A note on Gaussian correlation inequalities for nonsymmetric sets

Adrian P. C. Lim[a], Dejun Luo[a,b*]

[a]UR Mathématiques, Université de Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg
[b]Key Lab of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

Abstract

We consider the Gaussian correlation inequality for nonsymmetric convex sets. More precisely, if $A \subset \mathbb{R}^d$ is convex and the origin $0 \in A$, then for any ball B centered at the origin, it holds $\gamma_d(A \cap B) \geq \gamma_d(A)\gamma_d(B)$, where γ_d is the standard Gaussian measure on $\mathbb{R}^d$. This generalizes Proposition 1 in [Arch. Rational Mech. Anal. 161 (2002), 257–269].

Keywords: Correlation inequality, Gaussian measure, convexity, log-concavity, optimal transport

MSC 2000: 60E15, 26A51

1 Introduction

Let $A, B \subset \mathbb{R}^d$ be symmetric convex subsets. The Gaussian correlation inequality claims that

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

where γ_d is the standard Gaussian measure on $\mathbb{R}^d$. The case $d = 1$ is trivial, since A and B are centered intervals, hence one is contained in the other. The case $d = 2$ was proved by Pitt [7]. For higher dimensional cases, there are only partial results. For instance, in [8] it was shown that (1.1) holds if A and B are ellipsoids, which was soon generalized by Hargé [3] to allow one of them to be an arbitrary symmetric convex set. Hargé's proof relies on the modified Ornstein-Uhlenbeck semigroup and the properties of log-concave functions. A rather short proof of Hargé's result was presented in [2], by making use of the deep results in the theory of optimal transport, that is, the optimal transport map, which pushes forward the Gaussian measure γ_d to a probability measure ν having a log-concave density with respect to γ_d, is a contraction (see [1]). In [4], the author obtained a correlation inequality for the Gaussian measure via a formula for Itô-Wiener chaos expansion. Li W.V. [5] presented a weaker form of the correlation inequality (1.1), which is useful to show the existence of small ball constants. For a more detailed survey of the studies on (1.1), see [6, Section 2.4].

In this note we consider two special cases of the correlation inequality. It is clear that we only need to consider bounded subsets of $\mathbb{R}^d$. In the sequel, we always assume that the sets are bounded and closed. First we prove the following result.

*Email: luodj@amss.ac.cn

Theorem 1.1. *Let* $\mathrm{d}\mu = \rho(|x|)\,\mathrm{d}x$ *be a probability measure on* $\mathbb{R}^d$ *with* $\rho \in C(\mathbb{R}_+, (0,\infty))$. *Suppose* $A \subset \mathbb{R}^d$ *is convex and the origin* $0 \in A$. *Then for any ball* B *centered at the origin, we have*

$$\mu(A \cap B) \geq \mu(A)\mu(B).$$

Clearly the Gaussian measure γ_d is a special case of μ considered above. The new point here is that the set A does not have to be symmetric, at the price of the regularity on B. Theorem 1.1 is a slight generalization of [2, Proposition 1]; the latter requires that the origin 0 is the unique fixed point of all the isometries which leave A (globally) invariant. Remark also that our proof (see Section 2) uses purely elementary analysis, while the one in [2] relies on the result in the theory of optimal transport.

If we want to prove the correlation inequality for more general sets B other than the balls (e.g. the ellipsoids), then some additional conditions have to be imposed on the set A.

Theorem 1.2. *Assume that* μ *is a product probability measure:* $\mu = \prod_{i=1}^d \mu_i$, *where* $\mathrm{d}\mu_i = \rho_i(|x_i|)\,\mathrm{d}x_i$ *with* $\rho_i \in C(\mathbb{R}_+, (0,\infty))$. *Let* $A \subset \mathbb{R}^d$ *be a convex set with the following property:* $x \in A$ *implies that its projections on all the coordinate hyperplanes also belong to* A. *Then for any ellipsoid*

$$B = \left\{ x \in \mathbb{R}^d : \frac{x_1^2}{a_1^2} + \cdots + \frac{x_d^2}{a_d^2} \leq 1 \right\},$$

where $a_1, \cdots, a_d$ *are positive constants, we have* $\mu(A \cap B) \geq \mu(A)\,\mu(B)$.

