathcal{M}_p : \|\alpha x \beta\|_n \leq \|\alpha\| \|x\|_n \|\beta\|$ for $x \in M_n(V)$ and $\alpha, \beta \in M_n^p$.

Let V and W be p -operator spaces. We say that a linear map $\varphi : V \rightarrow W$ is p -completely bounded if

$$\|\varphi\|_{pcb} = \sup_{n \in \mathbb{Z}_{>0}} \{\|\varphi_n\|\} < \infty,$$

where $\varphi_n : [x_{ij}] \in M_n(V) \rightarrow [\varphi(x_{ij})] \in M_n(W)$ is the induced map from $M_n(V)$ to $M_n(W)$. We say that φ is a p -complete contraction (respectively, a p -complete isometry) if $\|\varphi\|_{pcb} \leq 1$ (respectively, φ_n is an isometry for each $n \in \mathbb{Z}_{>0}$).

Let E be an L^p -space and let n be a positive integer. Then

$$E^n = l^p(\{1, 2, \dots, n\}, E)$$

with the norm $\|[x_i]\| = (\sum_{i=1}^n \|x_i\|^p)^{\frac{1}{p}}$ is again an L^p -space. We can obtain a norm $\|\cdot\|_n$ on the matrix space $M_n(\mathcal{B}(E))$ by the canonical identification $M_n(\mathcal{B}(E)) \cong \mathcal{B}(E^n)$. Then it follows from [27] that $\mathcal{B}(E)$ is a p -operator space. On the other hand, Le Merdy proves that every p -operator space is p -completely isometrically isomorphic to a norm-closed subspace of $\mathcal{B}(E)$ for some $E \in SQ_p$ (see [29, Theorem 4.1]), where SQ_p is the collection of subspaces of quotients of L^p -spaces. The reader is referred to [1; 10; 30] for more research on p -operator spaces.

The p -operator space projective tensor norm $\|\cdot\|_{\wedge_p, n}$ on $M_n(V \otimes W)$ is defined by $\|u\|_{\wedge_p, n} = \inf\{\|\alpha\| \|v\| \|w\| \|\beta\| : u = \alpha(v \otimes w)\beta$

$$\text{for } \alpha \in M_{n,kl}^p, v \in M_k(V), w \in M_l(W) \text{ and } \beta \in M_{kl,n}^p\}.$$

We let $V \hat{\otimes}^p W$ denote the completion of $V \otimes W$ with respect to this matrix norm, and call $V \hat{\otimes}^p W$ the p -operator space projective tensor product of V and W .

The following lemma is a functorial property of the p -operator space projective tensor product.

Lemma 2.4 [1, p. 938]. *Let V_1, V_2, W_1 and W_2 be p -operator spaces. If*

$$u_i : V_i \rightarrow W_i, \quad i = 1, 2,$$

are p -complete contractions, then the corresponding mapping

$$u_1 \otimes u_2 : V_1 \otimes V_2 \rightarrow W_1 \otimes W_2$$

extends to a p -complete contraction

$$u_1 \hat{\otimes}^p u_2 : V_1 \hat{\otimes}^p V_2 \rightarrow W_1 \hat{\otimes}^p W_2.$$

We let $\mathcal{CB}_p(V, W)$ denote the space of p -completely bounded maps from V to W . It follows from Le Merdy's characterization theorem that $\mathcal{CB}_p(V, W)$ is a p -operator space with the matrix norm given by

$$M_n(\mathcal{CB}_p(V, W)) = \mathcal{CB}_p(V, M_n(W)).$$

In particular, the dual space $V' = \mathcal{CB}_p(V, \mathbb{C})$ has a natural p -operator space structure given by

$$M_n(V') = \mathcal{CB}_p(V, M_n^p).$$

Let V and W be p -operator spaces. There exists an injective embedding

$$\theta : x \otimes y \in V \otimes W \hookrightarrow \theta(x \otimes y) \in \mathcal{CB}_p(V', W)$$

given by $\theta(x \otimes y)(f) = f(x)y$ for $f \in V'$. The completion $V \overset{\vee}{\otimes} W$ of $V \otimes W$ in $\mathcal{CB}_p(V', W)$ is a p -operator subspace of $\mathcal{CB}_p(V', W)$. We call $V \overset{\vee}{\otimes} W$ the p -operator space injective tensor product of V and W . Let $M_m(V')_1$ and $M_k(W')_1$ denote the closed unit ball of $M_m(V')$ and $M_k(W')$, respectively. It follows from [1] that for each $u \in M_n(V \otimes W)$, the p -operator space injective tensor norm $\|u\|_{\vee_p, n}$ can be expressed by

$$\|u\|_{\vee_p, n} = \sup \{ \|(\varphi \otimes \psi)_n(u)\| : \varphi \in M_m(V')_1, \psi \in M_k(W')_1, m, k \in \mathbb{Z}_{>0} \}.$$

The following lemma is a functorial property of the p -operator space injective tensor product.

Lemma 2.5 [1, p. 942]. *Let V_1, V_2, W_1 and W_2 be p -operator spaces. If*

$$u_i : V_i \rightarrow W_i, \quad i = 1, 2,$$

are p -complete contractions, then the corresponding mapping

$$u_1 \otimes u_2 : V_1 \otimes V_2 \rightarrow W_1 \otimes W_2$$

extends to a p -complete contraction

$$u_1 \overset{\vee}{\otimes} u_2 : V_1 \overset{\vee}{\otimes} V_2 \rightarrow W_1 \overset{\vee}{\otimes} W_2.$$

2.3. Spatial L^p -operator tensor products and p -completely bounded maps of L^p -operator algebras. Let $p \geq 1$. Let (X, μ) and (Y, ν) be two measure spaces, there is an L^p -tensor product $L^p(X, \mu) \otimes_p L^p(Y, \nu)$ which can be canonical identified with $L^p(X \times Y, \mu \times \nu)$ via $\xi \otimes \eta(x, y) = \xi(x)\eta(y)$ for all $\xi \in L^p(X, \mu)$ and $\eta \in L^p(Y, \nu)$. If $a \in \mathcal{B}(L^p(X, \mu))$ and $b \in \mathcal{B}(L^p(Y, \nu))$, then there is a corresponding tensor product operator $a \otimes b \in \mathcal{B}(L^p(X \times Y, \mu \times \nu))$. Let $A \subset \mathcal{B}(L^p(X, \mu))$ and $B \subset \mathcal{B}(L^p(Y, \nu))$ be two norm-closed subalgebras. Define an algebra

$$A \otimes_p B \subset \mathcal{B}(L^p(X \times Y, \mu \times \nu))$$

to be the closed linear span of all $a \in A$ and $b \in B$. Then $A \otimes_p B$ is an L^p -operator algebra, and it is called the spatial L^p -operator tensor product of A and B .

Remark 2.6. Let $A \subset \mathcal{B}(L^p(X, \mu))$ and $B \subset \mathcal{B}(L^p(Y, \nu))$ be L^p -operator algebras. Then it follows from [1, Theorem 3.3] that $A \overset{\vee}{\otimes} B$ is p -completely isometric to $A \otimes_p B$.

Given a norm-closed subalgebra A of $\mathcal{B}(L^p(X, \mu))$, the spatial tensor product $M_n^p \otimes_p A$ is the L^p -matrix algebra. Clearly, each element of $M_n^p \otimes_p A$ is of form $[a_{i,j}]_{1 \leq i,j \leq n}$ with $a_{i,j} \in A$, which is also written as $\sum_{i,j=1}^n e_{i,j} \otimes a_{i,j}$, where $\{e_{i,j}\}_{1 \leq i,j \leq n}$ are the canonical matrix units of M_n^p .

Definition 2.7. Let A be a closed subalgebra of $\mathcal{B}(L^p(X, \mu))$, B be a closed subalgebra of $\mathcal{B}(L^p(Y, \nu))$ and φ be a linear map $\varphi : A \rightarrow B$. We denote by φ_n the map from $M_n^p \otimes_p A$ to $M_n^p \otimes_p B$ defined by

$$\varphi_n \left(\sum_{i,j=1}^n e_{i,j} \otimes a_{i,j} \right) = \sum_{i,j=1}^n e_{i,j} \otimes \varphi(a_{i,j})$$

for $\sum_{i,j=1}^n e_{i,j} \otimes a_{i,j} \in M_n^p \otimes_p A$. We denote

$$\|\varphi\|_{pcb} = \sup_{n \in \mathbb{Z}_{>0}} \|\varphi_n\|.$$

We say that φ is p -completely bounded if $\|\varphi\|_{pcb} \leq C$ for some positive constant C , say that φ is p -completely contractive if $\|\varphi\|_{pcb} \leq 1$, and say that φ is p -completely isometric if φ_n is isometric for all positive integer n .

2.4. p -nuclearity and p -semidiscreteness. We now define p -nuclearity.

Definition 2.8 [1, Proposition 5.1(a)]. Let (X, \mathcal{B}, μ) be a measure space, and let $A \subset \mathcal{B}(L^p(X, \mu))$ be a norm-closed subalgebra. We say that A is p -nuclear if there exist nets of p -completely contractive maps $\varphi_\alpha : A \rightarrow M_{n(\alpha)}^p$ and $\psi_\alpha : M_{n(\alpha)}^p \rightarrow A$ such that

$$\|\psi_\alpha \circ \varphi_\alpha(a) - a\| \rightarrow 0$$

for all $a \in A$.

Definition 2.9. The p -pseudomeasure algebra $PM_p(G)$ is p -semidiscrete if there exist nets of weak* continuous p -completely contractive maps $\varphi_\alpha : PM_p(G) \rightarrow M_{n(\alpha)}^p$ and $\psi_\alpha : M_{n(\alpha)}^p \rightarrow PM_p(G)$ such that

$$\langle \psi_\alpha \circ \varphi_\alpha(a) - a, f \rangle \rightarrow 0$$

for all $a \in PM_p(G)$ and $f \in A_p(G)$.

