Bernoulli **25**(4B), 2019, 3978–4006
https://doi.org/10.3150/19-BEJ1117

Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one

ADRIEN SAUMARD

CREST, ENSAI, Université Bretagne Loire, Campus de ker-lann, Rue Blaise Pascal – BP 37203 35712 Bruz Cedex, France. E-mail: asaumard@gmail.com

We investigate links between the so-called Stein's density approach in dimension one and some functional and concentration inequalities. We show that measures having a finite first moment and a density with connected support satisfy a weighted Poincaré inequality with the weight being the Stein kernel, that indeed exists and is unique in this case. Furthermore, we prove weighted log-Sobolev and asymmetric Brascamp–Lieb type inequalities related to Stein kernels. We also show that existence of a uniformly bounded Stein kernel is sufficient to ensure a positive Cheeger isoperimetric constant. Then we derive new concentration inequalities. In particular, we prove generalized Mills' type inequalities when a Stein kernel is uniformly bounded and sub-gamma concentration for Lipschitz functions of a variable with a sub-linear Stein kernel. More generally, when some exponential moments are finite, the Laplace transform of the random variable of interest is shown to bounded from above by the Laplace transform of the Stein kernel. Along the way, we prove a general lemma for bounding the Laplace transform of a random variable, that may be of independent interest. We also provide density and tail formulas as well as tail bounds, generalizing previous results that where obtained in the context of Malliavin calculus.

Keywords: concentration inequality; covariance identity; isoperimetric constant; Stein kernel; tail bound; weighted log-Sobolev inequality; weighted Poincaré inequality

1. Introduction

Since its introduction by Charles Stein (Stein [52,53]), the so-called Stein's method is a corpus of techniques that revealed itself very successful in studying probability approximation and convergence in law (see, for instance, Chen, Goldstein and Shao [23], Chatterjee [21], Ley, Reinert and Swan [40] and references therein). Much less is known regarding the interplay between Stein's method and functional inequalities. Recently, a series of papers (Ledoux, Nourdin and Peccati [38], Ledoux, Nourdin and Peccati [39], Fathi and Nelson [31], Courtade, Fathi and Pananjady [26]) started to fill this gap.

More precisely, Ledoux, Nourdin and Peccati [38] provide some improvement of the log-Sobolev inequality and Talagrand's quadratic transportation cost inequality through the use of a Stein kernel and in particular, the Stein discrepancy that measures the closeness of the Stein kernels to identity. In a second paper Ledoux, Nourdin and Peccati [39], these authors also provide a lower bound of the deficit in the Gaussian log-Sobolev inequality in terms Stein's characterization of the Gaussian distribution. Recently, Fathi and Nelson [31] also consider free Stein kernel

1350-7265

and use it to improve the free log-Sobolev inequality. Finally, Courtade, Fathi and Pananjady [26] proved that the existence of a reversed weighted Poincaré inequality is sufficient to ensure existence of a Stein kernel. To do so, they use an elegant argument based on the Lax–Milgram theorem. They also provide bounds on the Stein discrepancy and application to a quantitative central limit theorem.

Particularizing to dimension one, the present paper aims at pursuing investigations about the relations between Stein's method – especially Stein kernels – and some functional inequalities, together with some concentration inequalities. The limitation to dimension one comes, for most of the results, from the one-dimensional nature of the covariance identities given in Section 2, that are instrumental for the rest of the paper and which crucially rely on the use of the properties of the cumulative distribution function.

We prove that a measure ν having a finite first moment and a density with connected support satisfies a weighted Poincaré inequality in the sense of Bobkov and Ledoux [11], with the weight being the Stein kernel τ_ν (see the definition in Section 2 below), that is unique in this case. More precisely, for any $f \in L_2(\nu)$, absolutely continuous, we have

$$\operatorname{Var}\bigl(f(X)\bigr) \leq \mathbb{E}\bigl[\tau_\nu(X)\bigl(f'(X)\bigr)^2\bigr]. \tag{1}$$

The latter inequality allows us to recover by different techniques some weighted Poincaré inequalities previously established in Bobkov and Ledoux [12] for the Beta distribution or in Bonnefont, Joulin and Ma [14] for the generalized Cauchy distribution and to highlight new ones, considering for instance Pearson's class of distributions.

It is also well known that Muckenhoupt-type criteria characterize (weighted) Poincaré and log-Sobolev inequalities on the real line (Ané et al. [1], Bobkov and Götze [6]). We indeed recover, up to a multiplicative constant, inequality (1) from the classical Muckenhoupt criterion. Furthermore, using the criterion first established by Bobkov and Götze [6] to characterize log-Sobolev inequalities on the real line, we prove that, under the conditions ensuring the weighted Poincaré inequality (1), together with some asymptotic assumptions on the behavior of the Stein kernel around the edges of the support of the measure ν, the following inequality holds,

$$\operatorname{Ent}_\nu\bigl(g^2\bigr) \leq C_\nu \int \tau_\nu^2\bigl(g'\bigr)^2\, d\nu, \tag{2}$$

for some constant $C_\nu > 0$ and with $\operatorname{Ent}_\nu(g^2) = \int g^2 \log g^2\, d\nu - \int g^2\, d\nu \log \int g^2\, d\nu$. More precisely, if the support of ν is compact, then inequality (2) is valid if τ_ν^{-1} is integrable at the edges of the support. If on contrary, an edge of the support is infinite, then a necessary condition for inequality (2) to hold is that the Stein kernel does not tend to zero around this edge. In the way, we provide an improvement in dimension one of a weighted log-Sobolev inequality for the Cauchy measure due to Bobkov and Ledoux [11], that is of independent interest.

We also show that existence of a uniformly bounded Stein kernel is nearly sufficient to ensure a positive Cheeger isoperimetric constant. In addition, we derive asymmetric Brascamp–Lieb type inequalities related to the Stein kernel.

There is also a growing literature, initiated by Chatterjee [19], about the links between Stein's method and concentration inequalities. Several approaches are considered, from the method of exchangeable pairs (Chatterjee [19], Chatterjee and Dey [22], Mackey et al. [43], Paulin, Mackey and Tropp [48]), to the density approach coupled with Malliavin calculus (Nourdin and Viens [47], Viens [55], Eden and Viens [29], Treilhard and Mansouri [54]), size biased coupling (Ghosh and Goldstein [32,33], Ghosh, Goldstein and Raič [34], Cook, Goldstein and Johnson [24]), zero bias coupling (Goldstein and Işlak [35]) or more general Stein couplings (Barbour, Ross and Wen [4]). As emphasized for instance, in the survey by Chatterjee [21], one major strength of Stein-type methodologies applied to concentration of measure is that it often allows to deal with dependent and complex system of random variables, finding for instance applications in statistical mechanics or in random graph theory.

In the present work, we investigate relationships between Stein kernels and concentration of measure by building upon ideas and exporting techniques about the use of covariance identities for Gaussian concentration from Bobkov, Götze and Houdré [7].

Considering first the case where a Stein kernel is uniformly bounded, we recover the well-known fact that the associated random variable admits a sub-Gaussian behavior. But we also prove in this setting some refined concentration inequalities, that we call generalized Mills' type inequalities, in reference to the classical Mills' inequality for the normal distribution (see for instance Dümbgen [28]). Assume also that a Stein kernel τ_v exists for the measure ν, is uniformly bounded, and denote $c = \|\tau_v\|_\infty^{-1}$. Then the function $T_g(r) = e^{cr^2/2}\mathbb{E}(g - \mathbb{E}g)\mathbf{1}_{\{g-\mathbb{E}g\geq r\}}$ is non-increasing in $r \geq 0$. Consequently, for all $r > 0$,

$$\mathbb{P}(g - \mathbb{E}g \geq r) \leq \mathbb{E}(g - \mathbb{E}g)_+ \frac{e^{-cr^2/2}}{r}. \tag{3}$$

In particular, Beta distributions have a bounded Stein kernel and our concentration inequalities improve on previously best known concentration inequalities for Beta distributions, recently due to Bobkov and Ledoux [12].

Furthermore, we consider the situation where a Stein kernel has a sub-linear behavior, recovering and extending in this case sub-Gamma concentration previously established by Nourdin and Viens [47]. We also prove some generalized Mills' type inequalities in this case. More generally, we prove that the Laplace transform of a Stein kernel controls the Laplace transform of a Lipschitz function taken on the related distribution. Take f a 1-Lipschitz function with mean zero with respect to ν and assume that f has an exponential moment with respect to ν, that is there exists $a > 0$ such that $\mathbb{E}[e^{af(X)}] < +\infty$. Then for any $\lambda \in (0, a)$,

$$\mathbb{E}\big[e^{\lambda f(X)}\big] \leq \mathbb{E}\big[e^{\lambda^2 \tau_v(X)}\big]. \tag{4}$$

It is worth noting that to prove such a result, we state a generic lemma – Lemma 15 Section 4 – allowing to bound the Laplace transform of a random variable. We believe that this lemma has an interest by itself, as it may be convenient when dealing with Chernoff's method in general.

We also obtain lower tail bounds without the need of Malliavin calculus, thus extending previous results due to Nourdin and Viens [47] and Viens [55].

The paper is organized as follows. In Section 2, we introduce some background material, by discussing some well-known and new formulas for Stein kernels and Stein factors in connection with Menz and Otto's covariance identity. We also provide formulas involving the Stein operator, for densities and tails. Then we prove in Section 3 some (weighted) functional inequalities linked to the behavior of the Stein kernel. In Section 4, we make use of some covariance inequalities to derive various concentration inequalities for Lipschitz functions of a random variable having a Stein kernel. Finally, we prove some tail bounds related to the behavior of the Stein kernel – assumed to be unique – in Section 5. Some further results, proofs and comments are gathered in a supplemental article Saumard [49].

2. On covariance identities and the Stein kernel

Take a real random variable X of distribution ν with density p with respect to the Lebesgue measure on $\mathbb{R}$ and cumulative distribution function F. Assume that the mean of the distribution ν exists and denote it by $\mu = \mathbb{E}[X]$. Denote also $\mathrm{Supp}(\nu) = \overline{\{x \in \mathbb{R} : p(x) > 0\}} \subset \bar{\mathbb{R}}(:= \mathbb{R} \cup \{-\infty, +\infty\})$ the support of the measure ν, defined as the closure of the set where the density is positive and assume that this support is connected. We denote by $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{+\infty\}$, $a < b$, the edges of $\mathrm{Supp}(\nu)$. For convenience, we also denote by $I(\nu) = \mathrm{Int}(\mathrm{Supp}(\nu))$ the interior of the support of ν. The distribution ν is said to have a Stein kernel τ_ν, if the following identity holds true,

$$\mathbb{E}\big[(X - \mu)\varphi(X)\big] = \mathbb{E}\big[\tau_\nu(X)\varphi'(X)\big], \tag{5}$$

with φ being any differentiable test function such that the functions $x \mapsto (x - \mu)\varphi(x)$ and $x \mapsto \tau_\nu(x)\varphi'(x)$ are ν-integrable and $[\tau_\nu p\varphi]_a^b = 0$. It is well-known (Ledoux, Nourdin and Peccati [38], Courtade, Fathi and Pananjady [26], Ley, Reinert and Swan [40]), that under our assumptions the Stein kernel τ_ν exists, is unique up to sets of ν-measure zero and a version of the latter is given by the following formula,

$$\tau_\nu(x) = \frac{1}{p(x)} \int_x^\infty (y - \mu) p(y)\, dy, \tag{6}$$

for any $x \in I(\nu)$. Formula (6) comes from a simple integration by parts. Notice that τ_ν is almost surely positive on the interior of the support of ν.

