ALEA, Lat. Am. J. Probab. Math. Stat. 21, 627–660 (2024)
DOI: 10.30757/ALEA.v21-25

Covariance inequalities for convex and log-concave functions

Michel Bonnefont, Erwan Hillion and Adrien Saumard

(M. Bonnefont) UMR CNRS 5251, Institut de Mathématiques de Bordeaux, Université Bordeaux, France
E-mail address: michel.bonnefont@math.u-bordeaux.fr

(E. Hillion) Aix Marseille Univ, CNRS, I2M, Marseille, France
E-mail address: erwan.hillion@univ-amu.fr

(A. Saumard) Univ. Rennes, ENSAI, CNRS, CREST - UMR 9194, F-35000 Rennes, France
E-mail address: adrien.saumard@ensai.fr

Abstract. Extending results of Hargé and Hu for the Gaussian measure, we prove inequalities for the covariance $\mathrm{Cov}_\mu(f, g)$ where μ is a general product probability measure on $\mathbb{R}^d$ and $f, g : \mathbb{R}^d \to \mathbb{R}$ satisfy some convexity or log-concavity assumptions, with possibly some symmetries.

1. Introduction

If μ is a probability measure on $\mathbb{R}^d$ and if $f, g \in L^2(d\mu)$ are two square integrable functions with respect to μ, their covariance is defined by

$$\mathrm{Cov}_\mu(f, g) = \int \left(f - \int f d\mu\right)\left(g - \int g d\mu\right) d\mu$$

and is a measure of the joint variability of the two functions. Here and in all the sequel, we make the assumptions that f and g have enough integrability and regularity conditions, so that all the written quantities are well defined.

Lying at the intersection of probability, analysis and geometry, covariance identities and inequalities provide a variety of tools. Without trying to be exhaustive, let us cite some of them: FKG inequalities (Fortuin et al. (1971)), (asymmetric) Brascamp-Lieb inequalities (Menz and Otto (2013); Carlen et al. (2013); Arnaudon et al. (2018)), Stein kernels (Chatterjee (2007); Nourdin and Viens (2009); Ledoux et al. (2015); Courtade et al. (2019); Fathi (2019); Saumard (2019)), concentration inequalities (Bobkov et al. (2001); Houdré and Privault (2002), Ledoux (2001a, Section 5.5)).

The proof techniques of these covariance identities and inequalities vary from semi-group techniques, other types of integration by parts, measure transportation or stochastic calculus. Gaussian measures offer a particularly fruitful framework in this perspective and in Theorem 1.1 below, we

Received by the editors July 21st, 2023; accepted January 25th, 2024.

MB is partially supported by the French ANR-18-CE40-0012 RAGE project and the French ANR-19-CE40-0010-01 QuAMProcs project.

recall famous covariance inequalities known for the standard Gaussian measure. The main point of this work is to discuss and extend partially these results beyond the Gaussian assumption, to the case of general product measures.

Theorem 1.1. *Let γ be the standard Gaussian distribution on $\mathbb{R}^d$.*

(1) *Hu (1997); Hargé (2004) Let f and g be two convex functions on $\mathbb{R}^d$, then*

$$\mathrm{Cov}_\gamma(f,g) \geq \mathrm{Cov}_\gamma(f,x) \cdot \mathrm{Cov}_\gamma(g,x) \tag{1.1}$$

where $\cdot$ denotes the standard scalar product on $\mathbb{R}^d$.

(2) *Hargé (2004) Let f be a log-concave function and g be a convex function. Assume moreover that f is orthogonal to the linear functions – that is $\mathrm{Cov}_\gamma(f,x) = 0$ –, then*

$$\mathrm{Cov}_\gamma(f,g) \leq 0. \tag{1.2}$$

(3) *Royen (2014) Let f and g be some quasi-concave functions that are both even, then*

$$\mathrm{Cov}_\gamma(f,g) \geq 0. \tag{1.3}$$

The first point of Theorem 1.1 is due to Hu (1997) and was recovered by Hargé (2004). Hu's proof is based on some Itô-Wiener chaos decomposition. This decomposition is based on the interpolation of the covariance by the standard heat semi-group. Hargé's proof of the second point is based on optimal transport theory and Caffarelli's contraction theorem. Hargé obtained in fact an inequality when f is not necessarily orthogonal to the linear functions, which by a limiting argument recovers (1).

Point (3) was proven by Royen (2014). It is known as the Gaussian correlation inequality and was an open question during decades. We refer to Latała and Matlak (2017) and Barthe (2019) for history of this result. Royen proved his result in its geometric form, for symmetric convex bodies, by approximation with finite intersections of symmetric slabs. The main ingredients are then an interpolation of some dependent and independent Gaussian measures through their covariance matrix and clever computations of the Laplace transform of multivariate Gamma distributions. Royen thus proves its result for some family of multivariate Gamma distributions. Subsequently, Eskenazis et al. (2018) noticed that Theorem 1.1(3) still holds for product measures whose marginals are mixtures of centered Gaussian measures. In the Appendix, we also show that Theorem 1.1(2) is true in the latter situation.

The first main new results of this paper are devoted to dimension one. In dimension one, the covariance inequalities of Theorem 1.1 are not limited to the Gaussian context but actually hold for *any* probability measure on $\mathbb{R}$ having a finite variance.

Theorem 1.2. *Let μ be any probability measure on $\mathbb{R}$ admitting a second moment.*

(1) *For any convex functions f and g, one has*

$$\mathrm{Var}(\mu)\,\mathrm{Cov}_\mu(f,g) \geq \mathrm{Cov}_\mu(f,x)\,\mathrm{Cov}_\mu(g,x). \tag{1.4}$$

(2) *Let f be a log-concave function and g be a convex function. Assume moreover that f is orthogonal to the linear function x, then*

$$\mathrm{Cov}_\mu(f,g) \leq 0. \tag{1.5}$$

(3) *Let f and g be some quasi-concave functions that are both even, then*

$$\mathrm{Cov}_\mu(f,g) \geq 0. \tag{1.6}$$

The fact that Theorem 1.2 holds for any probability measure whereas Theorem 1.1 seems limited to the Gaussian setting is striking and rises the following natural question: what about general product measures? Before trying to answer this question, we shall introduce some notations and the hypotheses.

Notations and hypotheses: In all the sequel of the paper, we consider $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ to be a probability product measure on $\mathbb{R}^d$. Moreover for each $1 \leq k \leq d$, we denote by $a_k(x_k)$ a positive function on $\mathbb{R}$ and by A_k its primitive, centered with respect to μ_k and we assume that $A_k \in L^2(\mu_k)$. When we apply the results with $a_k \equiv 1$, we thus implicitly assume that the probability measure μ_k admits a second moment. We assume that $f, g \in L^2(\mu)$. In order to apply the tensorization method and to exchange derivative and integral, we also assume that all the first and second partial derivatives of f and g are integrable with respect to μ.

Remark 1.3. It is actually possible to weaken the regularity assumptions on the second order partial derivatives. It is indeed sufficient to assume that the first derivatives are monotonic on $\mathbb{R}^d$, at the price of standard approximation arguments. This is particularly transparent with the pure determinantal approach. But for clarity and simplicity, we prove the theorems by using the second order partial derivatives.

Arguably, the first basic idea to investigate general product measures is to use a tensorization argument. This allows us to obtain the following extension of Theorem 1.2(1) to the higher dimensional case.

Theorem 1.4. *Let μ be a product measure on $\mathbb{R}^d$.*
Let f and g be two functions on $\mathbb{R}^d$ such that for each pair $1 \leq i, j \leq d$, the signs of

$$\partial_j \left(\frac{\partial_i f(x)}{a_i(x_i)} \right) \text{ and } \partial_j \left(\frac{\partial_i g(x)}{a_i(x_i)} \right) \tag{1.7}$$

are constant on $\mathbb{R}^d$ and equal. Then

$$\operatorname{Cov}_\mu(f, g) \geq \sum_{i=1}^{d} \frac{1}{\operatorname{Var}_{\mu_i}(A_i)} \operatorname{Cov}_\mu(f(x), A_i(x_i)) \operatorname{Cov}_\mu(g(x), A_i(x_i)).$$

Taking the functions $a_i \equiv 1$ gives the following corollary.

Corollary 1.5. *Let μ be a product measure on $\mathbb{R}^d$. Let f and g be two functions on $\mathbb{R}^d$ such that for each couple $1 \leq i, j \leq d$, the signs of*

$$\partial_{i,j} f(x) \text{ and } \partial_{i,j} g(x) \tag{1.8}$$

are constant and equal. Then

$$\operatorname{Cov}_\mu(f, g) \geq \sum_{i=1}^{d} \frac{1}{\operatorname{Var}(\mu_i)} \operatorname{Cov}_\mu(f(x), x_i) \operatorname{Cov}_\mu(g(x), x_i).$$

A striking point is that Corollary 1.5 is not limited to the Gaussian setting, but holds for any product measure with marginals having a finite second moment. Particularizing to the Gaussian case, where $\mu = \gamma$, the conclusion of Corollary 1.5 is the same as in Theorem 1.1(1), but under different assumptions on the functions f and g. Even if they coincide in dimension one, the two assumptions are different in higher dimensions and are not included one into another. The assumption of Corollary 1.5 seems less classical from a geometric point of view than the convexity assumption. Actually, such assumption on the sign of the second partial derivatives also appears in the context of Gaussian comparison theorems, see for instance Ledoux and Talagrand (2011, Theorem 3.11), that implies Slepian's lemma and Gordon's min-max theorem. Note finally that in the Gaussian setting, even if the statement of Corollary 1.5 seems to be new, its proof could be also deduced from the arguments developed in the proof of Hu (1997).

Remark 1.6. The conditions stated in (1.7) can also be written as the conjunction of Conditions (1.9) and (1.10) below: for each $1 \leq i \leq d$, the signs of

$$\partial_i \left(\frac{\partial_i f(x)}{a_i(x_i)} \right) \text{ and } \partial_i \left(\frac{\partial_i g(x)}{a_i(x_i)} \right) \tag{1.9}$$

are constant and equal and for each couple $1 \leq i \neq j \leq d$, the signs of

$$\partial_{ij} f(x) \text{ and } \partial_{ij} g(x) \tag{1.10}$$

are constant and equal. The condition described by Equation (1.9) can be interpreted as follows: let $B_i : \mathbb{R}^d \to \mathbb{R}^d$ be the inverse bijection of

$$(x_1, \ldots, x_d) \mapsto (x_1, \ldots, A_i(x_i), \ldots, x_d),$$

then Condition (1.9) means that the functions

$$x_i \mapsto (f \circ B_i)(x_1, \ldots, x_d) \ , \ x_i \mapsto (g \circ B_i)(x_1, \ldots, x_d)$$

are both convex or both concave. In the case where $a_i \equiv 1$ for all $i = 1, \ldots, d$, and if moreover, all the signs in (1.9) are the same, then the functions f and g are both coordinatewise convex or both coordinatewise concave.

We now want to investigate what happens when the functions are assumed to satisfy some symmetries. As we shall see, the good notion that fits with the tensorization argument is quite strong and is the unconditionality of (at least) one function.

We recall that a function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *unconditional* if it is symmetric with respect to each hyperplan of coordinates : for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$f(x_1, \ldots, x_d) = f(\varepsilon_1 x_1, \ldots, \varepsilon_d x_d)$$

holds for each choice of signs $(\varepsilon_1, \ldots, \varepsilon_d) \in \{-1, 1\}^d$.

Theorem 1.7 below is a multi-dimensional extension of Theorem 1.2(1) with a symmetry assumption.

Theorem 1.7. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product probability measure on $\mathbb{R}^d$ and assume that for each $1 \leq i \leq d$, the measure μ_i is even. Furthermore, the positive functions a_i are assumed to be even.*

Let f and g be two functions on $\mathbb{R}^d$ such that for each $1 \leq i \leq d$ and all $x \in \mathbb{R}^d$, the signs of

$$\partial_i \left(\frac{\partial_i f(x)}{a_i(x_i)} \right) \text{ and } \partial_i \left(\frac{\partial_i g(x)}{a_i(x_i)} \right) \tag{1.11}$$

are constant and equal. Assume moreover that one of the functions is unconditional. Then

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

The following corollary is directly obtained by setting again the functions a_i to be identically equal to 1.

Corollary 1.8. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $\mathbb{R}^d$ and assume that for $1 \leq i \leq d$, the measures μ_i are even. Let f and g be two functions on $\mathbb{R}^d$ such that for each $1 \leq i \leq d$, the signs of*

$$\partial_{i,i} f(x) \text{ and } \partial_{i,i} g(x) \tag{1.12}$$

are constant and equal. Assume moreover that one of the functions is unconditional. Then

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

With these symmetries, the tensorization method also leads to the following extension of Theorem 1.2(2) and (3).

Theorem 1.9. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $\mathbb{R}^d$.*

(1) *Assume that for $1 \le i \le d$, the marginals μ_i are even and log-concave. Let f be an unconditional positive log-concave function and g be a coordinatewise convex function on $\mathbb{R}^d$, then*

$$\mathrm{Cov}_\mu(f,g) \le 0.$$

(2) *Assume that f and g are both unconditional and coordinatewise quasi-concave. Then*

$$\mathrm{Cov}_\mu(f,g) \ge 0.$$

A drawback of this tensorization approach is arguably that in Theorem 1.7 and Corollary 1.8, we assume a strong symmetry property: the unconditionality of at least one function.

In order to require less symmetry assumptions, it is natural to try to use, instead of the tensorization argument, a more global approach. A first attempt would be to use the interpolation with the associated Markov semi-group and the covariance representation given in (8.4). Actually, we shall provide a slightly different covariance representation, based on an argument of "duplication" of random variables (Lemma 7.1). The main reason for this choice is that the latter approach is much simpler than the semi-group approach and is also more effective. See more comments in Section 8. As expected, this approach allows us to reduce drastically the symmetries required on the functions, but at prize of considering some convexity type assumptions that are less common.

Theorem 1.10. *Let μ be a product measure on $\mathbb{R}^d$.*

(1) *Let $f = e^{-\phi}$ and g be two functions on $\mathbb{R}^d$ such that all $x \in \mathbb{R}^d$,*

$$\partial_i \left(\frac{\partial_i \phi(x)}{a_i(x_i)}\right) \le 0 \text{ and } \partial_i \left(\frac{\partial_i g(x)}{a_i(x_i)}\right) \ge 0 \text{ for all } 1 \le i \le d, \tag{1.13}$$

and

$$\partial_{i,j}\phi(x) \le 0 \text{ and } \partial_{i,j} g(x) \ge 0 \text{ for all } 1 \le i \ne j \le d. \tag{1.14}$$

Assume moreover that f is orthogonal to the functions $A_i(x_i)$ for all $1 \le i \le d$, then

$$\mathrm{Cov}_\mu(f,g) \ge 0.$$

(2) *Let $f = e^{-\phi}$ and $g = e^{-\psi}$ be two functions on $\mathbb{R}^d$ such that for all $x \in \mathbb{R}^d$,*

$$\partial_i \left(\frac{\partial_i \phi(x)}{a_i(x_i)}\right) \le 0 \text{ and } \partial_i \left(\frac{\partial_i \psi(x)}{a_i(x_i)}\right) \le 0 \text{ for all } 1 \le i \le d, \tag{1.15}$$

and

$$\partial_{i,j}\phi(x) \le 0 \text{ and } \partial_{i,j} \psi(x) \le 0 \text{ for all } 1 \le i \ne j \le d. \tag{1.16}$$

Assume moreover that the product measure μ is symmetric and the functions a_i are even for $1 \le i \le d$ and that also both f and g are even, then

$$\mathrm{Cov}_\mu(f,g) \ge 0.$$

As before, the choice $a_k \equiv 1$ is worth looking at and gives the following corollary.

Corollary 1.11. *Let μ be a product measure on $\mathbb{R}^d$.*

(1) *Let $f = e^{-\phi}$ and g be two functions on $\mathbb{R}^d$ such that for all $1 \le i,j \le d$ and all $x \in \mathbb{R}^d$,*

$$\partial_{i,j}\phi(x) \le 0 \text{ and } \partial_{i,j} g(x) \ge 0. \tag{1.17}$$

Assume moreover that f is orthogonal to the coordinate functions x_i for all $1 \le i \le d$, then

$$\mathrm{Cov}_\mu(f,g) \ge 0.$$

(2) *Let $f = e^{-\phi}$ and $g = e^{-\psi}$ be two functions on $\mathbb{R}^d$ such that for all $1 \leq i, j \leq d$ and all $x \in \mathbb{R}^d$,*

$$\partial_{i,j}\phi(x) \leq 0 \text{ and } \partial_{i,j}\psi(x) \leq 0. \tag{1.18}$$

Assume moreover that the product measure μ is symmetric and that both f and g are even, then

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

In the case where $d = 1$, Corollary 1.11 reads: if f is log-convex, g is convex and f is orthogonal to x, then $\mathrm{Cov}_\mu(f, g) \geq 0$. This can be seen as a consequence of Theorem 1.2(1), as log-convexity implies convexity.

