



**HAL**  
open science

# Regularization in $L_1$ for the Ornstein-Uhlenbeck semigroup

Joseph Lehec

► **To cite this version:**

Joseph Lehec. Regularization in  $L_1$  for the Ornstein-Uhlenbeck semigroup. Annales de la Faculté des Sciences de Toulouse. Mathématiques., 2016, 25 (1), pp.191-204. 10.5802/afst.1492 . hal-01428315

**HAL Id: hal-01428315**

**<https://hal.science/hal-01428315>**

Submitted on 6 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Regularization in $L_1$ for the Ornstein–Uhlenbeck semigroup

Joseph Lehec \*

July 26, 2015

## Abstract

Let  $\gamma_n$  be the standard Gaussian measure on  $\mathbb{R}^n$  and let  $(Q_t)$  be the Ornstein–Uhlenbeck semigroup. Eldan and Lee recently established that for every non–negative function  $f$  of integral 1 and any time  $t$  the following tail inequality holds true:

$$\gamma_n(\{Q_t f > r\}) \leq C_t \frac{(\log \log r)^4}{r\sqrt{\log r}}, \quad \forall r > 1$$

where  $C_t$  is a constant depending on  $t$  but not on the dimension. The purpose of the present paper is to simplify parts of their argument and to remove the  $(\log \log r)^4$  factor.

## Résumé

Soit  $\gamma_n$  la mesure Gaussienne standard sur  $\mathbb{R}^n$  et soit  $(Q_t)$  le semi-groupe d’Ornstein–Uhlenbeck. Eldan et Lee ont montré récemment que pour toute fonction positive  $f$  d’intégrale 1 et pour temps  $t$  la queue de distribution de  $Q_t f$  vérifie

$$\gamma_n(\{Q_t f > r\}) \leq C_t \frac{(\log \log r)^4}{r\sqrt{\log r}}, \quad \forall r > 1$$

où  $C_t$  est une constante dépendant seulement de  $t$  et pas de la dimension. L’objet de cet article est de simplifier en partie leur démonstration et d’éliminer le facteur  $(\log \log r)^4$ .

## 1 Introduction

Let  $\gamma_n$  be the standard Gaussian measure on  $\mathbb{R}^n$  and let  $(Q_t)$  be the Ornstein–Uhlenbeck semigroup: for every test function  $f$

$$Q_t f(x) = \int_{\mathbb{R}^n} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) \gamma_n(dy). \quad (1)$$

---

\*CEREMADE (UMR CNRS 7534) Université Paris–Dauphine.

Nelson [6] established that if  $p > 1$  and  $t > 0$  then  $Q_t$  is a contraction from  $L_p(\gamma_n)$  to  $L_q(\gamma_n)$  for some  $q > p$ , namely for

$$q = 1 + e^{2t}(p - 1).$$

The semigroup  $(Q_t)$  is said to be *hypercontractive*. This turns out to be equivalent to the logarithmic Sobolev inequality (see the classical article by Gross [3]). In this paper we establish a regularity property of  $Q_t f$  assuming only that  $f$  is in  $L^1(\gamma_n)$ .

Let  $f$  be a non-negative function satisfying

$$\int_{\mathbb{R}^n} f d\gamma_n = 1,$$

and let  $t > 0$ . Since  $Q_t f \geq 0$  and

$$\int_{\mathbb{R}^n} Q_t f d\gamma_n = \int_{\mathbb{R}^n} f d\gamma_n = 1,$$

Markov inequality gives

$$\gamma_n(\{Q_t f \geq r\}) \leq \frac{1}{r},$$

for all  $r \geq 1$ . Now Markov inequality is only sharp for indicator functions and  $Q_t f$  cannot be an indicator function, so it may be the case that this inequality can be improved. More precisely one might conjecture that for any fixed  $t > 0$  (or at least for  $t$  large enough) there exists a function  $\alpha$  satisfying