Although we will focus only on dimension one, it is worth noting that the definition of a Stein kernel extends to higher dimension, where it is matrix-valued. The question of existence of the Stein kernel for a particular multi-dimensional measure ν is nontrivial and only a few general results are known related to this problem (see, for instance, Ledoux, Nourdin and Peccati [38], Courtade, Fathi and Pananjady [26] and Fathi [30]). In particular, Courtade, Fathi and Pananjady [26] proves that the existence of a Stein kernel is ensured whenever a (converse weighted) Poincaré inequality is satisfied for the probability measure ν. Recently, Stein kernels that are positive definite matrices have been constructed in Fathi [30] using transportation techniques.

In this section, that essentially aims at stating some background results that will be instrumental for the rest of the paper, we will among other things recover Identity (6) and introduce a new formula for the Stein kernel by means of a covariance identity recently obtained in Menz and Otto [44] and further developed in Saumard and Wellner [50]. It actually appears that Menz and Otto's covariance identity is a consequence of an old result by Hoeffding (see the discussion in Saumard and Wellner [51]).

We define a non-negative and symmetric kernel k_ν on $\mathbb{R}^2$ by

$$k_\nu(x,y) = F(x \wedge y) - F(x)F(y), \quad \text{for all } (x,y) \in \mathbb{R}^2. \tag{7}$$

For any $p \in [1,+\infty]$, we denote by $L_p(\nu)$ the space of measurable functions f such that $\|f\|_p^p = \int |f|^p \, d\nu < +\infty$ for $p \in [1,+\infty)$ and $\|f\|_\infty = \operatorname{ess\,sup}_{x\in\mathbb{R}} |f(x)| < +\infty$ for $p = +\infty$. If $f \in L_p(\nu)$, $g \in L_q(\nu)$, $p^{-1} + q^{-1} = 1$, we also write

$$\operatorname{Cov}(f,g) = \int \left(f - \int f \, d\nu \right) g \, d\nu$$

the covariance of f and g with respect to ν. For $f \in L_2(\nu)$, we write $\operatorname{Var}(f) = \operatorname{Cov}(f,f)$ the variance of f with respect to ν. For a random variable X of distribution ν, we will also write $\mathbb{E}[h(X)] = \mathbb{E}[h] = \int h \, d\nu$.

Proposition 1 (Corollary 2.2, Saumard and Wellner [50]). *If g and h are absolutely continuous and $g \in L_p(\nu)$, $h \in L_q(\nu)$ for some $p \in [1,\infty]$ and $p^{-1} + q^{-1} = 1$, then*

$$\operatorname{Cov}(g,h) = \iint_{\mathbb{R}^2} g'(x) k_\nu(x,y) h'(y) \, dx \, dy. \tag{8}$$

It is worth mentioning that the covariance identity (8) heavily relies on dimension one, since it uses the properties of the cumulative distribution function F through the kernel k_ν. In dimension greater than or equal to 2, a covariance identity of the form of (8) – with derivatives replaced by gradients, – would actually imply that the measure ν is Gaussian (for more details, see Bobkov, Götze and Houdré [7] and also Remark 2 below).

Remark 2. In the context of goodness-of-fit tests, Liu, Lee and Jordan [41] introduce the notion of kernelized Stein discrepancy as follows. If $K(x,y)$ is a kernel on $\mathbb{R}^2$, p and q are two densities and (X,Y) is a pair of independent random variables distributed according to p, then the kernelized Stein discrepancy $\mathbb{S}_K(p,q)$ between p and q related to K is

$$\mathbb{S}_K(p,q) = \mathbb{E}\big[\delta_{q,p}(X) K(X,Y) \delta_{q,p}(Y)\big],$$

where $\delta_{q,p}(x) = (\log q(x))' - (\log p(x))'$ is the difference between scores of p and q. This notion is in fact presented in Liu, Lee and Jordan [41] in higher dimension and is used as an efficient tool to assess the proximity of the laws p and q. From formula (8), we see that if we take $K_\nu(x,y) = k_\nu(x,y) p_\nu(x)^{-1} p_\nu(y)^{-1}$, then we get the following formula, valid in dimension one,

$$\mathbb{S}_{K_\nu}(p,q) = \operatorname{Var}_\nu\left(\log\left(\frac{p}{q} \right) \right).$$

In higher dimension, Bobkov, Götze and Houdré [7] proved that the Gaussian measures satisfy a covariance identity of the same form as in (8) above, with derivatives replaced by gradients. More precisely, let (X,Y) be a pair of independent normalized Gaussian vectors in $\mathbb{R}^d$, let μ_α

be the measure of the pair $(X, \alpha X + \sqrt{1-\alpha^2}Y)$ and let $p_N(x, y)$ be the density associated the measure $\int_0^1 \mu_\alpha \, d\alpha$. Then we have

$$\mathrm{Cov}\big(g(X), h(X)\big) = \iint_{\mathbb{R}^2} \nabla g(x)^T p_N(x, y) \nabla h(y) \, dx \, dy.$$

This gives that for a kernel $K_N(x, y) = p_N(x, y)\varphi^{-1}(x)\varphi^{-1}(y)$, where φ is the standard normal density on $\mathbb{R}^d$, we also have

$$\mathbb{S}_{K_N}(p, q) = \mathrm{Var}\left(\log\left(\frac{p}{q}\right)(X)\right).$$

The following formulas will also be useful. They can be seen as special instances of the previous covariance representation formula.

Corollary 3 (Corollary 2.1, Saumard and Wellner [50]). *For an absolutely continuous function $h \in L_1(F)$,*

$$F(z) \int_{\mathbb{R}} h \, d\nu - \int_{-\infty}^{z} h \, d\nu = \int_{\mathbb{R}} k_\nu(z, y) h'(y) \, dy \tag{9}$$

and

$$-\big(1 - F(z)\big) \int_{\mathbb{R}} h \, d\nu + \int_{(z,\infty)} h \, d\nu = \int_{\mathbb{R}} k_\nu(z, y) h'(y) \, dy. \tag{10}$$

By combining Theorem 1 and Corollary 3, we get the following covariance identity.

Proposition 4. *Let ν be a probability measure on $\mathbb{R}$ and $p, q \geq 1$ such that $p^{-1} + q^{-1} = 1$. Denote $\mathcal{L}h(x) = \int_x^\infty h \, d\nu - (1 - F(x)) \int_{\mathbb{R}} h \, d\nu = F(x) \int_{\mathbb{R}} h \, d\nu - \int_{-\infty}^{x} h \, d\nu$ for every $x \in \mathbb{R}$. If $g \in L_p(\nu)$ and $h \in L_q(\nu)$ are absolutely continuous and if $g' \mathcal{L}h$ is integrable with respect to the Lebesgue measure, then*

$$\mathrm{Cov}(g, h) = \int_{\mathbb{R}} g'(x) \mathcal{L}h(x) \, dx. \tag{11}$$

Furthermore, if ν has a density p with respect to the Lebesgue measure that has a connected support, then

$$\mathrm{Cov}(g, h) = \int_{\mathbb{R}} g'(x) \bar{\mathcal{L}}h(x) p(x) \, dx = \mathbb{E}\big[g'(X) \bar{\mathcal{L}}h(X)\big], \tag{12}$$

where, for every $x \in I(\nu)$,

$$\bar{\mathcal{L}}h(x) = p(x)^{-1} \mathcal{L}h(x) = \frac{1}{p(x)} \int_x^\infty h \, d\nu - \frac{1 - F(x)}{p(x)} \mathbb{E}[h]. \tag{13}$$

If $x \notin I(\nu)$, we take $\bar{\mathcal{L}}h(x) = 0$.

Proof. Identity (11) consists in applying Fubini theorem in the formula of Theorem 1 and then using Corollary 3. If ν has a density p with respect to the Lebesgue measure that has a connected support, then for every $x \notin I(\nu)$ we have $\mathcal{L}h(x) = 0$. Consequently, from Identity (11) we get, for $g \in L_\infty(\nu)$, $h \in L_1(\nu)$ absolutely continuous,

$$\operatorname{Cov}(g,h) = \int_{I(\nu)} g'(x)\mathcal{L}h(x)\,dx = \int_{I(\nu)} g'(x)\bar{\mathcal{L}}h(x)p(x)\,dx$$

and so Identity (12) is proved. □

From Proposition 4, we can directly recover formula (6) for the Stein kernel, when it is assumed that the measure has a connected support and finite first moment. Indeed, by taking $h(x) = x - \mu$, we have h ν-integrable and differentiable and so, for any absolutely continuous function $g \in L_\infty(\nu)$ such that $g'\bar{\mathcal{L}}h$ is ν-integrable, applying Identity (12) – since $g'\mathcal{L}h = g'\bar{\mathcal{L}}hp$ a.s. is Lebesgue integrable – yields

$$\operatorname{Cov}(g,h) = \int_{\mathbb{R}} (x-\mu)g(x)p(x)\,dx = \int_{\mathbb{R}} g'\bar{\mathcal{L}}h\,d\nu. \tag{14}$$

As by a standard approximation argument, identity (14) can be extended to any g such that the functions $x \mapsto (x-\mu)g(x)$ and $x \mapsto \tau_\nu(x)g'(x)$ are ν-integrable and $[\tau_\nu pg]_a^b = 0$, we deduce that a version of the Stein kernel τ_ν is given by $\bar{\mathcal{L}}h$, which is nothing but the right-hand side of Identity (6).

Following the nice recent survey Ley, Reinert and Swan [40] related to the Stein method in dimension one, identity (12) is exactly the so-called "generalized Stein covariance identity", written in terms of the inverse of the Stein operator rather than the Stein operator itself. Indeed, it is easy to see that the inverse $\mathcal{T}_\nu$ of the operator $\bar{\mathcal{L}}$ acting on integrable functions with mean zero is given by the following formula

$$\mathcal{T}_\nu f = \frac{(fp)'}{p}\mathbf{1}_{I(\nu)},$$

which is exactly the Stein operator (see Definition 2.1 of Ley, Reinert and Swan [40]).

It is also well known, see again Ley, Reinert and Swan [40], that the inverse of the Stein operator, that is $\bar{\mathcal{L}}$, is highly involved in deriving bounds for distances between distributions. From Corollary 3, we have the following seemingly new formula for this important quantity,

$$\mathcal{T}_\nu^{-1}h(x) = \bar{\mathcal{L}}h(x) = \frac{1}{p(x)}\int_{\mathbb{R}} k_\nu(x,y)h'(y)\,dy. \tag{15}$$

Particularizing the latter identity with $h(x) = x - \mu$, we obtain the following identity for the Stein kernel,

$$\tau_\nu(x) = \frac{1}{p(x)}\int_{\mathbb{R}} k_\nu(x,y)\,dy. \tag{16}$$

A consequence of (16) that will be important in Section 3 when deriving weighted functional inequalities is that for any $x \in I(\nu)$ the function $y \mapsto k_\nu(x,y)(p(x)\tau_\nu(x))^{-1}$ can be seen as the density – with respect to the Lebesgue measure – of a probability measure, since it is nonnegative and integrates to one.

We also deduce from (15) the following upper bound,

$$\left|\mathcal{T}_\nu^{-1} h(x)\right| \leq \frac{\|h'\|_\infty}{p(x)} \int_{\mathbb{R}} k_\nu(x,y)\,dy = \frac{\|h'\|_\infty}{p(x)}\left(F(x)\int_{\mathbb{R}} x\,d\nu(x) - \int_{-\infty}^{x} x\,d\nu(x)\right),$$

which is exactly the formula given in Proposition 3.13(a) of Döbler [27].