Unfortunately, with this global method of Theorem 1.10, we were not able to really extend Theorem 1.2(2) and (3) to some relevant enough "generalized log-concave" situation without any unconditionality assumption. This comes from the fact that the signs of $\partial_{i,j}\phi$ for $i \neq j$ have to be non-negative in order that the corresponding measure satisfies the Holley condition (see Proposition 7.8).

Outline. The paper is organized as follows. The case of the dimension one is investigated in Sections 2, 3 and 4. In Sections 2 and 3, we produce two different proofs of Theorem 1.2. The first one, given in Section 2, is based on the use of determinants and the so-called Andreev's formula. The second one, detailed in Section 3, is based on a covariance identity due to Hoeffding and the use on $\mathbb{R}^2$ of the classical FKG inequality. In Section 4, we notice that more structure is actually present in dimension one: the kernel k in Hoeffding's covariance identity is indeed *totally positive* in the sense of Karlin (1968). Consequently, determinantal covariance inequalities for general Chebyshev systems follow (see Theorem 4.2 for the precise statement). The latter inequalities are also recovered without using Hoeffding's covariance identity, through a direct approach with determinants and Andreev's formula.

The tensorization method and the proofs of Theorems 1.4, 1.7 and 1.9 are given in Section 5, except for the proofs of Theorem 1.2(3) and Theorem 1.9(2), that pertain to the hypothesis of quasi-concavity and are detailed in Section 6. Indeed, the method for proving Theorem 1.2(3) in dimension one is very specific and independent from the rest of the paper. Theorem 1.9(2) is then obtained by tensorization.

In addition, a generalization of Hoeffding's covariance identity for product measures, obtained by a duplication argument, is provided in Section 7. A second proof of Theorem 1.4 and the proof of Theorem 1.10 are then given. As explained above, another natural generalization of Hoeffding's covariance identity would be given through the standard semi-group interpolation. Comments on the difficulty of using this covariance representation are provided in Section 8. Some possible examples of applications are given in Section 9. Finally, in the Appendix, we also prove that Theorem 1.1(2) is true for product measures whose marginals are mixtures of centered Gaussian measures.

2. A determinantal approach in dimension one

This section is devoted to a first proof of Theorem 1.2(1) and (2). The proof is based on properties of determinants, in particular on the intertwining between determinants and the integral operator, a property known as Andreev's formula. Similar arguments will be used in Section 4 in the more general framework of Chebyshev systems.

2.1. *Convexity and determinants.*

Definition 2.1. A pair of real-valued functions (u, U) is said to satisfy Assumption $(\mathcal{C})$ if for any triple $(x_1, x_2, x_3) \in \mathbb{R}^3$ with $x_1 < x_2 < x_3$, one has

$$D(x_1, x_2, x_3) = \begin{vmatrix} 1 & 1 & 1 \\ u(x_1) & u(x_2) & u(x_3) \\ U(x_1) & U(x_2) & U(x_3) \end{vmatrix} \geq 0.$$

In other terms, the couple (u, U) satisfies Assumption $(\mathcal{C})$ if and only if the triple $(1, u, U)$ forms a Chebyshev system (see Definition 4.1). From elementary properties of determinants, it follows that:

Proposition 2.2. *Let (u, U) be satisfying Assumption $(\mathcal{C})$ and $D : \mathbb{R}^3 \to \mathbb{R}$ be as defined above. Let $(x_1, x_2, x_3) \in \mathbb{R}^3$. Let $\sigma \in S_3$ be a permutation of $\{1, 2, 3\}$ such that $x_{\sigma(1)} \leq x_{\sigma(2)} \leq x_{\sigma(3)}$. Then*

$$\varepsilon(\sigma) D(x_1, x_2, x_3) \geq 0$$

Proposition 2.3. *A function $f : \mathbb{R} \to \mathbb{R}$ is convex if and only if, for any $x \in \mathbb{R}$, the pair $(x, f(x))$ satisfies $(\mathcal{C})$.*

Proof: Take $(x_1, x_2, x_3) \in \mathbb{R}^3$ with $x_1 < x_2 < x_3$. Expanding the determinant $D(x_1, x_2, x_3)$ gives:

$$D(x_1, x_2, x_3) = (x_2 - x_1)(f(x_3) - f(x_2)) - (x_3 - x_2)(f(x_2) - f(x_1)).$$

Dividing by the positive quantity $(x_3 - x_2)(x_2 - x_1) > 0$, we obtain that $D(x_1, x_2, x_3) \geq 0$ if and only if

$$\frac{f(x_3) - f(x_2)}{x_3 - x_2} \geq \frac{f(x_2) - f(x_1)}{x_2 - x_1},$$

which is the slope inequality equivalent to convexity of f. □

Corollary 2.4. *If U is an increasing bijection between $\mathbb{R}$ and some interval I, then (u, U) satisfies $(\mathcal{C})$ if and only if $u \circ U^{-1}$ is concave on I.*

Proof: We notice that, for $x_1 < x_2 < x_3$ we have

$$D(x_1, x_2, x_3) = D(U^{-1}(y_1), U^{-1}(y_2), U^{-1}(y_3))$$

for some triple $y_1 < y_2 < y_3 \in I$. Elementary properties of determinants then give:

$$D(x_1, x_2, x_3) = \begin{vmatrix} 1 & 1 & 1 \\ y_1 & y_2 & y_3 \\ (-u \circ U^{-1})(y_1) & (-u \circ U^{-1})(y_2) & (-u \circ U^{-1})(y_3) \end{vmatrix}$$

and we conclude by applying Proposition 2.3. □

Let us now consider some positive function f. Let F be a primitive of f. It is known (see Bobkov (1996)) that f is log-concave if and only if $f \circ F^{-1}$ is concave. From the above, we deduce the following proposition.

Proposition 2.5. *A positive function f is log-concave if and only if the pair (f, F) satisfies $(\mathcal{C})$.*

2.2. *An Andreev-type formula.* A key point in this approach is the following Andreev-type formula which exchanges expectation and determinants.

Proposition 2.6. *Let $(f_i)_{1 \leq i \leq n}$ and $(g_i)_{1 \leq i \leq n}$ be two n-uples of functions in $L^2(\mu)$. We have:*

$$\det\left(\mathbb{E}_\mu\left[f_i(X) g_j(X)\right]\right) = \frac{1}{n!} \mathbb{E}_{\mu \otimes \cdots \otimes \mu}\left[\det\left(f_i(X_j)\right) \det\left(g_i(X_j)\right)\right].$$

Proof: An elementary formula for determinant asserts that:

$$n!\det\left(\mathbb{E}_\mu\left[f_i(X)g_j(X)\right]\right)=\sum_{\sigma,\sigma'\in S_n}\varepsilon(\sigma)\varepsilon(\sigma')\prod_{i=1}^n\mathbb{E}_\mu\left[f_{\sigma(i)}(X)g_{\sigma'(i)}(X)\right].$$

Fubini's theorem allows us to write:

$$n!\det\left(\mathbb{E}_\mu\left[f_i(X)g_j(X)\right]\right)=\sum_{\sigma,\sigma'\in S_n}\varepsilon(\sigma)\varepsilon(\sigma')\mathbb{E}_{\mu\otimes\cdots\otimes\mu}\left[\prod_{i=1}^n f_{\sigma(i)}(X_i)g_{\sigma'(i)}(X_i)\right].$$

We thus have:

$$\begin{aligned}
n!\det\left(\mathbb{E}_\mu\left[f_i(X)g_j(X)\right]\right) &= \mathbb{E}_{\mu\otimes\cdots\otimes\mu}\left[\sum_{\sigma,\sigma'\in S_n}\varepsilon(\sigma)\varepsilon(\sigma')\prod_{i=1}^n f_{\sigma(i)}(X_i)g_{\sigma'(i)}(X_i)\right]\\
&= \mathbb{E}_{\mu\otimes\cdots\otimes\mu}\left[\left(\sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{i=1}^n f_{\sigma(i)}(X_i)\right)\left(\sum_{\sigma'\in S_n}\varepsilon(\sigma')\prod_{i=1}^n g_{\sigma'(i)}(X_i)\right)\right]\\
&= \mathbb{E}_{\mu\otimes\cdots\otimes\mu}\left[\det\left(f_j(X_i)\right)\det\left(g_j(X_i)\right)\right]\\
&= \mathbb{E}_{\mu\otimes\cdots\otimes\mu}\left[\det\left(f_i(X_j)\right)\det\left(g_i(X_j)\right)\right].
\end{aligned}$$

□

Note that with the particular choice $n=2$, $f_1=1$, $f_2=f$, $g_1=1$, $g_2=g$, Proposition 2.6 gives the so-called "Chebyshev's other inequality":

Proposition 2.7 (Chebyshev)**.** *If $f,g\in L^2(\mu)$ are both non-increasing or both non-decreasing, then* $\mathrm{Cov}_\mu(f,g)\geq 0$. *Furthermore, if f is increasing (resp. decreasing) on a set of positive $\mu-$measure and g is increasing (resp. decreasing) on $\mathbb{R}$, then* $\mathrm{Cov}_\mu(f,g)>0$.

2.3. *A first proof of Theorem 1.2.* The first proof of Theorem 1.2 will be deduced from the following more general result.

Theorem 2.8. *Let (u,U) and (v,V) be two pairs of functions satisfying Assumption $(\mathcal{C})$. Then*

$$\mathrm{Cov}_\mu(U,V)\mathrm{Cov}_\mu(u,v)\geq\mathrm{Cov}_\mu(u,V)\mathrm{Cov}_\mu(U,v).$$

Proof: We want to show that $D\geq 0$, where

$$D=\begin{vmatrix}\mathrm{Cov}_\mu(u,v) & \mathrm{Cov}_\mu(u,V)\\ \mathrm{Cov}_\mu(U,v) & \mathrm{Cov}_\mu(U,V)\end{vmatrix}.$$

But we also have:

$$D=\begin{vmatrix}1 & \mathbb{E}_\mu[v] & \mathbb{E}_\mu[V]\\ \mathbb{E}_\mu[u] & \mathbb{E}_\mu[uv] & \mathbb{E}_\mu[uV]\\ \mathbb{E}_\mu[U] & \mathbb{E}_\mu[Uv] & \mathbb{E}_\mu[UV]\end{vmatrix}.$$

The latter equality can be proven by simply expanding the determinant. We now apply Proposition 2.6 with $f_1=1, f_2=u, f_3=U$ and $g_1=1, g_2=v, g_3=V$. This gives

$$D=\int_{\mathbb{R}^3}\begin{vmatrix}1 & 1 & 1\\ u(x_1) & u(x_2) & u(x_3)\\ U(x_1) & U(x_2) & U(x_3)\end{vmatrix}\begin{vmatrix}1 & 1 & 1\\ v(x_1) & v(x_2) & v(x_3)\\ V(x_1) & V(x_2) & V(x_3)\end{vmatrix}d\mu(x_1)d\mu(x_2)d\mu(x_3).$$

Let $(x_1,x_2,x_3)\in\mathbb{R}^3$ and $\sigma\in S_3$ be such that $x_{\sigma(1)}<x_{\sigma(2)}<x_{\sigma(3)}$. As both pairs (u,U) and (v,V) satisfy $(\mathcal{C})$, we apply Proposition 2.2 as follows,

$$\begin{vmatrix}1 & 1 & 1\\ u(x_1) & u(x_2) & u(x_3)\\ U(x_1) & U(x_2) & U(x_3)\end{vmatrix}\begin{vmatrix}1 & 1 & 1\\ v(x_1) & v(x_2) & v(x_3)\\ V(x_1) & V(x_2) & V(x_3)\end{vmatrix}$$

$$= \left(\varepsilon(\sigma)\begin{vmatrix} 1 & 1 & 1 \\ u(x_1) & u(x_2) & u(x_3) \\ U(x_1) & U(x_2) & U(x_3) \end{vmatrix}\right)\left(\varepsilon(\sigma)\begin{vmatrix} 1 & 1 & 1 \\ v(x_1) & v(x_2) & v(x_3) \\ V(x_1) & V(x_2) & V(x_3) \end{vmatrix}\right) \geq 0,$$

from which we deduce $D \geq 0$, as wanted. □

As we prove now, points (1) and (2) of Theorem 1.2 are particular cases of Theorem 2.8, for a suitable choice for the pairs (u, U) and (v, V).

Proof of Theorem 1.2(1) and (2): The first item is a direct consequence of Theorem 2.8 and Proposition 2.3 for the particular choice $u(x) = x$, $U(x) = f(x)$, $v(x) = x$ and $V(x) = g(x)$.

For the second item, we set $u(x) = f(x)$, $U(x) = \int_0^x f(t)dt$, $v(x) = x$, $V(x) = g(x)$. Propositions 2.5 and 2.3 show that the pairs (u, U) and (v, V) both satisfy Assumption $(\mathcal{C})$. By Theorem 2.8, we thus have:

$$\mathrm{Cov}_\mu(f, x)\mathrm{Cov}_\mu(U, g) \geq \mathrm{Cov}_\mu(f, g)\mathrm{Cov}_\mu(U, x).$$

The orthogonality assumption gives $\mathrm{Cov}_\mu(f, x) = 0$. Moreover, as f we can assume without loss of generality that f is positive on a set of positive measure, and as the function U is increasing on $\mathbb{R}$, Proposition 2.7 then gives $\mathrm{Cov}_\mu(U, x) > 0$, so that the inequality $\mathrm{Cov}_\mu(f, g) \leq 0$ holds, as desired. □

3. The Hoeffding covariance identity approach in dimension one

This section is devoted to a second proof of Theorem 1.2(1) and (2). The proof will follow from the Hoeffding covariance identity and the use of the FKG inequality for a new probability measure on $\mathbb{R}^2$. A key point is that the kernel of the Hoeffding representation is *totally positive.*

3.1. *Hoeffding's covariance identity.* We start by recalling the following representation formula for the covariance, which is a consequence of a slightly more general covariance identity due to Hoeffding. See Saumard and Wellner (2018) for more details about Hoeffding's covariance identity.

Theorem 3.1. *Let μ be a probability measure on $\mathbb{R}$ and denote by F_μ its cumulative distribution function, then for all functions f and g in $L^2(\mu)$ and absolutely continuous, one has*

$$\mathrm{Cov}_\mu(f, g) = \iint f'(x)k_\mu(x, y)g'(y)dxdy, \tag{3.1}$$

with

$$k_\mu(x, y) = F_\mu(x \wedge y) - F_\mu(x)F_\mu(y)$$

and $x \wedge y = \min(x, y)$.

For simplicity, when it is clear from context, we shall write $k = k_\mu$ in the sequel.

We now recall some properties of the kernel $k : \mathbb{R}^2 \to [0, +\infty)$. Taking $f(\cdot) = \mathbf{1}_{[x,\infty[}(\cdot)$, $g(\cdot) = \mathbf{1}_{[y,\infty[}(\cdot)$ and X a random variable with distribution μ, one sees that the kernel k is necessarily unique and can also be written

$$k(x, y) = \mathrm{Cov}_\mu\left(1_{\{X \leq x\}}, 1_{\{X \leq y\}}\right).$$

This kernel is non-negative, bounded, continuous if μ is assumed to be a continuous measure, but it is not differentiable on the line $y = x$.

Let us emphasize the fact that the kernel k is *totally positive* in the sense of Karlin (1968). This result should be classical but we could not find a reference of it in the literature.

Theorem 3.2. *For all $n \geq 2$, $s_1 \leq \cdots \leq s_n \in \mathbb{R}$ and $t_1 \leq \cdots \leq t_n \in \mathbb{R}$,*

$$\det\left(k(s_i, t_j)\right)_{1 \leq i,j \leq n} \geq 0.$$

Proof: The proof follows from Theorem 3.1 in Karlin (1968), or Theorem 4.2 in Pinkus (2010), by showing that the matrix $(k(s_i, t_j))_{1 \leq i,j \leq n}$ is a Green matrix. One can also directly use Corollary 3.1 in Karlin (1968) by writing

$$k(x, y) = \begin{cases} \phi(x)\psi(y) \text{ if } x \geq y \\ \psi(x)\phi(y) \text{ if } x \leq y \end{cases}$$

with $\phi(x) = F_\mu(x)$ non-decreasing and $\psi(y) = 1 - F_\mu(y)$ non-increasing. □

In the case $n = 2$, Theorem 3.2 provides the following inequality.