When $p = 2$ and A is a C^* -algebra, R. R. Smith proved that p -nuclearity is equivalent to nuclearity (see [39, Theorem 1.1]).

Remark 2.10. The p -nuclearity is not equivalent to the amenability of L^p -operator algebras. The reader is referred to [43, Remark 1.4(iii)] for some examples.

Example 2.11 [1; 43]. Let $p \in [1, \infty)$. The following are examples of p -nuclear L^p -operator algebras:

- (i) $C(X)$, where X is a compact metric space;
- (ii) M_n^p and $\overline{\bigcup_{n=1}^{\infty} M_n^p}$;
- (iii) the reduced group L^p -operator algebra $F_\lambda^p(G)$, where G is a discrete amenable group;
- (iv) the L^p -Cuntz algebra \mathcal{O}_d^p ;
- (v) the rotation L^p -operator algebras $F^p(\mathbb{Z}, F^p(\mathbb{Z}), \beta_\theta)$ and $F^p(\mathbb{Z}, S^1, \alpha_\theta)$.

3. Proof of Theorem 1.4

We show (i) \Rightarrow (ii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (vii) \Rightarrow (i), then (i) \Rightarrow (iii) \Rightarrow (v); this will prove Theorem 1.4.

(i) \Rightarrow (ii): This follows from [1, Proposition 5.1].

(ii) \Rightarrow (iv): Let $\varphi_\alpha : F_\lambda^p(G) \rightarrow M_{n(\alpha)}^p$ and $\psi_\alpha : M_{n(\alpha)}^p \rightarrow F_\lambda^p(G)$ be the nets of p -completely contractive maps such that

$$\|\psi_\alpha \circ \varphi_\alpha(a) - a\| \rightarrow 0$$

for all $a \in F_\lambda^p(G)$. Since $M_{n(\alpha)}^p$ has p -OAP, it follows from [1, Theorem 3.12] that there exists an isomorphism Ψ from $M_{n(\alpha)}^p \overset{\vee}{\otimes} F_\lambda^p(G)$ to $M_{n(\alpha)}^p \overset{\wedge}{\otimes} F_\lambda^p(G)$. By Lemmas 2.4 and 2.5, we have

$$\begin{array}{ccc} F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G) & \xrightarrow{\varphi_\alpha \overset{\vee}{\otimes} \text{Id}_{F_\lambda^p(G)}} & M_{n(\alpha)}^p \overset{\vee}{\otimes} F_\lambda^p(G) \\ & & \downarrow \Psi \\ F_\lambda^p(G) \overset{\wedge}{\otimes} F_\lambda^p(G) & \xleftarrow{\psi_\alpha \overset{\wedge}{\otimes} \text{Id}_{F_\lambda^p(G)}} & M_{n(\alpha)}^p \overset{\wedge}{\otimes} F_\lambda^p(G) \end{array}$$

Define $\Phi_\alpha = (\psi_\alpha \overset{\wedge}{\otimes} \text{Id}_{F_\lambda^p(G)}) \circ \Psi \circ (\varphi_\alpha \overset{\vee}{\otimes} \text{Id}_{F_\lambda^p(G)})$. Then Φ_α is a bounded linear map from $F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G)$ to $F_\lambda^p(G) \overset{\wedge}{\otimes} F_\lambda^p(G)$. We denote by $\text{Id}_{F_\lambda^p(G)} \otimes \text{Id}_{F_\lambda^p(G)} : F_\lambda^p(G) \otimes F_\lambda^p(G) \rightarrow F_\lambda^p(G) \otimes F_\lambda^p(G)$ the algebraic tensor product map. Since $\|\Phi_\alpha(x) - \text{Id}_{F_\lambda^p(G)} \otimes \text{Id}_{F_\lambda^p(G)}(x)\| \rightarrow 0$ for all $x \in F_\lambda^p(G) \otimes F_\lambda^p(G)$, it follows that $\text{Id}_{F_\lambda^p(G)} \otimes \text{Id}_{F_\lambda^p(G)}$ is a bounded linear map from the p -operator space injective tensor norm on $F_\lambda^p(G) \otimes F_\lambda^p(G)$ to the p -operator space projective tensor norm on $F_\lambda^p(G) \otimes F_\lambda^p(G)$. Hence it extends to a bounded linear map $\Phi : F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G) \rightarrow F_\lambda^p(G) \overset{\wedge}{\otimes} F_\lambda^p(G)$. Since $\|\cdot\|_{\wedge, p}$ is the largest p -operator space norm [10, Proposition 4.8], it follows that Φ is an isomorphism. This proves (iv).

(iv) \Rightarrow (v): Recall that $\rho_p : G \rightarrow \mathcal{B}(l^p(G))$ is the right regular representation. We denote by $\lambda_p \cdot \rho_p$ the biregular representation $(s, t) \rightarrow \lambda_p(s)\rho_p(t)$ of $G \times G$ on $l^p(G)$. Since $\|\cdot\|_{\wedge_p}$ is the largest p -operator space norm [10, Proposition 4.8], it follows that the canonical linear map $h : F_\lambda^p(G) \otimes F_\lambda^p(G) \rightarrow B(l^p(G))$ defined by

$$h(\lambda_p(s) \otimes \lambda_p(t)) = \lambda_p(s)\rho_p(t)$$

has a continuous extension on $F_\lambda^p(G) \hat{\otimes} F_\lambda^p(G)$ and it is denoted by $(\lambda_p \cdot \rho_p)_\lambda$. By (iv), there exists a bounded linear map from $F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G)$ to $B(l^p(G))$, which is denoted by $(\widetilde{\lambda_p \cdot \rho_p})_\lambda$. This proves (v).

(v) \Rightarrow (vi): We recall that σ_p is the conjugacy representation $s \mapsto \lambda_p(s)\rho_p(s)$ of G on $l^p(G)$. By (iv), the following diagram is commutative:

$$\begin{CD} F^p(G) @>\sigma_p>> B(l^p(G)) \\ @V\lambda_pVV @A(\widetilde{\lambda_p \cdot \rho_p})_\lambda AA \\ F_\lambda^p(G) @>\iota>> F_\lambda^p(G) \overset{\vee}{\otimes} F_\lambda^p(G). \end{CD}$$

Here $\iota(s) = \lambda_p(s) \otimes \lambda_p(s)$ for each $s \in G$. Let $\theta = (\widetilde{\lambda_p \cdot \rho_p})_\lambda \circ \iota$. Then $\sigma_p = \theta \circ \lambda_p$.

Claim 1: $\|\theta\| \leq 1$. To see this, recall that $\rho_p(s)\delta_t = \delta_{ts^{-1}}$ for all $s \in G$. Then

$$\|\lambda_p(f)\| = \|\rho_p(f)\|$$

for all $f \in C_c(G)$. In fact, let $V : l^p(G) \rightarrow l^p(G)$ be the invertible isometry given by $V\delta_t = \delta_{t^{-1}}$. Then one can check that $V\lambda_p(f)V^{-1}\delta_t = \rho_p(f)\delta_t$. This shows that $\|\lambda_p(f)\| = \|\rho_p(f)\|$.

For any $f \in C_c(G)$ with $\|\lambda_p(f)\| \leq 1$ and $\xi \in l^p(G)$, we have

$$\|\theta(f)\xi\| = \|\lambda_p(f)\rho_p(f)\xi\| \leq \|\lambda_p(f)\| \cdot \|\rho_p(f)\| \cdot \|\xi\| = \|\lambda_p(f)\|^2 \cdot \|\xi\|.$$

Hence $\|\theta(f)\| \leq \|\lambda_p(f)\|^2 \leq 1$, and therefore $\|\theta\| \leq 1$. This proves Claim 1.

Now we will prove (vi). Since $\sigma_p = \theta \circ \lambda_p$ and $\|\theta\| \leq 1$, it follows that

$$\|\sigma_p(f)\| = \|\theta \circ \lambda_p(f)\| \leq \|\lambda_p(f)\|.$$

It is easy to check that $\sigma_p(s)\delta_e = \delta_e$. Hence

$$\|\lambda_p(f)\| \geq \|\sigma_p(f)\| \geq \|\sigma_p(f)\delta_e\| = \left| \sum_{t \in G} f(t) \right|.$$

This proves (vi).

(vi) \Rightarrow (vii): For any finite subset $E \subset G$, by (vi), we have $\|\sum_{t \in E} \lambda_p(t)\| \geq |E|$. Obviously, $\|\sum_{t \in E} \lambda_p(t)\| \leq |E|$. This proves (vii).

(vii) \Rightarrow (i): For any finite subset $E \subset G$, we can assume that $e \in E$, where e is the unit of G . By (v), we have $\|\sum_{t \in E} \lambda_p(t)/|E|\| = 1$. Then there exists a sequence

(ξ_i) in $l^p(G)$ such that $\xi_i \geq 0$, $\|\xi_i\|_p = 1$ and $\|\sum_{t \in E} \lambda_p(t)\xi_i/|E|\| \rightarrow 1$. Since $l^p(G)$ is a uniformly convex Banach space for $p \in (1, \infty)$, it follows from [14] that $l^p(G)$ is a full k -convex Banach space for all positive integer $k \geq 2$. Then

$$\|\lambda_p(s)\xi_i - \lambda_p(t)\xi_i\|_p \rightarrow 0$$

for all $s, t \in E$. Since $e \in E$, it follows that

$$\|\lambda_p(s)\xi_i - \xi_i\|_p \rightarrow 0$$

for all $s \in E$. Then there exists a net (η_α) in $l^p(G)$ such that $\eta_\alpha \geq 0$, $\|\eta_\alpha\|_p = 1$ and

$$\|\lambda_p(s)\eta_\alpha - \eta_\alpha\|_p \rightarrow 0$$

for all $s \in G$. By Definition 2.2, we have that G is amenable.

