A martingale bound for the entropy associated with a trimmed filtration on $\mathbb{R}^d$

Alexei Kulik[a,*], **Taras Tymoshkevych**[b]

[a]*Institute of Mathematics, National Academy of Science, Kiev, Ukraine*
[b]*Kyiv National Taras Shevchenko University, Ukraine*

kulik.alex.m@gmail.com (A. Kulik), tymoshkevych@gmail.com (T. Tymoshkevych)

Abstract   Using martingale methods, we provide bounds for the entropy of a probability measure on $\mathbb{R}^d$ with the right-hand side given in a certain integral form. As a corollary, in the one-dimensional case, we obtain a weighted log-Sobolev inequality.

1   Introduction

A probability measure μ on $\mathbb{R}^d$ is said to satisfy the *log-Sobolev inequality* if for every smooth compactly supported function $f : \mathbb{R}^d \to \mathbb{R}$, the *entropy* of f^2, which by definition equals

$$\mathbf{Ent}_\mu \, f^2 = \int_{\mathbb{R}^d} f^2 \log f^2 \, d\mu - \left(\int_{\mathbb{R}^d} f^2 \, d\mu \right) \log\left(\int_{\mathbb{R}^d} f^2 \, d\mu \right),$$

possesses a bound

$$\mathbf{Ent}_\mu \, f^2 \le 2c \int_{\mathbb{R}^d} \|\nabla f\|^2 \, d\mu \tag{1}$$

---

*Corresponding author.

VEX

with some constant c. The least possible constant c such that (1) holds for every compactly supported smooth f is called the log-Sobolev constant for the measure μ; the multiplier 2 in (1) is chosen in such a way that for the standard Gaussian measure on $\mathbb{R}^d$, its log-Sobolev constant equals 1.

The *weighted log-Sobolev inequality* has the form

$$\mathbf{Ent}_\mu \, f^2 \le 2 \int_{\mathbb{R}^d} \|W\nabla f\|^2 \, d\mu, \qquad (2)$$

where the function W, taking values in $\mathbb{R}^{d\times d}$, has the meaning of a *weight*. Clearly, one can consider (1) as a particular case of (2) with constant weight W equal to $\sqrt{c}$ multiplied by the identity matrix. The problem of giving explicit conditions on μ that ensure the log-Sobolev inequality or its modifications is intensively studied in the literature, in particular, because of numerous connections between these inequalities with measure concentration, semigroup properties, and so on (see, e.g., [8]). Motivated by this general problem, in this paper, we propose an approach that is based mainly on martingale methods and provides explicit bounds for the entropy with the right-hand side given in a certain integral form.

Our approach is motivated by the well-known fact that, on a path space of a Brownian motion, the log-Sobolev inequality possesses a simple proof based on fine martingale properties of the space (cf. [1, 6]). We observe that a part of this proof is, to a high extent, insensitive w.r.t. the structure of the probability space; we formulate a respective martingale bound for the entropy in Section 1.1. To apply this general bound on a probability space of the form $(\mathbb{R}^d, \mu)$, one needs a proper martingale structure therein. In Section 2, we introduce such a structure in terms of a *trimming filtration*, defined in terms of a set of *trimmed regions* in $\mathbb{R}^d$. This leads to an integral bound for the entropy on $(\mathbb{R}^d, \mu)$. In Section 3, we show the way how this bound can be used to obtain a weighted log-Sobolev inequality; this is made in the one-dimensional case $d = 1$, although we expect that similar arguments should be effective for the multidimensional case as well; this is a subject of our further research.

1.1   A martingale bound for the entropy

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space with filtration $\mathbb{F} = \{\mathcal{F}_t, t \in [0, 1]\}$, which is right-continuous and complete, that is, every $\mathcal{F}_t$ contains all $\mathbf{P}$-null sets from $\mathcal{F}$. Let $\{M_t, t \in [0, 1]\}$ be a nonnegative square-integrable martingale w.r.t. $\mathbb{F}$ on this space, with càdlàg trajectories. We will use the following standard facts and notation (see [4]).

The martingale M has unique decomposition $M = M^c + M^d$, where M^c is a continuous martingale, and M^d is a purely discontinuous martingale (see [4], Definition 9.20). Denote by $\langle M^c \rangle$ the *quadratic variation* of M^c, by

$$[M]_t = \langle M^c \rangle_t + \sum_{s \le t} (M_s - M_{s-})^2$$

the *optional quadratic variation* of M, and by $\langle M \rangle$ the *predictable quadratic variation* of M, that is, the projection of $[M]$ on the set of $\mathbb{F}$-predictable processes. Alternatively, $\langle M \rangle$ is identified as the $\mathbb{F}$-predictable process that appears in the Doob–Meyer

decomposition for M^2, that is, the $\mathbb{F}$-predictable nondecreasing process A such that $A_0 = 0$ and $M^2 - A$ is a martingale.

For a nonnegative r.v. ξ, define its entropy by $\mathbf{Ent}\,\xi = \mathbf{E}\xi \log \xi - \mathbf{E}\xi \log(\mathbf{E}\xi)$ with the convention $0 \log 0 = 0$.

