msp

Geometry & Topology 22 (2018) 3501–3573

A formal Riemannian structure on conformal classes and uniqueness for the σ_2–Yamabe problem

MATTHEW GURSKY
JEFFREY STREETS

We define a new formal Riemannian metric on a conformal classes of four-manifolds in the context of the σ_2–Yamabe problem. Exploiting this new variational structure we show that solutions are unique unless the manifold is conformally equivalent to the round sphere.

58J05; 53C44, 58B20

1 Introduction

1.1 Background

In [20], we defined a formal Riemannian metric on the space of conformal metrics on surfaces of positive (or negative) Gauss curvature. Our goal in this paper is to show that one can extend this definition to conformal classes of metrics on four-manifolds, and to explore the geometric properties of this metric and their applications. The definition we give can be extended to higher (even) dimensions, but this will be pursued in a subsequent article since there are technical issues that do not arise in two or four dimensions; see Gursky and Streets [21].

In addition to verifying the formal properties of this metric we prove a remarkable geometric consequence: namely, solutions of the σ_2–Yamabe problem — whose existence follows from our positivity assumption and Chang, Gursky and Yang [6] — are *unique*, unless the manifold is conformally equivalent to the sphere. This is a surprising departure from the classical (or σ_1–) Yamabe problem, where explicit examples of nonuniqueness are known (see Remarks 1.6 and 1.7 below). Thus, positive conformal classes on four-manifolds have a unique conformal representative whose σ_2–curvature is constant; moreover, the value of this constant (after normalizing the volume) can be expressed in terms of the Euler characteristic and the L^2–norm of the Weyl tensor (see

Published: 23 September 2018 DOI: 10.2140/gt.2018.22.3501

the introduction of Chang, Gursky and Yang [7]). We also remark that this representative has positive Ricci curvature.

To give a more detailed description it will be helpful to return to the setting of surfaces. Let (M, g_0) be a compact Riemannian surface with positive Gauss curvature $K_0 > 0$, and let $[g_0]$ denote the conformal class of g_0. Define

$$\mathcal{C}^+ = \{g_u = e^{2u} g_0 \in [g_0] : K_u = K_{g_u} > 0\}. \tag{1-1}$$

Formally, the tangent space to $[g_0]$ at any metric $g_u \in [g_0]$ is given by $C^\infty(M)$. For $\phi, \psi \in C^\infty(M) \cong T_u([g_0])$ we define

$$\langle\langle \phi, \psi \rangle\rangle_u = \int_M \phi \psi K_u \, dA_u, \tag{1-2}$$

where K_u is the Gauss curvature and dA_u is the area form of g_u.

The definition (1-2) is inspired by the Mabuchi–Semmes–Donaldson metric [27; 32; 12; 13] of Kähler geometry, wherein a formal Riemann metric is put on a Kähler class by imposing on the tangent space to a given Kähler potential the L^2 metric with respect to the associated Kähler metric. As observed in [27], this metric enjoys many nice formal properties, for instance nonpositive sectional curvature. Moreover, it has a profound relationship to natural functionals in Kähler geometry such as the Mabuchi K–energy and the Calabi energy, as well as their gradient flow, the Calabi flow.

In [20] we established a number of analogous properties for the metric defined by (1-2). For example, $\mathcal{C}^+$ endowed with the metric in (1-5) has nonpositive curvature in the sense of Alexandrov. We also showed that the normalized Liouville energy $F \colon W^{1,2} \to \mathbb{R}$, defined by

$$F[u] = \int_M |\nabla_0 u|^2 \, dA_0 + 2 \int_M K_0 u \, dA_0 - \left(\int_M K_0 \, dA_0 \right) \log \left(\fint_M e^{2u} \, dA_0 \right), \tag{1-3}$$

is *geodesically convex*. Recall that critical points of F, which are precisely the conformal metrics of constant Gauss curvature, are minimizers and unique up to Möbius transformation. Many of these global geometric properties are based on existence and partial regularity results for geodesics in $\mathcal{C}^+$ (see Section 4 of [20] for precise statements).

In this paper we study a natural generalization of the inner product (1-5). For an n–dimensional Riemannian manifold ($n \geq 3$), we denote the Schouten tensor by

$$A = \frac{1}{n-2} \left(\mathrm{Ric} - \frac{1}{2(n-1)} R g \right),$$

where Ric is the Ricci tensor and R is the scalar curvature. Let $\sigma_k(g^{-1}A)$ denote the k^{th} symmetric function of the eigenvalues of the $(1,1)$–tensor obtained by raising an index of A, ie

$$A_i^j = g^{jk}A_{ik}.$$

The quantity $\sigma_k(g^{-1}A)$ is called the *σ_k–curvature* or the *k–scalar curvature*. For example,

$$\sigma_1(g^{-1}A) = \frac{R}{2(n-1)}. \tag{1-4}$$

For $1 \le k \le n$, we write $A = A_g \in \Gamma_k^+$ if $\sigma_j(g^{-1}A) > 0$ on M^n for all $1 \le j \le k$. By (1-4), we have $A_g \in \Gamma_1^+$ if g has positive scalar curvature, while $A_g \in \Gamma_n^+$ if the Schouten tensor of g is positive definite.

We will be interested in the case where $n = 4$ and $k = 2$. To this end, let (M^4, g_0) be a compact Riemannian four-manifold such that $A_{g_0} \in \Gamma_2^+$. Given $u \in C^\infty(M)$, let A_u denote the Schouten tensor of the conformal metric $g_u = e^{-2u}g_0$. We will say that u is *admissible* if $A_u \in \Gamma_2^+$. Let

$$\mathcal{C}^+ = \mathcal{C}^+([g_0]) = \{g_u \in [g_0] \mid A_u \in \Gamma_2^+\}.$$

By a result of Guan, Viaclovsky and Wang [18], if $g_u \in \mathcal{C}^+$ then g_u has positive Ricci curvature. As noted above, the tangent space to $\mathcal{C}^+$ at any point is given by $C^\infty(M)$. Thus, in analogy with (1-5) we define, for $\phi, \psi \in C^\infty(M)$,

$$\langle \phi, \psi \rangle_u = \int_M \phi\psi\sigma_2(g_u^{-1}A_u)\,dV_u. \tag{1-5}$$

Remark 1.1 To simplify the notation we will write $\sigma_2(A)$ instead of $\sigma_2(g^{-1}A)$. Since we will be working with conformal metrics, we will also need to distinguish between $g^{-1}A_u$ and $g_u^{-1}A_u$, ie whether we are using g or g_u to raise an index. Therefore, we will adopt the usual convention that $\sigma_2(A_u) = \sigma_2(g^{-1}A_u)$, but write $\sigma_2(g_u^{-1}A_u)$ when we are using g_u to raise an index. Note that

$$\sigma_2(g_u^{-1}A_u) = e^{4u}\sigma_2(A_u). \tag{1-6}$$

In particular,

$$\sigma_2(g_u^{-1}A_u)\,dV_u = \sigma_2(A_u)\,dV.$$

Remark 1.2 There is a sharp characterization of conformal classes for which $\mathcal{C}^+$ is nonempty. In view of the conformal invariance of the integral

$$\sigma := \int \sigma_2(g^{-1}A_g)\, dV_g,$$

a necessary condition for $[g]$ to admit a metric $g_u \in [g]$ with $A_u \in \Gamma_2^+$ is the positivity of the Yamabe invariant and the positivity of σ. In Chang, Gursky and Yang [7] these conditions were shown to be sufficient. Thus we have an exact parallel with the case of two dimensions, since a conformal class of metrics on a surface admits a metric of positive Gauss curvature if and only if the total Gauss curvature is positive.

1.2 Formal metric properties

We begin by establishing in Section 3 some fundamental formal properties of the metric defined in (1-5). We first introduce a formal path derivative which can be regarded as the Levi-Civita connection associated to the metric. Using this we compute the curvature tensor, and furthermore show that the curvature is nonpositive:

Theorem 1.3 *Given (M^4, g) a compact Riemannian manifold, with $A_g \in \Gamma_2^+$. Then* (1-5) *defines a metric with nonpositive sectional curvature on $\mathcal{C}^+$.*

Next, we derive the geodesic equation. Formal calculations derived using either the path derivative or variations of the length functional yield that a one-parameter family of conformal factors is a geodesic if and only if

$$u_{tt} - \frac{1}{\sigma_2(A_u)}\langle T_1(A_u), \nabla u_t \otimes \nabla u_t\rangle = 0, \tag{1-7}$$

where T_1 is the Newton transform and $\langle \cdot, \cdot \rangle$ denotes the inner product on tensor bundles induced by g (the background metric). This is a degenerate fully nonlinear equation, which is related to a σ_2–type problem for the spacetime Hessian of u, in direct analogy to the $(n+1)$–dimensional degenerate Monge–Ampère interpretation of the Mabuchi geodesic equation in Kähler geometry. We also show that one parameter families of conformal transformations are automatically geodesics (Proposition 3.12). This is again in analogy with the fact that one-parameter families of biholomorphisms generate families of Kähler potentials which are Mabuchi geodesics.

In the Kähler setting, the Mabuchi metric and its geodesics are intimately related to Mabuchi's K–energy functional. This is a "relative functional" defined via path integration of a closed 1–form on a Kähler class. It was shown in Mabuchi [26; 27] that this functional is geodesically convex, leading to the conjecture that extremal Kähler metrics are unique up to biholomorphism in a fixed Kähler class. Confirming this conjecture requires extensive existence and regularity results for the geodesic equation. An initial theory of $C^{1,1}$ was developed in Błocki [1], Calabi and Chen [5] and Chen [9], and eventually a more refined regularity theory was developed and the conjecture finally confirmed in Chen and Tian [10].

In our setting there is a natural analogue of Mabuchi's functional. For surfaces it is given by the Liouville energy, or regularized determinant (1-3). In four dimensions this functional was written down by Chang and Yang [8] (although it appears implicitly in Chang, Gursky and Yang [7]):

$$F[u] = \int \{2\Delta u|\nabla u|^2 - |\nabla u|^4 - 2\,\mathrm{Ric}(\nabla u,\nabla u) + R|\nabla u|^2 - 8u\sigma_2(A_g)\}\,dV \\ - 2\Big(\int \sigma_2(A_g)\,dV\Big)\log\Big(⨍ e^{-4u}\,dV\Big). \tag{1-8}$$

After this, Brendle and Viaclovsky [4] give a path-integration derivation of this functional which makes clearer the analogy between it and the Mabuchi functional in Kähler geometry. We will not need the precise formula, only the fact that it provides a conformal primitive for $\sigma_2(A)$; ie if u_s is a path with $du_s/ds|_{s=0} = u'$, then

$$\frac{d}{ds}F[u_s]\Big|_{s=0} = \int u'[-\sigma_2(g_u^{-1}A_u) + \overline{\sigma}]\,dV_u. \tag{1-9}$$

Consequently, u is a critical point of F if and only if $g_u = e^{-2u}g$ is a solution of the σ_2–Yamabe problem:

$$\sigma_2(g_u^{-1}A_u) \equiv \text{const.} \tag{1-10}$$

In four dimensions the existence of solutions to (1-10) in conformal classes with $\mathcal{C}^+ \neq \varnothing$ was first proved by Chang, Gursky and Yang [6] (for surveys on solving the σ_k–Yamabe problem for general $2 \le k \le n$ see Viaclovsky [37] and Sheng, Trudinger and Wang [33]). In particular, if $\mathcal{C}^+([g])$ is nonempty, then $[g]$ always admits a critical point of F. Our next result gives us deeper insight into the variational structure of F:

Theorem 1.4 *The functional F in* (1-8) *is geodesically convex.*

The proof of this theorem requires the use of a sharp curvature-weighted Poincaré inequality due to Andrews (unpublished). In fact, it follows from Andrews' inequality that F is strictly convex, up to one-parameter families of conformal automorphisms on the round sphere. This sharp characterization naturally leads one to conjecture that critical points of F are unique, except in the case of the sphere. We are able to confirm this surprising fact:

Theorem 1.5 *Let (M^4, g) be a compact Riemannian manifold such that $\mathcal{C}^+([g]) \neq \varnothing$.*

(1) *If (M^4, g) is not conformal to (S^4, g_{S^4}), then there exists a unique solution to the σ_2–Yamabe problem in $[g]$.*

(2) *In $[g_{S^4}]$, all solutions to the σ_2–Yamabe problem are round metrics.*

Remark 1.6 This uniqueness property is in stark contrast to the Yamabe problem, in which generic conformal classes admit arbitrarily many distinct solutions (see Pollack [29]). In dimensions $n \geq 25$ the solution space may even be noncompact; see Brendle and Marques [2; 3].

Remark 1.7 Explicit examples of nonuniqueness for the Yamabe problem were constructed by Schoen [31], in which he constructed Delaunay-type solutions on $S^{n-1} \times S^1$. By lifting to the universal cover $S^{n-1} \times \mathbb{R}$ and imposing symmetry, he reduced the Yamabe equation to an ODE and studied the phase portrait. Interestingly, Viaclovsky [35] carried out a similar construction for solutions of the σ_k–Yamabe problem when $k < \frac{n}{2}$. However, once $k \geq \frac{n}{2}$ the construction fails, since the admissibility condition implies the Ricci curvature of any solution would have to be positive, and $S^{n-1} \times S^1$ does not admit a metric with positive Ricci curvature.

The proof of Theorem 1.5 consists of two main phases. First we develop a weak existence/regularity theory for the geodesic equation (1-7). In general for degenerate Monge–Ampère equations one typically expects at best $C^{1,1}$ control, and indeed this is verified in the Kähler setting by Chen [9] (with compliments due to Błocki [1]). Where Mabuchi geodesics can be interpreted as solutions of a degenerate complex Monge–Ampère equation, our geodesics are solutions to a degenerate σ_2–equation (Proposition 4.1), and so one at best again expects $C^{1,1}$ regularity. However, due to some technical issues arising from the presence of first-order terms in the Schouten tensor, we are not able to establish such estimates. Rather, we are forced to regularize the equation by rendering the right-hand side positive (which is a standard trick), but also perturbing the coefficients on the time direction term, to further break the nondegeneracy. This leads to full C^∞ regularity, but only the C^1 estimates persist as the regularization parameters go to zero.

Given this, one cannot directly rigorously establish properties of F related to the geodesic convexity.[1] Nonetheless we are able to improve the regularity of an approximate geodesic connecting any two solutions to the σ_2–problem by smoothing via the parabolic flow introduced by Guan and Wang [19]. In particular we are able to take a sequence of approximate geodesics connecting two critical points for F, smooth them for a short time with this flow, and then show that this process yields a path of critical points for F, although not necessarily a geodesic. Combining this with arguments using the geodesic convexity shows that the existence of this path implies that the critical points are all round metrics on S^4, finishing the proof.

1.3 Outline

In Section 2 we establish notation and record some basic properties of the Schouten tensor and of elementary symmetric polynomials. Next, in Section 3 we establish the basic properties of the σ_2 metric defined in (1-5). In particular we prove Theorem 1.3 and establish the geodesic convexity of the functional F. Then, in Section 4 we develop estimates for approximate solutions to the geodesic equation, leading to a weak existence theory. In Section 5 we show a short-time smoothing result, which we will use to improve the regularity of approximate geodesics connecting any two critical points of the F–functional. We combine these two main technical tools in Section 6 to establish Theorem 1.5.

2 Background

In this section we establish our notation and some basic formulas. Although we are primarily interested in four dimensions, we will state most of the standard results for symmetric functions we will need for general n and k.

2.1 The Schouten tensor

Given a Riemannian manifold (M^n, g) let A denote the Schouten tensor of g. Given a conformal metric $g_u = e^{-2u} g$, the tensor A transforms according to

$$A_u = A + \nabla^2 u + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g. \tag{2-1}$$

[1] Recently Weiyong He [22] has established the existence of $C^{1,1}$ geodesics, leading to a more direct proof of the uniqueness statement. This work also corrects a technical problem with an earlier version of this paper (compare Theorem 4.18).

Let $g_u = e^{-2u(t)} g$ be a one-parameter family of conformal metrics. Then using formula (2-1) it follows that

$$\frac{\partial}{\partial t}(g_u^{-1}A_u)_i^j = 2\Big(\frac{\partial u}{\partial t}\Big)(g_u^{-1}A_u)_i^j + \Big(\nabla_u^2 \frac{\partial u}{\partial t}\Big)_i^j, \tag{2-2}$$

where the Hessian is with respect to g_u. A direct calculation [30] yields

$$\frac{\partial}{\partial t}\sigma_k(g_u^{-1}A_u) = \Big\langle T_{k-1}(g_u^{-1}A_u), \nabla_u^2 \frac{\partial u}{\partial t}\Big\rangle_{g_u} + 2k\frac{\partial u}{\partial t}\sigma_k(g_u^{-1}A_u), \tag{2-3}$$

where T_{k-1} is the Newton transform. Since the Newton transform is a $(1,1)$–tensor, for the pairing in (2-3) we *lower* an index of $T_{k-1}(g_u^{-1}A_u)$ and view it as a $(0,2)$–tensor, and use the inner product induced by g_u. For example, if $n=4$ and $k=2$,

$$T_1(g_u A_u) = -A_u + \sigma_1(g_u^{-1}A_u) g_u. \tag{2-4}$$

Combining (2-3) with the variation of the volume form yields

$$\begin{aligned} \frac{\partial}{\partial t}&[\sigma_k(g_u^{-1}A_u)\, dV_u] \\ &= \Big\langle T_{k-1}(g_u^{-1}A_u), \nabla_u^2 \frac{\partial u}{\partial t}\Big\rangle_{g_u} dV_u + (n-2k)\frac{\partial u}{\partial t}\sigma_k(g_u^{-1}A_u)\, dV_u. \end{aligned} \tag{2-5}$$

A key property we will use throughout is the following:

Lemma 2.1 *If $k=2$ or if the manifold is locally conformally flat, then $T_{k-1}(g^{-1}A)$ is divergence-free.*

Remark 2.2 This was proved in [34]. The essential idea also appears in [30], where the Schouten tensor is replaced with the second fundamental form of a hypersurface of a space of constant curvature. In both cases one needs that the tensor is Codazzi, ie

$$\nabla_k A_{ij} = \nabla_j A_{ik}.$$

Note that the conformal invariance of the integral

$$\sigma = \int_M \sigma_2(g_u^{-1}A_u)\, dV_u$$

follows from the variational formula (2-5) and Lemma 2.1. We denote the average value by

$$\overline{\sigma} = \sigma V_u^{-1}. \tag{2-6}$$

2.2 Properties of elementary symmetric polynomials

We record some lemmas concerning elementary symmetric polynomials and Newton transforms. To begin we record basic facts which are well known from Garding's theory of hyperbolic polynomials [15]. We use these to derive some further properties of generalized Newton transforms required for our estimates of the geodesic equation. First, given $A \in \Gamma_k^+$ we let $\sigma_k(A)$ denote the k^{th} elementary polynomial in the eigenvalues of A. Moreover, given $A_1, \dots, A_k$ we define the generalized Newton transformation by

$$[T_k]_{ij}(A_1, \dots, A_k) := \frac{1}{k!}\delta^{i,i_1,\dots,i_k}_{j,j_1,\dots,j_k}(A_1)_{i_1 j_1} \cdots (A_k)_{i_k j_k},$$

where here δ denotes the generalized Kronecker delta function. Moreover, we set

$$\Sigma_k(A_1, \dots, A_k) = \frac{1}{(k-1)!}\delta^{i_1,\dots,i_k}_{j_1,\dots,j_k}(A_1)_{i_1 j_1} \cdots (A_k)_{i_k j_k}.$$

Lemma 2.3 *One has:*

(1) *If $A_1, \dots, A_k \in \Gamma_k^+$, then $[T_k]_{ij}(A_1, \dots, A_k) > 0$.*

(2) *If $A_1, \dots, A_k \in \Gamma_k^+$, then $\Sigma(A_1, \dots, A_k) > 0$.*

(3) *If $A - B \in \Gamma_k^+$ and $A_2, \dots, A_k \in \Gamma_k^+$, then*

$$\Sigma(B, A_2, \dots, A_k) < \Sigma(A, A_2, \dots, A_k).$$

Lemma 2.4 *Given $A, B \in \Gamma_k^+$ $A < B$, one has $T_{k-1}(A) < T_{k-1}(B)$.*

Proof From Lemma 2.3, for $A_i \in \Gamma_k$ one has $T_k(A_1, \dots, A_k) > 0$. Now consider $M_t = A + t(B - A)$. Since $B - A$ is positive definite, certainly it lies in Γ_k^+. It follows that

$$\frac{d}{dt}T_k(M_t) = \frac{d}{dt}[T_k](M_t, \dots, M_t) = \sum_{j=1}^{k}[T_k](M_t, \dots, B - A, \dots M_t) \geq 0.$$

The result follows. □

Lemma 2.5 *Given A a symmetric matrix and X a vector, one has for $k \geq 1$,*

$$\langle T_k(A - X \otimes X), X \otimes X\rangle = \langle T_k(A), X \otimes X\rangle,$$
$$\sigma_k(A - X \otimes X) = \sigma_k(A) - \langle T_{k-1}(A), X \otimes X\rangle.$$

Proof If we express the matrix $B_t = A - tX \otimes X$ in a basis where X is the first basis vector, it is clear that the function

$$f(t) = \sigma_k(B_t)$$

is a linear function of t. It follows that its time derivative is constant, hence

$$C = f'(t) = -\langle T_{k-1}(A - tX \otimes X), X \otimes X \rangle.$$

Hence,

$$\langle T_{k-1}(A), X \otimes X \rangle = -f'(0) = -f'(1) = \langle T_{k-1}(A - X \otimes X), X \otimes X \rangle.$$

Moreover, this shows that

$$\sigma_k(A - X \otimes X) = f(1) = f(0) + \int_0^1 f'(s)\,ds = \sigma_k(A) - \langle T_{k-1}(A), X \otimes X \rangle. \quad \square$$

Lemma 2.6 *Given $A, B \in \mathrm{Sym}^2(\mathbb{R}^4)$ with $A, B \in \Gamma_2^+$ one has*

$$\langle T_1(B), A \rangle^2 \geq 4\sigma_2(A)\sigma_2(B).$$

Proof We compute that

$$\begin{aligned}
\frac{\sigma_1(A)}{\sigma_1(B)}\langle T_1(B), A \rangle &= -\frac{\sigma_1(A)}{\sigma_1(B)}\langle B, A \rangle + \sigma_1(A)^2 \\
&\geq -\frac{1}{2}\left[\frac{\sigma_1(A)}{\sigma_1(B)}\right]^2 |B|^2 - \frac{1}{2}|A|^2 + [\sigma_1(A)]^2 \\
&= -\frac{1}{2}\sigma_1(A)^2\left[\frac{|B|^2 - \sigma_1(B)^2 + \sigma_1(B)^2}{\sigma_1(B)^2}\right] + \sigma_2(A) + \frac{1}{2}\sigma_1(A)^2 \\
&= \frac{\sigma_1(A)^2}{\sigma_1(B)^2}\sigma_2(B) + \sigma_2(A).
\end{aligned}$$

Rearranging this and applying Cauchy–Schwarz yields

$$\sigma_2(A) \leq \frac{\sigma_1(A)}{\sigma_1(B)}\langle T_1(B), A \rangle - \frac{\sigma_1(A)^2}{\sigma_1(B)^2}\sigma_2(B) \leq \frac{1}{4\sigma_2(B)}\langle T_1(B), A \rangle^2,$$

as required. $\square$

3 The σ_2 metric

In this section we define the σ_2 metric and establish fundamental properties of this metric concerning connections, torsion, curvature and distance. We end by showing the crucial geodesic convexity property of the functional F of Chang and Yang.

3.1 Metric, connection and curvature

As in the introduction, let

$$\mathcal{C}^+ = \mathcal{C}^+([g]) = \{g_u = e^{-2u} g : A_u \in \Gamma_2^+\}.$$

Definition 3.1 Let (M^4, g) be a compact Riemannian four-manifold. The σ_k *metric* is the formal Riemannian metric defined for $g_u \in \mathcal{C}^+([g]) = \mathcal{C}^+$ and $\alpha, \beta \in T_u\mathcal{C}^+ \cong C^\infty(M)$ via

$$\langle \alpha, \beta \rangle_u = \frac{1}{\sigma} \int_M \alpha\beta\sigma_2(g_u^{-1} A_u)\, dV_u.$$

Moreover, given a path u_t in $\mathcal{C}^+$ and a one-parameter family α_t of tangent vectors with $\alpha_t \in T_{u_t}\mathcal{C}^+$, we define the directional derivative along the path u_t by

$$\begin{aligned} \frac{D}{\partial t}\alpha &:= \alpha_t - \sigma_2(g_u^{-1}A_u)^{-1} \langle T_1(g_u^{-1}A_u), \nabla\alpha \otimes \nabla u_t \rangle_{g_u} \\ &= \alpha_t - \sigma_2(A_u)^{-1} \langle T_1(A_u), \nabla\alpha \otimes \nabla u_t \rangle, \end{aligned} \tag{3-1}$$

where we have used (1-6), and the convention that $T_1(g^{-1}A_u) = T_1(A_u)$.

Lemma 3.2 *The connection defined by* (3-1) *is metric-compatible and torsion-free.*

Proof First we check metric compatibility. We compute, using (2-5) and Lemma 2.1,

$$\begin{aligned} \frac{d}{dt}\langle \alpha_t, \beta_t \rangle_{u_t} &= \frac{d}{dt}\int_M \alpha\beta\sigma_2(g_u^{-1}A_u)\, dV_u \\ &= \langle \dot\alpha, \beta \rangle + \langle \alpha, \dot\beta \rangle + \int_M \alpha\beta \Big\langle T_1(g_u^{-1}A_u), \nabla_u^2 \frac{\partial u}{\partial t} \Big\rangle\, dV_u \\ &= \langle \dot\alpha, \beta \rangle + \langle \alpha, \dot\beta \rangle - \int_M \Big\langle T_1(g_u^{-1}A_u), (\alpha\nabla\beta + \beta\nabla\alpha) \otimes \nabla_u \frac{\partial u}{\partial t} \Big\rangle\, dV_u \\ &= \Big\langle \frac{D}{\partial t}\alpha, \beta \Big\rangle + \Big\langle \alpha, \frac{D}{\partial t}\beta \Big\rangle. \end{aligned}$$

Next, to compute the torsion, let $u_{s,t}$ be a two parameter family of conformal factors. Then

$$\begin{aligned}\frac{D}{\partial s}\frac{\partial u}{\partial t}-\frac{D}{\partial t}\frac{\partial u}{\partial s}&=\frac{\partial^2 u}{\partial s\partial t}-\sigma_2(g_u^{-1}A_u)^{-1}\Big\langle T_1(g_u^{-1}A_u),\nabla\frac{\partial u}{\partial s}\otimes\nabla\frac{\partial u}{\partial t}\Big\rangle_u\\&\qquad-\frac{\partial^2 u}{\partial s\partial t}+\sigma_2(g_u^{-1}A_u)^{-1}\Big\langle T_1(g_u^{-1}A_u),\nabla\frac{\partial u}{\partial t}\otimes\nabla\frac{\partial u}{\partial s}\Big\rangle_u\\&=0.\end{aligned}$$

The lemma follows. □

Next we compute the sectional curvature, and conclude that it is nonpositive. We first record an integral identity in Lemma 3.3 and a certain general quadratic inequality in Lemma 3.4. We then obtain the curvature inequality by exploiting these identities.

