ANALYSIS & PDE

Volume 16 No. 5 2023

PAATA IVANISVILI AND RYAN RUSSELL

EXPONENTIAL INTEGRABILITY IN GAUSS SPACE

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1271–1288
DOI: 10.2140/apde.2023.16.1271

EXPONENTIAL INTEGRABILITY IN GAUSS SPACE

PAATA IVANISVILI AND RYAN RUSSELL

Talagrand showed that finiteness of $\mathbb{E}\, e^{|\nabla f(X)|^2/2}$ implies finiteness of $\mathbb{E}\, e^{f(X)-\mathbb{E}f(X)}$, where X is the standard Gaussian vector in $\mathbb{R}^n$ and f is a smooth function. However, in this paper we show that finiteness of $\mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-1}$ implies finiteness of $\mathbb{E}\, e^{f(X)-\mathbb{E}f(X)}$, and we also obtain quantitative bounds

$$\log \mathbb{E}\, e^{f-\mathbb{E}f} \le 10\, \mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-1}.$$

Moreover, the extra factor $(1+|\nabla f|)^{-1}$ is the best possible in the sense that there is a smooth f with $\mathbb{E}\, e^{f-\mathbb{E}f} = \infty$ but $\mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-c} < \infty$ for all $c > 1$. As an application we show corresponding dual inequalities for the discrete time dyadic martingales and their quadratic variations.

1. Introduction

Bobkov and Götze [1999] showed that for a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ with $\mathbb{E} f(X) = 0$ we have

$$\mathbb{E} e^{f(X)} \le (\mathbb{E} e^{\alpha|\nabla f(X)|^2})^{1/(2\alpha-1)} \quad \text{for any } \alpha > \tfrac{1}{2}, \tag{1-1}$$

for the class of random vectors X in $\mathbb{R}^n$ satisfying the log-Sobolev inequality *with constant* 1. In particular, the estimate (1-1) holds true when $X \sim \mathcal{N}(0, I_{n\times n})$ is the standard Gaussian vector in $\mathbb{R}^n$ and $I_{n\times n}$ is the identity matrix. The inequality implies the measure concentration inequality $\mathbb{P}(f(X) > \lambda) \le e^{-\lambda^2/2}$ for all $\lambda \ge 0$ provided that $|\nabla f| \le 1$ and $\mathbb{E} f(X) = 0$. In [Bobkov and Götze 1999] it was asked what happens in the endpoint case when $\alpha = \frac{1}{2}$, i.e., does finiteness of $\mathbb{E} e^{|\nabla f(X)|^2/2}$ imply finiteness of $\mathbb{E} e^{f(X)}$ even for $n = 1$ and $X \sim \mathcal{N}(0, 1)$?

From the aforementioned paper, it is not hard to see that the Bobkov–Götze exponential inequality (1-1) is optimal in terms of the powers, i.e., one cannot replace $1/(2\alpha - 1)$ with $1/(c\alpha - 1)$ for some $c < 2$, and one cannot replace $e^{\alpha|\nabla f|^2}$ with $e^{c\alpha|\nabla f|^2}$ for some $c < 1$. Notice that the finiteness of $\mathbb{E} e^{\beta|\nabla f(X)|^2}$ for some $\beta \in \left(0, \frac{1}{2}\right)$ does not imply finiteness of $\mathbb{E} e^{f(X)}$; for instance, consider $X \sim \mathcal{N}(0, 1)$ and $f(x) = \frac{1}{2}(x^2 - 1)$. Therefore, perhaps

$$\mathbb{E} e^{f(X)} < h(\mathbb{E} e^{|\nabla f(X)|^2/2})$$

is the *best possible* inequality one may seek for some $h : [1, \infty) \to [0, \infty)$.

According to a discussion on page 8 in [Bobkov and Götze 1999], Talagrand showed that even though (1-1) *fails* at the endpoint exponent $\alpha = \frac{1}{2}$, surprisingly, the finiteness of $\mathbb{E} e^{|\nabla f(X)|^2/2}$ still implies finiteness of $\mathbb{E} e^{f}$ for $X \sim \mathcal{N}(0, I_{n\times n})$. We are not aware of Talagrand's proof as it was never published; we do not know if he solved the problem only for $n = 1$ or for all $n \ge 1$.

MSC2020: 26D10, 35E10, 42B35.
Keywords: exponential integrability, heat flow, Gauss space, measure concentration.

In this paper we show that the finiteness of $\mathbb{E}e^{|\nabla f|/2}(1+|\nabla f(X)|)^{-1}$ implies the finiteness of $\mathbb{E}e^{f(X)}$ for all $n \geq 1$, and the extra factor $(1+|\nabla f|)^{-1}$ is the best possible in the sense that it cannot be replaced by $(1+|\nabla f|)^{-c}$ for some $c > 1$. Moreover, we provide quantitative bounds.

Theorem 1.1. *For any $n \geq 1$, we have*

$$\log \mathbb{E}e^{f(X)-\mathbb{E}f} \leq 10\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-1} \tag{1-2}$$

for all $f \in C_0^\infty(\mathbb{R}^n)$, where $X \sim \mathcal{N}(0, I_{n\times n})$.

To see the sharpness of the factor $(1+|\nabla f|)^{-1}$ in (1-2), let $n=1$, and let $f(x)=\frac{1}{2}x^2$. Then $\mathbb{E}e^{f(X)}=\infty$. On the other hand,

$$\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-c} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{dx}{(1+|x|)^c} < \infty$$

for all $c > 1$. It remains to multiply f by a smooth cut-off function $\mathbb{1}_{|x|\leq R}$ and take the limit $R \to \infty$.

Using standard mass transportation arguments the exponential integrability (1-2) may be extended to random vectors X having *uniformly* log-concave densities.

Corollary 1.2. *Let X be an arbitrary random vector in $\mathbb{R}^n$ with density $e^{-u(x)}\,dx$ such that* $\operatorname{Hess} u \geq R I_{n\times n}$ *for some $R > 0$. Then*

$$\log \mathbb{E}e^{f(X)-\mathbb{E}f} \leq 10\mathbb{E}e^{|\nabla f(X)|^2/(2R)}(1+R^{-1/2}|\nabla f(X)|)^{-1} \tag{1-3}$$

for all $f \in C_0^\infty(\mathbb{R}^n)$.

Exponential integrability has been studied for other random vectors X as well. Let us briefly record some known results where we assume f to be real-valued with $\mathbb{E}f(Y)=0$. In all examples Y is uniformly distributed on the set where it is given.

$$\log \mathbb{E}e^{f(Y)} \leq \mathbb{E}\tfrac{1}{4}|\nabla_{\mathbb{S}^2} f(Y)|^2, \qquad Y \in \mathbb{S}^2 = \{\|x\|=1,\ x \in \mathbb{R}^3\}, \tag{1-4}$$

$$\log \mathbb{E}e^{f(Y)} \leq 1+\mathbb{E}\tfrac{1}{16}|\nabla f(Y)|^2, \quad Y \in \mathbb{D} = \{\|x\| \leq 1,\ x \in \mathbb{R}^2\}, \tag{1-5}$$

$$\log \mathbb{E}e^{f(Y)} \leq \log \mathbb{E}e^{D(f)^2(Y)}, \qquad Y \in \{-1,1\}^n, \tag{1-6}$$

$$\log \mathbb{E}e^{f(Y)} \leq \log \mathbb{E}e^{4|\nabla f|^2(Y)}, \qquad Y \in [-1,1]^n, \qquad \text{(only for convex } f\text{)}, \tag{1-7}$$

where in (1-6) by the symbol $D(f)^2$ we denote the *discrete gradient*; see [Bobkov and Götze 1999]. The estimate (1-4), also known as the Mozer–Trudinger inequality (with the best constants due to Onofri), has been critical for geometric applications [Moser 1971; Onofri 1982]. A slightly weaker version of (1-6), namely,

$$\mathbb{E}e^{f(Y)} \leq \mathbb{E}e^{\pi^2 D(f)^2(Y)/8},$$

was obtained by Efraim and Lust-Piquard [2008].

