GEOMETRICAL PROPERTIES OF THE DIFFUSION SEMIGROUPS AND CONVEX INEQUALITIES

ALEXANDER V. KOLESNIKOV

Abstract. We prove that the diffusion semigroup T_t generated by the second-order elliptic operator $\Delta + \langle b, \nabla \rangle$, where b is smooth, antisymmetric and globally Lipschitz, preserves Sherman functions on $\mathbb{R}^d$, where $d \geq 2$, if and only if b is a linear mapping. The class of Sherman functions includes all even quasi-concave functions. In particular, if T_t admits Sherman transition probabilities, then b is linear. We also discuss relations with the Gaussian correlation conjecture and prove some related inequalities.

1. Introduction

In this paper we study some special geometric properties of the diffusion semigroups which are closely related to the notion of convexity. We also discuss applications of the PDE methods to the Gaussian correlation conjecture. This paper continues the author's research [20], where the following result was established. Consider a diffusion semigroup T_t on $\mathbb{R}^d$ generated by the second-order elliptic operator $a^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + b^i \frac{\partial}{\partial x_i}$ with smooth coefficients (here and after we mean the summation in indices i, j). Suppose that $T_t f$ is a log-concave function for every log-concave f (we recall that a function f is log-concave, if it has the form $f = e^{-V}$, where V is convex). Then every a^{ij} is constant and b^i is affine.

This study was motivated in part by a well known problem from the theory of Gaussian measures, the so-called Gaussian correlation conjecture (see a detailed discussion in Section 3). The Gaussian correlation conjecture states that the standard Gaussian measure γ on $\mathbb{R}^d$ satisfies the following inequality:

$$\gamma(A \cap B) \geq \gamma(A)\gamma(B), \tag{1.1}$$

where A and B are symmetric convex sets. This inequality has been proved so far only for some certain pairs of sets (see [5], [13], [14], [25], [28], [30]). In particular, (1.1) holds for the two-dimensional sets. This inequality has been proved in [28] under some restriction on the size of the sets.

It was shown in [13] that (1.1) holds if A is an ellipsoid. An important point of the proof was the property of the Ornstein-Uhlenbeck semigroup to preserve log-concave functions. One can hope that applying some other semigroups with such a property one can extend this result to other measures or functions. Unfortunately, by a result from [20] only the Gaussian semigroups (i.e. the semigroups with the Gaussian transition probabilities) preserve log-concavity. In Section 2 we consider another important class of functions which includes all symmetric quasi-concave functions (the so-called Sherman functions). This class is preserved not only by Gaussian semigroups. As a consequence we get examples of non-Gaussian measures satisfying the correlation inequality (see also a more direct approach in Section 3). However, even this class of semigroups is not very rich. We prove that if we restrict ourselves to diffusions with a constant diffusion matrix and antisymmetric, globally Lipschitz and smooth drift, then the corresponding semigroups turn out to be Gaussian if $d \geq 2$.

The geometrical properties of transition probabilities of the elliptic and parabolic equations have been considered in many works ([6], [7], [8], [9], [10], [15], [17], [18], [19], [22]). We note that the property of semigroups to preserve some classes of functions (e.g. log-concave or Sherman) can characterize the geometric properties of the corresponding transition probabilities. For example, the semigroups with log-concave transition probabilities preserve log-concavity and semigroups with symmetric and quasi-concave transition probabilities preserve Sherman functions (see Corollary 2.17). In particular, every semigroup with a constant diffusion matrix and a smooth Lipschitz drift which admits symmetric quasi-concave transition probabilities (or preserves even quasi-concave functions) is Gaussian.

In Section 3 we consider some applications of the Sherman functions. In particular, we prove some special cases of the Gaussian correlation inequality and study some correlation properties of measures with Sherman densities. In addition we establish some relations with the well-known Brascamp-Lieb inequality.

2. Semigroups preserving Sherman functions.

In this section we consider a class of functions which we call Sherman functions. Apparently they have been considered first by S. Sherman in [29].

First we recall the well-known Prékopa theorem (see [26] for the original proof and [1], [2], [23] for further developments and applications).

Prékopa's Theorem. *For every log-concave integrable function $f(x,y)$ on $\mathbb{R}^d \times \mathbb{R}^k$, the integral $\int_{\mathbb{R}^k} f(x,y)\,dy$ is a log-concave function of x.*

Definition 2.1. *We say that $f : \mathbb{R}^d \to [0,\infty)$ is a Sherman function if f is a limit of functions of the type*

$$\sum_{i=1}^{k} c_i I_{A_i}(x), \quad c_i \ge 0,$$

where every $A_i \subset \mathbb{R}^d$ is convex, bounded and symmetric about the origin, with respect to the norm $\operatorname{ess\ sup}_{x\in\mathbb{R}^d} |f(x)|$. *The set of these functions will be denoted by $\mathcal{S}h^d$.*

Sherman functions enjoy a lot of nice properties. It follows directly from the definition that they are stable under compositions with linear mappings. Clearly, a product and a sum of two Sherman functions is Sherman.

We call a function f symmetric (antisymmetric) if $f(-x) = f(x)$ $\bigl(f(-x) = -f(x)\bigr)$. It follows from the definition, that every Sherman function is symmetric.

Definition 2.2. *We call a bounded function f quasi-concave if all the level sets $\{f > c\}$ are convex.*

Remark 2.3. *Obviously, every log-concave function is quasi-concave. It is easy to check that every even quasi-concave function satisfying $\lim_{R\to\infty} |f|_{L^\infty(\mathbb{R}^d\setminus B_R)} = 0$ is a Sherman function. Indeed, it suffices to consider approximations of the type*

$$\Bigl(\sum_{i=1}^{n} (a_{i+1} - a_i) I_{\{f>a_i\}}\Bigr) I_{\{|x|\le R\}},$$

where

$$0 = a_1 < a_2 < \cdots < a_{n+1} = \max_{x\in\mathbb{R}_d} f(x)$$

is a partition of $\bigl[0, \max_{x\in\mathbb{R}_d} f(x)\bigr]$ into the equal parts.

It follows from the Prékopa theorem, Remark 2.3 and Definition 2.1 that the Prékopa theorem holds for Sherman functions.

Prékopa's theorem for Sherman functions: If $f \in \mathcal{S}h^{d+k} \cap L^1(\mathbb{R}^d)$, then

$$\int_{\mathbb{R}^k} f(x_1, \cdots, x_{d+k})\,dx^1 \cdots dx^k \in \mathcal{S}h^d.$$

The more special properties are the following (2)-3) hold for some version of f):

1) It follows directly from the Prékopa theorem for Sherman functions that the Sherman functions are stable under convolutions.

2) If $f \in \mathcal{S}h^d$, then $[0,\infty) \ni t \to f(tv)$ is non-increasing for every $v \in \mathbb{R}^d$.

3) One can easily check that for $f \in \mathcal{S}h^d$ and every two points x, y and $0 \le \lambda \le 1$ the following inequality holds: $f(0) + f(\lambda x + (1-\lambda)y) \ge f(x) + f(y)$. In particular, if $I_A \in \mathcal{S}h^d$, then $I_A = I_{A'}$ a. e., where A' is symmetric and convex.

4) The following inequality holds:

$$\int_{\mathbb{R}^d} v^i v^j f_{x_i}(x) g_{x_j}(x)\, dx \ge 0, \tag{2.2}$$

where $f, g \in \mathcal{S}h^d$ are $C_0^\infty(\mathbb{R}^d)$ Sherman functions and $v = (v^1, \cdots, v^d) \in \mathbb{R}^d$.

Proof. Indeed, $q = f * g = \int_{\mathbb{R}^d} f(x-y)g(x)\, dx$ is a Sherman function and attains its maximum in the origin. Consequently, the matrix $D_2 q$ of the second derivatives of q is non-positive in the origin. Hence

$$0 \ge v^i v^j q_{x_i x_j}(0) = \int_{\mathbb{R}^d} v^i v^j f_{x_i x_j}(x) g(x)\, dx = -\int_{\mathbb{R}^d} v^i v^j f_{x_i}(x) g_{x_j}(x)\, dx,$$

which completes the proof. □

In contrast to the log-concave functions, there exist non-Gaussian diffusion semigroups preserving Sherman functions. Let G be a finite-dimensional Lie group and let $R : [0,\infty) \times G \to \mathrm{GL}(d)$ be a mapping with values in the space $GL(d)$ of $d \times d$ matrices with the following properties:

$$1) R(t+s, g_1 g_2) = R(t, g_1) R(s, g_2) \quad 2) R(0,0) = Id. \tag{2.3}$$

Let μ_t be a convolution semigroup of probability measures on G, i.e. a family of measures depending on $t \ge 0$ and satisfying

$$\mu_0 = \delta_e, \quad \mu_t * \mu_s = \mu_{t+s}.$$

Let us consider a semigroup of the following type:

$$T_t f(x) = \int_G f(R(t,g)x)\, d\mu_t(g), \quad x \in \mathbb{R}^d, \quad f : \mathbb{R}^d \to \mathbb{R}. \tag{2.4}$$

One can easily check that T_t is a semigroup indeed. Obviously, T_t preserves Sherman functions on $\mathbb{R}^d$.

The semigroups of such a type were studied by Hunt in [16]. Hunt gave the full description of the generators of these semigroups. It turns out that the diffusion part of the generator of T_t has the form

$$L = a^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + b^i(x) \frac{\partial}{\partial x_i},$$

where every a^{ij} is a quadratic function and every b^i is linear. The non-diffusion part can be described in terms of Lévy measures on G.