(i) \Rightarrow (iii): It follows from [1, Proposition 5.1].

(iii) \Rightarrow (v): Let $\varphi_\alpha : PM_p(G) \rightarrow M_{n(\alpha)}^p$ and $\psi_\alpha : M_{n(\alpha)}^p \rightarrow PM_p(G)$ be the nets of weak* continuous p -completely contractive maps such that

$$\langle \psi_\alpha \circ \varphi_\alpha(a) - a, f \rangle \rightarrow 0$$

for all $a \in PM_p(G)$ and $f \in A_p(G)$.

Since $M_{n(\alpha)}^p$ has p -OAP, it follows from [1, Theorem 3.12] that there exists an isomorphism Ψ from $M_{n(\alpha)}^p \overset{\vee}{\otimes} PM_p(G)$ to $M_{n(\alpha)}^p \overset{\wedge}{\otimes} PM_p(G)$. By Lemmas 2.4 and 2.5, we have

$$\begin{array}{ccc} PM_p(G) \overset{\vee}{\otimes} PM_p(G) & \xrightarrow{\varphi_\alpha \overset{\vee}{\otimes} \text{Id}_{PM_p(G)}} & M_{n(\alpha)}^p \overset{\vee}{\otimes} PM_p(G) \\ & & \downarrow \Psi \\ PM_p(G) \overset{\wedge}{\otimes} PM_p(G) & \xleftarrow{\psi_\alpha \overset{\wedge}{\otimes} \text{Id}_{PM_p(G)}} & M_{n(\alpha)}^p \overset{\wedge}{\otimes} PM_p(G) \end{array}$$

Since $\|\cdot\|_{\wedge_p}$ is the largest p -operator space norm [10, Proposition 4.8], it follows that the canonical linear map

$$h : PM_p(G) \otimes PM_p(G) \rightarrow B(l^p(G))$$

defined by

$$h(\lambda_p(s) \otimes \lambda_p(t)) = \lambda_p(s)\rho_p(t)$$

has a continuous extension on $PM_p(G) \overset{\wedge}{\otimes} PM_p(G)$ and it is denoted by $(\lambda_p \cdot \rho_p)_\lambda$.

Then we can define a net of bounded linear maps

$$\Phi_\alpha : PM_p(G) \overset{\vee}{\otimes} PM_p(G) \rightarrow \mathcal{B}(l^p(G))$$

by $\Phi_\alpha = ((\lambda_p \cdot \rho_p)_\lambda) \circ (\psi_\alpha \overset{\wedge}{\otimes} \text{Id}_{PM_p(G)}) \circ \Psi \circ (\varphi_\alpha \overset{\vee}{\otimes} \text{Id}_{PM_p(G)})$. Since $\mathcal{B}(l^p(G)) = \mathcal{N}(l^p(G))'$, it follows from [4, Theorem 1.3.7] that there exists a point-weak* cluster

point Φ of the net (Φ_α) . Then we get a bounded linear map

$$\Phi : PM_p(G) \overset{\vee}{\otimes} PM_p(G) \rightarrow \mathcal{B}(l^p(G)).$$

Claim 2: *The bounded linear map Φ extends the map $h : PM_p(G) \otimes PM_p(G) \rightarrow \mathcal{B}(l^p(G))$ given by $h(\lambda_p(s) \otimes \lambda_p(t)) = \lambda_p(s)\rho_p(t)$.*

Indeed, since $PM_p(G)$ is p -semidiscrete, we have

$$\langle \psi_\alpha \circ \varphi_\alpha(\lambda_p(s)) - \lambda_p(s), f \rangle = \sum_{n=1}^{\infty} \langle \xi_n, (\psi_\alpha \circ \varphi_\alpha(\lambda_p(s)) - \lambda_p(s))(\eta_n) \rangle \rightarrow 0,$$

for all $f = \sum_{n=1}^{\infty} \xi_n * \check{\eta}_n \in A_p(G)$ and $s, t \in G$. Then, for any

$$g = \sum_{n=1}^{\infty} x_n \otimes y_n \in l^{p'}(G) \widehat{\otimes} l^p(G),$$

we have

$$\begin{aligned} & \langle \Phi_\alpha(\lambda_p(s) \otimes \lambda_p(t)) - \lambda_p(s)\rho_p(t), g \rangle \\ &= \langle (\lambda_p \cdot \rho_p)_\lambda(\psi_\alpha \circ \varphi_\alpha(\lambda_p(s)) \otimes \lambda_p(t)) - \lambda_p(s)\rho_p(t), g \rangle \\ &= \langle \psi_\alpha \circ \varphi_\alpha(\lambda_p(s))\rho_p(t) - \lambda_p(s)\rho_p(t), g \rangle \\ &= \langle (\psi_\alpha \circ \varphi_\alpha(\lambda_p(s)) - \lambda_p(s))\rho_p(t), g \rangle \\ &= \sum_{n=1}^{\infty} \langle x_n, (\psi_\alpha \circ \varphi_\alpha(\lambda_p(s)) - \lambda_p(s))\rho_p(t)y_n \rangle \\ &\rightarrow 0. \end{aligned}$$

It follows that

$$\begin{aligned} & | \langle \Phi(\lambda_p(s) \otimes \lambda_p(t)) - \lambda_p(s)\rho_p(t), g \rangle | \\ &\leq | \langle \Phi(\lambda_p(s) \otimes \lambda_p(t)) - \Phi_\alpha(\lambda_p(s) \otimes \lambda_p(t)), g \rangle | \\ &\quad + | \langle \Phi_\alpha(\lambda_p(s) \otimes \lambda_p(t)) - \lambda_p(s)\rho_p(t), g \rangle | \\ &\rightarrow 0. \end{aligned}$$

Hence $\Phi(\lambda_p(s) \otimes \lambda_p(t)) = h(\lambda_p(s) \otimes \lambda_p(t))$, proving Claim 2. Then (v) follows easily. □

Acknowledgements

The author is supported by National Natural Science Foundation of China (grant number 12201240) and Research Start-up Funding Program of Hangzhou Normal University (grant number 4235C50224204070).

References

[1] G. An, J.-J. Lee, and Z.-J. Ruan, “On p -approximation properties for p -operator spaces”, *J. Funct. Anal.* **259**:4 (2010), 933–974. MR

- [2] C. Anantharaman-Delaroche, “On tensor products of group C^* -algebras and related topics”, pp. 1–35 in *Limits of graphs in group theory and computer science*, EPFL Press, Lausanne, 2009. MR
- [3] D. P. Blecher and N. C. Phillips, “ L^p -operator algebras with approximate identities, I”, *Pacific J. Math.* **303**:2 (2019), 401–457. MR
- [4] N. P. Brown and N. Ozawa, *C^* -algebras and finite-dimensional approximations*, Graduate Studies in Mathematics **88**, Amer. Math. Soc., 2008. MR
- [5] M. D. Choi and E. G. Effros, “Nuclear C^* -algebras and the approximation property”, *Amer. J. Math.* **100**:1 (1978), 61–79. MR
- [6] Y. Choi, E. Gardella, and H. Thiel, “Rigidity results for L^p -operator algebras and applications”, *Adv. Math.* **452** (2024), art. id. 109747, 47 pp. MR
- [7] Y. C. Chung and K. Li, “Rigidity of ℓ^p Roe-type algebras”, *Bull. Lond. Math. Soc.* **50**:6 (2018), 1056–1070. MR
- [8] A. Connes, “On the cohomology of operator algebras”, *J. Functional Analysis* **28**:2 (1978), 248–253. MR
- [9] H. G. Dales, P. Aiena, J. Eschmeier, K. Laursen, and G. A. Willis, *Introduction to Banach algebras, operators, and harmonic analysis*, London Mathematical Society Student Texts **57**, Cambridge University Press, 2003. MR
- [10] M. Daws, “ p -operator spaces and Figà–Talamanca–Herz algebras”, *J. Operator Theory* **63**:1 (2010), 47–83. MR
- [11] M. Daws and N. Spronk, “On convoluters on L^p -spaces”, *Studia Math.* **245**:1 (2019), 15–31. MR
- [12] A. Derighetti, *Convolution operators on groups*, Lecture Notes of the Unione Matematica Italiana **11**, Springer, 2011. MR
- [13] E. M. Elkiær, “Symmetrized pseudofunction algebras from L^p -representations and amenability of locally compact groups”, *Expo. Math.* **43**:4 (2025), art. id. 125685, 20 pp. MR
- [14] K. Fan and I. Glicksberg, “Some geometric properties of the spheres in a normed linear space”, *Duke Math. J.* **25** (1958), 553–568. MR
- [15] E. Gardella, “A modern look at algebras of operators on L^p -spaces”, *Expo. Math.* **39**:3 (2021), 420–453. MR
- [16] E. Gardella and M. Lupini, “Representations of étale groupoids on L^p -spaces”, *Adv. Math.* **318** (2017), 233–278. MR
- [17] E. Gardella and H. Thiel, “Banach algebras generated by an invertible isometry of an L^p -space”, *J. Funct. Anal.* **269**:6 (2015), 1796–1839. MR
- [18] E. Gardella and H. Thiel, “Group algebras acting on L^p -spaces”, *J. Fourier Anal. Appl.* **21**:6 (2015), 1310–1343. MR
- [19] E. Gardella and H. Thiel, “Representations of p -convolution algebras on L^q -spaces”, *Trans. Amer. Math. Soc.* **371**:3 (2019), 2207–2236. MR
- [20] E. Gardella and H. Thiel, “Isomorphisms of algebras of convolution operators”, *Ann. Sci. Éc. Norm. Supér. (4)* **55**:5 (2022), 1433–1471. MR
- [21] U. Haagerup, “All nuclear C^* -algebras are amenable”, *Invent. Math.* **74**:2 (1983), 305–319. MR
- [22] C. Herz, “The theory of p -spaces with an application to convolution operators”, *Trans. Amer. Math. Soc.* **154** (1971), 69–82. MR
- [23] C. Herz, “Harmonic synthesis for subgroups”, *Ann. Inst. Fourier (Grenoble)* **23**:3 (1973), 91–123. MR