The proof of the main theorem follows from *heat flow* arguments. We construct a certain increasing quantity $A(s)$ with respect to a parameter $s \in [0,1]$. We will see that

$$\log \mathbb{E}e^{f(X)} = A(0) \leq A(1) \leq \mathbb{E}f(X) + 10\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-1}.$$

To describe the expression for $A(t)$, let $\Phi(t) = \mathbb{P}(X_1 \le t)$ be the Gaussian cumulative distribution function, and set $k(x) = -\log(\Phi'(t)/\Phi(t))$. Our main object will be a certain function $F : [0, \infty) \to [0, \infty)$ defined as

$$F(x) = \int_0^x e^{k((k')^{-1}(t))}\, dt \quad \text{for all } x \in [0, \infty), \tag{1-8}$$

where $(k')^{-1}$ is the inverse function to k' (it will be explained in the next section why F is well defined). For $g : \mathbb{R}^n \to (0, \infty)$, we consider its heat flow $U_s g(y) := \mathbb{E} g(y + \sqrt{s} X)$, where $s \in [0, 1]$. Then

$$A(s) := U_s\left[\log U_{1-s} g + F\left(\frac{\sqrt{s}|\nabla U_{1-s} g|}{U_{1-s} g}\right)\right](0)$$

will have the desired properties: $A'(s) \ge 0$, $A(0) = \log \mathbb{E} g$, and $A(1) = \mathbb{E} \log g + \mathbb{E} F(|\nabla g|/g)$. The argument gives the inequality

$$\log \mathbb{E} g - \mathbb{E} \log g \le \mathbb{E} F\left(\frac{|\nabla g|}{g}\right). \tag{1-9}$$

If we set $g(x) = e^{f(x)}$ with $f : \mathbb{R}^n \to \mathbb{R}$ and use the chain rule $|\nabla g|/g = |\nabla f|$, we obtain

$$\log \mathbb{E} e^{f - \mathbb{E} f} \le \mathbb{E} F(|\nabla f|). \tag{1-10}$$

The last step is to show the pointwise estimate $F(s) \le 10 e^{s^2/2}(1+s)^{-1}$ for all $s \ge 0$. We remark that the obtained inequality (1-10) is stronger than (1-2) and it should be considered as a corollary of (1-10); however, due to a complicated expression for F we decided to state the main result in the form of (1-2).

The computation of $A'(s)$ is technical and is done in Section 2C, where we also explain how the expression $A(t)$ was "discovered". We should note that the main reason that makes $A' \ge 0$ is the fact that $k'/k'' > 0$ and the inequality[1]

$$1 - k'' - k' e^k \ge 0,$$

which for $k = -\log(\Phi'(t)/\Phi(t))$ serendipitously turns out to be an equality.

Sections 2A and 2B are technical and can be skipped when reading the paper for the first time. In these sections we show that $F \in C^2([0, \infty))$ is an increasing convex function with values $F(0) = F'(0) = 0$ and $F''(0) = 1$. Furthermore, the *modified* hessian matrix of

$$M(x, y) := \log x + \int_0^{y/x} e^{k((k')^{-1}(t))}\, dt \tag{1-11}$$

is positive semidefinite:

$$\begin{pmatrix} M_{xx} + \frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \ge 0 \quad \text{for all } (x, y) \in (0, \infty) \times [0, \infty). \tag{1-12}$$

In Section 2C we demonstrate that the condition (1-12) implies the inequality

$$M(\mathbb{E} g(X), 0) \le \mathbb{E} M(g(X), |\nabla g(X)|) \tag{1-13}$$

for all smooth bounded $g : \mathbb{R}^n \to (0, \infty)$. At the end of Section 2C, we deduce Theorem 1.1 and Corollary 1.2 from (1-13).

[1] It is an equality for $k = -\log(\Phi'(t)/\Phi(t))$ yet an inequality would be sufficient for our purposes.

As an application, in Section 3 we show that the dual inequality to (1-9), in the sense of duality described in Section 3.2 of [Ivanisvili et al. 2018], corresponds to the following theorem.

Theorem 1.3. *For any positive martingale* $\{\xi_n\}_{n\geq 0}$ *on a probability space* $([0,1],\mathcal{B},dx)$ *adapted to a discrete time dyadic filtration* $([0,1),\varnothing)=\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots$ *such that* $\xi_N=\xi_{N+1}=\cdots=\xi_\infty>0$ *for a sufficiently large* N, *we have*

$$\log\mathbb{E}\xi_\infty-\mathbb{E}\log\xi_\infty\leq\mathbb{E}G\biggl(\frac{\xi_\infty}{[\xi_\infty]^{1/2}}\biggr), \tag{1-14}$$

where $[\xi_\infty]=\sum_{k\geq 0}(\xi_{k+1}-\xi_k)^2$ *is the quadratic variation, and* $G(t):=\int_t^\infty\int_s^\infty r^{-2}e^{(s^2-r^2)/2}\,dr\,ds$.

In Lemma 3.2 we obtain the two-sided estimate

$$\tfrac{1}{3}\log(1+t^{-2})\leq G(t)\leq\log(1+t^{-2})\quad\text{for all } t\geq 0.$$

In particular, (1-14) implies that

$$\log\mathbb{E}\xi_\infty-\mathbb{E}\log\xi_\infty\leq\mathbb{E}\log\biggl(1+\frac{[\xi_\infty]}{\xi_\infty^2}\biggr). \tag{1-15}$$

Estimate (1-15) shows how well $\log\xi_\infty$ is concentrated around $\log\mathbb{E}\xi_\infty$ provided that one can control the quadratic variation of ξ_∞. Theorem 1.3 posits a duality approach developed in [Ivanisvili et al. 2018]. This may be considered as complementary to the e-entropy bound

$$\mathbb{E}e^{\xi_\infty-\mathbb{E}\xi_\infty}\leq\frac{e^{-\varepsilon}}{1-\varepsilon}$$

which holds for all discrete time simple martingales ξ_n (not necessarily positive and dyadic) provided that $[\xi_\infty]\leq\varepsilon^2$; see Corollary 1.12 in [Stolyarov et al. 2022].

The proof of (1-14) uses the special function

$$N(p,t):=\log(p)+\int_{p/\sqrt{t}}^\infty\int_s^\infty r^{-2}e^{(s^2-r^2)/2}\,dr\,ds$$

which we find by dualizing $M(x,y)=\log x+F(y/x)$. We deduce that N is *heat convex*, i.e.,

$$2N(p,t)\leq N(p+a,t+a^2)+N(p-a,t+a^2) \tag{1-16}$$

for all reals p, a, t such that $p\pm a\geq 0$ and $t\geq 0$. Finally, after iterating (1-16), we recover (1-14).

2. Proofs of Theorem 1.1 and Corollary 1.2

2A. *Step 1*: *an implicit function F and its properties.*

Let

$$k(x):=-\log(\log\Phi(x))'=\tfrac{1}{2}x^2+\log\biggl(\int_{-\infty}^x e^{-s^2/2}\,ds\biggr)\quad\text{for all } x\in\mathbb{R}.$$

Define a real-valued function F as

$$F(k'(t))=\int_{-\infty}^t k''(s)e^{k(s)}\,ds\quad\text{for all } t\in\mathbb{R}. \tag{2-1}$$

The following lemma, in particular, shows that F is well defined.

Lemma 2.1. *We have*

(1) $k'(-\infty)=0,\ k'(x)\sim x$ *as* $x\to\infty$, *and* $k''>0$ (*and hence* $k'>0$);

(2) $F:[0,\infty)\to[0,\infty),\ F(0)=F'(0)=0,\ F''(0)=1,\ F'(k')=e^k$, *and* $F''(k')=(k'/k'')e^k$.