The semigroups of this type are called "semigroups with multiplicative noise", in contrast to the semigroups with "additive noise" (generalized Mehler semigroup), which have the form

$$P_t f = \int_{\mathbb{R}^d} f(p_t x + y)\, d\nu_t(dy),$$

where p_t is a semigroup of linear operators on $\mathbb{R}^d$ and ν_t satisfies the following relation:

$$\nu_{t+s} = \left(\nu_t \circ p_s^{-1}\right) * \nu_s.$$

More on the generalized Mehler semigroups see in [4], [11]. One can easily verify that the Ornstein-Uhlenbeck semigroup and the heat semigroup belong to this class.

Example 2.4. *Let us consider the following d-dimensional stochastic differential equation*

$$d\xi_t(x) = \sum_{i=1}^{d} \sigma_i \xi_t(x) dW_t^i + b\xi_t(x)\, dt, \quad \xi_0 = x$$

where $(W_t^1, \cdots, W_t^d)$ is a d-dimensional Wiener process, σ_i and b are constant matrices. If the matrices σ_i, b are mutually commutative, the solution of this equation can be written down by the following explicit formula:

$$\xi_t(x) = \exp\Big(\sum_{i=1}^{d} \sigma_i W_t^i + t\Big[b - \frac{1}{2}\sum_{i=1}^{d} \sigma_i \sigma_i^T\Big]\Big) x.$$

The corresponding semigroup has the form:

$$T_t f(x) = \int_{\mathbb{R}^d} f\Big(\exp\big(\sum_{i=1}^{d} \sigma_i y_i + t\big[b - \frac{1}{2}\sum_{i=1}^{d} \sigma_i \sigma_i^T\big]\big) x\Big)\, d\gamma_t(y_1) \cdots d\gamma_t(y_d),$$

where γ_t is a distribution of the one-dimensional Wiener process at t. Obviously, T_t is a semigroup of the type (2.4). According to the Ito formula the generator of T_t has the form $\frac{1}{2}\alpha^{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \beta^i \frac{\partial}{\partial x_i}$ where $A = (a^{ij})$ and $A = \Sigma\Sigma^T$, where the i-th column of Σ equals to $\sigma_i x$ and $\beta = bx$.

We recall the well-known Trotter formula (see, for example, [27]). Let T_t^1, T_t^2 be contraction semigroups on a Banach space B, generated by operators L_1 and L_2. Assume that the semigroup T_t, generated by the closure of $(L_1 + L_2)|_{D(L_1)\cap D(L_2)}$ is a contraction semigroup. Then it satisfies the relation

$$T_t x = \lim_{n\to\infty} (T_{t/n}^1 T_{t/n}^2)^n x, \quad x \in B. \tag{2.5}$$

Example 2.5. *The heat semigroup $e^{t\Delta}$ preserves Sherman functions by the Prékopa theorem for Sherman functions. It is easy to check that every semigroup generated by $\frac{1}{2}x_i^2\Delta - \alpha x_i \partial_{x_i}$ belongs to the type of semigroups described in Example 2.4, hence preserves Sherman functions. Then by the Trotter formula the semigroup generated by*

$$L_0 = \Big(1 + \frac{1}{2}\sum_{i=1}^{d} x_i^2\Big)\Delta - \alpha\big\langle x, \nabla\big\rangle \tag{2.6}$$

preserves Sherman functions. Note that e^{tL_0} has the following invariant measure:

$$\mu = C(\alpha, d)\frac{dx}{(1 + \frac{1}{2}\sum_{i=1}^{d} x_i^2)^{\alpha+1}}.$$

Then it follows from the Pitt's inequality that μ satisfies the correlation inequality if $d = 2$ (see Section 3 for detailed discussions).

Another example is given by the semigroup generated by

$$L = \frac{1}{2}\sum_{i=1}^{d} x_i^2 \frac{\partial^2}{\partial x_i^2} - \alpha\big\langle x, \nabla\big\rangle. \tag{2.7}$$

(in Example 2.4 one can set: $\sigma_i = diag\{\delta_{ij}\}$, $1 \le j \le d$.) By the same reason as above $e^{t(\Delta+L)}$ preserves Sherman functions. The corresponding invariant measure has the form

$$\mu = Z(\alpha, d)\frac{dx}{\prod_{i=1}^{d}(1 + \frac{1}{2}x_i^2)^{\alpha+1}}.$$

We note that the property of $e^{t(\Delta+L)}$ to preserve Sherman functions can be proved by the Feynman–Kac formula. We consider below a more general case, namely the case when the semigroup is generated by the sum of additive and multiplicative noises.

Example 2.6. The Feynman–Kac formula.

Let us define

$$q_t f(x) = \int_{\mathbb{R}^d} f(x+y)\nu_t(dy), \quad x \in \mathbb{R}^d,\ t > 0, \tag{2.8}$$

$$p_t f(x) = \int_{\mathbb{R}^d} f(R(t,y)x)\mu_t(dy), \quad x \in \mathbb{R}^d,\ t > 0. \tag{2.9}$$

Here μ_t and ν_t are convolution semigroups on $\mathbb{R}^d$ and $R(t,y)$ satisfies (2.3).

We recall that by the Lévy-Khinchine theorem (see [11], [24]) the Fourier transform of ν_t admits the following representation:

$$\mathcal{F}(\nu_t) = e^{-t\lambda(a)}$$

where

$$\lambda(a) = -i\langle a,b\rangle + \frac{1}{2}\langle a, Qa\rangle - \int_{\mathbb{R}^d}\Big(e^{i\langle a,y\rangle} - 1 - \frac{i\langle a,y\rangle}{1+\|y\|^2}\Big)M(dy). \tag{2.10}$$

Here $b \in \mathbb{R}^d$, $Q : \mathbb{R}^d \to \mathbb{R}^d$ is linear, symmetric and non-negative and M is a Lévy measure on $\mathbb{R}^d$, i.e. a Borel measure satisfying

$$M(\{0\}) = 0, \qquad \int_{\mathbb{R}^d} \min(1, ||y||^2)M(dy) < \infty.$$

Let us construct the semigroup $(qp)_t$ generated by the sum of generators of q_t and p_t. To this end we apply the Feynman–Kac formula. First we consider the adjoint semigroups acting on the space of measures:

$$q_t^* : m \to m * \nu_t,$$

$$p_t^* : m \to \int_{\mathbb{R}^d} m \circ R^{-1}(t,y)\mu_t(dy),$$

where m is some Borel finite measure on $\mathbb{R}^d$. We define new semigroups $\mathcal{F}q_t^*$, $\mathcal{F}p_t^*$ acting on functions with the help of Fourier transform

$$\mathcal{F}q_t^* : \hat{m} \to \mathcal{F}(q_t^* m) = \hat{m}(a)e^{-t\lambda(a)}, \tag{2.11}$$

$$\mathcal{F}p_t^* : \hat{m} \to \mathcal{F}(p_t^* m) = \int_{\mathbb{R}} \hat{m}(R^*(t,y)a)\mu_t(\,dy). \tag{2.12}$$

Here $\hat{m} = \mathcal{F}m$ denotes the Fourier transform of m. Let w_t be a stochastic Markov process with values in $\mathbb{R}^d$ with distribution (μ_t) in time t. Let us denote by $\mathcal{P}(w_.)$ the corresponding measure on the space $(\mathbb{R}^d)^{[0,\infty)}$. By the Feynman–Kac formula the semigroup generated by the sum of generators of $\mathcal{F}q_t^*$ and $\mathcal{F}p_t^*$ has the form:

$$\hat{m} \to \int_{(\mathbb{R}^d)^{[0,\infty)}} \hat{m}(R^*(t,w_t)a)e^{-\int_0^t \lambda(R^*(s,w_s)a)\,ds}\,d\mathcal{P}(w_.). \tag{2.13}$$

Note that $e^{-\int_0^t \lambda(R^*(s,w_s)a)\,ds}$ is a Fourier transform of some measure $\nu_{t,w_.}$, depending on a path $w_. \in (\mathbb{R}^d)^{[0,\infty)}$:

$$\mathcal{F}(\nu_{t,w_.}) = e^{-\int_0^t \lambda(R^*(s,w_s)a)\,ds}. \tag{2.14}$$

We get from (2.13) the following expression for $((qp)_t)^*$:

$$((qp)_t)^* : m \to \int_{(\mathbb{R}^d)^{[0,\infty)}} m \circ (R(t,w_t)^{-1}) * \nu_{t,w_.}\,d\mathcal{P}(w_.)$$

and finally for $(qp)_t$:

$$(qp)_t : f \to \int_{(\mathbb{R}^d)^{[0,\infty)}} \int_{\mathbb{R}^d} f(R(t,w_t)x+y)\nu_{t,w_\cdot}(\,dy)\,d\mathcal{P}(w_\cdot). \tag{2.15}$$

It can be seen directly from (2.15) that if μ_t, ν_t are symmetric measures and every ν_t admits Sherman density, then $(qp)_t$ preserves Sherman functions. Indeed, for every path $\omega_\cdot$ the Fourier transform (2.14) can be approximated by the product of the functions of the type $\exp\bigl(-(t_{i+1}-t_i)\lambda(R^*(t_i,\omega_{t_i})a)\bigr)$. Consequently, the measure ν_{t,ω_t} can be approximated by the convolutions of the measures of the type $\nu_{t_{i+1}-t_i}\circ R^{-1}(t_i,\omega_{t_i})$, hence admits Sherman density. The claim follows from the Prékopa theorem for Sherman functions. $\square$

Remark 2.7. *Since even quasi-concave functions generates Sherman functions, every semigroup preserving even quasi-concave functions preserves Sherman functions.*

Now we discuss the case of semigroups generated by a constant diffusion part and a non-linear drift.