Theorem 1. *Let the σ-algebra $\mathcal{F}_0$ be degenerate. Then for any nonnegative square-integrable martingale $\{M_t, t \in [0, 1]\}$ with càdlàg trajectories,*

$$\mathbf{Ent}\,M_1 \le \mathbf{E}\int_0^1 \frac{1}{M_{t-}}d\langle M\rangle_t.$$

Proof. Consider first the case where

$$c_1 \le M_t \le c_2, \quad t \in [0, 1], \tag{3}$$

with some positive constants c_1, c_2. Consider a smooth function Φ, bounded with all its derivatives, such that

$$\Phi(x) = x \log x, \quad x \in [c_1, c_2].$$

Then by the Itô formula (see [4], Theorem 12.19),

$$\Phi(M_1) - \Phi(M_0) = \int_0^1 \Phi'(M_{t-})\,dM_t + \frac{1}{2}\int_0^1 \Phi''(M_{t-})\,d\langle M^c\rangle_t$$
$$+ \sum_{0<t\le 1}\big[\Phi(M_t) - \Phi(M_{t-}) - \Phi'(M_{t-})(M_t - M_{t-})\big].$$

Clearly,

$$\mathbf{E}\int_0^1 \Phi'(M_{t-})\,dM_t = 0.$$

Because $\mathcal{F}_0$ is assumed to be degenerate, $M_0 = \mathbf{E}[M_1|\mathcal{F}_0] = \mathbf{E}M_1$ a.s., and hence

$$\mathbf{Ent}\,M_1 = \mathbf{E}\big(\Phi(M_1) - \Phi(M_0)\big)$$
$$= \frac{1}{2}\mathbf{E}\int_0^1 \Phi''(M_{t-})\,d\langle M^c\rangle_t$$
$$+ \mathbf{E}\sum_{0<t\le 1}\big[\Phi(M_t) - \Phi(M_{t-}) - \Phi'(M_{t-})(M_t - M_{t-})\big].$$

For $x \in [c_1, c_2]$, we have $\Phi'(x) = 1 + \log x$ and $\Phi''(x) = 1/x$. Observe that for any x, δ such that $x, x + \delta \in [c_1, c_2]$,

$$\Phi(x + \delta) - \Phi(x) - \Phi'(x)\delta = (x + \delta)\log(x + \delta) - x \log x - \delta(1 + \log x)$$
$$= (x + \delta)\log\left(1 + \frac{\delta}{x}\right) - \delta \le (x + \delta)\frac{\delta}{x} - \delta = \frac{\delta^2}{x}.$$

Then

$$\mathbf{Ent}\,M_1 \le \frac{1}{2}\mathbf{E}\int_0^1 \frac{1}{M_{t-}}d\langle M^c\rangle_t + \mathbf{E}\sum_{0<t\le 1}\frac{(M_t - M_{t-})^2}{M_{t-}} \le \mathbf{E}\int_0^1 \frac{1}{M_{t-}}d[M]_t.$$

Because the process M_{t-}, $t \in [0, 1]$, is $\mathbb{F}$-predictable, we have

$$\mathbf{E} \int_0^1 \frac{1}{M_{t-}} d[M]_t = \mathbf{E} \int_0^1 \frac{1}{M_{t-}} d\langle M \rangle_t,$$

which completes the proof of the required bound under assumption (3).

The upper bound in this assumption can be removed using the following standard localization procedure. For $N \geq 1$, define

$$\tau_N = \inf\{t \in [0, 1] : M_t \geq N\}$$

with the convention $\inf \varnothing = 1$. Then, repeating the above argument, we get

$$\mathbf{Ent}\, M_{\tau_N} \leq \mathbf{E} \int_0^{\tau_N} \frac{1}{M_{t-}} d\langle M \rangle_t \leq \mathbf{E} \int_0^1 \frac{1}{M_{t-}} d\langle M \rangle_t.$$

We have $M_{\tau_N} \to M_1$, $N \to \infty$ a.s. On the other hand, $\mathbf{E} M_{\tau_N}^2 \leq \mathbf{E} M_1^2$, and

$$x \log x = o(x^2), \quad x \to +\infty.$$

Hence, the family $\{M_{\tau_N} \log M_{\tau_N}, N \geq 1\}$ is uniformly integrable, and

$$\mathbf{Ent}\, M_{\tau_N} \to \mathbf{Ent}\, M_1, \quad N \to \infty.$$

Passing to the limit as $N \to \infty$, we obtain the required statement under the assumption $M_t \geq c_1 > 0$. Taking $M_t + (1/n)$ instead of M_t and then passing to the limit as $n \to \infty$, we complete the proof of the theorem. $\qquad\Box$

We further give two examples where the shown martingale bound for the entropy is applied. In these examples, it would be more convenient to assume that t varies in $[0, \infty)$ instead of $[0, 1]$; a respective version of Theorem 1 can be proved by literally the same argument.

Example 1 (Log-Sobolev inequality on a Brownian path space; [1, 6]). Let B_t, $t \geq 0$, be a Wiener process on $(\Omega, \mathcal{F}, \mathbf{P})$ such that $\mathcal{F} = \sigma(B)$. Let $\{\mathcal{F}_t\}$ be the natural filtration for B. Then for every $\zeta \in L_2(\Omega, \mathbf{P})$, the following martingale representation is available:

$$\zeta = \mathbf{E}\zeta + \int_0^\infty \eta_s \, dB_s \qquad (4)$$

with the Itô integral of a (unique) square-integrable $\{\mathcal{F}_t\}$-adapted process $\{\eta_t\}$ in the right-hand side (cf. [3]). Take $\xi \in L_4(\Omega, \mathbf{P})$ and put $\zeta = \xi^2$ and

$$M_t = \mathbf{E}[\zeta | \mathcal{F}_t] = \mathbf{E}\zeta + \int_0^t \eta_s \, dB_s, \quad t \geq 0.$$

Then the calculation from the proof of Theorem 1 gives the bound

$$\mathbf{Ent}\, \xi^2 \leq \frac{1}{2}\mathbf{E} \int_0^1 \frac{1}{M_{t-}} d\langle M^c \rangle_t = \frac{1}{2}\mathbf{E} \int_0^1 \frac{\eta_t^2}{M_t} dt = \frac{1}{2}\mathbf{E} \int_0^1 \frac{\eta_t^2}{\mathbf{E}[\xi^2 | \mathcal{F}_t]} dt.$$

Note the extra term $1/2$, which appears because the martingale M is continuous.