Lemma 3.3 *If $\phi,\psi\in C^\infty(M)$, then*

$$\begin{aligned}\int\{\nabla^2\phi(\nabla\psi,\nabla\psi)-\Delta\phi|\nabla\psi|^2-\nabla^2\psi(\nabla\psi,\nabla\phi)+\Delta\psi\langle\nabla\psi,\nabla\phi\rangle\}\phi\,dV\\=\int\{-|\langle\nabla\phi,\nabla\psi\rangle|^2+|\nabla\phi|^2|\nabla\psi|^2\}\,dV.\end{aligned}$$

Proof Consider the vector field

$$X_i=\langle\nabla\phi,\nabla\psi\rangle\nabla_i\psi-|\nabla\psi|^2\nabla_i\phi.$$

Taking the divergence gives

$$\begin{aligned}\delta X&=\nabla_i X_i\\&=\nabla^2\phi(\nabla\psi,\nabla\psi)+\nabla^2\psi(\nabla\phi,\nabla\psi)+\Delta\psi\langle\nabla\phi,\nabla\psi\rangle-2\nabla^2\psi(\nabla\psi,\nabla\phi)-\Delta\phi|\nabla\psi|^2\\&=\nabla^2\phi(\nabla\psi,\nabla\psi)-\Delta\phi|\nabla\psi|^2-\nabla^2\psi(\nabla\psi,\nabla\phi)+\Delta\psi\langle\nabla\psi,\nabla\phi\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}I&\equiv\int\{\nabla^2\phi(\nabla\psi,\nabla\psi)-\Delta\phi|\nabla\psi|^2-\nabla^2\psi(\nabla\psi,\nabla\phi)+\Delta\psi\langle\nabla\psi,\nabla\phi\rangle\}\phi\,dV\\&=\int(\delta X)\phi\,dV.\end{aligned}$$

On the other hand, integrating by parts gives

$$I=\int(\delta X)\phi\,dV=-\int\langle X,\nabla\phi\rangle\,dV=\int\{-|\langle\nabla\phi,\nabla\psi\rangle|^2+|\nabla\phi|^2|\nabla\psi|^2\}\,dV,$$

as claimed. □

Lemma 3.4 *Let $T_1 = T_1(A)$ denote the first Newton transformation of the symmetric linear map $A\colon V \to V$, where V is a real inner product space of dimension four. Assume $A \in \Gamma_2^+$. Then, for all $X, Y \in V$,*

$$-T_1(X,X)T_1(Y,Y) + T_1(X,Y)^2 + \sigma_2(A)[|X|^2|Y|^2 - \langle X,Y\rangle^2] \le 0.$$

Proof Choose an orthonormal basis for V which diagonalizes T_1, and let $\{\lambda_1,\dots,\lambda_4\}$ denote the eigenvalues of T_1. Note by our assumption on A we know that $\lambda_i \ge 0$ for each i. With respect to this orthonormal basis, write $X = (x_1,\dots,x_4)$ and $Y = (y_1,\dots,y_4)$. Then, expanding and collecting terms, we get

$$\begin{aligned}
-T_1&(X,X)T_1(Y,Y) + T_1(X,Y)^2 \\
&= -\{\lambda_1 x_1^2 + \dots + \lambda_4 x_4^2\}\{\lambda_1 y_1^2 + \dots + \lambda_4 y_4^2\} + \{\lambda_1 x_1 y_1 + \dots + \lambda_4 x_4 y_4\}^2 \\
&= -\lambda_1\lambda_2(x_1^2 y_2^2 + x_2^2 y_1^2 - 2x_1x_2y_1y_2) - \lambda_1\lambda_3(x_1^2 y_3^2 + x_3^2 y_1^2 - 2x_1x_3y_1y_3) \\
&\qquad\qquad - \dots - \lambda_3\lambda_4(x_3^2 y_4^2 + x_4^2 y_3^2 - 2x_3x_4y_3y_4).
\end{aligned}$$

Next, let

$$Z = X \wedge Y,$$

whose components are

$$z_{ij} = x_i y_j - x_j y_i.$$

In terms of Z, we can rewrite the above as

$$-T_1(X,X)T_1(Y,Y) + T_1(X,Y)^2 = -\lambda_1\lambda_2 z_{12}^2 - \lambda_1\lambda_3 z_{13}^2 - \dots - \lambda_3\lambda_4 z_{34}^2.$$

At the same time,

$$|X|^2|Y|^2 - \langle X,Y\rangle^2 = \tfrac{1}{2}|Z|^2 = z_{12}^2 + z_{13}^2 + \dots + z_{34}^2.$$

Therefore,

$$\begin{aligned}
&-T_1(X,X)T_1(Y,Y) + T_1(X,Y)^2 + \sigma_2(A)[|X|^2|Y|^2 - \langle X,Y\rangle^2] \\
&\qquad = -\lambda_1\lambda_2 z_{12}^2 - \lambda_1\lambda_3 z_{13}^2 - \dots - \lambda_3\lambda_4 z_{34}^2 + \sigma_2(A)[z_{12}^2 + z_{13}^2 + \dots + z_{34}^2].
\end{aligned} \tag{3-2}$$

We need to express $\sigma_2(A)$ in terms of the eigenvalues of T_1. Since

$$T_1 = -A + \sigma_1(A)\cdot I, \tag{3-3}$$

taking the trace it follows that

$$\lambda_1 + \dots + \lambda_4 = 3\sigma_1(A).$$

Also, taking the norm-squared in (3-3),

$$|T_1|^2 = |A|^2 + 2\sigma_1(A)^2.$$

Therefore,

$$\sigma_2(A) = \tfrac{1}{3}(-\lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_3\lambda_4).$$

Substituting this into (3-2),

$$\begin{aligned}(3\text{-}4)\quad & -T_1(X,X)T_1(Y,Y) + T_1(X,Y)^2 + \sigma_2(A)[|X|^2|Y|^2 - \langle X,Y\rangle^2] \\ &= -\lambda_1\lambda_2 z_{12}^2 - \lambda_1\lambda_3 z_{13}^2 - \cdots - \lambda_3\lambda_4 z_{34}^2 \\ &\qquad + \tfrac{1}{3}(-\lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_3\lambda_4)[z_{12}^2 + z_{13}^2 + \cdots + z_{34}^2] \\ &= \tfrac{1}{3}(-\lambda_1^2 - \cdots - \lambda_4^2 - 2\lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_3\lambda_4) z_{12}^2 \\ &\qquad + \tfrac{1}{3}(-\lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1\lambda_2 - 2\lambda_1\lambda_3 + \lambda_1\lambda_4 + \cdots + \lambda_3\lambda_4) z_{13}^2 \\ &\qquad\qquad + \cdots + \tfrac{1}{3}(-\lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1\lambda_2 + \cdots + \lambda_2\lambda_4 - 2\lambda_3\lambda_4) z_{34}^2.\end{aligned}$$

We claim that the coefficients of the z_{ij}^2–terms are all nonpositive. To see this, consider the first one:

$$\begin{aligned}(3\text{-}5)\quad & -\lambda_1^2 - \cdots - \lambda_4^2 - 2\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4 \\ &= -(\lambda_1+\lambda_2)^2 - \lambda_3^2 - \lambda_4^2 + (\lambda_1+\lambda_2)\lambda_3 + (\lambda_1+\lambda_2)\lambda_4 + \lambda_3\lambda_4 \\ &\le -(\lambda_1+\lambda_2)^2 - \lambda_3^2 - \lambda_4^2 + \tfrac{1}{2}(\lambda_1+\lambda_2)^2 + \tfrac{1}{2}\lambda_3^2 + \tfrac{1}{2}(\lambda_1+\lambda_2)^2 \\ &\qquad + \tfrac{1}{2}\lambda_4^2 + \tfrac{1}{2}\lambda_3^2 + \tfrac{1}{2}\lambda_4^2 \\ &= 0. \qquad \square\end{aligned}$$

Finally, we prove the required curvature inequality, which is a more precise statement of Theorem 1.3.

Theorem 3.5 *Let (M^4, g) be a compact Riemannian manifold such that $A_g \in \Gamma_2^+$. Given $u \in \Gamma_2^+$ and $\phi, \psi \in T_u\Gamma_2^+$, we have*

$$\begin{aligned}&K(\phi,\psi) \\ &= \int \frac{1}{\sigma_2(g_u^{-1}A_u)} \big\{ -\langle T_1(g_u^{-1}A_u), \nabla\phi\otimes\nabla\phi\rangle\langle T_1(g_u^{-1}A_u), \nabla\psi\otimes\nabla\psi\rangle \\ &\qquad + \langle T_1(g_u^{-1}A_u), \nabla\phi\otimes\nabla\psi\rangle^2 + \sigma_2(g_u^{-1}A_u)|\nabla\phi|^2|\nabla\psi|^2 \\ &\qquad\qquad - \sigma_2(g_u^{-1}A_u)|\langle\nabla\phi,\nabla\psi\rangle|^2 \big\}\, dV_u \\ &\le 0,\end{aligned}$$

where the inner products are with respect to g_u

Proof Let $u(s,t)$ be a 2–parameter family of conformal factors, and $\alpha=\alpha(s,t)\in T_{u(s,t)}\mathcal{C}^+$.

For economy, we use the notation

$$\mu=\frac{1}{\sigma_2(g_u^{-1}A_u)}.$$

Using the formula for the directional derivative in (3-1), we have

$$\begin{aligned}
(3\text{-}6)\quad \frac{D}{\partial s}\frac{D}{\partial t}\alpha &= \frac{\partial}{\partial s}\Big(\frac{D}{\partial t}\alpha\Big)-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{D}{\partial t}\alpha\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\\
&=\frac{\partial}{\partial s}\Big\{\frac{\partial\alpha}{\partial t}-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\\
&\qquad\qquad -\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{D}{\partial t}\alpha\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\\
&=\frac{\partial^2\alpha}{\partial s\partial t}+\mu^2\Big\langle T_1(g_u^{-1}A_u),\nabla^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\\
&\quad-\mu\Big\langle\frac{\partial}{\partial s}T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\\
&\quad-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial\alpha}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)+\nabla\alpha\otimes\nabla\Big(\frac{\partial^2 u}{\partial s\partial t}\Big)\Big\rangle_u\\
&\quad-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{D}{\partial t}\alpha\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u.
\end{aligned}$$

In the above, we have used the fact that the inner product on symmetric 2–tensors satisfies

$$\frac{\partial}{\partial s}\langle\cdot,\cdot\rangle_u=4\frac{\partial u}{\partial s}\langle\cdot,\cdot\rangle_u.$$

For the last term in (3-6), we have

$$\begin{aligned}
&-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{D}{\partial t}\alpha\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\\
&\quad=-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\frac{\partial\alpha}{\partial t}-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\\
&\quad=-\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial\alpha}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\\
&\qquad\qquad+\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u.
\end{aligned}$$

By (2-4) and (2-2),

$$\frac{\partial}{\partial s}T_1(g_u^{-1}A_u) = \frac{\partial}{\partial s}\{-A_u + \sigma_1(g_u^{-1}A_u)g_u\} = -\nabla_u^2\Big(\frac{\partial u}{\partial s}\Big) + \Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u.$$

Substituting this into (3-6), we get

$$\begin{aligned}
&\frac{D}{\partial s}\frac{D}{\partial t}\alpha = \\
&\frac{\partial^2\alpha}{\partial s\partial t} + \mu\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad + \Big\langle \nabla_u^2\Big(\frac{\partial u}{\partial s}\Big) - \Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u, \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad - \Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial \alpha}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad + \Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial^2 u}{\partial s\partial t}\Big)\Big\rangle_u \\
&\qquad - \Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial \alpha}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u \\
&\qquad\quad + \Big\langle T_1(g_u^{-1}A_u), \nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}.
\end{aligned}$$

Next, we rearrange the terms into two groups: those symmetric in s and t, and those that are not:

$$\begin{aligned}
&\frac{D}{\partial s}\frac{D}{\partial t}\alpha = \\
&\frac{\partial^2\alpha}{\partial s\partial t} + \mu\Big\{-\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial \alpha}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad\qquad - \Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial \alpha}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u \\
&\qquad\qquad\qquad + \Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial^2 u}{\partial s\partial t}\Big)\Big\rangle_u\Big\} \\
&\quad + \mu\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad + \Big\langle \nabla_u^2\Big(\frac{\partial u}{\partial s}\Big) - \Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u, \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad + \Big\langle T_1(g_u^{-1}A_u), \nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}.
\end{aligned}$$

Therefore,

$$
\begin{aligned}
(3\text{-}7)\quad & \Big(\frac{D}{\partial s}\frac{D}{\partial t}-\frac{D}{\partial t}\frac{D}{\partial s}\Big)\alpha \\
&= \mu\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\quad -\mu\Big\langle T_1(g_u^{-1}A_u),\nabla^2\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u \\
&\quad +\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial s}\Big)-\Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u,\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\quad -\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial t}\Big)-\Delta_u\Big(\frac{\partial u}{\partial t}\Big)g_u,\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u \\
&\quad +\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u \\
&\quad -\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\alpha\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}.
\end{aligned}
$$

To compute the sectional curvature of the plane spanned by $\{\frac{\partial u}{\partial s},\frac{\partial u}{\partial t}\}$, we take $\alpha=\frac{\partial u}{\partial t}$ in the formula above, then take the inner product with $\frac{\partial u}{\partial s}$:

$$
\begin{aligned}
&\Big\langle\Big(\frac{D}{\partial s}\frac{D}{\partial t}-\frac{D}{\partial t}\frac{D}{\partial s}\Big)\frac{\partial u}{\partial t},\frac{\partial u}{\partial s}\Big\rangle_u \\
&= \int\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla_u^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&\quad -\mu\Big\langle T_1(g_u^{-1}A_u),\nabla_u^2\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&\quad +\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial s}\Big)-\Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u,\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&\quad -\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial t}\Big)-\Delta_u\Big(\frac{\partial u}{\partial t}\Big)g_u,\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&\quad +\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&\quad -\Big\langle T_1(g_u^{-1}A_u),\nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u),\nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\frac{\partial u}{\partial s}\Big\}\,dV_u.
\end{aligned}
$$

Consider the last two lines above. Integrating by parts and using the fact that $T_1(g_u^{-1}A_u)$ is divergence-free, we get

$$\begin{aligned}
\int \Big\{ &\Big\langle T_1(g_u^{-1}A_u), \nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&-\Big\langle T_1(g_u^{-1}A_u), \nabla\Big\{\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\frac{\partial u}{\partial s}\Big\}\,dV_u \\
= \int \Big\{ &-\mu\Big\langle T_1(g_u^{-1}A_u), \nabla_u^2\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&-\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&+\mu\Big\langle T_1(g_u^{-1}A_u), \nabla_u^2\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\frac{\partial u}{\partial s} \\
&+\mu\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u^2\Big\}\,dV_u.
\end{aligned}$$

Substituting this into (3-7) we find that the first two lines there cancel, and we arrive at

$$\begin{aligned}
&\Big\langle\Big(\frac{D}{\partial s}\frac{D}{\partial t}-\frac{D}{\partial t}\frac{D}{\partial s}\Big)\frac{\partial u}{\partial s},\frac{\partial u}{\partial t}\Big\rangle_u \\
&= \int\Big\{\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial s}\Big)-\Delta_u\Big(\frac{\partial u}{\partial s}\Big)g_u, \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad -\Big\langle\nabla_u^2\Big(\frac{\partial u}{\partial t}\Big)-\Delta_u\Big(\frac{\partial u}{\partial t}\Big)g_u, \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\}\frac{\partial u}{\partial s}\,dV_u \\
&\quad +\int\mu\Big\{-\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad +\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u^2\Big\}\,dV_u.
\end{aligned}$$

From Lemmas 3.3 and 3.4 we conclude

$$\begin{aligned}
&\Big\langle\Big(\frac{D}{\partial s}\frac{D}{\partial t}-\frac{D}{\partial t}\frac{D}{\partial s}\Big)\frac{\partial u}{\partial s},\frac{\partial u}{\partial t}\Big\rangle_u \\
&= \int\mu\Big\{-\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial s}\Big)\Big\rangle_u\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial t}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u \\
&\qquad +\Big\langle T_1(g_u^{-1}A_u), \nabla\Big(\frac{\partial u}{\partial s}\Big)\otimes\nabla\Big(\frac{\partial u}{\partial t}\Big)\Big\rangle_u^2 \\
&\qquad +\sigma_2(g_u^{-1}A_u)\Big|\nabla\frac{\partial u}{\partial s}\Big|_u^2\Big|\nabla\frac{\partial u}{\partial t}\Big|_u^2-\sigma_2(g_u^{-1}A_u)\Big|\Big\langle\nabla\frac{\partial u}{\partial s},\nabla\frac{\partial u}{\partial t}\Big\rangle_u\Big|_u\Big\}\,dV_u \\
&\le 0,
\end{aligned}$$

as required. □

Remark 3.6 The Mabuchi metric turns out to be formally an infinite-dimensional symmetric space, evidenced by the sectional curvatures admitting an interpretation as the square norm of the Poisson bracket of the two tangent vector functions. There does not seem to be such an interpretation in this setting.

3.2 Formal metric space structure

In this subsection we observe some fundamental properties of lengths of curves and distances in the σ_2 metric.

Definition 3.7 Given a path $u\colon [a,b] \to \mathcal{C}^+$, the *length of u* is

$$\mathcal{L}(u) := \int_a^b \langle \alpha, \beta \rangle^{\frac{1}{2}}\, dt = \int_a^b \left[\int_M \left(\frac{\partial u}{\partial t}\right)^2 \sigma_2(g_u^{-1} A_u)\, dV_u \right]^{\frac{1}{2}} dt.$$

A curve is a *geodesic* if it is a critical point for L.

Lemma 3.8 *A curve $u_t \in \mathcal{C}^+$ is a geodesic if and only if*

$$u_{tt} - \frac{1}{\sigma_2(A_u)} \langle T_1(A_u), \nabla u_t \otimes \nabla u_t \rangle = 0. \tag{3-8}$$

Proof Formally, by Lemma 3.2 the connection is indeed the Riemannian connection and so a curve is a geodesic if and only if

$$0 = \frac{D}{\partial t}\frac{\partial u}{\partial t} = u_{tt} - \frac{1}{\sigma_2(A_u)} \langle T_1(A_u), \nabla u_t \otimes \nabla u_t \rangle.$$

This can also be derived by directly taking the first variation of the length functional. □

Remark 3.9 We observe a canonical isometric splitting of $T_u\mathcal{C}^+$ with respect to the σ_k metric. In particular, the real line $\mathbb{R} \subset T_u\mathcal{C}^+$ given by constant functions is orthogonal to

$$T_u^0\mathcal{C}^+ := \left\{ \alpha \;\middle|\; \int_M \alpha \sigma_2(g_u^{-1} A_u)\, dV_u = 0 \right\}.$$

In the next lemma we show two basic properties of geodesics, namely that they preserve this isometric splitting, and are automatically parametrized with constant speed.

Lemma 3.10 *Let u_t be a solution to* (3-8). *Then*

$$\frac{d}{dt} \int_M u_t \sigma_2(g_u^{-1} A_u)\, dV_u = 0,$$
$$\frac{d}{dt} \int_M u_t^2 \sigma_2(g_u^{-1} A_u)\, dV_u = 0.$$

Proof Differentiating and using (2-5),

$$
\begin{aligned}
\frac{d}{dt}\int_M & u_t\sigma_2(g_u^{-1}A_u)\,dV_u \\
&= \int_M (u_{tt}\sigma_2(g_u^{-1}A_u) + u_t\langle T_1(g_u^{-1}A_u), \nabla^2 u_t\rangle_u)\,dV_u \\
&= \int_M (u_{tt} - \sigma_2(g_u^{-1}A_u)^{-1}\langle T_1(g_u^{-1}A_u), \nabla u_t \otimes \nabla u_t\rangle_u)\sigma_2(g_u^{-1}A_u)\,dV_u \\
&= 0.
\end{aligned}
$$

Next,

$$
\begin{aligned}
\frac{d}{dt}\int_M & u_t^2\sigma_2(g_u^{-1}A_u)\,dV_u \\
&= \int_M [2\sigma_2(g_u^{-1}A_u)u_{tt}u_t + u_t^2\langle T_1(g_u^{-1}A_u), \nabla^2 u_t\rangle_u]\,dV_u \\
&= 2\int_M \sigma_2(g_u^{-1}A_u)u_t\left[u_{tt} - \frac{1}{\sigma_2(g_u^{-1}A_u)}\langle T_1(g_u^{-1}A_u), \nabla u_t \otimes \nabla u_t\rangle_u\right]dV_u \\
&= 0. \qquad \square
\end{aligned}
$$

Proposition 3.11 *Given $u_0, u_1 \in C^\infty(M)$ and $u_t\colon [0,1] \to \mathcal{C}^+$ a geodesic, one has*

$$
\mathcal{L}(u) \geq \sigma^{-\frac{1}{2}} \max\left\{\int_{u_1>u_0} (u_1-u_0)\sigma_2(g_{u_1}^{-1}A_{u_1})\,dV_{u_1}, \int_{u_0>u_1} (u_0-u_1)\sigma_2(g_{u_0}^{-1}A_{u_0})\,dV_{u_0}\right\}.
$$

Proof Observe that the geodesic equation implies $u_{tt} \geq 0$, and so we obtain the pointwise inequality

$$
u_t(0) \leq u_1 - u_0 \leq u_t(1).
$$

Thus using Hölder's inequality we have

$$
\begin{aligned}
E(1) = \left(\int_M u_t^2\sigma_2(g_{u_1}^{-1}A_{u_1})\,dV_{u_1}\right)^{\frac{1}{2}} &\geq \sigma^{-\frac{1}{2}}\int_M |u_t|\sigma_2(g_{u_1}^{-1}A_{u_1})\,dV_{u_1} \\
&\geq \sigma^{-\frac{1}{2}}\int_{u_1>u_0} (u_1-u_0)\sigma_2(g_{u_1}^{-1}A_{u_1})\,dV_{u_1}.
\end{aligned}
$$

A similar argument yields

$$
E(0) \geq \sigma^{-\frac{1}{2}}\int_{u_0>u_1} (u_0-u_1)\sigma_2(g_{u_0}^{-1}A_{u_0})\,dV_{u_0}.
$$

Since geodesics are automatically constant speed by Lemma 3.10, the result follows. $\square$

3.3 Geodesics and the conformal group of the sphere

As in the two-dimensional case, we will show that the one-parameter family of transformations that generate the conformal group of the sphere are geodesics. In anticipation of our forthcoming article on the higher-dimensional case we will prove a more general result.

Let (S^n, g_0) denote the round sphere. Using stereographic projection $\sigma\colon S^n \setminus \{N\} \to \mathbb{R}^n$, where $N \in S^n$ denotes the north pole, one can define a one-parameter of conformal maps of S^n by conjugating the dilation map $\delta_\alpha\colon x \mapsto \alpha^{-1}x$ on $\mathbb{R}^n$ with σ:

$$\phi_\alpha = \sigma^{-1} \circ \delta_\alpha \circ \sigma\colon S^n \to S^n.$$

Taking $\alpha(t) = e^{\lambda t}$, where λ is a fixed real number, we can define the path of conformal metrics

$$g(t) = e^{-2u} g_0 = \phi_\alpha^* g_0 = \left[\frac{2\alpha(t)}{(1+\xi) + \alpha(t)^2(1-\xi)}\right]^2, \tag{3-9}$$

where $\xi = x^{n+1}$ is the $(n+1)$–coordinate function, ie $N = (0, \dots, 0, 1)$ (see [24]).

Proposition 3.12 *If $k = \frac{n}{2}$, the path $g(t) = e^{-2u(t)} g_0\colon (-\infty, +\infty) \to \mathcal{C}^+$ satisfies*

$$u_{tt} - \frac{1}{\sigma_k(A_u)} \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle = 0. \tag{3-10}$$

In particular, when $n = 4$, this path defines a geodesic.

Proof The proof is by a direct calculation. There is a more geometric approach, which is used in the two-dimensional case (see [20, Proposition 3.7]). However, this would require us to introduce a certain gradient flow on paths, and verify that the length is nonincreasing under the flow. To avoid the additional machinery we will just verify that (3-10) holds.

By (3-9),

$$u = u(t) = -\log 2\alpha + \log[(1+\xi) + \alpha^2(1-\xi)].$$

This yields

$$u_t = -\frac{\dot\alpha}{\alpha} + \frac{2\alpha\dot\alpha(1-\xi)}{(1+\xi) + \alpha^2(1-\xi)}$$

and hence

$$u_{tt} = -\frac{\alpha_{tt}}{\alpha} + \left(\frac{\alpha_t}{\alpha}\right)^2 + \frac{[(1+\xi)+\alpha^2(1-\xi)](2\alpha\alpha_{tt} + 2\alpha_t^2)(1-\xi) - 4\alpha^2\alpha_t^2(1-\xi)^2}{[(1+\xi)+\alpha^2(1-\xi)]^2}.$$

Since $\alpha(t) = e^{\lambda t}$, we have

$$u_{tt} = 4\lambda^2 e^{2\lambda t} \frac{1-\xi^2}{[(1+\xi)+\alpha^2(1-\xi)]^2}. \tag{3-11}$$

Also,

$$\begin{aligned}
\nabla u_t &= -\frac{2\alpha\alpha_t \nabla\xi}{(1+\xi)+\alpha^2(1-\xi)} - \frac{2\alpha\alpha_t(1-\xi)}{[(1+\xi)+\alpha^2(1-\xi)]^2}[(1-\alpha^2)\nabla\xi] \\
&= \frac{-2\alpha\alpha_t\nabla\xi}{[(1+\xi)+\alpha^2(1-\xi)]^2}[(1+\xi)+\alpha^2(1-\xi)+(1-\xi)(1-\alpha^2)] \\
&= \frac{-4\alpha\alpha_t\nabla\xi}{[(1+\xi)+\alpha^2(1-\xi)]^2} \\
&= \frac{-4\lambda e^{2\lambda t}\nabla\xi}{[(1+\xi)+\alpha^2(1-\xi)]^2}.
\end{aligned}$$

On S^n, the Schouten tensor is a multiple of the identity; in fact, $A(g_0) = \frac{1}{2} g_0$. Therefore, using standard identities for the symmetric functions,

$$\frac{1}{\sigma_k(g(t)^{-1}A_{g(t)})} T_1(g(t)^{-1}A_{g(t)}) = \frac{2k}{n} g(t) = g(t),$$

since $k = \frac{n}{2}$. Thus,

$$\begin{aligned}
\frac{1}{\sigma_k(g(t)^{-1}A_{g(t)})} &\langle T_{k-1}(g(t)^{-1}A_{g(t)}), \nabla u_t \otimes \nabla u_t\rangle \\
&= 4\lambda^2 e^{2\lambda t} \frac{|\nabla\xi|^2}{[(1+\xi)+\alpha^2(1-\xi)]^2}.
\end{aligned} \tag{3-12}$$

Since $|\nabla\xi|^2 = 1-\xi^2$, comparing (3-11) and (3-12) we see that u satisfies (3-10). $\square$

Remark 3.13 We do not expect conformal vector fields on general backgrounds to generate nontrivial geodesics, and thus nonuniqueness of solutions. It follows from a result of Lelong-Ferrand [25] and Obata [28] that if (M^n, g) is not conformally equivalent to the round sphere, then any conformal Killing field is a Killing field for a conformally related metric. Expressed with respect to this background metric, pullback by a family of isometries will result in no change on the level of conformal factors.