Case $d = 1$: One can easily verify with the help of maximum principle that every diffusion semigroup with symmetric transition probabilities preserves Sherman functions. Note that a bounded f is a Sherman function on the line if and only if it is symmetric, non-increasing on $[0,\infty)$ and $\lim_{x\to\infty} f(x) = 0$.

Case $d > 1$: The situation in higher dimensions is more complicated. We don't have the full description of the semigroups preserving Sherman functions in higher dimensions, but, surprisingly, only the Gaussian diffusions enjoys this property in many partial cases. Since the Sherman functions are even, we restrict ourselves to the case when the generator of T_t has the form $L = c\Delta + b^i\frac{\partial}{\partial x_i}$, b is antisymmetric, i.e. $b(-x) = -b(x)$, and c is some constant. Under this assumption T_t preserves even functions. In addition we suppose throughout the paper that b is globally Lipschitz. According to the classical result from the theory of diffusion processes (see [12]) there exists a process $\xi_t(x)$, $t \in [0,\infty)$ satisfying the following stochastic differential equation:

$$\xi_t(x) = x + \sqrt{2}cW_t + \int_0^t b(\xi_s)\,ds.$$

Here W_t is the standard d-dimensional Wiener process. We consider the semigroup

$$T_t : f \to \mathbb{E} f(\xi_t(x)).$$

The corresponding transition probabilities are denoted by $p(t,x,y)$. We consider all the semigroups in the space $C_0(\mathbb{R}^d)$ of continuous functions tending to zero as $x \to \infty$ and equipped with the sup-norm. It is well-known that T_t is a contraction semigroup on this space.

Theorem 2.8. *Let $d \geq 2$, T_t be a diffusion semigroup generated by $L = c\Delta + b^i\frac{\partial}{\partial x_i}$ where c is constant, every b^i is a globally Lipschitz C^2-function and $b^i(-x) = -b^i(x)$. If T_t preserves Sherman functions, then every b^i is linear.*

For every non-degenerated operator B we consider its action on the differential operator L. L^B is obtained from

$$L = a^{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \bigl\langle b(x), \nabla \bigr\rangle$$

by changing the coordinate system $x \mapsto Bx$, e.g.

$$L^B = c^{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \bigl\langle b(Bx), (B^T)^{-1}\nabla \bigr\rangle,$$

where $C = B^{-1}A(B^T)^{-1}$. Suppose that the semigroup $T_t = e^{tL}$ generated by L preserves $\mathcal{S}h^d$. Then by the invariance of Sherman functions under linear transformations the semigroup $T_t = e^{tL^B}$ also preserves $\mathcal{S}h^d$. Consider the probability Haar measure χ_d on the group of orthogonal operators $O(d)$ in d-dimensional space (recall that U is orthogonal if $UU^T = U^TU = I$) and the "average" semigroup $T_t^{O(d)}$ generated by $L^{O(d)} = \int_{O(d)} L^U \, d\chi_d(U)$.

Lemma 2.9. *If L has constant second-order coefficients, then $L^{O(d)}$ has the form*

$$L^{O(d)} = c\Delta + \varphi(\|x\|^2)\langle x, \nabla\rangle. \tag{2.16}$$

Lemma 2.10. *Suppose that T_t generated by $L = c\Delta + b^i \frac{\partial}{\partial x_i}$ preserves Sherman functions. Then the semigroup generated by $L^{O(d)}$ preserves Sherman functions.*

Proof. It follows from the Trotter formula that the semigroup generated by the operators of the type $\sum_{i=1}^n \alpha_i L^{B_i}$, where every B_i is linear and $\alpha_i > 0$, preserves Sherman functions. Approximating $L^{O(d)}$ by operators of this type and using the representation of T_t by stochastic differential equations, we easily get the claim from the classical results on dependence on parameter of the solution of the stochastic differential equations with Lipschitz coefficients (see [12]). □

Lemma 2.11. *Let T_t be generated by (2.16), where φ is twice continuously differentiable. Suppose that T_t preserves $\mathcal{S}h^d$. Then the function $x \to \varphi(\|x\|^2)$ is convex.*

Proof.
Let $p = p(x_1, x_2, \cdots x_d)$ and $f = f(x_1)$ be $C_0^\infty(\mathbb{R}^d)$ Sherman functions. Then inequality (2.2) implies that

$$g(t) = -\int_{\mathbb{R}^d} (T_t f(x_1))_{x_i x_i} p(x) \, dx$$

is non-negative for every i, $t \geq 0$. Take $i > 1$. Since $g(0) = 0$, we get $0 \leq g'(0) = -\int_{\mathbb{R}^d} (Lf)_{x_i x_i} p(x) \, dx$. Taking into account that f does not depend on x_i, we get:

$$-\int_{\mathbb{R}^d} x_1 f_{x_1} \varphi(\|x\|^2)_{x_i x_i} p(x) \, dx \geq 0. \tag{2.17}$$

Let us fix $\alpha > 0$ and construct a sequence of functions $f^\varepsilon(x_1)$ such that $f^\varepsilon(x_1) \to I_{[-\alpha,\alpha]}(x_1)$ almost everywhere and the measures $-x_i (f^\varepsilon)'(x_i) \, dx$ converge weakly to the measure $\alpha(\delta_\alpha + \delta_{-\alpha})$ on the line when $\varepsilon \to 0$. Taking into account that p is symmetric, we get from (2.17)

$$\int_{L_\alpha} \Big[\varphi(\|x\|^2)_{x_i x_i} p(x)\Big]\Big|_{x_1=\alpha} dx_2 \cdots dx_d \geq 0,$$

where $L_\alpha = \{x_1 = \alpha\}$. Note that for every $\tilde{x} \in L_\alpha$ and $\varepsilon > 0$ one can choose p in such a way that $\mathrm{supp}(p) \cap L_\alpha$ is contained in the ε-neighborhood of $\tilde{x}$. This implies that $\big(\varphi(\|x\|^2)_{x_i x_i} \geq 0$ on L_α. Since $\alpha > 0$ is arbitrary and $\big(\varphi(\|x\|^2)_{x_i x_i}$ is continuous, we get that $\big(\varphi(\|x\|^2)_{x_i x_i} \geq 0$ everywhere, $i > 1$. Obviously, in the similar way we can show that $\big(\varphi(\|x\|^2)_{vv} \geq 0$ for every vector $v \in \mathbb{R}^d$. The proof is complete. □

Lemma 2.12. *Suppose that $f : \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function satisfying the following property $\int_{O(d)} f(AUx) \, d\chi_d(U) = 0$ for every linear A. Then $f(-x) = -f(x)$ for every x.*

Proof. Let E be an arbitrary ellipsoid with the center in the origin. Choose a linear operator A such that E is the image of the ball $B = \{x : \|x\| \leq 1\}$ under A. One has

$$\int_E f(x) \, dx = \int_{O(d)} \Big(\int_E f(x) \, dx\Big) \, d\chi_d(U) = |\det A| \int_{O(d)} \Big(\int_B f(Ax) \, dx\Big) \, d\chi_d(U) =$$

$$|\det A| \int_{O(d)} \Big(\int_B f(AUx) \, dx\Big) \, d\chi_d(U) = |\det A| \int_B \Big(\int_{O(d)} f(AUx) \, d\chi_d(U)\Big) \, dx = 0.$$

So, we get $\int_E f(x)\,dx = 0$ for an arbitrary symmetric ellipsoid E. This implies that $\int_{\partial E} f(x)\,d\mathcal{H}_{d-1} = 0$ for the Haussdorf measure $\mathcal{H}_{d-1}$ on ∂E. Consider the sequence of ellipsoids

$$E_\varepsilon = \Big\{(x_1, x_2, \cdots x_d) : \sum_{i=1}^{d-1} x_i^2 + \frac{x_d^2}{\varepsilon^2} = 1\Big\}.$$

Note that the sequence of Haussdorf measures on ∂E_ε tends weakly to the doubled $(d-1)$-dimensional Hausdorff measure on the $(d-1)$-dimensional ellipsoid

$$E_{d-1} = \{x : x_d = 0; \sum_{i=1}^{d-1} x_i^2 \le 1\}.$$

Hence $\int_{E_{d-1}} f(x)\,d\mathcal{H}^{d-1} = 0$. Obviously, the same holds for every symmetric $(d-1)$-dimensional ellipsoid. By the induction method we finally arrive at the dimension 1 and get that for every $v \in \mathbb{R}^d$

$$\int_{-1}^{1} f(tv)\,dt = 0.$$

Hence

$$\int_0^1 \Big(f(tv) + f(-tv)\Big)\,dt = 0, \quad \forall v \in \mathbb{R}^d.$$

This implies that $f(-x) = -f(-x)$. □

Lemma 2.13. *Let b be a continuously differentiable antisymmetric vector field such that for every linear operator A and every $v \in \mathbb{R}^d$*

$$\int_{O(d)} (U)^{-1} A^{-1} b(AUv)\,d\chi(U) = 0.$$

Then $\mathrm{div}(b) = 0$.