Next, recall the Ocone representation [10] for the process $\{\eta_t\}$, which is valid if ζ possesses the Malliavin derivative $D\zeta = \{D_t\zeta, t \geq 0\}$:

$$\eta_t = \mathbf{E}[D_t\zeta|\mathcal{F}_t], \quad t \geq 0. \tag{5}$$

We omit the details concerning the Malliavin calculus, referring the reader, if necessary, to [9]. Because the Malliavin derivative possesses the chain rule, we have

$$\eta_t^2 = 4\big(\mathbf{E}[\xi\, D_t\xi|\mathcal{F}_t]\big)^2 \leq 4\mathbf{E}\big[\xi^2\big|\mathcal{F}_t\big]\mathbf{E}\big[(D_t\xi)^2\big|\mathcal{F}_t\big],$$

and consequently the following log-Sobolev-type inequality holds:

$$\mathbf{Ent}\,\xi^2 \leq 2\mathbf{E}\int_0^1 \mathbf{E}\big[(D_t\xi)^2\big|\mathcal{F}_t\big]\,dt = 2\mathbf{E}\|D\xi\|_H^2, \tag{6}$$

where $D\xi$ is considered as a random element in $H = L_2(0, \infty)$. By a proper approximation procedure one can show that (7) holds for every $\xi \in L_2(\Omega, \mathbf{P})$ that has a Malliavin derivative $D\xi \in L_2(\Omega, \mathbf{P}, H)$.

The previous example is classic and well known. The next one apparently is new, which is a bit surprising because the main ingredients therein (the Malliavin calculus on the Poisson space and the respective analogue of the Clark–Ocone representation (4), (5)) are well known (cf. [2, 5]).

Example 2 (Log-Sobolev inequality on the Poisson path space). Let N_t, $t \geq 0$, be a Poisson process with intensity λ, and $\mathcal{F} = \sigma(N)$. Denote by τ_k, $k \geq 1$, the moments of consequent jumps of the process N, and by $\mathcal{F}_t = \sigma(N_s, s \leq t)$, $t \geq 0$, the natural filtration for N. For any variable of the form

$$\xi = F(\tau_1, \ldots, \tau_n)$$

with some $n \geq 1$ and some compactly supported $F \in C^1(\mathbb{R}^n)$, define the random element $D\xi$ in $H = L_2(0, \infty)$ by

$$D\xi = -\sum_{k=1}^n F_k'(\tau_1, \ldots, \tau_n)\mathbf{1}_{[0, \tau_k]}.$$

Denote by the same symbol D the closure of D, considered as an unbounded operator $L_2(\Omega, \mathbf{P}) \to L_2(\Omega, \mathbf{P}, H)$. Then the following analogue of the Clark–Ocone representation (4), (5) is available ([5]): for every ζ that possesses the stochastic derivative $D\zeta$, the following martingale representation holds:

$$\zeta = \mathbf{E}\zeta + \frac{1}{\lambda}\int_0^\infty \eta_s\, d\tilde{N}_s,$$

where $\tilde{N}_t = N_t - \lambda t$ denotes the compensated Poisson process corresponding to N, and $\{\eta_t\}$ is the projection in $L_2(\Omega, \mathbf{P}, H)$ of $D\xi$ on the subspace generated by the $\{\mathcal{F}_t\}$-predictable processes.

Proceeding in the same way as we did in the previous example, we obtain the following log-Sobolev-type inequality on the Poisson path space:

$$\mathbf{Ent}\,\xi^2 \leq \frac{4}{\lambda^2}\mathbf{E}\|D\xi\|_H^2. \tag{7}$$

2 Trimmed regions on $\mathbb{R}^d$ and associated integral bounds for the entropy

Let μ be a probability measure on $\mathbb{R}^d$ with Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$. Our further aim is to apply the general martingale bound from Theorem 1 in the particular setting $(\Omega, \mathcal{F}, \mathbf{P}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$. To this end, we first *construct* a filtration $\{\mathcal{F}_t, t \in [0, 1]\}$.

In what follows, we denote $\mathcal{N}_\mu = \{A \in \mathcal{F} : \mu(A) = 0\}$ (the class of μ-null Borel sets).

Fix the family $\{D_t, t \in [0, 1]\}$ of closed subsets of $\mathbb{R}^d$ such that:

(i) $D_s \subset D_t$, $s \le t$;

(ii) $D_0 \in \mathcal{N}_\mu$, $\mu(D_t) < 1$ for $t < 1$, and $D_1 = \mathbb{R}^d$;

(iii) for every $t > 0$,

$$D_t \setminus \left(\bigcup_{s<t} D_s \right) \in \mathcal{N}_\mu,$$

and for every $t < 1$,

$$D_t = \bigcap_{s>t} D_s.$$

We call the sets D_t, $t \in [0, 1]$, *trimmed regions*, following the terminology used frequently in the multivariate analysis (cf. [7]). Given the family $\{D_t\}$, we define the respective *trimmed filtration* $\{\mathcal{F}_t\}$ by the following convention. Denote $Q_t = \mathbb{R}^d \setminus D_t$. Then, by definition, a set $A \in \mathcal{F}$ belongs to $\mathcal{F}_t$ if either $A \cap Q_t \in \mathcal{N}_\mu$ or $Q_t \setminus A \in \mathcal{N}_\mu$.

By the construction, $\mathbb{F} = \{\mathcal{F}_t\}$ is complete. It is also clear that, by property (ii) of the family $\{D_t\}$, the σ-algebra $\mathcal{F}_0$ is degenerate and, by property (iii), the filtration $\mathbb{F}$ is continuous. Hence, we can apply Theorem 1.