$$\lim_{r \rightarrow +\infty} \alpha(r) = 0$$

and

$$\gamma_n(\{Q_t f \geq r\}) \leq \frac{\alpha(r)}{r}, \tag{2}$$

for every  $r \geq 1$  and for every non-negative function  $f$  of integral 1. The function  $\alpha$  should be independent of the dimension  $n$ , just as the hypercontractivity result stated above. Such a phenomenon was actually conjectured by Talagrand in [7] in a slightly different context. He conjectured that the same inequality holds true when  $\gamma_n$  is replaced by the uniform measure on the discrete cube  $\{-1, 1\}^n$  and the Orstein-Uhlenbeck semigroup is replaced by the semigroup associated to the random walk on the discrete cube. The Gaussian version of the conjecture would follow from Talagrand's discrete version by the central limit theorem. In this paper we will only focus on the Gaussian case.

In [1], Ball, Barthe, Bednorz, Oleszkiewicz and Wolff showed that in dimension 1 the inequality (2) holds with decay

$$\alpha(r) = \frac{C}{\sqrt{\log r}},$$

where the constant  $C$  depends on the time parameter  $t$ . Moreover the authors provide an example showing that the  $1/\sqrt{\log r}$  decay is sharp. They also have

a result in higher dimension but they lose a factor  $\log \log r$  and, more importantly, their constant  $C$  then tends to  $+\infty$  (actually exponentially fast) with the dimension. The deadlock was broken recently by Eldan and Lee who showed in [2] that (2) holds with function

$$\alpha(r) = C \frac{(\log \log r)^4}{\sqrt{\log r}},$$

with a constant  $C$  that is independent of the dimension. Again up to the  $\log \log$  factor the result is optimal.

In this article we revisit the argument of Eldan and Lee. We shall simplify some steps of their proof and short cut some others. As a result, we are able to remove the extra  $\log \log$  factor. We would like to make clear though that this note does not really contain any new idea and that the core of our argument is all Eldan and Lee's.

## 2 Main results

Recall that  $\gamma_n$  is the standard Gaussian measure and that  $(Q_t)$  is the Ornstein–Uhlenbeck semigroup, defined by (1). Here is our main result.

**Theorem 1.** *Let  $f$  be a non-negative function on  $\mathbb{R}^n$  satisfying  $\int_{\mathbb{R}^n} f d\gamma_n = 1$  and let  $t > 0$ . Then for every  $r > 1$*

$$\gamma_n(\{Q_t f > r\}) \leq C \frac{\max(1, t^{-1})}{r\sqrt{\log r}},$$

where  $C$  is a universal constant.

As in Eldan and Lee's paper, the Ornstein–Uhlenbeck semigroup only plays a rôle through the following lemma.

**Lemma 2.** *Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}_+$ . For every  $t > 0$ , we have*

$$\nabla^2 \log(Q_t f) \geq -\frac{1}{2t} \text{id},$$

pointwise.

*Proof.* This is really straightforward: observe that (1) can be rewritten as

$$Q_t f(x) = (f * g_{1-\rho})(\rho x),$$

where  $\rho = e^{-t}$  and  $g_{1-\rho}$  is the density of the Gaussian measure with mean 0 and covariance  $(1-\rho)\text{id}$ . Then differentiate twice and use the Cauchy–Schwarz inequality. Details are left to the reader.  $\square$

What we actually prove is the following, where  $Q_t$  does not appear anymore.

**Theorem 3.** Let  $f$  be a positive function on  $\mathbb{R}^n$  satisfying  $\int_{\mathbb{R}^n} f d\gamma_n = 1$ . Assume that  $f$  is smooth and satisfies

$$\nabla^2 \log f \geq -\beta \text{id} \quad (3)$$

pointwise, for some  $\beta \geq 0$ . Then for every  $r > 1$

$$\gamma_n(\{f > r\}) \leq \frac{C \max(\beta, 1)}{r\sqrt{\log r}},$$

where  $C$  is a universal constant.