Proof. The derivative Db of b satisfies the following relation:

$$\int_{O(d)} (U)^{-1} A^{-1} Db(AUv) AU\,d\chi(U) = 0.$$

Since $\mathrm{Tr}(U)^{-1} A^{-1} Db(AUv)AU = \mathrm{Tr}Db(AUv)$ for every U, one gets

$$\int_{O(d)} \mathrm{Tr}Db(AUv)\,d\chi(U) = 0.$$

By Lemma 2.12 the function $\mathrm{Tr}Db = \mathrm{div}(b)$ is antisymmetric. But Db is symmetric, hence $\mathrm{div}(b) = 0$. □

Now we prove some lemmas in the two-dimensional case. It is convenient to consider another group of operators on $\mathbb{R}^2$ — the group of orthogonal operators $O^+(2)$ which preserve orientation, i.e. $U \in O(2)$ and $\det U = 1$. The $O^+(2)$-averages of second-order differential operators with constant diffusion matrix have the following form:

Lemma 2.14.

$$L^{O^+(2)} = c\Delta + \varphi(x^2+y^2)\Big(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\Big) + q(x^2+y^2)\Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big). \tag{2.18}$$

Lemma 2.15. *Assume that the semigroup generated by*

$$L = c\Delta + q(x^2+y^2)\Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)$$

where q is twice continuously differentiable, preserves Sherman functions. Then q is constant.

Proof. Exactly in the same way as in Lemma 2.11 one proves that

$$(q(x^2+y^2)y)_{yy} \geq 0.$$

But changing coordinates $x \mapsto y,\ y \mapsto x$ one gets that $e^{t\tilde{L}}$, where

$$\tilde{L} = c\Delta - q(x^2+y^2)\Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)$$

also preserves Sherman functions and $(q(x^2+y^2)y)_{yy} \leq 0$. Hence $(q(x^2+y^2)y)_{yy} = 0$ and we get the claim. □

Lemma 2.16. *Let $b : \mathbb{R}^2 \to \mathbb{R}^2$ be a smooth antisymmetric vector field such that $div(b) = 0$ and for every linear operator A and every $x \in \mathbb{R}^2$*

$$\int_{O^+(2)} (U)^{-1}A^{-1}b(AUx)\, d\chi(U) = 0.$$

Then $b = 0$.

Proof. For every smooth compactly supported vector field ω

$$\int_{\mathbb{R}^2} \langle \omega(x), \int_{O^+(2)} (U)^{-1}A^{-1}b(AUx)\, d\chi(U)\rangle\, dx = 0.$$

By the Fubini theorem and change of variables $z = Ax$ one obtains

$$\int_{\mathbb{R}^2} \Big\langle (A^T)^{-1} \int_{O^+(2)} U\omega\big(U^{-1}A^{-1}z\big)\chi(\, dU), b(z)\Big\rangle\, dz = 0.$$

Set: $\omega = q(x_1^2+x_2^2)\Big(x_2, -x_1\Big)$. One can prove by direct computations that for every linear transformation of the form $x = Cz$, where $z = (u,v)$ and

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

one has

$$C^T\omega(Cz) = \big(ad - bc\big)q\big((au+bv)^2 + (cu+dv)^2\big)\Big(v, -u\Big). \tag{2.19}$$

In particular $U\omega(U^{-1}x) = \omega(x)$ for every orthogonal U. Finally, taking $A^{-1} = C$ we get that for every non-degenerated C

$$\int_{\mathbb{R}^2} q(\|Cz\|^2)\big(b^1(z)v - b^2(z)u\big)\, dz = 0.$$

Since q is arbitrary and b is antisymmetric, we prove by arguments of Lemma 2.12 that

$$b^1(u,v)v = b^2(u,v)u.$$

Since $\mathrm{div}(b) = 0$, one can find a symmetric function F such that $F_v = b^1,\ F_u = -b^2$, hence $F_v v + F_u u = 0$. This implies that $F = G(u/v)$ for some G. Obviously, every non-constant function of this type can not be smooth. Hence F is constant and $b = 0$. □

Proof of Theorem 2.8. First we assume that $d = 2$. One can easily prove that every semigroup generated by the first-order differential operator $\big\langle Ax, \nabla\big\rangle$ with linear coefficients preserves Sherman functions. By the Trotter formula the class of semigroups preserving Sherman functions is stable under perturbations by this kind of operators. Subtracting $Db(0)x$ from b one can assume without loss of generality that $Db(0) = 0$. Suppose that b is non-linear. By Lemma 2.10 $T_t = \exp\big(t(L^A)^{O(2)}\big)$ preserves Sherman functions for every linear A. Since b is globally Lipschitz, by Lemma 2.11 and (2.16) the drift of $L_A^{O(2)}$ is linear and, moreover, equals to zero, since $Db(0) = 0$. By Lemma 2.13 $\mathrm{div}(b) = 0$.

In the same way as in Lemma 2.10 one proves that the semigroup $\exp(t(L^A)^{O_2^+})$ preserves Sherman functions. We note that every vector field of the type $A^{-1}b(Ax)$ is a divergence-free field. Indeed, by the property of trace

$$\mathrm{Tr} D\big(A^{-1}b(Ax)\big) = \mathrm{Tr} A^{-1}Db(Ax)A = 0.$$

Hence the drift of $L_A^{O_2^+}$ is divergence-free. Then it follows from Lemma 2.14 that

$$L_A^{O_2^+} = c(A)\Delta + q(A, x_1^2 + x_2^2)\Big(x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}\Big).$$

Then by Lemma 2.15 $q(A, x^2+y^2) = 0$, hence

$$\int_{O^+(2)} (U)^{-1}A^{-1}b(AUx)\, d\chi(U) = 0.$$

Finally, $b = 0$ by Lemma 2.16.

Now let us consider the case $d \geq 3$. Represent $\mathbb{R}^d$ as a direct product $\mathbb{R}^d = X_2 \times X_{d-2}$, where $X_2 = \mathrm{span}\{e_1, e_2\}$, $X_{d-2} = \mathrm{span}\{e_3, \cdots, e_d\}$.

In the same way as above we consider the groups of orthogonal operators $O(2)$ and $O^+(2)$ acting on X_2. Note that

$$\beta(x_1, x_2) = \big(b^1, b^2\big)|_{\{x_i=0,\ i\geq 3\}}$$

is a Lipschitz antisymmetric mapping on X_2. Suppose that β is non-linear. Then according to the two-dimensional result there exists a linear operator A_2 on X_2 such that the average $\mathcal{L}$ of $\big\langle A^{-1}\beta \circ A, \nabla\big\rangle$ with respect to $O(2)$ or $O^+(2)$ has the form

$$\mathcal{L} = \beta_A(x_1^2 + x_2^2)\Big(x_1\frac{\partial}{\partial x_1} + x_2\frac{\partial}{\partial x_2}\Big)$$

or

$$\mathcal{L} = \beta_A(x_1^2 + x_2^2)\Big(x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}\Big)$$

accordingly and β_A is non-constant. We introduce an extension $A : \mathbb{R}^d \to \mathbb{R}^d$ of A_2 to the whole space setting $A = A_2 \otimes \mathrm{Id}_{d-2}$, where Id_{d-2} is the identity operator on X_{d-2}. The groups $O(2)$ and $O^+(2)$ can be extended on $\mathbb{R}^d$ in the same way. Let

$$\tilde{L}^A = \alpha\Delta + \sum_{i=1}^{d} \tilde{\beta}_i \frac{\partial}{\partial x_i}$$

be the average of L^A with respect the same group which provides a non-constant β_A. Obviously

$$\sum_{i=1}^{2} \tilde{\beta}_i\big|_{\{x_i=0,\ i\geq 3\}} \frac{\partial}{\partial x_i} = \mathcal{L}. \tag{2.20}$$

Now we prove exactly as in Lemma 2.11 that for all Sherman functions $f(x_1)$, p

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^d} \big[P_t(f(x_1))\big]_{x_2x_2} p\, dx\big|_{t=0} \leq 0,$$

where $P_t = e^{t\tilde{L}^A}$. Chose p is such a way that $p = p_1(x_1, x_2)p_{2,\varepsilon}(X)$, where $X = (x_3, \cdots .x_d)$. Let $p_{2,\varepsilon}(X)$ be a sequence of Gaussian probability densities on $\mathbb{R}^2$ such that

$$p_{2,\varepsilon}(X)\, dX \to \delta_0$$

weakly. Note that $\sum_{i=3}^{d} \tilde{\beta}_i \frac{\partial f}{\partial x_i} = 0$. Then using (2.20) we arrive in the limit at the following relation for the two-dimensional differential operator $\mathcal{L}$

$$\int_{\mathbb{R}^2} \big[\mathcal{L}f(x_1)\big]_{x_2x_2} p_1(x_1, x_2)\, dx_1\, dx_2 \leq 0.$$

So, we are again in the two-dimensional situation. Hence by the two-dimensional arguments (see Lemma 2.11 and Lemma 2.15) we get that β_A is constant and β is linear.