Fix a Borel-measurable function $g : \mathbb{R}^d \to \mathbb{R}^+$ that is square-integrable w.r.t. μ. Consider it as a random variable on $(\Omega, \mathcal{F}, \mathbf{P}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$ and define

$$g_t = \mathbf{E}[g | \mathcal{F}_t], \quad t \in [0, 1].$$

Since the σ-algebra possesses an explicit description, we can calculate every g_t directly; namely, for $t > 0$ and μ-a.a. x, we have

$$g_t(x) = \begin{cases} g(x), & x \in D_t, \\ G_t, & x \in Q_t, \end{cases} \tag{8}$$

where we denote

$$G_t = \frac{1}{\mu(Q_t)} \int_{Q_t} g(y) \, \mu(dy). \tag{9}$$

Note that $\mu(Q_t) > 0$ for $t < 1$ and the function $G : [0, 1) \to \mathbb{R}^+$ is continuous. In what follows, we consider the modification of the process $\{g_t\}$ defined by (8) for *every* $x \in \mathbb{R}^d$. Its trajectories can be described as follows. Denote

$$\tau(x) = \inf\{t : x \in D_t\}; \tag{10}$$

then by property (iii) of the family $\{D_t\}$ we have $\tau(x) = \min\{t : x \in D_t\}$, and by property (ii) we have $\tau(x) < 1$, $x \in \mathbb{R}^d$, $\tau(x) = 0 \Leftrightarrow x \in D_0$. Then, for a fixed $x \in \mathbb{R}^d$, we have

$$g_t(x) = g(x)1_{t \geq \tau(x)} + G_t 1_{t < \tau(x)}, \quad t \in [0, 1],$$

which is a càdlàg function because $\{G_t\}$ is continuous on $[0, 1)$.

Theorem 2. *Let g be a Borel-measurable function $g : \mathbb{R}^d \to \mathbb{R}^+$, square-integrable w.r.t. μ. Let $\{D_t\}$ be a family of trimmed regions that satisfy* (i)–(iii).
Then

$$\mathbf{Ent}_\mu \, g \leq \int_{\mathbb{R}^d} \frac{(g(x) - G_{\tau(x)})^2}{G_{\tau(x)}} \mu(dx),$$

where the functions G and τ are defined by (9) *and* (10), *respectively.*

Proof. We have already verified the assumptions of Theorem 1: the filtration $\{\mathcal{F}_t\}$ is complete and right continuous, and the square-integrable martingale $\{g_t\}$ has càdlàg trajectories. Because $g_1 = g$ a.s. and $\mathcal{F}_0$ is degenerate, by Theorem 1 we have the bound

$$\mathbf{Ent}_\mu \, g \leq \mathbf{E} \int_0^1 \frac{1}{g_{t-}} d\langle g \rangle_t.$$

Hence, we only have to specify the integral in the right-hand side of this bound. Namely, our aim is to prove that

$$\mathbf{E} \int_0^1 \frac{1}{g_{t-}} d\langle g \rangle_t = \int_{\mathbb{R}^d} \frac{(g(x) - G_{\tau(x)})^2}{G_{\tau(x)}} \mu(dx). \tag{11}$$

First, we observe the following.

Lemma 1. *Let $0 < s < t < 1$, and let α be a bounded $\mathcal{F}_s$-measurable random variable. Then*

$$\mathbf{E}\big[\alpha(\langle g \rangle_t - \langle g \rangle_s)\big] = \int_{D_t \setminus D_s} \alpha(x)\big(g(x) - G_{\tau(x)}\big)^2 \mu(dx).$$

Proof. By the definition of $\langle g \rangle$,

$$\mathbf{E}\big[\alpha(\langle g \rangle_t - \langle g \rangle_s)\big] = \mathbf{E}\big[\alpha(g_t^2 - g_s^2)\big] = \mathbf{E}\big[\alpha\big(\mathbf{E}(g_t^2|\mathcal{F}_s) - g_s^2\big)\big].$$

We have

$$g_s^2(x) = \begin{cases} g^2(x), & x \in D_s, \\ G_s^2, & x \in Q_s, \end{cases} \qquad g_t^2(x) = \begin{cases} g^2(x), & x \in D_t, \\ G_t^2, & x \in Q_t, \end{cases}$$

and applying formula (8) with $g = g_t^2$ and $t = s$, we get

$$\mathbf{E}\big(g_t^2|\mathcal{F}_s\big)(x) - g_s^2(x) = \begin{cases} 0, & x \in D_s, \\ \frac{H_{t,s}}{\mu(Q_s)}, & x \in Q_s, \end{cases}$$

$$H_{t,s} = \left(\int_{D_t \setminus D_s} \left(g^2(x) - G_s^2 \right) \mu(dx) + \int_{Q_t} \left(G_t^2 - G_s^2 \right) \mu(dx) \right).$$

Because α is $\mathcal{F}_s$-measurable, it equals a constant on Q_s μ-a.s. Denote this constant by A; then the previous calculation gives

$$\mathbf{E}\big[\alpha\big(\langle g \rangle_t - \langle g \rangle_s\big)\big] = A H_{t,s}.$$

Write $H_{t,s}$ in the form

$$H_{t,s} = \int_{D_t \setminus D_s} g^2(x) \, \mu(dx) + \mu(Q_t) G_t^2 - \mu(Q_s) G_s^2.$$

Denote

$$\mu_t = \mu(Q_t), \qquad I_t = \int_{Q_t} g \, d\mu;$$

then

$$\mu(Q_t) G_t^2 = \mu_t G_t^2 = \frac{I_t^2}{\mu_t}.$$

Observe that the functions $\mu_t, t \in [0, 1]$ and $I_t, t \in [0, 1]$, are continuous functions of a bounded variation and $\mu_t > 0$, $t < 1$. Then

$$\mu(Q_t) G_t^2 - \mu(Q_s) G_s^2 = \int_s^t d\left(\frac{I_v^2}{\mu_v} \right) = \int_s^t \left(-\frac{I_v^2}{\mu_v^2} d\mu_v + 2\frac{I_v}{\mu_v} dI_v \right)$$

$$= \int_s^t \left(-G_v^2 \, d\mu_v + 2 G_v \, dI_v \right).$$

It is easy to show that

$$-\int_s^t G_v^2 \, d\mu_v = \int_{D_t \setminus D_s} G_{\tau(x)}^2 \, \mu(dx). \qquad (12)$$