Proof. Let us investigate the asymptotic behavior of k and its derivatives at $x=-\infty$. Let $x<0$, and for $m\ge 0$ define

$$I_m := e^{x^2/2}\int_{-\infty}^{x} e^{-s^2/2}s^{-m}\,ds.$$

Integration by parts reveals $I_m=-x^{-(m+1)}-(m+1)I_{m+2}$. By iterating we obtain

$$\begin{aligned} e^{x^2/2}\int_{-\infty}^{x} e^{-s^2/2}\,ds = I_0 &= -x^{-1}+x^{-3}-3\cdot x^{-5}+3\cdot 5\cdot x^{-7}+3\cdot 5\cdot 7\cdot I_8\\ &= -x^{-1}+x^{-3}-3x^{-5}+O(|x|^{-7}) \qquad \text{as } x\to-\infty, \end{aligned}$$

because $|I_8|\le\int_{-\infty}^{x}s^{-8}\,ds\le|x|^{-7}$. Thus, as $x\to-\infty$ we have

$$\begin{aligned} e^{k(x)} &= I_0 = -x^{-1}+x^{-3}-3x^{-5}+O(|x|^{-7}), & e^{-k(x)} &= -x-x^{-1}+2x^{-3}+O(|x|^{-5}),\\ k'(x) &= x+e^{-k(x)} = -x^{-1}+2x^{-3}+O(|x|^{-5}), & k''(x) &= 1-k'(x)e^{-k(x)} = x^{-2}+O(|x|^{-4}), \end{aligned}$$

$$k''(x)e^{k(x)} = -x^{-3}+O(|x|^{-5}), \qquad \text{and} \qquad \frac{k'(x)e^{k(x)}}{k''(x)} = 1+O(|x|^{-2}).$$

The claim

$$k'(x)=x+\frac{1}{e^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds}\sim x \quad \text{as } x\to\infty$$

is trivial. Next, we show that $k''>0$. By elementary calculus we have

$$k''=1-\frac{xe^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds+1}{\left(e^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds\right)^2}=\frac{e^{x^2}}{e^{2k(x)}}\left[\left(\int_{-\infty}^{x}e^{-s^2/2}\,ds\right)^2-xe^{-x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds-e^{-x^2}\right].$$

If we let $h(x):=e^{-x^2/2}$ and $H(x):=\int_{-\infty}^{x}e^{-t^2/2}\,dt$, then it suffices to show

$$u(x):=H^2-xhH-h^2>0.$$

Clearly $H'=h$ and $h'=-xh$. Next

$$\begin{aligned} u' &= 2Hh-hH+x^2hH-xh^2+2xh^2 = Hh+x^2hH+xh^2\\ &= (H+x^2H+xh)h = \left(H+h\frac{x}{1+x^2}\right)(1+x^2)h. \end{aligned}$$

Let $v(x)=H(x)+h(x)x/(1+x^2)$. Then, we have

$$v'(x)=h-h\frac{x^2}{1+x^2}+h\frac{1-x^2}{(1+x^2)^2}=\left(\frac{1}{1+x^2}+\frac{1-x^2}{(1+x^2)^2}\right)h=\frac{2}{(1+x^2)^2}h>0.$$

Since $v(-\infty)=0$ and $v'>0$, we obtain $v(x)>0$ for all $x\in\mathbb{R}$. In particular $u'>0$, and taking into account that $u(-\infty)=0$, we conclude $u(x)>0$ for all $x\in\mathbb{R}$.

To verify the second part of the lemma notice that $F(0)=0$ by considering the limit as $t\to-\infty$ in (2-1). Taking the derivative in t of (2-1) and dividing both sides by $k''>0$ we obtain $F'(k')=e^k$. Considering the limit as $t\to-\infty$ we realize $F'(0)=0$. Taking the second derivative gives $F''(k')=(k'/k'')e^k$. We have $(k'/k'')e^k=1+O(x^{-2})\to 1$ as $x\to-\infty$. Hence $F''(0)=1$, proving the lemma. □

It follows that $k'>0$ and $k':\mathbb{R}\to[0,\infty)$. Thus, we may consider the inverse map $t\mapsto k'(t)$ denoted by $(k')^{-1}:[0,\infty)\to\mathbb{R}$. After a suitable change of variables in (2-1), we write

$$F(x)=\int_{-\infty}^{(k')^{-1}(x)} k''(s)e^{k(s)}\,ds$$
$$=\int_0^x e^{k((k')^{-1}(u))}\,du, \qquad (\text{by } s=(k')^{-1}(u)),$$

which coincides with the expression announced in (1-8).

Lemma 2.2. *We have* $F(x)\le 10e^{x^2/2}(1+x)^{-1}$ *for all* $x\ge 0$.

Proof. Notice that $k'(u)=u+e^{-k(u)}\ge u$ (for all $u\in\mathbb{R}$) and $k''>0$. Therefore, $u\ge (k')^{-1}(u)$ for $u\ge 0$, so the inequality $k(u)\ge k((k')^{-1}(u))$ follows from the fact that $k'>0$. Thus

$$F(x)\le\int_0^x e^{k(u)}\,du=\int_0^x e^{u^2/2}\int_{-\infty}^u e^{-s^2/2}\,ds\,du\le\sqrt{2\pi}\int_0^x e^{u^2/2}\,du.$$

Next, we claim the simple chain of inequalities

$$\int_0^x e^{u^2/2}\,du \overset{(A)}{\le} \frac{2x}{1+x^2}e^{x^2/2} \overset{(B)}{\le} \frac{3}{1+x}e^{x^2/2}.$$

Indeed, inequality (A) follows from the fact that it is true at $x=0$ and

$$\frac{d}{dx}\left(\frac{2x}{1+x^2}e^{x^2/2}-\int_0^x e^{u^2/2}du\right)=e^{x^2/2}\left(1-\frac{4x^2}{(1+x^2)^2}\right)\ge e^{x^2/2}\left(1-\frac{4x^2}{(2x)^2}\right)=0.$$

In contrast, inequality (B) is immediate. Therefore, we conclude that

$$F(x)\le 3\sqrt{2\pi}e^{x^2/2}(1+x)^{-1}\le 10e^{x^2/2}(1+x)^{-1} \quad \text{for all } x\ge 0.$$ □

2B. *Step 2*: *Monge–Ampère type PDE.* Define

$$M(x,y)=\log x+F(y/x) \quad \text{for all } (x,y)\in(0,\infty)\times[0,\infty). \tag{2-2}$$

Clearly $M\in C^2$ and $M_y(x,0)=0$, where $M_x=\partial M/\partial x$ and $M_y=\partial M/\partial y$. Next, let us consider the matrix

$$A(x,y):=\begin{pmatrix} M_{xx}+\frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy}\end{pmatrix}. \tag{2-3}$$

We claim the following:

Lemma 2.3. *For each* $(x,y)\in(0,\infty)\times[0,\infty)$, *the matrix* $A(x,y)$ *is positive semidefinite with* $\det(A)=0$.

Proof. Let us calculate the partial derivatives of M. Let $t := yx^{-1}$. We have

$$M_x = x^{-1} - yx^{-2}F'(yx^{-1}) = x^{-1}(1 - tF'(t)), \quad M_y = x^{-1}F'(t),$$
$$M_{xx} = -x^{-2} + 2yx^{-3}F'(yx^{-1}) + (yx^{-2})^2 F''(yx^{-1}) = x^{-2}(-1 + 2tF'(t) + t^2F''(t)),$$
$$M_{yx} = -x^{-2}(F'(t) + tF''(t)), \quad M_{yy} = x^{-2}F''(t) \overset{\text{Lemma 2.1}}{>} 0.$$

To see that $A(x, y)$ is positive semidefinite, it suffices (due to the inequality $M_{yy} > 0$) to check that $\det(A) = 0$. We have

$$\det(A) = M_{xx}M_{yy} - M_{xy}^2 + \frac{M_y M_{yy}}{y} = x^{-4}\left[(-1 + 2tF' + t^2F'')F'' - (F' + tF'')^2 + \frac{F'F''}{t}\right]$$
$$= x^{-4}\left[-F'' - (F')^2 + \frac{F'F''}{t}\right].$$

Next, for $t = k'$ we have $F'(k') = e^k$ and $F''(k') = k'e^k/k''$ by Lemma 2.1. Therefore

$$-F'' - (F')^2 + \frac{F'F''}{t} = -\frac{k'e^k}{k''} - e^{2k} + \frac{e^{2k}}{k''} = \frac{e^{2k}}{k''}(1 - k'' - k'e^{-k}) = 0,$$

as $k'(x) = x + e^{-k(x)}$ (and hence $k'' = 1 - k'e^{-k}$). □

2C. *Step 3*: *the heat flow argument.* First we would like to give an explanation for how the flow is constructed. For simplicity consider $n = 1$. If we succeed in proving the inequality

$$M(\mathbb{E}g(\xi), 0) \le \mathbb{E}M(g(\xi), |g'(\xi)|), \quad g : \mathbb{R} \to (0, \infty), \tag{2-4}$$

where $\xi \sim \mathcal{N}(0, 1)$ and $M(x, y) = \log x + F(y/x)$, then we obtain

$$\log \mathbb{E}g + F(0) \le \mathbb{E}\log g + \mathbb{E}F(|g'|/g),$$

which for $g = e^f$ coincides with (1-10). So the goal is to prove (2-4). We consider a discrete approximation of ξ, namely, let

$$\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m),$$

where the ε_j are i.i.d. symmetric Bernoulli ± 1 random variables. By the central limit theorem,

$$\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}} \xrightarrow{d} \xi \quad \text{as } m \to \infty.$$