This result will be proved in Section 3. It is easy to see that the set A considered in Theorem 1.2 contains the origin 0. An example for the set A is $\{x = (x_1, \cdots, x_d) \in \mathbb{R}^d : \forall\, i = 1, \cdots, d,\ x_i \geq 0 \text{ and } \sum_{i=1}^d x_i \leq 1\}$. We would like to mention that Theorem 1.2 still holds for more general set B, see Remark 3.3.

A nonnegative function $f : \mathbb{R}^d \to \mathbb{R}_+$ is called log-concave if for any $x, y \in \mathbb{R}^d$ and $0 < \lambda < 1$, $f(\lambda x + (1-\lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda}$. For any convex set $A \subset \mathbb{R}^d$, one easily concludes that the indicator function $\mathbf{1}_A$ is log-concave. The Gaussian correlation inequality (1.1) has the following functional version: for any log-concave and symmetric functions f, g, it holds

$$\gamma_d(fg) \geq \gamma_d(f)\gamma_d(g). \tag{1.2}$$

Here $\gamma_d(f) = \int_{\mathbb{R}^d} f\,\mathrm{d}\gamma_d$. Following the method of [2, Section 3], we show in the last section that (1.2) holds if f is log-concave and $g = \varphi(\langle \Sigma x, x\rangle)$, where $\varphi \in C(\mathbb{R}_+, \mathbb{R}_+)$ is a decreasing function and Σ is a positive definite matrix. So we give an alternative proof to [3, Theorem 2]. By an approximation argument, we obtain again Hargé's result.

2 Proof of Theorem 1.1

In this section we prove Theorem 1.1. In fact we will prove a more general result. To this end, we introduce a class of functions on $\mathbb{R}^d$:

$$\mathcal{C}_d = \big\{ f \in C_c(\mathbb{R}^d, \mathbb{R}_+) : \forall\, c > 0, \{f > c\} \text{ is convex and } \forall\, x \in \mathbb{R}^d, f(x) \leq f(0) \big\}. \tag{2.1}$$

Let S^{d-1} be the unit sphere in $\mathbb{R}^d$. For a bounded measurable function g on $\mathbb{R}^d$, define $\mu(g) = \int_{\mathbb{R}^d} g\,\mathrm{d}\mu$. We have the following simple observations.

Lemma 2.1. *Let* $f \in \mathcal{C}_d$ *and* $f \neq 0$. *Then*

(i) *for any* $\theta \in S^{d-1}$, *the function* $t \mapsto f(t\theta)$ *is decreasing on* $\mathbb{R}_+$;

(ii) $f(0) > \mu(f)$.

Proof. (i) Suppose that there are $t_1 < t_2$ such that $f(t_1\theta) < f(t_2\theta)$. Consider the set $E := \big\{f > [f(t_1\theta) + f(t_2\theta)]/2\big\}$. Then $0, t_2\theta \in E$ but $t_1\theta \in E^c$, which contradicts the fact that E is convex.

(ii) By the definition of the class $\mathcal{C}_d$,

$$f(0) - \mu(f) = \int_{\mathbb{R}^d} (f(0) - f(x))\,\mathrm{d}\mu(x) \geq 0.$$

If $f(0) = \mu(f)$, then $f(x) \equiv f(0)$ for all $x \in \mathbb{R}^d$, which is impossible. Hence $f(0) > \mu(f)$. $\square$

Now we prove

Theorem 2.2. *Assume that $\mathrm{d}\mu = \rho(|x|)\,\mathrm{d}x$ is a probability measure on $\mathbb{R}^d$ with $\rho \in C(\mathbb{R}_+, (0,\infty))$. For any $f \in \mathcal{C}_d$ and any ball B centered at the origin, it holds*

$$\mu(f\mathbf{1}_B) \geq \mu(f)\,\mu(B).$$

Proof. Obviously we can assume $\mu(f) > 0$. For $t \geq 0$, let B_t be the ball centered at the origin with radius t. Define the function

$$\Phi(t) = \mu(f\mathbf{1}_{B_t}) - \mu(f)\mu(B_t), \quad t \geq 0.$$

First we show that Φ is positive when t is sufficiently small and large. By Lemma 2.1(ii), there is $t_0 > 0$ such that for all $x \in B_{t_0}$, $f(x) > \mu(f)$. Thus for any $t \in (0, t_0)$,

