

Geometric influences II: Correlation inequalities and noise sensitivity

Nathan Keller^{a,1}, Elchanan Mossel^{b,2} and Arnab Sen^{c,3}

^a*Department of Mathematics, Bar Ilan University, Ramat Gan, Israel. E-mail: nathan.keller@math.biu.ac.il*

^b*Statistics Department, University of Pennsylvania, 400 Huntsman Hall, Philadelphia, PA 19104, USA and Departments of Statistics and Computer Science, U.C. Berkeley, 367 Evans Hall, Berkeley, CA 94720, USA. E-mail: mossel@stat.berkeley.edu*

^c*Department of Mathematics, University of Minnesota, 238 Vincent Hall, 206 Church St SE, Minneapolis, MN 55455, USA. E-mail: a.sen@statslab.cam.ac.uk*

Received 10 June 2012; accepted 17 March 2013

Abstract. In a recent paper, we presented a new definition of influences in product spaces of continuous distributions, and showed that analogues of the most fundamental results on discrete influences, such as the KKL theorem, hold for the new definition in Gaussian space. In this paper we prove Gaussian analogues of two of the central applications of influences: Talagrand's lower bound on the correlation of increasing subsets of the discrete cube, and the Benjamini–Kalai–Schramm (BKS) noise sensitivity theorem. We then use the Gaussian results to obtain analogues of Talagrand's bound for all discrete probability spaces and to reestablish analogues of the BKS theorem for biased two-point product spaces.

Résumé. Dans un papier récent, nous avons présenté une nouvelle définition de l'influence dans des produits d'espaces de fonctions continues et montré que des résultats analogues aux résultats les plus importants sur les influences discrètes, comme le théorème KKL, sont valables pour la nouvelle définition dans des espaces gaussiens. Dans cet article, nous prouvons des analogues gaussiens de deux des applications principales des influences : la borne inférieure de Talagrand sur la corrélation de sous-ensembles croissants du cube discret et le théorème de Benjamini–Kalai–Schramm (BKS) sur la sensibilité au bruit. Ensuite nous utilisons les résultats gaussiens pour obtenir des analogues de la borne de Talagrand pour tous les espaces de probabilités discrets et pour retrouver l'analogue du théorème BKS pour des espaces produits biaisés à deux points.

MSC: 60C05; 05D40

Keywords: Influences; Geometric influences; Noise sensitivity; Correlation between increasing sets; Talagrand's bound; Gaussian measure; Isoperimetric inequality

1. Introduction

Definition 1.1. Consider the discrete cube $\{-1, 1\}^n$ endowed with the uniform measure $\nu^{\otimes n} = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$, and let $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. The influence of the i th coordinate on f is defined as

$$I_i(f) := \mathbb{E}_\nu \left[|f(X) - f(X^{[i]})| \right], \quad (1.1)$$

¹Part of the work was done while the author was with the Weizmann Institute of Science and was supported by the Koshland Center for Basic Research.

²Supported by NSF Grant DMS 1106999 and CCF 1320105 DOD ONR Grant N00014-14-1-0823, and A & G Simons collaborative Grant #328025.

³Supported by EPSRC grant EP/G055068/1.

where $X = (X_1, \dots, X_n)$ is a random vector in $\{-1, 1\}^n$ distributed according to the measure $\nu^{\otimes n}$, and $X^{[i]}$ denotes the vector obtained from X by replacing X_i by $-X_i$ and leaving the other coordinates unchanged. The subscript ν in \mathbb{E}_ν emphasizes the fact that the expectation is taken w.r.t. the measure $\nu^{\otimes n}$. For a subset A of the discrete cube $\{-1, 1\}^n$, we write $I_i(A)$ as a shorthand for $I_i(1_A)$, and refer to it as the influence of the i th coordinate on A .

The notion of influences of variables on Boolean functions is one of the central concepts in the theory of discrete harmonic analysis. In the last two decades it found several applications in diverse fields, including Combinatorics, Theoretical Computer Science, Statistical Physics, Social Choice Theory, etc. (see, for example, the survey articles [16,27]).

Two of the central applications are Talagrand’s lower bound on the correlation between increasing subsets of the discrete cube [29] and the Benjamini–Kalai–Schramm (BKS) theorem on noise sensitivity [3].

Talagrand’s result is an improvement over the classical Harris–Kleitman correlation inequality [13,23] stating that any two increasing (see Definition 4.1 below) subsets of the discrete cube are non-negatively correlated.

Theorem 1.2 (Talagrand). *For any pair of increasing subsets $A, B \subset \{-1, 1\}^n$,*

$$\nu^{\otimes n}(A \cap B) - \nu^{\otimes n}(A)\nu^{\otimes n}(B) \geq c\varphi\left(\sum_{i=1}^n I_i(A)I_i(B)\right),$$

where $\varphi(x) = x / \log(e/x)$, and $c > 0$ is a universal constant.

The BKS theorem deals with the sensitivity of a Boolean function (or equivalently, a subset of the discrete cube) to a small random perturbation of its input.

Definition 1.3. *For a function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$, and for $\eta \in (0, 1)$, let*

$$Z(f, \eta) = \mathbb{E}[f(X)f(X^\eta)],$$

where $X = (X_1, \dots, X_n)$ is uniformly distributed in $\{-1, 1\}^n$ and $X^\eta = (X_1^\eta, \dots, X_n^\eta)$ is a $(1 - \eta)$ -correlated copy of X . (This means that for $j \in \{1, 2, \dots, n\}$, $X_j^\eta = X_j$ with probability $1 - \eta$ and $X_j^\eta = X'_j$ with probability η , independently for distinct j ’s, where $X' = (X'_1, \dots, X'_n)$ is an i.i.d. copy of X .) Following Benjamini, Kalai and Schramm [3], we denote

$$\text{VAR}(f, \eta) = Z(f, \eta) - \mathbb{E}[f(X)]^2.$$

For a set $B \subseteq \{-1, 1\}^n$, and for $\eta \in (0, 1)$, we write

$$Z(B, \eta) = Z(1_B, \eta) \quad \text{and} \quad \text{VAR}(B, \eta) = \text{VAR}(1_B, \eta).$$

A sequence of sets $B_\ell \subseteq \{-1, 1\}^{\ell}$ is said to be asymptotically noise sensitive if

$$\lim_{\ell \rightarrow \infty} \text{VAR}(B_\ell, \eta) = 0 \quad \text{for each } \eta \in (0, 1). \tag{1.2}$$

In a seminal paper, Benjamini, Kalai and Schramm [3] proved that a sequence of sets $B_\ell \subseteq \{-1, 1\}^{\ell}$ is asymptotically noise sensitive if the sum of the squares of the influences $\sum_{i=1}^{\ell} I_i(B_\ell)^2$ goes to zero as $\ell \rightarrow \infty$. Recently, Keller and Kindler [20] obtained a quantitative version of the BKS theorem.