Corollary 3.3. *For all $s_1 \leq s_2$ and $t_1 \leq t_2$,*

$$k(s_1, t_1)k(s_2, t_2) \geq k(s_1, t_2)k(s_2, t_1). \tag{3.2}$$

The conclusion of Corollary 3.3 is well known in the literature under different names. Inequality (3.2), here in the case of $\mathbb{R}^2$, is sometimes referred to as the *Holley condition* or the *strong FKG condition*. The kernel k is also called *log-supermodular* or *multivariate totally positive of order 2*. We shall also need the following extension:

Corollary 3.4. *Let a and b be two positive functions on $\mathbb{R}$ and define the kernel $k_{a,b}$ on $\mathbb{R}^2$ by*

$$k_{a,b}(x, y) = a(x)k(x, y)b(y).$$

Then the kernel is totally positive and thus satisfies inequality (3.2).

We recall now the classical result, due to Fortuin et al. (1971), which asserts that the Holley condition implies the FKG inequality. We first state the definition of the FKG inequality in $\mathbb{R}^d$.

Definition 3.5. Let $d \geq 1$. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *coordinate increasing* if it is non-decreasing along each coordinate, that is if:

$$\text{for all } x, y \in \mathbb{R}^d, \text{ satisfying } x_i \leq y_i, 1 \leq i \leq d \text{ one has } f(x) \leq f(y). \tag{3.3}$$

A probability measure ν on $\mathbb{R}^d$ is said to satisfy the *FKG inequality* if for all functions f and g *coordinate increasing*, one has:

$$\text{Cov}_\nu(f, g) \geq 0. \tag{3.4}$$

Theorem 3.6. *Let ν be a probability measure $\mathbb{R}^d$ with density k with respect to the Lebesgue measure. Assume that for all $x, y \in \mathbb{R}^d$,*

$$k(x \wedge y)k(x \vee y) \geq k(x)k(y), \tag{3.5}$$

where $x \wedge y = (\min(x_1, y_1), \ldots, \min(x_d, y_d))$ and $x \vee y = (\max(x_1, y_1), \ldots, \max(x_d, y_d))$. Then ν satisfies the FKG inequality.

Remark 3.7. Writing $k = e^H$, the condition (3.5) writes:

$$H(x \wedge y) + H(x \vee y) \geq H(x) + H(y). \tag{3.6}$$

In the case where k is smooth - more precisely when H is of class $\mathcal{C}^2$ here -, inequality (3.6) is equivalent to the following condition on the second order cross-derivatives of H:

$$\frac{\partial^2}{\partial x_i \partial x_j} H(x) \geq 0 \text{ for } 1 \leq i \neq j \leq d.$$

In this case, Bakry and Michel (1992) proved the slightly stronger result that the associated semigroup (see Section 8) preserves the class of coordinate increasing functions. Finally note that the kernel k_μ given in Theorem 3.1 above, is not smooth on the diagonal of $\mathbb{R}^2$ and that $\partial^2_{x,y} \ln k_\mu(x, y) = 0$ for all $x \neq y \in \mathbb{R}^2$.

Remark 3.8. Condition (3.5) is not equivalent to the FKG inequality. In the Gaussian setting, for a Gaussian vector with non-singular matrix covariance Γ on $\mathbb{R}^d$, the Holley condition (3.5) is equivalent to $(\Gamma^{-1})_{i,j} \le 0$ for $1 \le i \ne j \le d$. But as proven by Pitt (1982) in the Gaussian setting, the FKG inequality is equivalent to $\Gamma_{i,j} \ge 0$ (see also Tong (1990)). The condition on the coefficient of Γ^{-1} implies the one for Γ. But the converse does not hold. The example 4.3.2 in Tong (1990) provides a covariance matrix for $d \ge 3$ such that $\Gamma_{i,j} \ge 0$ for all $1 \le i, j \le d$ but not $(\Gamma^{-1})_{i,j} \le 0$ for all $1 \le i \ne j \le d$.

3.2. *Hoeffding's formula as a relation between covariances.* The main result of this section is Lemma 3.9 where we express the quantities appearing in Theorem 1.2 as a covariance of the derivatives of the functions with respect to a new probability measure on $\mathbb{R}^2$. Let μ be a probability measure on $\mathbb{R}$ admitting a second moment. We recall that k is the non-negative kernel:

$$k(x,y) = F_\mu(x \wedge y) - F_\mu(x)F_\mu(y).$$

and that from Theorem 3.1, it satisfies

$$\iint k(x,y)dxdy = \mathrm{Var}_\mu(x) = \mathrm{Var}(\mu). \tag{3.7}$$

By assumption, this last quantity is finite and we denote by $\mu^{(1)}$ the following probability measure on $\mathbb{R}^2$:

$$d\mu^{(1)}(x,y) = \frac{k(x,y)}{\iint k(x',y')dx'dy'}dxdy.$$

In the case where f and g are some positive functions, we also denote,

$$d\mu_f^{(1)}(x,y) = \frac{f(x)k(x,y)}{\iint f(x')k(x',y')dx'dy'}dxdy$$

and

$$d\mu_{f,g}^{(1)}(x,y) = \frac{f(x)k(x,y)g(y)}{\iint f(x')k(x',y')g(y')dx'dy'}dxdy.$$

The main result here is the following relation between the covariances of μ and $\mu^{(1)}$. It consists essentially in a rewriting of Hoeffding's covariance identity (3.1) and to highlight the slight difference, we call it "Hoeffding's covariance relation".

Lemma 3.9 (Hoeffding's covariance relation)**.** *Let μ be a probability measure on $\mathbb{R}$ admitting a second moment, with $\mathrm{Var}(\mu) > 0$. Let $f, g : \mathbb{R} \to \mathbb{R}$ be some absolutely continuous functions that belong to $L^2(\mu)$.*

(1) *Then,*

$$\frac{\mathrm{Cov}_\mu(f(x),g(x))}{\mathrm{Var}(\mu)} - \frac{\mathrm{Cov}_\mu(f(x),x)}{\mathrm{Var}(\mu)}\frac{\mathrm{Cov}_\mu(g(x),x)}{\mathrm{Var}(\mu)} = \mathrm{Cov}_{\mu^{(1)}}(f'(x),g'(y)). \tag{3.8}$$

(2) *If moreover, $f = e^{-\phi}$ is positive:*

$$\frac{\mathrm{Cov}_\mu(f(x),g(x))}{Z_f} - \frac{1}{Z_f^2}\mathrm{Cov}_\mu(f(x),x)\,\mathrm{Cov}_\mu(F(x),g(x)) = \mathrm{Cov}_{\mu_f^{(1)}}(-\phi'(x),g'(y)). \tag{3.9}$$

(3) *If moreover $g = e^{-\psi}$ is positive:*

$$\frac{\mathrm{Cov}_\mu(f(x),g(x))}{Z_{f,g}} - \frac{1}{Z_{f,g}^2}\mathrm{Cov}_\mu(f(x),G(x))\,\mathrm{Cov}_\mu(F(x),g(x)) = \mathrm{Cov}_{\mu_{f,g}^{(1)}}(\phi'(x),\psi'(y)). \tag{3.10}$$

Here F and G are primitives of f and g and

$$Z_f = \iint f(x)k(x,y)dxdy = \mathrm{Cov}_\mu(F(x), x) > 0,$$

and

$$Z_{f,g} = \iint f(x)k(x,y)g(y)dxdy = \mathrm{Cov}_\mu(F(x), G(x)) > 0.$$

Note that this approach is also linked to determinants in the sense that the left hand sides of the equalities (3.8), (3.9) (3.10) can be written as determinants. For example, formula (3.8) can be written as

$$\mathrm{Var}(\mu)^2 \mathrm{Cov}_{\mu^{(1)}}(f'(x), g'(y)) = \det \begin{pmatrix} \mathrm{Var}(\mu) & \mathrm{Cov}_\mu(x, f(x)) \\ \mathrm{Cov}_\mu(x, g(x)) & \mathrm{Cov}_\mu(f, g) \end{pmatrix}. \tag{3.11}$$

Proof: Using several times the covariance representation of Theorem 3.1, one has:

$$\begin{aligned} \frac{\mathrm{Cov}_\mu(f(x), g(x))}{\mathrm{Var}(\mu)} &= \iint f'(x) \frac{k(x,y)}{\iint k(x',y')dx'dy'} g'(y)dxdy \\ =& \mathrm{Cov}_{\mu^{(1)}}(f'(x), g'(y)) \\ &+ \left(\iint f'(x)\frac{k(x,y)}{\iint k(x',y')dx'dy'}dxdy\right)\left(\iint g'(y)\frac{k(x,y)}{\iint k(x',y')dx'dy'}dxdy\right) \\ =& \mathrm{Cov}_{\mu^{(1)}}(f'(x), g'(y)) + \frac{\mathrm{Cov}_\mu(f(x), x)}{\mathrm{Var}(\mu)} \frac{\mathrm{Cov}_\mu(g(x), x)}{\mathrm{Var}(\mu)}. \end{aligned}$$

Similarly, if $f = e^{-\phi}$,

$$\begin{aligned} \frac{\mathrm{Cov}_\mu(f(x), g(x))}{Z_f} &= \iint (-\phi'(x)) \frac{f(x)k(x,y)}{Z_f} g'(y)dxdy \\ =& \mathrm{Cov}_{\mu_f^{(1)}}(-\phi'(x), g'(y)) + \frac{1}{Z_f^2}\left(\iint f'(x)k(x,y)dxdy\right)\left(\iint f(x)k(x,y)g'(y)dxdy\right) \\ =& \mathrm{Cov}_{\mu_f^{(1)}}(-\phi'(x), g'(y)) + \frac{1}{Z_f^2}\mathrm{Cov}_\mu(f(x), x)\,\mathrm{Cov}_\mu(F(x), g(x)) \end{aligned}$$

and if moreover $g = e^{-\psi}$,

$$\begin{aligned} \frac{\mathrm{Cov}_\mu(f(x), g(x))}{Z_{f,g}} &= \iint (-\phi'(x)) \frac{f(x)k(x,y)g(y)}{Z_{f,g}} (-\psi'(y))dxdy \\ =& \mathrm{Cov}_{\mu_{f,g}^{(1)}}(\phi'(x), \psi'(y)) + \frac{1}{Z_{f,g}^2}\left(\iint f'(x)k(x,y)g(y)dxdy\right)\left(\iint f(x)k(x,y)g'(y)dxdy\right) \\ =& \mathrm{Cov}_{\mu_{f,g}^{(1)}}(\phi'(x), \psi'(y)) + \frac{1}{Z_{f,g}^2}\mathrm{Cov}_\mu(f(x), G(x))\,\mathrm{Cov}_\mu(F(x), g(x)). \end{aligned}$$

□

3.3. *A second proof of Theorem 1.2.* We are now ready to turn to the second proof of Theorem 1.2(1) and (2) pertaining to dimension one. The last ingredient will be the use of the FKG inequality.

Proofs of Theorem 1.2(1) and (2): Let μ be any probability measure $\mathbb{R}$ admitting a second moment and let f and g be two convex functions on $\mathbb{R}$. Using the first covariance relation of Lemma 3.9, it is equivalent to prove that:

$$\mathrm{Cov}_{\mu^{(1)}}(f'(x), g'(y)) \geq 0.$$

By Corollary 3.3 and Theorem 3.6, the probability measure $\mu^{(1)}$ on $\mathbb{R}^2$ satisfies the FKG inequality. Since f and g are convex, the functions $(x,y) \to f'(x)$ and $(x,y) \to g'(y)$ on $\mathbb{R}^2$ are in particular increasing along coordinates in $\mathbb{R}^2$, which implies the desired inequality.

As for the second item, let $f = e^{-\phi}$ be a log-concave function and g be a convex function on $\mathbb{R}$ and such that f is orthogonal to the linear function x. By the second covariance formula of Lemma 3.9, and since $\mathrm{Cov}_\mu(f(x), x) = 0$,

$$\mathrm{Cov}_\mu(f,g) = -\mathrm{Cov}_{\mu_f^{(1)}}(\phi'(x), g'(y)).$$

The second point follows similarly as above, since by Corollary 3.4 the kernel of the probability measure $\mu_f^{(1)}$ is also totally positive. □

4. More covariance inequalities in dimension one for Chebyshev systems

In this section, we consider some generalizations in dimension one of Theorem 1.2, involving formulations through determinants for Chebyshev systems. As for Theorem 1.2, we provide two proofs, one based only through determinantal identities and the other taking advantage of the strong fact that the kernel k_μ is totally positive.

4.1. *Covariance inequalities for Chebyshev systems.* Let us first define Chebyshev systems.

Definition 4.1. A r-uple of functions $(f_1, \ldots, f_r)$ with $f_i : \mathbb{R} \to \mathbb{R}$ is said to form a *Chebyshev system* (of order r) if for all $t_1 \leq \cdots \leq t_r \in \mathbb{R}$,

$$\det(f_i(t_j))_{1\leq i,j\leq r} \geq 0. \tag{4.1}$$

The main result of this section is the following:

Theorem 4.2. *Let μ be a probability measure on $\mathbb{R}$. Let $n \geq 1$ and $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ and $g_1, \ldots, g_n : \mathbb{R} \to \mathbb{R}$ be some functions in $L^2(\mu)$ such that both the $(n+1)$-uples $(1, f_1, \ldots, f_n)$ and $(1, g_1, \ldots, g_n)$ form two Chebyshev systems. Denote $F(x) = \begin{pmatrix} f_1(x) \\ \ldots \\ f_n(x) \end{pmatrix}$ and $G(x) = \begin{pmatrix} g_1(x) \\ \ldots \\ g_n(x) \end{pmatrix}$, then:*

$$\det(\mathrm{Cov}_\mu(F,G)) \geq 0. \tag{4.2}$$

4.2. *A first proof using the determinantal approach.* We produce here a proof of Theorem 4.2 which is based on the methods introduced in Section 2.

We first claim that

$$\det\left(\mathrm{Cov}_\mu(F,G)\right)_{1\leq i,j\leq n} = \det\left(\mathbb{E}_\mu\left[f_i g_j\right]_{0\leq i,j\leq n}\right), \tag{4.3}$$

where we set $f_0(x) = 1$ and $g_0(x) = 1$. Indeed, let A be the matrix $(\mathbb{E}_\mu[f_i g_j])_{0\leq i,j\leq n}$. Let $C_0, \ldots, C_n$ denote the columns of the matrix A. Let us consider the matrix B obtained by replacing, for every $1 \leq j \leq n$, the column C_j by $C_j - \mathbb{E}[g_j]C_0$. As B is obtained from A only by elementary operations, they have the same determinant. Moreover, B can be written in block form:

$$B = \begin{pmatrix} 1 & 0 \\ \mathbb{E}_\mu(F) & \mathrm{Cov}_\mu(F,G) \end{pmatrix}.$$

This proves that $\det(A) = \det\left(\mathrm{Cov}_\mu(F,G)\right)$ and (4.3) follows.

Proof of Theorem 4.2: By the equality (4.3) and the Andreev-type formula of Proposition 2.6, one has

$$\det\left(\mathrm{Cov}_\mu(F,G)\right) = \frac{1}{n!}\int_{\mathbb{R}^{n+1}} \det\left(f_i(x_j)\right)_{0\leq i,j\leq n} \det\left(g_i(x_j)\right)_{0\leq i,j\leq n} d\mu(x_0)\cdots d\mu(x_n).$$

Let us fix $(x_0, \dots, x_n) \in \mathbb{R}^{n+1}$. There exists a permutation $\sigma \in S_n$ such that $x_{\sigma(0)} \leq \cdots \leq x_{\sigma(n)}$. As $(f_0, \dots, f_n)$ and $(g_0, \dots, g_n)$ are Chebyshev systems, one has

$$\varepsilon(\sigma) \det \left(f_i(x_j)\right)_{0 \leq i,j \leq n} \geq 0 \ , \ \varepsilon(\sigma) \det \left(g_i(x_j)\right)_{0 \leq i,j \leq n} \geq 0.$$

One then has $\det \left(f_i(x_j)\right)_{0 \leq i,j \leq n} \det \left(g_i(x_j)\right)_{0 \leq i,j \leq n} \geq 0$, so

$$\det \left(\mathrm{Cov}_\mu(F, G)\right) \geq 0$$

and the result follows. □

4.3. *A second proof with the Hoeffding covariance identity.* We turn now to the proof of the covariance inequality of Theorem 4.2 using Hoeffding's covariance identity (3.1). The main point will be the use of a bivariate Andreev-type formula for bilinear integral operators. Another key point is to transfer the assumption on the functions to an assumption on their derivatives.