$$\Phi(t) = \int_{B_t} \big[f(x) - \mu(f)\big]\,\mathrm{d}\mu(x) > 0.$$

When t is big enough such that $\mathrm{supp}(f) \subset B_t$, we have

$$\Phi(t) = \mu(f) - \mu(f)\mu(B_t) = \mu(f)\mu(B_t^c) > 0.$$

Next we compute the derivative $\Phi'(t)$. We have for $h > 0$,

$$\mu(f\mathbf{1}_{B_{t+h}}) - \mu(f\mathbf{1}_{B_t}) = \int_{B_{t+h}\setminus B_t} f(x)\,\mathrm{d}\mu(x) = \int_{B_{t+h}\setminus B_t} f(x)\rho(|x|)\,\mathrm{d}x.$$

Using the spherical coordinate, the above equality can be written as

$$\begin{aligned}
\mu(f\mathbf{1}_{B_{t+h}}) - \mu(f\mathbf{1}_{B_t}) &= \int_t^{t+h} \bigg(\int_{S^{d-1}} f(r\theta)\rho(|r\theta|)\,\mathrm{d}\sigma(\theta)\bigg) r^{d-1}\,\mathrm{d}r \\
&= \int_t^{t+h} \bigg(\int_{S^{d-1}} f(r\theta)\,\mathrm{d}\sigma(\theta)\bigg)\rho(r) r^{d-1}\,\mathrm{d}r,
\end{aligned}$$

where σ is the volume measure on S^{d-1}. Since the functions f and ρ are continuous, dividing both sides by h and letting $h \to 0$ lead to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu(f\mathbf{1}_{B_t}) = \rho(t)t^{d-1}\int_{S^{d-1}} f(t\theta)\,\mathrm{d}\sigma(\theta).$$

Similarly we have $\frac{\mathrm{d}}{\mathrm{d}t}\mu(B_t) = \rho(t)t^{d-1}\sigma(S^{d-1})$. Therefore

$$\begin{aligned}
\Phi'(t) &= \rho(t)t^{d-1}\int_{S^{d-1}} f(t\theta)\,\mathrm{d}\sigma(\theta) - \mu(f)\rho(t)t^{d-1}\sigma(S^{d-1}) \\
&= \rho(t)t^{d-1}\int_{S^{d-1}} \big[f(t\theta) - \mu(f)\big]\,\mathrm{d}\sigma(\theta). \qquad (2.2)
\end{aligned}$$

From this expression, it is clear that Φ' is continuous. For $t > 0$ small enough, we conclude from (2.2) and Lemma 2.1(ii) that $\Phi'(t) > 0$. Let $t_1 = \inf\{t > 0 : \Phi'(t) = 0\}$. Then $\int_{S^{d-1}} \big[f(t_1\theta) - \mu(f)\big]\, \mathrm{d}\sigma(\theta) = 0$. By Lemma 2.1(i), for any $t > t_1$,

$$\int_{S^{d-1}} \big[f(t\theta) - \mu(f)\big]\, \mathrm{d}\sigma(\theta) \le \int_{S^{d-1}} \big[f(t_1\theta) - \mu(f)\big]\, \mathrm{d}\sigma(\theta) = 0.$$

Hence $\Phi'(t) \le 0$. It follows that $\Phi(t)$ is increasing on $[0, t_1]$ and decreasing on $[t_1, \infty)$. Combining this with the fact that $\Phi(t) > 0$ when t is sufficiently small and large, we complete the proof. $\square$

Now we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. We will construct a sequence of functions, belonging to $\mathcal{C}_d$, which converge to $\mathbf{1}_A$. Let $\mathrm{dist}(\cdot, A)$ be the distance function to A. For $n \ge 1$, define

$$f_n(x) = 1 - n\big[n^{-1} \wedge \mathrm{dist}(x, A)\big], \quad x \in \mathbb{R}^d. \tag{2.3}$$