By the same arguments the orthogonal projection on every two-dimensional subspace $Y \subset \mathbb{R}^d$ of the vector field $b|_Y$ is linear. Let us show that this fact implies linearity of b. We apply induction in d. For $d = 2$ the claim is trivial. Now suppose that the claim is proved for n-dimensional space. Consider the case $d = n + 1$. First we note that the orthogonal projections p_L of $b|_L$ on every n-dimensional subspace L admits the same property, hence $p_L(b|_L)$ is linear for n-dimensional subspaces by the step of induction. Consider the linear vector field

$$\tilde{b}(x) = \sum_{i=1}^{n+1} x_i b(e_i)$$

for some orthonormal basis $\{e_i\}$. Let us prove that $c = b - \tilde{b} = 0$. Note that $p_L(c|_L)$ is also linear for n-dimensional subspaces. In addition, $c(e_i) = 0$ for every i. Set $L_i = \{x : x_i = 0\}$. Fix $v \in L_i$. One has $p_{L_i}(c(v)) = \sum_{j \neq i} v_j p_{L_i}(c(e_j)) = 0$. Hence $c(v) = \lambda(v)e_i$. In particular, $c|_{L_i \cap L_j} = 0$ for $i \neq j$. Take some $x \in \mathbb{R}^{n+1}$ with $x_i \neq 0$ for all i. One has $x = x_i e_i + v_i$, $v_i \in L_i$. Let L_{i,v_i} be the two-dimensional subspace generated by e_i and v_i. One has $p_{L_{i,v_i}} c(x) = p_{L_{i,v_i}} c(x_i e_i) + p_{L_{i,v_i}} c(v_i) = p_{L_{i,v_i}}(\lambda(v)e_i) = \lambda(v)e_i$. Hence $c(x)$ is orthogonal to $v_i = x - x_i e_i$. Since the same holds for every i, one can easily get that $c(x) = 0$. By the continuity reason $c \equiv 0$. The proof is complete. □

Corollary 2.17. *Under assumptions of Theorem 2.8 the transition probabilities $p(t,x,y)$ of the semigroup T_t for every non-linear drift b are not Sherman functions. In particular, the level sets $\{(x,y) : p(t,x,y) > c\}$ of $p(t,x,y)$ are not convex. In addition, T_t does not preserves even quasi-concave functions.*

Proof. Follows from the Prékopa theorem for Sherman functions and Remark 2.7. □

3. Remarks on the Gaussian correlation inequality

In this section we discuss some applications of Sherman functions to the Gaussian correlation inequality. It is easy to verify that the Gaussian correlation conjecture for sets is equivalent to the following conjecture for functions:

$$\mathrm{Cov}(f,g) \geq 0,$$

where

$$\mathrm{Cov}(f,g) = \int_{\mathbb{R}} fg \, d\gamma - \int_{\mathbb{R}^d} f \, d\gamma \int_{\mathbb{R}^d} g \, d\gamma$$

and $f, g \in \mathcal{S}h^d$.

This problem has been attacked by different methods — geometrical, analytical, variational and probabilistic. Let us briefly discuss below the semigroup approach. Recall that the Ornstein-Uhlenbeck semigroup is given by :

$$T_t f(x) = \int_{\mathbb{R}^n} f(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y) \, d\gamma(y).$$

Using invariance of γ with respect to T_t one obtains the following identity:

$$\mathrm{Cov}(f,g) = -\int_0^\infty \frac{d}{dt} \int_{\mathbb{R}^d} T_t f(x) g(x) \, d\gamma(x) \, dt = \int_0^\infty \int_{\mathbb{R}^d} \langle \nabla T_t f(x), \nabla g(x) \rangle \, d\gamma(x) \, dt. \tag{3.21}$$

Let f, g be log-concave. By the Prékopa theorem $T_t f$ is log-concave. It was shown by Pitt [25] that for every pair of even convex functions V and W on $\mathbb{R}^2$ and every decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ the following inequality holds:

$$\int_{\mathbb{R}^2} \langle \nabla e^{-V}, \nabla e^{-W} \rangle \psi(x^2 + y^2) \, dx \, dy \geq 0.$$

Obviously, the Pitt's inequality, Prékopa theorem and formula (3.21) imply the correlation inequality in the two-dimensional case. The same inequality for $\mathbb{R}^d$ would imply the positive solution of the Gaussian correlation conjecture. Note that if the Pitt's inequality holds, then it obviously holds for Sherman functions. We prove below some special cases of the Pitt's inequality. In particular, we show some links with the well-known Brascamp-Lieb inequality and prove the Gaussian correlation conjecture for some special class of functions.

Finally we note that many interesting results concerning correlation conjecture have been recently obtained by method of optimal transportation (see [31] for details). For example, it was shown in [14] that an even log-concave function e^{-W} and an even convex function V are negatively correlated, i.e.

$$\int_{\mathbb{R}^d} e^{-W} V \, d\gamma \le \int_{\mathbb{R}^d} e^{-W} \, d\gamma \int_{\mathbb{R}^d} V \, d\gamma.$$

The main result of this section is the following:

Theorem 3.1. *Let $e^{-W} : \mathbb{R}^d \to \mathbb{R}$ be even and log-concave and $V(x) = \sum_{i=1}^d f(x_i)$, where every $f_i : \mathbb{R} \to \mathbb{R}^+$ is convex and even. Then for every non-increasing and convex function $G : \mathbb{R}^+ \to \mathbb{R}^+$ the following inequality holds:*

$$\int_{\mathbb{R}^d} e^{-W} G(V) \, d\gamma \ge \int_{\mathbb{R}^d} e^{-W} \, d\gamma \int_{\mathbb{R}^d} G(V) \, d\gamma. \tag{3.22}$$

Unfortunately, we are unable to show inequality (3.22) without the assumption of convexity of G. Otherwise we would get the correlation inequality for every pair of sets A, B, where B is an arbitrary symmetric convex set and A is a set of the type $\{V \le c\}$.

Lemma 3.2. *Let $e^{-W} : \mathbb{R}^d \to \mathbb{R}^+$ be even and log-concave, $\psi : \mathbb{R}^d \to \mathbb{R}^+$ be even and quasi-concave. Suppose in addition that*

$$\psi(x_1, \cdots, x_{i-1}, -x_i, x_{i+1}, \cdots, x_d) = \psi(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d).$$

Let $f(x_i) : \mathbb{R} \to \mathbb{R}^+$ be an even function, increasing on $[0, \infty)$. Then

$$\int_{\mathbb{R}^d} f'(x_i) \frac{\partial W}{\partial x_i} e^{-W} \psi \, dx \ge 0.$$

Proof. It is enough to prove Lemma for the case $\psi = I_A$, where A is a convex set, symmetric about the origin and the hyperplane $\{x_i = 0\}$. Assume without loss of generality that A is given by

$$A = \{x : |x_i| \le \varphi(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_d)\},$$

where φ is a smooth concave function on $\mathbb{R}^{d-1}$. Define $\tilde{\varphi}(x) := \varphi(pr_i x) : \mathbb{R}^d \to \mathbb{R}$, where

$$pr_i : x \to (x_1, \cdots, x_{i-1}, x_{i+1}, x_d).$$

Note that for every $a > 0$ the function

$$\big(I_{[-a,a]}(x_i) - I_{[-a,a]} \circ \tilde{\varphi}\big) I_{(A \cap \mathbb{R}^{d-1}) \times \mathbb{R}}$$

is an indicator of a convex set

$$[-a, a] \times \{\{x_i = a\} \cap A\},$$

hence log-concave. Since $e^{-sf(x_i)}$, $s > 0$ can be uniformly approximated by functions of the type $\sum_{i=1}^k c_i I_{[-a_i, a_i]}$, $c_i, a_i > 0$, we get that

$$F_s = (e^{-sf(x_i)} - e^{-sf \circ \tilde{\varphi}}) I_{(A \cap \mathbb{R}^{d-1}) \times \mathbb{R}}$$

can be approximated by sums of even log-concave functions, hence $F_s \in \mathcal{S}h^d$ and

$$g(t) := \int_A e^{-W(x_1, \cdots, x_{i-1}, x_i + t, x_{i+1}, \cdots, x_d)} F_s \, dx$$

is a Sherman function. Hence $g(t)$ attains its maximum at zero. Consequently

$$0 \geq \int_A \frac{\partial^2 e^{-W}}{\partial^2 x_i} F_s \, dx.$$

Since $\partial A = \{|x_1| = \varphi\}$, we get that $F_s = 0$ on ∂A and one can apply integration by parts:

$$0 \geq -\int_A \frac{\partial e^{-W}}{\partial x_i} \frac{\partial F_s}{\partial x_i} \, dx = -s \int_A \frac{\partial W}{\partial x_i} f'(x_i) e^{-s f(x_i)} e^{-W} \, dx.$$

Dividing on s and tending s to zero we get the desired inequality. □

Lemma 3.2 implies some partial cases of the Pitt's inequality for $\mathbb{R}^d$. In particular, it extends the result of Hargé in ([13]), where he has shown that

$$\int_{\mathbb{R}^d} \langle \nabla e^{-W}, \nabla e^{-Q} \rangle \, d\gamma \geq 0,$$

where e^{-W} is an arbitrary even log-concave function and Q is a non-negative quadratic form. Indeed, choosing an appropriate orthogonal basis for Q we can assume that Q has the diagonal form. Then the claim follows immediately from Lemma 3.2. Another interesting case of the Pitt's inequality is proved in Corollary 3.6.

Proof of Theorem 3.1 Let us apply formula (3.21) to $f = e^{-W}$, $g = G(V)$. By the Prékopa Theorem $T_t f = e^{-W_t}$, where W_t is convex. Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function. Note that the function $\varphi(V) I_{B_R}$, where $B_R = \{\|x\| \leq R\}$ is quasi-concave. Lemma 3.2 implies that

$$\int_{B_R} \frac{\partial V}{\partial x_i} \frac{\partial W_t}{\partial x_i} e^{-W_t} \varphi(V) \, dx \geq 0 \quad \forall i.$$

Clearly, I_{B_R} can be replaced in this inequality by a function $\psi(\|x\|)$, where $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ is a decreasing function. Hence taking $\varphi = -G'$ we get

$$-\int_{\mathbb{R}^d} \langle \nabla V, \nabla W_t \rangle e^{-W_t} G'(V) \, d\gamma \geq 0$$

and the left-hand side of (3.21) is non-negative. The proof is complete. □

Example 3.3. *Let $e^{-W} : \mathbb{R}^d \to \mathbb{R}$ be even and log-concave and $V(x) = \sum_{i=1}^d f(x_i)$, where every $f_i : \mathbb{R} \to \mathbb{R}^+$ is convex and even. Then*

$$\int_{\mathbb{R}^d} e^{-W} e^{-\phi(V)} \, d\gamma \geq \int_{\mathbb{R}^d} e^{-W} \, d\gamma \int_{\mathbb{R}^d} e^{-\phi(V)} \, d\gamma,$$

where $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is concave in increasing.