- [24] C. Herz, “On the asymmetry of norms of convolution operators, I”, *J. Functional Analysis* **23**:1 (1976), 11–22. MR
- [25] B. E. Johnson, *Cohomology in Banach algebras*, Memoirs of the Amer. Math. Soc. **127**, Amer. Math. Soc., 1972. MR
- [26] E. Kirchberg, “ C^* -nuclearity implies CPAP”, *Math. Nachr.* **76** (1977), 203–212. MR
- [27] S. Kwapien, “On operators factorizable through L_p space”, *Bull. Soc. Math. France* **100** (1972), 215–225. MR
- [28] C. Lance, “On nuclear C^* -algebras”, *J. Functional Analysis* **12** (1973), 157–176. MR
- [29] C. Le Merdy, “Factorization of p -completely bounded multilinear maps”, *Pacific J. Math.* **172**:1 (1996), 187–213. MR
- [30] J. J. Lee, *On p -operator spaces and their applications*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2010, <https://www.proquest.com/docview/863835462>. MR
- [31] N. C. Phillips, “Analogues of Cuntz algebras on L^p spaces”, preprint, 2012. arXiv 1201.4196
- [32] N. C. Phillips, “Crossed products of L^p operator algebras and the K-theory of Cuntz algebras on L^p spaces”, preprint, 2013. arXiv 1309.6406
- [33] N. C. Phillips, “Isomorphism, nonisomorphism, and amenability of L^p UHF algebras”, preprint, 2013. arXiv 1309.3694v2
- [34] N. C. Phillips, “Simplicity of UHF and Cuntz algebras on L^p spaces”, preprint, 2013. arXiv 1309.0115
- [35] N. C. Phillips, “Open problems related to operator algebras on L^p spaces”, preprint, 2014, <https://pdfs.semanticscholar.org/0823/5038ec45079e7721a59021a4492da2c2b1a3.pdf>.
- [36] N. C. Phillips, “Operator algebras on L^p spaces which “look like” C^* -algebras”, lecture notes, 2014, https://pages.uoregon.edu/ncp/Talks/20140527_GPOTS/LpOpAlgs_TalkSummary.pdf.
- [37] N. C. Phillips and M. G. Viola, “Classification of spatial L^p AF algebras”, *Internat. J. Math.* **31**:13 (2020), art. id. 2050088, 41 pp. MR
- [38] G. Pisier, *Tensor products of C^* -algebras and operator spaces: the Connes–Kirchberg problem*, London Mathematical Society Student Texts **96**, Cambridge University Press, 2020. MR
- [39] R. R. Smith, “Completely contractive factorizations of C^* -algebras”, *J. Funct. Anal.* **64**:3 (1985), 330–337. MR
- [40] M. Takesaki, “On the cross-norm of the direct product of C^* -algebras”, *Tohoku Math. J. (2)* **16** (1964), 111–122. MR
- [41] Q. Wang and Z. Wang, “Notes on the ℓ^p -Toeplitz algebra on $\ell^p(\mathbb{N})$ ”, *Israel J. Math.* **245**:1 (2021), 153–163. MR
- [42] Z. Wang and S. Zhu, “On the Takai duality for L^p operator crossed products”, *Math. Z.* **304**:4 (2023), art. id. 54, 23 pp. MR
- [43] Z. Wang and S. Zhu, “ p -nuclearity of L^p -operator crossed products”, *Israel J. Math.* (published online November 2025).

Received November 23, 2025. Revised February 25, 2026.

ZHEN WANG
 SCHOOL OF MATHEMATICS
 HANGZHOU NORMAL UNIVERSITY
 HANGZHOU, ZHEJIANG
 CHINA
 wangzhen@hznu.edu.cn

A HIGHER-RANK ANALOG OF THE STRONG OPENNESS PROPERTY

JINGCAO WU

We study a strong openness property for singular Hermitian vector bundles (E, h) that are Griffiths-semipositive.

1. Introduction

The strong openness conjecture is an important subject in complex geometry. It was completely solved by Q. Guan and X. Zhou as follows:

Theorem 1.1 [11, Theorem 1.1]. *Let Δ^n be the unit polydisc, and let (L, φ) be a singular Hermitian line bundle on Δ^n . Assume that the associated curvature satisfies $i\Theta_{L, \varphi} \geq 0$, and that F is a holomorphic section of L which satisfies*

$$\int_{\Delta^n} |F|^2 e^{-\varphi} d\lambda < \infty,$$

where $d\lambda$ is the Lebesgue measure. After shrinking Δ^n if necessary, there exists a positive ε such that

$$\int_{\Delta^n} |F|^2 e^{-(1+\varepsilon)\varphi} d\lambda < \infty.$$

Equivalently, the following equality concerning multiplier ideal sheaves holds:

$$(1) \quad \bigcup_{\varepsilon > 0} \mathcal{I}((1+\varepsilon)\varphi) = \mathcal{I}(\varphi).$$

It has led to fruitful developments, such as [3; 1; 2; 13; 9; 12]. Among them, we mention the following variant for later use.

Corollary 1.1 [3, Corollary B.2]. *Let φ, ψ be plurisubharmonic functions on a domain $U \subseteq \mathbb{C}^n$. Assume that ψ has analytic singularities. Then for any compact subset $K \Subset U$, there exists a positive ε such that*

$$e^{\psi-\varphi} \in L_{\text{loc}}^1(U, K) \quad \text{if and only if} \quad e^{\psi-(1+\varepsilon)\varphi} \in L_{\text{loc}}^1(U, K).$$

This research was supported by NSFC (grant 12201381) and Shanghai Scientific Sailing Program (grant 22YF1412500).

MSC2020: 32L10, 32U05, 32D15.

Keywords: strong openness, singular metric, vector bundle.

Here $L^1_{\text{loc}}(U, K)$ denotes the set of Lebesgue measurable functions f defined in a neighborhood of K in U such that $|f|$ is locally integrable in a neighborhood of every point $x \in K$.

It is natural to ask if the *strong openness property* holds for the higher-rank vector bundles. There are some discussions in [16; 18; 19], and we should make a brief review first. For an arbitrary holomorphic vector bundle E and a singular Hermitian metric h , they considered the multiplier submodule sheaf $\mathcal{E}(h)$ defined as

$$\mathcal{E}(h)_x := \{F_x \in E_x \mid |F_x|_h^2 \text{ is integrable around } x\}.$$

Suppose that $\{h_j\}$ is a sequence of singular metrics decreasing to h . Then, under a certain Nakano-type positivity (on $\{h_j\}$ or h), they concluded that

$$(2) \quad \bigcup_j \mathcal{E}(h_j) = \mathcal{E}(h).$$

In this paper, we investigate this problem from another point of view. More precisely, we apply the deep relationship between the Hermitian metrics on E and its tautological line bundle $\mathcal{O}_E(1)$, to reduce everything to the rank one case. This theory is known as the Finsler geometry, which will be recalled later. Then we obtain that

Theorem 1.2. *Let E be a holomorphic vector bundle of rank r over Δ^n , and let h be a singular Hermitian metric on E such that (E, h) is Griffiths semipositive. Assume that $\det h$ has analytic singularities. If F is a holomorphic section of E satisfying*

$$\int_{\Delta^n} |F|_h^2 d\lambda < \infty,$$

then after shrinking Δ^n if necessary, there exists a positive ε (independent of F) such that

$$\int_{\Delta^n} |F|_h^{2(1+\varepsilon)} d\lambda < \infty.$$

In principle, the Griffiths-type positivity is strictly weaker than the Nakano-type positivity. But we should also notice that [16; 18; 19] made no restrictions on the singularity of $\det h$. Hence, there is no direct relationship between [16; 18; 19] and Theorem 1.2.

Next, let us consider general singular metrics. Remember that in the line bundle case, if we furthermore suppose that $\mathcal{S}(\varphi) = \mathcal{O}_{\Delta^n}$, the equality (1) is then called the *openness property*. It was proved by [4], and certainly can be seen as a special case of Theorem 1.1. We generalize it to the higher-rank vector bundles as follows:

Theorem 1.3. *Let E be a holomorphic vector bundle of rank r over Δ^n_R , where Δ^n_R is a polydisc with radius R slightly larger than 1. Let h be a singular Hermitian*

metric on E such that (E, h) is Griffiths semipositive, and let φ be the metric on $\mathcal{O}_E(1)$ induced by h . Assume that

$$(3) \quad \mathcal{I}((r + 1)\varphi) = \mathcal{O}_{\mathbb{P}(E^*)}.$$

Then there exist positive numbers \hat{R} and ε such that

$$\int_{\Delta_{\hat{R}}^n} |F|_h^{2(1+\varepsilon)} d\lambda < \infty$$

for any holomorphic section F of E .

As we will see in Section 4, (3) implies that

$$\int_{\Delta_R^n} |F|_h^2 d\lambda < \infty$$

for any F , namely

$$\mathcal{E}(h) = E.$$

But the converse seems not obvious.

This paper is organized as follows: in Section 2 we will review the basic materials of singular metrics and Finsler geometry, and in Section 3 we give the proof of Theorem 1.2. In the end we prove Theorem 1.3.

2. Preliminary

Let X be a complex manifold of dimension n , and let $p : E \rightarrow X$ be a holomorphic vector bundle over X of rank r .

2.1. Singular Hermitian metrics.

Definition 2.1 [5; 21; 22]. A singular Hermitian metric on E is a map h that associates to every point $x \in X$ a singular Hermitian inner product $|\cdot|_{h,x} : E_x \rightarrow [0, +\infty]$ on the complex vector space E_x , subject to the following two conditions:

1. h is finite and positive definite almost everywhere, meaning that for all x outside a set of Lebesgue measure zero, $|\cdot|_{h,x}$ is a Hermitian inner product on E_x ;
2. h is measurable, meaning that the function

$$|F|_h : U \rightarrow [0, +\infty], \quad x \mapsto |F(x)|_{h,x},$$

is measurable for any open $U \subseteq X$ and $F \in \Gamma(U, E)$.