Then it is clear that $f_n \in C_c(\mathbb{R}^d, \mathbb{R}_+)$, $0 \le f_n \le 1$ and the restriction $f_n|_A \equiv 1$. Since $0 \in A$, we have for all $x \in \mathbb{R}^d$, $f_n(x) \le 1 = f_n(0)$. It remains to show that for any $c \in [0, 1)$, $\{f_n > c\}$ is convex. In fact,

$$\{f_n > c\} = \{x \in \mathbb{R}^d : \mathrm{dist}(x, A) < (1 - c)/n\}.$$

If $x, y \in \{f_n > c\}$, then $\mathrm{dist}(x, A) \vee \mathrm{dist}(y, A) < (1 - c)/n$. Thus there are $x_0, y_0 \in A$ such that $|x - x_0| \vee |y - y_0| < (1 - c)/n$. For any $\lambda \in (0, 1)$, we have by the convexity of A, $\lambda x_0 + (1 - \lambda)y_0 \in A$. Moreover,

$$\big|(\lambda x + (1 - \lambda)y) - (\lambda x_0 + (1 - \lambda)y_0)\big| \le \lambda|x - x_0| + (1 - \lambda)|y - y_0| < (1 - c)/n.$$

Therefore $\mathrm{dist}\big(\lambda x + (1 - \lambda)y, A\big) < (1 - c)/n$; equivalently, $\lambda x + (1 - \lambda)y \in \{f_n > c\}$. This means that $\{f_n > c\}$ is convex.

Now applying Theorem 2.2 to f_n, we have

$$\mu(f_n \mathbf{1}_B) \ge \mu(f_n)\, \mu(B), \quad \text{for all } n \ge 1.$$

Since $f_n \downarrow \mathbf{1}_A$ on $\mathbb{R}^d$, by the monotone convergence theorem, letting $n \to \infty$ completes the proof. $\square$

3 Proof of Theorem 1.2

In order to prove Theorem 1.2, we introduce another family of functions:

$$\begin{aligned} \bar{\mathcal{C}}_d = \big\{ f \in C_c(\mathbb{R}^d, \mathbb{R}_+) : \forall\, i \in \{1, \cdots, d\} \text{ and } (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_d) \in \mathbb{R}^{d-1} \text{ fixed,} \\ \text{the function } x_i \mapsto f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d) \in \mathcal{C}_1 \big\}, \end{aligned}$$

where $\mathcal{C}_1$ is the class of functions defined in (2.1) for $d = 1$. We have

Theorem 3.1. *Assume that μ is a product probability measure: $\mu = \prod_{i=1}^d \mu_i$, where $\mathrm{d}\mu_i = \rho_i(|x_i|)\, \mathrm{d}x_i$ with $\rho_i \in C(\mathbb{R}_+, (0, \infty))$. Let B be an ellipsoid:*

$$B = \left\{ x \in \mathbb{R}^d : \frac{x_1^2}{a_1^2} + \cdots + \frac{x_d^2}{a_d^2} \le 1 \right\},$$

where $a_1, \cdots, a_d$ are positive constants. Then for any $f \in \bar{\mathcal{C}}_d$, the following inequality holds:

$$\mu(f \mathbf{1}_B) \ge \mu(f)\, \mu(B).$$

Proof. We will prove this result by induction on the dimension d. When $d=1$, this theorem is a special case of Theorem 1.1. Now suppose that the assertion is true in the $d-1$ dimensional case. Denote by $\mu_{(d-1)} = \prod_{i=1}^{d-1} \mu_i$ the product measure on $\mathbb{R}^{d-1}$. By Fubini's theorem,

$$\begin{aligned}\mu(f\mathbf{1}_B) &= \int_B f \, \mathrm{d}(\mu_{(d-1)} \times \mu_d) \\ &= \int_{-a_d}^{a_d} \mathrm{d}\mu_d(x_d) \int_{B_{d-1}(x_d)} f(x_1, \cdots, x_{d-1}, x_d) \, \mathrm{d}\mu_{(d-1)}(x_1, \cdots, x_{d-1}),\end{aligned} \tag{3.1}$$

where $B_{d-1}(x_d)$ is a $d-1$ dimensional ellipsoid:

$$B_{d-1}(x_d) = \bigg\{ (x_1, \cdots, x_{d-1}) \in \mathbb{R}^{d-1} : \frac{x_1^2}{a_1^2} + \cdots + \frac{x_{d-1}^2}{a_{d-1}^2} \leq 1 - \frac{x_d^2}{a_d^2} \bigg\}.$$