Proposition 4.3. *Let $n \geq 1$ and $f_1, \dots, f_n : \mathbb{R} \to \mathbb{R}$ be some $\mathcal{C}^1$ functions. The following assertions are equivalent:*

(1) *The $(n+1)$-uple $(1, f_1, \dots, f_n)$ forms a Chebyshev system.*
(2) *The n-uple $(f_1', \dots, f_n')$ forms a Chebyshev system.*

Proof: Assume (1), let $x_0 < x_1 < \cdots < x_n$ and set $f_0 = 1$. Then, replacing the i-th column C_i by $C_i - C_{i-1}$ for $1 \leq i \leq n$, yields

$$\begin{aligned} 0 \leq \det(f_i(x_j))_{0 \leq i,j \leq n} &= \det(f_i(x_j) - f_i(x_{j-1}))_{1 \leq i,j \leq n} \\ &= \prod_{j=1}^{n} (x_j - x_{j-1}) \det \left(\frac{f_i(x_j) - f_i(x_{j-1})}{x_j - x_{j-1}} \right)_{1 \leq i,j \leq n}. \end{aligned}$$

Letting successively x_j tend to x_{j-1} for $j = 1, ..., n$ gives

$$\det(f_i'(x_{j-1}))_{1 \leq i,j \leq n} \geq 0$$

and (2) follows.

Now assume (2), for $x_0 < x_1 < \cdots < x_n$. Since one has $f_0 = 1$, replacing the i-th column C_i by $C_i - C_{i-1}$ for $1 \leq i \leq n$, one gets

$$\begin{aligned} \det(f_i(x_j))_{0 \leq i,j \leq n} &= \det(f_i(x_j) - f_i(x_{j-1}))_{1 \leq i,j \leq n} \\ &= \det \left(\int_{x_{j-1}}^{x_j} f_i'(u_j) du_j \right)_{1 \leq i,j \leq n} \\ &= \int_{x_{n-1}}^{x_n} \cdots \int_{x_1}^{x_2} \int_{x_0}^{x_1} \det(f_i'(u_j))_{1 \leq i,j \leq n} \, du_1 du_2 \dots du_n \end{aligned}$$

where the last line follows from an Andreev-type formula. Since $u_1 \leq u_2 \leq \cdots \leq u_n$, we have $\det(f_i'(u_j)) \geq 0$ by (2), and (1) follows. □

The second key point is the following bivariate Andreev-type formula for bilinear kernel integral operators on $\mathbb{R}^{2n}$. It is stated here without a proof.

Proposition 4.4. *With the same notation as in Theorem 4.2,*

$$\begin{aligned} \det(\mathrm{Cov}_\mu(F, G)) &= \det \left(\iint_{x,y \in \mathbb{R}} f_i'(x) k(x, y) g_j'(y) d\mu(x) d\mu(y) \right)_{1 \leq i,j \leq n} \\ &= \iint_{\mathcal{D}} \det \left(f_i'(x_j)\right) \det \left(k(x_i, y_j)\right) \det \left(g_i'(y_j)\right) dx_1 \dots dx_n dy_1 \dots dy_n \end{aligned}$$

where $\mathcal{D}$ is the domain of $\mathbb{R}^{2n}$ defined by

$$\mathcal{D} = \Big\{ (x_1, \dots, x_n, y_1, \dots, y_n) \in \mathbb{R}^{2n} \;:\; x_1 < \dots < x_n \;,\; y_1 < \dots < y_n \Big\} .$$

A second proof of Theorem 4.2 is then immediate.

Second proof of Theorem 4.2: By hypothesis and since the kernel k is totally positive, the three determinants in the integral on the second line in the equality of Proposition 4.4 are non-negative and the result follows. ☐

4.4. *Some applications.* We shall use Theorem 4.2 under the following particular form.

Corollary 4.5. *Let $n \geq 1$ and $\phi_1, \dots, \phi_n : \mathbb{R} \to \mathbb{R}$ and $f, g : \mathbb{R} \to \mathbb{R}$ be some functions and denote $F(x) = \begin{pmatrix} \phi_1(x) \\ \cdots \\ \phi_n(x) \\ f(x) \end{pmatrix}$ and $G(x) = \begin{pmatrix} \phi_1(x) \\ \cdots \\ \phi_n(x) \\ g(x) \end{pmatrix}$. Assume that $(1, \phi_1, \dots, \phi_n, f)$ and $(1, \phi_1, \dots, \phi_n, g)$ form two Chebyshev systems, then:*

$$\det(\mathrm{Cov}_\mu(F, G)) \geq 0. \tag{4.4}$$

It is well known that, if f is smooth and if we choose more precisely $\phi_1(x) = x, \dots, \phi_n(x) = x^n$, the condition that $(1, \phi_1, \dots, \phi_n, f)$ forms a Chebyshev system is equivalent to $f^{(n+1)}(x) \geq 0$, for all $x \in \mathbb{R}$. This is a well known generalization of Proposition 2.3, see Karlin (1968, Chapter 6 Example 4).

In the case where we consider $n = 1$, we recover Theorem 1.2(1). Indeed, in dimension one, by Proposition 2.3, the convexity assumptions on f and g are equivalent to the fact that $(1, x, f)$ and $(1, x, g)$ both form a Chebyshev system.

We now describe the result for $n = 2$ and $\phi_1(x) = x, \phi_2(x) = x^2$.

Corollary 4.6. *Let μ be probability measure on $\mathbb{R}$ admitting a fourth moment. Assume that $f^{(3)}(x) \geq 0$ and $g^{(3)}(x) \geq 0$ for all $x \in \mathbb{R}$, then*

$$\begin{aligned} &\mathrm{Cov}_\mu(f, g) \left(\mathrm{Var}_\mu(x^2) \mathrm{Var}_\mu(x) - \mathrm{Cov}_\mu(x, x^2)^2 \right) \\ &\geq \left(\mathrm{Cov}_\mu(x, f) \quad \mathrm{Cov}_\mu(x^2, f) \right) \begin{pmatrix} \mathrm{Var}_\mu(x^2) & -\mathrm{Cov}_\mu(x, x^2) \\ -\mathrm{Cov}_\mu(x, x^2) & \mathrm{Var}_\mu(x) \end{pmatrix} \begin{pmatrix} \mathrm{Cov}_\mu(x, g) \\ \mathrm{Cov}_\mu(x^2, g) \end{pmatrix} . \end{aligned}$$

If moreover $\int_{\mathbb{R}} x d\mu = \int_{\mathbb{R}} x^3 d\mu = 0$, the latter inequality writes

$$\mathrm{Cov}(f, g) \geq \frac{1}{\mathrm{Var}_\mu(x)} \mathrm{Cov}(x, f) \cdot \mathrm{Cov}(x, g) + \frac{1}{\mathrm{Var}_\mu(x^2)} \mathrm{Cov}(x^2, f) \cdot \mathrm{Cov}(x^2, g).$$

5. The tensorization method for product measures

We investigate here probability product measures on $\mathbb{R}^d$ with $d \geq 2$, through the use of the tensorization argument. This method consists in decomposing the covariance for the product measure by the one-dimensional covariances of the marginals and then applying the covariance inequalities previously obtained in dimension one.

5.1. *The tensorization decomposition of the covariance.*

Lemma 5.1. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a probability product measure. For a function $f : \mathbb{R}^d \to \mathbb{R}$, set*

$$f_k(x_1, \dots x_k) = \iint f(x_1, \dots x_d) d\mu_{k+1}(x_{k+1}) \dots d\mu_d(x_d),$$

for $1 \leq k \leq d$, and set $f_0 = \int f d\mu$. Then it holds,

$$\mathrm{Cov}_\mu(f, g) = \sum_{k=1}^{d} \iint \mathrm{Cov}_{\mu_k}(f_k, g_k) d\mu_1(x_1) \dots d\mu_{k-1}(x_{k-1}). \tag{5.1}$$

In the above lemma, the function f_k is the conditional expectation of f knowing $(x_1, \dots, x_k)$ and $\mathrm{Cov}_{\mu_k}(f_k, g_k)$ is the covariance with respect to the one-dimensional marginal μ_k of f_k and g_k ; that is the function depending on $(x_1, \dots, x_{k-1})$ given by:

$$\mathrm{Cov}_{\mu_k}(x \mapsto f_k(x_1, \dots, x_{k-1}, x), x \mapsto g_k(x_1, \dots, x_{k-1}, x)).$$

This decomposition of the covariance is well known and implies the famous tensorization property of the Poincaré inequality; see e.g. Bakry et al. (2014). To be complete, (5.1) is stated for the variance in Ledoux (2001b), but it also applies to the covariance due to the following polarization identity:

$$4 \,\mathrm{Cov}(f, g) = \mathrm{Var}(f + g) - \mathrm{Var}(f - g).$$

5.2. *A weighted Hoeffding's covariance relation in dimension one.* For the sequel, we shall need a slight generalization of the covariance relation of Lemma 3.9 in dimension one, that we describe now.

Let μ be a probability measure on $\mathbb{R}$. Let a be a positive function on $\mathbb{R}$ and let A be the (centered) primitive of a, $A = \int a dx + c$. We assume that a is such that $\mathrm{Var}_\mu(A) < +\infty$. With these notations and with also the same notations as in Section 3.2, we define

$$k_{a,a}(x, y) := a(x) k(x, y) a(y)$$

and we set $Z_{a,a} = \iint k_{a,a}(x, y) dx dy$. By Theorem 3.1, one gets $Z_{a,a} = \mathrm{Var}_\mu(A)$. We will also consider the measure $\mu^{(1)}_{a,a}$, defined by

$$d\mu^{(1)}_{a,a}(x, y) = \frac{k_{a,a}(x, y)}{\iint k_{a,a}(x', y') dx' dy'} dx dy.$$

Since the kernel k is totally positive, by Corollary 3.3, one gets that the kernels $k_{a,a}$ are also totally positive.

Let μ be a probability measure on $\mathbb{R}$. Let a, b be some positive function on $\mathbb{R}$ and let A and B be the (centered) primitive of a and B: $A = \int a dx + c$, $B = \int b dx + c'$. We assume that a and b are such that $\mathrm{Var}_\mu(A) < +\infty$ and $\mathrm{Var}_\mu(B) < +\infty$. With these notations and with also the same notations as in Section 3.2, we define

$$k_{a,b}(x, y) := a(x) k(x, y) b(y)$$

and we set $Z_{a,b} = \iint k_{a,b}(x, y) dx dy$. By Theorem 3.1, one gets $Z_{a,a} = \mathrm{Var}_\mu(A)$. and $Z_{a,b} = \mathrm{Cov}_\mu(A, B)$. We will also consider the measure $\mu^{(1)}_{a,b}$, defined by

$$d\mu^{(1)}_{a,b}(x, y) = \frac{k_{a,a}(x, y)}{\iint k_{a,a}(x', y') dx' dy'} dx dy.$$

Since the kernel k is totally positive, by Corollary 3.3, one gets that the kernels $k_{a,b}$ are also totally positive. The following lemma provides a generalization of Lemma 3.9, which corresponds to the case $a \equiv 1$.

Lemma 5.2. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be some absolutely continuous functions that belong to $L^2(\mu)$.*

(1) *Then,*

$$\begin{aligned}&\mathrm{Cov}_\mu(f(x), g(x))\\ =&Z_{a,a}\left[\mathrm{Cov}_{\mu^{(1)}_{a,a}}\left(\frac{f'(x)}{a(x)}, \frac{g'(y)}{a(y)}\right) + \frac{\mathrm{Cov}_\mu(f(x), A(x))}{Z_{a,a}}\frac{\mathrm{Cov}_\mu(g(x), A(x))}{Z_{a,a}}\right].\end{aligned} \tag{5.2}$$

(2) *If moreover $f = e^{-\phi}$, then*

$$\begin{aligned}&\mathrm{Cov}_\mu(f(x), g(x))\\ =&Z_{af,a}\left[\mathrm{Cov}_{\mu^{(1)}_{af,a}}\left(\frac{-\phi'(x)}{a(x)}, \frac{g'(y)}{a(y)}\right) + \frac{\mathrm{Cov}_\mu(f(x), A(x))}{Z_{af,a}}\frac{\mathrm{Cov}_\mu(F_a(x), g(x))}{Z_{af,a}}\right].\end{aligned} \tag{5.3}$$

(3) *If moreover $g = e^{-\psi}$, then*

$$\begin{aligned}&\mathrm{Cov}_\mu(f(x), g(x))\\ =&Z_{af,ag}\left[\mathrm{Cov}_{\mu^{(1)}_{af,ag}}\left(\frac{\phi'(x)}{a(x)}, \frac{\psi'(y)}{a(y)}\right) + \frac{\mathrm{Cov}_\mu(f(x), G_a(x))}{Z_{af,ag}}\frac{\mathrm{Cov}_\mu(F_a(x), g(x))}{Z_{af,ag}}\right].\end{aligned} \tag{5.4}$$

Here F_a and G_a denotes respectively the primitives of af and ag and the constants $Z_{af,a}$ and $Z_{af,ag}$ are defined as in Lemma 3.9.

5.3. *Proofs of Theorems 1.4, 1.7 and 1.9.* We first consider the proof of Theorem 1.4. We then state a similar result in Theorem 5.3, but with slightly different assumptions.

Proof of Theorem 1.4: First, we assume that f and g are such that all the quantities in (1.9) and (1.10) are non-negative. By the tensorization of the covariance (5.1) and the first covariance relation of Lemma 5.2, one has

$$\begin{aligned}&\mathrm{Cov}_\mu(f, g)\\ =&\sum_{k=1}^{d}\frac{1}{Z_{k,a_k,a_k}}\iint \mathrm{Cov}_{\mu_k}(f_k, A_k(x_k))\mathrm{Cov}_{\mu_k}(g_k, A_k(x_k))d\mu_1(x_1)\dots d\mu_{k-1}(x_{k-1}) \qquad (5.5)\\ &+\sum_{k=1}^{d} Z_{k,a_k,a_k}\iint \mathrm{Cov}_{\mu^{(1)}_{k,(a_k,a_k)}}\left(\frac{\partial_k f_k(x_k)}{a_k(x_k)}, \frac{\partial_k g_k(y_k)}{a_k(y_k)}\right)d\mu_1(x_1)\dots d\mu_{k-1}(x_{k-1}), \qquad (5.6)\end{aligned}$$

where more precisely

$$\partial_k f_k(x_k) = \partial_k f_k(x_1, \dots, x_{k-1}, x_k) \text{ and } \partial_k g_k(y_k) = \partial_k g_k(x_1, \dots, x_{k-1}, y_k).$$

We first prove that the sum in display (5.6) is non-negative. The assumption (1.9) implies that both $(x_k, y_k) \mapsto \frac{\partial_k f(x_k)}{a_k(x_k)}$ and $(x_k, y_k) \mapsto \frac{\partial_k g(y_k)}{a_k(y_k)}$ are coordinatewise increasing on $\mathbb{R}^2$ for every fixed $x_1, \dots, x_{k-1}$. Since by Corollary 3.4, the measure $\mu^{(1)}_{k,(a_k,a_k)}$ satisfies the FKG criterion on $\mathbb{R}^2$, the term (5.6) is non-negative. We now turn to the sum in display (5.5). With a similar notation for g, we set

$$F_{k,a_k}(x_1, \dots x_{k-1}) = \mathrm{Cov}_{\mu_k}(f_k, A_k(x_k)).$$

Since the hypotheses allow to exchange derivation and integrals, one has for $1 \le i \le (k-1)$,

$$\begin{aligned}&\partial_i \mathrm{Cov}_{\mu_k}(f_k, A_k(x_k))\\ =&\iint_{x_k,y_k}\left(\iint_{x_{k+1},\dots,x_d}\partial_{i,k} f(x)d\mu_{k+1}(x_{k+1})\dots d\mu_d(x_d)\right)k_{\mu_k}(x_k, y_k)a_k(y_k)dx_k dy_k.\end{aligned}$$

In particular, the assumptions (1.10) give that the above integrands are non-negative. This implies that the functions F_{k,a_k} and G_{k,a_k} are both coordinate increasing on $\mathbb{R}^{k-1}$. By the standard FKG inequality for product measures, one gets that

$$\begin{aligned}
&\iint F_{k,a_k}\, G_{k,a_k} d\mu_1(x_1)\ldots d\mu_{k-1}(x_{k-1}) \\
&\geq \iint F_{k,a_k} d\mu_1(x_1)\ldots d\mu_{k-1}(x_{k-1}) \cdot \iint G_{k,a_k} d\mu_1(x_1)\ldots d\mu_{k-1}(x_{k-1}) \\
&= \mathrm{Cov}_\mu(f(x), A_k(x_k)) \cdot \mathrm{Cov}_\mu(g(x), A_k(x_k)).
\end{aligned}$$

Summing over the index k ends the proof in the case of non-negative signs in assumptions (1.9) and (1.10). Finally, analyzing the above proof, one sees that it is still valid for general signs. Indeed, under the general case of assumption (1.9), the functions $(x_k, y_k) \mapsto \frac{\partial_k f(x_k)}{a_k(x_k)}$ and $(x_k, y_k) \mapsto \frac{\partial_k g(y_k)}{a_k(y_k)}$ are either both coordinatewise increasing or both coordinatewise decreasing on $\mathbb{R}^2$ and thus have a non-negative covariance with respect to the measure $\mu^{(1)}_{k,(a_k,a_k)}$. Secondly, the last argument relies on the FKG inequality for product measures. In this case, the FKG inequality is in fact also valid if the functions F_{k,a_k} and G_{k,a_k} are monotone along coordinates, with the same monotonicity along each coordinate. This is the case under the general assumption (1.10) and the result follows.