Let us note $\varphi(x) = -\log p(x)$ when $p(x) > 0$ and $+\infty$ otherwise, the so-called potential of the density p. If on $I(\nu)$, φ has derivative $\varphi' \in L_1(\nu)$ absolutely continuous, then Corollary 2.3 in Saumard and Wellner [50] gives

$$\int_{\mathbb{R}} k_\nu(x,y)\varphi''(y)\,dy = p(x).$$

Using the latter identity together with (15), we deduce the following upper-bound: if p is strictly log-concave (that is $\varphi'' > 0$ on $I(\nu)$), then

$$\sup_{x\in I(\nu)} \left|\mathcal{T}_\nu^{-1} h(x)\right| \leq \sup_{x\in I(\nu)} \frac{|h'(x)|}{\varphi''(x)}. \tag{17}$$

In particular, if p is c-strongly log-concave, meaning that $\varphi'' \geq c > 0$ on $\mathbb{R}$, then the Stein kernel is uniformly bounded and $\|\tau_\nu\|_\infty \leq c^{-1}$. For more about the Stein method related to (strongly) log-concave measures, see, for instance, Mackey and Gorham [42].

Furthermore, by differentiating (15), we obtain for any $x \in I(\nu)$,

$$\begin{aligned}\left(\mathcal{T}_\nu^{-1} h\right)'(x) &= \varphi'(x)\mathcal{T}_\nu^{-1} h(x) - h(x) - \int_{\mathbb{R}} F(y)h'(y)\,dy \\ &= \varphi'(x)\mathcal{T}_\nu^{-1} h(x) - h(x) + \mathbb{E}\big[h(X)\big],\end{aligned}$$

that is

$$\left(\mathcal{T}_\nu^{-1} h\right)'(x) - \varphi'(x)\mathcal{T}_\nu^{-1} h(x) = -h(x) + \mathbb{E}\big[h(X)\big].$$

This is nothing but the so-called Stein equation associated to the Stein operator.

We conclude this section with the following formulas, that are available when considering a density with connected support and that will be useful in the rest of the paper (see in particular Sections 3.2 and 5).

Proposition 5. *Assume that X is a random variable with distribution ν having a density p with connected support with respect to the Lebesgue measure on $\mathbb{R}$. Take $h \in L_1(\nu)$ with $\mathbb{E}[h(X)] = 0$ and assume that the function $\bar{\mathcal{L}}h$ defined in (13) is ν-almost surely strictly positive. We have, for any $x_0, x \in I(\nu)$,*

$$p(x) = \frac{\mathbb{E}[h(X)1_{\{X\geq x_0\}}]}{\bar{\mathcal{L}}h(x)} \exp\left(-\int_{x_0}^{x} \frac{h(y)}{\bar{\mathcal{L}}h(y)}\,dy\right). \tag{18}$$

Consequently, if X has a finite first moment, for any $x \in I(\nu)$,

$$p(x) = \frac{\mathbb{E}[|X-\mu|]}{2\tau_\nu(x)} \exp\left(-\int_\mu^x \frac{y-\mu}{\tau_\nu(y)}\,dy\right). \tag{19}$$

By setting $T_h(x) = \exp(-\int_{x_0}^x \frac{h(y)}{\bar{\mathcal{L}}h(y)}\,dy)$ and $I(\nu) = (a,b)$, if the function h is ν-almost positive, differentiable on (x,b) and if the ratio $T_h(y)/h(y)$ tends to zero when y tends to b^-, then we have, for any $x_0, x \in I(\nu)$,

$$\mathbb{P}(X \geq x) = \mathbb{E}\big[h(X)1_{\{X\geq x_0\}}\big]\left(\frac{T_h(x)}{h(x)} - \int_x^b \frac{h'(y)}{h^2(y)} T_h(y)\,dy\right). \tag{20}$$

Formula (18) can also be found in Döbler [27], Equation (3.11), under the assumption that h is decreasing and for a special choice of x_0. Since $\mathbb{E}[h(X)] = 0$, it is easily seen through its definition (13), that if $h \neq 0$ ν-a.s. then $\bar{\mathcal{L}}h > 0$ ν-a.s. When $h = \mathrm{Id} - \mu$, formulas (19) and (20) were first proved respectively in Nourdin and Viens [47] and Viens [55], although with assumption that the random variable X belongs to the space $\mathbf{D}^{1,2}$ of square integrable random variables with the natural Hilbert norm of their Malliavin derivative also square integrable.

In order to take advantage of formulas (18) and (20), one has to use some information about $\bar{\mathcal{L}}h$. The most common choice is $h = \mathrm{Id} - \mu$, which corresponds to the Stein kernel $\bar{\mathcal{L}}(\mathrm{Id} - \mu) = \tau_\nu$.

Proof. Begin with Identity (18). As $x_0 \in I(\nu)$ and the function $\mathcal{L}h$ defined in (4) is ν-almost surely positive, we have for any $x \in I(\upsilon)$,

$$\mathcal{L}h(x) = \mathcal{L}h(x_0)\exp\left(\int_{x_0}^x \big(\ln(\mathcal{L}h)\big)'(y)\,dy\right).$$

To conclude, note that $\mathcal{L}h(x_0) = \mathbb{E}[h(X)1_{\{X\geq x_0\}}]$ and $(\ln(\mathcal{L}h))' = -h/\bar{\mathcal{L}}h$. To prove (19), simply remark that is follows from (18) by taking $h = \mathrm{Id} - \mu$ and $x_0 = \mu$.

It remains to prove (20). We have from (18), $p = \mathbb{E}[h(X)1_{\{X\geq x_0\}}]T_h/\bar{\mathcal{L}}h$ and by definition of T_h, $T_h' = -hT_h/\bar{\mathcal{L}}h$. Hence, integrating between x and b gives

$$\begin{aligned}
\mathbb{P}(X \geq x) &= \mathbb{E}\big[h(X)1_{\{X\geq x_0\}}\big]\int_x^b \frac{T_h(y)}{\bar{\mathcal{L}}h(y)}\,dy \\
&= \mathbb{E}\big[h(X)1_{\{X\geq x_0\}}\big]\int_x^b \frac{-T_h'(y)}{h(y)}\,dy \\
&= \mathbb{E}\big[h(X)1_{\{X\geq x_0\}}\big]\left(\left[\frac{-T_h}{h}\right]_x^b - \int_x^b \frac{h'(y)T_h(y)}{h^2(y)}\,dy\right) \\
&= \mathbb{E}\big[h(X)1_{\{X\geq x_0\}}\big]\left(\frac{T_h(x)}{h(x)} - \int_x^b \frac{h'(y)T_h(y)}{h^2(y)}\,dy\right).
\end{aligned}$$

□

3. Some weighted functional inequalities

Weighted functional inequalities appear naturally when generalizing Gaussian functional inequalities such as Poincaré and log-Sobolev inequalities. They were put to emphasis for the generalized Cauchy distribution and more general κ-concave distributions by Bobkov and Ledoux [10,11], also in connection with isoperimetric-type problems, weighted Cheeger-type inequalities and concentration of measure. Then several authors proved related weighted functional inequalities (Bakry, Cattiaux and Guillin [2], Bonnefont and Joulin [13], Bonnefont, Joulin and Ma [14,15], Cattiaux et al. [18], Cattiaux, Guillin and Wu [17], Cordero-Erausquin and Gozlan [25], Gozlan [36]). In the following, we show the strong connection between Stein kernels and the existence of weighted functional inequalities. Note that a remarkable first result in this direction was recently established by Courtade, Fathi and Pananjady [26] who proved that a reversed weighted Poincaré inequality is sufficient to ensure the existence of a Stein kernel in $\mathbb{R}^d$, $d \geq 1$.

Our results are derived in dimension one. Indeed, the proofs of weighted Poincaré inequalities (Section 3.1) and asymmetric Brascamp–Lieb type inequalities (Section 1.3, Saumard [49]) rely on covariance identities stated in Section 2, that are based on properties of cumulative distribution functions available in dimension one. Recently, Fathi [30] generalized in higher dimension the weighted Poincaré inequalities derived in Section 3.1 using transportation techniques. We also derive in Section 3.2 some weighted log-Sobolev inequalities, that are derived through the use of some Muckenhoupt-type criteria that are only valid in dimension one. However, it is natural to conjecture a multi-dimensional generalization, especially using tools developed in Fathi [30], but this remains an open question. Finally, using a formula for the isoperimetric constant in dimension one due to Bobkov and Houdré [8], we prove that a uniformly bounded Stein kernel is essentially sufficient to ensure a positive isoperimetric constant. Again, the problem in higher dimension seems much more involved and is left as an interesting open question.

3.1. Weighted Poincaré-type inequality

According to Bobkov and Ledoux [11], a measure ν on $\mathbb{R}$ is said to satisfy a weighted Poincaré inequality if there exists a nonnegative measurable weight function ω such that for any smooth function $f \in L_2(\nu)$,

$$\operatorname{Var}\big(f(X)\big) \leq \mathbb{E}\big[\omega(X)\big(f'(X)\big)^2\big]. \tag{21}$$

The following theorem shows that a probability measure having a finite first moment and density with connected support on the real line satisfies a weighted Poincaré inequality, with the weight being its Stein kernel.

Theorem 6. *Take a real random variable X of distribution ν with density p with respect to the Lebesgue measure on $\mathbb{R}$. Assume that $\mathbb{E}[|X|] < +\infty$, p has a connected support and denote τ_ν the Stein kernel of ν. Take $f \in L_2(\nu)$, absolutely continuous. Then*

$$\operatorname{Var}\big(f(X)\big) \leq \mathbb{E}\big[\tau_\nu(X)\big(f'(X)\big)^2\big]. \tag{22}$$

The preceding inequality is optimal whenever ν *admits a finite second moment, that is* $\mathbb{E}[X^2] < +\infty$, *since equality is reached for* $f = \mathrm{Id}$, *by definition of the Stein kernel.*

Proof. We have

$$\begin{aligned}
\operatorname{Var}\big(f(X)\big) &= \mathbb{E}\left[\sqrt{\tau_\nu(X)} f'(X) \frac{\bar{\mathcal{L}} f(X)}{\sqrt{\tau_\nu(X)}}\right] \\
&\le \sqrt{\mathbb{E}\big[\tau_\nu(X)\big(f'(X)\big)^2\big]} \sqrt{\mathbb{E}\left[\tau_\nu(X)\left(\frac{\bar{\mathcal{L}} f(X)}{\tau_\nu(X)}\right)^2\right]}.
\end{aligned}$$

By the use of Jensen's inequality, for any $x \in I(\nu)$,

$$\begin{aligned}
\left(\frac{\bar{\mathcal{L}} f(x)}{\tau_\nu(x)}\right)^2 &= \left(\int f'(y) \frac{k_\nu(x,y)}{\int k_\nu(x,z)\,dz}\,dy\right)^2 \\
&\le \int \big(f'(y)\big)^2 \frac{k_\nu(x,y)}{\int_z k_\nu(x,z)\,dz}\,dy \\
&= \int \frac{k_\nu(x,y)}{\tau_\nu(x)p(x)} \big(f'(y)\big)^2\,dy.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\tau_\nu(X)\left(\frac{\bar{\mathcal{L}} f(X)}{\tau_\nu(X)}\right)^2\right] &\le \int \tau_\nu(x)p(x)\left(\int \frac{k_\nu(x,y)}{\tau_\nu(x)p(x)}\big(f'(y)\big)^2\,dy\right)dx \\
&= \iint k_\nu(x,y)\big(f'(y)\big)^2\,dx\,dy \\
&= \int \tau_\nu(y)\big(f'(y)\big)^2 p(y)\,dy,
\end{aligned}$$

which concludes the proof. □

It is worth mentioning that the famous Brascamp–Lieb inequality provides another weighted Poincaré inequality in dimension one: if ν is strictly log-concave, of density $p = \exp(-\varphi)$ with a smooth potential φ, then for any smooth function $f \in L_2(\nu)$,

$$\operatorname{Var}\big(f(X)\big) \le \mathbb{E}\big[\big(\varphi''(X)\big)^{-1}\big(f'(X)\big)^2\big]. \tag{23}$$

In particular, if ν is strongly log-concave (that is $\varphi'' \ge c > 0$ for some constant $c > 0$), then both the Brascamp–Lieb inequality (23) and inequality (22) – combined with the estimate $\|\tau_\nu\|_\infty \le c^{-1}$ coming from inequality (17) – imply the Poincaré inequality $\operatorname{Var}(f(X)) \le c^{-1}\mathbb{E}[(f'(X))^2]$, that also follows from the Bakry–Émery criterion. However, in general, the Stein kernel appearing in (22) may behave differently from the inverse of the second derivative of the potential and there is no general ordering between the right-hand sides of inequalities (22) and (23).