Indeed, because G is continuous on $[0, 1)$, the left-hand side integral can be approximated by the integral sum

$$\sum_{k=1}^m G_{v_k}^2 \left(\mu_{v_{k-1}} - \mu_{v_k} \right),$$

where $s = v_0 < \cdots < v_m = t$ is some partition of $[s, t]$. This sum equals

$$\sum_{k=1}^m G_{v_k}^2 \mu(D_{v_k} \setminus D_{v_{k-1}}).$$

For $x \in D_{v_k} \setminus D_{v_{k-1}}$, we have $\tau(x) \in [v_{k-1}, v_k]$. Hence, this sum equals

$$\sum_{k=1}^m \int_{D_{v_k} \setminus D_{v_{k-1}}} G_{\tau(x)}^2 \, \mu(dx) = \int_{D_t \setminus D_s} G_{\tau(x)}^2 \, \mu(dx)$$

up to a residue term that is dominated by

$$\sup_{u,v\in[s,t],|u-v|\leq\max_k(v_k-v_{k-1})}\left|G_u^2 - G_v^2\right|$$

and tends to zero as the size of the partition tends to zero. This proves (12). Similarly, we can show that

$$\int_s^t G_v\,dI_v = -\int_{D_t\setminus D_s} G_{\tau(x)}g(x)\,\mu(dx).$$

We can summarize this calculation as follows:

$$\mathbf{E}\big[\alpha\big(\langle g\rangle_t - \langle g\rangle_s\big)\big] = A\int_{D_t\setminus D_s}\big(g(x) - G_{\tau(x)}\big)^2\mu(dx).$$

Because $\alpha(x) = A$ for μ-a.a. $x \notin D_s$, this completes the proof. □

Let us continue with the proof of (11). Assume first that $g \geq c$ with some $c > 0$. Then $g_t \geq c$, and consequently the process $1/g_{t-}$ is left continuous and bounded. In addition, the function $G_t = I_t/\mu_t$ is bounded on every segment $[0, T] \subset [0, 1)$.

Fix $T < 1$ and take a sequence $\{\lambda^n\}$ of dyadic partitions of $[0, T]$,

$$\lambda^n = \big\{t_k^n, k = 0, \ldots, 2^n\big\}, \quad t_k^n = \frac{Tk}{2^n},$$

and define

$$g_t^n = g_0 1_{t=0} + \sum_{k=1}^{2^n} g_{t_{k-1}^n} 1_{t\in(t_{k-1}^n, t_k^n]}.$$

For every fixed $t > 0$, the value g_t^n equals the value of g at some (dyadic) point $t_n < t$, and $t_n \to t-$. Hence,

$$\frac{1}{g_t^n} \to \frac{1}{g_{t-}}, \quad n \to \infty,$$

pointwise. In addition, because of the additional assumption $g \geq c$, this sequence is bounded by $1/c$. Hence, by the dominated convergence theorem,

$$\mathbf{E}\int_0^T \frac{1}{g_{t-}}d\langle g\rangle_t = \lim_{n\to\infty} \mathbf{E}\sum_{k=1}^{2^n} \frac{1}{g_{t_{k-1}^n}}\big(\langle g\rangle_{t_k^n} - \langle g\rangle_{t_{k-1}^n}\big);$$

here we take into account that the point $t = 0$ in the left-hand side integral is negligible because $g_t \to \mathbf{E}g$, $t \to 0+$, in L_2, and consequently $\langle g\rangle_t \to 0, t \to 0+$, in L_1. By Lemma 1,

$$\mathbf{E}\sum_{k=1}^{2^n} \frac{1}{g_{t_{k-1}^n}}\big(\langle g\rangle_{t_k^n} - \langle g\rangle_{t_{k-1}^n}\big) = \mathbf{E}\sum_{k=1}^{2^n} \int_{D_{t_k^n}\setminus D_{t_{k-1}^n}} \frac{(g(x) - G_{\tau(x)})^2}{G_{t_{k-1}^n}}\mu(dx);$$

recall that $g_{t_{k-1}^n}(x) = G_{t_{k-1}^n}$ for $x \notin D_{t_{k-1}^n}$. Next, for $x \in D_{t_k^n}\setminus D_{t_{k-1}^n}$, we have $|\tau(x) - t_{k-1}^n| \leq 2^{-n}$. Because $G_t, t \in [0, T]$, is uniformly continuous and separated

from zero, and $G_{\tau(x)}$, $x \in D_T$ is bounded, we obtain that

$$\mathbf{E} \int_0^T \frac{1}{g_{t-}} d\langle g\rangle_t = \lim_{n\to\infty} \mathbf{E} \sum_{k=1}^{2^n} \int_{D_{t_k^n}\backslash D_{t_{k-1}^n}} \frac{(g(x) - G_{\tau(x)})^2}{G_{t_{k-1}^n}} \mu(dx)$$

$$= \int_{D_T} \frac{(g(x) - G_{\tau(x)})^2}{G_\tau(x)} \mu(dx).$$

Taking $T \to 1-$ and applying the monotone convergence theorem to both sides of the above identity, we get (11).

To remove the additional assumption $g \geq c$, consider the family $g_t^n = g_t + 1/n$. Then $\langle g^n \rangle = \langle g \rangle$, $g^n(x) - G_{\tau(x)}^n = g(x) - G_{\tau(x)}$, $g_{t-}^n = g_{t-} + (1/n)$, $G_{\tau(x)}^n = G_{\tau(x)} + (1/n)$. Hence, we can write (11) for g^n, apply the monotone convergence theorem to both sides of this identity, and get (11) for g. $\qquad\square$

3 One corollary: a weighted log-Sobolev inequality on $\mathbb{R}$

In this section, we show the way how the integral bound for the entropy, established in Theorem 2, can be used to obtain weighted log-Sobolev inequalities. Consider a continuous probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and denote by p_μ the density of its absolutely continuous part. Fix a family of segments $D_t = [a_t, b_t]$, $t \in [0, 1)$, where $a_0 = b_0$, the function a_t is continuous and decreasing to $-\infty$ as $t \to 1-$, and the function $b.$ is continuous and increasing to $+\infty$ as $t \to 1-$. Then the family

$$D_t = [a_t, b_t], \quad t \in [0, 1), \qquad D_1 = \mathbb{R},$$

satisfies the assumptions imposed before. Hence, Theorem 2 is applicable.