3.4 The F–functional and geodesic convexity

We now derive the geodesic convexity of the F–functional of Chang and Yang. The crucial input is a sharp curvature-weighted Poincaré inequality due to Andrews:

Proposition 3.14 (Andrews; see [11, page 517]) *Let (M^n, g) be a closed Riemannian manifold with positive Ricci curvature. Given $\phi \in C^\infty(M)$ such that $\int_M \phi\, dV = 0$,*

$$\frac{n}{n-1} \int_M \phi^2\, dV \le \int_M (\mathrm{Ric}^{-1})^{ij} \nabla_i \phi \nabla_j \phi\, dV,$$

with equality if and only if $\phi \equiv 0$ or (M^n, g) is isometric to the round sphere.

The convexity of F will follow from a weaker form of this inequality:

Corollary 3.15 *Let (M^4, g) be a closed Riemannian manifold such that $A_g \in \Gamma_2^+$. Given $\phi \in C^\infty(M)$ such that $\int_M \phi\, dV = 0$,*

$$\int_M \frac{1}{\sigma_2(A_g)} T_1(A_g)^{ij} \nabla_i \phi \nabla_j \phi\, dV_g \ge 4 \int_M \phi^2\, dV_g - \left(\frac{4}{\int_M dV_g}\right)\left(\int_M \phi\, dV_g\right)^2,$$

with equality if and only if $\phi \equiv 0$ or (M^n, g) is isometric to the round sphere.

Proof We assume $\int_M \phi\, dV_g = 0$. By Andrews' Poincaré inequality we have

$$\frac{4}{3} \int_M \phi^2\, dV_g \le \int_M (\mathrm{Ric}^{-1})^{ij} \nabla_i \phi \nabla_j \phi\, dV_g.$$

To show the claim it suffices to show that

$$3\,\mathrm{Ric}^{-1}(X, X) \le \frac{1}{\sigma_2(A)} T_1(X, X).$$

Since Ric and $T_1(A)$ commute, it suffices to show that $\mathrm{Ric} \circ T_1 \ge 3\sigma_2(A) g$. Since $\mathrm{Ric} = 2A + \sigma_1(A) g$, this is equivalent to

$$-2A \circ A + \sigma_1(A) A + \sigma_1(A)^2 g \ge 3\sigma_2(A) g.$$

Now let $Z = A - \frac{1}{4}\sigma_1(A) g$; then we can rewrite this as

$$-2Z^2 + \tfrac{9}{8}\sigma_1(A)^2 g \ge 3\sigma_2 g.$$

Now, a Lagrange multiplier argument shows that

$$Z \circ Z \le \tfrac{3}{4}|Z|^2 g.$$

Thus,

$$-2Z^2 + \tfrac{9}{8}\sigma_1(A)^2 g \ge -\tfrac{3}{2}|Z|^2 g + \tfrac{9}{8}\sigma_1(A)^2 g = 3\sigma_2(A) g. \qquad \square$$

Proposition 3.16 *The functional F is geodesically convex.*

Proof It follows from [8] that, for a path of conformal metrics $u = u(t)$,

$$\frac{d}{dt}F[u] = \int_M u_t[-\sigma_2(g_u^{-1}A_u) + \overline{\sigma}]\,dV_u. \tag{3-13}$$

Assuming the path is a geodesic, then differentiating again and using Lemma 3.10 we have

$$\begin{aligned}
\frac{d^2}{dt^2}F[u] &= \frac{d}{dt}\int_M u_t[-\sigma_2(g_u^{-1}A_u) + \overline{\sigma}]\,dV_u \\
&= \sigma\frac{d}{dt}\int_M u_t V_u^{-1}\,dV_u \\
&= \sigma\int_M \left[u_{tt}V_u^{-1} + V_u^{-2}u_t\left(\int_M 4u_t\,dV_u\right) - 4V_u^{-1}u_t^2\right]dV_u \\
&= \sigma V_u^{-1}\Bigg[\int_M \frac{1}{\sigma_2(g_u^{-1}A_u)}\langle T_1(g_u^{-1}A_u), \nabla u_t \otimes \nabla u_t\rangle_u\,dV_u \\
&\qquad\qquad -4\left(\int_M u_t^2\,dV_u - V_u^{-1}\left(\int_M u_t\,dV_u\right)^2\right)\Bigg] \\
&\geq 0,
\end{aligned}$$

where the last line follows from Corollary 3.15. □

4 Estimates of the geodesic equation

In this section we establish several fundamental properties of the geodesic equation (3-8). Once again, for future reference we will consider a more general equation which reduces to (3-8) when $n = 4$ and $k = 2$:

$$u_{tt} = \frac{1}{\sigma_k(A_u)}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle.$$

To begin, we define a certain regularization of this equation. In particular, let

$$\Phi(u) := u_{tt}\sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle.$$

Furthermore, let

$$\Phi_\epsilon(u) = (1+\epsilon)u_{tt}\sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle.$$

We will fix two parameters ϵ and s, and study a priori estimates for

$$\Phi_\epsilon(u(\cdot,\cdot,s)) = sf.$$

To obtain estimates though we will simply fix a function $f \in C^\infty(M \times [0,1])$ and study the equation

$$(\star_{\epsilon,f}) \qquad \mathcal{G}_f^\epsilon(u) = \Phi_\epsilon(u) - f = 0.$$

As remarked above, in the setting of Mabuchi geodesics, as observed by Semmes [32] if one complexifies the time direction, the equation admits an interpretation as a certain modification of the tensor A will show up naturally in the linearized operator. Let

$$E = E_u^\epsilon = (1+\epsilon)u_{tt}A_u - \nabla u_t \otimes \nabla u_t. \tag{4-1}$$

Proposition 4.1 *A path $u \in C^2$ satisfies $(\star_{\epsilon,f})$ if and only if*

$$[(1+\epsilon)u_{tt}]^{1-k}\sigma_k(E_u^\epsilon) = f. \tag{4-2}$$

Proof Using Lemma 2.5 and homogeneity properties of elementary symmetric polynomials, we compute

$$\begin{aligned}\sigma_k(E_u^\epsilon) &= \sigma_k((1+\epsilon)u_{tt}A_u - \nabla u_t \otimes \nabla u_t)\\ &= \sigma_k((1+\epsilon)u_{tt}A_u) - \langle T_{k-1}((1+\epsilon)u_{tt}A_u), \nabla u_t \otimes \nabla u_t\rangle\\ &= [(1+\epsilon)u_{tt}]^{k-1}[(1+\epsilon)u_{tt}\sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle].\end{aligned}$$

The proposition follows. □

We will say that a solution u of $(\star_{\epsilon,f})$ is *admissible* if $E_u^\epsilon \in \Gamma_k^+$. As we will see below, $(\star_{\epsilon,f})$ is elliptic for admissible solutions.

Lemma 4.2 *Let $u = u(s,\cdot) \in C^\infty(M \times [0,1])$ be a one-parameter family of smooth functions such that $\frac{d}{ds}u(s,\cdot)\big|_{s=0} = v$. Then*

$$\frac{d}{ds}u_{tt}^{1-k}\sigma_k(E_{u(s,\cdot)}^\epsilon)\Big|_{s=0} = \mathcal{L}(v),$$

where

$$\begin{aligned}\mathcal{L}(v) = (1+\epsilon)^{k-1}u_{tt}^{-1}f v_{tt}&\\ +u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}(\nabla^2 v + \nabla v \otimes \nabla u + \nabla u \otimes \nabla v - \langle \nabla v, \nabla u\rangle g)&\\ -\nabla v_t \otimes \nabla u_t - \nabla u_t \otimes \nabla v_t + u_{tt}^{-1}v_{tt}\nabla u_t \otimes \nabla u_t\big\rangle.&\end{aligned} \tag{4-3}$$

Proof We compute

$$\begin{aligned}
\text{(4-4)}\quad & \frac{d}{ds} u_{tt}^{1-k} \sigma_k(E_{u_s}^{\epsilon}) \\
&= (1-k) u_{tt}^{-k} \sigma_k(E_u^{\epsilon}) v_{tt} + u_{tt}^{1-k} \Big\langle T_{k-1}(E_u^{\epsilon}), \frac{d}{ds} E_u^{\epsilon} \Big\rangle \\
&= (1-k) u_{tt}^{-k} \sigma_k(E_u^{\epsilon}) v_{tt} \\
&\qquad + u_{tt}^{1-k} \Big\langle T_{k-1}(E_u^{\epsilon}), \\
&\qquad\qquad (1+\epsilon) v_{tt} A_u + (1+\epsilon) u_{tt} \frac{d}{ds} A_u - \nabla v_t \otimes \nabla u_t - \nabla u_t \otimes \nabla v_t \Big\rangle.
\end{aligned}$$

The second term can be simplified using Lemma 2.5 to

$$\begin{aligned}
\text{(4-5)}\quad & (1+\epsilon) u_{tt}^{1-k} \langle T_{k-1}(E_u^{\epsilon}), v_{tt} A_u \rangle \\
&= v_{tt} (1+\epsilon) u_{tt}^{1-k} [u_{tt}^{-1} (1+\epsilon)^{-1} \langle T_{k-1}(E_u^{\epsilon}), E_u^{\epsilon} + \nabla u_t \otimes \nabla u_t \rangle] \\
&= v_{tt} u_{tt}^{-k} [k \sigma_k(E_u^{\epsilon}) + \langle T_{k-1}(E_u), \nabla u_t \otimes \nabla u_t \rangle] \\
&= k v_{tt} u_{tt}^{-k} \sigma_k(E_u^{\epsilon}) + v_{tt} u_{tt}^{-1} (1+\epsilon)^{k-1} \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle \\
&= k v_{tt} u_{tt}^{-k} \sigma_k(E_u^{\epsilon}) + v_{tt} [(1+\epsilon)^k \sigma_k(A_u) - f (1+\epsilon)^{k-1} u_{tt}^{-1}] \\
&= v_{tt} [u_{tt}^{-k} (k-1) \sigma_k(E_u^{\epsilon}) + (1+\epsilon)^k \sigma_k(A_u)].
\end{aligned}$$

Hence, the overall term involving v_{tt} in (4-4) is $v_{tt}(1+\epsilon)^k \sigma_k(A_u)$. However, we can furthermore express, again using the geodesic equation and Lemma 2.5, that

$$\begin{aligned}
(1+\epsilon)^k \sigma_k(A_u) &= (1+\epsilon)^{k-1} u_{tt}^{-1} f + (1+\epsilon)^{k-1} u_{tt}^{-1} \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle \\
&= (1+\epsilon)^{k-1} u_{tt}^{-1} f + u_{tt}^{-k} \langle T_{k-1}(E_u^{\epsilon}), \nabla u_t \otimes \nabla u_t \rangle.
\end{aligned}$$

Likewise, we simplify the third term of (4-4) as

$$(u_{tt}+\epsilon)^{1-k} \langle T_{k-1}(E_u), (1+\epsilon) u_{tt} (\nabla^2 v + \nabla v \otimes \nabla u + \nabla u \otimes \nabla v - \langle \nabla v, \nabla u \rangle g) \rangle.$$

Collecting these calculations yields the result. □

Lemma 4.3 *Given $f \geq 0$, equation $(\star_{\epsilon,f})$ for admissible u is strictly elliptic for $\epsilon > 0$, and weakly elliptic for $\epsilon = 0$.*

Proof We compute the principal symbol of $\mathcal{L}$. We will ignore the first term of (4-3), which has weakly positive symbol. Now fix a vector $V = (\lambda, X) \in T[0,1] \times TM$. It follows from (4-3) that the principal symbol of $\mathcal{L}$ acts via

$$\begin{aligned}
L(V,V) = u_{tt}^{1-k} \Big\langle T_{k-1}(E_u^{\epsilon}), (1+\epsilon) u_{tt} X \otimes X - \nabla u_t \otimes (\lambda X) - (\lambda X) \otimes \nabla u_t \\
+ u_{tt}^{-1} \nabla u_t \otimes \nabla u_t (\lambda^2) \Big\rangle.
\end{aligned}$$

It follows from the Cauchy–Schwarz inequality that for any $\rho > 0$, as an inequality of matrices one has

$$-\lambda X \otimes \nabla u_t - \lambda \nabla u_t \otimes X \leq \rho X \otimes X + \rho^{-1}\lambda^2 \nabla u_t \otimes \nabla u_t.$$

Applying this inequality with $\rho = \left(1 + \frac{\epsilon}{2}\right)u_{tt}$ yields

$$\begin{aligned}(1+\epsilon)u_{tt} X \otimes X - \nabla u_t \otimes (\lambda X) - (\lambda X) \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t (\lambda^2)\\ \geq \tfrac{1}{2}\epsilon u_{tt} X \otimes X + \tfrac{1}{2}\epsilon u_{tt}^{-1}\lambda^2.\end{aligned}$$

Since u is admissible, we have $T_{k-1}(E_u^\epsilon) > 0$, and the result follows. $\square$

4.1 C^0 estimate

To prove a C^0 estimate we begin with two technical lemmas:

Lemma 4.4 *Suppose $\phi = \phi(t)$. Then*

$$\mathcal{L}\phi = \phi_{tt}(1+\epsilon)^k \sigma_k(A_u).$$

Proof We directly compute using (4-3), Lemma 2.5 and the geodesic equation that

$$\begin{aligned}\mathcal{L}\phi &= \phi_{tt}\{(1+\epsilon)^{k-1} u_{tt}^{-1} f + u_{tt}^{1-k}\langle T_{k-1}(E_u^\epsilon), u_{tt}^{-1}\nabla u_t \otimes \nabla u_t\rangle\}\\ &= \phi_{tt}\{(1+\epsilon)^{k-1} u_{tt}^{-1} f + u_{tt}^{1-k}\langle T_{k-1}((1+\epsilon)u_{tt}A_u), u_{tt}^{-1}\nabla u_t \otimes \nabla u_t\rangle\}\\ &= \phi_{tt}(1+\epsilon)^{k-1}\{u_{tt}^{-1} f + u_{tt}^{-1}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle\}\\ &= \phi_{tt}(1+\epsilon)^k \sigma_k(A_u).\end{aligned}$$ $\square$

Lemma 4.5 *Let u be an admissible solution to $(\star_{\epsilon,f})$. Then*

$$\mathcal{L}u = (k+1)(1+\epsilon)^{k-1} f + (1+\epsilon)u_{tt}^{2-k}\left\langle T_{k-1}(E_u), -A + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\right\rangle.$$

Proof To begin we directly compute using (4-3) that

$$\begin{aligned}\mathcal{L}u = (1+\epsilon)^{k-1} f + u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}(\nabla^2 u + 2\nabla u \otimes \nabla u - |\nabla u|^2 g)\\ - \nabla u_t \otimes \nabla u_t\big\rangle.\end{aligned}$$

For the second term we simplify

$$\begin{aligned}
&(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u^{\epsilon}),\nabla^2 u\rangle\\
&\quad= (1+\epsilon)u_{tt}^{2-k}\bigl\langle T_{k-1}(E_u), A_u - A - \nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr\rangle\\
&\quad= u_{tt}^{2-k}\langle T_{k-1}(E_u), u_{tt}^{-1}[E_u+\nabla u_t\otimes\nabla u_t]\rangle\\
&\qquad\qquad + (1+\epsilon)u_{tt}^{2-k}\bigl\langle T_{k-1}(E), -A - \nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr\rangle\\
&\quad= k u_{tt}^{1-k}\sigma_k(E) + u_{tt}^{1-k}\langle T_{k-1}(E_u),\nabla u_t\otimes\nabla u_t\rangle\\
&\qquad\qquad + (1+\epsilon)u_{tt}^{2-k}\bigl\langle T_{k-1}(E_u), -A - \nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr\rangle\\
&\quad= k(1+\epsilon)^{k-1} f + u_{tt}^{1-k}\langle T_{k-1}(E_u),\nabla u_t\otimes\nabla u_t\rangle\\
&\qquad\qquad + (1+\epsilon)u_{tt}^{2-k}\bigl\langle T_{k-1}(E_u), -A - \nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr\rangle.
\end{aligned}$$

Combining these calculations yields the result. □

Proposition 4.6 *Let u be an admissible solution to $(\star_{\epsilon,f})$. Then*

$$\sup_{M\times[0,1]}|u| \le C\bigl(u_{|M\times\{0,1\}}, \max_M f\bigr).$$

Proof We first observe that an admissible solution to $(\star_{\epsilon,f})$ satisfies $u_{tt}\ge 0$, and hence by convexity one has $\sup_{M\times[0,1]} u \le \sup_{M\times\{0,1\}} u$. To obtain the lower bound, fix a constant Λ and let

$$\Psi = u + \Lambda t(1-t).$$

Observe that at an interior spacetime minimum of Ψ one has

$$0 = \nabla u,\quad \nabla^2 u > 0.$$

Using this and Lemma 4.5 yields, at such a spacetime minimum,

$$\begin{aligned}
\mathcal{L}\Psi &= (k+1)(1+\epsilon)^{k-1} f - (1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u), A\rangle\\
&\qquad - 2\Lambda[(1+\epsilon)^{k-1}u_{tt}^{-1} f + u_{tt}^{1-k}\langle T_{k-1}(E_u^{\epsilon}), u_{tt}^{-1}\nabla u_t\otimes\nabla u_t\rangle].
\end{aligned}$$

Next we claim

$$\Psi_{tt}\nabla^2\Psi - \nabla\Psi_t\otimes\nabla\Psi_t \ge 0.$$

Since we are at a minimum for Ψ, the matrix $\Psi_{tt}\nabla^2\Psi$ is positive semidefinite. The expression above is thus the difference between a positive semidefinite matrix and a negative definite rank 1 matrix. The lemma follows if we establish positivity in the nondegenerate direction of the rank 1 matrix we subtracted, ie $\nabla\Psi_t$. In particular, it then suffices to show

$$\Psi_{tt}\nabla^k\nabla^l\Psi\nabla_k\Psi_t\nabla_l\Psi_t - |\nabla\Psi_t|^4 \geq 0.$$

To establish this we use that Ψ is actually a spacetime minimum. This implies that the spacetime Hessian is positive semidefinite. Testing this condition against the vector

$$-\sqrt{\Psi_{tt}}\nabla\Psi_t \oplus \frac{|\nabla\Psi_t|^2}{\sqrt{\Psi_{tt}}}\frac{\partial}{\partial t}$$

yields

$$0 \leq \Psi_{tt}\nabla^k\nabla^l\Psi\nabla_k\Psi_t\nabla_l\Psi_t - 2|\nabla\Psi_t|^4 + |\nabla\Psi_t|^4,$$

as required. However, using the explicit form of Ψ we see that this implies

$$(u_{tt} - \Lambda)\nabla^2 u - \nabla u_t \otimes \nabla u_t \geq 0,$$

which, since $\nabla^2 u > 0$, implies

$$u_{tt}\nabla^2 u - \nabla u_t \otimes \nabla u_t \geq 0.$$

Hence $E_u \geq u_{tt}A$, and then we obtain using Lemma 2.3 that

$$\begin{aligned} u_{tt}^{2-k}\langle T_{k-1}(E_u), A\rangle &= u_{tt}^{2-k}\Sigma(E_u, \dots, E_u, A) \\ &\geq u_{tt}^{2-k}\Sigma(u_{tt}A, \dots, u_{tt}A, A) \\ &= u_{tt}\sigma_k(A) \\ &\geq 0. \end{aligned}$$

We can also simplify

$$\begin{aligned} u_{tt}^{1-k}\langle T_{k-1}(E_u^\epsilon), u_{tt}^{-1}\nabla u_t \otimes \nabla u_t\rangle &= (1+\epsilon)^{k-1}u_{tt}^{-1}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle \\ &= -(1+\epsilon)^{k-1}u_{tt}^{-1}f + (1+\epsilon)^{k-1}\sigma_k(A_u). \end{aligned}$$

Combining these observations yields, at the interior minimum,

$$\begin{aligned} \mathcal{L}\Psi &\leq (k+1)(1+\epsilon)^{k-1}f - 2\Lambda(1+\epsilon)^{k-1}\sigma_k(A_u) \\ &\leq (k+1)(1+\epsilon)^{k-1}f - 2\Lambda(1+\epsilon)^{k-1}\sigma_k(A) \\ &\leq Cf - 2\delta\Lambda \end{aligned}$$

for some constants C and δ depending only on the background data and maximum of f. Choosing Λ sufficiently large with respect to these constants yields $\mathcal{L}\Psi < 0$. Hence, Ψ cannot have an interior minimum, and the result follows. $\square$

Remark 4.7 In the following estimates, all bounds on solutions be understood to depend on

$$\max_M \left\{ f + \frac{|f_t|}{f} + \frac{|\nabla f|}{f} + \frac{|f_{tt}|}{f} + \frac{|\nabla^2 f|}{f} \right\},$$

but this dependence will be suppressed to simplify the exposition.

4.2 C^1 estimates

Proposition 4.8 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$\sup_{M\times[0,1]} |u_t| \le C.$$

Proof First we observe that, since $u_{tt} \ge 0$, there is a constant such that $u_t(0) \le C$ by direct integration. Now fix constants Λ_1 and Λ_2 and consider

$$\Phi(x,t) = u(x,t) - u(x,0) - \Lambda_1 t^2 + \Lambda_2 t,$$

where Λ_1 is chosen large below, and Λ_2 is chosen still larger so that $\Phi(x,1) \ge 0$. First note using (4-3) that

$$\mathcal{L}u_0 = u_{tt}^{1-k} \big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}(\nabla^2 u_0 + \nabla u_0 \otimes \nabla u + \nabla u \otimes \nabla u_0 - \langle \nabla u_0, \nabla u\rangle g)\big\rangle.$$

Combining this with Lemmas 4.4 and 4.5 we obtain

$$\begin{aligned}
\mathcal{L}\Phi &= \mathcal{L}u - \mathcal{L}u_0 - \Lambda_1 \mathcal{L}t^2 \\
&= (1+\epsilon)u_{tt}^{2-k} \\
&\qquad \times \big\langle T_{k-1}(E_u), -A - \nabla^2 u_0 + \nabla u \otimes \nabla u - 2\nabla u \otimes \nabla u_0 - \tfrac{1}{2}|\nabla u|^2 g + \langle \nabla u_0, \nabla u\rangle g\big\rangle \\
&\qquad + (k+1)(1+\epsilon)^{k-1} f - 2\Lambda_1 (1+\epsilon)^{k-1} u_{tt}^{-1} f - 2\Lambda_1 u_{tt}^{-k} \langle T_{k-1}(E_u^\epsilon), \nabla u_t \otimes \nabla u_t\rangle.
\end{aligned}$$

Also we have $\nabla u = \nabla u_0$ at the minimum, so we can simplify to

$$\begin{aligned}
\mathcal{L}\Phi &= -u_{tt}^{2-k} \big\langle T_{k-1}(E), A + \nabla^2 u_0 + \nabla u_0 \otimes \nabla u_0 - \tfrac{1}{2}|\nabla u_0|^2 g\big\rangle \\
&\qquad\qquad + (k+1)(1+\epsilon)^{k-1} f - 2\Lambda_1 (1+\epsilon)^k \sigma_k(A_u) \\
&= -u_{tt}^{2-k} \langle T_{k-1}(E), A_{u_0}\rangle + (k+1)(1+\epsilon)^{k-1} f - 2\Lambda_1(1+\epsilon)^k \sigma_k(A_u).
\end{aligned}$$

At a spacetime minimum for Φ we have $\nabla^2(u - u_0) \ge 0$, and hence

$$\begin{aligned}
0 &\le \Phi_{tt} \nabla^2 \Phi - \nabla \Phi_t \otimes \nabla \Phi_t \\
&= (u_{tt} - 2\Lambda_1)\nabla^2(u - u_0) - \nabla u_t \otimes \nabla u_t \\
&\le u_{tt} \nabla^2(u - u_0) - \nabla u_t \otimes \nabla u_t.
\end{aligned}$$

Using this yields

$$\begin{aligned} E_u &= [(1+\epsilon)u_{tt}A_u - \nabla u_t \otimes \nabla u_t] \\ &= \big[(1+\epsilon)u_{tt}\big(A + \nabla^2 u + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\big) - \nabla u_t \otimes \nabla u_t\big] \\ &\geq \big[(1+\epsilon)u_{tt}\big(A + \nabla^2 u_0 + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\big)\big] \\ &= \big[(1+\epsilon)u_{tt}\big(A + \nabla^2 u_0 + \nabla u_0 \otimes \nabla u_0 - \tfrac{1}{2}|\nabla u_0|^2 g\big)\big]. \end{aligned}$$

It follows from Lemma 2.4 that

$$\langle T_{k-1}(E), A\rangle \geq 0.$$

A similar calculation shows that at the minimum point under consideration we have

$$\sigma_k(A_u) \geq \sigma_k(A_{u_0}).$$

Putting these estimates together yields

$$\mathcal{L}\Phi \leq (k+1)(1+\epsilon)^{k-1} f - 2\Lambda_1 (1+\epsilon)^k \sigma_k(A_{u_0}).$$

If we choose Λ_1 sufficiently large with respect to the positive lower bound for $\sigma_k(A_{u_0})$ and the maximum of f, we obtain $L\Phi < 0$, and hence Φ cannot have an interior minimum. Thus, it follows that $\Phi_t(x,0) \geq 0$ for all x, and thus the lower bound for $u_t(0)$ follows. A very similar estimate yields a two-sided bound for $u_t(1)$. Since $u_{tt} \geq 0$ everywhere we have a two-sided bound for u_t everywhere. □

We next proceed to obtain the interior spatial gradient estimate. To do this we need two preliminary calculations.