□

As announced in the beginning of this section, we also obtain with this tensorization approach, a similar result under slightly different conditions.

Theorem 5.3. *Let μ be a product measure on $\mathbb{R}^d$. Assume that the marginals μ_k are absolutely continuous with respect to the Lebesgue measure, with positive densities e^{-V_k}, for smooth potentials V_k. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ and assume that for each $1 \leq k \leq d$, the signs of*

$$\partial_k \left(\frac{\partial_k f(x)}{a_k(x_k)}\right) \text{ and } \partial_k \left(\frac{\partial_k g(x)}{a_k(x_k)}\right) \tag{5.7}$$

are constant and equal and that for each (j,k) with $1 \leq j < k \leq d$, the signs of

$$\partial_{j,k} \left(f(x)\frac{A_k(x_k)}{V_k'(x_k)}\right) \text{ and } \partial_{j,k} \left(g(x)\frac{A_k(x_k)}{V_k'(x_k)}\right) \tag{5.8}$$

are also constant and equal. Assume furthermore the following technical assumption:

$$\lim_{x_k \to \pm\infty} f_k(x_1, \ldots, x_k)\frac{A_k(x_k)}{V_k'(x_k)} e^{-V_k(x_k)} = 0, \tag{5.9}$$

where f_k is defined in Lemma 5.1. Then, it holds

$$\mathrm{Cov}_\mu(f,g) \geq \sum_{k=1}^d \frac{1}{\mathrm{Var}_{\mu_k}(A_k)} \mathrm{Cov}_\mu(f(x), A_k(x_k)) \cdot \mathrm{Cov}_\mu(g(x), A_k(x_k)).$$

Note that since each A_k vanishes exactly in one point, the hypothesis in (5.8) forces each V_k to be unimodal (in the sense that V_k has only one zero). In particular, if each potential V_k is strictly convex, one can specialize the result to the case $A_k = V_k'$ or equivalently $a_k = V_k''$. With this specific choice, Theorems 5.3 and 1.4 coincide.

Proof of Theorem 5.3: The proof is similar to the one of Theorem 1.4. The only difference is that we use integration by parts to get a different representation of F_{k,a_k}. We obtain

$$F_{k,a_k}(x_1, \ldots x_{k-1}) = \mathrm{Cov}_{\mu_k}\left(f_k, V_k'(x_k)\frac{A_k(x_k)}{V_k'(x_k)}\right) = \int \partial_k \left(f_k \frac{A_k(x_k)}{V_k'(x_k)}\right) d\mu_k(x_k).$$

Notice that the bracket terms in the integration by parts are zero due to Assumption (5.9). Furthermore, by Assumption (5.8), the two functions F_{k,a_k} and G_{k,a_k} are coordinatewise monotone with the same kind of monotony along each coordinate. The result follows. □

We now add the symmetry assumptions and turn to the proofs of Theorem 1.7 and 1.9.

Proof of Theorem 1.7: We start by the same covariance formulae, given in (5.5), (5.6), as in the proof of Theorem 1.4. As previously, in view of assumption (1.11), all terms in (5.6) are non-negative. Now, we show that under the symmetry assumptions, all the terms in (5.5) vanish. Indeed, since the product measure μ is symmetric, the function f is unconditional and the function a_k are even, we have that, for any $1 \leq k \leq d$, and any $x_1, \ldots x_{k-1}$, the functions

$$x_k \mapsto f_k(x_1, \ldots x_{k-1}, x_k)$$

are even and that the primitive functions A_k are odd. Hence, it holds that

$$\mathrm{Cov}_{\mu_k}(f_k, A_k(x_k)) = 0 \text{ for each } 1 \leq k \leq d \text{ and all } x \in \mathbb{R}^d$$

and all the terms in (5.5) vanish. The result follows. □

Proof of Theorem 1.9(1): With the same notations as above, writing $\phi_k = -\ln f_k$ and using the second point of Lemma 5.2, one has

$$\begin{aligned}
&\mathrm{Cov}_\mu(f, g) \\
&= \sum_{k=1}^{d} \iint \frac{1}{Z_{k,f_k,1}} \mathrm{Cov}_{\mu_k}(f_k, x_k) \left(\iint f_k(x_k) k_{\mu_k}(x_k, y_k) \partial_k g_k(y_k) dx_k dy_k \right) d\mu_1 \ldots d\mu_{k-1} \\
&+ \sum_{k=1}^{d} \iint Z_{k,f_k,1} \mathrm{Cov}_{\mu^{(1)}_{k,(f_k,1)}} \left(-\partial_k \phi_k(x_k), \partial_k g_k(y_k)\right) d\mu_1 \ldots d\mu_{k-1}.
\end{aligned}$$

Since f is unconditional, for each fixed $(x_1, \ldots, x_{k-1})$, the function f_k is even, and thus

$$\mathrm{Cov}_{\mu_k}(f_k, x_k) = 0$$

and the terms of the sum in the right-hand side of the above equality vanish. Furthermore, as f and the μ_i are log-concave, by stability through marginalization of log-concavity (Prékopa's theorem), the functions f_k are log-concave, meaning that the functions ϕ_k are convex. Since by Corollary 3.4, the measures $\mu^{(1)}_{k,(f_k,1)}$ satisfy the FKG inequality on $\mathbb{R}^2$, one has for each fixed $(x_1, \ldots, x_{k-1}) \in \mathbb{R}^{k-1}$,

$$\mathrm{Cov}_{\mu^{(1)}_{k,(f_k,1)}} \left(-\partial_k \phi_k(x_1, \ldots, x_{k-1}, x_k), \partial_k g_k(x_1, \ldots, x_{k-1}, y_k)\right) \leq 0$$

and the proof is complete. □

The proof of Theorem 1.9(2) is given in the next section. One can note that we do not state a version of Theorem 1.9 with the function a_k or A_k. The reason is that we do not know a natural hypothesis on f that would induce a sign for the quantities $\partial_k \left(\frac{\partial_k \phi_k}{a_k} \right)$.

6. The quasi-concave case

This section is devoted to the proof of Theorem 1.2(3) and Theorem 1.9(2), related to the quasi-concave case. This assumption indeed requires different techniques than in the rest of the paper. The result in dimension one is obtained through the so-called layer-cake representation of the functions (see (6.1)) and the result in dimension $d \geq 2$ is then obtained by tensorisation. A similar result already appears in Schechtman et al. (1998), but as far as we know, the statement of Theorem 1.9(2) is new.

Recall that a quasi-concave function f on $\mathbb{R}^d$ is a real-valued function that satisfies, for any $x, y \in \mathbb{R}^d$ and any $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}.$$

An equivalent formulation of quasi-concavity consists in requiring that the upper level sets of the function are convex. In the following, we make use of a weaker notion than quasi-concavity, that we term "coordinatewise quasi-concavity":

Definition 6.1. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *coordinatewise quasi-concave* if for all $(x_1, ..., x_{i-1}, x_{i+1}, ..., x_d) \in \mathbb{R}^{d-1}$, the functions

$$x_i \in \mathbb{R} \mapsto f(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_d) \in \mathbb{R}$$

are quasi-concave.

Another characterization is thus that for any $\lambda \in [0,1]$, any $(x_1, ..., x_{i-1}, x_{i+1}, ..., x_d) \in \mathbb{R}^{d-1}$ and any $x_i, y_i \in \mathbb{R}$,

$$\begin{aligned}&f(x_1, ..., x_{i-1}, \lambda x_i + (1-\lambda)y_i, x_{i+1}, ..., x_d)\\ &\geq \min\{f(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_d), f(x_1, ..., x_{i-1}, y_i, x_{i+1}, ..., x_d)\}.\end{aligned}$$

The above definition and its characterization directly imply that quasi-concave functions are coordinatewise quasi-concave, but the converse is not true.

Note that the interpretation in terms of convex upper level sets does not hold anymore for the notion of coordinatewise quasi-concavity. But still, the upper level sets of a coordinatewise quasi-concave function are connected sets.

We now turn to the proof of Theorem 1.2 (3) in dimension one.

Proof of Theorem 1.2 (3): Let f and g be two quasi-concave non-negative even functions on $\mathbb{R}$. We write, for $x \in \mathbb{R}$,

$$f(x) = \int_0^\infty \mathbf{1}_{A_t}(x)dt \text{ and } g(x) = \int_0^\infty \mathbf{1}_{B_t}(x)dt \tag{6.1}$$

where A_t and B_t for $t \geq 0$ are the level sets of f and g defined by

$$A_t := \{x \in \mathbb{R}, f(x) \geq t\} \text{ and } B_t := \{x \in \mathbb{R}, g(x) \geq t\}.$$

The key point is here that since f and g are quasi-concave and even, the sets A_s and B_t are symmetric intervals on $\mathbb{R}$ and therefore, for each $s, t \geq 0$, $A_s \subset B_t$ or $B_t \subset A_s$. Therefore by Fubini-Tonelli, one has

$$\begin{aligned}\int f(x)g(x)d\mu(x) &= \int_x \int_{s=0}^\infty \int_{t=0}^\infty \mathbf{1}_{A_s}(x)\mathbf{1}_{B_t}(x)dsdtd\mu(x)\\ &= \int_{s=0}^\infty \int_{t=0}^\infty \mu(A_s \cap B_t)dsdt\\ &= \int_{s=0}^\infty \int_{t=0}^\infty \min(\mu(A_s), \mu(B_t))dsdt\\ &\geq \int_{s=0}^\infty \int_{t=0}^\infty \mu(A_s)\mu(B_t)dsdt\\ &= \int f(x)d\mu(x) \int g(x)d\mu(x);\end{aligned}$$

which is precisely the desired inequality. The general case, where f and g are not supposed non-negative, follows easily from this special case : by truncation and dominated convergence. □

We now prove Theorem 1.9(2) by the tenzorisation method. The main argument is ensured by the following lemma, which states the stability of unconditional coordinatewise quasi-concavity by marginalization. Its proof can be found below.

Lemma 6.2. *Consider an integer $d \geq 2$ and take $k \in \{1, ..., d-1\}$. Assume that a function f on $\mathbb{R}^d$ is unconditional and coordinatewise quasi-concave. Then the function*

$$f_k(x_1, \dots x_k) = \int f(x_1, \dots x_d) d\mu_{k+1}(x_{k+1}) \dots d\mu_d(x_d)$$

is coordinatewise quasi-concave and unconditional.

Proof of Theorem 1.9(2): Let f and g be unconditional and coordinatewise quasi-concave functions. Let us first recall the standard tensorization formula:

$$\mathrm{Cov}_\mu(f, g) = \sum_{k=1}^{d} \iint \mathrm{Cov}_{\mu_k}(f_k, g_k) d\mu_1 \dots d\mu_{k-1}.$$

By Lemma 6.2 above, for any $(x_1, ..., x_{k-1}) \in \mathbb{R}^{k-1}$ the functions $f_k(x_1, ..., x_{k-1}, \cdot)$ and $g_k(x_1, ..., x_{k-1}, \cdot)$ are even and quasi-concave on $\mathbb{R}$. By Theorem 1.2 (3), one has $\mathrm{Cov}_{\mu_k}(f_k, g_k) \geq 0$ and the result follows. □

Proof of Lemma 6.2: Unconditionality of f_k directly follows from unconditionality of f. As for the coordinatewise quasi-concavity, we will make the reasoning for the first coordinate x_1 and the arguments readily extend to the other coordinates. Take a pair (x_1, y_1) such that $|x_1| \leq |y_1|$. Assume without loss of generality that $y_1 \geq 0$ (otherwise replace it by $-y_1$). As f is unconditional and coordinatewise quasi-concave, for any $y \in [-y_1, y_1]$ and for any $(x_2, ..., x_d) \in \mathbb{R}^{d-1}$, we have

$$\begin{aligned} f(y, x_2, \dots, x_d) &\geq \min\{f(-y_1, x_2, ..., x_d), f(y_1, x_2, ..., x_d)\} \\ &= f(y_1, x_2, ..., x_d), \end{aligned}$$

where the latter equality follows from unconditionality of f. In particular, as $x_1 \in [-y_1, y_1]$, we have for any $(x_2, ..., x_d) \in \mathbb{R}^{d-1}$,

$$f(x_1, x_2, ..., x_d) \geq f(y_1, x_2, ..., x_d).$$

This gives, for any $\lambda \in [0, 1]$,

$$\begin{aligned} f_k(\lambda x_1 + (1-\lambda) y_1, x_2, \dots x_k) &= \int f(\lambda x_1 + (1-\lambda) y_1, x_2, \dots x_d) d\mu_{k+1} \dots d\mu_d \\ &\geq \int f(y_1, x_2, \dots x_d) d\mu_{k+1} \dots d\mu_d \\ &= \min\{f_k(x_1, x_2, \dots x_k), f_k(y_1, x_2, \dots x_k)\}. \end{aligned}$$

By symmetry, the case $|y_1| \leq |x_1|$ follows, which finishes the proof. □

7. A global approach for product measures

We provide in this section another proof of Theorem 1.4 and we provide the proof of Theorem 1.10. The first main ingredient that will be used is a generalization of Hoeffding's covariance identity (3.1) for product measures. The two other ingredients are a generalization to product measures of the Hoeffding's covariance relations of Lemmas 3.9 and 5.2 and the use of FKG inequalities.

7.1. *Duplication and a generalization of Hoeffding's covariance identity.* We first present in Lemma 7.1 a duplication argument for the covariance of a product measure. Similar duplication representations are well known, see e.g. Chatterjee (2007). We then deduce in Proposition 7.2 a generalization of Hoeffding's covariance identity for product measures.