Notice that for fixed $x_d \in [-a_d, a_d]$, $f(\cdot, x_d) \in \bar{\mathcal{C}}_{d-1}$. Using the induction hypothesis, we have

$$\int_{B_{d-1}(x_d)} f(x_1, \cdots, x_{d-1}, x_d) \, \mathrm{d}\mu_{(d-1)}(x_1, \cdots, x_{d-1}) \geq \mu_{(d-1)}(f(\cdot, x_d)) \, \mu_{(d-1)}(B_{d-1}(x_d)).$$

Therefore by (3.1),

$$\mu(f\mathbf{1}_B) \geq \int_{-a_d}^{a_d} \mu_{(d-1)}(f(\cdot, x_d)) \, \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d).$$

The function $[-a_d, a_d] \ni x_d \mapsto \mu_{(d-1)}(B_{d-1}(x_d))$ is even and $\mu_{(d-1)}(B_{d-1}(a_d)) = 0$. We extend it to a function on $\mathbb{R}$ by setting $\mu_{(d-1)}(B_{d-1}(x_d)) \equiv 0$ for $|x_d| > a_d$. Then the above inequality becomes

$$\begin{aligned}\mu(f\mathbf{1}_B) &\geq \int_{-\infty}^{\infty} \mu_{(d-1)}(f(\cdot, x_d)) \, \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d) \\ &= \bigg(\int_{-\infty}^{0} + \int_{0}^{\infty} \bigg) \mu_{(d-1)}(f(\cdot, x_d)) \, \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d).\end{aligned} \tag{3.2}$$

We denote by I_1 and I_2 the two integrals on the right hand side of (3.2). Note that the even function $x_d \mapsto \mu_{(d-1)}(B_{d-1}(x_d))$ is decreasing on $\mathbb{R}_+$. On the other hand, for any fixed $(x_1, \cdots, x_{d-1}) \in \mathbb{R}^{d-1}$, by the definition of the class $\bar{\mathcal{C}}_d$ and Lemma 2.1(i), the function $x_d \mapsto f(x_1, \cdots, x_{d-1}, x_d)$ is decreasing (resp. increasing) on $\mathbb{R}_+$ (resp. $\mathbb{R}_- = (-\infty, 0]$). Hence the same is true for $x_d \mapsto \mu_{(d-1)}(f(\cdot, x_d))$. Applying the FKG inequality (see Lemma 3.2 below) to $2\mu_d$ on $(-\infty, 0]$ leads to

$$\begin{aligned}I_1 &\geq \frac{1}{2} \bigg(2 \int_{-\infty}^{0} \mu_{(d-1)}(f(\cdot, x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \bigg(2 \int_{-\infty}^{0} \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \\ &= \bigg(\int_{-\infty}^{0} \mu_{(d-1)}(f(\cdot, x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \bigg(\int_{-\infty}^{\infty} \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d) \bigg),\end{aligned} \tag{3.3}$$

where the equality follows from the symmetry of the integrand the the measure μ_d. Similarly we have

$$I_2 \geq \bigg(\int_{0}^{\infty} \mu_{(d-1)}(f(\cdot, x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \bigg(\int_{-\infty}^{\infty} \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d) \bigg).$$

Combining this with (3.2) and (3.3), we conclude that

$$\begin{aligned}\mu(f\mathbf{1}_B) &\geq \bigg(\int_{-\infty}^{\infty} \mu_{(d-1)}(f(\cdot, x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \bigg(\int_{-\infty}^{\infty} \mu_{(d-1)}(B_{d-1}(x_d)) \, \mathrm{d}\mu_d(x_d) \bigg) \\ &= \mu(f) \, \mu(B).\end{aligned}$$

Therefore the result holds as well in the d dimensional case. The proof is complete. □

Lemma 3.2 (FKG inequality). *Let $-\infty \le a, b \le \infty$ and ν be a probability measure on $[a,b]$. Assume f and g are two bounded increasing (or decreasing) functions on $[a,b]$, then*

$$\int_a^b fg\,\mathrm{d}\nu \ge \Big(\int_a^b f\,\mathrm{d}\nu\Big)\Big(\int_a^b g\,\mathrm{d}\nu\Big).$$