Lemma 4.9 *Let u be an admissible solution to $(\star_{\epsilon,f})$. Then*

$$\mathcal{L}e^{-\lambda u} \geq -\lambda e^{-\lambda u}\mathcal{L}u + \tfrac{1}{2}\lambda^2 e^{-\lambda u} u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), \nabla u \otimes \nabla u\rangle - C\lambda^2 e^{-\lambda u}\sigma_k(A_u)u_t^2.$$

Proof To begin we directly compute using (4-3) that

$$\begin{aligned} \mathcal{L}e^{-\lambda u} = (1+\epsilon)^{k-1} u_{tt}^{-1} f (e^{-\lambda u})_{tt} \\ + u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}(\nabla^2 e^{-\lambda u} + \nabla e^{-\lambda u} \otimes \nabla u + \nabla u \otimes \nabla e^{-\lambda u} \\ - \langle \nabla e^{-\lambda u}, \nabla u\rangle) - \nabla (e^{-\lambda u})_t \otimes \nabla u_t \\ - \nabla u_t \otimes \nabla (e^{-\lambda u})_t + u_{tt}^{-1}\nabla u_t \otimes \nabla u_t (e^{-\lambda u})_{tt}\big\rangle \end{aligned}$$

$$
\begin{aligned}
&= -\lambda e^{-\lambda u}\mathcal{L}u + (1+\epsilon)^{k-1}u_{tt}^{-1} f\lambda^2 e^{-\lambda u}u_t^2 \\
&\qquad + \lambda^2 e^{-\lambda u} u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}\nabla u\otimes\nabla u - u_t\nabla u\otimes\nabla u_t \\
&\qquad\qquad\qquad\qquad\qquad\qquad - u_t\nabla u_t\otimes\nabla u + u_{tt}^{-1}u_t^2\nabla u_t\otimes\nabla u_t\big\rangle.
\end{aligned}
$$

Next we observe using the Cauchy–Schwarz inequality and equation $(\star_{\epsilon,f})$ that

$$
\begin{aligned}
&u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}\nabla u\otimes\nabla u - u_t\nabla u\otimes\nabla u_t - u_t\nabla u_t\otimes\nabla u + u_{tt}^{-1}u_t^2\nabla u_t\otimes\nabla u_t\big\rangle \\
&\qquad = \sigma_k(A_u)u_t^2 - 2u_t u_{tt}^{1-k}\langle T_{k-1}(E_u), \nabla u_t\otimes\nabla u\rangle + u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla u\otimes\nabla u\rangle \\
&\qquad \geq -C\sigma_k(A_u)u_t^2 + \tfrac{1}{2}u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla u\otimes\nabla u\rangle.
\end{aligned}
$$

Combining these calculations yields the result. □

Lemma 4.10 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$
\mathcal{L}u_t^2 = 2u_t f_t + 2(1+\epsilon)^{k-1} f u_{tt} + 2\epsilon u_{tt}^{2-k} T_{k-1}(E)^{jk}\nabla_j u_t\nabla_k u_t.
$$

Proof It follows directly from the definition of $\mathcal{L}$ that $\mathcal{L}u_t = f_t$. It follows that

$$
\begin{aligned}
\mathcal{L}u_t^2 &= 2u_t\mathcal{L}u_t + 2(1+\epsilon)^{k-1} f u_{tt} \\
&\qquad + 2u_{tt}^{1-k}T_{k-1}(E)^{jk}\{(1+\epsilon)u_{tt}\nabla_j u_t\nabla_k u_t - 2\nabla_j u_t\nabla_k u_t u_{tt} + \nabla_j u_t\nabla_k u_t u_{tt}\} \\
&= 2u_t f_t + 2(1+\epsilon)^{k-1} f u_{tt} + 2\epsilon u_{tt}^{2-k}T_{k-1}(E)^{jk}\nabla_j u_t\nabla_k u_t,
\end{aligned}
$$

as required. □

Lemma 4.11 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$
\begin{aligned}
\mathcal{L}|\nabla u|^2 &= 2u_{tt}^{1-k}T_{k-1}(E)^{jk}\big\{(1+\epsilon)u_{tt}\nabla_i\nabla_j u\nabla_i\nabla_k u - 2\nabla_i\nabla_j u\nabla_k u_t\nabla_i u_t \\
&\qquad\qquad\qquad\qquad\qquad\qquad + u_{tt}^{-1}\nabla_j u_t\nabla_k u_t|\nabla u_t|^2\big\} \\
&\qquad + 2(1+\epsilon)^{k-1}u_{tt}^{-1} f|\nabla u_t|^2 + 2\langle\nabla f, \nabla u\rangle \\
&\qquad - 2(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E), \nabla^i u\nabla_i A + R_{ijk}^{l}\nabla^i u\nabla_l u\rangle.
\end{aligned}
$$

Proof To begin we take the gradient of the geodesic equation to yield

$$
\begin{aligned}
\nabla_i f &= \nabla_i[u_{tt}^{1-k}\sigma_k(E_u^\epsilon)] \\
&= (1-k)u_{tt}^{-k}\nabla_i u_{tt}\sigma_k(E_u^\epsilon) + u_{tt}^{1-k}\langle T_{k-1}(E_u^\epsilon), \nabla_i E_u^\epsilon\rangle \\
&= (1-k)u_{tt}^{-k}\nabla_i u_{tt}\sigma_k(E_u) \\
&\qquad + u_{tt}^{1-k}\big\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)\nabla_i u_{tt}A_u + (1+\epsilon)u_{tt}\nabla_i A_u - \nabla_i\nabla u_t\otimes\nabla u_t \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \nabla u_t\otimes\nabla_i\nabla u_t\big\rangle.
\end{aligned}
$$

A calculation similar to (4-5) shows that

$$\begin{aligned}(1-k)u_{tt}^{-k}\nabla_i u_{tt}\sigma_k(E_u)+u_{tt}^{1-k}\langle T_{k-1}(E_u),(1+\epsilon)\nabla_i u_{tt}A_u\rangle \\ =(1+\epsilon)^{k-1}u_{tt}^{-1}f\nabla_i u_{tt}+u_{tt}^{-k}\langle T_{k-1}(E),\nabla u_t\otimes\nabla u_t\rangle\nabla_i u_{tt}.\end{aligned}$$

Next we simplify via

$$\begin{aligned}&\nabla_i(A_u)_{jk}\\ &\quad=\nabla_i\big[A_{jk}+\nabla_j\nabla_k u+\nabla_j u\nabla_k u-\tfrac{1}{2}|\nabla u|^2 g_{jk}\big]\\ &\quad=\nabla_i A_{jk}+\nabla_i\nabla_j\nabla_k u+\nabla_i\nabla_j u\nabla_k u+\nabla_j u\nabla_i\nabla_k u-\tfrac{1}{2}\nabla_i|\nabla u|^2 g_{jk}\\ &\quad=\nabla_i A_{jk}+\nabla_j\nabla_k\nabla_i u+R_{ijk}^l\nabla_l u+\nabla_i\nabla_j u\nabla_k u+\nabla_j u\nabla_i\nabla_k u-\tfrac{1}{2}\nabla_i|\nabla u|^2 g_{jk}.\end{aligned}$$

Hence, we obtain the identity

$$\mathcal{L}\nabla_i u=\nabla_i f-(1+\epsilon)u_{tt}^{2-k}T_{k-1}(E)^{jk}\{\nabla_i A_{jk}+R_{ijk}^l\nabla_l u\}. \tag{4-6}$$

On the other hand, using (4-3) we have

$$\begin{aligned}\mathcal{L}|\nabla u|^2&=2\langle\mathcal{L}\nabla u,\nabla u\rangle+2(1+\epsilon)^{k-1}u_{tt}^{-1}f|\nabla u_t|^2\\ &\qquad+2u_{tt}^{1-k}T_{k-1}(E)^{jk}\big\{(1+\epsilon)u_{tt}\nabla_i\nabla_j u\nabla_i\nabla_k u-2\nabla_i\nabla_j u\nabla_k u_t\nabla_i u_t\\ &\qquad\qquad+u_{tt}^{-1}\nabla_j u_t\nabla_k u_t|\nabla u_t|^2\big\}\\ &=2u_{tt}^{1-k}T_{k-1}(E)^{jk}\big\{(1+\epsilon)u_{tt}\nabla_i\nabla_j u\nabla_i\nabla_k u-2\nabla_i\nabla_j u\nabla_k u_t\nabla_i u_t\\ &\qquad\qquad+u_{tt}^{-1}\nabla_j u_t\nabla_k u_t|\nabla u_t|^2\big\}\\ &\qquad+2(1+\epsilon)^{k-1}u_{tt}^{-1}f|\nabla u_t|^2+2\langle\nabla f,\nabla u\rangle\\ &\qquad-2(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E),\nabla^i u\nabla_i A+R_{ijk}^l\nabla^i u\nabla_l u\rangle,\end{aligned}$$

as required. □

Proposition 4.12 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$\sup_{M\times[0,1]}|\nabla u|^2\le C.$$

Proof Without loss of generality we can assume $u<0$. Choose $\lambda_1,\lambda_2,\lambda_3\in\mathbb{R}$ and let

$$\Phi=|\nabla u|^2+\lambda_1 u_t^2+e^{-\lambda_2 u}+\lambda_3 t(t-1).$$

Lemmas 4.4, 4.9, 4.10 and 4.11 show that

$$\begin{aligned}\mathcal{L}\Phi\ge\mathcal{L}|\nabla u|^2+2\lambda_1[f_t u_t+(1+\epsilon)^{k-1}f u_{tt}+\epsilon u_{tt}^{2-k}T_{k-1}(E)^{jk}\nabla_j u_t\nabla_k u_t]-\lambda_2\mathcal{L}ue^{-\lambda_2 u}\\ +\tfrac{1}{2}\lambda_2^2 e^{-\lambda_2 u}u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon),\nabla u\otimes\nabla u\rangle-C\lambda_2^2 e^{-\lambda_2 u}\sigma_k(A_u)u_t^2+\lambda_3\sigma_k(A_u)\end{aligned}$$

$$\begin{aligned}
&\geq 2\langle \nabla f, \nabla u\rangle + 2\sigma_k(A_u)|\nabla u_t|^2 + 2u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla_j \nabla_i u \nabla_k \nabla_i u\rangle \\
&\quad - 4u_{tt}^{1-k}\langle T_{k-1}(E_u)_{ij}, \nabla_i u_t \nabla_k u_t \nabla_j \nabla_k u\rangle \\
&\quad - 2u_{tt}^{2-k}\langle T_{k-1}(E_u)_{jk}\nabla_i u, \nabla_i A_{jk} + R_{ijk}^l \nabla_l u\rangle - C\lambda_1 f_t + \lambda_1 f u_{tt} \\
&\quad - \lambda_2 e^{-\lambda_2 u}\big[f - u_{tt}^{2-k}\langle T_{k-1}(E_u), A\rangle \\
&\qquad\qquad + u_{tt}^{2-k}\big[\langle T_{k-1}(E_u), \nabla u \otimes \nabla u\rangle - \tfrac{1}{2}\operatorname{tr} T_{k-1}(E_u)|\nabla u|^2\big]\big] \\
&\quad + \tfrac{1}{2}\lambda_2^2 e^{-\lambda_2 u} u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), \nabla u \otimes \nabla u\rangle - C\lambda_2^2 e^{-\lambda_2 u}\sigma_k(A_u)u_t^2 + \lambda_3 \sigma_k(A_u).
\end{aligned}$$

First we observe that, using the Cauchy–Schwarz inequality and Lemma 2.5,

$$\begin{aligned}
&4u_{tt}^{1-k}\langle T_{k-1}(E_u)_{ij}, \nabla_i u_t \nabla_k u_t \nabla_j \nabla_k u\rangle \\
&\qquad = 4u_{tt}^{1-k}[\langle T_{k-1}(E_u)^{\frac{1}{2}} \cdot \nabla^2 u, T_{k-1}(E_u)^{\frac{1}{2}} \cdot \nabla u_t \otimes \nabla u_t\rangle] \\
&\qquad \leq 2u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u\rangle + 2u_{tt}^{-k}\langle T_{k-1}(E_u), \nabla u_t \otimes \nabla u_t\rangle|\nabla u_t|^2 \\
&\qquad = 2u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u\rangle + 2u_{tt}^{-1}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t\rangle|\nabla u_t|^2 \\
&\qquad = 2u_{tt}^{2-k}\langle T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u\rangle + 2[\sigma_k(A_u) - f u_{tt}^{-1}]|\nabla u_t|^2.
\end{aligned}$$

Observe the preliminary inequality

$$\begin{aligned}
u_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u) &= u_{tt}^{2-k}\sigma_{k-1}(E_u) \\
&\geq u_{tt}^{2-k}[\sigma_k(E_u)^{(k-1)/k}] \\
&= u_{tt}^{2-k}[f u_{tt}^{k-1}]^{(k-1)/k} \\
&= f^{(k-1)/k} u_{tt}^{2-k+(k^2-2k+1)/k} \\
&= f^{(k-1)/k} u_{tt}^{1/k}.
\end{aligned}$$

Next observe the estimate

$$\begin{aligned}
\langle \nabla f, \nabla u\rangle &\leq C f u_{tt}^{-1/k} + C f u_{tt}^{1/k}|\nabla u|^2 \\
&\leq C f u_{tt}^{-1} + C f u_{tt} + C f u_{tt}^{1/k}|\nabla u|^2 \\
&\leq C f u_{tt}^{-1} + C f u_{tt} + C f^{(k-1)/k} u_{tt}^{1/k}|\nabla u|^2.
\end{aligned}$$

Next observe that

$$-2u_{tt}^{2-k}\langle T_{k-1}(E_u)_{jk}\nabla_i u, \nabla_i A_{jk} + R_{ijk}^l \nabla_l u\rangle \geq -C u_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u)[1 + |\nabla u|^2].$$

Combining these preliminary observations and using Proposition 4.8 yields

$$\begin{aligned}
L\Phi &\geq (\lambda_3 - C) f u_{tt}^{-1} + (\lambda_1 - C) f u_{tt} + \big(\tfrac{1}{4}\lambda_2 e^{-\lambda_2 u} - C\big) f^{(k-1)/k} u_{tt}^{1/k} |\nabla u|^2 - C\lambda_1 f \\
&\qquad + e^{-\lambda_2 u} u_{tt}^{2-k} \big[\tfrac{1}{2}\lambda_2^2 - \lambda_2\big] \langle T_{k-1}(E_u), \nabla u \otimes \nabla u\rangle \\
&\qquad\quad + u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u)\big[-C - C|\nabla u|^2 + \tfrac{1}{4}\lambda_2 e^{-\lambda_2 u}|\nabla u|^2\big] + \sigma_k(A_u)[\lambda_3 - C\lambda_2^2] \\
&\geq \tfrac{1}{2}\lambda_3 f u_{tt}^{-1} + \tfrac{1}{2}\lambda_1 f u_{tt} - C\lambda_1 f + u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u)[-C + |\nabla u|^2] \\
&\geq u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u)[-C + |\nabla u|^2],
\end{aligned}$$

where the second inequality follows by choosing λ_1 and λ_2 large with respect to universal constants and noting that $e^{-\lambda_2 u} > 1$ for every choice of λ_2, and then choosing λ_3 large with respect to these choices. The third inequality follows by choosing λ_3 large with respect to λ_1. Using the previously establishing a priori estimates for u and u_t, at a sufficiently large maximum of Φ we will have $|\nabla u|^2 \geq C$, and hence we see that $L\Phi > 0$ at a sufficiently large maximum, a contradiction. The a priori estimate for $|\nabla u|^2$ follows. □

4.3 C^2 estimates

Lemma 4.13 *Given an admissible solution u of $(\star_{\epsilon,f})$, we have*

$$\begin{aligned}
\mathcal{L}u_{tt} &= -k f^{\frac{k}{k-1}} u_{tt} (1+\epsilon)^k \mathcal{F}^{ij,kl}[(E_u)_t]_{ij}[(E_u)_t]_{kl} \\
&\quad + u_{tt}^{1-k}\big\langle T_{k-1}(E), 2(1+\epsilon) u_{tt}^{-2} u_{ttt}^2 \nabla u_t \otimes \nabla u_t - 4u_{tt}^{-1} u_{ttt} \nabla u_{tt} \otimes \nabla u_t \\
&\qquad\qquad + 2\nabla u_{tt} \otimes \nabla u_{tt} - 2u_{tt} \nabla u_t \otimes \nabla u_t + (1+\epsilon) u_{tt} |\nabla u_t|^2 g\big\rangle \\
&\quad + (1+\epsilon)^{k-1} k f^{\frac{k-1}{k}} (f^{\frac{1}{k}})_{tt} + 2(k-1)(1+\epsilon)^{k-1} u_{tt}^{-1} f^{\frac{k-1}{k}} (f^{\frac{1}{k}})_t u_{ttt} \\
&\quad - 2(1+\epsilon)^{k-1} u_{tt}^{-1} u_{ttt} f_t + (1+\epsilon)^{k-1} \frac{k+1}{k} f u_{tt}^{-2} u_{ttt}^2.
\end{aligned}$$

Proof First we compute using (4-3) that

$$\begin{aligned}
\text{(4-7)} \quad \mathcal{L}u_{tt} &= (1+\epsilon)^{k-1} u_{tt}^{-1} f u_{tttt} \\
&\quad + u_{tt}^{1-k}\big\langle T_{k-1}(E), (1+\epsilon) u_{tt} (\nabla^2 u_{tt} + \nabla u_{tt} \otimes \nabla u + \nabla u \otimes \nabla u_{tt} \\
&\qquad - \langle \nabla u_{tt} \nabla u\rangle g) - \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_{ttt} \\
&\qquad\qquad + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t u_{tttt}\big\rangle.
\end{aligned}$$

To simplify notation we adopt the following (standard) conventions: for an $n \times n$ symmetric matrix $r = r_{ij}$ we write

$$\mathcal{F}(r) = \sigma_k(r)^{1/k},$$

and denote the derivatives of $\mathcal{F}$ with respect to the entries of r by

$$\frac{\partial}{\partial r_{pq}}\mathcal{F}(r)=\mathcal{F}(r)^{pq},$$

$$\frac{\partial^2}{\partial r_{pq}\partial r_{rs}}\mathcal{F}(r)=\mathcal{F}(r)^{pq,rs}.$$

We next need to differentiate the equation, which we can rewrite as

$$c_\epsilon f^{\frac{1}{k}}u_{tt}^{\frac{k-1}{k}}=\sigma_k(E_u)^{\frac{1}{k}}=\mathcal{F}(E_u),$$

where $c_\epsilon=(1+\epsilon)^{\frac{k-1}{k}}$. Differentiating this yields

$$\begin{aligned} c_\epsilon(f^{\frac{1}{k}})_t u_{tt}^{\frac{k-1}{k}}+c_\epsilon\frac{k-1}{k}f^{\frac{1}{k}}u_{tt}^{-\frac{1}{k}}u_{ttt}&=\mathcal{F}^{ij}\Big[\frac{\partial}{\partial t}E_u\Big]_{ij}\\ &=\frac{1}{k}\sigma_k(E_u)^{\frac{1-k}{k}}\langle T_{k-1}(E_u),(E_u)_t\rangle. \end{aligned}$$

Differentiating again yields

$$\begin{aligned} &\mathcal{F}^{ij}[(E_u)_{tt}]_{ij}+\mathcal{F}^{ij,kl}[(E_u)_t]_{ij}[(E_u)_t]_{kl}\\ &\quad=c_\epsilon\bigg[(f^{\frac{1}{k}})_{tt}u_{tt}^{\frac{k-1}{k}}+2\frac{k-1}{k}(f^{\frac{1}{k}})_t u_{tt}^{-\frac{1}{k}}u_{ttt}-\frac{1}{k}\Big(\frac{k-1}{k}\Big)f^{\frac{1}{k}}u_{tt}^{-\frac{1+k}{k}}u_{ttt}^2\\ &\qquad\qquad+\frac{k-1}{k}f^{\frac{1}{k}}u_{tt}^{-\frac{1}{k}}u_{tttt}\bigg]. \end{aligned}\tag{4-8}$$

Next we want to get an explicit formula for $(E_u)_{tt}$, which we build up to in stages. We first observe the preliminary computation

$$\begin{aligned} (1+\epsilon)(A_u)_t&=[u_{tt}^{-1}E_u+u_{tt}^{-1}\nabla u_t\otimes\nabla u_t]_t\\ &=-u_{tt}^{-2}u_{ttt}E_u+u_{tt}^{-1}(E_u)_t-u_{tt}^{-2}u_{ttt}\nabla u_t\otimes\nabla u_t\\ &\qquad+u_{tt}^{-1}\nabla u_{tt}\otimes\nabla u_t+u_{tt}^{-1}\nabla u_t\otimes\nabla u_{tt}. \end{aligned}\tag{4-9}$$

Next we compute that

$$\begin{aligned} [(E_u)_t]&=(1+\epsilon)u_{ttt}A_u+(1+\epsilon)u_{tt}(A_u)_t-\nabla u_{tt}\otimes\nabla u_t-\nabla u_t\otimes\nabla u_{tt}\\ &=(1+\epsilon)u_{ttt}A_u+(1+\epsilon)u_{tt}[\nabla^2u_t+\nabla u_t\otimes\nabla u+\nabla u\otimes\nabla u_t-\langle\nabla u_t,\nabla u\rangle g]\\ &\qquad-\nabla u_{tt}\otimes\nabla u_t-\nabla u_t\otimes\nabla u_{tt}. \end{aligned}$$

Next we have, using (4-9),

$$\begin{aligned}
[(E_u)_{tt}] &= (1+\epsilon)u_{tttt}A_u + 2(1+\epsilon)u_{ttt}(A_u)_t + (1+\epsilon)u_{tt}(A_u)_{tt} \\
&\qquad - \nabla u_{ttt}\otimes\nabla u_t - 2\nabla u_{tt}\otimes\nabla u_{tt} - \nabla u_t\otimes\nabla u_{ttt} \\
&= (1+\epsilon)u_{tttt}A_u + 2(1+\epsilon)u_{ttt}(A_u)_t \\
&\qquad + (1+\epsilon)u_{tt}\big[\nabla^2 u_{tt} + \nabla u_{tt}\otimes\nabla u + 2\nabla u_t\otimes\nabla u_t + \nabla u\otimes\nabla u_{tt} \\
&\qquad\qquad - |\nabla u_t|^2 g - \langle\nabla u,\nabla u_{tt}\rangle g\big] \\
&\qquad - \nabla u_{ttt}\otimes\nabla u_t - 2\nabla u_{tt}\otimes\nabla u_{tt} - \nabla u_t\otimes\nabla u_{ttt} \\
&= (1+\epsilon)u_{tttt}A_u \\
&\qquad + 2u_{ttt}\big[-u_{tt}^{-2}u_{ttt}E_u + u_{tt}^{-1}(E_u)_t - u_{tt}^{-2}u_{ttt}\nabla u_t\otimes\nabla u_t \\
&\qquad\qquad + u_{tt}^{-1}\nabla u_{tt}\otimes\nabla u_t + u_{tt}^{-1}\nabla u_t\otimes\nabla u_{tt}\big] \\
&\qquad + (1+\epsilon)u_{tt}\big[\nabla^2 u_{tt} + \nabla u_{tt}\otimes\nabla u + 2\nabla u_t\otimes\nabla u_t + \nabla u\otimes\nabla u_{tt} \\
&\qquad\qquad - |\nabla u_t|^2 g - \langle\nabla u,\nabla u_{tt}\rangle g\big] \\
&\qquad - \nabla u_{ttt}\otimes\nabla u_t - 2\nabla u_{tt}\otimes\nabla u_{tt} - \nabla u_t\otimes\nabla u_{ttt}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&k\sigma_k(E_u)^{\frac{k-1}{k}}u_{tt}^{1-k}\mathcal{F}^{ij}[(E_u)_{tt}]_{ij} \\
&\quad = u_{tt}^{1-k}\big\langle T_{k-1}(E_u), (1+\epsilon)u_{tttt}A_u - 2u_{tt}^{-2}u_{ttt}^2E_u + 2u_{tt}^{-1}u_{ttt}(E_u)_t \\
&\qquad\qquad - 2u_{tt}^{-2}u_{ttt}^2\nabla u_t\otimes\nabla u_t + 4u_{tt}^{-1}u_{ttt}\nabla u_{tt}\otimes\nabla u_t \\
&\qquad\qquad + (1+\epsilon)\big\{u_{tt}\nabla^2 u_{tt} + 2u_{tt}\nabla u_{tt}\otimes\nabla u + 2u_{tt}\nabla u_t\otimes\nabla u_t \\
&\qquad\qquad\qquad - u_{tt}|\nabla u_t|^2 g - u_{tt}\langle\nabla u,\nabla u_{tt}\rangle g\big\} \\
&\qquad\qquad - 2\nabla u_{ttt}\otimes\nabla u_t - 2\nabla u_{tt}\otimes\nabla u_{tt}\big\rangle \\
&\quad = \sum_{i=1}^{12} A_i.
\end{aligned}$$

Comparing against (4-7) yields

$$\begin{aligned}
\mathcal{L}u_{tt} &= A_1 + A_6 + A_7 + A_{10} + A_{11} + u_{tttt}u_{tt}^{1-k}\langle T_{k-1}(E), -(1+\epsilon)A_u + u_{tt}^{-1}\nabla u_t\otimes\nabla u_t\rangle \\
&\qquad + (1+\epsilon)^{k-1}u_{tt}^{-1}fu_{tttt} \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + u_{tttt}[u_{tt}^{-k}\langle T_{k-1}(E), -E\rangle + (1+\epsilon)^{k-1}u_{tt}^{-1}f] \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + u_{tttt}[ku_{tt}^{-k}\sigma_k(E) + (1+\epsilon)^{k-1}u_{tt}^{-1}f] \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + f(1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt}.
\end{aligned}$$

Hence, we obtain

$$
\begin{aligned}
\text{(4-10)}\quad & \mathcal{L}u_{tt} \\
&= k\sigma_k(E_u)^{\frac{k-1}{k}} u_{tt}^{1-k} \mathcal{F}^{ij}[(E_u)_{tt}]_{ij} \\
&\quad - u_{tt}^{1-k} \langle T_{k-1}(E_u), -2u_{tt}^{-2}u_{ttt}^2 E_u + 2u_{tt}^{-1}u_{ttt}(E_u)_t \\
&\qquad -2(1+\epsilon)u_{tt}^{-2}u_{ttt}^2 \nabla u_t \otimes \nabla u_t + 4u_{tt}^{-1}u_{ttt}\nabla u_{tt}\otimes\nabla u_t \\
&\qquad +2u_{tt}\nabla u_t\otimes\nabla u_t - (1+\epsilon)u_{tt}|\nabla u_t|^2 g - 2\nabla u_{tt}\otimes\nabla u_{tt}\rangle \\
&\quad + f(1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt} \\
&= k\sigma_k(E_u)^{\frac{k-1}{k}} u_{tt}^{1-k} \\
&\quad \times \Big[-\mathcal{F}^{ij,kl}[(E_u)_t]_{ij}[(E_u)_t]_{kl} \\
&\qquad + c_\epsilon \Big[(f^{\frac{1}{k}})_{tt} u_{tt}^{\frac{k-1}{k}} + 2\frac{k-1}{k}(f^{\frac{1}{k}})_t u_{tt}^{-\frac{1}{k}} u_{ttt} - \frac{1}{k}\Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} u_{ttt}^2 \\
&\qquad\qquad + \frac{k-1}{k} f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} u_{tttt} \Big]\Big] \\
&\quad + u_{tt}^{1-k}\langle T_{k-1}(E), 2u_{tt}^{-2}u_{ttt}^2 E_u - 2u_{tt}^{-1}u_{ttt}(E_u)_t \\
&\qquad +2(1+\epsilon)u_{tt}^{-2}u_{ttt}^2\nabla u_t\otimes\nabla u_t - 4u_{tt}^{-1}u_{ttt}\nabla u_{tt}\otimes\nabla u_t \\
&\qquad -2u_{tt}\nabla u_t\otimes\nabla u_t + (1+\epsilon)u_{tt}|\nabla u_t|^2 g + 2\nabla u_{tt}\otimes\nabla u_{tt}\rangle \\
&\quad + f(1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt} \\
&= \sum_{i=1}^{13} A_i.
\end{aligned}
$$

We now clean up some of the lower-order terms. In particular we express

$$
\begin{aligned}
k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k} &= k[f u_{tt}^{k-1}(1+\epsilon)^{k-1}]^{\frac{k-1}{k}} u_{tt}^{1-k} \\
&= k f^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k}.
\end{aligned}
$$

Then observe

$$
\begin{aligned}
A_2 &= (k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k})((1+\epsilon)^{\frac{k-1}{k}} (f^{\frac{1}{k}})_{tt} u_{tt}^{\frac{k-1}{k}}) \\
&= (k f^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k})((1+\epsilon)^{\frac{k-1}{k}} (f^{\frac{1}{k}})_{tt} u_{tt}^{\frac{k-1}{k}}) \\
&= (1+\epsilon)^{k-1} k f^{\frac{k-1}{k}} (f^{\frac{1}{k}})_{tt}.
\end{aligned}
$$