Theorem 1.4 (Quantitative BKS theorem). *For any n , for any function $f : \{-1, 1\}^n \rightarrow [0, 1]$, and for any $\eta \in (0, 1)$,*

$$\text{VAR}(f, \eta) \leq c_1 \cdot \left(\sum_{i=1}^n I_i(f)^2\right)^{c_2 \cdot \eta},$$

where c_1, c_2 are positive universal constants.

The basic results on influences were obtained for functions on the discrete cube, but some applications required generalization of the results to more general product spaces. Unlike the discrete case, where there exists a single natural definition of influence, for general product spaces several definitions were presented in different papers, see for example [7,14,18]. In [21], we presented a new notion of influences in product spaces of continuous distributions, which we called geometric influences, and proved analogues of the fundamental results on influences, such as the Kahn–Kalai–Linial (KKL) theorem [15] and Talagrand’s influence sum bound [28], for geometric influences.

In this paper we prove analogues of Talagrand’s lower bound on the correlation of increasing sets (Theorem 1.2 above) and of the quantitative BKS theorem (Theorem 1.4 above), that hold for the standard Gaussian measure in \mathbb{R}^n with respect to geometric influences.

Definition 1.5. Let $\mu(dx) = (1/\sqrt{2\pi}) \exp(-x^2/2) dx$ be the standard Gaussian measure on \mathbb{R} . Let ϕ (resp. Φ) be the density (resp. distribution function) of the Gaussian measure μ on \mathbb{R} , and denote $\bar{\Phi}(x) = 1 - \Phi(x)$. Given a Borel-measurable set $A \subseteq \mathbb{R}$, its lower Minkowski content $\mu^+(A)$ is defined as

$$\mu^+(A) := \liminf_{r \downarrow 0} \frac{\mu(A + [-r, r]) - \mu(A)}{r}.$$

For any Borel-measurable set $A \subseteq \mathbb{R}^n$, for each $1 \leq i \leq n$ and an element $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, the restriction of A along the fiber of x in the i th direction is given by

$$A_i^x := \{y \in \mathbb{R} : (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \in A\}.$$

The geometric influence of the i th coordinate on A is

$$I_i^{\mathcal{G}}(A) := \mathbb{E}_x[\mu^+(A_i^x)],$$

that is, the expectation of $\mu^+(A_i^x)$ when x is chosen according to the measure μ .

We note that the geometric meaning of the influence is that for a monotone (either increasing or decreasing) set A , the sum of influences of A is equal to the size of its boundary with respect to a uniform enlargement (see [21]).

In the sequel, whenever we talk about sets or functions in \mathbb{R}^n , we implicitly assume that they are Borel measurable. Our first result is a lower bound on the correlation between two increasing bounded functions in the Gaussian space.

Theorem 1.6. Let $\varphi(x) = x/\log(e/x)$. There exists a universal constant $c > 0$ such that for any $n \in \mathbb{N}$ and for any two increasing subsets A and B of \mathbb{R}^n , we have

$$\mu^{\otimes n}(A \cap B) - \mu^{\otimes n}(A)\mu^{\otimes n}(B) \geq c\varphi\left(\sum_{i=1}^n I_i^{\mathcal{G}}(A)I_i^{\mathcal{G}}(B)\right).$$

We show that the assertion of the theorem is tight, up to the constant factor. The proof of Theorem 1.6 uses Talagrand’s result for the discrete cube, along with appropriate limit arguments. By appealing to direct Gaussian arguments, we obtain another lower bound on the correlation between a pair of increasing subsets in the Gaussian space.

Theorem 1.7. There exists a universal constant $c > 0$ such that for any $n \in \mathbb{N}$ and for any two increasing subsets A and B of \mathbb{R}^n , we have

$$\mu^{\otimes n}(A \cap B) - \mu^{\otimes n}(A)\mu^{\otimes n}(B) \geq c \sum_{i=1}^n \frac{I_i^{\mathcal{G}}(A)I_i^{\mathcal{G}}(B)}{\sqrt{\log(e/I_i^{\mathcal{G}}(A)) \log(e/I_i^{\mathcal{G}}(B))}}.$$

In fact, we prove functional versions of the above two theorems (see Theorem 2.1 and Theorem 3.1), which, with a little bit of extra work, can then be applied to deduce the results for the characteristic functions of increasing sets.

Theorem 1.7 is neither uniformly stronger nor uniformly weaker than Theorem 1.6, as there are cases where each one beats the other. It should be noted that while Talagrand’s lower bound uses the classical Bonami–Beckner hypercontractive inequality [2,4], the proof of Theorem 1.7 uses Borell’s reverse hypercontractive inequality [5]. It will be interesting to find out whether hypercontractivity and reverse hypercontractivity can be combined to obtain a new lower bound that will enjoy the benefits of both Theorems 1.6 and 1.7.

Recall that the classical Gaussian FKG inequality [10] asserts that for any pair of coordinate-wise increasing functions $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$, we have

$$\mathbb{E}_\mu[fg] \geq \mathbb{E}_\mu[f]\mathbb{E}_\mu[g].$$

Hence, Theorems 1.6 and 1.7 (or more appropriately their functional versions) provide quantitative versions of the Gaussian FKG inequality.

Our second result is a Gaussian analogue of the noise sensitivity results of Benjamini–Kalai–Schramm [3].

Definition 1.8. Let W, W' be i.i.d. standard Gaussian vectors on \mathbb{R}^n and let $W^\rho = \sqrt{1 - \rho^2}W + \rho W'$. For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, and for $\rho \in (0, 1)$, let

$$Z^{\mathcal{G}}(f, \rho) = \mathbb{E}[f(W)f(W^\rho)],$$

provided $\mathbb{E}[|f(W)|^2] < \infty$. Denote

$$\text{VAR}^{\mathcal{G}}(f, \rho) = Z^{\mathcal{G}}(f, \rho) - \mathbb{E}[f(W)]^2.$$

For a set $A \subset \mathbb{R}^n$, and for $\rho \in (0, 1)$, we write

$$Z^{\mathcal{G}}(A, \rho) = Z(1_A, \rho), \quad \text{and} \quad \text{VAR}^{\mathcal{G}}(A, \rho) = \text{VAR}^{\mathcal{G}}(1_A, \rho).$$

A sequence of sets $A_\ell \subseteq \mathbb{R}^{n_\ell}$ is said to be asymptotically Gaussian noise-sensitive if

$$\lim_{\ell \rightarrow \infty} \text{VAR}^{\mathcal{G}}(A_\ell, \rho) = 0 \quad \text{for each } \rho \in (0, 1). \tag{1.3}$$

Theorem 1.9. For any $n \geq 1$, for any set $A \subset \mathbb{R}^n$, and for any $\rho \in (0, 1)$,

$$\text{VAR}^{\mathcal{G}}(A, \rho) \leq C_1 \cdot \left(\sum_{i=1}^n I_i^{\mathcal{G}}(A)^2 \right)^{C_2 \rho^2},$$

where C_1, C_2 are positive universal constants.