Obviously, Theorem 3 and Lemma 2 altogether yield Theorem 1.

Let us comment on the optimality of Theorem 1 and Theorem 3. In dimension 1, consider the function

$$f_\alpha(x) = e^{\alpha x - \alpha^2/2}.$$

Observe that  $f_\alpha \geq 0$  and that  $\int_{\mathbb{R}} f_\alpha d\gamma_1 = 1$ . Note that for every  $t \geq 1$  we have

$$\gamma_1([t, +\infty)) \geq \frac{c e^{-t^2/2}}{t},$$

where  $c$  is a universal constant. So if  $\alpha > 0$  and  $r \geq e$  then

$$\gamma_1(\{f_\alpha \geq r\}) \geq \frac{c \exp\left(-\frac{1}{2}\left(\frac{\log r}{\alpha} + \frac{\alpha}{2}\right)^2\right)}{\frac{\log r}{\alpha} + \frac{\alpha}{2}}.$$

Choosing  $\alpha = \sqrt{2 \log r}$  we get

$$\gamma_1(\{f_\alpha \geq r\}) \geq \frac{c'}{r\sqrt{\log r}}.$$

Since  $(\log f_\alpha)'' = 0$  this shows that the dependence in  $r$  in Theorem 3 is sharp. Actually this example also shows that the dependence in  $r$  in Theorem 1 is sharp. Indeed, it is easily seen that

$$Q_t f_\alpha = f_{\alpha e^{-t}},$$

for every  $\alpha \in \mathbb{R}$  and  $t > 0$ . This implies that  $f_\alpha$  always belongs to the image  $Q_t$ . Of course, this example also works in higher dimension: just replace  $f_\alpha$  by

$$f_u(x) = e^{\langle u, x \rangle - |u|^2/2}$$

where  $u$  belongs to  $\mathbb{R}^n$ .

**Theorem 4.** Let  $X$  be a random vector having density  $f$  with respect to the Gaussian measure, and assume that  $f$  satisfies (3). Then for every  $r > 1$

$$\mathbb{P}(f(X) \in (r, er]) \leq C \frac{\max(\beta, 1)}{\sqrt{\log r}}.$$

Theorem 4 easily yields Theorem 3.

*Proof of Theorem 3.* Let  $G$  be standard Gaussian vector on  $\mathbb{R}^n$  and let  $X$  be a random vector having density  $f$  with respect to  $\gamma_n$ . Then using Theorem 4

$$\begin{aligned}
\mathbb{P}[f(G) > r] &= \sum_{k=0}^{+\infty} \mathbb{P}(f(G) \in (e^k r, e^{k+1} r]) \\
&\leq \sum_{k=0}^{+\infty} (e^k r)^{-1} \mathbb{E} [f(G) \mathbf{1}_{\{f(G) \in (e^k r, e^{k+1} r]\}}] \\
&= \sum_{k=0}^{+\infty} (e^k r)^{-1} \mathbb{P}(f(X) \in (e^k r, e^{k+1} r]) \\
&\leq \sum_{k=0}^{+\infty} (e^k r)^{-1} C \frac{\max(\beta, 1)}{\sqrt{\log(e^k r)}} \\
&\leq C \frac{e-1}{e-1} \frac{1}{r} \frac{\max(\beta, 1)}{\sqrt{\log r}},
\end{aligned}$$

which is the result.  $\square$

The rest of the note is devoted to the proof of Theorem 4.

### 3 Preliminaries: the stochastic construction

Let  $\mu$  be a probability measure on  $\mathbb{R}^n$  having density  $f$  with respect to the Gaussian measure. We shall assume that  $f$  is bounded away from 0, that  $f$  is  $\mathcal{C}^2$  and that  $\nabla f$  and  $\nabla^2 f$  are bounded. A simple density argument shows that we do not lose generality by adding these technical assumptions.