Indeed, let us discuss the situation for a classical class of examples in functional inequalities, namely the class of Subbotin densities $p_\alpha(x) = Z_\alpha^{-1} \exp(-|x|^\alpha/\alpha)$ for $x \in \mathbb{R}$, where $Z_\alpha > 0$ is the normalizing constant. Recall that densities p_α are not strongly log-concave and do not satisfy the Bakry–Émery criterion – except for $\alpha = 2$ which corresponds to the normal density – but they satisfy a Poincaré inequality if and only if $\alpha \geq 1$ and a log-Sobolev inequality if and only if $\alpha \geq 2$ (see Latała and Oleszkiewicz [37] and also Bonnefont and Joulin [13] for a thorough discussion on optimal constants in these inequalities). We restrict our discussion to the condition $\alpha > 1$ for which p_α is strictly log-concave, so that the Brascamp–Lieb inequality applies. More precisely, by writing $p_\alpha = \exp(-\varphi_\alpha)$, we have $(\varphi'')^{-1}(x) = (\alpha - 1)^{-1}|x|^{2-\alpha}$. Using the explicit formula (16) for the Stein kernel and an integration by parts, one can easily check that if $\alpha \in (1, 2)$, then $\tau_\alpha(x) < |x|^{2-\alpha}$ where τ_α is the Stein kernel associated to p_α. We thus get, for $\alpha \in (1, 2)$, $(\varphi'')^{-1}(x) > \tau_\alpha(x)$ for any $x \in \mathbb{R}$ and so the Brascamp–Lieb inequality is less accurate than Theorem 6 in this case. Furthermore, if $\alpha > 2$ then $\tau_\alpha(x) > |x|^{2-\alpha}$ and so $(\varphi'')^{-1}(x) < \tau_\alpha(x)$ for any $x \in \mathbb{R}$, which means that the Brascamp–Lieb inequality is more accurate than Theorem 6 for $\alpha > 2$. However, we can not recover through the use of Theorem 6 – or the Brascamp–Lieb inequality – the existence of a spectral gap for $\alpha \in [1, 2)$. Finally, Theorem 6 gives us the existence of some weighted Poincaré inequalities for any $\alpha > 0$, whereas Brascamp–Lieb inequality only applies for $\alpha > 1$.

It is also worth mentioning that Theorem 6 has been recently generalized to higher dimension by Fathi [30], for a Stein kernel that is a positive definite matrix and that is defined through the use of a so-called moment map. The proof of this (non-trivial) extension of our result is actually based on the Brascamp–Lieb inequality itself, but applied to a measure also defined through the moment map problem.

Let us now detail some classical examples falling into the setting of Theorem 6.

The beta distribution $B_{\alpha,\beta}$, $\alpha, \beta > 0$ is supported on $(0, 1)$, with density $p_{\alpha,\beta}$ given by

$$p_{\alpha,\beta}(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < x < 1. \tag{24}$$

The normalizing constant $B(\alpha, \beta)$ is the classical beta function of two variables. The beta distribution has been for instance recently studied in Bobkov and Ledoux [12] in connection with the analysis of the rates of convergence of the empirical measure on $\mathbb{R}$ for some Kantorovich transport distances. The Stein kernel $\tau_{\alpha,\beta}$ associated to the Beta distribution is given by $\tau_{\alpha,\beta}(x) = (\alpha + \beta)^{-1}x(1 - x)$ for $x \in (0, 1)$ (see, for instance, Ley, Reinert and Swan [40]) and thus Theorem 6 allows to exactly recover Proposition B.5 of Bobkov and Ledoux [12] (which is optimal for linear functions as noticed in Bobkov and Ledoux [12]). Our techniques are noticeably different since the weighted Poincaré inequality is proved in Bobkov and Ledoux [12] by using orthogonal (Jacobi) polynomials. Notice also that the beta density $p_{\alpha,\beta}$ is strictly log-concave on the interior of its support. Indeed, by writing $p_{\alpha,\beta} = \exp(-\varphi_{\alpha,\beta})$, we get, for any $x \in (0, 1)$,

$$\left(\varphi''\right)^{-1}(x) = \frac{x^2(1-x)^2}{(\alpha - 1)(1 - x)^2 + (\beta - 1)x^2}.$$

Remark that, for instance, $(\varphi'')^{-1}(x) \sim_{x\to 0^+} x^2/(\alpha - 1)$, whereas $\tau_{\alpha,\beta}(x) \sim_{x\to 0^+} x/(\alpha + \beta)$, so the weights in the Brascamp–Lieb inequality (23) and in inequality (22) are not of the same order at 0 (or by symmetry at 1), although it is also easy to show that there exists a constant $c_{\alpha,\beta} > 0$ – $c_{\alpha,\beta} = (\alpha+\beta)/(2\sqrt{(\alpha-1)(\beta-1)})$ works – such that $(\varphi'')^{-1}(x) \leq c_{\alpha,\beta}\tau_{\alpha,\beta}(x)$ for any $x \in (0,1)$.

Note that considering Laguerre polynomials, that are eigenfunctions of the Laguerre operator for which the Gamma distribution is invariant and reversible, one can also show an optimal weighted Poincaré inequality for the Gamma distribution, which include as a special instance the exponential distribution (see Bobkov and Ledoux [9] and also Bakry, Gentil and Ledoux [3], Section 2.7). Theorem 6 also gives an optimal weighted Poincaré inequality for the Gamma distribution and more generally for Pearson's class of distributions (see below).

Note also that the beta distribution seems to be outside of the scope of the weighted Poincaré inequalities described in Bonnefont, Joulin and Ma [14] since it is assumed in the latter article that the weight of the considered Poincaré-type inequalities is positive on $\mathbb{R}$, which is not the case for the beta distribution. Furthermore, Bobkov and Ledoux [12] also provides some weighted Cheeger inequality for the Beta distribution, but such a result seems outside the scope of our approach based on covariance identity (8). When considering concentration properties of beta distributions in Section 4 below, we will however provide some improvements compared to the results of Bobkov and Ledoux [12].

Furthermore, it has also been noticed that the generalized Cauchy distribution satisfies a weighted Poincaré distribution, which also implies in this case a reverse weighted Poincaré inequality (see Bobkov and Ledoux [11], Bonnefont, Joulin and Ma [14]). In fact, Bobkov and Ledoux [11] shows that the generalized Cauchy distribution plays a central role when considering functional inequalities for κ-concave measures, with $\kappa < 0$.

The generalized Cauchy distribution ν_β of parameter $\beta > 1/2$ has density $p_\beta(x) = Z_\beta^{-1}(1+x^2)^{-\beta}$ for $x \in \mathbb{R}$ and normalizing constant $Z_\beta > 0$. Its Stein kernel τ_β exists for $\beta > 1$ and writes $\tau_\beta(x) = (1+x^2)/(2(\beta-1))$. This allows us to recover in the case where $\beta > 3/2$ – that is ν_β has a finite second moment – the optimal weighted Poincaré inequality also derived in Bonnefont, Joulin and Ma [14], Theorem 3.1. Note that Theorem 3.1 of Bonnefont, Joulin and Ma [14] also provides the optimal constant in the weighted Poincaré inequality with a weight proportional to $1+x^2$ in the range $\beta \in (1/2, 3/2]$.

Let us conclude this short list of examples by mentioning Pearson's class of distributions, for which the density p is solution to the following differential equation,

$$\frac{p'(x)}{p(x)} = \frac{\alpha - x}{\beta_2(x-\lambda)^2 + \beta_1(x-\lambda) + \beta_0}, \tag{25}$$

for some constants λ, α, β_j, $j = 0,1,2$. This class of distributions, that contains for instance Gaussian, Gamma, Beta and Student distributions, has been well studied in the context of Stein's method, see Ley, Reinert and Swan [40] and references therein. In particular, if a density satisfies (25) with $\beta_2 \neq 1/2$, then the corresponding distribution ν has a Stein kernel $\tau_\nu(x) = (1-2\beta_2)^{-1}(\beta_0 + \beta_1 x + \beta_2 x^2)$, for any $x \in I(\nu)$. Particularizing to the Student distribution t_α with density p_α proportional to $(\alpha + x^2)^{-(1+\alpha)/2}$ on $\mathbb{R}$ for $\alpha > 1$, we get that for any smooth function $f \in L_2(t_\alpha)$,

$$\operatorname{Var}_{t_\alpha}(f) \leq \frac{1}{\alpha-1}\int \left(x^2+\alpha\right) f'^2(x)\, dt_\alpha(x).$$

Some further comments related to some concentration inequalities that can be proved from weighted Poincaré inequalities and to links between converse weighted Poincaré inequalities and existence of a Stein kernel can be found in Saumard [49], Section 1.1.

3.2. Links with Muckenhoupt-type criteria

It is well known that the Muckenhoupt criterion (Muckenhoupt [46]), which provides a necessary and sufficient condition for a (weighted) Hardy inequalty to hold on the real line, can be used to sharply estimate the best constant in Poincaré inequalities (see, for instance, Ané et al. [1] and Miclo [45] and references therein). The following theorem, providing sharp estimates, is given in Miclo [45].

Theorem 7 (Miclo [45]). *Let η be a probability measure on $\mathbb{R}$ with median m and let χ be a measure on $\mathbb{R}$ with Radon–Nikodym derivative with respect to the Lebesgue measure denoted by n. The best constant C_P such that, for every locally Lipschitz $f:\mathbb{R}\to\mathbb{R}$ it holds*

$$\mathrm{Var}_\eta(f)\le C_P\int\left(f'\right)^2 d\chi,$$

verifies $\max\{B_+,B_-\}\le C_P\le 4\max\{B_+,B_-\}$ *where*

$$B_+=\sup_{x>m}\eta\bigl([x,+\infty)\bigr)\int_m^x\frac{dt}{n(t)}\quad \text{and}\quad B_-=\sup_{x<m}\eta\bigl((-\infty,x]\bigr)\int_x^m\frac{dt}{n(t)}.$$

Considering Theorem 6, a natural question is: can we recover (up to a constant) Inequality (22) from Theorem 7 above? The answer is positive. Indeed, we want to show that $\max\{B_+,B_-\}$ is finite. We will only discuss computations for B_+ since B_- can be treated symmetrically. With the notations of Theorems 6 and 7, we take $\eta=\nu$ and $n(t)=\tau_\nu(t)p(t)$. This gives

$$B_+=\sup_{x>m}\nu\bigl([x,+\infty)\bigr)\int_m^x\frac{dt}{\tau_\nu(t)p(t)}=\sup_{x>m}\nu\bigl([x,+\infty)\bigr)\int_m^x\frac{dt}{\int_t^\infty(y-\mu)p(y)\,dy}.$$

Furthermore, there exists $\delta>0$ such that $\nu([x_0,+\infty))>0$ for $x_0=\max\{m,\mu\}+\delta$. Hence, for any $x\ge x_0$,

$$\nu\bigl([x,+\infty)\bigr)\int_m^x\frac{dt}{\int_t^\infty(y-\mu)p(y)\,dy}\le\frac{\nu([x,+\infty))}{\int_x^\infty(y-\mu)p(y)\,dy}\le\frac{1}{\delta}.$$

As

$$\sup_{x\in(m,x_0)}\nu\bigl([x,+\infty)\bigr)\int_m^x\frac{dt}{\int_t^\infty(y-\mu)p(y)\,dy}\le\frac{x_0-m}{\int_m^\infty(y-\mu)p(y)\,dy\wedge\int_{x_0}^\infty(y-\mu)p(y)\,dy}<+\infty,$$

we get $B_+<+\infty$. Consequently, we quickly recover under the assumptions of Theorem 6 the fact that the measure ν satisfies a weighted Poincaré inequality of the form of (22), although with a multiplicative constant at the right-hand side that a priori depends on the measure ν. Using (22) together with Theorem 7, we have in fact $\max\{B_+, B_-\} \le 1$, but we couldn't achieve this bound – even up to a numerical constant – by direct computations. It is maybe worth mentioning that the proof of Theorem 7 in Miclo [45] is technically involved and that the bound $\max\{B_+, B_-\} \le 1$ in our case might be rather difficult to establish by direct computations.