Proof. For any $x, y \in [a,b]$, since both f and g are increasing (or decreasing) functions on $[a,b]$, we have

$$(f(x)-f(y))(g(x)-g(y)) \ge 0.$$

As the two functions are bounded, we can integrate the above inequality on $[a,b]^2$ with respect to $\nu \times \nu$ and obtain

$$\int_{[a,b]^2} (f(x)-f(y))(g(x)-g(y))\,\mathrm{d}(\nu\times\nu)(x,y) \ge 0.$$

Expanding the product gives the desired result. $\square$

Remark 3.3. *The proof of Theorem 3.1 works for more general set B. Indeed, for $i = 1, \cdots, d$, let $f_i \in C(\mathbb{R}_+, \mathbb{R}_+)$ be a strictly increasing function such that $f_i(0) = 0$. Then the result of Theorem 3.1 still holds for the set*

$$B = \{x \in \mathbb{R}^d : f_1(|x_1|) + \cdots + f_d(|x_d|) \le 1\}.$$

Notice that B can even be non-convex. For example, when $d = 2$ and $f_1(t) = f_2(t) = \sqrt{t}$ for $t \ge 0$, then $B = \{x \in \mathbb{R}^2 : \sqrt{|x_1|} + \sqrt{|x_2|} \le 1\}$ is clearly not convex.

Now we are in the position to prove Theorem 1.2. We focus on the case $d \ge 2$ (the case $d = 1$ has been proved in Theorem 1.1).

Proof of Theorem 1.2. Consider the approximations f_n of the indicator function $\mathbf{1}_A$ defined in (2.3). In order to apply Theorem 3.1, we have to show that for every $n \ge 1$, $f_n \in \bar{\mathcal{C}}_d$. For simplicity of notations, we assume $i = 1$, that is, for any $x' = (x_2, \cdots, x_d) \in \mathbb{R}^{d-1}$ fixed, we need to prove that the function $x_1 \mapsto f_n(x_1, x') \in \mathcal{C}_1$, where $\mathcal{C}_1$ is defined in (2.1). For any $c > 0$,

$$I := \{x_1 \in \mathbb{R} : f_n(x_1, x') > c\} = \big\{x_1 \in \mathbb{R} : \mathrm{dist}\big((x_1, x'), A\big) < (1-c)/n\big\}.$$

If $x_1, \bar{x}_1 \in I$, $x_1 < \bar{x}_1$, then $\mathrm{dist}\big((x_1, x'), A\big) \vee \mathrm{dist}\big((\bar{x}_1, x'), A\big) < (1-c)/n$. Since A is convex, as in the proof of Theorem 1.1, we can show that

$$\mathrm{dist}\big(\lambda(x_1, x') + (1-\lambda)(\bar{x}_1, x'), A\big) < (1-c)/n$$

for all $\lambda \in (0,1)$. That is, $\mathrm{dist}\big((\lambda x_1 + (1-\lambda)\bar{x}_1, x'), A\big) < (1-c)/n$. Consequently, for all $\lambda \in (0,1)$, $\lambda x_1 + (1-\lambda)\bar{x}_1 \in I$. This means that I is an interval, hence it is convex.

Next we prove that $0 \in I$ whenever I is nonempty. Indeed, if $x_1 \in I$, then $\mathrm{dist}\big((x_1, x'), A\big) < (1-c)/n$. Hence there is $y = (y_1, y') \in A$ such that $|(x_1, x') - (y_1, y')| < (1-c)/n$. By the property of A, we have $(0, y') \in A$. Moreover

$$|(0, x') - (0, y')| \le |(x_1, x') - (y_1, y')| < (1-c)/n.$$

Therefore $\mathrm{dist}\big((0, x'), A\big) < (1-c)/n$, that is, $0 \in I$. Now if there is $x_1 \in \mathbb{R}$ such that $f_n(x_1, x') > f_n(0, x')$, then consider the interval

$$\tilde{I} = \big\{f_n(\cdot, x') > \big(f_n(x_1, x') + f_n(0, x')\big)/2\big\}.$$

We have $x_1 \in \tilde{I}$ but $0 \in \tilde{I}^c$, which is a contradiction with the result that we have just proved. Hence $f_n(0, x') \ge f_n(x_1, x')$ for all $x_1 \in \mathbb{R}$. Therefore the function $x_1 \mapsto f_n(x_1, x') \in \mathcal{C}_1$. Summing up these arguments, we conclude that $f_n \in \bar{\mathcal{C}}_d$.