Eldan and Lee's argument is based on a stochastic construction which we describe now. Let  $(B_t)$  be a standard  $n$ -dimensional Brownian motion and let  $(P_t)$  be the associated semigroup:

$$P_t h(x) = \mathbb{E}[h(x + B_t)],$$

for all test functions  $h$ . Note that  $(P_t)$  is the heat semigroup, not the Ornstein-Uhlenbeck semigroup. Consider the stochastic differential equation

$$\begin{cases} X_0 = 0 \\ dX_t = dB_t + \nabla \log(P_{1-t} f)(X_t) dt, \quad t \in [0, 1]. \end{cases} \quad (4)$$

The technical assumptions made on  $f$  insure that the map

$$x \mapsto \nabla \log P_{1-t} f(x)$$

is Lipschitz, with a Lipschitz norm that does not depend on  $t \in [0, 1]$ . So the equation (4) has a strong solution  $(X_t)$ . In our previous work [4] we study the

process  $(X_t)$  in details and we give some applications to functional inequalities. Let us recap here some of these properties and refer to [4, section 2.5] for proofs. Recall that if  $\mu_1, \mu_2$  are two probability measures, the relative entropy of  $\mu_1$  with respect to  $\mu_2$  is defined by

$$H(\mu_1 | \mu_2) = \int \log \left( \frac{d\mu_1}{d\mu_2} \right) d\mu_1,$$

if  $\mu_1$  is absolutely continuous with respect to  $\mu_2$  (and  $H(\mu_1 | \mu_2) = +\infty$  otherwise). Also in the sequel we call *drift* any process  $(u_t)$  taking values in  $\mathbb{R}^n$  which is adapted to the natural filtration of  $(B_t)$  (this means that  $u_t$  depends only on  $(B_s)_{s \leq t}$ ) and satisfies

$$\int_0^1 |u_t|^2 ds < +\infty,$$

almost surely. Let  $(v_t)$  be the drift

$$v_t = \nabla \log P_{1-t} f(X_t).$$

Using Itô's formula it is easily seen that

$$d \log P_{1-t} f(X_t) = \langle v_t, dB_t \rangle + \frac{1}{2} |v_t|^2 dt.$$

Therefore, for every  $t \in [0, 1]$

$$P_{1-t} f(X_t) = \exp \left( \int_0^t \langle v_s, dB_s \rangle + \frac{1}{2} \int_0^t |v_s|^2 ds \right). \quad (5)$$

Combining this with the Girsanov change of measure theorem one can show that the random vector  $X_1$  has law  $\mu$  (again we refer to [4] for details). Moreover we have the equality

$$H(\mu | \gamma_n) = \frac{1}{2} \mathbb{E} \left[ \int_0^1 |v_s|^2 ds \right]. \quad (6)$$

Also, if  $(u_t)$  is any drift and if  $\nu$  is the law of

$$B_1 + \int_0^1 u_t dt,$$

then

$$H(\nu | \gamma_n) \leq \frac{1}{2} \mathbb{E} \left[ \int_0^1 |u_s|^2 ds \right]. \quad (7)$$

So the drift  $(v_t)$  is in some sense optimal. Lastly, and this will play a crucial rôle in the sequel, the process  $(v_t)$  is a martingale.

Eldan and Lee introduce a perturbed version of the process  $(X_t)$ , which we now describe. From now on we fix  $r > 1$  and we let

$$T = \inf \{ t \in [0, 1], P_{1-t} f(X_t) > r \} \wedge 1$$

be the first time the process  $(P_{1-t}f(X_t))$  hits the value  $r$ , with the convention that  $T = 1$  if it does not ever reach  $r$  (note that our definition of  $T$  differs from the one of Eldan and Lee). Now given  $\delta > 0$  we let  $(X_t^\delta)$  be the process defined by

$$X_t^\delta = X_t + \delta \int_0^{T \wedge t} v_s ds.$$

Since  $T$  is a stopping time, this perturbed process is still of the form Brownian motion plus drift:

$$X_t^\delta = B_t + \int_0^t (1 + \delta \mathbf{1}_{\{s \leq T\}}) v_s ds.$$

So letting  $\mu^\delta$  be the law of  $X_1^\delta$  and using (7) we get

$$\begin{aligned} \mathbb{H}(\mu^\delta \mid \gamma_n) &\leq \frac{1}{2} \mathbb{E} \left[ \int_0^1 (1 + \delta \mathbf{1}_{\{s \leq T\}})^2 |v_s|^2 ds \right] \\ &= \frac{1}{2} \mathbb{E} \left[ \int_0^1 |v_s|^2 ds \right] + \left( \delta + \frac{\delta^2}{2} \right) \mathbb{E} \left[ \int_0^T |v_s|^2 ds \right]. \end{aligned} \quad (8)$$

## 4 Proof of the main result

The proof can be decomposed into two steps. Recall that  $r$  is fixed from the beginning and that  $X_1^\delta$  actually depends on  $r$  through the stopping time  $T$ .

The first step is to prove that if  $\delta$  is small then  $\mu$  and  $\mu^\delta$  are not too different.

**Proposition 5.** *Assuming (3), we have*

$$d_{TV}(\mu, \mu^\delta) \leq \delta \sqrt{(\beta + 1) \log r},$$

for every  $\delta > 0$ , where  $d_{TV}$  denotes the total variation distance.

The second step is to argue that  $f(X_1^\delta)$  tends to be bigger than  $f(X_1)$ . An intuition for this property is that the difference between  $X_1^\delta$  and  $X_1$  is somehow in the direction of  $\nabla f(X_1)$ .

**Proposition 6.** *Assuming (3), we have*

$$\mathbb{P}(f(X_1^\delta) \leq r^{1+2\delta} e^{-4}) \leq \mathbb{P}(f(X_1) \leq r) + (\beta + 4)\delta^2 \log(r),$$

for all  $\delta > 0$ .

*Remark.* Note that both propositions use the convexity hypothesis (3).

It is now very easy to prove Theorem 4. Since  $X_1$  has law  $\mu$ , all we need to prove is

$$\mathbb{P}(f(X_1) \in (r, er]) \leq C \frac{\max(\beta, 1)}{\sqrt{\log r}}.$$

We choose

$$\delta = \frac{5}{2 \log r}.$$



For this value of  $\delta$ , Proposition 6 gives

$$\mathbb{P}(f(X_1^\delta) \leq er) \leq \mathbb{P}(f(X_1) \leq r) + \frac{25}{4} \frac{\beta + 4}{\log r},$$

whereas Proposition 5 yields

$$\begin{aligned} \mathbb{P}(f(X_1) \leq er) &\leq \mathbb{P}(f(X_1^\delta) \leq er) + d_{TV}(\mu, \mu^\delta) \\ &\leq \mathbb{P}(f(X_1^\delta) \leq er) + \frac{5}{2} \left( \frac{\beta + 1}{\log r} \right)^{1/2}. \end{aligned}$$

Combining the two inequalities we obtain

$$\mathbb{P}(f(X_1) \leq er) \leq \mathbb{P}(f(X_1) \leq r) + C \frac{\max(\beta, 1)}{\sqrt{\log r}},$$

which is the result.

*Remark.* We actually prove the slightly stronger statement:

$$\mathbb{P}(f(X_1) \in (r, er]) \leq C \max \left( \frac{\max(\beta, 1)}{\log r}, \left( \frac{\max(\beta, 1)}{\log r} \right)^{1/2} \right).$$

## 5 Proof of the total variation estimate

We actually bound the relative entropy of  $\mu^\delta$  with respect to  $\mu$ . Recall that  $\log f$  is assumed to be weakly convex: there exists  $\beta \geq 0$  such that

$$\nabla^2 \log f \geq -\beta \text{id}, \tag{9}$$

pointwise.

**Proposition 7.** *Assuming (9), we have*

$$\mathbb{H}(\mu^\delta | \mu) \leq \delta^2 (\beta + 1) \log r,$$

for all  $\delta > 0$ .