The proof of Theorem 1.9 again relies upon an appropriate limit argument and uses Theorem 1.4 as a blackbox.

Theorems 1.6 and 1.7 allow us to obtain analogues of Talagrand’s lower bounds for any discrete product probability space (see Theorem 5.2), where the lower bound involves a discrete variant of the geometric influence, called h -influence. Theorem 1.9 can be used to obtain an analogue of the BKS theorem in the case of the discrete hypercube $\{0, 1\}^n$ endowed with a biased product measure (see Theorem 5.4). We note that for the biased product measures on the discrete hypercube, these results were previously obtained in [19,20] by different methods. Comparison of our results with the results of [19,20] suggests that, in some sense, the h -influence obtained from the geometric influence is more natural than the notion of influences used for the biased measure in previous works.

This paper is organized as follows. In Section 2 we prove functional versions of Theorem 1.6 and Theorem 1.9. In Section 3, we present a functional version of Theorem 1.7 using the Ornstein–Uhlenbeck semigroup theory. In Section 4 we give an argument to suitably approximate the characteristic functions of monotone sets by smooth functions and apply it to deduce Theorems 1.6, 1.7 and 1.9 from their functional counterparts. We also discuss how Theorem 1.6 and Theorem 1.7 compare against each other. Finally, we deduce the analogous statements for discrete product probability spaces in Section 5, and conclude the paper with a few open problems in Section 6.

2. Refined Gaussian FKG inequality and Gaussian BKS theorem

The main goal of this section is to we prove the following two theorems which are functional forms of Theorem 1.6 and Theorem 1.9. Note that the role of Gaussian influences is now played by the L^1 norm of the partial derivatives of the functions.

Theorem 2.1. *Let $\varphi(x) = x/\log(e/x)$. There exists a universal constant $c > 0$ such that for any $n \geq 1$ and for any two increasing continuously differentiable functions $f, g : \mathbb{R}^n \rightarrow [-1, 1]$, we have*

$$\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g] \geq c\varphi\left(\sum_{i=1}^n \mathbb{E}_\mu[|\partial_i f|]\mathbb{E}_\mu[|\partial_i g|]\right),$$

where \mathbb{E}_μ stands for integration w.r.t. $\mu^{\otimes n}$.

Theorem 2.2. *For any $n \geq 1$, for any continuously differentiable function $f : \mathbb{R}^n \rightarrow [-1, 1]$, and for any $\rho \in (0, 1)$,*

$$\text{VAR}^G(f, \rho) \leq C_1 \cdot \left(\sum_{i=1}^n \mathbb{E}_\mu[|\partial_i f|]^2\right)^{C_2\rho^2},$$

where C_1, C_2 are positive universal constants.

The proof strategy is to approximate the functions in the ‘‘Gaussian world’’ by sequences of functions defined on the discrete cubes $\{-1, 1\}^{n_\ell}$ (where $n_\ell \rightarrow \infty$), and to deduce the assertions of the theorems by an appropriate limit argument from the corresponding theorems in the ‘‘discrete world.’’

For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, we construct a sequence $\{\check{f}_m\}_{m=1}^\infty$ of functions as follows. For each $m \in \mathbb{N}$, we denote elements in $\{-1, 1\}^{mn}$ by vectors (x_1, x_2, \dots, x_n) , where each $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ is a vector in $\{-1, 1\}^m$. We write $s_i = s_i(m)$ as a shorthand for $m^{-1/2} \sum_{j=1}^m x_{ij}$ and let $s = (s_1, \dots, s_n) \in \mathbb{R}^n$. Then, we define the function $\check{f}_m : \{-1, 1\}^{mn} \rightarrow \mathbb{R}$ by $\check{f}_m(x_1, \dots, x_n) = f(s_1, \dots, s_n)$. In order to simplify the notation, we leave the dependence of s on m implicit in some of the places, and alert the reader that in the sequel, s always depends on m . The next lemma is our main tool for transferring the results from the discrete world to the Gaussian world.

Lemma 2.3. *Fix $n \geq 1$ and $1 \leq i \leq n$. Let f and g be two continuously differentiable functions on \mathbb{R}^n such that the partial derivatives $\partial_i f$ and $\partial_i g$ are bounded. Then*

$$\sum_{j=1}^m I_{ij}(\check{f}_m)I_{ij}(\check{g}_m) \rightarrow 4\mathbb{E}_\mu[|\partial_i f|]\mathbb{E}_\mu[|\partial_i g|], \quad \text{as } m \rightarrow \infty.$$

Proof. Since the functions \check{f}_m and \check{g}_m are invariant under permutations of the coordinates $\{x_{ij}\}_{1 \leq j \leq m}$ for each fixed i , it follows that $\sum_{j=1}^m I_{ij}(\check{f}_m)I_{ij}(\check{g}_m) = mI_{i1}(\check{f}_m)I_{i1}(\check{g}_m)$. Thus, it suffices to show that $\sqrt{m}I_{i1}(\check{f}_m) \rightarrow 2\mathbb{E}_\mu[|\partial_i f|]$ and similarly for g . Without loss of generality, we take $i = 1$. We have

$$I_{11}(\check{f}_m) = \mathbb{E}_\nu[|f(s'_1 + m^{-1/2}, s_2, \dots, s_n) - f(s'_1 - m^{-1/2}, s_2, \dots, s_n)|],$$

where $s'_1 = m^{-1/2} \sum_{j=2}^m x_{1j}$. By the Mean Value Theorem,

$$\frac{f(s'_1 + m^{-1/2}, s_2, \dots, s_n) - f(s'_1 - m^{-1/2}, s_2, \dots, s_n)}{2m^{-1/2}} = \partial_1 f(s'_1 + \varepsilon_m, s_2, \dots, s_m),$$

where ε_m is an error term that depends on s'_1, s_2, \dots, s_n , and whose absolute value is bounded by $m^{-1/2}$. Therefore, we obtain

$$\sqrt{m}I_{11}(\check{f}_m) = 2\mathbb{E}_\nu[|\partial_1 f(s'_1 + \varepsilon_m, s_2, \dots, s_m)|].$$

Since $(s'_1 + \varepsilon_m, s_2, \dots, s_m)$ converges in distribution to $\mu^{\otimes n}$, and since $\partial_1 f$ is a continuous, bounded function, we conclude that

$$\lim_{m \rightarrow \infty} \sqrt{m} I_{11}(\check{f}_m) = 2\mathbb{E}_\mu[\partial_1 f].$$

The assertion of the lemma follows. □

To prove Theorem 2.1, we will need the following functional version of Talagrand’s inequality on the discrete cube.