We hope to prove the *hypercube analog* of (2-4), i.e.,

$$M(\mathbb{E}\tilde{g}(\vec{\varepsilon}), 0) \le \mathbb{E}M(\tilde{g}(\vec{\varepsilon}), |D\tilde{g}(\vec{\varepsilon})|), \quad \tilde{g}(\vec{\varepsilon}) = g\left(\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}}\right), \tag{2-5}$$

for all $m \ge 1$, where the *discrete* gradient $|D\tilde{g}(\vec{\varepsilon})| := \sqrt{\sum_{j=1}^m |D_j\tilde{g}(\vec{\varepsilon})|^2}$ is defined as follows:

$$D_j\tilde{g}(\varepsilon_1, \dots, \varepsilon_m) = \frac{\tilde{g}(\varepsilon_1, \dots, \varepsilon_j, \dots, \varepsilon_m) - \tilde{g}(\varepsilon_1, \dots, -\varepsilon_j, \dots, \varepsilon_m)}{2} \quad \text{for } j = 1, \dots, m.$$

One sees that as $m \to \infty$ we have

$$D_j\tilde{g}(\vec{\varepsilon}) = g'\left(\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}}\right)\frac{\varepsilon_j}{\sqrt{m}} + O\left(\frac{1}{m}\right)$$

and

$$|D\tilde{g}(\vec{\varepsilon})| = \sqrt{\left[g'\left(\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}}\right)\right]^2 + O\left(\frac{1}{\sqrt{m}}\right)},$$

at least for bounded smooth functions g with uniformly bounded derivatives. Thus taking the limit $m \to \infty$ we observe that the right-hand side of (2-5) converges to the right-hand side of (2-4); in particular, (2-5) implies (2-4).

Next, we take this one step further and consider the inequality (2-5) for all $\tilde{g} : \{-1, 1\}^m \to \mathbb{R}$ instead of the specific functions defined in (2-5); in doing so we are ever so slightly enlarging the class of test functions to include those that are not invariant with respect to permutations of $(\varepsilon_1, \ldots, \varepsilon_n)$. To prove that

$$M(\mathbb{E}h, 0) \leq \mathbb{E}M(h, |Dh|) \qquad \text{for all } h : \{-1, 1\}^m \to (0, \infty) \quad \text{and all } m \geq 1, \tag{2-6}$$

one trivial argument would be to invoke the product structure of $\{-1, 1\}^m$. For example, if we manage to show an *intermediate* "4-point" inequality

$$M(\mathbb{E}_{\varepsilon_1}h, |D\mathbb{E}_{\varepsilon_1}h|) \leq \mathbb{E}_{\varepsilon_1}M(h, |Dh|), \tag{2-7}$$

where $\mathbb{E}_{\varepsilon_1}$ averages only with respect to ε_1, then by iterating (2-7) we deduce the inequality

$$M(\mathbb{E}h, 0) = M(\mathbb{E}_{\varepsilon_m} \cdots \mathbb{E}_{\varepsilon_1}h, |D\mathbb{E}_{\varepsilon_m} \cdots \mathbb{E}_{\varepsilon_1}h|) \leq \mathbb{E}_{\varepsilon_1} \cdots \mathbb{E}_{\varepsilon_m}M(h, |Dh|) = \mathbb{E}M(h, |Dh|).$$

Upon closer inspection, we see that (2-7) follows[2] from the 4-point inequality

$$2M(x, y) \leq M\big(x + a, \sqrt{a^2 + (y + b)^2}\big) + M\big(x - a, \sqrt{a^2 + (y - b)^2}\big) \tag{2-8}$$

for all real numbers $x,\ y,\ a,\ b$ such that $x \pm a > 0$. To prove (2-8) for one specific M seems to be a possible task; however, if we take into account that M is defined by (2-2) which involves an implicitly defined F, the 4-point inequality (2-8) becomes complicated (see [Ivanisvili and Volberg 2020], where one such inequality was proved for $M(x, y) = -\Re(x + iy)^{3/2}$ by tedious computations involving high degree polynomials with integer coefficients).

Expanding (2-8) at the point $(a, b) = (0, 0)$ via Taylor series, one easily obtains a necessary assumption: the infinitesimal form of (2-8), i.e.,

$$\begin{pmatrix} M_{xx} + \frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \geq 0. \tag{2-9}$$

Of course, the infinitesimal condition (2-9) does not necessarily imply its global two-point inequality (2-8) (and in particular (2-6)). Also, it may seem implausible to believe that the positive semidefiniteness of (2-9) implies the inequality (2-4) in Gauss space. Surprisingly this last guess turns out to be correct, and

[2] In fact they are equivalent provided that $y \mapsto M(x, y)$ is nondecreasing.

perhaps the reason lies in the fact that one only needs to verify (2-5) as $m \to \infty$ (and only for symmetric functions $\tilde{g}$). Let us "take the limit" and see how the heat flow arises.

Let $\mathbb{E}_{m-k}$ be the average with respect to the variables $\varepsilon_1, \dots, \varepsilon_{m-k}$, and let $\mathbb{E}^k$ be the average with respect to the remaining variables $\varepsilon_{m-k+1}, \dots, \varepsilon_m$. Then the 4-point inequality (2-7) implies that

$$k \mapsto \mathbb{E}^k M(\mathbb{E}_{m-k}\tilde{g}, |D\mathbb{E}_{m-k}\tilde{g}|) \quad \text{is nondecreasing on } [0, m].$$

The expression $\mathbb{E}^k M(\mathbb{E}_{m-k}\tilde{g}, |D\mathbb{E}_{m-k}\tilde{g}|)$ we rewrite as $\mathbb{E}^k M(A, B)$, where

$$A = \mathbb{E}_{m-k} g\left(\frac{\sum_{j=1}^k \varepsilon_j}{\sqrt{k}}\sqrt{\frac{k}{m}} + \frac{\sum_{j=k+1}^m \varepsilon_j}{\sqrt{m-k}}\sqrt{1-\frac{k}{m}}\right),$$
$$B = \sqrt{\frac{k}{m}\left[\mathbb{E}_{m-k} g'\left(\frac{\sum_{j=1}^k \varepsilon_j}{\sqrt{k}}\sqrt{\frac{k}{m}} + \frac{\sum_{j=k+1}^m \varepsilon_j}{\sqrt{m-k}}\sqrt{1-\frac{k}{m}}\right)\right]^2 + O\left(\frac{k}{m^{3/2}}\right)}.$$

Taking $k, m \to \infty$ so that $\frac{k}{m} \to s \in [0, 1]$, one can conclude that

$$s \mapsto \mathbb{E}_X M(\mathbb{E}_Y g(X\sqrt{s} + Y\sqrt{1-s}), \sqrt{s}|\mathbb{E}_Y g'(X\sqrt{s} + Y\sqrt{1-s})|) \quad \text{is nondecreasing on } [0, 1],$$

where $X, Y \in \mathcal{N}(0, 1)$ are independent and $\mathbb{E}_X$ takes the expectation with respect to the random variable X. In other words, if we let $U_s g(y) = \mathbb{E} g(y + \sqrt{s}X)$ to be a heat flow defined as

$$\frac{\partial}{\partial s} U_s g = \frac{1}{2}\frac{\partial^2}{\partial x^2} U_s g, \qquad U_0 g = g,$$

then

$$s \mapsto U_s M(U_{1-s} g, \sqrt{s}|U_{1-s} g'|)(0) \quad \text{is nondecreasing on } [0, 1]. \tag{2-10}$$