We call a function $f : \mathbb{R} \to \mathbb{R}$ *symmetric* w.r.t. the family $\{D_t\}$ if

$$f(a_t) = f(b_t), \quad t \in [0, 1).$$

In the following proposition, we apply Theorem 2 to $g = f^2$, where f is smooth and symmetric.

Proposition 1. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a smooth function that is symmetric w.r.t. the family* $\{D_t\}$. *Then*

$$\mathbf{Ent}_\mu f^2 \leq 4 \int_{\mathbb{R}} W(x) \big(f'(x)\big)^2 \mu(dx),$$

where

$$W(x) = V^2(x) \log\left(\frac{1}{\mu_{\tau(x)}}\right), \qquad V(x) = \begin{cases} \frac{\mu((-\infty,x))}{p_\mu(x)}, & x \leq a_0, \\ \frac{\mu((x,\infty))}{p_\mu(x)}, & x > a_0. \end{cases}$$

Proof. Write

$$g(x) - G_{\tau(x)} = \frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} \big(g(x) - g(y)\big) \mu(dy) = \frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} \int_x^y g'(z)\, dz\, \mu(dy).$$

Let us analyze the expression in the right-hand side. Observe that now $Q_{\tau(x)}$ is the union of two intervals $(-\infty, a_{\tau(x)})$ and $(b_{\tau(x)}, +\infty)$. Denote

$$Q_t^+ = (b_t, \infty), \qquad Q_t^- = (-\infty, a_t), \qquad \mu_t^\pm = \mu(Q_t^\pm).$$

The point x equals either $a_{\tau(x)}$ or $b_{\tau(x)}$; hence, because $g = f^2$ is symmetric,

$$g(x) = g(a_{\tau(x)}) = g(b_{\tau(x)}).$$

Then we have

$$\int_x^y g'(z)\,dz = \begin{cases} \int_{b_{\tau(x)}}^y g'(z)\,dz, & y \in Q_{\tau(x)}^+, \\ \int_{a_{\tau(x)}}^y g'(z)\,dz, & y \in Q_{\tau(x)}^-. \end{cases}$$

Consequently,

$$\left| g(x) - G_{\tau(x)} \right| \le \frac{1}{\mu_{\tau(x)}} \left[\int_{Q_{\tau(x)}^+} \int_{Q_{\tau(x)}^+,\tau(z)\le\tau(y)} \left| g'(z) \right| dz\,\mu(dy) \right. $$
$$\left. + \int_{Q_{\tau(x)}^-} \int_{Q_{\tau(x)}^-,\tau(z)\le\tau(y)} \left| g'(z) \right| dz\,\mu(dy) \right].$$

Using Fubini's theorem, we get

$$\left| g(x) - G_{\tau(x)} \right| \le \frac{1}{\mu_{\tau(x)}} \left[\int_{Q_{\tau(x)}^+} \mu_{\tau(z)}^+ \left| g'(z) \right| dz + \int_{Q_{\tau(x)}^-} \mu_{\tau(z)}^- \left| g'(z) \right| dz \right]$$
$$\le \frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} V(z) \left| g'(z) \right| \mu(dz).$$

Because $g = f^2$ and hence $g' = 2ff'$, by the Cauchy inequality we then have

$$\left(g(x) - G_{\tau(x)} \right)^2$$
$$\le 4 \left(\frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} (V(z)f'(z))^2\,\mu(dz) \right) \left(\frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} (f(z))^2\,\mu(dz) \right)$$
$$= 4 \left(\frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} (V(z)f'(z))^2\,\mu(dz) \right) G_{\tau(x)}.$$

Observe that

$$z \in Q_{\tau(x)} \quad \Leftrightarrow \quad \tau(z) > \tau(x) \quad \Leftrightarrow \quad x \in D_{\tau(z)} \setminus \{a_{\tau(z)}, b_{\tau(z)}\}.$$

Hence, by Theorem 2 and Fubini's theorem we have

$$\mathbf{Ent}_\mu(f^2) \le 4 \int_{\mathbb{R}} \left(\frac{1}{\mu_{\tau(x)}} \int_{Q_{\tau(x)}} (V(z)f'(z))^2\mu(dz) \right) \mu(dx)$$
$$= 4 \int_{\mathbb{R}} \left(\int_{D_{\tau(z)}} \frac{\mu(dx)}{\mu_{\tau(x)}} \right) (V(z)f'(z))^2\,\mu(dz).$$

Similarly to the proof of (12), we can show that

$$\int_{D_t} \frac{\mu(dx)}{\mu_{\tau(x)}} = \log \mu_s \big|_{s=0}^{s=t} = \log\left(\frac{1}{\mu_t}\right);$$

the last identity holds because $\mu_0 = 1$. This completes the proof.   □

Next, we develop a *symmetrization* procedure in order to remove the restriction for f to be symmetric. For any $x \neq a_0$, one border point of the segment $D_\tau(x)$ equals x; let us denote $s(x)$ the other border point. Denote also $s(a_0) = a_0$. Define the σ-algebra $\hat{\mathcal{F}}$ of symmetric sets $A \in \mathcal{F}$, that is, such that $x \in A \Leftrightarrow s(x) \in A$. For a function $f \in L_2(\mathbb{R}, \mu)$, consider its L_2-*symmetrization*

$$\hat{f} = \big(\mathbf{E}_\mu[f^2|\hat{\mathcal{F}}]\big)^{1/2}.$$

It can be seen easily that there exists a measurable function $p : \mathbb{R} \to [0, 1]$ such that, for μ-a.a. $x \in \mathbb{R}$,

$$(\hat{f})^2(x) = p(x)f^2(x) + \big(1 - p(x)\big)f^2\big(s(x)\big) = \mathbf{E}_{\nu_x} f^2,$$

where we denote

$$\nu_x = p(x)\delta_x + \big(1 - p(x)\big)\delta_{s(x)}, \quad x \in \mathbb{R}.$$