Now applying Theorem 3.1 to f_n, we obtain $\mu(f_n \mathbf{1}_B) \ge \mu(f_n)\,\mu(B)$ for all $n \ge 1$. Letting $n \to \infty$ gives the desired result. $\square$

4 A special case of (1.2)

In the present section, we follow the method in [2] (see p.265) and prove that the inequality (1.2) holds if g is the composition of a decreasing function and a positive definite quadratic form. This gives an alternative proof to [3, Theorem 2].

Theorem 4.1. *Assume that $f \in C(\mathbb{R}^d, \mathbb{R}_+)$ is a log-concave and symmetric function. Let Σ be a positive definite matrix and $\varphi \in C(\mathbb{R}_+, \mathbb{R}_+)$ a decreasing function. Then*

$$\int_{\mathbb{R}^d} f(x)\,\varphi(\langle \Sigma x, x\rangle)\,\mathrm{d}\gamma_d(x) \geq \bigg(\int_{\mathbb{R}^d} f(x)\,\mathrm{d}\gamma_d(x)\bigg)\bigg(\int_{\mathbb{R}^d} \varphi(\langle \Sigma x, x\rangle)\,\mathrm{d}\gamma_d(x)\bigg). \tag{4.1}$$

Proof. Consider the Gaussian probability measure

$$\mathrm{d}\mu = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} e^{-\langle \Sigma^{-1}x, x\rangle/2}\mathrm{d}x,$$

where $\det(\Sigma)$ is the determinant of Σ. Then (4.1) is equivalent to

$$\int_{\mathbb{R}^d} f\big(\sqrt{\Sigma^{-1}}\,x\big)\varphi(|x|^2)\,\mathrm{d}\mu(x) \geq \bigg(\int_{\mathbb{R}^d} f\big(\sqrt{\Sigma^{-1}}\,x\big)\mathrm{d}\mu(x)\bigg)\bigg(\int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}\mu(x)\bigg). \tag{4.2}$$

Since f is log-concave and symmetric, it is easy to see that $f(x) \leq f(0)$ for all $x \in \mathbb{R}^d$, hence $C_f := \int_{\mathbb{R}^d} f\big(\sqrt{\Sigma^{-1}}\,x\big)\mathrm{d}\mu(x) < +\infty$. We introduce the probability measure μ_f defined by

$$\mathrm{d}\mu_f = \frac{1}{C_f} f\big(\sqrt{\Sigma^{-1}}\,x\big)\mathrm{d}\mu(x).$$

Hence it is sufficient to prove that

$$\int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}\mu_f(x) \geq \int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}\mu(x). \tag{4.3}$$

Now let T be the optimal transport map pushing forward the Gaussian measure μ to μ_f, i.e. $\mu_f = \mu \circ T^{-1}$. Since the density function $\frac{1}{C_f} f\big(\sqrt{\Sigma^{-1}}\,x\big)$ is also log-concave, we deduce from Caffarelli's result (see [1]) that T is a contraction. Moreover by the symmetry of the density function, we have $T(-x) = -T(x)$; particularly, $T(0) = 0$. Therefore, $|T(x)| \leq |x|$ for all $x \in \mathbb{R}^d$. As a result,

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}\mu_f(x) &= \int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}(\mu \circ T^{-1})(x)\\ &= \int_{\mathbb{R}^d} \varphi(|T(x)|^2)\,\mathrm{d}\mu(x) \geq \int_{\mathbb{R}^d} \varphi(|x|^2)\,\mathrm{d}\mu(x),\end{aligned}$$

where the last inequality follows from the fact that φ is a decreasing function. (4.3) is proved. □

If we take $\varphi(t) = e^{-t/2}$ for $t \geq 0$, then Theorem 4.1 reduces to [8, Proposition 2] (see p.352). Moreover by approximating the indicator function of a symmetric convex set, we can reprove Hargé's result [3].