Theorem 2.4. *For any $n \geq 1$ and for any pair of increasing functions $f, g : \{-1, 1\}^n \rightarrow [0, 1]$,*

$$\mathbb{E}_\nu[fg] - \mathbb{E}_\nu[f]\mathbb{E}_\nu[g] \geq c\varphi\left(\sum_{i=1}^n I_i(f)I_i(g)\right),$$

where $\varphi(x) = x / \log(e/x)$, and $c > 0$ is a universal constant.

This version is obtained by following Talagrand’s proof step-by-step, using the fact that for a monotone function f , $I_i(f)$ is equal in absolute value to the coefficient $\hat{f}(\{i\})$ in the standard Fourier–Walsh expansion of f . The exact proof (of a slightly more general statement) appears in [17].

Proof of Theorem 2.1. Note that since f, g are increasing and bounded, by the Fundamental Theorem of Calculus, $\partial_i f, \partial_i g$ are nonnegative and integrable. In particular, we have $0 \leq \mathbb{E}_\mu[\partial_i f], \mathbb{E}_\mu[\partial_i g] < \infty$ for all i . First we assume that f, g are increasing C^1 functions on \mathbb{R}^n such that both f, g take values in $[0, 1]$ and $\|\partial_i f\|_\infty, \|\partial_i g\|_\infty < \infty$ for all i . It follows from Theorem 2.4 that there exists a universal constant $c > 0$ such that for each $m \in \mathbb{N}$, we have

$$\int \check{f}_m \check{g}_m \, d\nu^{\otimes nm} - \int \check{f}_m \, d\nu^{\otimes nm} \int \check{g}_m \, d\nu^{\otimes nm} \geq c\varphi\left(\sum_{i=1}^n \sum_{j=1}^m I_{ij}(\check{f}_m)I_{ij}(\check{g}_m)\right). \tag{2.1}$$

By the Central Limit Theorem, $s(m) = (s_1, \dots, s_n)$ converges in distribution to $\mu^{\otimes n}$ as $m \rightarrow \infty$. Thus, the left hand side of (2.1) converges to $\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g]$ as $m \rightarrow \infty$. On the other hand, by letting $m \rightarrow \infty$ and applying Lemma 2.3 to the right hand side of (2.1), we obtain

$$\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g] \geq c\varphi\left(\sum_{i=1}^n \mathbb{E}_\mu[\partial_i f]\mathbb{E}_\mu[\partial_i g]\right). \tag{2.2}$$

We can easily extend the above inequality, with the constant c replaced by a new constant $c/(1 + \log 2)$, to increasing C^1 functions f, g such that both f, g take values in $[-1, 1]$ and $\|\partial_i f\|_\infty, \|\partial_i g\|_\infty < \infty$ for all i . To do that we apply (2.2) for the functions $(1 + f)/2, (1 + g)/2$ and note that $2\varphi(x/2) \geq \frac{1}{1+\log 2}\varphi(x)$ for all $x \in [0, 1]$.

Now we want to remove the condition that the partial derivatives of f, g are bounded. Let f, g be as given in the hypothesis of Theorem 2.1. For $K > 0$, set $J_K = [-K, K]^n, M_K = f(K, \dots, K)$, and $m_K = f(-K, \dots, -K)$. Since f is increasing, $M_K = \max_{x \in J_K} f(x)$ and $m_K = \min_{x \in J_K} f(x)$. Let $f_K = \min(\max(f, m_K), M_K)$. Hence, $f_K \equiv f$ inside J_K . Let $\eta \in C^\infty(\mathbb{R}^n)$ be the standard mollifier, that is, $\eta(x) = C \exp(\frac{1}{|x|^2 - 1}) \mathbf{1}_{|x| \leq 1}$, where the constant $C > 0$ is selected so that $\int_{\mathbb{R}^n} \eta(x) \, dx = 1$. For each $\varepsilon > 0$, set $\eta_\varepsilon(x) := \varepsilon^{-n} \eta(x/\varepsilon)$. Finally, define $f_{K,\varepsilon} = f_K * \eta_\varepsilon = \int_{\mathbb{R}^n} f_K(x - y) \eta_\varepsilon(y) \, dy$. From the standard properties of the mollifier, it follows that $f_{K,\varepsilon} \in C^\infty(\mathbb{R}^n)$, $f_{K,\varepsilon}$ is increasing and $|f_{K,\varepsilon}| \leq 1$. Note that for any $h \in \mathbb{R}$, for any $z \in \mathbb{R}^n$,

$$0 \leq \frac{f_K(z + e_i h) - f_K(z)}{h} \leq \frac{f(z + e_i h) - f(z)}{h},$$

e_i being the i th coordinate vector in \mathbb{R}^n . It follows that $0 \leq \partial_i f_{K,\varepsilon} \leq \partial_i (f * \eta_\varepsilon) = \partial_i f * \eta_\varepsilon$.

Given $\delta > 0$, we claim that there exist $K > 0$ and $\varepsilon > 0$ such that $\int_{\mathbb{R}^n} |f_{K,\varepsilon} - f|^2 d\mu^{\otimes n} < \delta$ and $\int_{\mathbb{R}^n} |\partial_i f_{K,\varepsilon} - \partial_i f| d\mu^{\otimes n} < \delta$. To prove the claim, first find $K > 0$ large such that $|\int_{J_{K-1/2}^c} \partial_i f d\mu^{\otimes n}| < \delta/3$. For $0 < \varepsilon < 1/2$, $\partial_i f_K = 0$ outside $J_{K+1/2}$ and we estimate

$$\begin{aligned} \int_{\mathbb{R}^n} |\partial_i f_{K,\varepsilon} - \partial_i f| d\mu^{\otimes n} &\leq \int_{J_{K+1/2}} |\partial_i f_{K,\varepsilon} - \partial_i f * \eta_\varepsilon| d\mu^{\otimes n} + \int_{J_{K+1/2}} |\partial_i f * \eta_\varepsilon - \partial_i f| d\mu^{\otimes n} \\ &\quad + \int_{J_{K+1/2}^c} \partial_i f d\mu^{\otimes n}. \end{aligned}$$

Note that whenever $\varepsilon \in (0, 1/2)$, $\partial_i f_{K,\varepsilon} = \partial_i f * \eta_\varepsilon$ on $J_{K-1/2}$. Hence,

$$\int_{J_{K+1/2}} |\partial_i f_{K,\varepsilon} - \partial_i f * \eta_\varepsilon| d\mu^{\otimes n} \leq \int_{J_{K+1/2} \cap J_{K-1/2}^c} |\partial_i f * \eta_\varepsilon| d\mu^{\otimes n}.$$

By the well-known property of the mollifier, $\partial_i f * \eta_\varepsilon \xrightarrow{L^p} \partial_i f$ for any $1 \leq p < \infty$ over compact sets. Thus, by choosing $\varepsilon > 0$ small we can make $\int_{J_{K+1/2}} |\partial_i f * \eta_\varepsilon - \partial_i f| d\mu^{\otimes n} < \delta/3$ and $\int_{J_{K+1/2} \cap J_{K-1/2}^c} |\partial_i f * \eta_\varepsilon| d\mu^{\otimes n} \leq \int_{J_{K+1/2} \cap J_{K-1/2}^c} |\partial_i f| d\mu^{\otimes n} + \delta/3$ and hence, $\int_{\mathbb{R}^n} |\partial_i f_{K,\varepsilon} - \partial_i f| d\mu^{\otimes n} < \delta$.