We have

$$\mathbf{E}_\mu f^2 = \mathbf{E}_\mu (\hat{f})^2$$

and, consequently,

$$\begin{aligned}
\mathbf{Ent}_\mu f^2 - \mathbf{Ent}_\mu (\hat{f})^2 &= \mathbf{E}_\mu f^2 \log f^2 - \mathbf{E}_\mu (\hat{f})^2 \log(\hat{f})^2 \\
&= \mathbf{E}_\mu \big(\mathbf{E}_\mu[f^2 \log f^2 - (\hat{f})^2 \log(\hat{f})^2|\hat{\mathcal{F}}]\big) \\
&= \int_{\mathbb{R}} \big(\mathbf{Ent}_{\nu_x} f^2\big) \mu(dx).
\end{aligned}$$

It is well known (cf. [8]) that for a Bernoulli measure $\nu = p\delta_1 + q\delta_{-1}$ $(p+q = 1)$, the following discrete analogue of the log-Sobolev inequality holds:

$$\mathbf{Ent}_\nu f^2 \leq C_p(Df)^2, \qquad C_p = \begin{cases} pq\frac{\log p - \log q}{p-q}, & p \neq q, \\ \frac{1}{2}, & p = q, \end{cases}$$

where we denote $Df = f(1) - f(-1)$. This yields the bound

$$\begin{aligned}
\mathbf{Ent}_\mu f^2 - \mathbf{Ent}_\mu (\hat{f})^2 &\leq \int_{\mathbb{R}} C_{p(x)}\big(f(x) - f\big(s(x)\big)\big)^2 \mu(dx) \\
&= \int_{\mathbb{R}} C_{p(x)} \bigg(\int_{D_{\tau(x)}} f'(z)\, dz\bigg)^2 \mu(dx).
\end{aligned}$$

By the Cauchy inequality,

$$\bigg(\int_{D_{\tau(x)}} f'(z)\, dz\bigg)^2 \leq \bigg(\int_{D_{\tau(x)}} \big(f'(z)\big)^2 \frac{\mu_{\tau(z)}^{3/2}}{p_\mu^2(z)} \mu(dz)\bigg)\bigg(\int_{D_{\tau(x)}} \frac{\mu(dz)}{\mu_{\tau(z)}^{3/2}}\bigg),$$

and, similarly to the proof of (12), we can show that

$$\int_{D_{\tau(x)}} \frac{\mu(dz)}{\mu_{\tau(z)}^{3/2}} = 2\big(\mu_{\tau(x)}^{-1/2} - 1\big) < 2\mu_{\tau(x)}^{-1/2}.$$

This yields the following bound for the difference $\mathbf{Ent}_\mu\, f^2 - \mathbf{Ent}_\mu(\hat{f})^2$, formulated in the terms of f':

$$\mathbf{Ent}_\mu\, f^2 - \mathbf{Ent}_\mu(\hat{f})^2 \le 2\int_{\mathbb{R}} \big(f'(z)\big)^2 U(z)\,\mu(dz),$$

$$U(z) = \frac{\mu_{\tau(z)}^{3/2}}{p_\mu^2(z)} \int_{Q_{\tau(z)}} C_{p(x)} \frac{\mu(dx)}{\mu_{\tau(x)}^{1/2}}.$$

Note that $C_p \le 1$ for any $p \in [0, 1]$, and hence we have

$$U(z) \le 2\left(\frac{\mu_{\tau(z)}}{p_\mu(z)}\right)^2.$$

Assuming that the bound from Proposition 1 is applicable to $\hat{f}$ (which is yet to be studied because $\hat{f}$ may fail to be smooth), we obtain the following inequality, valid without the assumption of symmetry of f:

$$\mathbf{Ent}_\mu\, f^2 \le \int_{\mathbb{R}} \big(4W(x)\big((\hat{f})'(x)\big)^2 + 2U(x)\big(f'(x)\big)^2\big)\,\mu(dx). \tag{13}$$

The right-hand side of this inequality contains the derivative of $\hat{f}$ and hence depends on the choice of the family of trimmed regions $\{D_t\}$. We further give a particular corollary, which appears when $\{D_t\}$ is the set of *quantile trimmed regions*. In what follows, we assume μ to possess a positive distribution density p_μ and choose $\{D_t = [a_t, b_t]\}$ in the following way. Denote $q_v = F_\mu^{-1}(v)$, that is, the quantile of μ of the level v, and put

$$a_t = q_{1/2-t/2}, \quad b_t = q_{1/2+t/2}, \quad t \in [0, 1).$$

In particular, $a_0 = b_0 = m$, the median of μ. Denote also $\hat{F}_\mu = \min(F_\mu, 1 - F_\mu)$; observe that now we have

$$\hat{F}_\mu(x) = \frac{1}{2}\mu_{\tau(x)}.$$

Theorem 3. *Let μ be a probability measure on $\mathbb{R}$ with positive distribution density p_μ. Then, for any absolutely continuous f, we have*

$$\mathbf{Ent}_\mu\, f^2 \le \int_{\mathbb{R}} K(x)\big(f'(x)\big)^2 \mu(dx), \qquad K(x) = 8\left(\frac{\hat{F}_\mu(x)}{p_\mu(x)}\right)^2\left(\log\frac{1}{2\hat{F}_\mu(x)} + 1\right).$$

Proof. First, observe that now the L_2-symmetrization of a function f has the form

$$\hat{f}(x) = \sqrt{\frac{1}{2}\big(f^2(x) + f^2(s(x))\big)}. \tag{14}$$

This identity is evident for functions f of the form $1_{(-\infty, F^{-1}(v))}$, $v \in (0, 1/2]$ and $1_{[F^{-1}(v), \infty)}$, $v \in [1/2, 1)$, and then easily extends to general f.