Definition 2.2 [22]. Let h be a singular Hermitian metric on E , which canonically induces a singular metric h^* on the dual bundle E^* .

- (1) (E, h) is called Griffiths-seminegative (or negatively curved) if, for any (local) holomorphic section F of E , the function $\log |F|_h^2$ is plurisubharmonic.

- (2) (E, h) is called Griffiths-semipositive (or positively curved) if (E^*, h^*) is Griffiths-seminegative.

Proposition 2.1 [22, Proposition 1.3]. *Let E be a holomorphic vector bundle over Δ^n , equipped with a singular Hermitian metric h .*

- (1) (See also [5, Proposition 3.1].) *If (E, h) is Griffiths-semipositive, then over any smaller polydisk there is a sequence of smooth, Griffiths positive metrics $\{h_\nu\}$ increasing pointwise to h .*
- (2) *If (E, h) is Griffiths-seminegative, then $\log \det h$ is a plurisubharmonic function.*

Definition 2.3 [6]. A plurisubharmonic function φ is said to have analytic singularities if φ can be written locally as

$$\varphi = \alpha \log(|f_1|^2 + \dots + |f_N|^2) + v,$$

where $\alpha \in \mathbb{R}_+$, v is a bounded function and the f_j are holomorphic functions.

When (E, h) is a Griffiths-semipositive singular Hermitian vector bundle, the function $-\log \det h$ is plurisubharmonic by Proposition 2.1(2). Then we can ask it has analytic singularities, and briefly say that $\det h$ has analytic singularities. Obviously, it not necessarily implies that h itself has analytic singularities.

2.2. Finsler geometry revisited. We only collect the necessary materials. One could refer to [8; 15] for a sophisticated comprehension.

Let $x = (x^1, \dots, x^n)$ be a local coordinate system in X and let $w = (w^0, \dots, w^{r-1})$ be the fiber coordinate system defined by a local holomorphic frame

$$W = \{W^0, \dots, W^{r-1}\}$$

of E . Let h be a smooth Hermitian metric on E . We write

$$h_i = \frac{\partial h}{\partial x^i}, \quad h_{\bar{j}} = \frac{\partial h}{\partial \bar{x}^j}, \quad h_{i\bar{j}} = \frac{\partial^2 h}{\partial x^i \partial \bar{x}^j}, \quad h_{i\alpha} = \frac{\partial^2 h}{\partial x^i \partial w^\alpha},$$

and so on, to denote the differentiation with respect to x^i, \bar{x}^j ($1 \leq i, j \leq n$) and w^α, \bar{w}^β ($0 \leq \alpha, \beta \leq r-1$).

Denote by $q : \mathbb{P}(E) \rightarrow X$ the natural projection from the projectivized bundle to the ambient space, and h induces a Hermitian metric q^*h on q^*E . Then as a subbundle,

$$\mathcal{O}_{\mathbb{P}(E)}(-1) := \{((x, [w]), Z) \in q^*E \mid Z = \lambda w, \lambda \in \mathbb{C}\}$$

inherits a metric from q^*E , whose weight function is denoted by $-\psi$. Accordingly, we can define the metrics $\psi, -\varphi$ and φ on $\mathcal{O}_{\mathbb{P}(E)}(1), \mathcal{O}_{\mathbb{P}(E^*)}(-1)$ and $\mathcal{O}_{\mathbb{P}(E^*)}(1)$ respectively in an obvious way. By definition, one easily verifies: when (E, h) is

Griffiths-semipositive, $(\mathcal{O}_{\mathbb{P}(E^*)}(1), \varphi)$ is semipositive. For this reason, we usually denote $\mathcal{O}_{\mathbb{P}(E^*)}(1)$ by $\mathcal{O}_E(1)$, and call it the tautological line bundle of E .

Remark 2.1. Following the same procedure, a singular Hermitian metric h on E induces singular metrics on $\mathcal{O}_{\mathbb{P}(E)}(-1)$, $\mathcal{O}_{\mathbb{P}(E)}(1)$, $\mathcal{O}_E(-1)$ and $\mathcal{O}_E(1)$ respectively.

Moreover, if (E, h) is Griffiths-semipositive, $\mathcal{O}_E(1)$ is pseudo-effective equipped with the corresponding metric as is shown in [21, Proposition 2.3.5].

Next let us recall the celebrated curvature formula in [15]. Remember that ψ is the weight function of the metric on $\mathcal{O}_{\mathbb{P}(E)}(1)$ induced by h . We first expand $i\partial\bar{\partial}\psi$ on $\mathbb{P}(E)$ as follows:

$$i\partial\bar{\partial}\psi = i(g_{i\bar{j}}dx^i \wedge d\bar{x}^j + g_{i\bar{\beta}}dx^i \wedge d\bar{e}^\beta + g_{\alpha\bar{j}}de^\alpha \wedge d\bar{x}^j + g_{\alpha\bar{\beta}}de^\alpha \wedge d\bar{e}^\beta).$$

Here (w^0, \dots, w^{r-1}) lifts as (e^1, \dots, e^{r-1}) , and we adopt the summation convention of Einstein. Note if we furthermore ask the holomorphic frame W to be normal with respect h at x , $ig_{\alpha\bar{\beta}}de^\alpha \wedge d\bar{e}^\beta$ is just the Fubini–Study metric on $\mathbb{P}(E)_x$. Thus the matrix $[g_{\alpha\bar{\beta}}]$ is invertible everywhere. Denote by $[g^{\beta\alpha}]$ the inverse matrix. Then we can define the conformal basis by

$$\{\delta e^\alpha = de^\alpha + g^{\beta\alpha}g_{i\bar{\beta}}dx^i, dx^i\}.$$

It is shown in [15] that on this basis we can rewrite $i\partial\bar{\partial}\psi$ as

$$(4) \quad i\partial\bar{\partial}\psi = -\Psi + \omega_{FS},$$

where

$$\Psi = i\Theta_{\alpha\bar{\beta}i\bar{j}} \frac{e^\alpha \bar{e}^\beta}{h_{\sigma\bar{\tau}} e^\sigma \bar{e}^\tau} dx^i \wedge d\bar{x}^j \quad \text{and} \quad \omega_{FS} = ig_{\alpha\bar{\beta}}\delta e^\alpha \wedge \delta \bar{e}^\beta.$$

Restricted on each $\mathbb{P}(E)_x$, ω_{FS} is just the Fubini–Study metric up to a coordinate transform.

We should also make a brief explanation of Ψ . $\Theta_{\alpha\bar{\beta}i\bar{j}}$ is the curvature tensor of E associated with h . Moreover, remember that (w^0, \dots, w^{r-1}) lifts to (e^1, \dots, e^{r-1}) . Locally, say on $\{w^0 \neq 0\}$, we have $e^\alpha = w^\alpha/w^0$ for $\alpha = 1, \dots, r-1$. In view of this relationship, Ψ is interpreted as

$$\Psi = i\Theta_{\alpha\bar{\beta}i\bar{j}} \frac{w^\alpha \bar{w}^\beta}{h_{\sigma\bar{\tau}} w^\sigma \bar{w}^\tau} dx^i \wedge d\bar{x}^j.$$

Since both $\Theta_{\alpha\bar{\beta}i\bar{j}}$ and $h_{\sigma\bar{\tau}}$ are independent of w^α and $w^\alpha \bar{w}^\beta / (h_{\sigma\bar{\tau}} w^\sigma \bar{w}^\tau)$ is homogeneous of degree 0, this expression is indeed a function of (e^1, \dots, e^{r-1}) . Therefore Ψ is well-defined.

In practice, we will apply (4) to the tautological bundle $(\mathcal{O}_E(1), \varphi)$ on $\mathbb{P}(E^*)$. It enables us to establish a crucial isometry:

$$K_{\mathbb{P}(E^*)/X}^{-1} \simeq \mathcal{O}_E(r) \otimes \pi^* \det E^*.$$

3. Strong openness

We briefly recall the L^2 -representation in [20; 17] for a (singular) Hermitian metric of a vector bundle, which helps us to reduce everything to the line bundle case; then we apply Corollary 1.1 to obtain our desired estimate.

We will work with the following setup: let X be a complex manifold, and let E be a holomorphic vector bundle of rank r over X . Let $\pi : \mathbb{P}(E^*) \rightarrow X$ be the natural projection. For a Hermitian metric h on E , we denote by φ the weight function of the induced metric h_L on $\mathcal{O}_E(1)$. Suppose that h is smooth, then $i\partial\bar{\partial}\varphi$ is a Fubini–Study-type metric along each fiber $\mathbb{P}(E^*)_x$. We briefly denote it by $\omega_x := (i\partial\bar{\partial}\varphi)|_{\mathbb{P}(E^*)_x}$, then the volume $\text{Vol}(\mathbb{P}(E^*)_x)$ against ω_x equals 1 for every x . This is standard in Finsler geometry, and one could refer to [7, Lemma 2.1] for a beautiful explanation.

3.1. The L^2 -representation. We have a canonical isomorphism $\pi_*\mathcal{O}_E(1) \simeq E$ [14, Chapter II, Proposition 7.11]. Hence, for any local holomorphic section F of E , $|F|^2e^{-\varphi}$ is understood in an obvious way. Combining with [20; 17], we obtain:

Proposition 3.1. *Up to a multiplication of a constant, for every $x \in X$ and $F \in E_x$ we have*

$$(5) \quad |F|_{h,x}^2 = \int_{\mathbb{P}(E^*)_x} |F|^2 e^{-\varphi} \omega_x^{r-1}.$$

Furthermore, φ induces an isometry between the canonical isomorphism:

$$K_{\mathbb{P}(E^*)/X}^{-1} \simeq \mathcal{O}_E(r) \otimes \pi^* \det E^*.$$

Namely, if (e^1, \dots, e^{r-1}) is a fiber coordinate system of $\mathbb{P}(E^*)$, then

$$(6) \quad \omega_x^{r-1} = i^{r-1} \pi^* \det h^* \cdot e^{-r\varphi} de^1 \wedge d\bar{e}^1 \wedge \dots \wedge de^{r-1} \wedge d\bar{e}^{r-1}$$

up to a constant.