On the other hand, note that

$$\int_{\mathbb{R}^n} |f_{K,\varepsilon} - f|^2 d\mu^{\otimes n} \leq 2 \int_{\mathbb{R}^n} |f_{K,\varepsilon} - f * \eta_\varepsilon|^2 d\mu^{\otimes n} + 2 \int_{\mathbb{R}^n} |f * \eta_\varepsilon - f|^2 d\mu^{\otimes n}.$$

For $\varepsilon > 0$ fixed, $\int_{\mathbb{R}^n} |f_{K,\varepsilon} - f * \eta_\varepsilon|^2 d\mu^{\otimes n} \rightarrow 0$ as $K \rightarrow \infty$ by dominated convergence. Since $f * \eta_\varepsilon \rightarrow f$ pointwise as $\varepsilon \rightarrow 0$, the second integral also goes to zero by dominated convergence. Thus we establish our claim.

Now note that (2.2) holds for functions $f_{K,\varepsilon}$ and $g_{K,\varepsilon}$. We complete the proof of the theorem by approximating the original functions f and g by $f_{K,\varepsilon}$ and $g_{K,\varepsilon}$ with suitably large K and small ε . \square

Now we prove Theorem 2.2, thus obtaining a Gaussian analogue of the quantitative BKS theorem [20].

Proof of Theorem 2.2. Assume first that f is continuously differentiable with bounded partial derivatives. We apply Theorem 1.4 to the approximating functions $\check{f}_m : \{-1, 1\}^{nm} \rightarrow [0, 1]$ to obtain, for any $\eta \in (0, 1)$ and any $m \geq 1$,

$$\text{VAR}(\check{f}_m, \eta) \leq c_1 \cdot \left(\sum_{i=1}^n \sum_{j=1}^m I_{ij}(\check{f}_m)^2 \right)^{c_2 \cdot \eta}, \quad (2.3)$$

where $c_1 > 0, c_2 > 0$ are universal constants. We claim that $\text{VAR}(\check{f}_m, 1 - \sqrt{1 - \rho^2}) \rightarrow \text{VAR}^{\mathcal{G}}(f, \rho)$ as $m \rightarrow \infty$. Let $\mathbf{x} = (x_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ and $\mathbf{y} = (y_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ be $\sqrt{1 - \rho^2}$ correlated vectors that are uniformly distributed in $\{-1, 1\}^{nm}$. Set $s(m) = (s_1, \dots, s_n)$ and $s^\rho(m) = (s_1^\rho, \dots, s_n^\rho)$ where $s_i = m^{-1/2} \sum_{j=1}^m x_{ij}$ and $s_i^\rho = m^{-1/2} \sum_{j=1}^m y_{ij}$. By definition, $Z(\check{f}_m, 1 - \sqrt{1 - \rho^2}) = \mathbb{E}_v[f(s(m))f(s^\rho(m))]$. By the Central Limit Theorem, $(s(m), s^\rho(m))$ converges in distribution to (W, W^ρ) as $m \rightarrow \infty$. Since the map $(z, z') \mapsto f(z)f(z')$ is bounded and continuous on \mathbb{R}^{2n} , it follows that $\lim_{m \rightarrow \infty} Z(\check{f}_m, 1 - \sqrt{1 - \rho^2}) = Z^{\mathcal{G}}(f, \rho)$. That $\mathbb{E}_v[\check{f}_m(\mathbf{x})] = \mathbb{E}_v[f(s(m))]$ converges to $\mathbb{E}_\mu[f(W)]$ as $m \rightarrow \infty$ is again an immediate consequence of the Central Limit Theorem. This proves the claim. By letting $m \rightarrow \infty$ in (2.3) with $\eta = 1 - \sqrt{1 - \rho^2} \in (0, 1)$ and by virtue of the above claim and Lemma 2.3, we obtain the following inequality for the function f with $C_2 = c_2$ and $C'_1 = 4c_1$,

$$\text{VAR}^{\mathcal{G}}(f, \rho) \leq C'_1 \cdot \left(\sum_{i=1}^n \mathbb{E}_\mu[|\partial_i f|^2] \right)^{C'_2(1 - \sqrt{1 - \rho^2})}. \quad (2.4)$$

Extending (2.4) to C^1 functions f with bounded partial derivatives which take values in $[-1, 1]$ instead of $[0, 1]$ is fairly straightforward and can be achieved (with $C_1 = 2C'_1$) by arguing with the function $(1 + f)/2$ which now takes

values in $[0, 1]$. If $\sum_{i=1}^n \mathbb{E}_\mu[|\partial_i f|^2] \leq 1$, then by observing the simple fact that $1 - \sqrt{1 - \rho^2} \geq \rho^2/2$ for all $\rho \in (0, 1)$, we get the desired inequality (with $C_2 = C'_2/2$) for the function f . On the other hand, if $\sum_{i=1}^n \mathbb{E}_\mu[|\partial_i f|^2] > 1$, then the assertion of the theorem trivially holds for f since $\text{VAR}^G(f, \rho) \leq |f| \leq 1$.

Now take a general C^1 function $f : \mathbb{R}^n \rightarrow [-1, 1]$. If $\mathbb{E}_\mu[|\partial_i f|] = \infty$ for some i , then the theorem holds trivially. So, assume that $\mathbb{E}_\mu[|\partial_i f|] < \infty$ for all i . Let η, η_ε and J_K be as above. Define $f_K = f \mathbf{1}_{J_K}$ and $f_{K,\varepsilon} = f_K * \eta_\varepsilon$. Clearly, $f_{K,\varepsilon}$ is C^∞ and $|f_{K,\varepsilon}| \leq 1$ and $|\partial_i f_{K,\varepsilon}|$ are bounded for all i (since $f_{K,\varepsilon}$ is compactly supported). Note that as $K \rightarrow \infty, \varepsilon \rightarrow 0$, $f_{K,\varepsilon}(z) \rightarrow f(z)$ pointwise, and hence by dominated convergence, $\text{VAR}^G(f_{K,\varepsilon}, \rho) \rightarrow \text{VAR}^G(f, \rho)$. Next we prove that $\mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} - \partial_i f|] \rightarrow 0$ as $K \rightarrow \infty, \varepsilon \rightarrow 0$. Towards this end, we bound