3.1. **The Brascamp-Lieb inequality and the Pitt's inequality.** We show below some interesting links between the well-known Brascamp-Lieb inequality (see [10]) and the Pitt's inequality.

Everywhere in this subsection $\mu = e^{-\Phi} \, dx$ is a probability measure on $\mathbb{R}^d$ such that Φ is uniformly convex, i.e. $D^2\Phi > cI$ for some positive c, and smooth. For every smooth function f such that $\int_{\mathbb{R}^d} f \, d\mu = 0$ and $\nabla f \in L^2(\mu)$ let us define

$$P_f = \int_0^\infty T_t f \, dt,$$

where T_t is the semigroup on $L^2(\mu)$ generated by $\mathcal{L} = \Delta - \langle \nabla\Phi, \nabla \rangle$. Using the well-known relations $\lim_{t\to\infty} T_t f = 0$, $\int_{\mathbb{R}^d} T_t f \, d\mu = 0$,

$$\left| \nabla T_t f \right| \leq e^{-\frac{t}{c}} T_t \left| \nabla f \right|$$

and the Poincaré inequality

$$\int_{\mathbb{R}^d} (T_t f)^2\, d\mu \le \frac{1}{c}\int_{\mathbb{R}^d} \bigl|\nabla T_t f\bigr|^2\, d\mu$$

(which follows, for example, from the Brascamp-Lieb inequality) one can easily show that P_f is well-defined. In addition, $\nabla P_f \in L^2(\mu)$ and

$$f = \Delta P_f - \bigl\langle \nabla\Phi, \nabla P_f \bigr\rangle = \mathcal{L} P_f.$$

Moreover, by the well-known coercivity identity

$$\int_{\mathbb{R}^d} \Bigl\langle D^2\Phi \nabla P_f, \nabla P_f \Bigr\rangle\, d\mu + \int_{\mathbb{R}^d} \|D^2 P_f\|_2^2\, d\mu = \int_{\mathbb{R}^d} (\mathcal{L}P_f)^2\, d\mu = \int_{\mathbb{R}^d} f^2\, d\mu.$$

Here $\|\cdot\|_2$ is the Hilbert-Schmidt norm. We will employ below the obvious observation that $P_{-\Phi_{x_i}} = x_i$.

The following lemma generalize the Brascamp-Lieb inequaity and implies some partial cases of the Pitt inequality (see Corollary 3.6).

Lemma 3.4. *The following formula holds (we denote for simplicity $P := P_f$):*

$$\begin{aligned}&\int_{\mathbb{R}^d} f^2\, d\mu + \int_{\mathbb{R}^d} \|D^2 P\|_2^2\, d\mu \\ &+ \int_{\mathbb{R}^d} \bigl\langle \bigl(D^2\Phi\bigr)^{-1}\bigl[\nabla(\Delta P) - D^2 P \nabla\Phi\bigr], \nabla(\Delta P) - D^2 P\nabla\Phi \bigr\rangle\, d\mu \\ &= \int_{\mathbb{R}^d} \bigl\langle \bigl(D^2\Phi\bigr)^{-1} \nabla f, \nabla f\bigr\rangle\, d\mu.\end{aligned}$$

In particular, the following inequality holds (Brascamp-Lieb inequality):

$$\int_{\mathbb{R}^d} \Bigl(g - \int_{R^d} g\, d\mu\Bigr)^2 d\mu \le \int_{\mathbb{R}^d} \bigl\langle \bigl(D^2\Phi\bigr)^{-1}\nabla g, \nabla g\bigr\rangle\, d\mu.$$

Proof. Note that

$$\begin{aligned}&\int_{\mathbb{R}^d} f^2\, d\mu - \int_{\mathbb{R}^d} \bigl(\Delta P\bigr)^2\, d\mu = \int_{\mathbb{R}^d} \bigl(f - \Delta P\bigr)\bigl(f + \Delta P\bigr)\, d\mu \\ &= -\int_{\mathbb{R}^d} \bigl\langle \nabla\Phi, \nabla P\bigr\rangle \bigl(2\Delta P - \bigl\langle \nabla\Phi, \nabla P\bigr\rangle\bigr)\, d\mu,\end{aligned}$$

$$\begin{aligned}&\int_{\mathbb{R}^d} \bigl\langle \nabla\Phi, \nabla P\bigr\rangle \Delta P\, d\mu = \int_{\mathbb{R}^d} \bigl\langle \nabla\Phi, \nabla P\bigr\rangle^2\, d\mu \\ &- \int_{\mathbb{R}^d} \bigl\langle D^2\Phi\nabla P, \nabla P\bigr\rangle\, d\mu - \int_{\mathbb{R}^d} \bigl\langle D^2 P\nabla\Phi, \nabla P\bigr\rangle\, d\mu.\end{aligned}$$

Hence

$$\begin{aligned}&\int_{\mathbb{R}^d} \bigl\langle \nabla\Phi, \nabla P\bigr\rangle \Delta P\, d\mu = \frac{1}{2}\int_{\mathbb{R}^d} \bigl\langle \nabla\Phi, \nabla P\bigr\rangle^2\, d\mu \\ &- \frac{1}{2}\int_{\mathbb{R}^d} \bigl\langle D^2\Phi\nabla P, \nabla P\bigr\rangle\, d\mu + \frac{1}{2}\int_{\mathbb{R}^d} \Bigl(\Delta P\bigl\langle \nabla\Phi, \nabla P\bigr\rangle - \bigl\langle D^2 P\nabla\Phi, \nabla P\bigr\rangle\Bigr)\, d\mu\end{aligned}$$

and

$$\begin{aligned}&\int_{\mathbb{R}^d} f^2\, d\mu - \int_{\mathbb{R}^d} \bigl(\Delta P\bigr)^2\, d\mu = \\ &\int_{\mathbb{R}^d} \bigl\langle D^2\Phi\nabla P, \nabla P\bigr\rangle\, d\mu - \int_{\mathbb{R}^d} \Bigl(\Delta P\bigl\langle \nabla\Phi, \nabla P\bigr\rangle\, d\mu - \bigl\langle D^2 P\nabla\Phi, \nabla P\bigr\rangle\Bigr)\, d\mu.\end{aligned}$$

Now let us note that

$$\int_{\mathbb{R}^d} f^2\,d\mu = \int_{\mathbb{R}^d} f\mathcal{L}P\,d\mu = -\int_{\mathbb{R}^d} \bigl\langle \nabla f, \nabla P\bigr\rangle\,d\mu.$$

Hence

$$\frac{1}{2}\int_{\mathbb{R}^d} f^2\,d\mu + \frac{1}{2}\int_{\mathbb{R}^d} (\Delta P)^2\,d\mu =$$
$$-\int_{\mathbb{R}^d} \bigl\langle \nabla f, \nabla P\bigr\rangle\,d\mu - \frac{1}{2}\int_{\mathbb{R}^d}\bigl\langle D^2\Phi\nabla P, \nabla P\bigr\rangle\,d\mu$$
$$\frac{1}{2}\int_{\mathbb{R}^d}\Bigl(\Delta P\bigl\langle\nabla\Phi, \nabla P\bigr\rangle\,d\mu - \bigl\langle D^2P\nabla\Phi, \nabla P\bigr\rangle\Bigr)\,d\mu$$

and

$$-2\int_{\mathbb{R}^d}\bigl\langle \nabla f, \nabla P\bigr\rangle\,d\mu - \int_{\mathbb{R}^d}\bigl\langle D^2\Phi\nabla P, \nabla P\bigr\rangle\,d\mu = \int_{\mathbb{R}^d}\bigl\langle (D^2\Phi)^{-1}\nabla f, \nabla f\bigr\rangle\,d\mu$$
$$-\int_{\mathbb{R}^d}\bigl\langle (D^2\Phi)^{-1}\bigl[\nabla f + D^2\Phi\nabla P\bigr], \nabla f + D^2\Phi\nabla P\bigr\rangle\,d\mu.$$

Since

$$\nabla f + D^2\Phi\nabla P = \nabla(\Delta P) - D^2P\nabla\Phi,$$

we obtain

$$\int_{\mathbb{R}^d} f^2\,d\mu + \int_{\mathbb{R}^d}(\Delta P)^2\,d\mu$$
$$+\int_{\mathbb{R}^d}\bigl\langle (D^2\Phi)^{-1}\bigl[\nabla(\Delta P) - D^2P\nabla\Phi\bigr], \nabla(\Delta P) - D^2P\nabla\Phi\bigr\rangle\,d\mu$$
$$= \int_{\mathbb{R}^d}\bigl\langle (D^2\Phi)^{-1}\nabla f, \nabla f\bigr\rangle\,d\mu + \int_{\mathbb{R}^d}\Bigl(\Delta P\bigl\langle\nabla\Phi, \nabla P\bigr\rangle\,d\mu - \bigl\langle D^2P\nabla\Phi, \nabla P\bigr\rangle\Bigr)\,d\mu.$$

Finally, integrating by parts one easily gets the following identity:

$$\int_{\mathbb{R}^d}\Bigl(\Delta P\bigl\langle\nabla\Phi, \nabla P\bigr\rangle\,d\mu - \bigl\langle D^2P\nabla\Phi, \nabla P\bigr\rangle\Bigr)\,d\mu = \int_{\mathbb{R}^d}(\Delta P)^2\,d\mu - \int_{\mathbb{R}^d}\|D^2P\|_2^2\,d\mu.$$