Next,

$$\begin{aligned} A_3 &= (k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k})\Big((1+\epsilon)^{\frac{k-1}{k}} 2\frac{k-1}{k}(f^{\frac{1}{k}})_t u_{tt}^{-\frac{1}{k}} u_{ttt}\Big) \\ &= (kf^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}}(1+\epsilon)^{(k-1)^2/k})\Big((1+\epsilon)^{\frac{k-1}{k}} 2\frac{k-1}{k}(f^{\frac{1}{k}})_t u_{tt}^{-\frac{1}{k}} u_{ttt}\Big) \\ &= 2(k-1)(1+\epsilon)^{k-1} u_{tt}^{-1} f^{\frac{k-1}{k}}(f^{\frac{1}{k}})_t u_{ttt}. \end{aligned}$$

Next,

$$\begin{aligned} A_4 &= (kf^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}}(1+\epsilon)^{(k-1)^2/k})\Big(-(1+\epsilon)^{\frac{k-1}{k}}\frac{1}{k}\Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} u_{ttt}^2\Big) \\ &= -(1+\epsilon)^{k-1}\Big(\frac{k-1}{k}\Big) f u_{tt}^{-2} u_{ttt}^2. \end{aligned}$$

Next note that

$$A_5 = k\sigma_k(E_u)^{\frac{k-1}{k}} u_{tt}^{1-k} c_\epsilon \frac{k-1}{k} f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} u_{tttt} = (k-1)(1+\epsilon)^{k-1} f u_{tt}^{-1} u_{tttt} = -A_{13}.$$

Also observe

$$\begin{aligned} A_6 &= u_{tt}^{1-k}\langle T_{k-1}(E), 2u_{tt}^{-2} u_{ttt}^2 E_u\rangle \\ &= 2k u_{tt}^{-1-k} u_{ttt}^2 \sigma_k(E) \\ &= 2k u_{tt}^{-1-k} u_{ttt}^2 [f u_{tt}^{k-1}(1+\epsilon)^{k-1}] \\ &= 2k(1+\epsilon)^{k-1} u_{tt}^{-2} u_{ttt}^2 f. \end{aligned}$$

Lastly,

$$\begin{aligned} A_7 &= -2u_{tt}^{1-k}\langle T_{k-1}(E), u_{tt}^{-1} u_{ttt}(E_u)_t\rangle \\ &= -2u_{tt}^{-k} u_{ttt}[\sigma_k(E)]_t \\ &= -2(1+\epsilon)^{k-1} u_{tt}^{-k} u_{ttt}[f u_{tt}^{k-1}]_t \\ &= -2(1+\epsilon)^{k-1} u_{tt}^{-k} u_{ttt}[f_t u_{tt}^{k-1} + (k-1) f u_{tt}^{k-2} u_{ttt}] \\ &= -2(1+\epsilon)^{k-1} u_{tt}^{-1} u_{ttt}[f_t + (k-1) f u_{tt}^{-1} u_{ttt}]. \end{aligned}$$

Inserting these simplifications into (4-9) yields the result. □

Proposition 4.14 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$\sup_{M\times[0,1]} u_{tt} \le C\epsilon^{-1}.$$

Proof Let's begin with a preliminary estimate for $\mathcal{L}u_{tt}$. Returning to Lemma 4.13 and considering the terms in order, one first observes by convexity of $\mathcal{F}$ that

$$-kf^{\frac{k}{k-1}} u_{tt}(1+\epsilon)^k \mathcal{F}^{ij,kl}[(E_u)_t]_{ij}[(E_u)_t]_{kl} \ge 0.$$

Also, by an application of the Cauchy–Schwarz inequality one has the matrix inequality

$$2u_{tt}^{-2}u_{ttt}^2\nabla u_t\otimes\nabla u_t-4u_{tt}^{-1}u_{ttt}\nabla u_{tt}\otimes\nabla u_t+2\nabla u_{tt}\otimes\nabla u_{tt}\geq 0.$$

Also, since u is an admissible solution we have

$$u_{tt}^{1-k}\langle T_{k-1}(E),u_{tt}|\nabla u_t|^2g\rangle=u_{tt}^{2-k}|\nabla u_t|^2\operatorname{tr}T_{k-1}(E)\geq 0.$$

Also we observe

$$(1+\epsilon)^{k-1}kf^{\frac{k-1}{k}}(f^{\frac{1}{k}})_{tt}\leq Cf^{\frac{k-1}{k}}[f^{\frac{1}{k}-1}f_{tt}+f^{\frac{1}{k}-2}f_t^2]\leq Cf.$$

Next,

$$\begin{aligned}2(k-1)(1+\epsilon)^{k-1}u_{tt}^{-1}f^{\frac{k-1}{k}}(f^{\frac{1}{k}})_tu_{ttt}&\leq Cf^{\frac{k-1}{k}}(f^{\frac{1}{k}-1}f_t)u_{tt}^{-1}u_{ttt}\\&\leq Cfu_{tt}^{-1}u_{ttt}\\&\leq C\delta^{-1}f+C\delta fu_{tt}^{-2}u_{ttt}^2.\end{aligned}$$

Also,

$$-2(1+\epsilon)^{k-1}u_{tt}^{-1}u_{ttt}f_t\leq Cfu_{tt}^{-1}u_{ttt}\leq C\delta^{-1}f+C\delta fu_{tt}^{-2}u_{ttt}^2.$$

Combining these estimates and choosing δ sufficiently small leads to the preliminary estimate

$$\mathcal{L}u_{tt}\geq -2u_{tt}^{2-k}\langle T_{k-1}(E),\nabla u_t\otimes\nabla u_t\rangle-Cf. \tag{4-11}$$

Similar considerations with the result of Lemma 4.10 lead to the preliminary estimate

$$\mathcal{L}u_t^2\geq -Cf+2fu_{tt}+2\epsilon u_{tt}^{2-k}\langle T_{k-1}(E),\nabla u_t\otimes\nabla u_t\rangle. \tag{4-12}$$

Now fix constants λ_i and let

$$\Phi=u_{tt}+\lambda_1\epsilon^{-1}u_t^2+\lambda_2t(t-1).$$

Choosing $\lambda_1\geq 1$, combining Lemma 4.4 with (4-11) and (4-12) yields

$$\begin{aligned}\mathcal{L}\Phi&\geq 2u_{tt}^{2-k}\langle T_{k-1}(E),(\lambda_1-1)\nabla u_t\otimes\nabla u_t\rangle-f(C+C\lambda_1\epsilon^{-1})+2\lambda_1\epsilon^{-1}fu_{tt}+\lambda_2fu_{tt}^{-1}\\&\geq f\big[(2\lambda_1\epsilon^{-1}-\delta(C+C\lambda_1\epsilon^{-1}))u_{tt}+(\lambda_2-\delta^{-1}(C+C\lambda_1\epsilon^{-1}))\big].\end{aligned}$$

If we now choose δ small above with respect to universal constants and then choose λ_2 large with respect to δ, we conclude

$$\mathcal{L}\Phi>0,$$

and hence Φ cannot have an interior maximum. The proposition follows. □

Lemma 4.15 *Given an admissible solution u of $(\star_{\epsilon,f})$, we have*

$$\begin{aligned}\mathcal{L}(\Delta u) = {} & -k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k} \mathcal{F}^{(ij),(kl)} \nabla_p (E_u)_{ij} \nabla_p (E_u)_{kl} \\ & + u_{tt}^{1-k} T_{k-1}(E)^{ij} \big\{ 2u_{tt}^{-2} |\nabla u_{tt}|^2 \nabla_i u_t \otimes \nabla_j u_t - 4u_{tt}^{-1} \nabla_p u_{tt} \nabla_i \nabla_p u_t \nabla_j u_t \\ & \qquad + 2\nabla_i \nabla_p u_t \nabla_j \nabla_p u_t - 2(1+\epsilon) u_{tt} \nabla_i \nabla_p u \nabla_j \nabla_p u \\ & \qquad + (1+\epsilon) u_{tt} |\nabla^2 u|^2 g_{ij} + u_{tt} \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) \big\} \\ & + k(1+\epsilon)^{k-1} f^{\frac{k-1}{k}} \Delta(f^{\frac{1}{k}}) - (1+\epsilon)^{k-1} \frac{2}{k} u_{tt}^{-1} \langle \nabla f, \nabla u_{tt} \rangle \\ & + (1+\epsilon)^{k-1} \Big(\frac{k+1}{k} \Big) f u_{tt}^{-2} |\nabla u_{tt}|^2. \end{aligned}$$

Proof To begin we compute, using (4-3),

$$\begin{aligned}\text{(4-13)} \quad \mathcal{L}(\Delta u) = {} & \\ & (1+\epsilon)^{k-1} u_{tt}^{-1} f \Delta u_{tt} \\ & + u_{tt}^{1-k} \big\langle T_{k-1}(E), \\ & \qquad (1+\epsilon) u_{tt} \big(\nabla^2 \Delta u + \nabla \Delta u \otimes \nabla u + \nabla u \otimes \nabla \Delta u - \langle \nabla \Delta u, \nabla u \rangle g \big) \\ & \qquad - \nabla \Delta u_t \otimes \nabla u_t - \nabla u_t \otimes \nabla \Delta u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \Delta u_{tt} \big\rangle. \end{aligned}$$

Next we differentiate the equation, which we rewrite as

$$c_\epsilon f^{\frac{1}{k}} u_{tt}^{\frac{k-1}{k}} = \sigma_k(E_u)^{\frac{1}{k}} =: \mathcal{F}(E_u).$$

Differentiating yields

$$c_\epsilon \nabla_p (f^{\frac{1}{k}}) u_{tt}^{\frac{k-1}{k}} + c_\epsilon \Big(\frac{k-1}{k} \Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} \nabla_p u_{tt} = \mathcal{F}^{ij} \nabla_p (E_u)_{ij}.$$

Differentiating again yields

$$\begin{aligned}& \mathcal{F}^{ij} (\Delta E_u)_{ij} + \mathcal{F}^{(ij),(kl)} \nabla_p (E_u)_{ij} \nabla_p (E_u)_{kl} \\ & \quad = c_\epsilon \Big[\Delta(f^{\frac{1}{k}}) u_{tt}^{\frac{k-1}{k}} + 2 \Big(\frac{k-1}{k} \Big) \langle \nabla(f^{\frac{1}{k}}), \nabla u_{tt} \rangle u_{tt}^{-\frac{1}{k}} \\ & \qquad\qquad - \frac{1}{k} \Big(\frac{k-1}{k} \Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} |\nabla u_{tt}|^2 + \Big(\frac{k-1}{k} \Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} \Delta u_{tt} \Big]. \end{aligned}$$

Next we have

$$\begin{aligned}\nabla_p (E_u)_{ij} & = \nabla_p [(1+\epsilon) u_{tt} (A_u)_{ij} - \nabla_i u_t \nabla_j u_t] \\ & = (1+\epsilon) \nabla_p u_{tt} (A_u)_{ij} + (1+\epsilon) u_{tt} \nabla_p (A_u)_{ij} - \nabla_p \nabla_i u_t \nabla_j u_t - \nabla_i u_t \nabla_p \nabla_j u_t. \end{aligned}$$

Differentiating again and commuting derivatives yields

$$\begin{aligned}(\Delta E_u)_{ij} &= (1+\epsilon)\Delta u_{tt}(A_u)_{ij} + 2(1+\epsilon)\nabla_p u_{tt}\nabla_p(A_u)_{ij} + (1+\epsilon)u_{tt}\Delta(A_u)_{ij} \\ &\quad - \nabla_i\Delta u_t\nabla_j u_t - \nabla_i u_t\nabla_j\Delta u_t - 2\nabla_i\nabla_p u_t\nabla_j\nabla_p u_t \\ &\quad - R_{ip}\nabla_p u_t\nabla_j u_t - R_{jp}\nabla_p u_t\nabla_i u_t.\end{aligned}$$

Differentiating the equation for the Schouten tensor yields

$$\nabla_p(A_u)_{ij} = \nabla_p A_{ij} + \nabla_p\nabla_i\nabla_j u + \nabla_i\nabla_p u\nabla_j u + \nabla_i u\nabla_j\nabla_p u - \tfrac{1}{2}\nabla_p|\nabla u|^2 g.$$

This implies

$$\begin{aligned}\Delta(A_u)_{ij} &= \Delta A_{ij} + \nabla_i\nabla_j\Delta u + \nabla_i\Delta u\nabla_j u + \nabla_i u\nabla_j\Delta u + 2\nabla_i\nabla_p u\nabla_j\nabla_p u \\ &\quad - |\nabla^2 u|^2 g_{ij} - \langle\nabla u,\nabla\Delta u\rangle g_{ij} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1).\end{aligned} \tag{4-14}$$

On the other hand it is also useful to express

$$\begin{aligned}(1+\epsilon)\nabla_p(A_u)_{ij} &= \nabla_p[u_{tt}^{-1}(E_u)_{ij} + u_{tt}^{-1}\nabla_i u_t\nabla_j u_t] \\ &= u_{tt}^{-1}\nabla_p(E_u)_{ij} - u_{tt}^{-2}(E_u)_{ij}\nabla_p u_{tt} - u_{tt}^{-2}\nabla_p u_{tt}\nabla_i u_t\nabla_j u_t \\ &\quad + u_{tt}^{-1}\nabla_i\nabla_p u_t\nabla_j u_t + u_{tt}^{-1}\nabla_i u_t\nabla_j\nabla_p u_t.\end{aligned}$$

Combining the above calculations yields

$$\begin{aligned}&\Delta(E_u)_{ij} \\ &= (1+\epsilon)\Delta u_{tt}(A_u)_{ij} \\ &\quad + 2\nabla_p u_{tt}\big[u_{tt}^{-1}\nabla_p(E_u)_{ij} - u_{tt}^{-2}(E_u)_{ij}\nabla_p u_{tt} - u_{tt}^{-2}\nabla_p u_{tt}\nabla_i u_t\nabla_j u_t \\ &\qquad + u_{tt}^{-1}\nabla_i\nabla_p u_t\nabla_j u_t + u_{tt}^{-1}\nabla_i u_t\nabla_j\nabla_p u_t\big] \\ &\quad + (1+\epsilon)u_{tt}\big[\nabla_i\nabla_j\Delta u + \nabla_i\Delta u\nabla_j u + \nabla_i u\nabla_j\Delta u + 2\nabla_i\nabla_p u\nabla_j\nabla_p u \\ &\qquad - |\nabla^2 u|^2 g_{ij} - \langle\nabla u,\nabla\Delta u\rangle g_{ij} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\big] \\ &\quad - \nabla_i\Delta u_t\nabla_j u_t - \nabla_i u_t\nabla_j\Delta u_t - 2\nabla_i\nabla_p u_t\nabla_j\nabla_p u_t \\ &= (1+\epsilon)\Delta u_{tt}(A_u)_{ij} + 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_p(E_u)_{ij} - 2u_{tt}^{-2}|\nabla u_{tt}|^2(E_u)_{ij} \\ &\quad - 2u_{tt}^{-2}|\nabla u_{tt}|^2\nabla_i u_t\otimes\nabla_j u_t + 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_i\nabla_p u_t\nabla_j u_t \\ &\quad + 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_j\nabla_p u_t\nabla_i u_t + (1+\epsilon)u_{tt}\nabla_i\nabla_j\Delta u \\ &\quad + (1+\epsilon)u_{tt}\nabla_i\Delta u\nabla_j u + (1+\epsilon)u_{tt}\nabla_i u\nabla_j\Delta u + 2(1+\epsilon)u_{tt}\nabla_i\nabla_p u\nabla_j\nabla_p u \\ &\quad - (1+\epsilon)u_{tt}|\nabla^2 u|^2 g_{ij} - (1+\epsilon)u_{tt}\langle\nabla u,\nabla\Delta u\rangle g_{ij} + u_{tt}\mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) \\ &\quad - \nabla_i\Delta u_t\nabla_j u_t - \nabla_i u_t\nabla_j\Delta u_t - 2\nabla_i\nabla_p u_t\nabla_j\nabla_p u_t.\end{aligned}$$

Thus,

$$
\begin{aligned}
&k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k}\mathcal{F}^{ij}(\Delta E_u)_{ij} \\
&\quad = u_{tt}^{1-k}\big\langle T_{k-1}(E), (1+\epsilon)\Delta u_{tt}(A_u)_{ij} + 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_p(E_u)_{ij} \\
&\qquad\qquad -2u_{tt}^{-2}|\nabla u_{tt}|^2(E_u)_{ij} - 2u_{tt}^{-2}|\nabla u_{tt}|^2\nabla_i u_t\otimes\nabla_j u_t \\
&\qquad\qquad +4u_{tt}^{-1}\nabla_p u_{tt}\nabla_i\nabla_p u_t\nabla_j u_t + (1+\epsilon)u_{tt}\nabla_i\nabla_j\Delta u \\
&\qquad\qquad +2(1+\epsilon)u_{tt}\nabla_i\Delta u\nabla_j u + 2(1+\epsilon)u_{tt}\nabla_i\nabla_p u\nabla_j\nabla_p u \\
&\qquad\qquad -(1+\epsilon)u_{tt}|\nabla^2 u|^2 g_{ij} - (1+\epsilon)u_{tt}\langle\nabla u,\nabla\Delta u\rangle g_{ij} \\
&\qquad\qquad -2\nabla_i\Delta u_t\nabla_j u_t - 2\nabla_i\nabla_p u_t\nabla_j\nabla_p u_t + u_{tt}\mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\big\rangle \\
&\quad = \sum_{i=1}^{13} A_i.
\end{aligned}
$$

Comparing this against (4-13) yields

$$
\begin{aligned}
\text{(4-15)}\quad \mathcal{L}(\Delta u) &= A_1 + A_6 + A_7 + A_{10} + A_{11} \\
&\qquad + u_{tt}^{1-k}\Delta u_{tt}\langle T_{k-1}(E), -(1+\epsilon)A_u + \nabla u_t\otimes\nabla u_t\rangle \\
&\qquad + (1+\epsilon)^{k-1}u_{tt}^{-1} f\Delta u_{tt} \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + u_{tt}^{-k}\Delta u_{tt}\langle T_{k-1}(E), -E\rangle \\
&\qquad + (1+\epsilon)^{k-1}u_{tt}^{-1} f\Delta u_{tt} \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + \Delta u_{tt}[-ku_{tt}^{-k}\sigma_k(E) + (1+\epsilon)^{k-1}u_{tt}^{-1} f] \\
&= A_1 + A_6 + A_7 + A_{10} + A_{11} + (1-k)(1+\epsilon)^{k-1}u_{tt}^{-1} f\Delta u_{tt}.
\end{aligned}
$$

Hence, collecting these calculations yields

$$
\begin{aligned}
\mathcal{L}(\Delta u) &= k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k}\mathcal{F}(\Delta E_u)_{ij} \\
&\quad - u_{tt}^{1-k}\big\langle T_{k-1}(E), 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_p(E_u)_{ij} - 2u_{tt}^{-2}|\nabla u_{tt}|^2(E_u)_{ij} \\
&\qquad - 2u_{tt}^{-2}|\nabla u_{tt}|^2\nabla_i u_t\otimes\nabla_j u_t + 4u_{tt}^{-1}\nabla_p u_{tt}\nabla_i\nabla_p u_t\nabla_j u_t \\
&\qquad + 2(1+\epsilon)u_{tt}\nabla_i\nabla_p u\nabla_j\nabla_p u - (1+\epsilon)u_{tt}|\nabla^2 u|^2 g_{ij} \\
&\qquad - 2\nabla_i\nabla_p u_t\nabla_j\nabla_p u_t + u_{tt}\mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\big\rangle \\
&\quad + (1-k)(1+\epsilon)^{k-1}u_{tt}^{-1} f\Delta u_{tt}
\end{aligned}
$$

$$
\begin{aligned}
&= -k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k} \mathcal{F}^{(ij),(kl)} \nabla_p (E_u)_{ij} \nabla_p (E_u)_{kl} \\
&\quad + c_\epsilon k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k} \\
&\qquad \times \Big[\Delta(f^{\frac{1}{k}}) u_{tt}^{\frac{k-1}{k}} + 2\Big(\frac{k-1}{k}\Big) \langle \nabla(f^{\frac{1}{k}}), \nabla u_{tt} \rangle u_{tt}^{-\frac{1}{k}} \\
&\qquad\qquad - \frac{1}{k}\Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} |\nabla u_{tt}|^2 + \Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} \Delta u_{tt} \Big] \\
&\quad + u_{tt}^{1-k} \big\langle T_{k-1}(E), -2u_{tt}^{-1} \nabla_p u_{tt} \nabla_p (E_u)_{ij} + 2u_{tt}^{-2} |\nabla u_{tt}|^2 (E_u)_{ij} \\
&\qquad + 2u_{tt}^{-2} |\nabla u_{tt}|^2 \nabla_i u_t \otimes \nabla_j u_t - 4u_{tt}^{-1} \nabla_p u_{tt} \nabla_i \nabla_p u_t \nabla_j u_t \\
&\qquad + 2\nabla_i \nabla_p u_t \nabla_j \nabla_p u_t - 2(1+\epsilon) u_{tt} \nabla_i \nabla_p u \nabla_j \nabla_p u \\
&\qquad + (1+\epsilon) u_{tt} |\nabla^2 u|^2 g_{ij} + u_{tt} \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) \big\rangle \\
&\quad + (1-k)(1+\epsilon)^{k-1} u_{tt}^{-1} f \Delta u_{tt} \\
&= \sum_{i=1}^{14} A_i.
\end{aligned}
$$

Now we simplify:

$$
\begin{aligned}
A_2 &= (k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k})(c_\epsilon u_{tt}^{\frac{k-1}{k}} \Delta(f^{\frac{1}{k}})) \\
&= (k f^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k})((1+\epsilon)^{\frac{k-1}{k}} u_{tt}^{\frac{k-1}{k}} \Delta(f^{\frac{1}{k}})) \\
&= k(1+\epsilon)^{k-1} f^{\frac{k-1}{k}} \Delta(f^{\frac{1}{k}}).
\end{aligned}
$$

Next,

$$
\begin{aligned}
A_3 &= (k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k}) \Big(2c_\epsilon \Big(\frac{k-1}{k}\Big) \langle \nabla(f^{\frac{1}{k}}), \nabla u_{tt} \rangle u_{tt}^{-\frac{1}{k}} \Big) \\
&= (k f^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k}) \Big(2(1+\epsilon)^{\frac{k-1}{k}} \Big(\frac{k-1}{k}\Big) \langle \nabla(f^{\frac{1}{k}}), \nabla u_{tt} \rangle u_{tt}^{-\frac{1}{k}} \Big) \\
&= 2(1+\epsilon)^{k-1} (k-1) f^{\frac{k-1}{k}} u_{tt}^{-1} \langle \nabla(f^{\frac{1}{k}}), \nabla u_{tt} \rangle \\
&= (1+\epsilon)^{k-1} \Big(2 - \frac{2}{k}\Big) u_{tt}^{-1} \langle \nabla f, \nabla u_{tt} \rangle.
\end{aligned}
$$

Next,

$$
\begin{aligned}
A_4 &= -(k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k}) \Big(c_\epsilon \frac{1}{k} \Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} |\nabla u_{tt}|^2 \Big) \\
&= -(k f^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k}) \Big((1+\epsilon)^{\frac{k-1}{k}} \frac{1}{k} \Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1+k}{k}} |\nabla u_{tt}|^2 \Big) \\
&= -(1+\epsilon)^{k-1} \Big(\frac{k-1}{k}\Big) f u_{tt}^{-2} |\nabla u_{tt}|^2.
\end{aligned}
$$

Next,

$$
\begin{aligned}
A_5 &= (k\sigma_k(E)^{\frac{k-1}{k}} u_{tt}^{1-k})\Big(c_\epsilon\Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} \Delta u_{tt}\Big) \\
&= (kf^{\frac{k-1}{k}} u_{tt}^{\frac{1-k}{k}} (1+\epsilon)^{(k-1)^2/k})\Big((1+\epsilon)^{\frac{k-1}{k}}\Big(\frac{k-1}{k}\Big) f^{\frac{1}{k}} u_{tt}^{-\frac{1}{k}} \Delta u_{tt}\Big) \\
&= (k-1)(1+\epsilon)^{k-1} f u_{tt}^{-1} \Delta u_{tt} \\
&= -A_{14}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
A_6 &= -2u_{tt}^{1-k} \nabla_p u_{tt} \langle T_{k-1}(E), u_{tt}^{-1} \nabla_p (E_u)_{ij}\rangle \\
&= -2u_{tt}^{-k} \nabla_p u_{tt} \nabla_p \sigma_k(E) \\
&= -2(1+\epsilon)^{k-1} u_{tt}^{-k} \nabla_p u_{tt} \nabla_p [f u_{tt}^{k-1}] \\
&= -2(1+\epsilon)^{k-1} u_{tt}^{-1} \langle \nabla f, \nabla u_{tt}\rangle - 2(1+\epsilon)^{k-1}(k-1) f u_{tt}^{-2} |\nabla u_{tt}|^2.
\end{aligned}
$$

Lastly,

$$
\begin{aligned}
A_7 &= 2u_{tt}^{1-k} u_{tt}^{-2} |\nabla u_{tt}|^2 \langle T_{k-1}(E), E\rangle \\
&= 2k u_{tt}^{1-k} u_{tt}^{-2} |\nabla u_{tt}|^2 \sigma_k(E) \\
&= 2k(1+\epsilon)^{k-1} f u_{tt}^{-2} |\nabla u_{tt}|^2.
\end{aligned}
$$

Collecting these simplifications yields the result. □

Proposition 4.16 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$
\sup_{M\times[0,1]} \Delta u \le C\epsilon^{-1}.
$$

Proof We begin with a preliminary estimate for $\mathcal{L}\Delta u$. Returning to Lemma 4.15 and considering the terms in order, one first observes by convexity of $\mathcal{F}$ that

$$
-kf^{\frac{k}{k-1}} u_{tt}(1+\epsilon)^k \mathcal{F}^{ij,kl} [\nabla_p(E_u)]_{ij} [\nabla_p(E_u)]_{kl} \ge 0.
$$

Also, by an application of the Cauchy–Schwarz inequality one has the matrix inequality

$$
2u_{tt}^{-2} |\nabla u_{tt}|^2 \nabla_i u_t \nabla_j u_t - 4u_{tt}^{-1} \nabla_p u_{tt} \nabla_i \nabla_p u_t \nabla_j u_t + 2\nabla_i \nabla_p u_t \nabla_j \nabla_p u_t \ge 0.
$$

Also we observe

$$
(1+\epsilon)^{k-1} k f^{\frac{k-1}{k}} \Delta(f^{\frac{1}{k}}) \le C f^{\frac{k-1}{k}} [f^{\frac{1}{k}-1} \Delta f + f^{\frac{1}{k}-2} |\nabla f|^2] \le Cf.
$$

Next,

$$
-\frac{2}{k}(1+\epsilon)^{k-1} u_{tt}^{-1} \langle \nabla f, \nabla u_{tt}\rangle \le Cf u_{tt}^{-1} |\nabla u_{tt}| \le C\delta^{-1} f + C\delta u_{tt}^{-2} |\nabla u_{tt}|^2.
$$

Combining these estimates and choosing δ sufficiently small leads to the preliminary estimate

$$\text{(4-16)}\quad \begin{aligned}\mathcal{L}\Delta u \geq -2(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E), \nabla_i\nabla_p u\nabla_j\nabla_p u\rangle \\ + u_{tt}^{2-k}\langle T_{k-1}(E), |\nabla^2 u|^2 g + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\rangle - Cf.\end{aligned}$$

Similar considerations applied to Lemma 4.11 yield

$$\text{(4-17)}\quad \mathcal{L}|\nabla u|^2 \geq 2\epsilon u_{tt}^{2-k} T_{k-1}(E)^{jk}\nabla_i\nabla_j u\nabla_i\nabla_k u - Cf - u_{tt}^{2-k}\langle T_{k-1}(E), \mathcal{O}(1)\rangle.$$

Now fix a constant $\lambda \in \mathbb{R}$ and consider

$$\Phi = \Delta u + \epsilon^{-1}[(1+\epsilon)|\nabla u|^2 + u_t^2 + \lambda t(t-1)].$$

Combining Lemma 4.4 with (4-12), (4-16) and (4-17) yields

$$\begin{aligned}\mathcal{L}\Phi \geq u_{tt}^{2-k}\langle T_{k-1}(E), |\nabla^2 u|^2 g + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) + \epsilon^{-1}\mathcal{O}(1)\rangle \\ - C\epsilon^{-1} f + 2\epsilon^{-1} f u_{tt} + \lambda\epsilon^{-1} f u_{tt}^{-1}.\end{aligned}$$

First we observe that at a sufficiently large maximum of Φ, the existing a priori estimates imply that Δu is also large. In particular, at a maximum for Φ where $|\nabla^2 u| \geq C\epsilon^{-\frac{1}{2}}$ we obtain

$$|\nabla^2 u|^2 g + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) + \epsilon^{-1}\mathcal{O}(1) \geq \tfrac{1}{2}|\nabla^2 u|^2 g,$$

and hence, since u is an admissible solution, we have

$$\begin{aligned}u_{tt}^{2-k}\langle T_{k-1}(E), |\nabla^2 u|^2 g + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) + \epsilon^{-1}\mathcal{O}(1)\rangle \\ \geq \tfrac{1}{2} u_{tt}^{2-k}|\nabla^2 u|^2 \operatorname{tr} T_{k-1}(E) \\ \geq 0.\end{aligned}$$

But then we can estimate

$$C\epsilon^{-1} f \leq \epsilon^{-1} f u_{tt} + C\epsilon^{-1} f u_{tt}^{-1}.$$

Hence, choosing λ sufficiently large we obtain, at a sufficiently large maximum for Φ which satisfies $\Delta u \geq C\epsilon^{-\frac{1}{2}}$, that one has

$$\mathcal{L}\Phi > 0,$$

a contradiction. The a priori estimate for Δu follows directly. □

4.4 Boundary estimates

By Proposition 4.8 we already have the boundary estimate

$$\sup_{M\times\{0,1\}} [|u|+|u_t|+|\nabla u|] \leq C.$$

In this section we prove boundary estimates for second-order derivatives. Our barrier function methods have a parallel with work of Guan [17].