$$\begin{aligned} \mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} - \partial_i f|] &\leq \mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} - \partial_i f \mathbf{1}_{J_{K-\varepsilon}}|] + \mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} - \partial_i f \mathbf{1}_{J_{K+\varepsilon} \cap J_{K-\varepsilon}^c}|] \\ &\quad + \mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} - \partial_i f \mathbf{1}_{J_{K+\varepsilon}^c}|] \\ &\leq \mathbb{E}_\mu[|\partial_i (f * \eta_\varepsilon) - \partial_i f \mathbf{1}_{J_K}|] + \mathbb{E}_\mu[|\partial_i f_{K,\varepsilon} \mathbf{1}_{J_{K+\varepsilon} \cap J_{K-\varepsilon}^c}|] + \mathbb{E}_\mu[|\partial_i f \mathbf{1}_{J_K^c}|]. \end{aligned} \tag{2.5}$$

Note that $\partial_i f_{K,\varepsilon}(x) = \int f_K(y) \partial_i \eta_\varepsilon(x - y) dy = \varepsilon^{-1} \int \partial_i \eta(z) f_K(x - \varepsilon z) dz$. Since $|f_K|$ is bounded by 1, and $\int |\partial_i \eta(z)| dz < \infty$, we have $|\partial_i f_{K,\varepsilon}| \leq C'\varepsilon$. Thus the second expectation in (2.5) can be bounded above by $C'\varepsilon^{-1} \mu^{\otimes n}(J_{K+\varepsilon} \cap J_{K-\varepsilon}^c) \leq C''\phi(K)$, where the constant C'' does not depend on K or ε . The third expectation in (2.5) can be made arbitrarily small by taking K sufficiently large and the first expectation can be made as small as we want choosing $\varepsilon > 0$ sufficiently small. Therefore, $\mathbb{E}_\mu[|\partial_i f_{K,\varepsilon}|] - \mathbb{E}_\mu[|\partial_i f|] \rightarrow 0$ as $K \rightarrow \infty, \varepsilon \rightarrow 0$.

Clearly, the statement of the theorem holds for each $f_{K,\varepsilon}$. Taking $K \rightarrow \infty, \varepsilon \rightarrow 0$, we obtain the desirable conclusion for the original function f . □

3. A direct Gaussian approach via the Ornstein–Uhlenbeck semigroup

In this section we prove a functional version of Theorem 1.7 (Theorem 3.1 below) and an inverse Gaussian BKS theorem using tools from the ‘‘Gaussian world’’ without appealing to the corresponding results for the discrete cube as we did in the previous section.

Theorem 3.1. *Let $f, g : \mathbb{R}^n \rightarrow [-1, 1]$ be increasing continuously differentiable functions. Then*

$$\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g] \geq c \sum_{i=1}^n \frac{\mathbb{E}_\mu[\partial_i f]\mathbb{E}_\mu[\partial_i g]}{\sqrt{\log(e/\mathbb{E}_\mu[\partial_i f])\log(e/\mathbb{E}_\mu[\partial_i g])}},$$

where $c > 0$ is a universal constant.

We start with a few standard definitions and simple lemmas related to the Ornstein–Uhlenbeck semigroup. For a more detailed treatment of these notions, the reader is referred to [9,24].

Definition 3.2. *Let $(P_t)_{t \geq 0}$ be the Ornstein–Uhlenbeck semigroup associated with the generator $L = \Delta - x \cdot \nabla$ on \mathbb{R}^n . This semigroup acts on the functions on \mathbb{R}^n as follows:*

$$P_t f(x) = \int f(e^{-t}x + \sqrt{1 - e^{-2t}}y) \mu^{\otimes n}(dy), \quad x \in \mathbb{R}^n.$$

It is well known that $(P_t)_{t \geq 0}$ is reversible with the invariant measure $\mu^{\otimes n}$. For $t > 0$, the operator P_t maps bounded measurable functions to C^∞ functions. It also maps an increasing function to an increasing function. The following simple properties of the operator P_t will be very useful for later purposes:

Observation 3.3. *Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a smooth function. Then:*

- i. $\partial_i P_t g = e^{-t} P_t \partial_i g \quad \forall t \geq 0.$ (3.1)

ii. If $|g(x)| \leq 1$ for all x , then

$$|\nabla P_t g| \leq \frac{1}{\sqrt{t}} \quad 0 < t \leq 1/2. \quad (3.2)$$

iii. If g is increasing, then

$$\partial_i P_t g \geq 0 \quad \forall t \geq 0. \quad (3.3)$$

Lemma 3.4. Let g be a smooth function with $|g(x)| \leq 1$ for all x , and let $t \in (0, 1/2]$.

(i) For $p \geq 1$, we have $\|\partial_i P_t g\|_p \leq t^{-(p-1)/2p} e^{-t/p} \|\partial_i g\|_1^{1/p}$.

(ii) Assume further that g is increasing. Then for $0 < p < 1$, we have $\|\partial_i P_t g\|_p \geq t^{(1-p)/2p} e^{-t/p} \|\partial_i g\|_1^{1/p}$.

Proof. (i) By (3.1) and (3.2), we have

$$\|\partial_i P_t g\|_p^p \leq t^{-(p-1)/2} \|\partial_i P_t g\|_1 = t^{-(p-1)/2} e^{-t} \|P_t(\partial_i g)\|_1 \leq t^{-(p-1)/2} e^{-t} \|\partial_i g\|_1,$$

where in the last inequality we use the fact that $P_t : L^1(\mu^{\otimes n}) \rightarrow L^1(\mu^{\otimes n})$ is a contraction.

(ii) Again using (3.1) and (3.2), we obtain

$$\|\partial_i P_t g\|_1 \leq t^{-(1-p)/2} \|\partial_i P_t g\|_p^p.$$

Note that since g is increasing,

$$\|\partial_i P_t g\|_1 = \mathbb{E}_\mu[\partial_i P_t g] = e^{-t} \mathbb{E}_\mu[P_t \partial_i g] = e^{-t} \mathbb{E}_\mu[\partial_i g] = e^{-t} \|\partial_i g\|_1.$$

Hence, we have $e^{-t} \|\partial_i g\|_1 \leq t^{-(1-p)/2} \|\partial_i P_t g\|_p^p$, as desired. \square

3.1. An alternative refined Gaussian FKG inequality

In order to prove Theorem 3.1, we need the following identity for the covariance of a pair of functions w.r.t. the Gaussian measure, which follows from [8], Lemma 3.3, using the polarization identity: $2 \operatorname{Cov}(f, g) = \operatorname{Var}(f + g) - \operatorname{Var}(f) - \operatorname{Var}(g)$.