Let us turn now to the important, natural question of the existence of weighted log-Sobolev inequalities under the existence of a Stein kernel.

Theorem 8. *Take a real random variable X of distribution ν with density p with respect to the Lebesgue measure on $\mathbb{R}$. Assume that $\mathbb{E}[|X|] < +\infty$, p has a connected support $[a, b] \subset \bar{\mathbb{R}}$ and denote τ_ν the Stein kernel of ν. Take g absolutely continuous. Then the following inequality holds*

$$\mathrm{Ent}_\nu\left(g^2\right) \le C_\nu \int \tau_\nu^2 \left(g'\right)^2 d\nu, \tag{26}$$

for some constant $C_\nu > 0$ if one of the following aymptotic condition holds at the supremum of its support, together with one of the symmetric conditions – that we don't write explicitely since they are obviously deduced – at the infinimum of its support:

- *$b < +\infty$ and τ_ν^{-1} is integrable at b^- with respect to the Lebesgue measure.*
- *$b = +\infty$ and $0 < c_- \le \tau_\nu(x) \le c_+ x^2 / \log x$ for some constants c_- and c_+ and for x large enough.*
- *$b = +\infty$ and $0 < c_- x \le \tau_\nu(x)$ for a constant c_- and for x large enough.*

Reciprocally, if $b = +\infty$ and $\tau_\nu \to_{x\to\pm\infty} 0$ then inequality (26) can not be satisfied for every smooth function g.

Theorem 8 gives a sufficient condition for the weighted log-Sobolev inequality (26) to hold when the support is a bounded interval: it suffices that the inverse of the Stein kernel is integrable at the edges of the support. Furthermore, when an edge is infinite, if the weighted log-Sobolev inequality (26) is valid, then the Stein kernel does not tend to zero around this edge.

The proof of Theorem 8, which is detailed in Saumard [49], Section 1.2, is based on a Muckenhoupt-type criterion due to Bobkov and Götze [6] (see also Barthe and Roberto [5] for a refinement) giving a necessary and sufficient condition for existence of (weighted) log-Sobolev inequalities on $\mathbb{R}$, together with the use of the formulas given in Propostion 5 above.

Consider the Subbotin densities, defined by $p_\alpha(x) = Z_\alpha^{-1} \exp(-|x|^\alpha/\alpha)$ for $\alpha > 0$ and $x \in \mathbb{R}$. There Stein kernels τ_α satisfy $\tau_\alpha(x) \sim_{x\to+\infty} x^{2-\alpha}$ (see the discussion in Section 3.1 above), so they achieve a weighted log-Sobolev inequality (26) if and only if $\alpha \in (0, 2]$. In the case where $\alpha \in [2, +\infty)$, they actually achieve a (unweighted) log-Sobolev inequality (Latała and Oleszkiewicz [37]).

It is reasonable to think that inequality (26) should be true under suitable conditions in higher dimension, with the right-hand side replaced by $\int |\tau_\nu \nabla f|^2 d\nu$ up to a constant, where τ_ν is the Stein kernel constructed by Fathi [30] using moment maps. However, some technical details remain to be solved at this point of our investigations.

Furthermore, notice that Theorem 8 allows to recover in dimension one and with a worst constant, a result due to Bobkov and Ledoux [10] in dimension $d \geq 1$, stating that generalized Cauchy distributions (see Section 3.1 above for a definition) verify the following weighted log-Sobolev inequality,

$$\mathrm{Ent}_{\nu_\beta}\left(g^2\right) \leq \frac{1}{\beta - 1} \int \left(1 + x^2\right)^2 \left(g'(x)\right)^2 d\nu_\beta(x), \quad \beta > 1. \tag{27}$$

Using the Muckenhoupt-type criterion due to Bobkov and Götze [6], we can actually sharpen and extend the previous inequality in the following way. There exists a constant C_β such that, for any smooth function g,

$$\mathrm{Ent}_{\nu_\beta}\left(g^2\right) \leq C_\beta \int \left(1 + x^2\right) \log\left(1 + x^2\right) \left(g'(x)\right)^2 d\nu_\beta(x), \quad \beta > 1/2. \tag{28}$$

Indeed, using the quantities defined in Saumard [49], Eq. (4), with $d\eta = d\nu_\beta = Z_\beta^{-1}(1 + x^2)^{-\beta} \, dx$ and $n(t) = Z_\beta^{-1}(1 + x^2)^{-\beta+1} \log(1 + x^2)$, we note that $m = 0$, $L_+ = L_-$ and by simple computations, we have for $\beta > 1/2$,

$$\int_m^x \frac{dt}{n(t)} = O_{x \to +\infty}\left(\frac{x^{2\beta-1}}{\log x}\right), \qquad \Lambda\left(\nu_\beta\left([x, +\infty)\right)\right) = O_{x \to +\infty}\left(x^{-2\beta+1} \log x\right).$$

This means that L_+ is finite and so, Inequality (28) is valid, which constitutes an improvement upon inequality (27). A natural question that remains open, is wether the weighted log-Sobolev inequality obtained for generalized Cauchy measures in Bobkov and Ledoux [10] can be also improved in dimension $d \geq 2$?

3.3. Isoperimetric constant

Let us complete this section about some functional inequalities linked to the Stein kernel by studying the isoperimetric constant. Recall that for a measure ν on $\mathbb{R}^d$, an isoperimetric inequality is an inequality of the form

$$\nu^+(A) \geq c \min\left\{\nu(A), 1 - \nu(A)\right\}, \tag{29}$$

where $c > 0$, A is an arbitrary measurable set in $\mathbb{R}^d$ and $\nu^+(A)$ stands for the ν-perimeter of A, defined to be

$$\nu^+(A) = \lim \inf_{r \to 0^+} \frac{\nu(A^r) - \nu(A)}{r},$$

with $A^r = \{x \in \mathbb{R}^d : \exists a \in A, |x - a| < r\}$ the r-neighborhood of A. The optimal value of $c = Is(\nu)$ in (29) is referred to as the isoperimetric constant of ν.

The next proposition shows that existence of a uniformly bounded Stein kernel is essentially sufficient for guaranteeing existence of a positive isoperimetric constant.

Proposition 9. *Assume that the probability measure ν has a connected support, finite first moment and continuous density p with respect to the Lebesgue measure. Assume also that its Stein kernel τ_ν is uniformly bounded on $I(\nu)$, $\|\tau_\nu\|_\infty < +\infty$. Then ν admits a positive isoperimetric constant $Is(\nu) > 0$.*

More precisely, from the proof of Proposition 9 (see below), we can extract a quantitative estimate of the isoperimetric constant. Indeed, under the assumptions of Proposition 9, denote by q_β the quantile of order $\beta \in (0, 1)$ of the measure ν and by μ its mean. Then, for any $\alpha \in (0, 1)$ such that $q_\alpha < \mu < q_{1-\alpha}$, we have

$$Is(\nu) \geq \min\left\{\alpha^{-1} \min_{x\in[q_\alpha, q_{1-\alpha}]}\{p(x)\}, \frac{\min\{\mu - q_\alpha, q_{1-\alpha} - \mu\}}{\|\tau_\nu\|_\infty}\right\}.$$

Measures having a uniformly bounded Stein kernel include strongly log-concave measures – as proved in Section 2 above, – but also smooth perturbations of the normal distribution that are Lipschitz and bounded from below by a positive constant (see Remark 2.9 in Ledoux, Nourdin and Peccati [38]). In addition, bounded perturbations of measures having a bounded Stein kernel given by formula (6) and a density that is bounded away from zero around its mean also have a bounded Stein kernel. Indeed, take $\varkappa$ a measure having a density $p_\varkappa$ with connected support $[a, b] \subset \bar{\mathbb{R}}$, $a < 0 < b$, mean zero and a finite first moment so that its Stein kernel $\tau_\varkappa$ is unique – up to sets of Lebesgue measure zero – and given by formula (6). Assume also that the density $p_\varkappa$ is bounded away from zero around zero, that is there exists $L, \delta > 0$ such that $[-\delta, \delta] \subset (a, b)$ and $\inf_{x\in[-\delta,\delta]} p_\varkappa(x) \geq L > 0$. Now, consider a function ρ on $\mathbb{R}$ such that $C^{-1} \leq \rho(x) \leq C$ for any $x \in (a, b)$ and for some constant $C > 0$, such that $p(x) = \rho(x)p_\varkappa(x)$ is the density of a probability measure ν with support $[a, b]$. As ν has a finite first moment, it also admits a Stein kernel τ_ν, given by formula (6). Furthermore, τ_ν is uniformly bounded and easy computations using the formula (6) give

$$\|\tau_\nu\|_\infty \leq C^2\left(\|\tau_\varkappa\|_\infty + |\mu| \max\left\{\frac{1}{L}, \frac{\|\tau_\varkappa\|_\infty}{\delta}\right\}\right),$$

where μ is the mean of ν.

It would be interesting to know if Proposition 9 also holds in higher dimension, but this question remains open. A further natural question would be: does a measure having a Stein kernel satisfy a weighted isoperimetric-type inequality, with a weight related to the Stein kernel? So far, we couldn't give an answer to this question. Note that Bobkov and Ledoux [10,11] proved some weighted Cheeger and weighted isoperimetric-type inequalities for the generalized Cauchy and for κ-concave distributions.

Proof. Let F be the cumulative distribution function of ν, μ be its mean and let $\varepsilon > 0$ be such that $[\mu - \varepsilon, \mu + \varepsilon] \subset I(\nu)$. Recall (Bobkov and Houdré [8], Theorem 1.3) that the isoperimetric constant associated to ν satisfies

$$Is(\nu) = \text{ess} \inf_{a<x<b} \frac{p(x)}{\min\{F(x), 1 - F(x)\}},$$

where $a < b$ are the edges of the support of ν. Take $x \in I(\nu)$ such that $x - \mu \geq \varepsilon/2$, then

$$\begin{aligned}\tau_\nu(x) &= \frac{1}{p(x)} \int_x^\infty (y-\mu)p(y)\,dy \\ &\geq \varepsilon \frac{1-F(x)}{2p(x)} \\ &\geq \frac{\varepsilon}{2} \frac{\min\{F(x), 1-F(x)\}}{p(x)}.\end{aligned}$$

The same estimate holds for $x \leq \mu - \varepsilon/2$ since $\tau_\nu(x) = p(x)^{-1} \int_{-\infty}^x (\mu - y)p(y)\,dy$. Hence,

$$\operatorname{ess} \inf_{x\in I(\nu), |x-\mu|\geq\varepsilon/2} \frac{p(x)}{\min\{F(x), 1-F(x)\}} \geq \frac{2}{\varepsilon \|\tau_\nu\|_\infty} > 0. \tag{30}$$

Furthermore, we have

$$\inf_{|x-\mu|\leq\varepsilon/2} \frac{p(x)}{\min\{F(x), 1-F(x)\}} \geq \frac{\inf_{|x-\mu|\leq\varepsilon/2} p(x)}{\min\{F(\mu-\varepsilon/2), 1-F(\mu+\varepsilon/2)\}} > 0. \tag{31}$$

The conclusion now follows from combining (30) and (31). □

4. Concentration inequalities

We state in this section some concentration inequalities related to Stein kernels in dimension one. Due to the use of covariance identities stated in Section 2, the proofs indeed heavily rely on dimension one. We can however notice that it is known from Ledoux, Nourdin and Peccati [38] that a uniformly bounded (multi-dimensional) Stein kernel ensures a sub-Gaussian concentration rate. We derive sharp sub-Gaussian inequalities in Theorem 11. It is also reasonable to think that results such as in Theorem 14 could be generalized to higher dimension, but this seems rather nontrivial and is left as an open question.