Corollary 4.2. *Let $A \subset \mathbb{R}^d$ be any symmetric convex set and B be the ellipsoid $\{x \in \mathbb{R}^d : \langle \Sigma x, x\rangle \leq 1\}$. Then*

$$\gamma_d(A \cap B) \geq \gamma_d(A)\gamma_d(B).$$

Proof. We consider again the sequence of approximating functions f_n defined in (2.3). First we show that f_n is log-concave for every $n \geq 1$. Since the set A is convex, it is easy to see that the distance function $\mathrm{dist}(\cdot, A)$ is also convex. Hence for any $x, y \in \mathbb{R}^d$ and $0 < \lambda < 1$,

$$\mathrm{dist}\big(\lambda x + (1-\lambda)y, A\big) \leq \lambda\, \mathrm{dist}(x, A) + (1-\lambda)\, \mathrm{dist}(y, A). \tag{4.4}$$

In order to show that $f_n(\lambda x + (1-\lambda)y) \geq f_n(x)^\lambda f_n(y)^{1-\lambda}$, it is enough to consider the case $f_n(x) \wedge f_n(y) > 0$, that is, $\mathrm{dist}(x, A) \vee \mathrm{dist}(y, A) < 1/n$. Therefore by (4.4),

$$1 - n\, \mathrm{dist}\big(\lambda x + (1-\lambda)y, A\big) \geq \lambda\, [1 - n\, \mathrm{dist}(x, A)] + (1-\lambda)\, [1 - n\, \mathrm{dist}(y, A)].$$

In other words,

$$f_n(\lambda x + (1-\lambda)y) \geq \lambda\, f_n(x) + (1-\lambda)\, f_n(y). \tag{4.5}$$

Now using Young's inequality (for any $a, b \geq 0$ and $p, q > 1$ such that $p^{-1} + q^{-1} = 1$, it holds $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$), we obtain

$$\lambda\, f_n(x) + (1-\lambda)\, f_n(y) \geq f_n(x)^\lambda f_n(y)^{1-\lambda}.$$

Combining this with (4.5), we obtain the log-concavity of f_n.

The symmetry of the set A implies that f_n is also symmetric. Now applying Theorem 4.1 to the functions f_n and letting $n \to \infty$, we arrive at

$$\int_{\mathbb{R}^d} \mathbf{1}_A(x)\, \varphi(\langle \Sigma x, x\rangle)\, \mathrm{d}\gamma_d(x) \geq \gamma_d(A) \bigg(\int_{\mathbb{R}^d} \varphi(\langle \Sigma x, x\rangle)\, \mathrm{d}\gamma_d(x) \bigg). \tag{4.6}$$

Next define

$$\varphi_n(t) = \begin{cases} 1, & t \in [0,1]; \\ 1 - n(t-1), & 1 < t < 1 + n^{-1}; \\ 0, & t \geq 1 + n^{-1}. \end{cases}$$

Then $\varphi_n(t) \downarrow \mathbf{1}_{[0,1]}(t)$ as $n \to \infty$. Replacing φ by φ_n in (4.6) and letting $n \to \infty$, we finally get the desired result. □

References

[1] Caffarelli L. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.* 214 (2000), 547–563.

[2] Cordero-Erausquin Dario. Some applications of mass transport to Gaussian-type inequalities. *Arch. Rational Mech. Anal.* 161 (2002), 257–269.

[3] Hargé Gilles. A particular case of correlation inequality for the Gaussian measure. *Ann. Probab.* 27 (1999), 1939–1951.

[4] Hu Yaozhong. Ito-Wiener chaos expansion with exact residual and correlation, variance inequalities. *J. Theoret. Probab.* 10 (1997), 835-848.

[5] Li Wenbo V. A Gaussian correlation inequality and its applications to small ball probabilities. *Elect. Comm. in Probab.* 4 (1999), 111–118.

[6] Li Wenbo V., Shao Qi-Man. Gaussian Processes: Inequalities, Small Ball Probabilities and Applications. *STOCHASTIC PROCESSES: THEORY AND METHODS.* Handbook of Statistics, Vol. 19, 533-598, Elsevier, New York (2001).

[7] Pitt L.D. A Gaussian correlation inequality for symmetric convex sets. *Ann. Probab.* 5 (1977), 470–474.

[8] Schechtman G., Schlumprecht T. and Zinn J. On the Gaussian measure of intersection of symmetric convex sets. *Ann. Probab.* 26 (1998), 346–357.