Luckily we may ignore all the steps by starting from the map (2-10) and taking its derivative in s to divine when it has nonnegative sign. Slightly abusing the notations, denote $D = \partial/\partial x$, and, for simplicity, let us work with the map $s \mapsto U_s M(U_{1-s} g, \sqrt{s} U_{1-s} g')$, where we omit the absolute value in the second argument of M. Let $b = U_{1-s} g$. Clearly $db/ds = -\frac{1}{2} D^2 b$. We have

$$\begin{aligned}\frac{d}{ds} U_s M(b, \sqrt{s} Db) &= \frac{1}{2} D^2 U_s M(b, \sqrt{s} Db) + U_s\left(-\frac{1}{2} D^2 b M_x + \left(\frac{1}{2\sqrt{s}} Db - \frac{\sqrt{s}}{2} D^3 b\right) M_y\right) \\ &= \frac{U_s}{2}\left(D(M_x Db + M_y \sqrt{s} D^2 b) - M_x D^2 b + \frac{M_y}{\sqrt{s}} Db - M_y \sqrt{s} D^3 b\right) \\ &= \frac{U_s}{2}\left(M_{xx}(Db)^2 + 2M_{xy}\sqrt{s}\, Db\, D^2 b + M_{yy} s (D^2 b)^2 + \frac{M_y}{\sqrt{s}} Db\right).\end{aligned}$$

Notice that

$$\begin{aligned}&M_{xx}(Db)^2 + 2M_{xy}\sqrt{s}\, Db\, D^2 b + M_{yy} s (D^2 b)^2 + \frac{M_y}{\sqrt{s}} Db \\ &\qquad = \begin{pmatrix} Db & \sqrt{s} D^2 b \end{pmatrix} \begin{pmatrix} M_{xx} + \frac{M_y}{\sqrt{s} Db} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \begin{pmatrix} Db \\ \sqrt{s} D^2 b \end{pmatrix} \ge 0.\end{aligned}$$

It remains to extend the argument to higher dimensions and put the absolute value back into the second argument of M.

Theorem 2.4. *Let* $M:(0,\infty)\times[0,\infty)\to\mathbb{R}$ *be such that* $M\in C^2$ *with* $M_y(x,0)=0$ *and*

$$\begin{pmatrix} M_{xx}+\frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \geq 0. \tag{2-11}$$

Then the map

$$s\mapsto U_s M(U_{1-s}g,\sqrt{s}|\nabla U_{1-s}g|) \quad \text{is nondecreasing on } [0,1] \tag{2-12}$$

for all smooth bounded functions $g:\mathbb{R}^n\to(0,\infty)$ *with uniformly bounded first and second derivatives.*

Proof. Let $M(x,y)=B(x,y^2)$. Let B_1 and B_2 be partial derivatives of B. Positive semidefiniteness of the matrix (2-11) in terms of B converts to

$$\begin{pmatrix} B_{11}(x,y^2)+2B_2(x,y^2) & 2yB_{12}(x,y^2) \\ 2yB_{12}(x,y^2) & 2B_2(x,y^2)+4y^2B_{22}(x,y^2) \end{pmatrix} \geq 0 \tag{2-13}$$

for all $x>0$ and all $y\geq 0$ (in fact this holds for all $y\in\mathbb{R}$). Next, let $G=U_{1-s}g$. Clearly $dG/ds=-\frac{1}{2}\Delta G$. We have

$$\frac{d}{ds}U_s B(U_{1-s}g, s|U_{1-s}\nabla g|^2)=\tfrac{1}{2}U_s[\Delta B(G,s|\nabla G|^2)-B_1\Delta G+2B_2|\nabla G|^2-2B_2 s\nabla G\cdot\nabla\Delta G].$$

Next, let $D_j=\partial/\partial x_j$. Then

$$\begin{aligned}
D_jB(G,s|\nabla G|^2)&=B_1D_jG+B_2sD_j|\nabla G|^2,\\
D_j^2B(G,s|\nabla G|^2)&=B_{11}(D_jG)^2+2B_{12}D_jGsD_j|\nabla G|^2+B_{22}s^2(D_j|\nabla G|^2)^2+B_1D_j^2G+B_2sD_j^2|\nabla G|^2,\\
\Delta B(G,s|\nabla G|^2)&=B_{11}|\nabla G|^2+2B_{12}\nabla G\cdot s\nabla|\nabla G|^2+B_{22}\big|s\nabla|\nabla G|^2\big|^2+B_1\Delta G+B_2s\Delta|\nabla G|^2.
\end{aligned}$$

Notice that $\Delta|\nabla G|^2=2\nabla G\cdot\nabla\Delta G+2\,\mathrm{Tr}(\mathrm{Hess}\,G)^2$. Therefore

$$\begin{aligned}
&\Delta B(G,s|\nabla G|^2)-B_1\Delta G+2B_2|\nabla G|^2-2B_2s\nabla G\cdot\nabla\Delta G\\
&\quad=B_{11}|\nabla G|^2+2B_{12}\nabla G\cdot s\nabla|\nabla G|^2+B_{22}\big|s\nabla|\nabla G|^2\big|^2+2B_2|\nabla G|^2+2B_2s\,\mathrm{Tr}(\mathrm{Hess}\,G)^2\\
&\quad\geq B_{11}|\nabla G|^2-2|B_{12}||\nabla G|\big|s\nabla|\nabla G|^2\big|+B_{22}\big|s\nabla|\nabla G|^2\big|^2+2B_2|\nabla G|^2+2B_2s\,\mathrm{Tr}(\mathrm{Hess}\,G)^2.
\end{aligned}\tag{2-14}$$

First we want to consider the case when $|\nabla G|=0$. We recall that $M(x,y)=B(x,y^2)$. Therefore $B_2(x,0)$ exists and is equal to $\frac{1}{2}M_{yy}(x,0)$ (due to the fact that $M_y(x,0)=0$). Also

$$\lim_{y\to 0}B_{12}(x,y^2)y=\tfrac{1}{2}M_{xy}(x,0)$$

and

$$\lim_{y\to 0}B_{22}(x,y^2)y^2=\lim_{y\to 0}\tfrac{1}{4}(M_{yy}(x,|y|)-2B_2(x,y^2))=0.$$

Therefore, if $|\nabla G| = 0$, then due to the inequality

$$\operatorname{Tr}(\operatorname{Hess} G)^2 |\nabla G|^2 = \sum_{j=1}^{n} |\nabla D_j G|^2 |\nabla G|^2 \geq \sum_{j=1}^{n} (\nabla D_j G \cdot \nabla G)^2 = \tfrac{1}{4}\big|\nabla |\nabla G|^2\big|^2, \tag{2-15}$$

the expression (2-14) simplifies to

$$2B_2(G,0)s\operatorname{Tr}(\operatorname{Hess} G)^2 = \tfrac{1}{2} M_{yy}(G,0)s \operatorname{Tr}(\operatorname{Hess} G)^2 \geq 0,$$

where the last inequality holds true by assumption (2-11), hence (2-14) is nonnegative.

If $|\nabla G| > 0$ then we proceed as follows: Assumption (2-11) implies $yM_{xx} + M_y \geq 0$. In particular, taking $y = 0$ we obtain $M_y(x,0) \geq 0$. Also it follows from (2-11) that $M_{yy} \geq 0$. Thus $M_y(x,y) \geq 0$ for all $y \geq 0$. In particular, $B_2(x,y^2) \geq 0$ for all $y > 0$ (and also for $y = 0$ as we just noticed that $B_2(x,0) = \frac{1}{2}M_{yy}(x,0) \geq 0$). Therefore, using inequality (2-15), we may estimate the last term in (2-14) from below as $B_2\big|s\nabla|\nabla G|^2\big|^2/(2s|\nabla G|^2)$. Finally,

$$B_{11}|\nabla G|^2 - 2|B_{12}||\nabla G|\big|s\nabla|\nabla G|^2\big| + B_{22}\big|s\nabla|\nabla G|^2\big|^2 + 2B_2|\nabla G|^2 + B_2\frac{\big|s\nabla|\nabla G|^2\big|^2}{2s|\nabla G|^2}$$
$$= \begin{pmatrix} -|\nabla G| & \frac{\sqrt{s}|\nabla|\nabla G|^2|}{2|\nabla G|}\end{pmatrix} \begin{pmatrix} B_{11}+2B_2 & 2\sqrt{s}|\nabla G||B_{12}| \\ 2\sqrt{s}|\nabla G||B_{12}| & 4s|\nabla G|^2 B_{22} + 2B_2 \end{pmatrix} \begin{pmatrix} -|\nabla G| \\ \frac{\sqrt{s}|\nabla|\nabla G|^2|}{2|\nabla G|}\end{pmatrix} \geq 0$$

by assumption (2-13) and the fact that B is evaluated at the point $(G, s|\nabla G|^2)$. □