Moreover, when (E, h) is a Griffiths-semipositive singular Hermitian vector bundle, the integral of the right hand side of (5) is still well-defined, and (5) holds almost everywhere.

Proof. The smooth version of the L^2 -representation (5) is nothing but a reformulation of [17, Theorem 7.1]. In fact, if we take $k = 1$ and $F = \mathcal{O}_X$, (3.3) in [17] defines a metric f on E via $\pi_*\mathcal{O}_E(1) \simeq E$, which is exactly

$$\frac{1}{(r-1)!} \int_{\mathbb{P}(E^*)_x} |F|^2 e^{-\varphi} \omega_x^{r-1}.$$

Then (7.3) there implies that $f = h$ up to multiplication by $\frac{(r+1)^{r-1}}{r!}$, which is the desired conclusion.

As for (6), it is enough to show that the associated curvature forms of both sides are equal. For this purpose, we apply (4). Keep the notation there. Without loss of generality, we can make $\{W^0, \dots, W^{r-1}\}$ normal for h^* at a fixed point x . Then at $(x; e^1, \dots, e^{r-1})$, ω_x is just the standard Fubini–Study metric on \mathbb{P}^{r-1} and

$$\Psi = i\Theta_{\alpha\bar{\beta}i\bar{j}} \frac{w^\alpha \bar{w}^\beta}{|w|^2} dx^i \wedge d\bar{x}^j.$$

Taking the integral against ω_x^{r-1} , we obtain

$$\begin{aligned} \pi_* \Psi &= i\Theta_{\alpha\bar{\beta}i\bar{j}} \left(\int_{\mathbb{P}(E^*)_x} \frac{w^\alpha \bar{w}^\beta}{|w|^2} \omega_x^{r-1} \right) dx^i \wedge d\bar{x}^j = i\Theta_{\alpha\bar{\beta}i\bar{j}} \frac{\delta_{\alpha\bar{\beta}}}{r} dx^i \wedge d\bar{x}^j \\ &= \frac{i}{r} \sum_{\alpha} \Theta_{\alpha\bar{\alpha}i\bar{j}} dx^i \wedge d\bar{x}^j = \frac{i}{r} \Theta_{\det E^*, \det h^*} \end{aligned}$$

at x , hence everywhere. On the other hand, remember that $\int_{\mathbb{P}(E^*)_x} \omega_x^{r-1} = 1$,

$$\int_{\mathbb{P}(E^*)/X} i\pi^* \Theta_{\det E^*, \det h^*} \wedge \omega_x^{r-1} = i\Theta_{\det E^*, \det h^*}.$$

Here $\int_{\mathbb{P}(E^*)/X} \cdot \wedge \omega_x^{r-1}$ refers to taking the integral along fibers. It exactly implies that there exists an element γ with

$$(7) \quad \Psi = \frac{i}{r} \pi^* \Theta_{\det E^*, \det h^*} + \gamma \quad \text{and} \quad \int_{\mathbb{P}(E^*)/X} \gamma \wedge \omega_x^{r-1} = 0.$$

Combining with the fact that ω_{FS} is the Fubini–Study metric along each fiber of π and the canonical isomorphism

$$K_{\mathbb{P}(E^*)/X}^{-1} \simeq \mathcal{O}_E(r) \otimes \pi^* \det E^*,$$

we conclude from (7) that

$$i\Theta_{\mathcal{O}_E(r), r\varphi} = ir\partial\bar{\partial}\varphi = -r\Psi + r\omega_{FS} = -i\pi^* \Theta_{\det E^*, \det h^*} + i\Theta_{K_{\mathbb{P}(E^*)/X}^{-1}, \omega_x^{r-1}}.$$

The proof of (6) is then finished. Observe that by setting

$$(G, h_G) := (\mathcal{O}_E(r+1) \otimes \pi^* \det E^*, h_L^{r+1} \otimes \pi^* \det h^*),$$

we can interpret (5) as

$$(8) \quad |F|_{h,x}^2 = \int_{\mathbb{P}(E^*)_x} |F|_{h_G}^2$$

in view of (6). The definition of $\int_{\mathbb{P}(E^*)_x} |F|_{h_G}^2$ is based on

$$\mathcal{O}_E(1) \simeq K_{\mathbb{P}(E^*)/X} \otimes \mathcal{O}_E(r+1) \otimes \pi^* \det E^*,$$

and is carefully explained in [20, Section 2.2].

Whereas the singular version of (5) is nothing but a reformulation of [20, Proposition 3.1]. Indeed, by [20] the integral of the right hand side of (8) is still well-defined when h is singular and (E, h) is Griffiths-semipositive. Moreover, (8) holds outside $V := \{\det h = \infty\}$. Note V must be a set of measure zero since h is finite almost everywhere, which concludes the last assertion. \square

We next restate and prove Theorem 1.2:

Theorem 3.1. *Let E be a holomorphic vector bundle of rank r over Δ^n , and let h be a singular Hermitian metric on E such that (E, h) is Griffiths-semipositive. Assume that $\det h$ has analytic singularities. If F is a holomorphic section of E satisfying*

$$\int_{\Delta^n} |F|_h^2 d\lambda < \infty,$$

then after shrinking Δ^n if necessary, there exists a positive ε (independent of F) such that

$$\int_{\Delta^n} |F|_h^{2(1+\varepsilon)} d\lambda < \infty.$$

Proof. Keep the notations of Proposition 3.1. Due to (the singular version of) (8) and Fubini’s theorem, it is enough to show that there exists a positive ε such that

$$\int_{\Delta^n} \left(\int_{\mathbb{P}(E^*)_x} |F|_{h_G}^2 \right)^{1+\varepsilon} d\lambda < \infty.$$

By hypothesis, $\pi^* \det h^* = e^\psi$, where ψ is a plurisubharmonic function with analytic singularities. So we have

$$\begin{aligned} \int_{\Delta^n} \left(\int_{\mathbb{P}(E^*)_x} |F|_{h_G}^2 \right)^{1+\varepsilon} d\lambda &= \int_{\Delta^n} (\det h^*)^{1+\varepsilon} \left(\int_{\mathbb{P}(E^*)_x} |F|^2 e^{-(r+1)\varphi} \right)^{1+\varepsilon} d\lambda \\ &\leq C \int_{\Delta^n} (\det h^*)^{1+\varepsilon} \left(\int_{\mathbb{P}(E^*)_x} (|F|^2 e^{-(r+1)\varphi})^{1+\varepsilon} \right) d\lambda \\ &= C \int_{\Delta^n} \int_{\mathbb{P}(E^*)_x} e^{(1+\varepsilon)(\log |F|^2 + \psi) - (1+\varepsilon)(r+1)\varphi} d\lambda. \end{aligned}$$

The inequality is a simple application of Hölder’s inequality, and C is a universal positive constant.

On the other hand, by assumption F a priori satisfies

$$\int_{\Delta^n} \int_{\mathbb{P}(E^*)_x} e^{\log |F|^2 + \psi - (r+1)\varphi} d\lambda < \infty.$$

So after shrinking Δ^n if necessary, by Corollary 1.1 there exists a positive ε_0 such that

$$(9) \quad \int_{\Delta^n} \int_{\mathbb{P}(E^*)_x} e^{\log |F|^2 + \psi - (1+\varepsilon_0)(r+1)\varphi} d\lambda < \infty.$$

Since $e^{\varepsilon_0(\log |F|^2 + \psi)}$ is bounded, (9) implies that

$$\int_{\Delta^n} \left(\int_{\mathbb{P}(E^*)_x} |F|_{h_G}^2 \right)^{1+\varepsilon_0} d\lambda < \infty.$$

In the end, notice that [23, Theorem 1.2] indicates that $\mathcal{E}(h)$ is coherent hence is locally finitely generated in our situation. So we can even obtain a uniform constant ε_0 for all holomorphic sections of $\mathcal{E}(h)$. The proof is complete. \square

4. Openness

Now we would like to remove the restriction in Theorem 3.1 that $\det h$ has analytic singularities. A natural idea is to approximate h by a sequence of smooth metrics $\{h_\nu\}$, whose existence is due to Proposition 2.1. Then we apply Theorem 3.1 on each h_ν (which gives a corresponding ε_ν), and take the limit to obtain the desired estimate. For this purpose, we need a universal lower bound for these ε_ν . The related topic is called the effectiveness of the *strong openness property*. It is much involved even in the line bundle case (see [10, Theorem 1.3]). As a compromise, we will instead apply the following effective version of the *openness property* in [4].

The line bundle case revisited.

Theorem 4.1 [4, Theorem 1.1]. *Let φ be a plurisubharmonic function in B^n with $\varphi \leq 0$. Assume that*

$$\int_{B^n} e^{-\varphi} d\lambda < \infty.$$

Then there is a number $\varepsilon > 0$ such that

$$\int_{B^n/2} e^{-(1+\varepsilon)\varphi} d\lambda < \infty.$$

Here B^n and $B^n/2$ refer to the n -dimensional balls of radii 1 and 1/2 respectively. Moreover, ε can be taken so that

$$\varepsilon \geq \frac{\delta_n}{\int_{B^n/2} e^{-\varphi} d\lambda},$$

where δ_n depends only on the dimension.