Proposition 3.5. Let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ be two absolutely continuous functions and suppose that $\|\nabla f\|_2^2, \|\nabla g\|_2^2 \in L^2(\mu^{\otimes n})$. Then

$$\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g] = \sum_{i=1}^n \int_0^\infty e^{-t} \mathbb{E}_\mu[\partial_i f P_t \partial_i g] dt. \quad (3.4)$$

Note that if f, g are increasing, then the RHS is clearly non-negative, and hence, $\mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g] \geq 0$. This already implies the Gaussian FKG inequality [10]. Moreover, the proposition gives a precise expression for $\operatorname{Cov}(f, g)$. However, as the precise expression is not so convenient to work with, we replace it by a more convenient lower bound to obtain Theorem 3.1.

Proof of Theorem 3.1. First of all, note that since f are increasing and $|f| \leq 1$, we have $\int_{\mathbb{R}^n} \partial_i f(z) dz \leq 2$ for all i . Hence, $\mathbb{E}_\mu[\partial_i f] \leq 1$. The same conclusion also holds for g .

To prove the theorem, we will use Borell's reverse hypercontractive inequality [5] which implies the following result. (See Corollary 3.3 of [26] for a discrete version of the result. The Gaussian version presented here follows immediately by a CLT argument.) Let $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be smooth bounded functions, then for any $p, q \in (0, 1)$ such that $e^{-2t} \leq (1-p)(1-q)$, the following inequality holds:

$$\mathbb{E}_\mu[f_1 P_t f_2] \geq \|f_1\|_p \|f_2\|_q. \quad (3.5)$$

Here the norms are taken w.r.t. the Gaussian measure $\mu^{\otimes n}$. Fix $1 \leq i \leq n$. Using (3.1) and the fact that P_t is reversible w.r.t. $\mu^{\otimes n}$, we have

$$\begin{aligned} \int_0^\infty e^{-t} \mathbb{E}_\mu[\partial_i f P_t \partial_i g] dt &\geq \int_1^\infty e^{-(t-1)} \mathbb{E}_\mu[\partial_i P_{1/2} f P_{t-1} \partial_i P_{1/2} g] dt \\ &= \int_0^\infty e^{-t} \mathbb{E}_\mu[\partial_i P_{1/2} f P_t \partial_i P_{1/2} g] dt \\ &= \int_0^1 \mathbb{E}_\mu[\partial_i P_{1/2} f T_s \partial_i P_{1/2} g] dt \quad [T_s := P_{\log(1/s)}]. \end{aligned} \tag{3.6}$$

By (3.5) and Lemma 3.4, we deduce that

$$\mathbb{E}_\mu[\partial_i P_{1/2} f T_s \partial_i P_{1/2} g] \geq \|\partial_i P_{1/2} f\|_p \|\partial_i P_{1/2} g\|_q \geq (2e)^{-(1/(2p)+1/(2q))} \|\partial_i f\|_1^{1/p} \|\partial_i g\|_1^{1/q}, \tag{3.7}$$

for $s > 0$ such that $s^2 \leq (1-p)(1-q)$. Optimizing the RHS of (3.7) over $p, q \in (0, 1)$ satisfying $s^2 \leq (1-p)(1-q)$, we obtain

$$\mathbb{E}_\mu[\partial_i P_{1/2} f T_s \partial_i P_{1/2} g] \geq \exp\left(-\frac{1}{2} \frac{a_i^2 + 2sa_i b_i + b_i^2}{1-s^2}\right),$$

where $a_i, b_i > 0$ are such that $(2e)^{-1/2} \|\partial_i f\|_1 = e^{-a_i^2/2}$ and $(2e)^{-1/2} \|\partial_i g\|_1 = e^{-b_i^2/2}$. Hence, by (3.6),

$$\begin{aligned} \int_0^\infty e^{-t} \mathbb{E}_\mu[\partial_i f P_t \partial_i g] dt &\geq \int_0^1 \exp\left(-\frac{1}{2} \frac{a_i^2 + 2sa_i b_i + b_i^2}{1-s^2}\right) ds \\ &\geq \varepsilon \exp\left(-\frac{1}{2} \frac{a_i^2 + 2\varepsilon a_i b_i + b_i^2}{1-\varepsilon^2}\right), \end{aligned} \tag{3.8}$$

for any $\varepsilon \in (0, 1)$. We are interested in finding a lower bound of the RHS of (3.8) when a_i and b_i are large. Note that the derivative of the RHS of (3.8) as a function of ε vanishes approximately at $\varepsilon \approx 1/a_i b_i$. Plugging in $\varepsilon = 1/a_i b_i$ in (3.8), we obtain

$$\begin{aligned} \frac{1}{a_i b_i} \exp\left(-\frac{1}{2} \frac{a_i^2 + 2\varepsilon a_i b_i + b_i^2}{1-\varepsilon^2}\right) &\geq \frac{1}{a_i b_i} \exp\left(-\frac{1}{2} (a_i^2 + 2\varepsilon a_i b_i + b_i^2) (1 + \varepsilon^2 + O(\varepsilon^4))\right) \\ &\geq \frac{c_1}{a_i b_i} e^{-(a_i^2 + b_i^2)/2} \geq c_1 e \cdot \frac{\|\partial_i f\|_1 \|\partial_i g\|_1}{\sqrt{\log(e/\|\partial_i f\|_1) \log(e/\|\partial_i g\|_1)}}, \end{aligned} \tag{3.9}$$

where $c_1 > 0$ is a universal constant. In the second inequality above, we used the fact $\varepsilon^2(a_i^2 + b_i^2) = O(1)$. This is because a_i, b_i are bounded from below, which follows from the fact $\|\partial_i f\|_1, \|\partial_i g\|_1 \leq 1$. Now we conclude the proof by combining Proposition 3.5 and the bounds (3.8) and (3.9) and by taking $c = c_1 e$. \square

3.2. A direct approach towards inverse Gaussian BKS

In this subsection we aim to prove a Gaussian analogue of the inverse BKS theorem (see Proposition 1.3 of [3]).