Proof of Theorem 1.1. Notice that $U_s g(y) = \mathbb{E} g(y + \sqrt{s}X)$. Therefore, comparing the values of the map (2-12) at the endpoints $s = 0$ and $s = 1$ we obtain

$$M(\mathbb{E} g(X), 0) = U_0 M(U_1 g, \sqrt{0}|U_1\nabla g|)(0) \leq U_1 M(U_0 g, |U_0 \nabla g|)(0) = \mathbb{E} M(g(X), |\nabla g(X)|).$$

In particular, for $g = e^f$ where $f \in C_0^\infty(\mathbb{R}^n)$, we obtain

$$\log \mathbb{E} e^{f(X)} \leq \mathbb{E} f(X) + \mathbb{E} F(|\nabla f(X)|).$$

The pointwise inequality $F(x) \leq 10 e^{x^2/2}(1+x)^{-1}$ from Lemma 2.2 finishes the proof of Theorem 1.1 □

Proof of Corollary 1.2. Let $d\mu = e^{-u(x)}\,dx$ be the density of the log-concave random vector X with $\operatorname{Hess} u \geq R I_{n\times n}$ for some $R > 0$. It follows from [Caffarelli 2000] that there exists a convex function $\psi : \mathbb{R}^n \to \mathbb{R}$ such that the Brenier map $T = \nabla\psi$ pushes forward the Gaussian measure

$$d\gamma_n(x) = \frac{e^{-|x|^2/2}}{\sqrt{(2\pi)^n}}\,dx$$

onto $d\mu$ and such that $0 \leq \operatorname{Hess}\psi \leq (1/\sqrt{R}) I_{n\times n}$. Next, apply the inequality

$$\log \int_{\mathbb{R}^n} e^{f(x)}\,d\gamma_n(x) \leq \int_{\mathbb{R}^n} f(x)\,d\gamma_n(x) + \int_{\mathbb{R}^n} F(|\nabla f(x)|)\,d\gamma_n(x) \tag{2-16}$$

with $f(x) = h(\nabla\psi(x))$ for an arbitrary $h \in C_0^\infty(\mathbb{R}^n)$. Then notice that

$$|\nabla f(x)| = |\operatorname{Hess}\psi \nabla h(\nabla\psi(x))| \leq \frac{1}{\sqrt{R}} |\nabla h(\nabla\psi(x))|.$$

Since $F' > 0$, we conclude that

$$F(|\nabla f(x)|) \le F\left(\frac{1}{\sqrt{R}}|\nabla h(\nabla\psi(x))|\right) \le 10e^{|\nabla h(\nabla\psi(x))|^2/(2R)}(1+R^{-1/2}|\nabla h(\nabla\psi)|)^{-1}.$$

The preceding inequality together with (2-16) implies that

$$\log\int_{\mathbb{R}^n} e^{h(x)}\,d\mu(x) \le \int_{\mathbb{R}^n} h(x)\,d\mu(x) + 10\int_{\mathbb{R}^n} e^{|\nabla h(x)|^2/(2R)}(1+R^{-1/2}|\nabla h(x)|)^{-1}\,d\mu(x)$$

for all $h \in C_0^\infty(\mathbb{R}^n)$. This finishes the proof of Corollary 1.2. □

Remark 2.5. The transport map $T(x_1,\dots,x_n) = (\Phi(x_1),\dots,\Phi(x_n))$ pushes forward the standard Gaussian measure onto the uniform measure on $[0,1]^n$, and it is $(2\pi)^{-1/2}$ Lipschitz. Therefore, the inequality (2-16) applied to $f(x) = h(T(x))$ for a smooth $h : [0,1]^n \to \mathbb{R}$ implies that

$$\log \mathbb{E}e^{h(Y)-\mathbb{E}h(Y)} \le \mathbb{E}F((2\pi)^{-1/2}|\nabla h(Y)|) \le \mathbb{E}e^{|\nabla h(Y)|^2/(4\pi)}(1+(2\pi)^{-1/2}|\nabla h|),$$

where $Y \sim \mathrm{unif}([0,1]^n)$. We thank an anonymous referee for this remark.

3. Applications: the proofs of Theorem 1.3 and estimate (1-15)

Let us recall the definition of dyadic martingales. For each $n \ge 0$ we denote by $\mathcal{D}_n$ the dyadic intervals belonging to $[0,1)$ of level n, i.e.,

$$\mathcal{D}_n = \left\{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right),\ k = 0,\dots,2^n-1\right\}.$$

Given $\xi \in L^1([0,1], dx)$, define a *dyadic martingale* $\{\xi_k\}_{k\ge 0}$ as

$$\xi_n(x) := \sum_{I\in\mathcal{D}_n} \langle\xi\rangle_I \mathbb{1}_I(x), \quad n \ge 0,$$

where

$$\langle\xi\rangle_I = \frac{1}{|I|}\int_I \xi\,dx;$$

here $|I|$ denotes the Lebesgue length of I. If we let $\mathcal{F}_n$ be the σ-algebra generated by the dyadic intervals in $\mathcal{D}_n$, then $\xi_n = \mathbb{E}(\xi \mid \mathcal{F}_n)$ is the martingale with respect to the increasing filtration $\{\mathcal{F}_k\}_{k\ge 0}$. Next we define the quadratic variation

$$[\xi] = \sum_{n\ge 0} d_n^2,$$

where $d_n := \xi_n - \xi_{n-1}$ is the martingale difference sequence. In what follows, to avoid the issues with convergence of the infinite series we will be assuming that all but finitely many d_n are zero, i.e., $\xi_N = \xi_{N+1} = \cdots = \xi$ for N sufficiently large. Such martingales we call simple dyadic martingales; they are also known as Walsh–Paley martingales [Hytönen et al. 2016].

Let

$$N(p,t) := \log(p) + G(p/\sqrt{t})$$

for $p > 0,\ t \geq 0$, where

$$G(s) = \int_s^\infty \int_r^\infty u^{-2} e^{(r^2-u^2)/2}\, du\, dr, \quad s > 0.$$

Lemma 3.1. *For all real numbers $p,\ a,\ t$ we have*

$$N(p+a, t+a^2) + N(p-a, t+a^2) \geq 2N(p,t), \tag{3-1}$$

provided that $p \pm a > 0$ and $t \geq 0$.

Proof. First we verify that $N(p,t)$ satisfies the backward heat equation

$$\tfrac{1}{2} N_{pp} + N_t = 0. \tag{3-2}$$

Indeed, we have

$$\begin{aligned} N_{pp} + 2N_t &= -\frac{1}{p^2} + \frac{G''(p/\sqrt{t})}{t} - G'(p/\sqrt{t})pt^{-3/2} \\ &= \frac{1}{p^2}(-1 + s^2 G''(s) - s^3 G'(s)), \end{aligned} \tag{3-3}$$

where $s = p/\sqrt{t}$. Direct calculations show that

$$G'(s) = -e^{s^2/2} \int_s^\infty u^{-2} e^{-u^2/2}\, du \quad \text{and} \quad G''(s) = s^{-2} - s e^{s^2/2} \int_s^\infty u^{-2} e^{-u^2/2}\, du. \tag{3-4}$$

Substituting (3-4) into (3-3) we see that the expression in (3-3) is zero.

Next, we claim that $t \mapsto N(p,t)$ is concave. Indeed,

$$\begin{aligned} N_t &= -\frac{1}{2} pt^{-3/2} G'(p/\sqrt{t}), \\ N_{tt} &= \frac{1}{4} p^2 t^{-3} G''(p/\sqrt{t}) + \frac{3}{4} pt^{-5/2} G'(p/\sqrt{t}) \\ &= \frac{1}{4t^2}[s^2 G''(s) + 3s G'(s)]. \end{aligned}$$

Since $N_{pp} + 2N_t = 0$, we have $G'' = s^{-2} + sG'(s)$ by (3-3). Therefore, the sign of N_{tt} coincides with the sign of $1 + (s^3 + 3s)G'(s)$. Using (3-4), it suffices to show that

$$\varphi(s) := \frac{e^{-s^2/2}}{s^3 + 3s} - \int_s^\infty u^{-2} e^{-u^2/2}\, du \leq 0 \quad \text{for all } s \geq 0.$$

We have $\varphi(\infty) = 0$ and

$$\varphi'(s) = e^{-s^2/2}\left[-\frac{1}{3+s^2} - \frac{3+3s^2}{(3s+s^3)^2} + \frac{1}{s^2}\right] = \frac{6e^{s^2/2}}{(3s+s^2)^2} \geq 0,$$

thereby $\varphi(s) \leq 0$, and hence $t \mapsto N(p,t)$ is concave for $t \geq 0$.