Lemma 7.1. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $\mathbb{R}^d$. Under suitable integrable conditions one has*

$$\mathrm{Cov}_\mu(f,g) = \frac{1}{2}\sum_{i=1}^{d} \mathbb{E}[\Delta_i f(X,X')\tilde{\Delta}_i g(X,X')], \tag{7.1}$$

where X and X' are two independent random variables of law μ,

$$\Delta_i f(X,X') = f(X_1,\dots,X_i,\dots,X_d) - f(X_1,\dots,X_i',\dots,X_d)$$

and

$$\tilde{\Delta}_i g(X,X') = g(X_1,\dots,X_i,X_{i+1}'\dots,X_d') - g(X_1,\dots,X_i',X_{i+1}'\dots,X_d').$$

Proof of Lemma 7.1: Let X' be an independent copy of X with law μ. By symmetrization and then the use of a telescopic sum, one has

$$\begin{aligned}\mathrm{Cov}_\mu(f,g) &= \mathbb{E}[f(X)(g(X)-g(X'))] \\ &= \sum_{i=1}^{d}\mathbb{E}[f(X)\tilde{\Delta}_i g(X,X')] \\ &= \sum_{i=1}^{d}\mathbb{E}\left[U_i(X,X')\right],\end{aligned}$$

where we define $U_i(X,X') = f(X)\tilde{\Delta}_i g(X,X')$. Let us denote $(X,X')^{\{j\}}$ to be the random vector given by

$$(X,X')^{\{j\}} = \Big((X_1,\dots,X_{j-1},X_j',X_{j+1},\dots,X_d),(X_1',\dots,X_{j-1}',X_j,X_{j+1}',\dots,X_d')\Big)$$

We also write $(X,X')^{\{j\}} = \left(X^{\{j\}},X'^{\{j\}}\right)$ with the slight abuse of notation that $X^{\{j\}}$ depends on (X,X'). Since μ is a product measure, for each i, $(X,X')^{\{i\}}$ is also of law $\mu\otimes\mu$ and thus

$$\begin{aligned}\mathbb{E}\left[U_i(X,X')\right] &= \mathbb{E}\left[U_i\left((X,X')^{\{i\}}\right)\right] \\ &= -\mathbb{E}[f(X^{\{i\}})\tilde{\Delta}_i g(X,X')]\end{aligned}$$

since $\tilde{\Delta}_i g\left((X,X')^{\{i\}}\right) = -\tilde{\Delta}_i g(X,X')$ and thus

$$\mathbb{E}\left[U_i(X,X')\right] = \frac{1}{2}\mathbb{E}\left[U_i(X,X')\right] + \frac{1}{2}\mathbb{E}\left[U_i\left((X,X')^{\{i\}}\right)\right] = \frac{1}{2}\mathbb{E}[\Delta_i f(X,X')\tilde{\Delta}_i g(X,X')]$$

and the result follows. □

From the duplication argument, one obtains the following generalization to product measures of Hoeffding's covariance identity.

Proposition 7.2. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $\mathbb{R}^d$. Let $f,g:\mathbb{R}^d\to\mathbb{R}$ be some coordinatewise absolutely continuous functions in $L^2(\mu)$, then*

$$\mathrm{Cov}_\mu(f,g) = \sum_{i=1}^{d}\iint_{x,x'\in\mathbb{R}^d} \partial_i f(x) k_{\mu_i}(x_i,x_i')\partial_i g(\underline{x}_{i-1},\overline{x'}_i)dx_i dx_i' d\mu(x_{-i})d\mu(x_{-i}') \tag{7.2}$$

where for $x_i,x_i'\in\mathbb{R}$, k_{μ_i} is the standard Hoeffding kernel for the marginal μ_i:

$$k_{\mu_i}(x_i,x_i') = F_{\mu_i}(x_i\wedge x_i') - F_{\mu_i}(x_i)F_{\mu_i}(x_i')$$

and for $x,x'\in\mathbb{R}^d$, $(\underline{x}_{i-1},\overline{x'}_i) = (x_1,\dots,x_{i-1},x_i',\dots,x_d')$, $x_{-i} = (x_1,\dots,x_{i-1},x_{i+1},\dots x_d)$ and $d\mu(x_{-i}) = d\mu_1(x_1)\dots d\mu_{i-1}(x_{i-1})d\mu_{i+1}(x_{i+1})\dots d\mu_d(x_d)$.

Proof of Proposition 7.2: We consider one term in the sum of the covariance formula of Lemma 7.1. We have

$$\begin{aligned}
&\mathbb{E}[\Delta_i f(X,X')\tilde{\Delta}_i g(X,X')] \\
&= \iint_{x,x'\in\mathbb{R}^d} \begin{pmatrix} f(\underline{x}_{i-1},x_i,\overline{x}_{i+1}) \\ -f(\underline{x}_{i-1},x_i',\overline{x}_{i+1}) \end{pmatrix} \begin{pmatrix} g(\underline{x}_{i-1},x_i,\overline{x'}_{i+1}) \\ -g(\underline{x}_{i-1},x_i',\overline{x'}_{i+1}) \end{pmatrix} d\mu(x)d\mu(x') \\
&= \iint_{x,x'\in\mathbb{R}^d} \iint_{s_i,t_i\in\mathbb{R}} \partial_i f(\underline{x}_{i-1},s_i,\overline{x}_{i+1})\partial_i g(\underline{x}_{i-1},t_i,\overline{x'}_{i+1}) \\
&\qquad \left(\mathbf{1}_{\{s_i\le x_i\}} - \mathbf{1}_{\{s_i\le x_i'\}}\right)\left(\mathbf{1}_{\{t_i\le x_i\}} - \mathbf{1}_{\{t_i\le x_i'\}}\right) ds_i dt_i d\mu(x)d\mu(x').
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\iint_{x_i,x_i'\in\mathbb{R}} \left(\mathbf{1}_{\{s_i\le x_i\}} - \mathbf{1}_{\{s_i\le x_i'\}}\right)\left(\mathbf{1}_{\{t_i\le x_i\}} - \mathbf{1}_{\{t_i\le x_i'\}}\right) d\mu_i(x_i)d\mu_i(x_i') \\
&= 2\left(\mathbb{P}(X_i \ge \max(s_i,t_i)) - \mathbb{P}(X_i\ge s_i)\mathbb{P}(X_i\ge t_i)\right) \\
&= 2\left(F_{\mu_i}(s_i\wedge t_i) - F_{\mu_i}(s_i)F_{\mu_i}(t_i)\right) \\
&= 2k_{\mu_i}(s_i,t_i)
\end{aligned}$$

and the proof follows by Fubini theorem and by a change in the name of the letters in the integral. □

We now study some symmetry properties of this covariance representation.

Lemma 7.3. *Assume μ_i is a symmetric one dimensional measure, then the kernel k_{μ_i} is even, that is*

$$k_{\mu_i}(-s_i,-t_i) = k_{\mu_i}(s_i,t_i).$$

Proof: Without loss of generality assume that $s\le t$, then $-t\le -s$, and

$$\begin{aligned}
k_{\mu_i}(-s,-t) &= F_{\mu_i}(-t) - F_{\mu_i}(-s)F_{\mu_i}(-t) \\
&= (1-F_{\mu_i}(t)) - (1-F_{\mu_i}(s))(1-F_{\mu_i}(t)) \\
&= F_{\mu_i}(s) - F_{\mu_i}(s)F_{\mu_i}(t) \\
&= k_{\mu_i}(s,t).
\end{aligned}$$

□

As a consequence, one obtains the following result.

Lemma 7.4. *Assume that μ is a symmetric product measure on $\mathbb{R}^d$. Let $f,g:\mathbb{R}^n\to\mathbb{R}$ be two even functions. Then, for any $1\le i\le d$,*

$$\iint_{x,x'\in\mathbb{R}^d} \partial_i f(x)k_{\mu_i}(x_i,x_i')g(\underline{x}_{i-1},\overline{x'}_i)dx_i dx_i' d\mu(x_{-i})d\mu(x'_{-i}) = 0.$$

Proof: The result follows from using the change of variables $(a,b)=(-x,-x')$ on $\mathbb{R}^{2d}$ and the fact that $\partial_i f$ is odd, g is even and that the kernel k_{μ_i} is even. □

We also derive the following formulas, that will be instrumental in our proofs.

Lemma 7.5. *Assume μ is a product measure on $\mathbb{R}^d$. For each $1\le k\le d$, let $a_k(x_k)$ be a positive function on $\mathbb{R}$ and let A_k be a primitive, centered with respect to μ_k. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a coordinatewise absolutely continuous function. Then for any $1\le i\le d$, one has*

$$\iint_{x,x'\in\mathbb{R}^d} \partial_i f(x)k_{\mu_i}(x_i,x_i')a_i(x_i')dx_i dx_i' d\mu(x_{-i})d\mu(x'_{-i}) = \mathrm{Cov}_\mu(f, A_i(x_i)).$$

In particular,

$$\iint_{x,x'\in\mathbb{R}^d} \partial_i f(x) k_{\mu_i}(x_i,x_i') dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i}) = \mathrm{Cov}_\mu(f,x_i),$$

where, by a slight abuse of notation, x_i stands for the ith-coordinate function. It also holds

$$\iint_{x,x'\in\mathbb{R}^d} k_{\mu_i}(x_i,x_i') dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i}) = \mathrm{Var}_\mu(x_i) = \mathrm{Var}(\mu_i).$$

Proof: The proof is a direct application of Proposition 7.2 with $g(x) = A_i(x_i)$, noticing that only one term in the sum is different from zero. □

7.2. *Hoeffding's covariance relation for product measures.* The main result here is Lemma 7.6 where a similar relation as in Lemma 3.9 is given for product measures.

Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure and write $\Gamma = \Gamma_\mu$ its covariance matrix. Since μ is a product measure, it is diagonal with $\Gamma_{i,i} = \mathrm{Var}_\mu(x_i) = \mathrm{Cov}_\mu(x_i,x_i)$.

Since the kernels k_{μ_i} are non-negative, one can introduce the probability measures on $\mathbb{R}^{2d}$, defined for $1 \le i \le d$ by

$$d\mu^{(1)}_{(i)}(x,y) = \frac{1}{\Gamma_{i,i}} k_{\mu_i}(x_i,y_i) dx_i dy_i d\mu(x_{-i}) d\mu(y_{-i}).$$

If f and g are positive and integrable, we also introduce the following probability measures,

$$d\mu^{(1)}_{(i),f}(x,x') = \frac{1}{Z_{i,f}} f(x) k_{\mu_i}(x_i,x_i') dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i}),$$

with

$$Z_{i,f} = \iint_{x,x'} f(x) k_{\mu_i}(x_i,x_i') dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i})$$

and

$$d\mu^{(1)}_{(i),f,g}(x,x') = \frac{1}{Z_{i,f,g}} f(x) k_{\mu_i}(x_i,x_i') g(\underline{x}_{i-1},\overline{x'}_i) dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i}),$$

with

$$Z_{i,f,g} = \iint_{x,x'} f(x) k_{\mu_i}(x_i,x_i') g(\underline{x}_{i-1},\overline{x'}_i) dx_i dx_i' d\mu(x_{-i}) d\mu(x'_{-i}).$$

The quantity $Z_{i,f}$ can still be written as a covariance with respect to μ: $Z_{i,f} = \mathrm{Cov}_\mu(F_i(x),x_i)$ where F_i is a function such that $\partial_i F_i(x) = f(x)$. This is not anymore the case for $Z_{i,f,g}$.

In the case of a product measure μ, Lemma 3.9 generalizes as follows.

Lemma 7.6. *Let $f,g:\mathbb{R}^d \to \mathbb{R}$ be in $L^2(\mu)$ and coordinatewise absolutely continuous.*

(1) *Then,*

$$\begin{aligned}\mathrm{Cov}_\mu(f,g) = &\sum_{i=1}^d \Gamma_{i,i} \mathrm{Cov}_{\mu^{(1)}_{(i)}}(\partial_i f(x), \partial_i g(\underline{x}_{i-1},\overline{x'}_i)) \\ &+ \sum_{i=1}^d \frac{1}{\Gamma_{i,i}} \mathrm{Cov}_\mu(f(x),x_i)\mathrm{Cov}_\mu(g(x),x_i).\end{aligned}$$

(2) *If moreover $f = e^{-\phi}$, then*

$$\mathrm{Cov}_\mu(f(x),g(x)) = \sum_{i=1}^d Z_{i,f} \mathrm{Cov}_{\mu^{(1)}_{(i),f}}(-\partial_i\phi(x), \partial_i g(\underline{x}_{i-1},\overline{x'}_i))$$

$$+\sum_{i=1}^{d}\mathrm{Cov}_\mu(f(x),x_i)$$
$$\times\left(\iint f(x)\frac{k^{(1)}_{\mu_i}(x_i,x_i')}{Z_{i,f}}\partial_i g(\underline{x}_{i-1},\overline{x'}_i))dx_idx_i'd\mu(x_{-i})d\mu(x'_{-i})\right).$$

In particular, if moreover f is orthogonal to the linear functions x_i, $1\le i\le d$,

$$\mathrm{Cov}_\mu(f(x),g(x))=\sum_{i=1}^{d}Z_{i,f}\mathrm{Cov}_{\mu^{(1)}_{(i),f}}(-\partial_i\phi(x),\partial_i g(\underline{x}_{i-1},\overline{x'}_i)).$$

(3) *If $f=e^{-\phi}$ and $g=e^{-\psi}$,*

$$\begin{aligned}&\mathrm{Cov}_\mu(f(x),g(x))\\ =&\sum_{i=1}^{d}Z_{i,f,g}\mathrm{Cov}_{\mu^{(1)}_{(i),f,g}}(\partial_i\phi(x),\partial_i\psi(\underline{x}_{i-1},\overline{x'}_i))\\ &+\sum_{i=1}^{d}Z_{i,f,g}\left(\iint\partial_i f(x)\frac{k^{(1)}_{\mu_i}(x_i,x_i')}{Z_{i,f,g}}g(\underline{x}_{i-1},\overline{x'}_i))dx_idx_i'd\mu(x_{-i})d\mu(x'_{-i})\right)\\ &\qquad\times\left(\iint f(x)\frac{k^{(1)}_{\mu_i}(x_i,x_i')}{Z_{i,f,g}}\partial_i g(\underline{x}_{i-1},\overline{x'}_i))dx_idx_i'd\mu(x_{-i})d\mu(x'_{-i})\right).\end{aligned}$$

In particular, if the measure μ is symmetric and if both f and g are even, then

$$\mathrm{Cov}_\mu(f(x),g(x))=\sum_{i=1}^{d}Z_{i,f,g}\mathrm{Cov}_{\mu^{(1)}_{(i),f,g}}(\partial_i\phi(x),\partial_i\psi(\underline{x}_{i-1},\overline{x'}_i)).$$

In fact, we shall use in the sequel the following slight weighted generalization, similar to the one of Lemma 5.2.

Lemma 7.7. *Let $f,g:\mathbb{R}^d\to\mathbb{R}$ be in $L^2(\mu)$ and coordinatewise absolutely continuous.*

(1) *Then,*

$$\begin{aligned}\mathrm{Cov}_\mu(f,g)=&\sum_{i=1}^{d}\mathrm{Var}_{\mu_i}(A_i)\ \mathrm{Cov}_{\mu^{(1)}_{(i),a_i,a_i}}\left(\frac{\partial_i f(x)}{a_i(x_i)},\frac{\partial_i g(\underline{x}_{i-1},\overline{x'}_i)}{a_i(x_i')}\right)\\ &+\sum_{i=1}^{d}\frac{1}{\mathrm{Var}_{\mu_i}(A_i)}\mathrm{Cov}_\mu(f(x),A_i(x_i))\mathrm{Cov}_\mu(g(x),A_i(x_i)).\end{aligned}$$

(2) *If moreover $f=e^{-\phi}$ and if f is orthogonal to the functions $A_i(x_i)$, $1\le i\le d$, then*

$$\mathrm{Cov}_\mu(f(x),g(x))=\sum_{i=1}^{d}Z_{i,a_if,a_i}\ \mathrm{Cov}_{\mu^{(1)}_{(i),a_if,a_i}}\left(-\frac{\partial_i\phi(x)}{a_i(x_i)},\frac{\partial_i g(\underline{x}_{i-1},\overline{x'}_i)}{a_i(x_i')}\right).$$

(3) *If moreover $f=e^{-\phi}$ and $g=e^{-\psi}$ and if the measure μ is symmetric, the function a_k are even and both f and g are even, then*

$$\mathrm{Cov}_\mu(f(x),g(x))=\sum_{i=1}^{d}Z_{i,a_if,a_ig}\ \mathrm{Cov}_{\mu^{(1)}_{(i),a_if,a_ig}}\left(\frac{\partial_i\phi(x)}{a_i(x_i)},\frac{\partial_i\psi(\underline{x}_{i-1},\overline{x'}_i)}{a_i(x_i')}\right).$$

Since the other points are somehow similar, we only do the proof for the first item of Lemma 7.6.