This yields Proposition 5 by Pinsker's inequality.

*Proof.* Observe that

$$\mathbb{H}(\mu^\delta | \mu) = \mathbb{H}(\mu^\delta | \gamma_n) - \int_{\mathbb{R}^n} \log(f) d\mu^\delta. \tag{10}$$

Now (9) gives

$$\begin{aligned} \log(f)(X_1^\delta) &\geq \log f(X_1) + \langle \nabla \log f(X_1), X_1^\delta - X_1 \rangle - \frac{\beta}{2} |X_1^\delta - X_1|^2, \\ &\geq \log f(X_1) + \delta \int_0^T \langle v_1, v_s \rangle ds - \frac{\beta \delta^2}{2} \int_0^T |v_s|^2 ds, \end{aligned} \tag{11}$$

almost surely. We shall use this inequality several times in the sequel. Recall that  $X_1$  has law  $\mu$  and that  $X_1^\delta$  has law  $\mu^\delta$ . Taking expectation in the previous inequality and using (10) we get

$$\begin{aligned} \mathbf{H}(\mu^\delta | \mu) &\leq \mathbf{H}(\mu^\delta | \gamma_n) - \mathbf{H}(\mu | \gamma_n) \\ &\quad - \delta \mathbf{E} \left[ \int_0^T \langle v_1, v_s \rangle ds \right] + \frac{\beta \delta^2}{2} \mathbf{E} \left[ \int_0^T |v_s|^2 ds \right]. \end{aligned}$$

Together with (6) and (8) we obtain

$$\mathbf{H}(\mu^\delta | \mu) \leq -\delta \mathbf{E} \left[ \int_0^T \langle v_1 - v_s, v_s \rangle ds \right] + \frac{(1 + \beta)\delta^2}{2} \mathbf{E} \left[ \int_0^T |v_s|^2 ds \right].$$

Now since  $(v_t)$  is a martingale and  $T$  a stopping time we have

$$\mathbf{E} [\langle v_1, v_s \rangle \mathbf{1}_{\{s \leq T\}}] = \mathbf{E} [|v_s|^2 \mathbf{1}_{\{s \leq T\}}]$$

for all time  $s \leq 1$ . This shows that the first term in the previous inequality is 0. To bound the second term, observe that the definition of  $T$  and the equality (5) imply that

$$\int_0^T \langle v_s, dB_s \rangle + \frac{1}{2} \int_0^T |v_s|^2 ds \leq \log r,$$

almost surely. Since  $(v_t)$  is a bounded drift, the process  $(\int_0^t \langle v_s, dB_s \rangle)$  is a martingale. Now  $T$  is a bounded stopping time, so by the optional stopping theorem

$$\mathbf{E} \left[ \int_0^T \langle v_s, dB_s \rangle \right] = 0.$$

Therefore, taking expectation in the previous inequality yields

$$\mathbf{E} \left[ \int_0^T |v_s|^2 ds \right] \leq 2 \log r,$$

which concludes the proof.  $\square$

## 6 Proof of Proposition 6

The goal is to prove that

$$\mathbf{P}(f(X_1^\delta) \leq r^{1+2\delta} e^{-4}) \leq \mathbf{P}(f(X_1) \leq r) + \delta^2(\beta + 4) \log r.$$

Obviously

$$\mathbf{P}(f(X_1^\delta) \leq r^{1+2\delta} e^{-4}) \leq \mathbf{P}(f(X_1) \leq r) + \mathbf{P}(f(X_1^\delta) \leq r^{1+2\delta} e^{-4}; f(X_1) > r)$$