Next, consider the process

$$X_s = N(p + B_s, t + s),$$

where B_s is the standard Brownian motion starting at zero. It follows from Ito's formula that X_s is a martingale. Indeed, we have

$$dX_s = N_s ds + N_p dB_s + \tfrac{1}{2} N_{pp} ds \overset{(3\text{-}2)}{=} N_p dB_s.$$

Define the stopping time

$$\tau = \inf\{s \geq 0 : B_s \notin (-a, a)\}.$$

Set $Y_s = X_{\min\{s,\tau\}}$ for $s \geq 0$. Clearly Y_s is a martingale. On the one hand $Y_0 = N(p, t)$. On the other hand

$$\begin{aligned} \mathbb{E} Y_\infty &= \mathbb{E} N(p+B_\tau, t+\tau) \\ &= \mathbb{E}(N(p-a, t+\tau) \mid B_\tau = -a)\mathbb{P}(B_\tau = -a) + \mathbb{E}(N(p+a, t+\tau) \mid B_\tau = -a)\mathbb{P}(B_\tau = -a) \\ &\leq \tfrac{1}{2}[N(p-a, t+\mathbb{E}(\tau \mid B_\tau = -a)) + N(p+a, t+\mathbb{E}(\tau \mid B_\tau = a))], \quad \text{(by concavity of } t \mapsto N(p,t)\text{)}. \end{aligned}$$

Finally, as $B_s^2 - s$ is a martingale, we have that $0 = \mathbb{E}(B_\tau^2 - \tau) = a^2 - \mathbb{E}\tau$. By symmetry we obtain $\mathbb{E}(\tau \mid B_\tau = -a) = \mathbb{E}(\tau \mid B_\tau = a) = a^2$. Thus the lemma follows from the optional stopping theorem. □

Before we complete the proof of Theorem 1.3 let us make a remark. If $N(p, t)$ is an arbitrary smooth function satisfying the backwards heat equation (3-2) and inequality (3-1), then $t \mapsto N(p, t)$ must be concave. In other words, the concavity of $t \mapsto N(p, t)$ is necessary and sufficient for the inequality (3-1) to hold provided that N solves the backwards heat equation. Indeed, let $r(a) = N(p + a, t + a^2)$. Then

$$\begin{aligned} r'(a) &= N_p + 2aN_t, \\ r''(a) &= N_{pp} + 4aN_{pt} + 2N_t + 4a^2 N_{tt} \overset{(3\text{-}2)}{=} 4aN_{pt} + 4a^2 N_{tt}, \\ r'''(a) &= 4N_{pt} + 4aN_{ppt} + 8a^2 N_{ptt} + 8aN_{tt} + 4a^2 N_{ttp} + 8a^3 N_{ttt} \overset{(3\text{-}2)}{=} 4N_{pt} + 12a^2 N_{ptt} + 8a^3 N_{ttt}, \\ r''''(a) &= 4N_{ptp} + 8aN_{ptt} + 24aN_{ptt} + 12a^2 N_{pttp} + 24a^3 N_{pttt} + 24a^2 N_{ttt} + 8a^3 N_{tttp} + 16a^4 N_{tttt} \\ &\overset{(3\text{-}2)}{=} 4N_{ppt} + 32aN_{ptt} + 32a^3 N_{pttt} + 16a^4 N_{tttt}. \end{aligned}$$

By Taylor's formula we have

$$\begin{aligned} N(p + a, t + a^2) + N(p - a, t + a^2) &= r(a) + r(-a) = 2r(0) + r''(0)a^2 + r''''(0)\frac{a^4}{12} + o(a^4) \\ &= 2N(p, t) + N_{ppt}(p, t)\frac{a^4}{3} + o(a^4) \\ &\overset{(3\text{-}2)}{=} 2N(p, t) - N_{tt}(p, t)\frac{2a^4}{3} + o(a^4). \end{aligned}$$

Thus it follows from (3-2) that

$$\lim_{a \to 0} \frac{N(p + a, t + a^2) + N(p - a, t + a^2) - 2N(p, t)}{a^4} = -\tfrac{2}{3} N_{tt}$$

is nonnegative, i.e., $t \mapsto N(p, t)$ is concave.

Now we are ready to complete the proof of Theorem 1.3. Let $N \geq 0$ be such that $\xi_N = \xi_{N+1} = \cdots = \xi$. We have

$$\begin{aligned}\mathbb{E}N(\xi,[\xi]) &= \mathbb{E}N(\xi_N,[\xi_N])\\ &= \mathbb{E}(\mathbb{E}(N(\xi_0+(\xi_1-\xi_0)+\cdots+(\xi_N-\xi_{N-1}),(\xi_1-\xi_0)^2+\cdots+(\xi_N-\xi_{N-1})^2)\mid\mathcal{F}_{N-1})).\end{aligned}$$

Notice that the random variables

$$\eta = \xi_0+(\xi_1-\xi_0)+\cdots+(\xi_{N-1}-\xi_{N-2}) \quad \text{and} \quad \zeta = (\xi_1-\xi_0)^2+\cdots+(\xi_{N-1}-\xi_{N-2})^2$$

are $\mathcal{F}_{N-1}$ measurable. Yet on each atom Q of $\mathcal{F}_{N-1}$ the random variable $\xi_{N-1}-\xi_N$ takes values $\pm A$ with equal probabilities $\frac{1}{2}|Q|$. Then it follows from (3-1) that

$$\mathbb{E}N(\xi_N,[\xi_N]) \geq \mathbb{E}N(\xi_{N-1},[\xi_{N-1}]).$$

Iterating this inequality and using the boundary value $N(p,0)=\log p$ for $p>0$, we obtain

$$\mathbb{E}N(\xi,[\xi]) \geq \mathbb{E}N(\xi_0,0) = \ln \mathbb{E}\xi.$$

This finishes the proof of Theorem 1.3. □

Inequality (1-15) follows from the following lemma.

Lemma 3.2. *We have that*

$$\log(1+y^{-2}) \geq \int_y^\infty \int_x^\infty e^{(-t^2+x^2)/2} t^{-2}\,dt\,dx \geq \tfrac{1}{3}\log(1+y^{-2}) \tag{3-5}$$

for all $y>0$.

Proof. For a positive constant $C>0$, consider a map

$$h(y;C) = \int_y^\infty \int_x^\infty e^{(-t^2+x^2)/2} t^{-2}\,dt\,dx - C\log(1+y^{-2}), \qquad y>0.$$

Notice that $h(\infty;C)=0$. To prove the second inequality in (3-5) (or the first inequality in (3-5)) it suffices to show that

$$h_y(y;C) = -\int_y^\infty \frac{e^{(-t^2+y^2)/2}}{t^2}\,dt + 2C\frac{1}{y^3+y} \leq 0 \tag{3-6}$$

for $C=\frac{1}{3}$ (or $h_y(y;C)\geq 0$ for $C=1$). Next, consider

$$\psi(y;C) = e^{-y^2/2}h'(y;C) = -\int_y^\infty e^{-t^2/2}t^{-2}\,dt + 2C\frac{e^{-y^2/2}}{y^3+y}.$$

Clearly $\psi(\infty;C)=0$. To show (3-6) for $C=\frac{1}{3}$ (or its reverse inequality when $C=1$), it suffices to verify that

$$\psi_y(y;C) = \frac{Ce^{-y^2/2}}{y^2}\left(\frac{1}{C} - 2\frac{(3y^2+1)}{(y^2+1)^2} - \frac{y^2}{y^2+1}\right) = \frac{Ce^{-y^2/2}}{y^2}\left(\frac{1}{C} - 1 + \frac{4}{(y^2+1)^2} - \frac{5}{y^2+1}\right) \geq 0$$

for $C = \frac{1}{3}$ (or the reverse inequality for $C = 1$). Let $s = (y^2+1)^{-1} \in [0,1]$. Then $-1+4t^2-5t$ is minimized on $[0,1]$ when $t = \frac{5}{8}$ and its minimal value is $-\frac{41}{16}$ (or maximized on $[0,1]$ when $t = 0$ and its maximal value is -1). The lemma is proved. □