From Proposition 4, Section 2, we get the following proposition.

Proposition 10. *Assume that ν has a a finite first moment and a density p with respect to the Lebesgue measure that has a connected support. If $g \in L_\infty(\nu)$ and $h \in L_1(\nu)$ are absolutely continuous and h is* 1*-Lipschitz, then*

$$\left|\operatorname{Cov}(g,h)\right| \leq \mathbb{E}\left[\left|g'\right| \cdot \tau_\nu\right], \tag{32}$$

where τ_ν is given in (6) and is the Stein kernel. Furthermore, if the Stein kernel is uniformly bounded, that is $\tau_\nu \in L_\infty(\nu)$, then

$$\left|\operatorname{Cov}(g,h)\right| \leq \|\tau_\nu\|_\infty \mathbb{E}\left[\left|g'\right|\right]. \tag{33}$$

Proof. Start from identity (12) and simply remark that, for $h_\nu(x) = x$,

$$\left|\bar{\mathcal{L}}h(x)\right| = \left|\frac{1}{p(x)} \int_{\mathbb{R}} k_\nu(x,y)h'(y)\,dy\right| \le \frac{\|h'\|_\infty}{p(x)} \int_{\mathbb{R}} k_\nu(x,y)\,dy = \left\|h'\right\|_\infty \tau_\nu(x). \qquad \square$$

Applying techniques similar to those developed in Bobkov, Götze and Houdré [7] for Gaussian vectors (see especially Theorem 2.2), we have the following Gaussian type concentration inequalities when the Stein kernel is uniformly bounded.

Theorem 11. *Assume that ν has a finite first moment and a density p with respect to the Lebesgue measure that has a connected support. Assume also that the Stein kernel τ_ν is uniformly bounded, $\tau_\nu \in L_\infty(\nu)$, and denote $c = \|\tau_\nu\|_\infty^{-1}$. Then the following concentration inequalities hold. For any 1-Lipschitz function g,*

$$\mathbb{P}\left(g \ge \int g\,d\nu + r\right) \le e^{-cr^2/2}. \tag{34}$$

Furthermore, the function $T_g(r) = e^{cr^2/2}\mathbb{E}(g - \mathbb{E}g)\mathbf{1}_{\{g-\mathbb{E}g \ge r\}}$ is non-increasing in $r \ge 0$. In particular, for all $r > 0$,

$$\mathbb{P}(g - \mathbb{E}g \ge r) \le \mathbb{E}(g - \mathbb{E}g)_+ \frac{e^{-cr^2/2}}{r}, \tag{35}$$

$$\mathbb{P}\left(|g - \mathbb{E}g| \ge r\right) \le \mathbb{E}|g - \mathbb{E}g| \frac{e^{-cr^2/2}}{r}. \tag{36}$$

Inequality (34) is closely related to Chatterjee's Gaussian coupling for random variables with bounded Stein kernel Chatterjee [20]. To our knowledge, refined concentration inequalities such as (35) and (36) are only available in the literature for Gaussian random variables or by extension, for strongly log-concave measures. Indeed, these inequalities can be established for strongly log-concave measures as an immediate consequence of the Caffarelli contraction theorem, which states that such measures can be realized as the pushforward of the Gaussian measure by a Lipschitz function. We refer to these inequalities as generalized Mills' type inequalities since taking $g = \mathrm{Id}$ in Inequality (36) allows to recover Mills' inequality (see, for instance, Dümbgen [28]): if Z is the normal distribution, then for any $t > 0$,

$$\mathbb{P}\left(|Z| > t\right) \le \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}.$$

Here the setting of a bounded Stein kernel is much larger and include for instance, smooth perturbations of the normal distribution that are Lipschitz and bounded away from zero (see Remark 2.9 in Ledoux, Nourdin and Peccati [38]) or bounded perturbations of measures having a bounded Stein kernel and a density bounded away from zero around its mean (see Section 3.3 above for more details).

Proof of Theorem 11. Take g to be 1-Lipschitz and mean zero with respect to ν, then for any $\lambda \geq 0$,

$$\mathbb{E}\big[g e^{\lambda g}\big] = \mathrm{Cov}\big(g, e^{\lambda g}\big) \leq \|\tau_v\|_\infty \mathbb{E}\big[\big|\big(e^{\lambda g}\big)'\big|\big] \leq \frac{\lambda}{c}\mathbb{E}\big[e^{\lambda g}\big].$$

Define $J(\lambda) = \log \mathbb{E}[e^{\lambda g}]$, $\lambda \geq 0$. We thus have the following differential inequality, $J'(\lambda) \leq \lambda/c$. Since $J(0) = 0$, this implies that $J(\lambda) \leq \lambda^2/(2c)$. Equivalently, $\mathbb{E}[e^{\lambda g}] \leq e^{\lambda^2/(2c)}$, which by the use of Chebyshev's inequality gives (34). Now, assume that as a random variable g has a continuous positive density p on the whole real line. Take $f = U(g)$ where U is a non-decreasing (piecewise) differentiable function on $\mathbb{R}$. Applying (33), we get

$$\mathbb{E}\big[gU(g)\big] \leq \mathbb{E}\big[U'(g)\big]/c. \tag{37}$$

Let G be the distribution function of g. Given $r > 0$ and $\varepsilon > 0$, applying (37) to the function $U(x) = \min\{(x-r)^+, \varepsilon\}$ leads to

$$\int_r^{r+\varepsilon} x(x-h)\,dG(x) + \varepsilon \int_{r+\varepsilon}^{+\infty} x\,dG(x) \leq \frac{G(r+\varepsilon) - G(r)}{c}.$$

Dividing by ε and letting ε tend to 0, we obtain, for all $h > 0$, $\int_r^\infty x\,dG(x) \leq p(r)/c$. Thus, the function $V(r) = \int_r^{+\infty} x\,dG(x) = \int_r^{+\infty} xp(x)\,dx$ satisfies the differential inequality $V(r) \leq -V'(r)/(cr)$, that is $(\log V(r))' \leq -c$, which is equivalent to saying that $\log V(r) + cr^2/2$ is non-increasing, and therefore the function $T_g(r)$ is non-increasing. □

We relax now the condition on Stein kernels, by assuming that it is "sub-linear". This condition is fulfilled by many important distributions, for instance by the Gaussian, Gamma or Beta distributions. We deduce a sub-Gamma behavior.

Theorem 12. *Assume that ν has a finite first moment and a density p with respect to the Lebesgue measure that has a connected support. Assume also that the Stein kernel τ_v is sub-linear, that is $\tau_v(x) \leq a|x-\mu| + b$, where μ is the mean value of ν. Then for any 1-Lipschitz function g and any $r > 0$,*

$$\mathbb{P}\left(g \geq \int g\,d\nu + r\right) \leq e^{-\frac{r^2}{2ar+2b}}. \tag{38}$$

When $g = \mathrm{Id}$, inequality (38) was proved by Nourdin and Viens [47] under the stronger condition that $\tau_v(x) \leq a(x-\mu) + b$ (which induces that the support of ν is bounded from below if $a > 0$).

Proof. Take g to be 1-Lipschitz and mean zero with respect to ν. Without loss of generality, we may assume that g is bounded (otherwise we approximate g by thresholding its largest values). Then for any $\lambda \geq 0$,

$$\mathbb{E}\big[g e^{\lambda g}\big] = \mathrm{Cov}\big(g, e^{\lambda g}\big) \leq \mathbb{E}\big[\big|\big(e^{\lambda g}\big)'\big|\tau_v\big] \leq \lambda \mathbb{E}\big[e^{\lambda g}\tau_v\big]. \tag{39}$$

Furthermore,

$$\mathbb{E}\big[e^{\lambda g}\tau_v\big] \le a\mathbb{E}\big[|X-\mu|e^{\lambda g(X)}\big] + b\mathbb{E}\big[e^{\lambda g}\big] \tag{40}$$

and $\mathbb{E}[|X-\mu|e^{\lambda g(X)}] = \mathbb{E}[(X-\mu)h(X)]$ where $h(x) = \operatorname{sign}(x-\mu)\exp(\lambda g(x))$ and $\operatorname{sign}(x-\mu) = 2\cdot\mathbf{1}\{x\ge\mu\}-1$. As $h'(x) = \operatorname{sign}(x-\mu)\lambda g'(x)\exp(\lambda g(x))$ a.s., we get

$$\mathbb{E}\big[|X-\mu|e^{\lambda g(X)}\big] = \lambda\mathbb{E}\big[\operatorname{sign}(X-\mu)g'(X)e^{\lambda g(X)}\tau_v(X)\big] \le \lambda\mathbb{E}\big[e^{\lambda g}\tau_v\big],$$

which gives, by combining with (40), $\mathbb{E}[e^{\lambda g}\tau_v] \le \lambda a\mathbb{E}[e^{\lambda g}\tau_v] + b\mathbb{E}[e^{\lambda g}]$. If $\lambda < 1/a$, this gives

$$\mathbb{E}\big[e^{\lambda g}\tau_v\big] \le \frac{b}{1-\lambda a}\mathbb{E}\big[e^{\lambda g}\big]. \tag{41}$$

Combining (39) and (41), we obtain, for any $\lambda < 1/a$, $\mathbb{E}[ge^{\lambda g}] \le \mathbb{E}[e^{\lambda g}]\lambda b/(1-\lambda a)$. Define $J(\lambda) = \log\mathbb{E}[e^{\lambda g}]$, $\lambda \ge 0$. We thus have the following differential inequality, $J'(\lambda) \le \lambda b/(1-\lambda a)$. Since $J(0) = 0$, this implies that $J(\lambda) \le \lambda^2 b/(2(1-\lambda a))$. Equivalently, $\mathbb{E}[e^{\lambda g}] \le e^{\lambda^2 b/(2(1-\lambda a))}$, which by the use of Chebyshev's inequality gives (34). □

Actually, particularizing to the variable X itself, we have the following concentration bounds, in the spirit of the generalized Mills' type inequalities obtained in Theorem 11.

Theorem 13. *Assume that ν has a finite first moment and a density p with respect to the Lebesgue measure that has a connected support. Assume also that the Stein kernel τ_ν is "sub-linear", that is $\tau_\nu(x) \le a|x-\mu| + b$ for some $a > 0$ and $b \ge 0$, where μ is the mean value of ν. Then the function*

$$T(r) = (ar+b)^{-b/a^2}e^{r/a}\mathbb{E}\big[(X-\mu)\mathbf{1}_{\{X-\mu\ge r\}}\big]$$

is non-increasing in $r \ge 0$. In particular, for all $r > 0$,

$$\mathbb{P}(X \ge \mu + r) \le \mathbb{E}(X-\mu)_+\frac{(ar+b)^{b/a^2}e^{-r/a}}{r}, \tag{42}$$

$$\mathbb{P}\big(|X-\mu| \ge r\big) \le \mathbb{E}|X-\mu|\frac{(ar+b)^{b/a^2}e^{-r/a}}{r}. \tag{43}$$

The concentration bounds (42) and (43), that seem to be new, have an interest for large values of r, where they improve upon Theorem 12 if $a > 0$, due to the factor 2 in front of the constant a in the right-hand side of (38).

Let us now state a more general theorem.