The proof is complete. □

Remark 3.5. *Note that result of Lemma 3.4 can be proved by optimal transportation methods. More precisely, Lemma 3.4 follows from the integral relation between the entropy $\int_{\mathbb{R}^d} F\log F\,d\mu$ and the quadratic optimal transportation mapping T sending a probability measure $F\cdot\mu$ to μ (see* [21] *Theorem 2.1, Theorem 2.2 and* [3]*). For the proof one applies this relation to the family of measures $\mu_\varepsilon = \exp\bigl(\varepsilon f - \log\int_{\mathbb{R}^d} e^{\varepsilon f}\bigr)\,d\mu$ and the corresponding family of optimal mappings T_ε pushing forward μ_ε to μ. One takes into account that $T_\varepsilon = \nabla V_\varepsilon$ for some convex V_ε solving a nonlinear Monge-Amperé equation. The result follows from the Taylor expansion in ε and linearization of the corresponding Monge-Amperé equation.*

Corollary 3.6. *Let γ be the standard Gaussian measure and V and W be smooth convex even functions such that $\|\nabla V\|^2 e^{-V}$, $\|\nabla W\|^2 e^{-W} \in L^1(\gamma)$. Suppose in addition that $Q(x)$ is a non-negative quadratic function and there exist numbers $\lambda_1, \lambda_2 > 0$ such that $\lambda_1 V + \lambda_2 W = Q(x)$. Then*

$$\int_{\mathbb{R}^d}\bigl\langle\nabla V, \nabla W\bigr\rangle e^{-V}e^{-W}\,d\gamma \geq 0.$$

Proof. Choosing an appropriate orthonormal basis one can assume without loss of generality that $Q(x) = \sum_{i=1}^d \alpha_i x_i^2$. We will apply the previous result to the probability measure $\mu = Ce^{-V}e^{-W}\cdot\gamma$, where C is a normalizing constant, and functions V_{x_i}, W_{x_i}.

Since $\lambda_1 V_{x_i} + \lambda_2 W_{x_i} = \alpha_i x_i$, it can be verified directly that

$$(\alpha_i+\lambda_1)P_{V_{x_i}} + (\alpha_i+\lambda_2)P_{W_{x_i}} = P_{(\alpha_i+\lambda_1)V_{x_i}+(\alpha_i+\lambda_2)W_{x_i}} = -\alpha_i x_i.$$

Hence

$$(\alpha_i+\lambda_1)D^2P_{V_{x_i}} = -(\alpha_i+\lambda_2)D^2P_{W_{x_i}}. \tag{3.23}$$

Applying Lemma 3.4 one gets

$$\begin{aligned}
&\int_{\mathbb{R}^d} V_{x_i}W_{x_i}\,d\mu + \int_{\mathbb{R}^d}\big\langle D^2P_{V_{x_i}}, D^2P_{W_{x_i}}\big\rangle_2\,d\mu \\
&+ \int_{\mathbb{R}^d}\Big\langle \big(D^2\Phi\big)^{-1}\big[\nabla(\Delta P_{V_{x_i}}) - D^2P_{V_{x_i}}\nabla\Phi\big], \nabla(\Delta P_{W_{x_i}}) - D^2P_{W_{x_i}}\nabla\Phi\Big\rangle\,d\mu \\
&= \int_{\mathbb{R}^d}\Big\langle \big(D^2\Phi\big)^{-1}\nabla V_{x_i}, \nabla W_{x_i}\Big\rangle\,d\mu,
\end{aligned}$$

where $\Phi = \frac{\|x\|^2}{2} + V(x) + W(x)$. It follows from (3.23) that

$$\int_{\mathbb{R}^d}\Big\langle \big(D^2\Phi\big)^{-1}\big[\nabla(\Delta P_{V_{x_i}}) - D^2P_{V_{x_i}}\nabla\Phi\big], \nabla(\Delta P_{W_{x_i}}) - D^2P_{W_{x_i}}\nabla\Phi\Big\rangle\,d\mu \le 0,$$

and

$$\int_{\mathbb{R}^d}\big\langle D^2P_{V_{x_i}}, D^2P_{W_{x_i}}\big\rangle_2\,d\mu \le 0.$$

Hence

$$\int_{\mathbb{R}^d} V_{x_i}W_{x_i}\,d\mu \ge \int_{\mathbb{R}^d}\Big\langle \big(I + D^2V + D^2W\big)^{-1}\nabla V_{x_i}, \nabla W_{x_i}\Big\rangle\,d\mu.$$

Finally

$$\int_{\mathbb{R}^d}\langle \nabla V, \nabla W\rangle e^{-V}e^{-W}\,d\gamma \ge \int_{\mathbb{R}^d}\mathrm{Tr} D^2V\big(I + D^2V + D^2W\big)^{-1}D^2W e^{-V}e^{-W}\,d\gamma.$$

Let us show that the latter is non-negative. It is enough to show that

$$\mathrm{Tr} A\big(I + A + B\big)^{-1}B \ge 0$$

if matrices A and B are non-negative. Without loss of generality we can assume that A and B are positive. Then

$$A\big(I + A + B\big)^{-1}B = \big(A^{-1} + \big(I + B\big)^{-1}\big)^{-1}\big(B^{-1} + I\big)^{-1}.$$

Since the inverse matrix of a positive matrix is also positive, this formula represents $A\big(I + A + B\big)^{-1}B$ as a product of two positive matrices $\tilde{A}$, $\tilde{B}$. One has $\mathrm{Tr}\tilde{A}\tilde{B} = \mathrm{Tr}\tilde{B}^{1/2}\tilde{A}\tilde{B}^{1/2} \ge 0$, since $\tilde{B}^{1/2}\tilde{A}\tilde{B}^{1/2} \ge 0$. The proof is complete. □

If in the proof above we replace the standard Gaussian measure by the Lebesgue measure we obtain the following inequality:

$$\int_{\mathbb{R}^d} V_{x_i}W_{x_i}e^{-V-W}\,dx \ge 0,$$

where W and V are convex even functions (this is a partial case of (2.2)).

Remark 3.7. *By Lemma 3.2*

$$\int_{\mathbb{R}^d}\langle \nabla V, Q\rangle e^{-V}e^{-Q}\,d\gamma \ge 0$$

for every convex even V and every nonnegative quadratic Q. Hence

$$\int_{\mathbb{R}^d}\langle \nabla V, \nabla W\rangle e^{-V}e^{-W}\,d\gamma \ge 0$$

holds for convex even functions V, W satisfying $W = V + \mu Q$ for some $\mu \geq 0$. Hence, the Pitt's inequality holds for every pair of convex even functions V, W satisfying $\lambda_1 V + \lambda_2 W = Q$ and any numbers λ_1, λ_2.

3.2. **Correlation inequality for mixtures of Gaussian measures.** We have already seen that there exist non-Gaussian measures satisfying the correlation inequality (see Example 2.5). We briefly discuss below a simple proof of this fact and prove some more general result in this direction.

Let us denote by $\mathcal{M}$ the set of symmetric probability measures satisfying the correlation inequality:

$$\mu(A \cap B) \geq \mu(A)\mu(B), \tag{3.24}$$

where A and B are symmetric convex sets. We denote by $\mathcal{M}_e$ the set of measures satisfying (3.24) under additional assumption that one of the sets is an ellipsoid. Recall that according to results of Pitt and Harge $\gamma \in \mathcal{M}$ for $d = 2$ and $\gamma \in \mathcal{M}_e$ for every d. Denote by $V(A)$ the standard Lebesgue volume of the set $A \subset \mathbb{R}^d$.

Denote by Λ the set of functions which admit the following representation:

$$f(\lambda) = \int_0^\infty e^{-\lambda t} \, d\nu(t),$$

i.e. f is a Laplace transform of a positive measure ν on $[0, \infty)$.

Proposition 3.8. *a) Let μ be a probability measure on $\mathbb{R}^2$ with density $f(x^2 + y^2)$ or $p(x^2)q(y^2)$, where $f, p, q \in \Lambda$. Then $\mu \in \mathcal{M}$.*

b) Let $\mu = \prod_{i=1}^k \mu_i$ be a product of probability measures μ_i. If every μ_i is a probability measure on $\mathbb{R}^{k_i}$ with density $f_i(\sum_{j=1}^{k_i} x_j^2)$, where $f_i \in \Lambda$, then $\mu \in \mathcal{M}_e$.

Lemma 3.9. *Let $\mu_1, \mu_2, \cdots, \mu_k$ be probability measures such that for a couple of sets A, B and for every i, j, $i < j$*

$$\mu_i(A \cap B) \geq \mu_i(A)\mu_i(B), \quad \mu_i(A) \leq \mu_j(A), \quad \mu_i(B) \leq \mu_j(B).$$

Then $\mu(A \cap B) \geq \mu(A)\mu(B)$, where

$$\mu = \sum_{i=1}^k \lambda_i \mu_i, \quad 0 < \lambda_i < 1 \quad \sum_{i=1}^k \lambda_i = 1$$

is a convex combination of μ_i.