Proof for the first item of Lemma 7.6: From Proposition 7.2 and Lemma 7.5, one has

$$\begin{aligned}
&\mathrm{Cov}_\mu(f(x),g(x))\\
&=\sum_{i=1}^{d}\iint_{x,x'\in\mathbb{R}^d}\partial_i f(x)k_{\mu_i}(x_i,x'_i)\partial_i g(\underline{x}_{i-1},\overline{x'}_i)dx_idx'_id\mu(x_{-i})d\mu(x'_{-i})\\
&=\sum_i\Gamma_{i,i}\mathrm{Cov}_{\mu^{(1)}_{(i)}}(\partial_i f(x),\partial_i g(\underline{x}_{i-1},\overline{x'}_i))\\
&\quad+\sum_i\Gamma_{i,i}\left(\iint\partial_i f(x)\frac{k^{(1)}_{\mu_i}(x_i,x'_i)}{\Gamma_{i,i}}dx_idx'_id\mu(x_{-i})d\mu(x'_{-i})\right)\\
&\qquad\times\left(\iint\partial_i g(\underline{x}_{i-1},\overline{x'}_i)\frac{k^{(1)}_{\mu_i}(x_i,x'_i)}{\Gamma_{i,i}}dx_idx'_id\mu(x_{-i})d\mu(x'_{-i})\right)\\
&=\sum_{i=1}^{d}\Gamma_{i,i}\,\mathrm{Cov}_{\mu^{(1)}_{(i)}}(\partial_i f(x),\partial_i g(y))+\sum_{i=1}^{d}\frac{1}{\Gamma_{i,i}}\,\mathrm{Cov}_\mu(f(x),x_i)\,\mathrm{Cov}_\mu(g(x),x_i).
\end{aligned}$$

□

7.3. *Another proof of Theorem 1.4 and a proof of Theorem 1.10.* Before we turn to the announced proofs, we highlight with the next statement that under our assumptions, the new probability measures on $\mathbb{R}^{2d}$ satisfy the Holley condition and thus the FKG inequality.

Recall that $\mu=\mu_1\otimes\cdots\otimes\mu_d$ is a product measure with marginals μ_k, $k=1,\dots,d$, admitting densities, denoted by $\exp(-V_k)$, with respect to the Lebesgue measure. For some index $i\in\{1,\dots,d\}$ and for f and g some positive functions on $\mathbb{R}^d$, the kernel $k_{(i),f,g}$ is defined on $\mathbb{R}^{2d}$ by

$$k_{(i),f,g}(x,x')=f(x)k_{\mu_i}(x_i,x'_i)g(x)\prod_{j\neq i}e^{-V_j(x_j)}\prod_{j\neq i}e^{-V_j(x'_j)}.$$

The measure $\mu^{(1)}_{(i),f,g}$ has a density on $\mathbb{R}^{2d}$ equal to

$$d\mu^{(1)}_{(i),f,g}(x,x')=\frac{1}{Z_{i,f,g}}k_{(i),f,g}(x,x')dxdx',$$

with

$$Z_{i,f,g}=\iint_{x,x'}f(x)k_{\mu_i}(x_i,x'_i)g(\underline{x}_{i-1},\overline{x'}_i)dx_idx'_id\mu(x_{-i})d\mu(x'_{-i}).$$

Proposition 7.8. *Let $\mu=\mu_1\otimes\cdots\otimes\mu_d$ be a probability product measure on $\mathbb{R}^d$ and grant the above notations. One has*

(1) *For all $1\le i\le d$, the measures $\mu^{(1)}_{(i)}$ and $\mu^{(1)}_{(i),a_i,a_i}$ satisfy the Holley condition (3.5). Moreover, for any choice of signs $(\varepsilon_1,\dots,\varepsilon_d)\in\{+1,-1\}^d$, the kernels $\tilde{k}_{(i)}$ and $\tilde{k}_{(i),a_i,a_i}$ defined by*

$$\tilde{k}_{(i)}(x,x')=k_{(i)}(\varepsilon x,\varepsilon x')\text{ and }\tilde{k}_{(i),a_i,a_i}(x,x')=k_{(i),a_i,a_i}(\varepsilon x,\varepsilon x')$$

with $k_{(i),a_i,a_i}$ the density - up to the constant factor $\mathrm{Var}_{\mu_i}(A_i)$ - of the measure $\mu^{(1)}_{(i),a_i,a_i}$ with respect to the Lebesgue measure on $\mathbb{R}^{2d}$ and

$$(\varepsilon x,\varepsilon x')=(\varepsilon_1x_1,\dots,\varepsilon_dx_d,\varepsilon_1x'_1,\dots,\varepsilon_dx'_d),$$

satisfy the Holley condition (3.5).

(2) *Assume that* $f = e^{-\phi}$ *and that for all* $1 \le i, j \le d$ *with* $i \ne j$,

$$\partial_{i,j}\phi(x) \le 0$$

then for all $1 \le i \le d$, *the measures* $\mu^{(1)}_{(i),f}$ *and* $\mu^{(1)}_{(i),a_i f,a_i}$ *satisfy the Holley condition* (3.5).

(3) *Assume that* $f = e^{-\phi}$ *and* $g = e^{-\psi}$ *and that for all* $1 \le i, j \le d$ *with* $i \ne j$,

$$\partial_{i,j}\phi(x) \le 0 \text{ and } \partial_{i,j}\psi(x) \le 0$$

then for all $1 \le i \le d$, *the measures* $\mu^{(1)}_{(i),f,g}$ *and* $\mu^{(1)}_{(i),a_i f,a_i g}$ *satisfiy the Holley condition* (3.5).

Note that in the latter proposition, the signs of the second-order cross derivatives for ϕ and ψ should be both non-positive.

Proof: The logarithm $H^{(1)}_{(i),a_i,a_i}$ of the density of $\mu^{(1)}_{(i),a_i,a_i}$ with respect to the Lebesgue measure on $\mathbb{R}^{2d}$ is given by

$$H^{(1)}_{(i),a_i,a_i}(x,x') = \ln k_{\mu_i}(x_i,x'_i) + \ln a_i(x_i) + \ln a_i(x'_i) - \sum_{j\ne i} V_j(x_j) - \sum_{j\ne i} V_j(x'_j).$$

Since k_{μ_i} is a totally positive kernel on $\mathbb{R}^2$, it follows easily that $H^{(1)}_{(i),a_i,a_i}$ satisfies (3.6). Now for $(\varepsilon_1,\ldots,\varepsilon_d) \in \{+1,-1\}^d$ fixed, the logarithm $\tilde{H}^{(1)}_{(i),a_i,a_i}$ of the kernel $\tilde{k}^{(1)}_{(i),a_i,a_i}$ is given by:

$$\tilde{H}^{(1)}_{(i),a_i,a_i}(x,x') = \ln k_{\mu_i}(\varepsilon_i x_i,\varepsilon_i x'_i) + \ln a_i(\varepsilon_i x_i) + \ln a_i(\varepsilon_i x'_i) - \sum_{j\ne i} V_j(\varepsilon_j x_j) - \sum_{j\ne i} V_j(\varepsilon_j x'_j).$$

Since the kernel $k_{\mu_i}(\varepsilon_i x_i,\varepsilon_i x'_i)$ is still totally positive on $\mathbb{R}^2$, the proof of the first point follows. We turn to the proof of the second point. The logarithm $H^{(1)}_{(i),a_i f,a_i}$ of the density of $\mu^{(1)}_{(i),a_i f,a_i}$ with respect to the Lebesgue measure on $\mathbb{R}^{2d}$ satisfies

$$H^{(1)}_{(i),a_i f,a_i}(x,x') = -\phi(x) + H^{(1)}_{(i),a_i,a_i}.$$

From assumption (1.17) and Remark 3.7, the function $x \to -\phi(x)$ satisfies (3.6) on $\mathbb{R}^d$ and thus clearly the function $(x,x') \to -\phi(x)$ also satisfies (3.6) on $\mathbb{R}^{2d}$. Finally, by summation, Inequality (3.6) is also valid on $\mathbb{R}^{2d}$ for $H^{(1)}_{(i),a_i f,a_i}$. The proof for the third point is similar and we omit the details. □

We now provide another proof of Theorem 1.4.

Another proof of Theorem 1.4: Let f and g be two functions on $\mathbb{R}^d$ satisfying (1.7). We first assume that all the signs of the second derivatives in Assumption 1.7 are non-negative. By Lemma 7.7, one has

$$\begin{aligned}&\mathrm{Cov}_\mu(f,g) - \sum_{i=1}^d \frac{1}{\mathrm{Var}_{\mu_i}(A_i)} \mathrm{Cov}_\mu(f(x),A_i(x_i))\mathrm{Cov}_\mu(g(x),A_i(x_i))\\ &= \sum_{i=1}^d \mathrm{Var}_{\mu_i}(A_i)\mathrm{Cov}_{\mu^{(1)}_{(i),a_i,a_i}}\left(\frac{\partial_i f(x)}{a_i(x_i)}, \frac{\partial_i g(\underline{x}_{i-1},\overline{x'}_i)}{a_i(x'_i)}\right).\end{aligned}$$

Furthermore, by Proposition 7.8(1), the measure $\mu^{(1)}_{(i),a_i,a_i}$, for $i \in \{1,\ldots,d\}$, satisfies the Holley condition (3.5). By condition (1.7) both functions $(x,x') \to \partial_i f(x)/a_i(x_i)$ and $(x,x') \to$

$\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)/a_i(x'_i)$ are coordinate increasing on $\mathbb{R}^{2d}$ and thus, for each $1 \le i \le d$,

$$\mathrm{Cov}_{\mu^{(1)}_{(i),a_i,a_i}} \left(\frac{\partial_i f(x)}{a_i(x_i)}, \frac{\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)} \right) \ge 0.$$

Summing these inequalities ends the proof in this specific case. In the general case, for any $\varepsilon = (\varepsilon_1, \dots, \varepsilon_d) \in \{+1,-1\}^d$, by the change of variable $(\tilde{x}, \tilde{x}') = (\varepsilon x, \varepsilon x')$, one has

$$\mathrm{Cov}_{\mu^{(1)}_{(i),a_i,a_i}} \left(\frac{\partial_i f(x)}{a_i(x_i)}, \frac{\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)} \right) = \mathrm{Cov}_{\tilde{\mu}^{(1)}_{(i),a_i,a_i}} \left(\frac{\partial_i f(\varepsilon x)}{a_i(\varepsilon_i x_i)}, \frac{\partial_i g(\underline{\varepsilon x}_{i-1}, \overline{\varepsilon x'}_i)}{a_i(\varepsilon_i x'_i)} \right)$$

and for each $1 \le i \le d$, it is possible to find some vector $\varepsilon = (\varepsilon_1, \dots, \varepsilon_d) \in \{+1,-1\}^d$ such that $\frac{\partial_i f(\varepsilon x)}{a_i(\varepsilon_i x_i)}$ and $\frac{\partial_i g(\underline{\varepsilon x}_{i-1}, \overline{\varepsilon x'}_i)}{a_i(\varepsilon_i x'_i)}$ are both coordinate increasing. More precisely, it suffices to take $\varepsilon_j = \mathrm{sign}\ \partial_j \left(\frac{\partial_i f}{a_i} \right)$. By Proposition 7.8(1), the measures $\tilde{\mu}^{(1)}_{(i),a_i,a_i}$ also satisfy the Holley condition and the result follows from the FKG inequality. □

We turn now to the proof of Theorem 1.10, where we add some symmetries.

Proof of Theorem 1.10: Let $f = e^{-\phi}$ and g be two functions on $\mathbb{R}^d$ satisfying (1.13) and (1.14) and assume that f is orthogonal to the functions A_i, $1 \le i \le d$. By Lemma 7.7(2), one has

$$\mathrm{Cov}_\mu(f(x), g(x)) = \sum_{i=1}^d Z_{i,a_i f,a_i} \mathrm{Cov}_{\mu^{(1)}_{(i),a_i f,a_i}} \left(-\frac{\partial_i \phi(x)}{a_i(x_i)}, \frac{\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)} \right).$$

Now for each i, since ϕ satisfies (1.14), by Proposition 7.8(2) the measure $\mu^{(1)}_{(i),a_i f,a_i}$ satisfies the Holley condition. Moreover adding condition (1.13) both functions $(x,x') \to \frac{-\partial_i \phi(x)}{a_i(x_i)}$ and $(x,x') \to \frac{\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)}$ are both coordinate increasing on $\mathbb{R}^{2d}$, and thus by the FKG inequality, for each $1 \le i \le d$, one has:

$$\mathrm{Cov}_{\mu^{(1)}_{(i),a_i f,a_i}} \left(-\frac{\partial_i \phi(x)}{a_i(x_i)}, \frac{\partial_i g(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)} \right) \ge 0.$$

Theorem 1.10(1) thus follows. The proof of Theorem 1.10(2) is similar, since by Lemma 7.7(3), the symmetry assumptions made on $f = e^{-\phi}$ and $g = e^{-\psi}$, give that

$$\mathrm{Cov}_\mu(f(x), g(x)) = \sum_{i=1}^d Z_{i,a_i f,a_i g} \mathrm{Cov}_{\mu^{(1)}_{(i),a_i f,a_i}} \left(\frac{\partial_i \phi(x)}{a_i(x_i)}, \frac{\partial_i \psi(\underline{x}_{i-1}, \overline{x'}_i)}{a_i(x'_i)} \right).$$

Finally, the assumptions (1.15) and (1.16) ensure that the measure $\mu^{(1)}_{i,f,g}$ satisfies the Holley condition and that the functions in the covariance are coordinate increasing, which gives the result. □

8. Comments on the standard semi-group interpolation

In this section, we explain what can be done using a standard covariance representation obtained by interpolation with the associated diffusion semi-group (see (8.4) below) and why we did not follow this natural approach, but rather used instead the covariance representation of Proposition 7.2.

We consider here a probability measure $\mu = e^{-V} dx$ with a smooth potential V. One can associate to it a diffusion semi-group with generator L defined for f smooth with compact support by

$$Lf = \Delta f - \nabla V \cdot \nabla f.$$

This diffusion operator is symmetric with respect to μ: for $f, g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$,

$$\int f L g d\mu = \int L f g d\mu = - \int \nabla f \cdot \nabla g d\mu.$$

Under mild conditions on V, one we can define alternatively the semi-group associated to L by the spectral theorem and functional calculus, or by a stochastic representation (see Bakry et al. (2014) for further details),

$$P_t f(x) = e^{tL}(f)(x) = \mathbb{E}[f(X_t^x)]$$

for some Markov diffusion process $(X_t^x)_{t\geq 0}$. We assume moreover that the operator $-\mathcal{L} + \mathrm{Hess}V$, with $\mathcal{L} = \mathrm{diag}(L, \dots, L)$ acting on gradients, is invertible. Note that this holds under some strong convexity of the potential V. In this situation, for $f, g : \mathbb{R}^d \to \mathbb{R}$ satisfying some integrability conditions on f and g, one has

$$\mathrm{Cov}_\mu(f,g) = \int_{\mathbb{R}^d} \nabla f(x) \cdot (-\mathcal{L} + \mathrm{Hess}V)^{-1} \nabla g(x) d\mu(x) \tag{8.1}$$

and thus

$$\mathrm{Cov}_\mu(f,g) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla f(x) K(x,y) \nabla g(y) dx dy$$

where K is the matricial kernel (with respect to the Lebesgue measure) of the operator $(-\mathcal{L} + \mathrm{Hess}V)^{-1}$. Moreover, the matricial kernel $K(x,y)$ admits the following stochastic Feynman-Kac representation,

$$K(x,y) = e^{-V(x)} \int_0^{+\infty} \mathbb{E}[Y_{t,x} | X_t = y] p_t(x,y) e^{-V(y)} dy, \tag{8.2}$$

where p_t stands for the heat kernel associated to P_t with respect to the measure μ and $Y_{t,x}$ is the matrix satisfying the following ordinary (random) differential equation,

$$\frac{d}{dt} Y_{t,x} = -Y_{t,x} \mathrm{Hess}V(X_t^x) \text{ for } t \geq 0;\ Y_{0,x} = Id. \tag{8.3}$$

In the case of a product measure, we can write $V(x) = V_1(x_1) + \cdots + V_d(x_d)$, for some real functions V_k. This gives the following generalization of Hoeffding's covariance identity,

$$\mathrm{Cov}_\mu(f,g) = \sum_{i=1}^d \iint_{x,y \in \mathbb{R}^d} \partial_i f(x)\, \kappa_i(x,y)\, \partial_i g(y) dx dy, \tag{8.4}$$

where for each $1 \leq i \leq d$, $k_i : \mathbb{R}^{2d} \to \mathbb{R}_+$ is the kernel defined by

$$\kappa_i(x,y) = \int_{t=0}^{\infty} \mathbb{E}\left[\exp\left(-\int_0^t V_i''(X_s^{x_i,i}) ds \right) | (X_t^{x_i,i} = y_i) \right] p_t(x,y)\, dt\, e^{-V(x)} e^{-V(y)}. \tag{8.5}$$

This kernel also writes as

$$\kappa_i(x,y) = \int_{t=0}^{\infty} \kappa_{i,t}\, dt$$

with

$$\kappa_{i,t}(x,y) := p_{t,i}^{V_i''}(x_i, y_i) \prod_{j=1, j\neq i}^d p_{t,j}(x_j, y_j) e^{-V(x)} e^{-V(y)},$$

where $p_{t,j}$ is the kernel of the one dimensional diffusion semi-group with generator given by $L_j f(x_j) := f''(x_j) - V_j'(x_j) f'(x_j)$ and where $p_{t,i}^{V_i''}$ is the kernel of the one dimensional Schrödinger semi-group, with generator given by $L_i^{V_i''} f(x_i) := L_i f(x_i) + V_i''(x_i) f(x_i)$. In the case of the standard Gaussian measure, one has $p_{t,i}^{V_i''} = e^{-t} p_{t,i}$.