Let us briefly recall some crucial steps in the proof of Theorem 4.1 for readers' benefit. For any $s \geq 0$ and holomorphic function h , let

$$\varphi_s := \max(\varphi + s, 0)$$

and

$$\|h\|_s^2 := \int_{B^n/2} |h|^2 e^{-2\varphi_s} d\lambda.$$

The factor 2 in the exponent is necessary, and we define the norm $\|\cdot\|_s$ here over $B^n/2$ (rather than B^n in [4]) for our purposes. Then a simple variant of Proposition 2.1 in [4] indicates that for $0 < \varepsilon < 1$,

$$(10) \quad \int_{B^n/2} e^{-(1+\varepsilon)\varphi} d\lambda = a_\varepsilon \int_0^\infty e^{(1+\varepsilon)s} \|1\|_s^2 ds + b_\varepsilon \text{Vol}(B^n).$$

Here

$$a_\varepsilon = \frac{1 - \varepsilon^2}{2} \quad \text{and} \quad b_\varepsilon = \frac{1 - \varepsilon}{2^{n+1}}.$$

Pick a universal constant δ_n so that if w is holomorphic and

$$\int_{B^n} |w|^2 d\lambda \leq 4\delta_n,$$

then

$$\sup_{B^n/2} |w| \leq \frac{1}{10}.$$

The existence of such a δ_n is due to the mean value property of holomorphic functions. In particular, δ_n depends only on the dimension.

Now $\int_{B^n} e^{-\varphi} d\lambda < \infty$ by hypothesis,

$$\int_0^\infty e^s \left(\int_{B^n} e^{-2\varphi_s} \right) ds \leq \int_{B^n} e^{-\varphi} d\lambda < \infty$$

due to (the original version of) Proposition 2.1 in [4]. Then by Theorem 3.3 of [4], for any $\varepsilon > 0$ and $s > \frac{1}{2\varepsilon}$ there is a holomorphic function h_s such that

$$(11) \quad \int_{B^n} |1 - h_s|^2 d\lambda \leq 4\varepsilon \int_{B^n} e^{-\varphi} d\lambda,$$

and

$$(12) \quad \int_{B^n} |h_s|^2 e^{-2\varphi_s} d\lambda \leq e^{-(1+2\varepsilon)s} \text{Vol}(B^n).$$

Take

$$\varepsilon = \frac{\delta_n}{\int_{B^n} e^{-\varphi} d\lambda}.$$

Then (11) implies that the holomorphic function $1 - h_s$ satisfies

$$\int_{B^n} |1 - h_s|^2 d\lambda \leq 4\delta_n,$$

hence $\sup_{B^n/2} |1 - h_s| \leq \frac{1}{10}$. Therefore h_s satisfies

$$1 < 2|h_s|^2$$

on $B^n/2$. Multiply by $e^{-2\varphi_s}$ and integrate on $B^n/2$, we obtain from this inequality

and (12) that

$$(13) \quad \|1\|_s^2 \leq 2 \int_{B^{n/2}} |h_s|^2 e^{-2\varphi_s} d\lambda \leq 2 \int_{B^n} |h_s|^2 e^{-2\varphi_s} d\lambda \leq 2e^{-(1+2\varepsilon)s} \text{Vol}(B^n).$$

Now multiply the both sides of (13) by $e^{(1+\varepsilon)s}$ and integrate from 0 to infinity. Combining with (10) we get

$$\begin{aligned} (14) \quad & \int_{B^{n/2}} e^{-(1+\varepsilon)\varphi} d\lambda \\ & \leq a_\varepsilon \int_0^{\frac{1}{2\varepsilon}} e^{(1+\varepsilon)s} \|1\|_s^2 ds + a_\varepsilon \int_{\frac{1}{2\varepsilon}}^\infty e^{(1+\varepsilon)s} 2e^{-(1+2\varepsilon)s} \text{Vol}(B^n) ds + b_\varepsilon \text{Vol}(B^n) \\ & \leq a_\varepsilon \text{Vol}(B^{n/2}) \int_0^{\frac{1}{2\varepsilon}} e^{(1+\varepsilon)s} ds + a_\varepsilon \text{Vol}(B^n) \int_{\frac{1}{2\varepsilon}}^\infty e^{-\varepsilon s} ds + b_\varepsilon \text{Vol}(B^n) \\ & = \frac{a_\varepsilon}{2^n} \text{Vol}(B^n) \frac{e^{\frac{1+\varepsilon}{2\varepsilon}} - 1}{1 + \varepsilon} + a_\varepsilon \text{Vol}(B^n) \frac{1}{\varepsilon\sqrt{e}} + b_\varepsilon \text{Vol}(B^n) \\ & = \left(\frac{a_\varepsilon (e^{\frac{1+\varepsilon}{2\varepsilon}} - 1)}{2^n(1 + \varepsilon)} + \frac{a_\varepsilon}{\varepsilon\sqrt{e}} + b_\varepsilon \right) \text{Vol}(B^n) \\ & =: c_\varepsilon \text{Vol}(B^n). \end{aligned}$$

Since (13) is valid when $s > \frac{1}{2\varepsilon}$, we have to divide the integral into two parts. The integral from 0 to $\frac{1}{2\varepsilon}$ is standard, since $0 \leq \varphi_s \leq \frac{1}{2\varepsilon}$; the integral from $\frac{1}{2\varepsilon}$ to infinity is estimated via (13), and it leads to the second inequality. The rest is routine.

Proof of Theorem 1.3. Now we are ready to prove the openness for the higher-rank case. Let E be a holomorphic vector bundle of rank r over Δ_R^n , where Δ_R^n is a polydisc with radius R slightly larger than 1. Let (e^1, \dots, e^{r-1}) be a fiber coordinate system of $\mathbb{P}(E^*)$, and let $de \wedge d\bar{e}$ be short for

$$de^1 \wedge d\bar{e}^1 \wedge \dots \wedge de^{r-1} \wedge d\bar{e}^{r-1}.$$

Take a finite coordinate chart $\{B_j^{n+r-1}\}$ of $\mathbb{P}(E^*)$. For an arbitrary singular Hermitian metric h on E , denote by φ the induced metric on $\mathcal{O}_E(1)$, and

$$\int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda := \max_j \left\{ \int_{B_j^{n+r-1}} e^{-(r+1)\varphi} d\lambda \right\}$$

with the convention that

$$\int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda = +\infty \quad \text{if some} \quad \int_{B_j^{n+r-1}} e^{-(r+1)\varphi} d\lambda = +\infty.$$

We restate and prove Theorem 1.3.

Theorem 4.2. *Assume that (E, h) is Griffiths-semipositive, and*

$$\mathcal{I}((r + 1)\varphi) = \mathcal{O}_{\mathbb{P}(E^*)}.$$

Then there exist positive numbers \hat{R} and ε such that

$$\int_{\Delta_{\hat{R}}^n} |F|_h^{2(1+\varepsilon)} d\lambda < \infty$$

for any holomorphic section F of E . Moreover, ε can be taken so that

$$\varepsilon \geq \frac{\delta_{n+r-1}}{\int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda},$$

where δ_{n+r-1} depends only on the dimension.

Proof. Pick an $R' \in (1, R)$. By Proposition 2.1(1), there exists a sequence of smooth, Griffiths positive metrics $\{h_\nu\}$ over $\Delta_{R'}^n$, increasing pointwise to h . Let φ_ν be the metric on $\mathcal{O}_E(1)$ induced by h_ν , and let $\omega_{x,\nu} := (i\partial\bar{\partial}\varphi_\nu)|_{\mathbb{P}(E^*)_x}$. Consequently, φ_ν decreases to φ . Since both φ and φ_ν are plurisubharmonic, we can assume $\varphi, \varphi_\nu \leq 0$ without loss of generality.

Now by assumption we have

$$\int_{B^{n+r-1}} e^{-(r+1)\varphi_\nu} d\lambda \leq \int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda < \infty.$$

Then apply Theorem 4.1 ((14), more precisely), we obtain

$$(15) \quad \int_{B^{n+r-1}/2} e^{-(1+\varepsilon)(r+1)\varphi_\nu} d\lambda \leq c_\varepsilon \text{Vol}(B^{n+r-1}),$$

where ε can be taken so that

$$\varepsilon \geq \frac{\delta_{n+r-1}}{\int_{B^{n+r-1}} e^{-(r+1)\varphi_\nu} d\lambda}.$$

Since φ_ν decreases to φ ,

$$\frac{\delta_{n+r-1}}{\int_{B^{n+r-1}} e^{-(r+1)\varphi_\nu} d\lambda} \searrow \frac{\delta_{n+r-1}}{\int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda}.$$

Let

$$\varepsilon = \frac{\delta_{n+r-1}}{\int_{B^{n+r-1}} e^{-(r+1)\varphi} d\lambda},$$

and take the limit of (15) with respect to ν . We finally obtain

$$\int_{B^{n+r-1}/2} e^{-(1+\varepsilon)(r+1)\varphi} d\lambda \leq c_\varepsilon \text{Vol}(B^{n+r-1}) < \infty.$$