4. Concluding remarks

One may ask how we guessed $N(p,t)$ which played an essential role in the proof of Theorem 1.3. There is a general argument [Ivanisvili et al. 2018] which informally says that *any estimate in Gauss space* (*or more generally on the hamming cube*) *involving f and its gradient has a corresponding* ***dual estimate*** *for a stopped Brownian motion and its quadratic variation* (*or more generally dyadic square function*). For example, to prove inequality (1-2), there was a certain function $M(x,y)$ used in the proof. This function satisfies the Monge–Ampere type PDE

$$\det \begin{pmatrix} M_{xx} + \frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} = 0 \tag{4-1}$$

with a boundary condition $M(x,0) = \log(x)$, so that the matrix in (4-1) is positive definite. Suppose we would like to solve the PDE (4-1) in general. Using exterior differential systems (see details in [Ivanisvili and Volberg 2018]), the PDE may be linearized to the backwards heat equation; namely, locally the solutions can be parametrized as

$$\begin{cases} M(x,y) = -px + \sqrt{t}\,y + u(p,t), \\ x = -u_p(p,t), \\ y = 2\sqrt{t}\,u_t(p,t), \end{cases}$$

where u satisfies the backwards heat equation

$$\begin{cases} u_t + \frac{1}{2}u_{pp} = 0, \\ u(M_x(x,0),0) = M(x,0) - xM_x(x,0), \end{cases}$$

with $t \geq 0$ and $p \in \Omega \subset \mathbb{R}$. An important observation is that if u happens to satisfy

$$u(p+a, t+a^2) + u(p-a, t+a^2) \geq 2u(p,t),$$

then under some additional assumptions on u, one expects an identity

$$M(x,y) = \sup_t \inf_p \{-px + \sqrt{t}\,y + u(p,t)\} = \inf_p \sup_t \{-px + \sqrt{t}\,y + u(p,t)\},$$

which, if true, implies that M satisfies the 4-point inequality (2-8); see [Ivanisvili et al. 2017; 2018] for more details. These functions $M(x,y)$ and $u(p,t)$ we call dual to each other. One may verify that for our particular M defined by (1-11), the corresponding dual $u(p,t)$ is

$$u(p,t) = 1 + \log(-p) + \int_{-p/\sqrt{t}}^{\infty} \int_s^{\infty} r^{-2} e^{(-r^2+s^2)/2}\, dr\, ds, \qquad p < 0, \quad t \geq 0,$$

which coincides with $N(p,t)$ after subtracting 1 and reflecting in the variable p.

Using this approach, one could try to prove the 4-point inequality (2-8), which would imply

$$\mathbb{E}M(g, |Dg|) \geq M(\mathbb{E}g, 0) \quad \text{for all } \ g : \{-1, 1\}^n \to \mathbb{R}_+.$$

So, one may hope to obtain Theorem 1.1 on the hamming cube after substituting $g = e^f$. However, we did not proceed with this path on the unfortunate grounds that the chain rule misbehaves on the hamming cube, i.e., the identity $|De^f|/e^f = |Df|$ does not hold. Therefore, to prove (1-2) on the hamming cube perhaps different ideas are needed.

Our last remark is that one may provide another proof of (1-15) using a simpler function compared to N (what we call *the supersolution*). Indeed, consider

$$N^{\sup}(p, t) = \tfrac{1}{2}\log(p^2 + t), \quad t \geq 0, \quad p > 0.$$

Notice that

$$N^{\sup}(p, 0) = \log(p), \tag{4-2}$$

$$\frac{N^{\sup}_{pp}}{2} + N^{\sup}_t = \frac{t + t^2}{2(p^2 + t)^2} \geq 0, \tag{4-3}$$

$$N^{\sup}_{tt} = -\frac{1}{2}\frac{1}{(t + p^2)^2} \leq 0. \tag{4-4}$$

Using the same argument as in the proof of (1-14) we verify that $N^{\sup}(p, t)$ satisfies (3-1). Notice that $N^{\sup}$ does not solve the backwards heat equation; however, due to inequality (4-3) the stochastic process Y_s constructed in the proof of (3-1) will be a submartingale which is sufficient for the proof of (3-1). Thus we obtain

$$\log \mathbb{E}\xi_\infty - \mathbb{E}\log \xi_\infty \leq \frac{1}{2}\mathbb{E}\log\left(1 + \frac{[\xi_\infty]}{\xi_\infty^2}\right),$$

which improves on (1-15) by a factor of $\frac{1}{2}$.

The supersolution $N^{\sup}(p, t)$ was guessed from the form of the inequality (1-15) by considering

$$\log(p) + C\log(1 + t/p^2)$$

and choosing an optimal constant C (in our case $C = \frac{1}{2}$ worked well). It was a good coincidence that such an $N^{\sup}$ satisfies (4-2), (4-3), and (4-4). However, if one tries to construct a supersolution to the inequality (1-2) one may hope that, by chance,

$$M(x, y) = \log(x) + Ce^{y^2/(2x^2)}(1 + y/x)^{-1}$$

may work for some positive C. A direct calculation shows that there is no positive constant C such that inequality (2-11) holds.

Acknowledgments

We would like to thank the referees for helpful suggestions and remarks. The authors also thank Sergey Bobkov and Changfeng Gui for helpful discussions and comments. Ivanisvili was supported in part by NSF grants DMS-2152346 and CAREER-DMS-2152401.

References

[Bobkov and Götze 1999] S. G. Bobkov and F. Götze, "Exponential integrability and transportation cost related to logarithmic Sobolev inequalities", *J. Funct. Anal.* **163**:1 (1999), 1–28. MR Zbl

[Caffarelli 2000] L. A. Caffarelli, "Monotonicity properties of optimal transportation and the FKG and related inequalities", *Comm. Math. Phys.* **214**:3 (2000), 547–563. MR Zbl

[Efraim and Lust-Piquard 2008] L. B. Efraim and F. Lust-Piquard, "Poincaré type inequalities on the discrete cube and in the CAR algebra", *Probab. Theory Related Fields* **141**:3-4 (2008), 569–602. MR Zbl

[Hytönen et al. 2016] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces, I: Martingales and Littlewood–Paley theory*, Ergebnisse der Mathematik (3) **63**, Springer, 2016. MR Zbl

[Ivanisvili and Volberg 2018] P. Ivanisvili and A. Volberg, "Isoperimetric functional inequalities via the maximum principle: the exterior differential systems approach", pp. 281–305 in 50 *years with Hardy spaces*, edited by A. Baranov and N. Nikolski, Oper. Theory Adv. Appl. **261**, Birkhäuser, Cham, 2018. MR Zbl

[Ivanisvili and Volberg 2020] P. Ivanisvili and A. Volberg, "Poincaré inequality $3/2$ on the Hamming cube", *Rev. Mat. Iberoam.* **36**:1 (2020), 79–97. MR Zbl

[Ivanisvili et al. 2017] P. Ivanisvili, F. Nazarov, and A. Volberg, "Hamming cube and martingales", *C. R. Math. Acad. Sci. Paris* **355**:10 (2017), 1072–1076. MR Zbl

[Ivanisvili et al. 2018] P. Ivanisvili, F. Nazarov, and A. Volberg, "Square functions and the Hamming cube: duality", *Discrete Anal.* **2018** (2018), art. id. 1. MR Zbl

[Moser 1971] J. Moser, "A sharp form of an inequality by N. Trudinger", *Indiana Univ. Math. J.* **20** (1971), 1077–1092. MR Zbl

[Onofri 1982] E. Onofri, "On the positivity of the effective action in a theory of random surfaces", *Comm. Math. Phys.* **86**:3 (1982), 321–326. MR Zbl

[Stolyarov et al. 2022] D. Stolyarov, V. Vasyunin, P. Zatitskiy, and I. Zlotnikov, "Sharp moment estimates for martingales with uniformly bounded square functions", *Math. Z.* **302**:1 (2022), 181–217. MR Zbl

Received 19 May 2021. Revised 25 Nov 2021. Accepted 20 Dec 2021.

PAATA IVANISVILI: pivanisv@uci.edu
Department of Mathematics, University of California, Irvine, Irvine, CA, United States

RYAN RUSSELL: rrussell14@g.ucla.edu
Current address: *Irvine, CA, United States*
Department of Mathematics and Statistics, California State University, Long Beach, CA, United States

Analysis & PDE

msp.org/apde

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

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 16 No. 5 2023