Proposition 4.17 *Given an admissible solution u to $(\star_{\epsilon,f})$, one has*

$$\sup_{M\times\{0,1\}} [|u_{tt}|+|\nabla u_t|+|\nabla^2 u|] \leq C.$$

Proof A bound for $|\nabla^2 u|$ is immediate. If we can prove a bound for the "mixed" term $|\nabla u_t|$, then restricting the equation for u to $t=0$ we have

$$\begin{aligned}(1+\epsilon)u_{tt}(\cdot\,,0)\sigma_k(A_{u(\cdot\,,0)}) &= \langle T_{k-1}(A_{u(\cdot\,,0)}), \nabla u_t(\cdot\,,0)\otimes\nabla u_t(\cdot\,,0)\rangle + f\\ &\leq C_1(1+|\nabla u_0|^2+|\nabla^2 u_0|)|\nabla u_t(\cdot\,,0)|^2 + C_2.\end{aligned}$$

Since u_0 is admissible,

$$\sigma_k(A_{u(\cdot\,,0)}) = \sigma_k(A_{u_0}) \geq \delta_0 > 0,$$

and it follows that

$$\sup_M u_{tt}(\cdot\,,0) \leq C_0(1+\sup_M |\nabla u_t(\cdot\,,0)|^2),$$

where C_0 depends on the second-order spacial derivatives of u_0. The same argument gives a corresponding bound for $u_{tt}(\cdot\,,1)$ in terms of the mixed derivative $|\nabla u_t(\cdot\,,1)|$.

To prove a bound on ∇u_t we consider the auxiliary function $\Psi\colon M\times[0,\tau]\to\mathbb{R}$, where $0<\tau<1$ will be chosen later,

$$\Psi = |\nabla(u-u_0)| + [e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda t(t-1),$$

where λ, Λ and Υ are constants yet to be determined. By making an appropriate choice of these constants, we claim that Ψ attains a nonpositive maximum on the boundary of $M\times[0,\tau]$. Assuming for the moment this is true, let us see how a bound for ∇u_t follows.

Choose a point $x_0 \in M$, and a unit tangent vector $X \in T_{x_0}M$. Let $\{x^i\}$ be a local coordinate system with $X = \partial/\partial x^1$ at x_0. Then

$$\begin{aligned}\frac{\partial}{\partial x^1}(u(x,t)-u_0(x)) + [e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda t(t-1)& \\ \leq |\nabla(u-u_0)(x,t)| + [e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda t(t-1)& \\ \leq 0.&\end{aligned}$$

Therefore,

$$\begin{aligned}0 &\geq \lim_{t\to 0+}\frac{1}{t}\left\{\frac{\partial}{\partial x^1}u(x,t) - \frac{\partial}{\partial x^1}u_0(x) + [e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda t(t-1)\right\} \\ &= \frac{\partial}{\partial x^1}u_t(x_0,0) + \frac{1}{t}[e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda(t-1).\end{aligned}$$

Since u_t is bounded, an upper bound on $\partial u_t/\partial x^1$ follows. Since $X = \partial/\partial x^1$ was arbitrary, we obtain a bound on $|\nabla u_t(x,0)|$.

To see that such a choice of λ, Λ, Υ and τ are possible, we first note that

$$\Psi(x,0) = 0.$$

Since $|\nabla u|$ is bounded,

$$\begin{aligned}\Psi(x,\tau) &= |\nabla u(x,\tau) - \nabla u_0(x)| + [e^{\lambda(u_0(x)-u(x,\tau)+\Upsilon)} - e^{\lambda\Upsilon}] + \Lambda\tau(\tau-1) \\ &\leq C_1 + |e^{\lambda(u_0(x)-u(x,\tau)+\Upsilon)} - e^{\lambda\Upsilon}| + \Lambda\tau(\tau-1).\end{aligned}$$

Since $|u_t|$ is also bounded,

$$|e^{\lambda(u_0(x)-u(x,\tau)+\Upsilon)} - e^{\lambda\Upsilon}| \leq C_2\lambda e^{C_2\lambda\tau+\Upsilon},$$

hence, if $0 < \tau < \frac{1}{2}$,

$$\Psi(x,\tau) \leq C_1 + C_2\tau\lambda e^{C_2\lambda\tau+\Upsilon} - \Lambda\tau(1-\tau) \leq C_1 + \left(C_2\lambda e^{\frac{1}{2}C_2\lambda+\Upsilon} - \tfrac{1}{2}\Lambda\right)\tau.$$

Therefore, if Λ is chosen large enough (depending on τ, C_1, C_2, λ and Υ), then

$$\Psi(x,\tau) \leq 0.$$

We conclude that $\Psi \leq 0$ on $\partial(M\times[0,\tau])$.

Assume the maximum of Ψ is attained at a point (x_0,t_0) which is interior (ie $0 < t_0 < \tau$). Let

$$\eta = \frac{\nabla(u-u_0)(x_0,t_0)}{|\nabla(u-u_0)(x_0,t_0)|}.$$

We can extend η locally via parallel transport along radial geodesics based at x_0. By construction,

$$\nabla\eta(x_0)=0,\quad |\nabla^2\eta(x_0)|\le C(g). \tag{4-18}$$

By using a cut-off function, we can assume η is globally defined and satisfies

$$|\eta|\le 1,$$

with $|\eta|=1$ in a neighborhood of x_0.

Define

$$H=\eta^\alpha\nabla_\alpha(u-u_0)+[e^{\lambda(u_0-u+\Upsilon)}-e^{\lambda\Upsilon}]+\Lambda t(t-1).$$

Since $|\eta|\le 1$,

$$H(x,t)\le\Psi(x,t),$$

and the max of H is attained at (x_0,t_0). Therefore,

$$\mathcal{L}H(x_0,t_0)\le 0.$$

To compute $\mathcal{L}H(x_0,t_0)$, let $\phi=\eta^\alpha\nabla_\alpha(u-u_0)$. Using (4-18), at (x_0,t_0) we have

$$\phi_t=\eta^\alpha\nabla_\alpha u_t,\quad \phi_{tt}=\eta^\alpha\nabla_\alpha u_{tt},\quad \nabla_k\phi_t=\eta^\alpha\nabla_k\nabla_\alpha u_t.$$

Also at (x_0,t_0),

$$\begin{aligned}\nabla_k\phi&=\eta^\alpha\nabla_k\nabla_\alpha(u-u_0)=\eta^\alpha\nabla_k\nabla_\alpha u+O(1),\\ \nabla_k\nabla_\ell\phi&=\nabla_k\nabla_\ell\eta^\alpha\nabla_\alpha(u-u_0)+\eta^\alpha\nabla_k\nabla_\ell\nabla_\alpha(u-u_0)=\eta^\alpha\nabla_k\nabla_\ell\nabla_\alpha u+O(1).\end{aligned}$$

Therefore, by the formula in (4-3), at (x_0,t_0) we have

$$\begin{aligned}\mathcal{L}\phi&=(1+\epsilon)^{k-1}u_{tt}^{-1}f\eta^\alpha\nabla_\alpha u_{tt}\\ &\quad+u_{tt}^{1-k}T_{k-1}(E_u^\epsilon)_{k\ell}\\ &\qquad\times\Big\{(1+\epsilon)u_{tt}\big[\eta^\alpha\nabla_k\nabla_\ell\nabla_\alpha u+\eta^\alpha\nabla_k\nabla_\alpha u\nabla_\ell u+\eta^\alpha\nabla_k u\nabla_\ell\nabla_\alpha u\\ &\qquad\qquad-(\eta^\alpha\nabla_m\nabla_\alpha v\nabla_m u)g_{k\ell}+O(1)g_{k\ell}\big]\\ &\qquad\qquad-\eta^\alpha\nabla_k\nabla_\alpha u_t\nabla_\ell u_t-\eta^\alpha\nabla_k u_t\nabla_\ell\nabla_\alpha u_t+\frac{\eta^\alpha\nabla_\alpha u_{tt}}{u_{tt}}\nabla_k u_t\nabla_\ell u_t\Big\}\\ &\ge\eta^\alpha\mathcal{L}\nabla_\alpha u-Cu_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon).\end{aligned}$$

Using the identity (4-6), we conclude

$$\mathcal{L}\phi\ge\langle\nabla f,\eta\rangle-Cu_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon)\ge -Cf-Cu_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon),$$

where the constants depend on $\max_M|\nabla f|/f$.

Next, we use Lemma 4.5 to calculate

$$
\begin{aligned}
\text{(4-19)}\quad \mathcal{L}(u-u_0)
&= (k+1)(1+\epsilon)^{k-1}f \\
&\quad + (1+\epsilon)u_{tt}^{2-k}\big\langle T_{k-1}(E_u^\epsilon), -A+\nabla u\otimes\nabla u-\tfrac{1}{2}|\nabla u|^2 g\big\rangle \\
&\quad - (1+\epsilon)u_{tt}^{2-k}\big\langle T_{k-1}(E_u^\epsilon), \nabla^2 u_0+\nabla u_0\otimes\nabla u+\nabla u\otimes\nabla u_0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \langle\nabla u_0,\nabla u\rangle g\big\rangle \\
&= (k+1)(1+\epsilon)^{k-1}f-(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), A+\nabla^2 u_0\rangle \\
&\quad + (1+\epsilon)u_{tt}^{2-k}\big[\langle T_{k-1}(E_u^\epsilon), \nabla u\otimes\nabla u\rangle-\tfrac{1}{2}\operatorname{tr} T_{k-1}(E_u^\epsilon)|\nabla u|^2\big] \\
&\quad - (1+\epsilon)u_{tt}^{2-k}[2\langle T_{k-1}(E_u^\epsilon), \nabla u\otimes\nabla u_0\rangle-\operatorname{tr} T_{k-1}(E_u^\epsilon)\langle\nabla u,\nabla u_0\rangle] \\
&= (k+1)(1+\epsilon)^{k-1}f \\
&\quad + (1+\epsilon)u_{tt}^{2-k}\big\langle T_{k-1}(E_u^\epsilon), -A_{u_0}+\nabla u_0\otimes\nabla u_0-\tfrac{1}{2}|\nabla u_0|^2 g\big\rangle \\
&\quad + (1+\epsilon)u_{tt}^{2-k}\big[\langle T_{k-1}(E_u^\epsilon), \nabla u\otimes\nabla u-2\nabla u\otimes\nabla u_0\rangle \\
&\qquad\qquad\qquad\qquad + \operatorname{tr} T_{k-1}(E_u^\epsilon)\big(-\tfrac{1}{2}|\nabla u|^2+\langle\nabla u,\nabla u_0\rangle\big)\big] \\
&= (k+1)(1+\epsilon)^{k-1}f \\
&\quad + (1+\epsilon)u_{tt}^{2-k} \\
&\qquad \times\big[-\langle T_{k-1}(E_u^\epsilon), A_{u_0}\rangle+\langle T_{k-1}(E_u^\epsilon), \nabla(u-u_0)\otimes\nabla(u-u_0)\rangle \\
&\qquad\qquad\qquad\qquad - \tfrac{1}{2}\operatorname{tr} T_{k-1}(E_u^\epsilon)|\nabla(u-u_0)|^2\big].
\end{aligned}
$$

Taking $v=e^{\lambda(u_0-u+\Upsilon)}-e^{\lambda\Upsilon}$ in Lemma 4.2, we also have

$$
\begin{aligned}
&\mathcal{L}(e^{\lambda(u_0-u+\Upsilon)}-e^{\lambda\Upsilon}) \\
&= e^{\lambda(u_0-u+\Upsilon)}\bigg\{(1+\epsilon)^{k-1}fu_{tt}^{-1}[-\lambda u_{tt}+\lambda^2u_t^2] \\
&\qquad +u_{tt}^{1-k}\bigg\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)u_{tt}\big[\lambda\nabla^2(u_0-u)+\lambda^2\nabla(u_0-u)\otimes\nabla(u_0-u) \\
&\qquad\qquad\qquad\qquad +\lambda\nabla(u_0-u)\otimes\nabla u+\lambda\nabla u\otimes\nabla(u_0-u) \\
&\qquad\qquad\qquad\qquad -\lambda\langle\nabla(u_0-u),\nabla u\rangle g\big] \\
&\qquad\qquad +\lambda\nabla u_t\otimes\nabla u_t+\lambda^2u_t\nabla(u_0-u)\otimes\nabla u_t \\
&\qquad\qquad +\lambda^2u_t\nabla u_t\otimes\nabla(u_0-u)+\lambda^2\frac{u_t^2}{u_{tt}}\nabla u_t\otimes\nabla u_t\bigg\rangle\bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&= -\lambda e^{\lambda(u_0-u+\Upsilon)}\mathcal{L}(u-u_0)\\
&\quad +\lambda^2 e^{\lambda(u_0-u+\Upsilon)}\\
&\qquad \times\biggl\{(1+\epsilon)^{k-1} f\frac{u_t^2}{u_{tt}}\\
&\qquad\qquad +u_{tt}^{2-k}\biggl\langle T_{k-1}(E_u^{\epsilon}), (1+\epsilon)\nabla(u-u_0)\otimes\nabla(u-u_0)+\frac{u_t}{u_{tt}}\nabla(u_0-u)\otimes\nabla u_t\\
&\qquad\qquad\qquad +\frac{u_t}{u_{tt}}\nabla u_t\otimes\nabla(u_0-u)+\frac{u_t^2}{u_{tt}^2}\nabla u_t\otimes\nabla u_t\biggr\rangle\biggr\}.
\end{aligned}
$$

We can estimate the term in braces as follows:

$$
\begin{aligned}
&(1+\epsilon)^{k-1} f\frac{u_t^2}{u_{tt}}+u_{tt}^{2-k}\biggl\langle T_{k-1}(E_u^{\epsilon}), (1+\epsilon)\nabla(u-u_0)\otimes\nabla(u-u_0)+\frac{u_t}{u_{tt}}\nabla(u_0-u)\otimes\nabla u_t\\
&\qquad\qquad +\frac{u_t}{u_{tt}}\nabla u_t\otimes\nabla(u_0-u)+\frac{u_t^2}{u_{tt}^2}\nabla u_t\otimes\nabla u_t\biggr\rangle\\
&\quad \geq (1+\epsilon)^{k-1} f\frac{u_t^2}{u_{tt}}+u_{tt}^{2-k}\biggl\langle T_{k-1}(E_u^{\epsilon}), \frac{1+\epsilon}{2}\nabla(u-u_0)\otimes\nabla(u-u_0)-\frac{u_t^2}{u_{tt}^2}\nabla u_t\otimes\nabla u_t\biggr\rangle.
\end{aligned}
$$

Using Lemma 2.5 and the regularized equation, the final term above can be rewritten:

$$
\begin{aligned}
u_{tt}^{2-k}&\biggl\langle T_{k-1}(E_u^{\epsilon}), -\frac{u_t^2}{u_{tt}^2}\nabla u_t\otimes\nabla u_t\biggr\rangle\\
&= -u_{tt}^{-k}u_t^2\langle T_{k-1}((1+\epsilon)u_{tt}A_u-\nabla u_t\otimes\nabla u_t), \nabla u_t\otimes\nabla u_t\rangle\\
&= -u_{tt}^{-k}u_t^2\langle T_{k-1}((1+\epsilon)u_{tt}A_u), \nabla u_t\otimes\nabla u_t\rangle\\
&= -(1+\epsilon)^{k-1}u_{tt}^{-1}u_t^2\langle T_{k-1}(A_u), \nabla u_t\otimes\nabla u_t\rangle\\
&= -(1+\epsilon)^{k-1}u_{tt}^{-1}u_t^2\{(1+\epsilon)u_{tt}\sigma_k(A_u)-f\}\\
&= -(1+\epsilon)^{k}u_t^2\sigma_k(A_u)+(1+\epsilon)^{k-1} f\frac{u_t^2}{u_{tt}}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\text{(4-20)}\quad &\mathcal{L}(e^{\lambda(u_0-u+\Upsilon)}-e^{\lambda\Upsilon})\\
&\geq -\lambda e^{\lambda(u_0-u+\Upsilon)}\mathcal{L}(u-u_0)\\
&\quad +\lambda^2 e^{\lambda(u_0-u+\Upsilon)}\biggl\{2(1+\epsilon)^{k-1} f\frac{u_t^2}{u_{tt}}-(1+\epsilon)^k u_t^2\sigma_k(A_u)\\
&\qquad\qquad +u_{tt}^{2-k}\biggl\langle T_{k-1}(E_u^{\epsilon}), \frac{1+\epsilon}{2}\nabla(u-u_0)\otimes\nabla(u-u_0)\biggr\rangle\biggr\}.
\end{aligned}
$$

Also, by (4-19),

$$
\begin{aligned}
\text{(4-21)}\quad -\lambda\mathcal{L}(u-u_0) = -\lambda(k+1)(1+\epsilon)^{k-1}f &+ \lambda(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), A_{u_0}\rangle \\
&+ u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), -\lambda(1+\epsilon)\nabla(u-u_0)\otimes\nabla(u-u_0)\rangle \\
&+ \tfrac{1}{2}(1+\epsilon)\lambda u_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u^\epsilon)|\nabla(u-u_0)|^2.
\end{aligned}
$$

Combining (4-20) and (4-21), we get

$$
\begin{aligned}
&\mathcal{L}(e^{\lambda(u_0-u+\Upsilon)} - e^{\lambda\Upsilon}) \\
&\geq e^{\lambda(u_0-u+\Upsilon)}\biggl\{-\lambda(k+1)(1+\epsilon)^{k-1}f + 2\lambda^2(1+\epsilon)^{k-1}f\frac{u_t^2}{u_{tt}} - \lambda^2(1+\epsilon)^k u_t^2\sigma_k(A_u) \\
&\qquad + u_{tt}^{2-k}\bigl\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)\bigl(\tfrac{1}{2}\lambda^2-\lambda\bigr)\nabla(u-u_0)\otimes\nabla(u-u_0)\bigr\rangle \\
&\qquad + \tfrac{1}{2}(1+\epsilon)\lambda u_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u^\epsilon)|\nabla(u-u_0)|^2 \\
&\qquad\qquad + \lambda(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), A_{u_0}\rangle\biggr\}.
\end{aligned}
$$

Next, using Lemma 4.4, we have

$$
\mathcal{L}(\Lambda t(1-t)) = 2\Lambda(1+\epsilon)^k\sigma_k(A_u).
$$

Combing the above, we conclude that at an interior maximum of H,

$$
\begin{aligned}
\mathcal{L}H \geq -Cf &- Cu_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u^\epsilon) + 2\Lambda(1+\epsilon)^k\sigma_k(A_u) \\
&+ e^{\lambda(u_0-u+\Upsilon)} \\
&\quad\times\biggl\{-\lambda(k+1)(1+\epsilon)^{k-1}f + 2\lambda^2(1+\epsilon)^{k-1}f\frac{u_t^2}{u_{tt}} - \lambda^2(1+\epsilon)^k u_t^2\sigma_k(A_u) \\
&\qquad + u_{tt}^{2-k}\bigl\langle T_{k-1}(E_u^\epsilon), (1+\epsilon)\bigl(\tfrac{1}{2}\lambda^2-\lambda\bigr)\nabla(u-u_0)\otimes\nabla(u-u_0)\bigr\rangle \\
&\qquad + \tfrac{1}{2}(1+\epsilon)\lambda u_{tt}^{2-k}\operatorname{tr} T_{k-1}(E_u^\epsilon)|\nabla(u-u_0)|^2 \\
&\qquad\qquad + \lambda(1+\epsilon)u_{tt}^{2-k}\langle T_{k-1}(E_u^\epsilon), A_{u_0}\rangle\biggr\}.
\end{aligned}
$$

Now note that since the cone Γ_k^+ is open and M is compact, there exists $\delta > 0$ depending only on u_0 such that $A_{u_0} - \delta g \in \Gamma_k^+$. It follows from Lemma 2.3 that

$$
\delta\operatorname{tr} T_{k-1}(E_u^\epsilon) = \Sigma(E_u^\epsilon,\dots,E_u^\epsilon,\delta g) < \Sigma(E_u^\epsilon,\dots,E_u^\epsilon,A_{u_0}) = \langle T_{k-1}(E_u^\epsilon), A_{u_0}\rangle.
$$

Therefore, if $\lambda \gg 2$ we have

$$\begin{aligned} \mathcal{L}H \geq \{&-C-\lambda(k+1)(1+\epsilon)^{k-1}e^{\lambda(u_0-u+\Upsilon)}\}f \\ &+\{2\Lambda(1+\epsilon)^k-\lambda^2(1+\epsilon)^k u_t^2 e^{\lambda(u_0-u+\Upsilon)}\}\sigma_k(A_u) \\ &+\{-C+\lambda(1+\epsilon)\delta\}u_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon). \end{aligned} \tag{4-22}$$

Observe that by choosing $\lambda=\lambda(\delta)$ large enough, we can assume the last term in (4-22) is bounded below by

$$\tfrac{1}{2}\lambda\delta u_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon). \tag{4-23}$$

By the Newton–Maclaurin inequality,

$$\begin{aligned} u_{tt}^{2-k}\operatorname{tr}T_{k-1}(E_u^\epsilon) &= (k-1)u_{tt}^{2-k}\sigma_{k-1}(E_u^\epsilon) \\ &\geq (k-1)u_{tt}^{2-k}\sigma_k(E_u^\epsilon)^{\frac{k-1}{k}} \\ &= (k-1)f^{\frac{k-1}{k}}u_{tt}^{\frac{1}{k}} \\ &\geq Cfu_{tt}^{\frac{1}{k}}. \end{aligned}$$

Combining this with (4-23) and substituting into (4-22), we get

$$\begin{aligned} \mathcal{L}H \geq \{&-C-\lambda(k+1)(1+\epsilon)^{k-1}e^{\lambda(u_0-u+\Upsilon)}+C\lambda\delta u_{tt}^{\frac{1}{k}}\}f \\ &+\{2\Lambda(1+\epsilon)^k-\lambda^2(1+\epsilon)^k u_t^2 e^{\lambda(u_0-u+\Upsilon)}\}\sigma_k(A_u). \end{aligned}$$

Let us fix the constant Υ so that

$$0\leq u_0-u+\Upsilon\leq C;$$

then

$$\mathcal{L}H\geq\{-C-C\lambda(k+1)+C\lambda\delta u_{tt}^{\frac{1}{k}}\}f+\{2\Lambda(1+\epsilon)^k-C\lambda^2u_t^2\}\sigma_k(A_u).$$

Next, we assume $\Lambda=\Lambda(\lambda,\max u_t^2)$ is chosen large enough that the coefficient of the second term above is

$$2\Lambda(1+\epsilon)^k-C\lambda^2u_t^2\geq\tfrac{1}{2}\lambda^2.$$

By the regularized equation,

$$\sigma_k(A_u)\geq\frac{f}{(1+\epsilon)u_{tt}}.$$

Therefore,

$$\mathcal{L}H\geq\left\{-C-C\lambda(k+1)+C\lambda\delta u_{tt}^{\frac{1}{k}}+\frac{1}{2(1+\epsilon)}\lambda^2u_{tt}^{-1}\right\}f.$$

If $u_{tt} > C(\delta)$ is large then the left-hand side is positive, which would be a contradiction at an interior maximum. On the other hand, if u_{tt} is small then as long as λ is chosen large enough, the last term in the braces will dominate and once again we conclude $\mathcal{L}H > 0$. It follows that H attains its maximum on the boundary, as claimed. □

4.5 Existence of approximate and regularizable geodesics

In this subsection we use the a priori estimates of the previous subsections to establish the existence of weak geodesics in the case $n = 4$.