In the end, observe that we can take a positive number $\hat{R} < R'$ such that the total

space of $\mathbb{P}(E^*)|_{\Delta_{\hat{R}}^n}$ is finitely covered by $B_j^{n+r-1}/2$. It exactly implies that

$$\begin{aligned} \int_{\Delta_{\hat{R}}^n} |F|_{h_\nu}^{2(1+\varepsilon)} d\lambda &= \int_{\Delta_{\hat{R}}^n} \left(\int_{\mathbb{P}(E^*)_x} |F|^2 e^{-\varphi_\nu} \omega_{x,\nu}^{r-1} \right)^{1+\varepsilon} d\lambda \\ &\leq \int_{\Delta_{\hat{R}}^n} \left(\int_{\mathbb{P}(E^*)_x} (|F|^2 e^{-\varphi_\nu})^{1+\varepsilon} \omega_{x,\nu}^{r-1} \right) \left(\int_{\mathbb{P}(E^*)_x} \omega_{x,\nu}^{r-1} \right)^{\frac{\varepsilon}{1+\varepsilon}} d\lambda \\ &= \int_{\Delta_{\hat{R}}^n} \left(\int_{\mathbb{P}(E^*)_x} (|F|^2 e^{-\left(\frac{r}{1+\varepsilon}+1\right)\varphi_\nu})^{1+\varepsilon} \cdot \pi^* \det h_\nu^* \cdot i^{r-1} de \wedge d\bar{e} \right) d\lambda \\ &\leq \sum_j \int_{B_j^{n+r-1}/2} |F|^{2(1+\varepsilon)} e^{-(r+1+\varepsilon)\varphi_\nu} \cdot \pi^* \det h_\nu^* d\lambda \\ &\leq C \sum_j \int_{B_j^{n+r-1}/2} e^{-(r+1+\varepsilon)\varphi_\nu} d\lambda \\ &\leq C \sum_j \int_{B_j^{n+r-1}/2} e^{-(r+1+\varepsilon)\varphi} d\lambda < \infty \end{aligned}$$

for any F . In the forth inequality we abuse the notation that $d\lambda$ refers to the Lebesgue measure on both $\Delta_{\hat{R}}^n$ and $B_j^{n+r-1}/2$. The fifth inequality is due to the fact that $\{\det h_\nu^*\}$ is a decreasing sequence of smooth functions. So $|F|^{2(1+\varepsilon)} \pi^* \det h_\nu^*$ is bounded by some positive constant C only depends on F . As ν tends to zero, we obtain

$$\int_{\Delta_{\hat{R}}^n} |F|_h^{2(1+\varepsilon)} d\lambda < \infty$$

for any F . The proof is complete. □

Let $\{U_j\}$ be a finite coordinate covering of $\mathbb{P}(E^*)$, and let $\{\rho_j\}$ be the associated partition of unity. When $\mathcal{S}((r+1)\varphi) = \mathcal{O}_{\mathbb{P}(E^*)}$, for any holomorphic section F we have

$$\begin{aligned} \int_{\Delta^n} |F|_{h_\nu}^2 d\lambda &= \int_{\Delta^n} \left(\int_{\mathbb{P}(E^*)_x} |F|^2 e^{-\varphi_\nu} \omega_{x,\nu}^{r-1} \right) d\lambda \\ &= \int_{\Delta^n} \left(\int_{\mathbb{P}(E^*)_x} |F|^2 e^{-(r+1)\varphi_\nu} \cdot \pi^* \det h_\nu^* \cdot i^{r-1} de \wedge d\bar{e} \right) d\lambda \\ &= \sum_j \int_{U_j} \rho_j |F|^2 e^{-(r+1)\varphi_\nu} \cdot \pi^* \det h_\nu^* d\lambda \\ &\leq C \sum_j \int_{U_j} e^{-(r+1)\varphi_\nu} d\lambda \\ &\leq C \sum_j \int_{U_j} e^{-(r+1)\varphi} d\lambda < \infty. \end{aligned}$$

As ν tends to zero, we obtain

$$\int_{\Delta^n} |F|_h^2 d\lambda < \infty$$

for any F . Hence

$$(16) \quad \mathcal{E}(h) = E.$$

However, if one attempts to deduce $\mathcal{S}((r+1)\varphi) = \mathcal{O}_{\mathbb{P}(E^*)}$ from (16), the influence of the zero locus of $\det h^*$ cannot be ignored. Hence the converse is not clear.

Acknowledgement

The author thanks Prof. Jixiang Fu for discussion and continuing encouragement. Thanks are also due to the referees for pointing out [20], which helped improve this paper.

References

- [1] S. J. Bao and Q. A. Guan, “ L^2 extension and effectiveness of strong openness property”, *Acta Math. Sin. (Engl. Ser.)* **38**:11 (2022), 1949–1964. MR
- [2] S. J. Bao and Q. A. Guan, “ L^2 extension and effectiveness of L^p strong openness property”, *Acta Math. Sin. (Engl. Ser.)* **39**:5 (2023), 814–826. MR
- [3] R. J. Berman, S. Boucksom, and M. Jonsson, “A variational approach to the Yau-Tian-Donaldson conjecture”, *J. Amer. Math. Soc.* **34**:3 (2021), 605–652. MR
- [4] B. Berndtsson, “The openness conjecture for plurisubharmonic functions”, preprint, 2013. arXiv 1305.5781
- [5] B. Berndtsson and M. Păun, “Bergman kernels and the pseudoeffectivity of relative canonical bundles”, *Duke Math. J.* **145**:2 (2008), 341–378. MR
- [6] J.-P. Demailly, *Analytic methods in algebraic geometry*, Surveys of Modern Mathematics **1**, International Press, Somerville, MA, 2012. MR
- [7] H. Feng, K. Liu, and X. Wan, “Chern forms of holomorphic Finsler vector bundles and some applications”, *Internat. J. Math.* **27**:4 (2016), art. id. 1650030, 22 pp. MR
- [8] P. A. Griffiths, “Hermitian differential geometry, Chern classes, and positive vector bundles”, pp. 185–251 in *Global analysis: papers in honor of K. Kodaira*, Princeton Univ. Press, 1969. MR
- [9] Q. Guan, “A sharp effectiveness result of Demailly’s strong openness conjecture”, *Adv. Math.* **348** (2019), 51–80. MR
- [10] Q. Guan and X. Zhou, “Effectiveness of Demailly’s strong openness conjecture and related problems”, *Invent. Math.* **202**:2 (2015), 635–676. MR
- [11] Q. Guan and X. Zhou, “A proof of Demailly’s strong openness conjecture”, *Ann. of Math. (2)* **182**:2 (2015), 605–616. MR
- [12] Q. Guan and X. Zhou, “Strong openness of multiplier ideal sheaves and optimal L^2 extension”, *Sci. China Math.* **60**:6 (2017), 967–976. MR
- [13] Q. Guan, Z. Li, and X. Zhou, “Stability of multiplier ideal sheaves”, *Chinese Ann. Math. Ser. B* **43**:5 (2022), 819–832. MR
- [14] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR
- [15] S. Kobayashi, “Negative vector bundles and complex Finsler structures”, *Nagoya Math. J.* **57** (1975), 153–166. MR

- [16] L. Lempert, “Modules of square integrable holomorphic germs”, pp. 311–333 in *Analysis meets geometry*, Birkhäuser, 2017. MR
- [17] K. Liu, X. Sun, and X. Yang, “Positivity and vanishing theorems for ample vector bundles”, *J. Algebraic Geom.* **22**:2 (2013), 303–331. MR
- [18] Z. Liu, B. Xiao, H. Yang, and X. Zhou, “Multiplier submodule sheaves and a problem of Lempert”, preprint, 2021. arXiv 2111.13452
- [19] Z. Liu, H. Yang, and X. Zhou, “New properties of multiplier submodule sheaves”, *C. R. Math. Acad. Sci. Paris* **360** (2022), 1205–1212. MR
- [20] Y. Liu, Z. Liu, H. Yang, and X. Zhou, “A Le Potier-type isomorphism twisted with multiplier submodule sheaves”, *J. Math. Pures Appl.* (9) **202** (2025), art. id. 103760, 38 pp. MR
- [21] M. Păun and S. Takayama, “Positivity of twisted relative pluricanonical bundles and their direct images”, *J. Algebraic Geom.* **27**:2 (2018), 211–272. MR
- [22] H. Raufi, “Singular hermitian metrics on holomorphic vector bundles”, *Ark. Mat.* **53**:2 (2015), 359–382. MR
- [23] Y. Zou, “Note on the singular Hermitian metrics with analytic singularities”, preprint, 2022. arXiv 2209.05053

Received March 19, 2025. Revised June 26, 2025.

JINGCAO WU
SCHOOL OF MATHEMATICS
SHANGHAI UNIVERSITY OF FINANCE AND ECONOMICS
SHANGHAI
CHINA
wujincao@shufe.edu.cn

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to journals+pjm@msp.org or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

At submission time a PDF file is required, and it must be derived from a \LaTeX source file in which the command `\documentclass[12pt]{amsart}` is active. No changes to horizontal or vertical margins, or line spacing, should be made. Please include in the preamble the command `\usepackage[pagewise]{lineno}\linenumbers` — this will spare referees tedious effort in line counting.

Carefully preserve all related files, such as \LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of \BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to journals+pjm@msp.org.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 342 No. 2 June 2026

On multilinear maximal operators along homogeneous curves	207
LARS BECKER and BEN KRAUSE	
The plain sphere number of a link	217
RYAN BLAIR, ALEXANDRA KJUCHUKOVA and ELLA PFAFF	
On the test properties of the Frobenius endomorphism	235
OLGUR CELIKBAS, ARASH SADEGHI and YONGWEI YAO	
Symplectic automorphisms of a surface with genus two fibration and their action on CH_0	259
JIABIN DU and WENFEI LIU	
Non-braid-positive hyperbolic L -space knots	275
KEISUKE HIMENO	
The Rankin–Selberg integral on GSp_2 for square-free levels	299
SEIJI KUGA and MASAO TSUZUKI	
Sum of the squares of the p' -character degrees	351
NGUYEN N. HUNG, J. MIQUEL MARTÍNEZ and GABRIEL NAVARRO	
On the first eigenvalue of the Hodge Laplacian of submanifolds	381
CHRISTOS-RAENT ONTI	
Howe duality for the dual pair $\mathrm{SL}_2(\mathbb{R}) \times F_{4,1}$: a ping pong of K -types	387
GORDAN SAVIN	
Liouville theorems and new gradient estimates for positive solutions to $\Delta_p u + au^q = 0$ on a complete manifold	395
YOUDE WANG and LIQIN ZHANG	
p -nuclearity of reduced group L^p -operator algebras	427
ZHEN WANG	
A higher-rank analog of the strong openness property	441
JINGCAO WU	