Theorem 14. *Assume that ν has a finite first moment, a density p with respect to the Lebesgue measure that has a connected support and denote τ_ν its Stein kernel. Set X a random variable of distribution ν. Take f a 1-Lipschitz function with mean zero with respect to ν and assume that*

f has an exponential moment with respect to ν, that is there exists $a > 0$ such that $\mathbb{E}[e^{af(X)}] < +\infty$. Then for any $\lambda \in (0, a)$,

$$\mathbb{E}\big[e^{\lambda f(X)}\big] \le \mathbb{E}\big[e^{\lambda^2 \tau_\nu(X)}\big]. \tag{44}$$

Consequently, if we denote $\psi_\tau(\lambda) = \ln \mathbb{E}[e^{\lambda^2 \tau_\nu(X)}] \in [0, +\infty]$ and $\psi_\tau^*(t) = \sup_{\lambda \in (0,a)}\{t\lambda - \psi_\tau(\lambda)\}$ the Fenchel–Legendre dual function of ψ_λ, then for any $r > 0$,

$$\mathbb{P}\big(f(X) > r\big) \vee \mathbb{P}\big(f(X) < -r\big) \le \exp\big(-\psi_\tau^*(r)\big). \tag{45}$$

Theorem 14 states that the concentration of Lipschitz functions taken on a real random variable with existing Stein kernel is controlled by the behavior of the exponential moments of the Stein kernel itself – if it indeed admits finite exponential moments.

Let us now briefly detail how to recover from Theorem 14 some results of Theorems 11 and 12, although with less accurate constants. If $\|\tau_\nu\|_\infty < +\infty$, then inequality (44) directly implies

$$\mathbb{E}\big[e^{\lambda f(X)}\big] \le e^{\lambda^2 \|\tau_\nu\|_\infty},$$

which gives

$$\mathbb{P}\big(f(X) > r\big) \vee \mathbb{P}\big(f(X) < -r\big) \le \exp\left(-\frac{r^2}{4\|\tau_\nu\|_\infty}\right).$$

The latter inequality takes the form of Inequality (34) of Theorem 11, although with a factor $1/2$ in the argument of the exponential in the right-hand side of the inequality.

Assume now, as in Theorem 12, that the Stein kernel τ_ν is sub-linear, that is there exist $a, b \in \mathbb{R}_+$ such that $\tau_\nu(x) \le a(x - \mu) + b$, where μ is the mean value of ν. Inequality (44) implies in this case,

$$\mathbb{E}\big[e^{\lambda f(X)}\big] \le \mathbb{E}\big[e^{a\lambda^2 (X-\mu)}\big] e^{b\lambda^2}. \tag{46}$$

The latter inequality being valid for any f being 1-Lipschitz and centered with respect to ν, we can apply it for $f(X) = X - \mu$. This gives

$$\mathbb{E}\big[e^{\lambda (X-\mu)}\big] \le \mathbb{E}\big[e^{a\lambda^2 (X-\mu)}\big] e^{b\lambda^2}.$$

Now, considering $\lambda < a^{-1}$, we have by Hölder's inequality, $\mathbb{E}[e^{a\lambda^2(X-\mu)}] \le \mathbb{E}[e^{\lambda(X-\mu)}]^{a\lambda}$. Plugging this estimate in the last inequality and rearranging the terms of the inequality gives

$$\mathbb{E}\big[e^{\lambda (X-\mu)}\big] \le e^{\frac{b\lambda^2}{1-\lambda a}}.$$

Going back to inequality (46), we obtain, for any $\lambda \in (0, a^{-1})$,

$$\mathbb{E}\big[e^{\lambda f(X)}\big] \le \mathbb{E}\big[e^{\lambda (X-\mu)}\big]^{a\lambda} e^{b\lambda^2} \le e^{b\lambda^2 (\frac{\lambda a}{1-\lambda a} + 1)} = e^{\frac{b\lambda^2}{1-\lambda a}}.$$

By the use of Cramèr–Chernoff method, this gives the result of Theorem 12, although with a constant $1/2$ in the argument of the exponential term controlling the deviations.

Proof. First note that Inequality (45) is direct consequence of Inequality (44) via the use of the Cramèr–Chernoff method (see for instance Section 2.2 of Boucheron, Lugosi and Massart [16]). To prove Inequality (44), also note that by Lemma 15 below, it suffices to prove that for any $\lambda \in (0, a)$,

$$\mathbb{E}\left[\lambda f(X) e^{\lambda f(X)}\right] \leq \mathbb{E}\left[\lambda^2 \tau_v(X) e^{\lambda f(X)}\right]. \tag{47}$$

Take $\lambda \in (0, a)$, it holds by identity (12),

$$\mathbb{E}\left[f(X) e^{\lambda f(X)}\right] = \operatorname{Cov}\left(f(X), e^{\lambda f(X)}\right) = \mathbb{E}\left[\lambda f'(X) \bar{\mathcal{L}}(\lambda f)(X) e^{\lambda f(X)}\right].$$

Hence, we obtain

$$\mathbb{E}\left[f(X) e^{\lambda f(X)}\right] \leq \lambda^2 \mathbb{E}\left[\left|f'(X)\right| \tau_v(X) e^{\lambda f(X)}\right] \leq \mathbb{E}\left[\lambda^2 \tau_v(X) e^{\lambda f(X)}\right].$$

Inequality (47) is thus proved, which completes the proof. □

Lemma 15. *Take X a random variable on a measurable space $(\mathcal{X}, \mathcal{T})$. Take g and h two measurable functions from $\mathcal{X}$ to $\mathbb{R}$ such that*

$$\mathbb{E}\left[g(X) e^{g(X)}\right] \leq \mathbb{E}\left[h(X) e^{g(X)}\right] < +\infty. \tag{48}$$

Then it holds,

$$\mathbb{E}\left[e^{g(X)}\right] \leq \mathbb{E}\left[e^{h(X)}\right]. \tag{49}$$

Lemma 15 summarizes the essence of the argument used in the proof of Theorem 2.3 of Bobkov, Götze and Houdré [7]. We could not find a reference in the literature for Lemma 15. We point out that Lemma 15 may have an interest by itself as it should be very handy when dealing with concentration inequalities using the Cramèr–Chernoff method. Its scope may thus go beyond our framework related to the behavior of the Stein kernel.

Proof. Note that if $\mathbb{E}[e^{h(X)}] = +\infty$ then Inequality (49) is satisfied. We assume now that $\mathbb{E}[e^{h(X)}] < +\infty$ and $\beta = \ln(\mathbb{E}[e^{h(X)}])$. By setting $U = h(X) - \beta$, we get $\mathbb{E}[e^U] = 1$ and so, by the duality formula for the entropy (see for instance Theorem 4.13 in Boucheron, Lugosi and Massart [16]), we have

$$\mathbb{E}\left[U e^{g(X)}\right] \leq \operatorname{Ent}\left(e^{g(X)}\right) = \mathbb{E}\left[g(X) e^{g(X)}\right] - \mathbb{E}\left[e^{g(X)}\right] \ln\left(\mathbb{E}\left[e^{g(X)}\right]\right).$$

Furthermore,

$$\mathbb{E}\left[g(X) e^{g(X)}\right] - \beta \mathbb{E}\left[e^{g(X)}\right] \leq \mathbb{E}\left[U e^{g(X)}\right].$$

Putting the above inequalities together, we obtain $\beta \geq \ln(\mathbb{E}[e^{g(X)}])$, which is equivalent to (49). □

5. Tail bounds

In the following theorem, we establish lower tail bounds when the Stein kernel is uniformly bounded away from zero. In particular, the support of the measure is $\mathbb{R}$ in this case, as can be seen from the explicit formula (6). Some further tail bounds can be found in Saumard [49], Section 3.

Theorem 16. *Take a real random variable X of distribution ν with density p with respect to the Lebesgue measure on $\mathbb{R}$. Assume that $\mathbb{E}[X]=0$, p has a connected support and denote τ_ν the Stein kernel of ν. If $\tau_\nu \geq \sigma_{\min}^2 > 0$ ν-almost surely, then the density p of ν is positive on $\mathbb{R}$ and the function*

$$R(x) = e^{x^2/2\sigma_{\min}^2} \int_x^{+\infty} yp(y)\,dy$$

is nondecreasing on $\mathbb{R}_+$. In particular, for any $x \geq 0$,

$$\int_x^{+\infty} yp(y)\,dy \geq \mathbb{E}\big[(X)_+\big] e^{-x^2/2\sigma_{\min}^2}. \tag{50}$$

By symmetry, for any $x \leq 0$,

$$-\int_{-\infty}^{x} yp(y)\,dy \geq \mathbb{E}\big[(X)_-\big] e^{-x^2/2\sigma_{\min}^2}. \tag{51}$$

Assume in addition that the function $L(x) = x^{1+\beta} p(x)$ is nonincreasing on $[s, +\infty)$, $s > 0$. Then for all $x \geq s$, it holds

$$\mathbb{P}(X \geq x) \geq \left(1 - \frac{1}{\beta}\right) \frac{\mathbb{E}[(X)_+]}{x} \exp\left(-\frac{x^2}{2\sigma_{\min}^2}\right). \tag{52}$$

Alternatively, assume that there exists $\alpha \in (0,2)$ such that $\limsup_{x\to+\infty} x^{-\alpha} \log \tau_\nu(x) < +\infty$. Then for any $\delta \in (0,2)$, there exist $L, x_0 > 0$ such that, for all $x > x_0$,

$$\mathbb{P}(X \geq x) \geq \frac{L}{x} \exp\left(-\frac{x^2}{(2-\delta)\sigma_{\min}^2}\right). \tag{53}$$

The results presented in Theorem 16 can be found in Nourdin and Viens [47] under the additional assumption, related to the use of Malliavin calculus, that the random variable $X \in \mathbf{D}^{1,2}$.

Proof. For any smooth function φ nondecreasing,

$$\mathbb{E}\big[X\varphi(X)\big] = \mathbb{E}\big[\tau_\nu(X)\varphi'(X)\big] \geq \sigma_{\min}^2 \mathbb{E}\big[\varphi'(X)\big].$$

Take $\varphi(x) = \min\{(x-c)_+, \varepsilon\}$, for some $c \in \mathbb{R}$ and $\varepsilon > 0$. Then

$$\mathbb{E}\big[X\varphi(X)\big] = \int_c^{c+\varepsilon} x(x-c)p(x)\,dx + \varepsilon \int_{c+\varepsilon}^{+\infty} xp(x)\,dx$$

and $\mathbb{E}[\varphi'(X)] = \mathbb{P}(X \in (c, c+\varepsilon])$. Dividing the latter two terms by ε and letting ε tend to zero gives,

$$\int_c^{+\infty} xp(x)\,dx \geq \sigma_{\min}^2 p(c). \tag{54}$$

Now set $V(c) = \int_c^{+\infty} xp(x)\,dx$. Inequality (54) writes, for any $c \geq 0$, $c\sigma_{\min}^{-2} V(c) \geq -V'(c)$. Then define, for any $c \geq 0$, $R(c) = V(c)\exp(c^2/(2\sigma_{\min}^2))$. We can differentiate R and we have

$$R'(c) = \left(V'(c) + \frac{c}{\sigma_{\min}^2} V(c)\right)\exp\left(\frac{c^2}{2\sigma_{\min}^2}\right) \geq 0.$$

In particular $R(c) \geq R(0)$, which gives (50). As $\tau_{-X}(x) = \tau_X(-x)$, we deduce by symmetry that (51) also holds. The proof of inequalities (52) and (53) follows from the same arguments as in the proof of points (ii) and (ii)$'$, Theorem 4.3, Nourdin and Viens [47]. We give them, with slight modifications, for the sake of completeness. By integration by parts, we have

$$V(c) = c\mathbb{P}(X \geq c) + \int_c^{+\infty} \mathbb{P}(X \geq x)\,dx.$$

We also have, for $x > 0$,

$$\mathbb{P}(X \geq x) = \int_x^{+\infty} \frac{y^{1+\beta} p(y)}{y^{1+\beta}}\,dy \leq x^{1+\beta} p(x) \int_x^{+\infty} \frac{dy}{y^{1+\beta}} = \frac{xp(x)}{\beta}.$$