Now recall the inequality (11) coming from the weak convexity of  $\log f$  and rewrite it as

$$\log f(X_1^\delta) \geq K_1 + 2\delta K_T + Y$$

where  $(K_t)$  is the process defined by

$$K_t = \log(P_{1-t})(f)(X_t) = \int_0^t \langle v_s, dB_s \rangle + \frac{1}{2} \int_0^t |v_s|^2 ds,$$

and  $Y$  is the random variable

$$Y = -2\delta \int_0^T \langle v_s, dB_s \rangle + \delta \int_0^T \langle v_1 - v_s, v_s \rangle ds - \frac{\beta\delta^2}{2} \int_0^T |v_s|^2 ds.$$

Recall that the stopping time  $T$  is the first time the process  $(K_t)$  exceeds the value  $\log r$  if it ever does, and  $T = 1$  otherwise. In particular, if

$$K_1 = \log f(X_1) > \log r$$

then  $K_T = \log r$ . So if  $f(X_1) > r$  then

$$f(X_1^\delta) > r^{1+2\delta} e^Y.$$

Therefore

$$\mathbb{P}(f(X_1^\delta) \leq r^{1+2\delta} e^{-4}; f(X_1) > r) \leq \mathbb{P}(Y \leq -4).$$

So we are done if we can prove that

$$\mathbb{P}(Y \leq -4) \leq (\beta + 4)\delta^2 \log r. \quad (12)$$

There are three terms in the definition of  $Y$ . The problematic one is

$$\delta \int_0^T \langle v_1 - v_s, v_s \rangle ds.$$

We know from the previous section that it has expectation 0. A natural way to get a deviation bound would be to estimate its second moment but it is not clear to us how to do this. Instead we make a complicated detour.

**Lemma 8.** *Let  $Z$  be an integrable random variable satisfying  $\mathbb{E}[e^Z] \leq 1$ . Then*

$$\mathbb{P}(Z \leq -2) \leq -\mathbb{E}[Z].$$

*Remark.* Note that  $\mathbb{E}[Z] \leq 0$  by Jensen's inequality.

*Proof.* Simply write

$$\begin{aligned} \mathbb{E}[e^Z] &\geq \mathbb{E}[e^Z \mathbf{1}_{\{Z > -2\}}] \\ &\geq \mathbb{E}[(Z + 1) \mathbf{1}_{\{Z > -2\}}] \\ &= \mathbb{E}[Z] - \mathbb{E}[Z \mathbf{1}_{\{Z \leq -2\}}] + 1 - \mathbb{P}(Z \leq -2) \\ &\geq \mathbb{E}[Z] + \mathbb{P}(Z \leq -2) + 1. \end{aligned}$$

So if  $\mathbb{E}[e^Z] \leq 1$  then  $\mathbb{P}(Z \leq -2) \leq -\mathbb{E}[Z]$ . □

**Lemma 9.** *Let  $Z$  be the variable*

$$Z = -\delta \int_0^T \langle v_s, dB_s \rangle + \delta \int_0^T \langle v_1 - v_s, v_s \rangle ds - \frac{(\beta + 1)\delta^2}{2} \int_0^T |v_s|^2 ds.$$

*Then*

$$\mathbb{P}(Z \leq -2) \leq \delta^2(\beta + 1) \log r.$$

*Proof.* As we have seen before the first two terms in the definition of  $Z$  have expectation 0 and

$$\mathbb{E}[Z] = -\frac{(\beta + 1)\delta^2}{2} \mathbb{E} \left[ \int_0^T |v_s|^2 ds \right] \geq -\delta^2(\beta + 1) \log r.$$

By Lemma 8 it is enough to show that  $\mathbb{E}[e^Z] \leq 1$ . To do so, we use the Girsanov change of measure formula. The process  $(X_t^\delta)$  is of the form Brownian motion plus drift:

$$\begin{aligned} X_t^\delta &= X_t + \delta \int_0^{T \wedge t} v_s ds \\ &= B_t + \int_0^t (1 + \delta \mathbf{1}_{\{s \leq T\}}) v_s ds, \end{aligned}$$