Theorem 4.18 *Given $u_0, u_1 \in \Gamma_2^+$, there exists $f \in C^\infty(M \times [0,1])$ with $f > 0$ and a smooth solution $u(x,t,s,\epsilon)$: $M \times [0,1] \times [0,1] \times (0,\epsilon_0] \to \mathbb{R}$ of $\mathcal{G}^\epsilon_{sf}(u_\epsilon) = 0$ such that:*

(1) *For each $\epsilon \in (0,\epsilon_0]$, $u_\epsilon = u(\cdot,\cdot,\cdot,\epsilon)$ satisfies*

$$u_\epsilon(x,0,s) = u_0(x), \quad u_\epsilon(x,1,s) = u_1(x).$$

(2) *There is a constant $C > 0$, independent of ϵ, such that*

$$|u_\epsilon| + |\nabla u_\epsilon| + |(u_\epsilon)_t| + \epsilon\{|\nabla^2 u_\epsilon| + |\nabla(u_\epsilon)_t| + |(u_\epsilon)_{tt}|\} \leq C.$$

Proof As the argument follows standard lines we provide only a sketch. Fix some $0 < \epsilon_0 < 1$, then choose an arbitrary $0 < \epsilon < \epsilon_0$. First we observe that it follows from [36, Proposition 3] that the path $u_t := tu_1 + (1-t)u_0$ lies in Γ_2^+. Moreover, there exists some constant Λ for which $w_t := u_t + \Lambda t(t-1)$ satisfies $E^\epsilon_u \in \Gamma_2^+$. Let $f := \Phi_\epsilon(w)$, and set

$$\mathcal{I} = \{s \in [0,1] : \text{there is a } u \in C^{4,\alpha} \cap \Gamma_2^+ \text{ that solves } (\star_{\epsilon,f})\}.$$

By construction, $1 \in \mathcal{I}$.

To verify that $\mathcal{I}$ is open, it suffices to study the linearized equation; ie given $\psi \in C^\infty(M \times [0,1])$, we need to solve for some $s \in \mathcal{I}$ the equation

$$\mathcal{L}_{u_\epsilon(\cdot,\cdot,s)}\phi = \psi$$

with ϕ satisfying Dirichlet boundary conditions. The solvability of this linear problem follows from [16, Theorem 6.13].

We claim that $\mathcal{I}$ is closed: Let $\{u_i = u_{s_i}\}$ be a sequence of admissible solutions with $s_i \geq s_0$. The preceding a priori estimates imply there is a constant C (independent of ϵ) such that

$$|u_i| + |\nabla u_i| + |(u_i)_t| + \epsilon\{|(u_i)_{tt}| + |(\nabla u_i)_t| + |\nabla^2 u_i|\} \leq C.$$

To obtain higher-order regularity, we need to verify the concavity of the operator. Observe that the equation can be rewritten as

$$\log\big((1+\epsilon)u_{tt}\sigma_2(A_u) - \langle T_1(A_u), \nabla u_t \otimes \nabla u_t\rangle\big) = f \tag{4-24}$$

for some smooth positive function f. This is a concave elliptic operator in the spacetime Hessian of u.[2] Thus by Evans [14] and Krylov [23] we conclude there is a constant $C = C(\epsilon, f)$ such that

$$\|u_i\|_{C^{2,\alpha}} \leq C.$$

Applying the Schauder estimates we obtain bounds on derivatives of all orders, and it follows that the set $\mathcal{I}$ is closed. Since $\mathcal{I}$ is open, closed and nonempty, it follows that $\mathcal{I} = [0, 1]$. The theorem follows. □

Definition 4.19 Given $u_0, u_1 \in \Gamma_k^+$, we say a one-parameter family of $C^{1,1}$ functions $u_\epsilon(x,t)\colon M \times [0,1] \to \mathbb{R}$ is an *ϵ–geodesic from u_0 to u_1* if

$$u_\epsilon(x,0) = u_0(x), \quad u_\epsilon(x,1,s) = u_1(x), \quad \mathcal{G}_0^\epsilon(u_\epsilon) = 0.$$

We furthermore will say that it is a *regularizable ϵ–geodesic* if there exists $f_0 \in C^\infty(M \times [0,1])$ with $f_0 > 0$ and a smooth function $u(x,t,s)\colon M \times [0,1] \times [0,1] \to \mathbb{R}$ with the following properties:

(i) For each $s \in [0,1]$, $u(\cdot,\cdot,s)$ satisfies

$$u(x,0,s) = u_0(x), \quad u(x,1,s) = u_1(x), \quad \mathcal{G}_{sf_0}(u) = 0.$$

(ii) There is a constant $C > 0$, independent of ϵ, such that

$$|u_\epsilon| + |\nabla u_\epsilon| + |(u_\epsilon)_t| + \epsilon\{|\nabla^2 u_\epsilon| + |\nabla(u_\epsilon)_t| + |(u_\epsilon)_{tt}|\} \leq C.$$

(iii) One has that $u(x,t,s) \to u(x,t)$ in the weak $C^{1,1}$ topology as $s \to 0$.

[2] An earlier draft of this paper noted that the formulation of Proposition 4.1 expressed the equation as a "σ_2–equation", and claimed that convexity of $\sigma_2^{1/2}$ thus sufficed to apply the Evans–Krylov regularity. This is false since the left-hand side of (4-2) is σ_2 of a nonlinear combination of second derivatives. Nonetheless, (4-24) is convex, as was shown by He [22, Theorem 4.1]. We thank the referee for pointing out this problem, and Weiyong He for correcting it.

Definition 4.20 Given $u_0, u_1 \in \Gamma_k^+$, we say a one-parameter family of C^1 functions $u(x,t)$ is a *regularizable geodesic from* u_0 *to* u_1 if there exists $f_0 \in C^\infty(M \times [0,1])$ with $f_0 > 0$ and a smooth function $u(x,t,s,\epsilon)$: $M \times [0,1] \times [0,1] \times [0,\epsilon_0] \to \mathbb{R}$ with the following properties:

(i) For each $\epsilon \in [0,\epsilon_0)$, $u_\epsilon = u(\cdot,\cdot,\cdot,\epsilon)$ satisfies

$$u_\epsilon(x,0,s) = u_0(x), \quad u_\epsilon(x,1,s) = u_1(x), \quad \mathcal{G}_{sf_0}(u_\epsilon) = 0.$$

(ii) There is a constant $C > 0$, independent of ϵ, such that

$$|u_\epsilon| + |\nabla u_\epsilon| + |(u_\epsilon)_t| + \epsilon\{|\nabla^2 u_\epsilon| + |\nabla (u_\epsilon)_t| + |(u_\epsilon)_{tt}|\} \leq C.$$

(iii) For each $0 < \alpha < 1$, $u_\epsilon \to u$ in $C^{0,\alpha}$ as $\epsilon, s \to 0$.

We can now show existence and uniqueness of a regularizable geodesic connecting any two points in Γ^+. The key issue for uniqueness is a comparison lemma.

Lemma 4.21 *Suppose* $u, \tilde{u} \in C^\infty$ *are admissible and satisfy*

$$\mathcal{G}^\epsilon_{f_1}(u) = 0, \quad \mathcal{G}^\epsilon_{f_2}(\tilde{u}) = 0,$$

where $f_1 \leq f_2$. *Assume further that on the boundary,*

$$u(x,0) = \tilde{u}(x,0), \quad u(x,1) = \tilde{u}(x,1).$$

Then, on $M \times [0,1]$,

$$u(x,t) \geq \tilde{u}(x,t).$$

We remark here also that the Lemma 4.21 can be used to exhibit uniqueness for solutions of the equation $\mathcal{G}^\epsilon_0(u) = 0$.

Corollary 4.22 *Given* $u_0, u_1 \in \Gamma_k^+$, *there exists a unique* ϵ*–geodesic from* u_0 *to* u_1.

Proof Let $u(x,t,\epsilon)$ and f be the data guaranteed by Theorem 4.18. Due to the a priori estimates, by Arzela–Ascoli there exists a $C^{1,1}$ limit as $s \to 0$. By definition this is an ϵ–geodesic. Now suppose $\tilde{u}$ is another regularizable geodesic connecting u_0 to u_1, with regularization $\tilde{u}(x,t,\epsilon)$ and auxiliary function $\tilde{f}$. Fixing some $\delta > 0$, for sufficiently small $\epsilon > 0$ Lemma 4.21 implies that $u(x,t,\epsilon) \geq \tilde{u}(x,t,\delta)$. Since the convergence is in $C^{0,\alpha}$, sending $\epsilon \to 0$ yields $u(x,t) \geq \tilde{u}(x,t,\delta)$. We can now send $\delta \to 0$ to obtain $u(x,t) \geq \tilde{u}(x,t)$. Since the roles of u and $\tilde{u}$ are interchangeable in that argument, it follows that $u(x,t) = \tilde{u}(x,t)$. $\square$

Corollary 4.23 *Given* $u_0, u_1 \in \Gamma_k^+$, *there exists a unique regularizable geodesic from* u_0 *to* u_1.

Proof Let $u(x,t,\epsilon)$ and f be the data guaranteed by Theorem 4.18. Due to the a priori estimates, by Arzela–Ascoli there exists a $C^{0,\alpha}$ limit as both $\epsilon \to 0$ and $s \to 0$. By definition this is a regularizable geodesic. Now suppose $\widetilde{u}$ is another regularizable geodesic connecting u_0 to u_1, with regularization $\widetilde{u}(x,t,\epsilon)$ and auxiliary function $\tilde{f}$. Fixing some $\delta > 0$, for sufficiently small $\epsilon > 0$ Lemma 4.21 implies that $u(x,t,\epsilon) \geq \widetilde{u}(x,t,\delta)$. Since the convergence is in $C^{0,\alpha}$, sending $\epsilon \to 0$ yields $u(x,t) \geq \widetilde{u}(x,t,\delta)$. We can now send $\delta \to 0$ to obtain $u(x,t) \geq \widetilde{u}(x,t)$. Since the roles of u and $\widetilde{u}$ are interchangeable in that argument, it follows that $u(x,t) = \widetilde{u}(x,t)$. □

5 Smoothing via Guan–Wang flow

In this section we develop a sharper picture (Theorem 5.13) of the short-time smoothing properties of a parabolic flow introduced by Guan and Wang [19]. This is used in the proof of Theorem 1.5 to smooth the approximate geodesics so that we can take strong limits to obtain a curve of critical points for F connecting any two given critical points.

In the first subsection we will derive a series of formulas for the evolution of various quantities. Since we will be quoting some of the formulas from the previous section, we will state these formulas for general dimensions. In the second subsection, where we derive some short-time estimates, we will specialize to the case $n = 4$ and $k = 2$.

First, we recall the definition of the flow introduced in [19]:

(5-1) $$\frac{\partial}{\partial t} u = \log \sigma_k(g_u^{-1} A_u) - V_u^{-1} \int_M \log \sigma_k(g_u^{-1} A_u)\, dV_g.$$

For technical simplicity we will instead study an unnormalized flow

(5-2) $$\frac{\partial}{\partial t} u = \log \sigma_k(g_u^{-1} A_u) = \log \sigma_k(A_u) + 2ku.$$

As we will be able to control the size of u along this flow, the renormalizing term will only change u by a controlled constant, and have no effect on the estimates. A fundamental property of the flow which we will exploit is monotonicity of the functional F.

Lemma 5.1 *Given a solution u to* (5-1)*, one has*

$$\frac{d}{dt} F[u] \leq 0.$$

Proof This is immediate from the flow equation (5-1) and the formula (3-13). □

5.1 Evolution equations

We remark that when the dimension n is greater than 4, Guan and Wang assumed the manifold was locally conformally flat. For the evolutionary formulas we are interested in this assumption will not be necessary.

Definition 5.2 Given an admissible solution u to (5-2), define

$$\begin{aligned} Lf &= \sigma_k(A_u)^{-1}\langle T_{k-1}(A_u), \nabla^2 f + \nabla u\otimes\nabla f + \nabla f\otimes\nabla u - \langle\nabla u,\nabla f\rangle g\rangle,\\ H &= \frac{\partial}{\partial t} - L,\end{aligned}$$

where the derivatives and inner products are with respect to g (the fixed background metric).

Lemma 5.3 *Let u be a solution to* (5-2). *Then*

$$Hu = \log\sigma_k(A_u) - k + \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), A - \nabla u\otimes\nabla u + \tfrac{1}{2}|\nabla u|^2 g\bigr\rangle + 2ku.$$

Proof We directly compute

$$\begin{aligned} Lu &= \sigma_k(A_u)^{-1}\langle T_{k-1}(A_u), \nabla^2 u + 2\nabla u\otimes\nabla u - |\nabla u|^2 g\rangle\\ &= \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), A_u - A + \nabla u\otimes\nabla u - \tfrac{1}{2}|\nabla u|^2 g\bigr\rangle\\ &= k + \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), -A + \nabla u\otimes\nabla u - \tfrac{1}{2}|\nabla u|^2 g\bigr\rangle.\end{aligned}$$

Combining this with (5-2) yields the result. □

Lemma 5.4 *Let u be a solution to* (5-2) *and $\lambda\in\mathbb{R}$. Then*

$$\begin{aligned} He^{\lambda u} = \lambda e^{\lambda u}\bigl[&\log\sigma_k(A_u) + 2ku - k\\ &+ \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), A - (1+\lambda)\nabla u\otimes\nabla u + \tfrac{1}{2}|\nabla u|^2 g\bigr\rangle\bigr].\end{aligned}$$

Proof Note

$$\frac{\partial}{\partial t}(e^{\lambda u}) = \lambda e^{\lambda u}(\log\sigma_k(A_u) + 2ku).$$

Also,

$$\begin{aligned} Le^{\lambda u} = \sigma_k(A_u)^{-1}\langle T_{k-1}(A_u), \lambda e^{\lambda u}\nabla^2 u + \lambda^2 e^{\lambda u}\nabla u\otimes\nabla u + 2\lambda e^{\lambda u}\nabla u\otimes\nabla u\\ - \lambda e^{\lambda u}|\nabla u|^2 g\rangle\end{aligned}$$

$$
\begin{aligned}
&= \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), \lambda e^{\lambda u}\bigl[A_u - A - \nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr] \\
&\qquad\qquad\qquad\qquad + \lambda(\lambda+2)e^{\lambda u}\nabla u\otimes\nabla u - \lambda e^{\lambda u}|\nabla u|^2 g\bigr\rangle \\
&= \lambda e^{\lambda u}\sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), A_u - A + (\lambda+1)\nabla u\otimes\nabla u - \tfrac12|\nabla u|^2 g\bigr\rangle \\
&= \lambda e^{\lambda u}\sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), -A + (\lambda+1)\nabla u\otimes\nabla u - \tfrac12|\nabla u|^2 g\bigr\rangle + \lambda k e^{\lambda u}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
He^{\lambda u} &= \frac{\partial}{\partial t}(e^{\lambda u}) - Le^{\lambda u} \\
&= \lambda e^{\lambda u}\bigl[\log\sigma_k(A_u) + 2ku - k \\
&\qquad + \sigma_k(A_u)^{-1}\bigl\langle T_{k-1}(A_u), A - (1+\lambda)\nabla u\otimes\nabla u + \tfrac12|\nabla u|^2 g\bigr\rangle\bigr]. \qquad \square
\end{aligned}
$$

Lemma 5.5 *Given a solution u to (5-2), one has*

$$
H|\nabla u|^2 = 2\sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\{-\nabla_i\nabla_p u\nabla_i\nabla_q u + \mathcal{O}(|\nabla u|^2+1)\} + 4k|\nabla u|^2.
$$

Proof We compute

$$
\begin{aligned}
\frac{\partial}{\partial t}\nabla_i u &= \nabla_i\log\sigma_k(A_u) + 2k\nabla_i u \\
&= \sigma_k(A_u)^{-1}\langle T_{k-1}(A_u), \nabla_i A_u\rangle + 2k\nabla_i u \\
&= \sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\bigl\{\nabla_i A_{pq} + \nabla_i\nabla_p\nabla_q u + 2\nabla_i\nabla_p u\nabla_q u - \nabla_i\nabla_j u\nabla_j u g_{pq}\bigr\} \\
&\qquad + 2k\nabla_i u \\
&= \sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq} \\
&\qquad\qquad \times\bigl\{\nabla_p\nabla_q\nabla_i u + 2\nabla_i\nabla_p u\nabla_q u - \nabla_i\nabla_j u\nabla_j u g_{pq} + (\nabla A + \mathrm{Rm} * \nabla u)_{ipq}\bigr\} \\
&\qquad + 2k\nabla_i u,
\end{aligned}
$$

hence

$$
\begin{aligned}
\frac{\partial}{\partial t}|\nabla u|^2 &= 2\sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq} \\
&\qquad\qquad \times\bigl\{\nabla_p\nabla_q\nabla_i u\nabla_i u + 2\nabla_i\nabla_p u\nabla_q u\nabla_i u - \nabla_i\nabla_j u\nabla_j u\nabla_i u g_{pq} \\
&\qquad\qquad\qquad\qquad + [(\nabla A + \mathrm{Rm} * \nabla u) * \nabla u]_{pq}\bigr\} \\
&\qquad + 4k|\nabla u|^2.
\end{aligned}
$$

Also,

$$
\begin{aligned}
L|\nabla u|^2 &= 2\sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\bigl\{\nabla_p\nabla_i u\nabla_q\nabla_i u + \nabla_p\nabla_q\nabla_i u\nabla_i u \\
&\qquad\qquad + 2\nabla_i\nabla_p u\nabla_q u\nabla_i u - \nabla_i\nabla_j u\nabla_j u\nabla_i u g_{pq}\bigr\}.
\end{aligned}
$$

It follows that

$$\begin{aligned}\frac{\partial}{\partial t}&|\nabla u|^2\\ &= L|\nabla u|^2 + 2\sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\{-\nabla_p\nabla_i u\nabla_q\nabla_i u + [(\nabla A + \mathrm{Rm} * \nabla u) * \nabla u]_{pq}\}\\ &\qquad + 4k|\nabla u|^2,\end{aligned}$$

which implies the result. □

Corollary 5.6 *Given a solution* u *to* (5-2), *one has*

$$H(e^{-4kt}|\nabla u|^2) = 2e^{-4kt}\sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\{-\nabla_i\nabla_p u\nabla_i\nabla_q u + \mathcal{O}(|\nabla u|^2+1)\}. \tag{5-3}$$

For the following lemma, for an $n\times n$ symmetric matrix $r = r_{ij}$ we write

$$\mathcal{F}(r) = \log\sigma_k(r),$$

and denote derivatives of $\mathcal{F}$ with respect to the entries of r by

$$\frac{\partial}{\partial r_{pq}}\mathcal{F}(r) = \mathcal{F}(r)^{pq}, \quad \frac{\partial^2}{\partial r_{pq}\partial r_{rs}}\mathcal{F}(r) = \mathcal{F}(r)^{pq,rs}.$$

Lemma 5.7 *Given a solution* u *to* (5-2), *one has*

$$\begin{aligned}H\Delta u = \mathcal{F}^{pq,rs}&\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs}\\ &+\sigma_k(A_u)^{-1}\big\langle T_{k-1}(A_u)_{ij}, 2\nabla_i\nabla_p u\nabla_j\nabla_p u - |\nabla^2 u|^2 g_{ij}\\ &\qquad + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2+1)\big\rangle.\end{aligned}$$

Proof We compute

$$\Delta\log\sigma_k(A_u) = \nabla_i[\mathcal{F}^{pq}\nabla_i(A_u)_{pq}] = \mathcal{F}^{pq,rs}\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs} + \mathcal{F}^{pq}\Delta(A_u)_{pq}.$$

Combining this with our prior calculation of ΔA_u (4-14) yields

$$\begin{aligned}\frac{\partial}{\partial t}\Delta u &= \Delta\log\sigma_k(A_u) + 2k\Delta u\\ &= \mathcal{F}^{pq,rs}\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs} + \sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}(\Delta A_u)_{pq} + 2k\Delta u\\ &= \mathcal{F}^{pq,rs}\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs}\\ &\quad + \sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq}\\ &\qquad \times\big\{\nabla_p\nabla_q(\Delta u) + \nabla_p\Delta u\nabla_q u + \nabla_p u\nabla_q\Delta u + 2\nabla_p\nabla_\ell u\nabla_q\nabla_\ell u - |\nabla^2 u|^2 g_{pq}\\ &\qquad\qquad - \langle\nabla u, \nabla\Delta u\rangle g_{pq} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\big\}\\ &\quad + 2k\Delta u\end{aligned}$$

$$
\begin{aligned}
&= L(\Delta u) + \mathcal{F}^{pq,rs}\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs} \\
&\qquad + \sigma_k(A_u)^{-1}T_{k-1}(A_u)^{pq} \\
&\qquad\qquad \times \{2\nabla_p\nabla_\ell u\nabla_q\nabla_\ell u - |\nabla^2 u|^2 g_{pq} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1)\},
\end{aligned}
$$

and the result follows. □

5.2 Estimates

In this section we specialize to the case $n = 4$ and $k = 2$, and use the evolutionary formulas from the preceding subsection to derive some short-time smoothing estimates. We begin with a standard result:

Lemma 5.8 *Let (M^4, g) be a compact Riemannian manifold such that $g \in \Gamma_2^+$.*

(i) *Given $u_0 \in \Gamma_2^+$, there exists a solution u of (5-2) on some small time interval $[0, \eta]$, where $\eta > 0$ depends on $|u_0|_{C^{4,\alpha}}$.*

(ii) *If $[0, T)$ is the maximal time interval on which the solution exists and $T < \infty$, then*

$$\limsup_{t\to T} \max_M \{|\log\sigma_2(A_u)| + |u| + |\nabla u| + |\nabla^2 u|\}(\cdot, t) = \infty.$$

Proof Part (i) appears in Proposition 3 of [19]. Although part (ii) is not explicitly stated, it is implicitly used in Section 3 of [19]; therefore, we provide a brief summary.

Suppose

$$\sup_{M\times[0,T)} \{|\log\sigma_2(A_u)| + |u| + |\nabla u| + |\nabla^2 u|\} \le C.$$

It then follows that (5-2) is strictly parabolic on $[0, T)$ with control over the parabolicity and the $C^{1,1}$ norm of u as $t \to T$. It follows from the Evans–Krylov estimates [14; 23] that there is a $C^{2,\alpha}$ estimate for u on $[0, T)$ which remains controlled as $t \to T$. Schauder estimates now imply that for any l and α there are $C^{l,\alpha}$ bounds on u on $[0, T)$ which remain controlled as $t \to T$, which by a standard compactness argument proves that the solution can be extended beyond time T if $T < \infty$. □

Proposition 5.9 *Suppose u is a smooth solution to (5-2) with $n = 4$ on $[0, T]$ with $T \le 1$. There is a constant $C = C(g)$ such that*

$$\sup_{M\times[0,T]} |u| \le C(1 + |u_0|_{C^0}).$$

Proof At a maximum for u, one has $A_u \le A$, and hence $\sigma_2(A_u) \le \sigma_2(A) < C$. By (5-2),

$$\frac{d}{dt}\max u \le C + 4\max u.$$

Integrating this inequality, we get an upper bound for u. Applying a similar argument at a minimum of u, we obtain a lower bound. □

Proposition 5.10 *Suppose u is a smooth solution to* (5-2) *with $n = 4$ on $[0,T]$ with $T \le 1$. There is a constant $C = C(g, |u_0|_{C^0}, |\nabla u_0|_{C^0})$ such that*

$$\sup_{M\times[0,T]} |\nabla u| \le C.$$

Proof Let

$$\Phi = e^{-8t}|\nabla u|^2 + \Lambda e^{-2u} - \mu t,$$

where $\Lambda, \mu > 0$ will be specified later. Combining Corollary 5.6 and Lemma 5.4, and using the fact that at a maximum of Φ we have $H\Phi \ge 0$, it follows that

$$\begin{aligned}
0 &\le H\Phi \\
&= 2\sigma_2(A_u)^{-1}T_1(A_u)^{pq} \\
&\qquad \times\bigl\{-e^{-8t}\nabla_i\nabla_p u\nabla_i\nabla_q u + e^{-8t}\mathcal{O}(1+|\nabla u|^2) - \Lambda e^{-2u}\bigl[\nabla_p u\nabla_q u + \tfrac{1}{2}|\nabla u|^2 g_{pq}\bigr]\bigr\} \\
&\qquad -2\Lambda e^{-2u}[\log\sigma_2(A_u) + 4u - 2 + \sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle] - \mu \\
&= I_1 + I_2 - \mu.
\end{aligned}$$

We can estimate the terms in braces in I_1 by

$$\begin{aligned}
-e^{-8t}\nabla_i\nabla_p u\nabla_i\nabla_q u + e^{-8t}\mathcal{O}(1+|\nabla u|^2) - \Lambda e^{-2u}\bigl[\nabla_p u\nabla_q u + \tfrac{1}{2}|\nabla u|^2 g_{pq}\bigr] \\
\le \bigl\{C + \bigl(C - \tfrac{1}{2}\Lambda e^{-2u}\bigr)|\nabla u|^2\bigr\} g_{pq}.
\end{aligned}$$

By Proposition 5.9, on the time interval under consideration we have a uniform bound on $|u|$ depending only on $|u_0|_{C^0}$, hence, if $\Lambda \gg 1$ is chosen large enough,

$$C + \bigl(C - \tfrac{1}{2}\Lambda e^{-2u}\bigr)|\nabla u|^2 \le C - |\nabla u|^2.$$

Again using that $|u|$ is controlled, at a sufficiently large maximum of Φ one has that $|\nabla u|$ must itself be arbitrarily large, thus at a sufficiently large maximum of Φ one has

$$I_1 \le 0.$$

To estimate I_2, we first consider the case where $\sigma_2(A_u) \geq 1$. Then $\log \sigma_2(A_u) \geq 0$ and the remaining terms in brackets are either bounded or nonnegative, hence

$$I_2 - \mu \leq C(\Lambda, \max |u|) - \mu \leq 0 \tag{5-4}$$

if μ is chosen large enough with respect to $|u_0|_{C^0}$. On the other hand, using Lemma 2.6 we see that

$$\sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle \geq \sigma_2(A_u)^{-1}\sigma_2(A_u)^{\frac{1}{2}}\sigma_2(A)^{\frac{1}{2}} = \frac{\sigma_2(A)^{\frac{1}{2}}}{\sigma_2(A_u)^{\frac{1}{2}}}.$$

It follows that there is a small constant $\delta = \delta(\sigma_2(A))$ such that if $0 < \sigma_2(A_u) \leq \delta$, then

$$\log \sigma_2(A_u) + \sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle \geq 0.$$

Then, arguing as we did in the case where $\sigma_2(A_u) \geq 1$, we can choose μ large enough to achieve (5-4) again. Finally, in the intermediate range $\delta \leq \sigma_2(A_u) \leq 1$, all the terms in the brackets in I_2 are bounded and nonpositive, and we again conclude that (5-4) holds once μ is chosen large enough. It follows that $H\Phi \leq 0$, and the result follows from the maximum principle. $\square$

Proposition 5.11 *Suppose u is a solution to* (5-2) *with $n = 4$ on $[0, T]$ with $T \leq 1$. There exists a constant $C = C(|u_0|_{C^0})$ such that for all $t \in [0, T]$, one has*

$$t\,|\log \sigma_2(A_u)| \leq C.$$

Proof We first note that by Proposition 5.9,

$$\sup_{M\times[0,T]} |u| \leq N, \tag{5-5}$$

where $N = N(|u_0|_{C^0})$.