Next, observe that

$$s(x) = F_\mu^{-1}\big(1 - F_\mu(x)\big), \tag{15}$$

and because F_μ is absolutely continuous and strictly increasing, $s(x)$ is absolutely continuous as well. Then $\hat{f}$ is absolutely continuous with

$$(\hat{f})'(x) = \frac{f(x)f'(x) + f(s(x))f'(s(x))s'(x)}{\sqrt{2(f^2(x) + f^2(s(x)))}};$$

here and below the derivatives are well defined for a.a. x. Using a standard localization/approximation procedure, we can show that Proposition 1 is well applicable to any absolutely continuous function. Hence, it is applicable to $\hat{f}$, and (13) holds.

We have

$$\big((\hat{f})'(x)\big)^2 \leq \frac{(f'(x))^2 + (f(s(x))f'(s(x))s'(x))^2}{f^2(x) + f^2(s(x))}$$

$$\leq \big(f'(x)f(x)\big)^2 + \big(f'(s(x))s'(x)\big)^2.$$

The function $W(x)$ in (13) now can be rewritten as

$$W(x) = \left(\frac{\hat{F}_\mu(x)}{p_\mu(x)}\right)^2 \log \frac{1}{2\hat{F}_\mu(x)};$$

hence,

$$\int_{\mathbb{R}} W(x)\big((\hat{f}'(x))\big)^2 \mu(dx)$$

$$\leq \int_{\mathbb{R}} W(x)\big(f'(x)\big)^2 \mu(dx) + \int_{\mathbb{R}} W(x)\big(f'(s(x))\big)^2\big(s'(x)\big)^2 \mu(dx).$$

Let us analyze the second integral in the right-hand side. By (15),

$$s'(x) = -\frac{p_\mu(x)}{p_\mu(s(x))}; \tag{16}$$

hence,

$$\int_{\mathbb{R}} W(x)\big(f'(s(x))\big)^2\big(s'(x)\big)^2 \mu(dx)$$

$$= \int_{\mathbb{R}} \big(f'(s(x))\big)^2 \left(\frac{\hat{F}_\mu(x)}{p_\mu(s(x))}\right)^2 \log \frac{1}{2\hat{F}_\mu(x)} p_\mu(x)\, dx.$$

Change the variables $y = s(x)$; observe that we have $x = s(y)$ and $\hat{F}_\mu(x) = \hat{F}_\mu(y)$. Then we finally get

$$\int_{\mathbb{R}} W(x)\big(f'(y)\big)^2 \big(s'(x)\big)^2 \, \mu(dx)$$

$$= \int_{\mathbb{R}} \big(f'(s(x))\big)^2 \left(\frac{\hat{F}_\mu(y)}{p_\mu(y)}\right)^2 \log \frac{1}{2\hat{F}_\mu(y)} \, p_\mu\big(s(y)\big) \, \frac{p_\mu(y)}{p_\mu(s(y))} dy$$

$$= \int_{\mathbb{R}} W(y)\big(f'(y)\big)^2 \, \mu(dy),$$

and therefore

$$\int_{\mathbb{R}} W(x)\big((\hat{f}')(x)\big)^2 \, \mu(dx) \le 2 \int_{\mathbb{R}} W(x)\big(f'(x)\big)^2 \, \mu(dx).$$

On the other hand, by identity (14) we have now $C_{p(x)} = 1/2$, and the function $U(x)$ in (13) can be rewritten as

$$U(x) = \left(\frac{\mu_{\tau(x)}}{p_\mu(x)}\right)^2 = 4\left(\frac{\hat{F}_\mu(x)}{p_\mu(x)}\right)^2,$$

which completes the proof of the statement. $\qquad\square$

References

[1] Capitaine, M., Hsu, E., Ledoux, M.: Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces. Electron. Commun. Probab. **2**, 71–81 (1997) MR1484557. doi:10.1214/ECP.v2-986

[2] Carlen, E., Pardoux, E.: Differential calculus and integration by parts on Poisson space. In: Stochastics, Algebra and Analysis in Classical and Quantum Dynamics, pp. 63–73. Kluwer, Dordrecht (1990) MR1052702

[3] Clark, I.M.C.: The representation of functionals of Brownian motion by stochastic integrals. Ann. Math. Stat. **41**(4), 46–52 (1970) MR0270448

[4] Elliott, R.J.: Stochastic Calculus and Applications. Appl. Math., vol. 18, Springer-Verlag, Berlin, Heidelberg, New York (1982) MR0678919

[5] Elliott, R.J., Tsoi, A.H.: Integration by parts for Poisson processes. J. Multivar. Anal. **44**, 179–190 (1993) MR1219202. doi:10.1006/jmva.1993.1010

[6] Gong, F.-Zh., Ma, Zh.-M.: Martingale representation and log-Sobolev inequality on loop space. C. R. Acad. Sci. Paris **326**, 749–753 (1998) MR1641711. doi:10.1016/S0764-4442(98)80043-8

[7] Koshevoy, G.A., Mosler, K.: Zonoid trimming for multivariate distributions. Ann. Stat. **25**(5), 1998–2017 (1997) MR1474078. doi:10.1214/aos/1069362382

[8] Ledoux, M.: Concentration of measure and logarithmic Sobolev inequalities. In: Seminaire de Probabilités XXXIII. Lect. Notes Math., vol. 1709, pp. 120–219 (1999) MR1767995. doi:10.1007/BFb0096511

[9] Nualart, D.: Analysis in Wiener space and anticipating stochastic calculus, Springer, Berlin, Heidelberg, Lect. Notes Math., vol. 1690, pp. 123–227 (1998) MR1668111. doi:10.1007/BFb0092538

[10] Ocone, D.: Malliavin's calcululs and stochastic integral representation of functionals of diffusion processes. Stochastics **8**, 161–185 (1984) MR0749372. doi:10.1080/17442508408833299