Note also that the drift term is bounded. Therefore, Girsanov's formula applies, see for instance [5, chapter 6] (beware that the authors oddly use the letter  $M$  to denote expectation). The process  $(D_t^\delta)$  defined by

$$D_t^\delta = \exp \left( - \int_0^t (1 + \delta \mathbf{1}_{\{s \leq T\}}) \langle v_s, dB_s \rangle - \frac{1}{2} \int_0^t |(1 + \delta \mathbf{1}_{\{s \leq T\}}) v_s|^2 ds \right)$$

is a non-negative martingale of expectation 1 and under the measure  $\mathbb{Q}^\delta$  defined by

$$d\mathbb{Q}^\delta = D_1^\delta d\mathbb{P}$$

the process  $(X_t^\delta)$  is a standard Brownian motion. In particular

$$\mathbb{E}[f(X_1^\delta) D_1^\delta] = \mathbb{E}[f(B_1)] = 1.$$

Now use inequality (11) once again and combine it with equality (5) written at time  $t = 1$ . This gives exactly

$$f(X_1^\delta) D_1^\delta \geq e^Z.$$

Therefore  $\mathbb{E}[e^Z] \leq 1$ , which concludes the proof.  $\square$

We now prove inequality (12). The idea being that the annoying term in  $Y$  is handled by the previous lemma. Observe that

$$Y = Z - \delta \int_0^T \langle v_s, dB_s \rangle - \frac{\delta^2}{2} \int_0^T |v_s|^2 ds.$$

So

$$\begin{aligned} \mathbb{P}(Y \leq -4) &\leq \mathbb{P}(Z \leq -2) + \mathbb{P}\left(\delta \int_0^T \langle v_s, dB_s \rangle \geq 1\right) \\ &\quad + \mathbb{P}\left(\frac{\delta^2}{2} \int_0^T |v_s|^2 ds \geq 1\right). \end{aligned}$$

Recall that  $\int_0^T \langle v_s, dB_s \rangle$  has mean 0 and observe that

$$\mathbb{E}\left[\left(\delta \int_0^T \langle v_s, dB_s \rangle\right)^2\right] = \delta^2 \mathbb{E}\left[\int_0^T |v_s|^2 ds\right] \leq 2\delta^2 \log r.$$

So by Tchebychev inequality

$$\mathbb{P}\left(\delta \int_0^T \langle v_s, dB_s \rangle \geq 1\right) \leq 2\delta^2 \log r.$$

Similarly by Markov inequality

$$\mathbb{P}\left(\frac{\delta^2}{2} \int_0^T |v_s|^2 ds \geq 1\right) \leq \delta^2 \log r.$$

Putting everything together we get (12), which concludes the proof.

### Acknowledgments

We would like to thank Ronen Eldan and James Lee for introducing us to the problem and communicating their work to us. We are also grateful to Michel Ledoux for some fruitful discussions at the final stage of this work.

### References

- [1] Ball, K.; Barthe, F.; Bednorz, W.; Oleszkiewicz, K.; Wolff, P. *L<sub>1</sub>-smoothing for the Ornstein–Uhlenbeck semigroup*. *Mathematika* 59 (2013), no. 1, 160–168.
- [2] Eldan, R.; Lee, J. *Regularization under diffusion and anti-concentration of temperature*. arXiv:1410.3887.
- [3] Gross, L. *Logarithmic Sobolev inequalities*. *Amer. J. Math.* 97 (1975), no. 4, 1061–1083.
- [4] Lehec, J. *Representation formula for the entropy and functional inequalities*. *Ann. Inst. Henri Poincaré Probab. Stat.* 49 (2013), no. 3, 885–899.

- [5] Liptser, R.; Shiryaev, A. Statistics of random processes. Vol I, general theory. 2nd edition. *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2001.
- [6] Nelson, E. *The free Markoff field*. J. Functional Analysis 12 (1973), 211–227.
- [7] Talagrand, M. *A conjecture on convolution operators, and a non-Dunford-Pettis operator on  $L_1$* . Israel J. Math. 68 (1989), no. 1, 82–88.