Next, by the evolution equations above,

$$\begin{aligned}
&\frac{\partial}{\partial t}\log \sigma_2(A_u)\\
&\quad= \sigma_2(A_u)^{-1}\Big\langle T_1(A_u), \frac{\partial}{\partial t}A_u\Big\rangle\\
&\quad= \sigma_2(A_u)^{-1}\big\langle T_1(A_u), \nabla^2 \log \sigma_2(A_u) + \nabla u \otimes \nabla \log \sigma_2(A_u) + \nabla \log \sigma_2(A_u) \otimes \nabla u\\
&\qquad\qquad - \langle \nabla u, \nabla \log \sigma_2(A_u)\rangle g + 4\nabla^2 u + 8\nabla u \otimes \nabla u - 4|\nabla u|^2 g\big\rangle\\
&\quad= L(\log \sigma_2(A_u)) + 4\sigma_2(A_u)^{-1}\langle T_1(A_u), \nabla^2 u + 2\nabla u \otimes \nabla u - |\nabla u|^2 g\rangle\\
&\quad= L(\log \sigma_2(A_u)) + 4\sigma_2(A_u)^{-1}\big\langle T_1(A_u), A_u - A + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\big\rangle\\
&\quad= L(\log \sigma_2(A_u)) + 8 + 4\sigma_2(A_u)^{-1}\big\langle T_1(A_u), -A + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\big\rangle,
\end{aligned}$$

hence

$$H(\log \sigma_2(A_u)) = 8 + 4\sigma_2(A_u)^{-1}\big\langle T_1(A_u), -A + \nabla u \otimes \nabla u - \tfrac{1}{2}|\nabla u|^2 g\big\rangle. \tag{5-6}$$

Set

$$\Phi := t \log \sigma_2(A_u) + \Lambda e^{-2u} - \mu t.$$

We will show that by choosing $\Lambda, \mu \gg 1$ sufficiently large (depending on N), $H\Phi \le 0$. This will give an upper bound on Φ depending only on the initial C^0–norm of u.

To begin, we combine (5-6) with Lemma 5.4 to get

$$\begin{aligned} H\Phi = -\mu + 8t + 4\Lambda(1-2u)e^{-2u} + (1-2\Lambda e^{-2u})\log \sigma_2(A_u) \\ +\sigma_2(A_u)^{-1}\big\langle T_1(A_u), -(4t+2\Lambda e^{-2u})A + (4t-2\Lambda e^{-2u})\nabla u \otimes \nabla u \\ -(2t+\Lambda e^{-2u})|\nabla u|^2 g\big\rangle. \end{aligned} \tag{5-7}$$

By choosing Λ large enough (depending on the constant N in (5-5)) we may assume the coefficient of the log term satisfies

$$1 - 2\Lambda e^{-2u} \le -1. \tag{5-8}$$

For t small (depending on N and Λ) the coefficients of the gradient terms in (5-7) are also nonpositive, so we have

$$\begin{aligned} H\Phi \le -\mu + 8t + 4\Lambda(1-2u)e^{-2u} + (1-2\Lambda e^{-2u})\log \sigma_2(A_u) \\ -(4t+2\Lambda e^{-2u})\sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle. \end{aligned} \tag{5-9}$$

If $\mu \gg 1$ is chosen large enough, the first three terms on the right-hand side of (5-7) can be bounded above by $-\frac{\mu}{2}$, and we conclude

$$H\Phi \le -\frac{\mu}{2} + (1-2\Lambda e^{-2u})\log \sigma_2(A_u) - (4t+2\Lambda e^{-2u})\sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle. \tag{5-10}$$

By Lemma 2.6 we have

$$\sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle \ge \sigma_2(A_u)^{-1}[4\sigma_2(A_u)^{\frac{1}{2}}\sigma_2(A)^{\frac{1}{2}}] \ge \delta\sigma_2(A_u)^{-\frac{1}{2}} > 0,$$

hence

$$-(4t+2\Lambda e^{-2u})\sigma_2(A_u)^{-1}\langle T_1(A_u), A\rangle \le -C_1\sigma_2(A_u)^{-\frac{1}{2}}. \tag{5-11}$$

If $\sigma_2(A_u) \ge 1$, it follows from (5-8), (5-10) and (5-11) that $H\Phi \le 0$. On the other hand, if $\sigma_2(A_u) < 1$, then

$$H\Phi \le -\frac{\mu}{2} - \log \sigma_2(A_u) - C_1\sigma_2(A_u)^{-\frac{1}{2}},$$

and by choosing $\mu \gg 1$ large enough (depending only on C_1) once again we have $H\Phi \le 0$.

To obtain a lower bound for $\log \sigma_2(A_u)$, we consider

$$\widetilde{\Phi} := -t \log \sigma_2(A_u) + \Lambda e^{-2u} - \mu t,$$

and apply a similar argument. We will omit the details. □

Proposition 5.12 *Suppose u is a solution to* (5-2) *with $n = 4$ on $[0, T]$ with $T \le 1$. There exists a constant $C = C(|u_0|_{C^0}, |u_0|_{C^1})$ such that for all $t \in [0, T]$, one has*

$$t\Delta u \le C.$$

Proof By Proposition 5.10, there is a $\Lambda = \Lambda(|u_0|_{C^0}, |u_0|_{C^1})$ such that

$$\sup_{M\times[0,T]} \{|\nabla u|^2 + |u|\} \le \Lambda.$$

Let

$$\Phi = t\Delta u + |\nabla u|^2.$$

A direct calculation using Lemmas 5.5 and 5.7 and some elementary estimates yields

$$\begin{aligned} H\Phi = \Delta u &+ t\mathcal{F}^{pq,rs}\nabla_i(A_u)_{pq}\nabla_i(A_u)_{rs} \\ &+ \sigma_2(A_u)^{-1}T_1(A_u)^{pq}\big\{2(t-1)\nabla_i\nabla_p u\nabla_i\nabla_q u - t|\nabla^2 u|^2 g_{pq} \\ &\qquad + \mathcal{O}(t|\nabla^2 u| + |\nabla u|^2 + 1)\big\}. \end{aligned} \tag{5-12}$$

If Φ attains a large space-time maximum, say $\Phi \ge B \ge 2A$, then

$$t\Delta u \ge B - A \ge \tfrac{1}{2}B,$$

hence

$$t|\nabla^2 u|^2 \ge \frac{B^2}{16t}.$$

Therefore, if $t \le 1$, the terms in braces in (5-12) can be estimated as

$$\begin{aligned} 2(t-1)\nabla_i\nabla_p u\nabla_i\nabla_q u - t|\nabla^2 u|^2 g_{pq} + \mathcal{O}(t|\nabla^2 u| + |\nabla u|^2 + 1) \\ &\le \{-t|\nabla^2 u|^2 + Ct|\nabla^2 u| + C(A)\}g_{pq} \\ &\le \Big\{-\frac{t}{2}|\nabla^2 u|^2 + C'\Big\}g_{pq} \\ &\le \Big\{-\frac{B^2}{32t} + C'\Big\}g_{pg} \\ &\le 0 \end{aligned}$$

if B is large enough. Thus we conclude $H\Phi < 0$ at a sufficiently large maximum, proving the result. □

Theorem 5.13 *Let* (M^4, g) *be a compact Riemannian manifold such that* $g \in \Gamma_2^+$. *Given* $u_0 \in \Gamma_2^+$, *there exists* $C = C(|u_0|, |\nabla u_0|)$ *such that the solution to* (5-1) *with initial condition* u_0 *exists on* $[0, 1]$ *and moreover satisfies*

$$-C \le t \log \sigma_2(A_u) \le C, \quad -C \le \Delta u \le \frac{C}{t}. \tag{5-13}$$

Furthermore, choosing $l \in \mathbb{N}$ *and* $0 < \alpha < 1$, *there exists* $C = C(|u_0|, |\nabla u_0|, l, \alpha)$ *such that*

$$|u_1|_{C^{l,\alpha}} \le C.$$

Proof By Lemma 5.8 there is a solution u of (5-2) on some small time interval $[0, \eta]$, where η depends on $|u_0|_{C^{4,\alpha}}$. We now argue that the solution can be extended smoothly to $[0, 1]$. Suppose $\eta \le T \le 1$ is the maximal smooth existence time of the flow. By Propositions 5.9 and 5.10, we have

$$\sup_{M \times [0,T)} \{|u| + |\nabla u|\} \le C(|u_0|_{C^0}, |u_0|_{C^1}).$$

In addition, by Propositions 5.11 and 5.12 the estimates (5-13) hold on $[0, T)$, where the constant is $C = C(|u_0|_{C^0}, |u_0|_{C^1})$.

By Lemma 5.8, if $T \le 1$, we must have

$$\limsup_{t \to T} \max_M \{|\log \sigma_2(A_u)| + |\nabla^2 u|\}(\cdot, t)| = \infty. \tag{5-14}$$

However, since $u \in \Gamma_2^+$ it follows that

$$|\nabla^2 u| \le C(1 + |\nabla u|^2 + |\Delta u|),$$

which combined with the estimates in (5-13) will contradict (5-14). It follows that $T > 1$. □

Remark 5.14 The proof above could be used to show long-time existence of the solution to (5-2), but the estimates of Propositions 5.9 and 5.10 degenerate as time approaches infinity, and thus one would not obtain convergence with these estimates. The short-time existence statement, together with a smoothing effect which is controlled by the C^1 norm of the smooth initial data, is a crucial tool in smoothing approximate geodesics to obtain Theorem 1.5.

6 Uniqueness of solutions to the σ_2–Yamabe problem

In this section we combine the previous results to establish Theorem 1.5. As described in the introduction, the proof consists of a few main steps. In particular, we use Theorem 4.18 to connect any two critical points for F by an ϵ–geodesic. Applying the geodesic convexity of F we obtain that the curve must consist of near-minimizers for F. We then smooth this approximate geodesic via Theorem 5.13. Taking the limit as $\epsilon \to 0$ of these smoothed paths yields a nontrivial one-parameter family of minimizers of F. Using our knowledge of the geodesic convexity of F we can show that this can only happen if the background conformal class is $[g_{S^4}]$ and the endpoints of the path are round metrics. Note that, unlike the Kähler setting, we are unable to show that the approximate geodesics converge directly to a nontrivial smooth geodesic due to the lack of stronger regularity results for the geodesics.

Lemma 6.1 *Given two admissible critical points u_0 and u_1 of F, one has $F[u_0]= F[u_1]$, and $F[u]\geq F[u_0]$ for all admissible u. Moreover, given f and $u=u(x,t,s,\epsilon)$, the approximate geodesics given by Theorem 4.18, one has, for any $t\in[0,1]$,*

$$\lim_{s,\epsilon\to 0} F[u(\,\cdot\,,t,s,\epsilon)] = F[u_0].$$

Proof Fix f, and let $u=u(x,t,s,\epsilon)$ be the approximate geodesics guaranteed by Theorem 4.18, connecting u_0 and u_1. To begin we repeat the calculation of Proposition 3.16 for these paths. Fix some s and ϵ and compute

$$\begin{aligned}
\frac{d^2}{dt^2}F[u] &= \frac{d}{dt}\int_M u_t[-\sigma_2(g_u^{-1}A_u)+\overline{\sigma}]\,dV_u \\
&= -\int_M [u_{tt}\sigma_2(g_u^{-1}A_u) + u_t\langle T_1(g_u^{-1}A_u), \nabla^2 u_t\rangle]\,dV_u \\
&\qquad\qquad + \sigma\int_M\left[u_{tt}V_u^{-1} + V_u^{-2}u_t\left(\int_M 4u_t\,dV_u\right) - 4V_u^{-1}u_t^2\right]dV_u \\
&= \int_M [\epsilon u_{tt}\sigma_2(g_u^{-1}A_u) - sf]\,dV_u + \sigma V_u^{-1}\int_M\left[\frac{1}{\sigma_2(g_u^{-1}A_u)}sf - \epsilon u_{tt}\right]dV_u \\
&\quad + \sigma V_u^{-1}\int_M\left[\frac{1}{\sigma_2(A)}\langle T_1(g_u^{-1}A_u), \nabla u_t\otimes\nabla u_t\rangle\right. \\
&\qquad\qquad\qquad \left. - 4\left(\int_M u_t^2\,dV_u - V_u^{-1}\left(\int_M u_t\,dV_u\right)^2\right)\right]dV_u.
\end{aligned}$$

Applying Corollary 3.15 to the above equation yields

$$\frac{d^2}{dt^2} F \geq -\int_M sf\, dV_u - \sigma V_u^{-1} \epsilon \int_M u_{tt}. \tag{6-1}$$

Now let us estimate using the uniform C^1 estimate:

$$\begin{aligned}\int_0^1 \int_M u_{tt}\, dV_u &= \int_0^1 \left[\frac{\partial}{\partial t} \int_M u_t dV_u - \int_M 4u_t^2 \right] dt \\ &= \int_M u_t\, dV_u \Big|_{t=0}^{t=1} - \int_0^1 \int_M 4u_t^2\, dV_u dt \\ &\leq C.\end{aligned}$$

Hence, integrating the inequality (6-1) and using that u_0 is a critical point yields

$$\frac{d}{dt} F[u](t) = \frac{d}{dt} F[u](t) - \frac{d}{dt} F[u](0) = \int_0^t \frac{d^2}{dt^2} F\, dt \geq -C(s+\epsilon).$$

Integrating this in time and sending $s, \epsilon \to 0$ yields

$$F[u_1] \geq F[u_0].$$

But since the roles of u_0 and u_1 are interchangeable, we obtain $F[u_0] = F[u_1]$. □

Lemma 6.2 *Fix (M^4, g) with $A_g \in \Gamma_2^+$, and suppose $u \in C^\infty(M)$ is an admissible critical point of F. Then either u is an isolated critical point for F or (M^4, g_u) is isometric to (S^4, g_{S^4}).*

Proof Suppose u is not an isolated critical point, so that there exists a sequence of admissible conformal factors $\{u_i\}$, with $u_i \neq u$, converging in C^∞ to u, normalized so that $\int_M (u - u_i)\, dV_u = 0$. We aim to use the convexity properties to show that the minimum eigenvalue of the linear operator

$$L(\phi) = -\langle T_1(g_u^{-1} A_u), \nabla_{g_u}^2 \phi \rangle_{g_u} - 4\overline{\sigma}\phi$$

is zero. Since u satisfies $\sigma_2(A_u) \equiv \overline{\sigma}$ and has unit volume, this lowest eigenvalue is characterized variationally as

$$\lambda_1 = \inf_{\{\phi \mid \int_M \phi\, dV_u = 0\}} \overline{\sigma} \int_M [\sigma_2(A_u)^{-1} \langle T_1(A_u), \nabla\phi \otimes \nabla\phi \rangle - 4\phi^2]\, dV_u.$$

It follows from Corollary 3.15 that $\lambda_1 \geq 0$, with equality if and only if (M^4, g_u) is isometric to (S^4, g_{S^4}). We suppose that $\lambda_1 > 0$ and derive a contradiction.

Fix a sufficiently large i that the path

$$w(x,t) = (1-t)u + tu_i$$

consists of admissible functions. Note that $w_{tt} = 0$, and, by construction,

$$\frac{dF(w(\cdot,t))}{dt}(0) = \frac{dF(w(\cdot,t))}{dt}(1) = 0.$$

It follows that for any i there exists $t_i \in [0,1]$ such that

$$\frac{d^2F(w(\cdot,t))}{dt^2}(t_i) = 0.$$

We aim to derive a contradiction from this setup. First we make a second variation calculation along this path using (1-9) and (2-5), yielding

$$\begin{aligned}
&\frac{d^2}{dt^2}F[w(\cdot,t)] \\
&\quad= \frac{d}{dt}\int_M w_t(-\sigma_2(g_w^{-1}A_w)+\overline{\sigma})\,dV_w \\
&\quad= \int_M w_{tt}(-\sigma_2(g_w^{-1}A_w)+\overline{\sigma})\,dV_w + \int_M [-w_t\langle T_1(g_w^{-1}A_w), \nabla^2 w_t\rangle - n\overline{\sigma}w_t^2]\,dV_w \\
&\quad= \int_M [\langle T_1(g_w^{-1}A_w), \nabla w_t \otimes \nabla w_t\rangle - n\overline{\sigma}w_t^2]\,dV_w \\
&\quad= \overline{\sigma}\int_M [\sigma_k(g_w^{-1}A_w)^{-1}\langle T_1(g_w^{-1}A_w), \nabla w_t \otimes \nabla w_t\rangle - nw_t^2]\,dV_w.
\end{aligned}$$

We next evaluate this at t_i. Using that $w^i := w(\cdot,t_i)$ converges to u as $i \to \infty$ yields

$$\begin{aligned}
0 &= \int_M [\langle T_1(g_{w^i}^{-1}A_{w^i}), \nabla w_t \otimes \nabla w_t\rangle - n\overline{\sigma}w_t^2]\,dV_{w^i} \\
&= \int_M [\langle (1-o(1))T_1(g_{u_0}^{-1}A_{u_0}), \nabla w_t \otimes \nabla w_t\rangle - n\overline{\sigma}w_t^2](1-o(1))\,dV_{u_0} \\
&= \overline{\sigma}\int_M [\sigma_k(g_{u_0}^{-1}A_{u_0})^{-1}\langle T_1(g_{u_0}^{-1}A_{u_0}), \nabla w_t \otimes \nabla w_t\rangle - nw_t^2]\,dV_{u_0} - o(1) \\
&\geq \overline{\sigma}\lambda_1\int_M w_t^2\,dV_{u_0} - o(1).
\end{aligned}$$

If $\lambda_1 > 0$ then, for sufficiently large i, this implies that $w_t = u_i - u = 0$, a contradiction. It follows that $\lambda_1 = 0$, and hence, by Corollary 3.15, (M^4, g_u) is isometric to (S^4, g_{S^4}). □

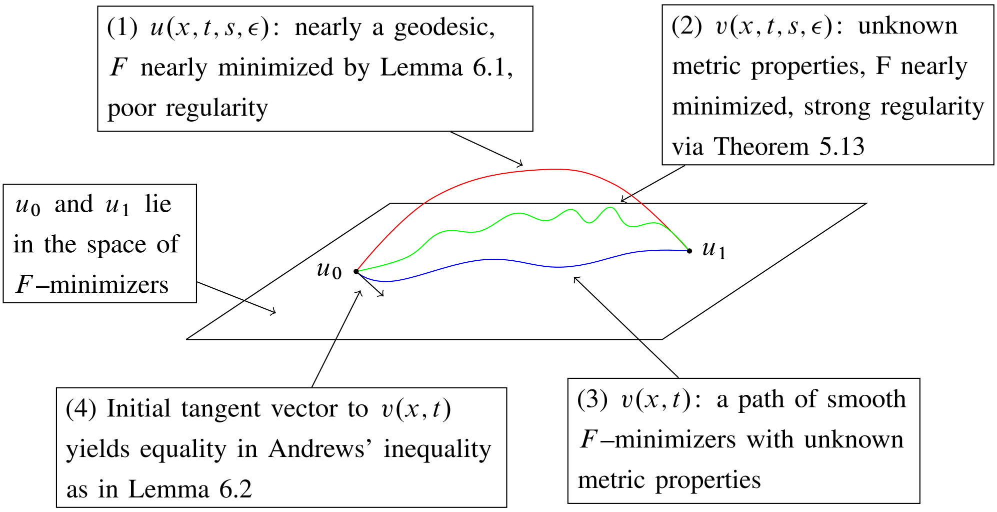

Figure 1: Scheme of the proof of Theorem 1.5

Proof of Theorem 1.5 See Figure 1 for a schematic outline of the argument. Suppose there exist two distinct solutions u_0 and u_1 to the σ_2–Yamabe problem. Let $u(x,t,s,\epsilon)$ be the family of approximate geodesics connecting u_0 to u_1 guaranteed by Theorem 4.18. Noting the a priori estimates on $|u|_{C^0}$ and $|\nabla u|_{C^0}$ are independent of s and ϵ, we have by Theorem 5.13 that the solution to the flow equation (5-1) with initial condition $u(\cdot,t,s,\epsilon)$ exists on some time interval [0, 1], and moreover the solution at time 1, denoted by $v(x,t,s,\epsilon)$, has uniform $C^{k,\alpha}$ estimates independent of s and ϵ and stays uniformly in the interior of Γ_2^+, in the sense that $T_1(g_v^{-1}A_v)$ has uniform upper and lower bounds. Due to these estimates we can obtain a one-parameter family of smooth functions $v(x,t) = \lim_{s,\epsilon\to 0} v(x,t,s,\epsilon)$ which is continuous in t. Moreover, by Lemmas 5.1 and 6.1 we see that $F[v(\cdot,t)] = F[u_0]$. It follows that $v(\cdot,t)$ is a nontrivial path of critical points for F through u_0, and hence by Lemma 6.2 we conclude that (M^4, g_u) is isometric to (S^4, g_{S^4}). □

Acknowledgements Gursky acknowledges the support of NSF grant DMS-1509633. Streets acknowledges the support of NSF Grant DMS-1454854.

References

[1] **Z Błocki**, *On geodesics in the space of Kähler metrics*, from "Advances in geometric analysis" (S Janeczko, J Li, D H Phong, editors), Adv. Lect. Math. 21, International, Somerville, MA (2012) 3–19 MR

[2] **S Brendle**, *Blow-up phenomena for the Yamabe equation*, J. Amer. Math. Soc. 21 (2008) 951–979 MR

[3] **S Brendle**, **F C Marques**, *Blow-up phenomena for the Yamabe equation, II*, J. Differential Geom. 81 (2009) 225–250 MR

[4] **S Brendle**, **J A Viaclovsky**, *A variational characterization for $\sigma_{n/2}$*, Calc. Var. Partial Differential Equations 20 (2004) 399–402 MR

[5] **E Calabi**, **X X Chen**, *The space of Kähler metrics, II*, J. Differential Geom. 61 (2002) 173–193 MR

[6] **S-Y A Chang**, **M J Gursky**, **P Yang**, *An a priori estimate for a fully nonlinear equation on four-manifolds*, J. Anal. Math. 87 (2002) 151–186 MR

[7] **S-Y A Chang**, **M J Gursky**, **P C Yang**, *An equation of Monge–Ampère type in conformal geometry, and four-manifolds of positive Ricci curvature*, Ann. of Math. 155 (2002) 709–787 MR

[8] **S-Y A Chang**, **P C Yang**, *The inequality of Moser and Trudinger and applications to conformal geometry*, Comm. Pure Appl. Math. 56 (2003) 1135–1150 MR

[9] **X Chen**, *The space of Kähler metrics*, J. Differential Geom. 56 (2000) 189–234 MR

[10] **X X Chen**, **G Tian**, *Geometry of Kähler metrics and foliations by holomorphic discs*, Publ. Math. Inst. Hautes Études Sci. 107 (2008) 1–107 MR

[11] **B Chow**, **P Lu**, **L Ni**, *Hamilton's Ricci flow*, Graduate Studies in Mathematics 77, Science, Beijing (2006) MR

[12] **S K Donaldson**, *Symmetric spaces, Kähler geometry and Hamiltonian dynamics*, from "Northern California Symplectic Geometry Seminar" (Y Eliashberg, D Fuchs, T Ratiu, A Weinstein, editors), Amer. Math. Soc. Transl. Ser. 2 196, Amer. Math. Soc., Providence, RI (1999) 13–33 MR

[13] **S K Donaldson**, *Conjectures in Kähler geometry*, from "Strings and geometry" (M Douglas, J Gauntlett, M Gross, editors), Clay Math. Proc. 3, Amer. Math. Soc., Providence, RI (2004) 71–78 MR

[14] **L C Evans**, *Classical solutions of fully nonlinear, convex, second-order elliptic equations*, Comm. Pure Appl. Math. 35 (1982) 333–363 MR

[15] **L Gårding**, *An inequality for hyperbolic polynomials*, J. Math. Mech. 8 (1959) 957–965 MR

[16] **D Gilbarg**, **N S Trudinger**, *Elliptic partial differential equations of second order*, Grundl. Math. Wissen. 224, Springer (1977) MR

[17] **B Guan**, *The Dirichlet problem for complex Monge–Ampère equations and regularity of the pluri-complex Green function*, Comm. Anal. Geom. 6 (1998) 687–703 MR

[18] **P Guan**, **J Viaclovsky**, **G Wang**, *Some properties of the Schouten tensor and applications to conformal geometry*, Trans. Amer. Math. Soc. 355 (2003) 925–933 MR

[19] **P Guan**, **G Wang**, *A fully nonlinear conformal flow on locally conformally flat manifolds*, J. Reine Angew. Math. 557 (2003) 219–238 MR

[20] **M J Gursky**, **J Streets**, *A formal Riemannian structure on conformal classes and the inverse Gauss curvature flow*, preprint (2015) arXiv

[21] **M J Gursky**, **J Streets**, *Variational structure of the v_k–Yamabe problem*, preprint (2016) arXiv

[22] **W He**, *The Gursky–Streets equations*, preprint (2017) arXiv

[23] **N V Krylov**, *Boundedly inhomogeneous elliptic and parabolic equations in a domain*, Izv. Akad. Nauk SSSR Ser. Mat. 47 (1983) 75–108 MR In Russian; translated in Math. USSR-Izv. 22 (1984) 67–97

[24] **J M Lee**, **T H Parker**, *The Yamabe problem*, Bull. Amer. Math. Soc. 17 (1987) 37–91 MR

[25] **J Lelong-Ferrand**, *Transformations conformes et quasiconformes des variétés riemanniennes; application à la démonstration d'une conjecture de A Lichnerowicz*, C. R. Acad. Sci. Paris Sér. A-B 269 (1969) A583–A586 MR

[26] **T Mabuchi**, *K–energy maps integrating Futaki invariants*, Tohoku Math. J. 38 (1986) 575–593 MR

[27] **T Mabuchi**, *Some symplectic geometry on compact Kähler manifolds, I*, Osaka J. Math. 24 (1987) 227–252 MR

[28] **M Obata**, *Certain conditions for a Riemannian manifold to be isometric with a sphere*, J. Math. Soc. Japan 14 (1962) 333–340 MR

[29] **D Pollack**, *Nonuniqueness and high energy solutions for a conformally invariant scalar equation*, Comm. Anal. Geom. 1 (1993) 347–414 MR

[30] **R C Reilly**, *Variational properties of functions of the mean curvatures for hypersurfaces in space forms*, J. Differential Geometry 8 (1973) 465–477 MR

[31] **R M Schoen**, *Variational theory for the total scalar curvature functional for Riemannian metrics and related topics*, from "Topics in calculus of variations" (M Giaquinta, editor), Lecture Notes in Math. 1365, Springer (1989) 120–154 MR

[32] **S Semmes**, *Complex Monge–Ampère and symplectic manifolds*, Amer. J. Math. 114 (1992) 495–550 MR

[33] **W Sheng**, **N S Trudinger**, **X-J Wang**, *The k–Yamabe problem*, from "Surveys in differential geometry" (H-D Cao, S-T Yau, editors), volume 17, International, Boston, MA (2012) 427–457 MR

[34] **J A Viaclovsky**, *Conformal geometry, contact geometry, and the calculus of variations*, Duke Math. J. 101 (2000) 283–316 MR

[35] **J A Viaclovsky**, *Conformally invariant Monge–Ampère equations: global solutions*, Trans. Amer. Math. Soc. 352 (2000) 4371–4379 MR

[36] **J A Viaclovsky**, *Estimates and existence results for some fully nonlinear elliptic equations on Riemannian manifolds*, Comm. Anal. Geom. 10 (2002) 815–846 MR

[37] **J Viaclovsky**, *Conformal geometry and fully nonlinear equations*, from "Inspired by S S Chern" (P A Griffiths, editor), Nankai Tracts Math. 11, World Sci., Hackensack, NJ (2006) 435–460 MR

Department of Mathematics, University of Notre Dame
Notre Dame, IN, United States

Department of Mathematics, University of California, Irvine
Irvine, CA, United States

`mgursky@nd.edu`, `jstreets@uci.edu`

Proposed: Simon Donaldson — Received: 12 June 2017
Seconded: Bruce Kleiner, Tobias H Colding — Revised: 29 September 2017