Proof. Let us prove the claim by induction in k. We have for $k = 2$:

$$\begin{aligned}
&\lambda\mu_1(A \cap B) + (1-\lambda)\mu_2(A \cap B) \\
&- \Big[\lambda\mu_1(A) + (1-\lambda)\mu_2(A)\Big]\Big[\lambda\mu_1(B) + (1-\lambda)\mu_2(B)\Big] \geq \\
&\lambda\mu_1(A)\mu_1(B) + (1-\lambda)\mu_2(A)\mu_2(B) \\
&- \Big[\lambda\mu_1(A) + (1-\lambda)\mu_2(A)\Big]\Big[\lambda\mu_1(B) + (1-\lambda)\mu_2(B)\Big] = \\
&\lambda(1-\lambda)\Big(\mu_1(A) - \mu_2(A)\Big)\Big(\mu_1(B) - \mu_2(B)\Big) \geq 0.
\end{aligned}$$

Let us prove the $k+1$ step. If $\mu = \sum_{i=1}^{k+1} \lambda_i \mu_i$, then $\mu = (\sum_{i=1}^k \lambda_i)\nu + \lambda_{k+1}\mu_{k+1}$, where $\nu = (\frac{1}{\sum_{i=1}^k \lambda_i}) \sum_{i=1}^k \lambda_i \mu_i$. By the k-th step of induction the correlation inequality holds for ν, A, B. In addition

$$\nu(A) \leq \frac{\sum_{i=1}^k \lambda_i \mu_i(A)}{\sum_{i=1}^k \lambda_i} \leq \mu_{k+1}(A).$$

We have reduced the proof to the case $k = 2$. □

Corollary 3.10. *Let $A, B \subset \mathbb{R}^d$ be symmetric sets and suppose that A is starshaped, i.e. contains all the sets of the type $\{tx, x \in A\}$ for all $0 \le t \le 1$. Let μ be a probability measure on $\mathbb{R}^d$ given by its density f such that every set $A_c = \{f > c\}$ has the form $A_c = k(c)A$ for some $k(c) \ge 0$. Suppose that the function $t \in [0, \infty) \to V(At \cap B)$ increases. Then the correlation inequality holds for A, B, μ.*

Proof. It is enough to verify the inequality for the case $\mu_{A_c} := \frac{1}{V(A_c)} I_{A_c}\, dx$ of the normalized Lebesque measure on A_c, approach μ by convex combination of μ_{A_c} and apply Lemma 3.9. □

For the proof of Proposition 3.8 we introduce the following partial order on the set of measures: we say that $\mu \le \nu$, where μ, ν are symmetric probability measures if $\mu(A) \le \nu(A)$ for every symmetric convex set A (or $\int_{\mathbb{R}^d} f\, d\mu \le \int_{\mathbb{R}^d} f\, d\nu$ for every even log-concave f).

Proof of Proposition 3.8:

a) Note that the correlation inequality holds for every Gaussian measure on the plane. In addition, if γ_{t_1} and γ_{t_2} are centered Gaussian measures with covariance matrices $t_1 E$ and $t_2 E$, then $\gamma_{t_1} \le \gamma_{t_2}$ if $t_1 \ge t_2$. Then the first claim follows from Lemma 3.9.

Let the measures μ_1 and μ_2 have the form $\mu_1 = p(x)q_1(y)$ and $\mu_2 = p(x)q_2(y)$ where p, q_1, q_2 are one-dimensional symmetric probability measures with log-concave densities. By the Prékopa theorem $\mu_1 \ge \mu_2$ if $q_1\, dy \ge q_2\, dy$. In particular, if μ_1, μ_2 are Gaussian, Lemma 3.9 implies that the correlation inequality holds for the probability measure with density $C(s)e^{-sx^2} \int_0^\infty e^{-ty^2}\, d\nu(t)$, where ν is some positive measure. We repeat this procedure with the term e^{-sx^2} and obtain the claim.

b) Note that $\mathcal{M}_e$ includes all finite-dimensional Gaussian measures. The proof is similar to a) and we omit it here. □

Example 3.11. *The following probability measures on $\mathbb{R}^d$ belong to $\mathcal{M}_e$ and to $\mathcal{M}$ if $d = 2$.*

$$C(d,\alpha)\frac{dx}{(1+r^2)^\alpha},\ \alpha > d\ (Student's\ distribution)$$

$$C(d,\alpha)e^{-r^{2\alpha}}\, dx,\ c^d e^{-\sum_{i=1}^d x_i^{2\alpha}}\, dx,\ 0 < \alpha < 1$$

$$\frac{1}{\pi^d}\prod_{i=1}^d \frac{dx}{1+x_i^2}$$

(i.e. the law of $(\xi_1, \cdots, \xi_d)$, where ξ_i are i.i.d. Cauchy distributions).

Symmetric stable measures also belong to $\mathcal{M}_e$ and to $\mathcal{M}$ if $d = 2$. It follows from the fact that they can be represented as mixtures of Gaussian measures (see [24]*).*

Acknowledgement

The author thanks Prof. Vladimir Bogachev for fruitful discussions and remarks. This research was supported by the RFBR project 04-01-00748, the DFG Grant 436 RUS 113/343/0(R). The paper was finished during the author's stay in Pisa, where he was supported by the Junior Visiting Position of the Centro di Ricerca Matematica Ennio De Giorgi, Scuole Normale Superiore di Pisa.

References

[1] Barthe F. Autour de l'inegalité de Brunn-Minkowski. — Ann. Fac. Sci. Toulouse Math. XII, 2003, p. 127–178.

[2] Bogachev V.I. Gaussian Measures, Amer. Math. Soc., Rhode Island, 1998.

[3] Bogachev V.I., Kolesnikov A.V., Medvedev K.V. Triangular transformations of measures. — Matem. Sbornik, 2005, v. 196, n. 3, p. 3–30 (in Russian); English transl.: Sbornik Math., 2005, v. 196, n. 3.

[4] Bogachev V., Röckner M., Schmuland B. Generalized Mehler semigroups and applications. — Probab. Theory Relat. Fields, 1996, v. 105, n. 2, p. 193–225.

[5] Borell C. A Gaussian correlation inequality for certain bodies in $\mathbb{R}^d$. — Math. Ann., 1981, v. 256, p. 569–573.

[6] Borell C. Greenian potential and concavity. — Math. Ann., 1985, v. 272, p. 155–160.

[7] Borell C. Brownian motion with negative drift and convex level sets in space-time. — Probab. Theory Rel. Fields, 1991, n. 87, p. 403–409.

[8] Borell C. Geometric properties of some familiar diffusions in $\mathbb{R}^n$. — Ann. Probab., 1993, v. 21, n. 3, p. 482–489.

[9] Borell C. A note on parabolic convexity and heat conduction. — Ann. Inst. Henry Poincaré, 1996, v. 32, n. 3, p. 387–393.

[10] Brascamp H., Lieb E.H. On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. — J. Funct. Anal., 1976, v. 22, p. 366–389.

[11] Fuhrmann M., Röckner M. Generalized Mehler semigroups: The Non-Gaussian Case. — Potential analysis, 2000, v. 12, p. 1–47.

[12] Gikhman I.I, Skorokhod A.V. Stochastic differential equations, Springer-Verlag, 1972.

[13] Hargé G. A particular case of correlation inequality for the Gaussian measure. — Ann. Probab., 1999, v. 27, p. 1939–1951.

[14] Hargé G. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. — Probab. Theory Related Fields, 2004, v. 13, n. 3, p. 415–440.

[15] Herbst I., Pitt L. Diffusion equation techniques in stochastic monotonicity and positive correlations. — Probab. Theory Rel.Fields, 1991, v. 87, p. 275–312.

[16] Hunt G. A. Semi-groups of measures on Lie groups. — Trans. Amer. math. Soc., 1956, v. 81, p. 264–293.

[17] Kawohl B. Starshapedness of level sets for the obtacle problem and for the capacitory potential problem. — Proc. Amer. Math. Soc.,1983, v. 89, n. 4, p. 637–640.

[18] Kawohl B. When are superharmonic functions concave? Application to the St. Venant torsion problem and to the fundamental mode of the clamped membrane. — Z. Angew. Math. Mech., 1984, v. 64, p. 364–366.

[19] Kennington A. Power concavity and boundary value problems. — Indiana Univ. Math. Jour.,1985, v. 34, n. 3, p. 687–704.

[20] Kolesnikov A.V. On the diffusion semigroups preserving the log-concavity. — J. Funct. Anal., 2001, v. 186, p. 196–205.

[21] Kolesnikov A.V. Convexity inequalities and optimal transport of infinite-dimensional measures. — J. Math. Pures Appl., 2004, v. 83, n. 11, p. 1373–1404.

[22] Korevaar N. J. Convex solutions to nonlinear elliptic and parabolic boundary value problem. — Indiana Univ. Math. Jour., 1983, v. 32, n. 4, p. 603–614.

[23] Ledoux M. The concentration of measure phenomenon. Mathematical Surveys and Monographs 89. Amer. Math. Soc., 2001.

[24] Linde W., Probability in Banach spaces – stable and infinitely divisible distributions, Chichester; New York : Wiley, 1986.

[25] Pitt L. D. A Gaussian correlation inequality for symmetric convex sets. — Ann. Probab., 1977, v. 5, p. 470–474.

[26] Prékopa A. On logarithmic concave measures and functions. — Acta Sci. Math., 1973, v. 34, p. 335–343.

[27] Reed M., Simon B. Methods of modern mathematical physics. II. Fourier analysis, self-adjointness. New York, Academic Press, 1975.

[28] Schechtman G., Schlumprecht T., Zinn J. On the Gaussian measure of the intersection of the symmetric, convex sets. — Ann. Probab., 1998, v. 26, p. 346–357.

[29] Sherman S. A theorem on convex sets with applications. — Ann. Math. Statist., 1955, v. 26, p. 763–766.

[30] Sǐdák Z., Rectangular confidence regions for the mean of multivariate normal distributions. — J. Amer. Statist. Assoc., 1967, v. 62, p. 626–623.

[31] Villani C. Topics in Optimal Transportation. Amer. Math. Soc., Providence, Rhode Island, 2003.

Scuola Normale Superiore, Centro di Ricerca Matematica Ennio De Giorgi, I-56100 Pisa, Italy, E-mail: sascha77@mail.ru