Hence,

$$V(c) \leq c\mathbb{P}(X \geq c) + \beta^{-1} \int_c^{+\infty} xp(x)\,dx = c\mathbb{P}(X \geq c) + \frac{V(c)}{\beta}$$

or equivalently, $\mathbb{P}(X \geq c) \geq (1 - 1/\beta)V(c)/c$. The conclusion follows by combining the latter inequality with inequality (50). It remains to prove (53). By formula (20) applied with $h(y) \equiv y$ – note that this is possible since by assumption $\tau_\nu > 0$ on $\mathbb{R}$, – it holds

$$p(x) = \frac{\mathbb{E}[|X|]}{2\tau_\nu(x)}\exp\left(-\int_0^x \frac{y}{\tau_\nu(y)}\,dy\right) \geq \frac{\mathbb{E}[|X|]}{2\tau_\nu(x)}\exp\left(-\frac{x^2}{2\sigma_{\min}^2}\right).$$

Let us fix $\varepsilon > 0$. By assumption on τ_ν, we get that there exists a positive constant C such that, for x large enough,

$$p(x) \geq C\exp\left(-\frac{x^2}{2\sigma_{\min}^2} - x^\alpha\right) \geq C\exp\left(-\frac{x^2}{(2-\varepsilon)\sigma_{\min}^2}\right).$$

Hence, for x large enough,

$$\mathbb{P}(X \geq x) \geq C\int_x^{+\infty} \exp\left(-\frac{y^2}{(2-\varepsilon)\sigma_{\min}^2}\right)dy.$$

The conclusion now easily follows from the following classical inequality: $\int_x^{+\infty} e^{-y^2/2}\,dy \geq (x/(1+x^2))\exp(-x^2/2)$. □

Acknowledgments

I owe thanks to Guillaume Poly for an instructive discussion about Stein's method, to Yvik Swan for a number of pointers to the literature and to Jon Wellner as well as Max Fathi for useful comments. I also warmly thank Michel Bonnefont and Aldéric Joulin for nice discussions and valuable comments related to some (weighted) functional inequalities. Finally, I express my gratitude to two anonymous referees for several insightful comments and suggestions that helped to complete the study and improve the quality of the paper.

Supplementary Material

Supplement to the article "Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one" (DOI: 10.3150/19-BEJ1117SUPP; .pdf). We provide additional results, comments and detail some proofs pertaining the work of this article.

References

[1] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C. and Scheffer, G. (2000). *Sur les Inégalités de Sobolev Logarithmiques*. *Panoramas et Synthèses* [*Panoramas and Syntheses*] **10**. Paris: Société Mathématique de France. With a preface by Dominique Bakry and Michel Ledoux. MR1845806

[2] Bakry, D., Cattiaux, P. and Guillin, A. (2008). Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.* **254** 727–759. MR2381160

[3] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209

[4] Barbour, A.D., Ross, N. and Wen, Y. (2018). Central moment inequalities using Stein's method. Preprint. Available at arXiv:1802.10225.

[5] Barthe, F. and Roberto, C. (2003). Sobolev inequalities for probability measures on the real line. *Studia Math.* **159** 481–497. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday (Polish). MR2052235

[6] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772

[7] Bobkov, S.G., Götze, F. and Houdré, C. (2001). On Gaussian and Bernoulli covariance representations. *Bernoulli* **7** 439–451. MR1836739

[8] Bobkov, S.G. and Houdré, C. (1997). Isoperimetric constants for product probability measures. *Ann. Probab.* **25** 184–205. MR1428505

[9] Bobkov, S. and Ledoux, M. (1997). Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields* **107** 383–400. MR1440138

[10] Bobkov, S.G. and Ledoux, M. (2009). On weighted isoperimetric and Poincaré-type inequalities. In *High Dimensional Probability V*: *The Luminy Volume*. *Inst*. *Math*. *Stat*. (*IMS*) *Collect*. **5** 1–29. Beachwood, OH: IMS. MR2797936

[11] Bobkov, S.G. and Ledoux, M. (2009). Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann*. *Probab*. **37** 403–427. MR2510011

[12] Bobkov, S. and Ledoux, M. (2014). One-dimensional empirical measures, order statistics, and Kantorovich transport distances. *Mem*. *Amer. Math*. *Soc*. To appear.

[13] Bonnefont, M. and Joulin, A. (2014). Intertwining relations for one-dimensional diffusions and application to functional inequalities. *Potential Anal*. **41** 1005–1031. MR3269712

[14] Bonnefont, M., Joulin, A. and Ma, Y. (2016). A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. *ESAIM Probab*. *Stat*. **20** 18–29. MR3519678

[15] Bonnefont, M., Joulin, A. and Ma, Y. (2016). Spectral gap for spherically symmetric log-concave probability measures, and beyond. *J*. *Funct*. *Anal*. **270** 2456–2482. MR3464047

[16] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities*. Oxford: Oxford Univ. Press. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR3185193

[17] Cattiaux, P., Guillin, A. and Wu, L.-M. (2011). Some remarks on weighted logarithmic Sobolev inequality. *Indiana Univ. Math*. *J*. **60** 1885–1904. MR3008255

[18] Cattiaux, P., Gozlan, N., Guillin, A. and Roberto, C. (2010). Functional inequalities for heavy tailed distributions and application to isoperimetry. *Electron*. *J*. *Probab*. **15** 346–385. MR2609591

[19] Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072

[20] Chatterjee, S. (2012). A new approach to strong embeddings. *Probab. Theory Related Fields* **152** 231–264. MR2875758

[21] Chatterjee, S. (2014). A short survey of Stein's method. In *Proceedings of the International Congress of Mathematicians – Seoul* 2014. *Vol*. *IV* 1–24. Seoul: Kyung Moon Sa. MR3727600

[22] Chatterjee, S. and Dey, P.S. (2010). Applications of Stein's method for concentration inequalities. *Ann*. *Probab*. **38** 2443–2485. MR2683635

[23] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications* (*New York*). Heidelberg: Springer. MR2732624

[24] Cook, N., Goldstein, L. and Johnson, T. (2018). Size biased couplings and the spectral gap for random regular graphs. *Ann*. *Probab*. **46** 72–125. MR3758727

[25] Cordero-Erausquin, D. and Gozlan, N. (2017). Transport proofs of weighted Poincaré inequalities for log-concave distributions. *Bernoulli* **23** 134–158. MR3556769

[26] Courtade, T.A., Fathi, M. and Pananjady, A. (2017). Existence of Stein Kernels under a Spectral Gap, and Discrepancy Bound. Preprint. Available at arXiv:1703.07707.

[27] Döbler, C. (2015). Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron*. *J*. *Probab*. **20** 109. MR3418541

[28] Dümbgen, L. (2010). Bounding Standard Gaussian Tail Probabilities. Technical Report.

[29] Eden, R. and Viens, F. (2013). General upper and lower tail estimates using Malliavin calculus and Stein's equations. In *Seminar on Stochastic Analysis*, *Random Fields and Applications VII*. *Progress in Probability* **67** 55–84. Basel: Birkhäuser/Springer. MR3380095

[30] Fathi, M. (2018). Stein kernels and moment maps. *Ann*. *Probab*. To appear. Available at arXiv:1804.04699.

[31] Fathi, M. and Nelson, B. (2017). Free Stein kernels and an improvement of the free logarithmic Sobolev inequality. *Adv*. *Math*. **317** 193–223. MR3682667

[32] Ghosh, S. and Goldstein, L. (2011). Applications of size biased couplings for concentration of measures. *Electron*. *Commun*. *Probab*. **16** 70–83. MR2763529

[33] Ghosh, S. and Goldstein, L. (2011). Concentration of measures via size-biased couplings. *Probab. Theory Related Fields* **149** 271–278. MR2773032

[34] Ghosh, S., Goldstein, L. and Raič, M. (2011). Concentration of measure for the number of isolated vertices in the Erdős–Rényi random graph by size bias couplings. *Statist. Probab. Lett.* **81** 1565–1570. MR2832913

[35] Goldstein, L. and Işlak, Ü. (2014). Concentration inequalities via zero bias couplings. *Statist. Probab. Lett.* **86** 17–23. MR3162712

[36] Gozlan, N. (2010). Poincaré inequalities and dimension free concentration of measure. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 708–739. MR2682264

[37] Latała, R. and Oleszkiewicz, K. (2000). Between Sobolev and Poincaré. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1745** 147–168. Berlin: Springer. MR1796718

[38] Ledoux, M., Nourdin, I. and Peccati, G. (2015). Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.* **25** 256–306. MR3320893

[39] Ledoux, M., Nourdin, I. and Peccati, G. (2017). A Stein deficit for the logarithmic Sobolev inequality. *Sci. China Math.* **60** 1163–1180. MR3665794

[40] Ley, C., Reinert, G. and Swan, Y. (2017). Stein's method for comparison of univariate distributions. *Probab. Surv.* **14** 1–52. MR3595350

[41] Liu, Q., Lee, J. and Jordan, M. (2016). A kernelized Stein discrepancy for goodness-of-fit tests. In *International Conference on Machine Learning* 276–284.

[42] Mackey, L. and Gorham, J. (2016). Multivariate Stein factors for a class of strongly log-concave distributions. *Electron. Commun. Probab.* **21** 56. MR3548768

[43] Mackey, L., Jordan, M.I., Chen, R.Y., Farrell, B. and Tropp, J.A. (2014). Matrix concentration inequalities via the method of exchangeable pairs. *Ann. Probab.* **42** 906–945. MR3189061

[44] Menz, G. and Otto, F. (2013). Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.* **41** 2182–2224. MR3098070

[45] Miclo, L. (2008). Quand est-ce que des bornes de Hardy permettent de calculer une constante de Poincaré exacte sur la droite? *Ann. Fac. Sci. Toulouse Math.* (6) **17** 121–192. MR2464097

[46] Muckenhoupt, B. (1972). Hardy's inequality with weights. *Studia Math.* **44** 31–38. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I. MR0311856

[47] Nourdin, I. and Viens, F.G. (2009). Density formula and concentration inequalities with Malliavin calculus. *Electron. J. Probab.* **14** 2287–2309. MR2556018

[48] Paulin, D., Mackey, L. and Tropp, J.A. (2016). Efron–Stein inequalities for random matrices. *Ann. Probab.* **44** 3431–3473. MR3551202

[49] Saumard, A. Supplement to "Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one." DOI:10.3150/19-BEJ1117SUPP.

[50] Saumard, A. and Wellner, J.A. (2018). Efron's monotonicity property for measures on $\mathbb{R}^2$. *J. Multivariate Anal.* **166** 212–224. MR3799644

[51] Saumard, A. and Wellner, J.A. (2018). On the isoperimetric constant, covariance inequalities and L_p-Poincaré inequalities in dimension one. *Bernoulli*, To appear. Available at arXiv:1711.00668.

[52] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California, Berkeley, Calif.*, 1970/1971), *Vol. II*: *Probability theory* 583–602. Berkeley, CA: Univ. California Press. MR0402873

[53] Stein, C. (1986). *Approximate Computation of Expectations. Institute of Mathematical Statistics Lecture Notes – Monograph Series* **7**. Hayward, CA: IMS. MR0882007

[54] Treilhard, J. and Mansouri, A.-R. (2015). Concentration inequalities via Malliavin calculus with applications. *Electron. Commun. Probab.* **20** 36. MR3352331

[55] Viens, F.G. (2009). Stein's lemma, Malliavin calculus, and tail bounds, with application to polymer fluctuation exponent. *Stochastic Process*. *Appl*. **119** 3671–3698. MR2568291

Received June 2018 and revised February 2019