We highlight that we do not know whether, in dimension $d \geq 2$, the probability measure with density proportional to $\kappa_i(x,y)$ satisfies the full FKG inequality on $\mathbb{R}^{2d}$. But one can also notice, that due to the *coincidence formula*, diffusion kernels and Schrödinger kernels in dimension one are totally positive (see Karlin (1968)). As a consequence, the kernels $\kappa_{i,t}$ satisfy the Holley condition. And if slightly differently, one has

$$\kappa_i(x,y) = \int_{t=0}^{\infty} \kappa_{i,t}\, d\nu(t)$$

for some probability measure ν on $\mathbb{R}_+$, one can use the following decomposition of the covariance,

$$\mathrm{Cov}_{\mu_{\kappa_i}}(u,v)=\int_0^\infty \mathrm{Cov}_{\mu_{\kappa_{i,t}}}(u,v)d\nu(t)+\mathrm{Cov}_\nu\left(t\to\int u\,d\mu_{\kappa_{i,t}},t\to\int v\,d\mu_{\kappa_{i,t}}\right) \quad (8.6)$$

and apply it with $u(x,y):=\partial_i f(x)$ and $v(x,y):=\partial_i g(y)$.

In view of proving Theorem 1.4, we were only able to pursue this approach when the marginals of $\mu_{\kappa_{i,t}}$ on $\mathbb{R}^d\times\mathbb{R}^d$ are constant for all $t>0$. In this situation, if moreover, $(x,y)\to u(x)$ and $(x,y)\to v(y)$, the term related to Cov_ν appearing in the right-hand side of 8.6 indeed vanishes and one obtains some *partial FKG inequalities* for the measure μ_{κ_i} on $\mathbb{R}^{2d}$. Here, the terms "partial" means that it is applied only to coordinate increasing functions of the form $(x,y)\to u(x)$ and $(x,y)\to v(y)$.

This property that the marginals $\mu_{\kappa_{i,t}}$ are constant, holds for the standard Gaussian measure and this approach may be pursued, with a second order covariance representation, to recover Theorem 1.1(1) for the standard Gaussian measure. Let us give some details.

First, the first order representation (8.4) is well known for the standard Gaussian measure (see Bobkov et al. (2001)). The measures $\gamma_{\kappa_{i,t}}$ are in fact independent of i, they are also Gaussian measures on $\mathbb{R}^{2d}$ and they have fixed marginals on $\mathbb{R}^d\times\mathbb{R}^d$. A second order covariance representation for γ thus means a first order covariance of the new measure(s) $\gamma_{\kappa_{i,t}}$ similar to (8.4). It can be obtained either by a change of variable since $\gamma_{\kappa_{i,t}}$ is still a Gaussian measure or by solving explicitly the stochastic Feynman-Kac representation. This method is similar to the one of Hu (1997), except that the latter approach specifically uses the fact that the Gaussian measure is the density at time 1 of the classical heat semi-group, whereas instead we use here the Orstein-Uhlenbeck operator.

Finally, for general product measures, the "constant marginal property" also holds for the modified kernels $k_{\mu_i,(a_i,a_i)}$, with the choice $a_i(x_i)=\frac{1}{g_i'(x_i)}$ where g_i is (if it exists) the first non-trivial eigenfunction associated to L. This leads to Theorem 1.4, but only for this specific choice. More importantly, this constant marginal property is valid for product measures under the symmetry assumptions of Theorem 1.10(2) and this route may also be taken to provide another proof Theorem 1.10(2).

9. Examples

In this final section, we provide a couple of examples of possible applications of our results.

First let μ be a product measure on $\mathbb{R}^d$ and for $\beta>0$ and consider the free energy, also known in the optimization community as the "soft max" function,

$$F_\beta(x):=\frac{1}{\beta}\ln\left(\sum_{i=1}^d e^{\beta x_i}\right).$$

By setting $p_i:=\frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}}$, it satisfies

$$\partial_i F_\beta=p_i\geq 0,$$
$$\partial_{ii}F_\beta(x)=\beta p_i(1-p_i)\geq 0,\quad \partial_{ij}F_\beta(x)=-\beta p_i p_j\leq 0,\quad i\neq j.$$

Thus, for any $\alpha,\beta>0$, Corollary 1.5 gives

$$\mathrm{Cov}_\mu(F_\alpha,F_\beta)\geq\sum_{i=1}^d\frac{1}{\mathrm{Var}(\mu_i)}\mathrm{Cov}(F_\alpha(x),x_i)\mathrm{Cov}(F_\beta(x),x_i). \quad (9.1)$$

Note that when $\alpha=\beta$, inequality (9.1) turns to the following Bessel inequality,

$$\mathrm{Var}_\mu(F_\beta)\geq\sum_{i=1}^d\frac{1}{\mathrm{Var}(\mu_i)}\mathrm{Cov}(F_\beta(x),x_i)^2.$$

We now turn to a second example. Let μ be a symmetric product measure on $\mathbb{R}^d$. Under some integrability condition, for $J \geq 0$, we consider the probability measure:

$$d\mu_J(x) = \frac{1}{Z_J} e^{J \sum_{i=1}^{d-1} x_i x_{i+1}} d\mu(x), \quad Z_J = \int_{\mathbb{R}^d} e^{J \sum_{i=1}^{d-1} x_i x_{i+1}} d\mu(x).$$

Let $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$ with $\theta_i \geq 0$, $1 \leq i \leq d$, then by Corollary 1.11, one has

$$\int_{\mathbb{R}^d} \langle x, \theta \rangle^2 d\mu_J(x) \geq \int_{\mathbb{R}^d} \langle x, \theta \rangle^2 d\mu(x). \tag{9.2}$$

10. Appendix

We consider here some product probability measures on $\mathbb{R}^d$ whose marginals are mixtures of centered Gaussian variables. This class of probability measures was investigated in Eskenazis et al. (2018), where the authors prove that they satisfy (1.3) and provide interesting examples. Here we show that those measures also satisfy (1.2).

We consider Gaussian mixtures of the form

$$\mu = \iint_{\sigma \in (0,\infty)^d} \gamma_{\Gamma_\sigma} d\nu(\sigma), \tag{10.1}$$

where γ_{Γ_σ} is the centered Gaussian random vector in $\mathbb{R}^d$ with covariance matrix $\Gamma_\sigma = diag(\sigma_1^2, \ldots, \sigma_d^2)$ and where ν is also a product measure on $(0,\infty)^d$.

Theorem 10.1. *Let μ be a product probability measure on $\mathbb{R}^d$, whose marginals are mixture of centered Gaussian variables. Then (1.2) holds.*

The proof relies on the following Lemma, the key point of which being that no symmetry assumption is required in the convex situation.

Lemma 10.2. *The following points hold.*

(1) *Let g be a convex function on $\mathbb{R}^d$, then the function*

$$(\sigma_1, \ldots, \sigma_d) \in (0,\infty)^d \to \int g(y) d\gamma_{\Gamma_\sigma}(y)$$

is coordinatewise increasing on $(0,\infty)^d$.

(2) *Let f be a quasi-concave and even function on $\mathbb{R}^d$, then the function*

$$(\sigma_1, \ldots, \sigma_d) \in (0,\infty)^d \to \int f(y) d\gamma_{\Gamma_\sigma}(y)$$

is coordinatewise decreasing on $(0,\infty)^d$.

Proof of Theorem 10.1: Let f be a log-concave and even function and let g be a convex function. Using the decomposition of the covariance (8.6), one has

$$\begin{aligned} \mathrm{Cov}_\mu(f,g) = & \iint_{(0,\infty)^d} \mathrm{Cov}_{\gamma_{\Gamma_\sigma}}(f,g) d\nu(\sigma) \\ & + \mathrm{Cov}_\nu \left(\sigma \in (0,\infty)^d \to \int f d\gamma_{\Gamma_\sigma}, \sigma \in (0,\infty)^d \to \int g d\gamma_{\Gamma_\sigma} \right). \end{aligned}$$

The rest of the proof consists in showing that the two terms in the right-hand side of the latter inequality are non-positive. Firstly, Hargé's result (1.2) also applies to any (centered) Gaussian distribution (see Hargé (2008)) and thus $\mathrm{Cov}_{\gamma_{\Gamma_\sigma}}(f,g) \leq 0$. Secondly, we use Lemma 10.2, since f is log-concave and even, it is also quasi-concave and even, and thus the two functions

$$\sigma \in (0,\infty)^d \to \int g d\gamma_{\Gamma_\sigma} \text{ and } \sigma \in (0,\infty)^d \to \int f d\gamma_{\Gamma_\sigma}$$

are respectively coordinatewise increasing and coordinatewise decreasing on $(0,\infty)^2$. The measure ν being a product measure, by the FKG inequality for product measure, the term $\mathrm{Cov}_\nu(\cdot,\cdot)$ is non-positive and the result follows. □

We turn now to the proof of Lemma 10.2.

Proof of Lemma 10.2: Let g be a convex function on $\mathbb{R}^d$. By a change of variable, one directly has

$$\int g(y)d\gamma_{\Gamma_\sigma}(y) = \int_{\mathbb{R}^d} g(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x),$$

where we recall that γ is the standard Gaussian distribution. To prove the desired property, we compute for $1\le l\le d$,

$$\begin{aligned}\frac{\partial}{\partial\sigma_l}\int_{\mathbb{R}^d} g(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x) &= \int_{\mathbb{R}^d} x_l\,\partial_l g(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x)\\ &= \mathrm{Cov}_\gamma(x_l,\partial_l g(\sigma_1 x_1,\ldots,\sigma_d x_d)).\end{aligned}$$

Furthermore, by the covariance representation (8.4) for the standard Gaussian measure, one has

$$\mathrm{Cov}_\gamma(x_l,\partial_l g(\sigma_1 x_1,\ldots,\sigma_d x_d)) = \iint_{x,y\in\mathbb{R}^d}\kappa(x,y)\sigma_l\partial_{ll} g(\sigma_1 y_1,\ldots,\sigma_d y_d)dxdy$$

and this quantity is non-negative since g is convex and $\kappa(x,y)\ge 0$. The result follows.
For f quasi-concave and even, we use the layer cake representation of f:

$$f(x) = \int_0^\infty \mathbf{1}_{A_t}(x)dt \text{ and } A_t := \{x\in\mathbb{R}^d, f(x)\ge t\}.$$

Here by assumption the A_t are convex and even. Since by Fubini,

$$\int_{\mathbb{R}^d} f(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x) = \int_0^\infty\int_{\mathbb{R}^d}\mathbf{1}_{A_t}(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x)dt,$$

the result follows from Eskenazis et al. (2018) where the authors prove the following property: for each $t\ge 0$,

$$(\sigma_1,\ldots,\sigma_d)\in(0,\infty)^d\to\int_{\mathbb{R}^d}\mathbf{1}_{A_t}(\sigma_1 x_1,\ldots,\sigma_d x_d)d\gamma(x)$$

is coordinatewise decreasing. □

Acknowledgment

The authors would like to thank the anonymous reviewer for their careful reading and their helpful comments which improves the quality of the paper.

References

Arnaudon, M., Bonnefont, M., and Joulin, A. Intertwinings and generalized Brascamp-Lieb inequalities. *Rev. Mat. Iberoam.*, **34** (3), 1021–1054 (2018). MR3850277.

Bakry, D., Gentil, I., and Ledoux, M. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, Cham (2014). ISBN 978-3-319-00226-2; 978-3-319-00227-9. MR3155209.

Bakry, D. and Michel, D. Sur les inégalités FKG. In *Séminaire de Probabilités, XXVI*, volume 1526 of *Lecture Notes in Math.*, pp. 170–188. Springer, Berlin (1992). MR1231994.

Barthe, F. L'inégalité de corrélation Gaussienne [d'après Thomas Royen]. In *Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135*, 407, pp. Exp. No. 1124, 117–133. Société Mathématique de France (2019). ISBN 978-2-85629-897-8. MR3939275.

Bobkov, S. Extremal properties of half-spaces for log-concave distributions. *Ann. Probab.*, **24** (1), 35–48 (1996). MR1387625.

Bobkov, S. G., Götze, F., and Houdré, C. On Gaussian and Bernoulli covariance representations. *Bernoulli*, **7** (3), 439–451 (2001). MR1836739.

Carlen, E. A., Cordero-Erausquin, D., and Lieb, E. H. Asymmetric covariance estimates of Brascamp-Lieb type and related inequalities for log-concave measures. *Ann. Inst. Henri Poincaré Probab. Stat.*, **49** (1), 1–12 (2013). MR3060145.

Chatterjee, S. Stein's method for concentration inequalities. *Probab. Theory Related Fields*, **138** (1-2), 305–321 (2007). MR2288072.

Courtade, T. A., Fathi, M., and Pananjady, A. Existence of Stein kernels under a spectral gap, and discrepancy bounds. *Ann. Inst. Henri Poincaré Probab. Stat.*, **55** (2), 777–790 (2019). MR3949953.

Eskenazis, A., Nayar, P., and Tkocz, T. Gaussian mixtures: entropy and geometric inequalities. *Ann. Probab.*, **46** (5), 2908–2945 (2018). MR3846841.

Fathi, M. Stein kernels and moment maps. *Ann. Probab.*, **47** (4), 2172–2185 (2019). MR3980918.

Fortuin, C. M., Kasteleyn, P. W., and Ginibre, J. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, **22**, 89–103 (1971). MR309498.

Hargé, G. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields*, **130** (3), 415–440 (2004). MR2095937.

Hargé, G. Reinforcement of an inequality due to Brascamp and Lieb. *J. Funct. Anal.*, **254** (2), 267–300 (2008). MR2376572.

Houdré, C. and Privault, N. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, **8** (6), 697–720 (2002). MR1962538.

Hu, Y. Itô-Wiener Chaos Expansion with Exact Residual and Correlation, Variance Inequalities. *J. Theoret. Probab.*, **10** (4), 835–848 (1997). DOI: 10.1023/A:1022654314791.

Karlin, S. *Total positivity. Vol. I.* Stanford University Press, Stanford, CA (1968). MR230102.

Latała, R. and Matlak, D. Royen's proof of the Gaussian correlation inequality. In *Geometric aspects of functional analysis*, volume 2169 of *Lecture Notes in Math.*, pp. 265–275. Springer, Cham (2017). MR3645127.

Ledoux, M. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2001a). ISBN 0-8218-2864-9. MR1849347.

Ledoux, M. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pp. 167–194. Springer, Berlin (2001b). MR1837286.

Ledoux, M., Nourdin, I., and Peccati, G. Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.*, **25** (1), 256–306 (2015). MR3320893.

Ledoux, M. and Talagrand, M. *Probability in Banach spaces. Isoperimetry and processes.* Classics in Mathematics. Springer-Verlag, Berlin (2011). ISBN 978-3-642-20211-7. MR2814399.

Menz, G. and Otto, F. Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.*, **41** (3B), 2182–2224 (2013). MR3098070.

Nourdin, I. and Viens, F. G. Density formula and concentration inequalities with Malliavin calculus. *Electron. J. Probab.*, **14**, no. 78, 2287–2309 (2009). MR2556018.

Pinkus, A. *Totally positive matrices*, volume 181 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-19408-2. MR2584277.

Pitt, L. D. Positively correlated normal variables are associated. *Ann. Probab.*, **10** (2), 496–499 (1982). MR665603.

Royen, T. A simple proof of the Gaussian correlation conjecture extended to some multivariate gamma distributions. *Far East J. Theor. Stat.*, **48** (2), 139–145 (2014). MR3289621.

Saumard, A. Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one. *Bernoulli*, **25** (4B), 3978–4006 (2019). MR4010979.

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

Schechtman, G., Schlumprecht, T., and Zinn, J. On the Gaussian measure of the intersection. *Ann. Probab.*, **26** (1), 346–357 (1998). MR1617052.

Tong, Y. L. *The Multivariate Normal Distribution.* Springer Series in Statistics. Springer New York, NY (1990). ISBN 978-1-4613-9657-4. DOI: 10.1007/978-1-4613-9655-0.