Proposition 3.6. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable increasing function such that $\|\nabla f\|_2^2 \in L^2(\mu^{\otimes n})$. Then for any $\rho \in (0, 1)$,*

$$\text{VAR}^G(f, \rho) \geq (1 - \rho^2) \sum_{i=1}^n \mathbb{E}_\mu[\partial_i f]^2.$$

Proof. First we show that for if f is as given in the proposition, then

$$\text{Var}_\mu(f) \geq \sum_{i=1}^n \mathbb{E}_\mu[\partial_i f]^2. \tag{3.10}$$

Taking $f = g$ in Proposition 3.5, we have

$$\text{Var}_\mu(f) = \sum_{i=1}^n \int_0^\infty e^{-t} \mathbb{E}_\mu[\partial_i f P_t \partial_i f] dt. \tag{3.11}$$

We claim that $\mathbb{E}_\mu[\partial_i f P_t \partial_i f]$ is a nonincreasing function of t . Indeed,

$$\frac{d}{dt} \mathbb{E}_\mu[g P_t g] = \mathbb{E}_\mu \left[g P_t \sum_{j=1}^n (\partial_j^2 g - x_j \partial_j g) \right] = \sum_{j=1}^n \mathbb{E}_\mu [g_t \partial_j^2 g - x_j g_t \partial_j g],$$

where $g_t = P_t g$. Integration by parts yields

$$\mathbb{E}_\mu [g_t \partial_j^2 g - x_j g_t \partial_j g] = \mathbb{E}_\mu [g \partial_j^2 g - \partial_j (g_t \partial_j g)] = -\mathbb{E}_\mu [\partial_j g_t \partial_j g] = -e^{-t} \mathbb{E}_\mu [\partial_j g P_t \partial_j g] \leq 0,$$

and hence,

$$\frac{d}{dt} \mathbb{E}_\mu [g P_t g] = \sum_{j=1}^n \mathbb{E}_\mu [g_t \partial_j^2 g - x_j g_t \partial_j g] \leq 0.$$

Therefore,

$$\mathbb{E}_\mu [\partial_i f P_t \partial_i f] \geq \mathbb{E}_\mu [\partial_i f P_\infty \partial_i f] = \mathbb{E}_\mu [\partial_i f]^2. \tag{3.12}$$

Combination of (3.11) with (3.12) yields (3.10).

By (3.10) and (3.1),

$$\text{Var}_\mu(P_t f) \geq \sum_{i=1}^n \mathbb{E}_\mu [\partial_i P_t f]^2 = \sum_{i=1}^n e^{-2t} \mathbb{E}_\mu [P_t \partial_i f]^2 = e^{-2t} \sum_{i=1}^n \mathbb{E}_\mu [\partial_i f]^2.$$

Note that by the definition of the Ornstein–Uhlenbeck operator, we have

$$\text{VAR}^G(f, \rho) = \mathbb{E}_\mu [f P_t f] - \mathbb{E}_\mu [f]^2 = \text{Var}_\mu(P_t f). \tag{3.13}$$

This completes the proof. □

As a corollary (which we will prove in the next section), we obtain an inverse Gaussian BKS theorem for increasing functions.

Corollary 3.7. *Let $\{A_\ell \subseteq \mathbb{R}^{n_\ell}\}$ be a sequence of increasing sets. If $\{A_\ell\}$ is asymptotically Gaussian noise sensitive, then $\sum_{i=1}^{n_\ell} I_i^G(A_\ell)^2 \rightarrow 0$ as $\ell \rightarrow \infty$.*

4. Smooth approximation of characteristic functions of monotone sets

In this section, we prove a result that connects the partial derivative of the characteristic function of an increasing set after being smoothed by the action of Ornstein–Uhlenbeck operator P_t to its geometric influence as $t \downarrow 0$. This will help us in deriving various theorems presented in the introduction, which involve sets, from the respective theorems involving C^1 functions.

Recall that as defined in the introduction, for any set $A \subseteq \mathbb{R}^n$, for each $1 \leq i \leq n$ and an element $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, the restriction of A along the fiber of x in the i th direction is given by

$$A_i^x := \{y \in \mathbb{R} : (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \in A\}.$$

Definition 4.1. A set $A \subset \mathbb{R}^n$ is called increasing (decreasing) if its characteristic function $\mathbf{1}_A$ is an increasing (decreasing) function in each coordinate. For any increasing set $A \subset \mathbb{R}^n$ and for any $x \in \mathbb{R}^n$, define

$$t_i(A; x^{(-i)}) := \inf\{y : y \in A_i^x\} \in [-\infty, \infty],$$

where $x^{(-i)} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathbb{R}^{n-1}$ and we use the convention that the infimum of the empty set is $+\infty$.

Note that $t_i(A; \cdot)$ is a decreasing function of $x^{(-i)}$ for any increasing set A . Also, for an increasing set A , its geometric influence is given by $I_i^{\mathcal{G}}(A) = \int_{\mathbb{R}^{n-1}} \phi(t_i(A; z^{(-i)})) \mu^{\otimes n-1}(dz^{(-i)})$.

Lemma 4.2. Let A be a monotone subset of \mathbb{R}^n . Then, for each $i \in \{1, \dots, n\}$, we have

$$\mathbb{E}_\mu[\partial_i P_t \mathbf{1}_A] \rightarrow I_i^{\mathcal{G}}(A) \quad \text{as } t \downarrow 0.$$

Remark 4.3. Lemma 4.2 does not hold in general without the monotonicity assumption. For example, take $n = 1$ and define $A = \mathbb{Q}$, the set of rational numbers. Then $P_t \mathbf{1}_A = 0$ for any $t > 0$ and hence $\lim_{t \rightarrow 0+} \mathbb{E}_\mu[\partial_1 P_t \mathbf{1}_A] = 0$ but it can be easily checked that $I_1^{\mathcal{G}}(A) = \infty$.

In order to prove Lemma 4.2 we need the following standard lemma. For sake of completeness, we present its proof.

Lemma 4.4. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a monotone function. Then the set of discontinuities of f has Lebesgue measure zero.

Proof. The n -dimensional space \mathbb{R}^n can be represented as a disjoint union of straight lines $\bigcup_{\{z \in \mathbb{R}^n : z_n = 0\}} l_z$, where each line is defined as $l_z = z + t(1, 1, \dots, 1)$, $t \in \mathbb{R}$. We would like to show that the set of discontinuities of f on each line l_z is of Lebesgue measure zero, and then, the assertion of the lemma would follow by a standard application of Fubini’s theorem.

For each such line l_z , the restriction of f to l_z can be represented by a one-dimensional function $f_z : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f_z(a) = f(z + (a, a, \dots, a))$. Note that if f is not continuous at some $x \in l_z$, then

$$\lim_{\varepsilon \rightarrow 0+} \left(\sup_{y \in x + [-\varepsilon, \varepsilon]^n} f(y) - \inf_{y \in x + [-\varepsilon, \varepsilon]^n} f(y) \right) > 0,$$

which implies, by monotonicity of f , that

$$\lim_{\varepsilon \rightarrow 0+} f(x + (\varepsilon, \dots, \varepsilon)) - f(x - (\varepsilon, \dots, \varepsilon)) \neq 0.$$

Hence, each discontinuity x of f corresponds to a discontinuity of the one-dimensional function f_z . Therefore, the set of discontinuities of f on a line l_z can be embedded into the set of discontinuities of the function f_z . However, for a fixed z , f_z is a monotone function on the real line, and thus, the set of its discontinuities is countable, and, in particular, of Lebesgue measure zero. Thus, the set of discontinuities of f on each line l_z is of Lebesgue measure zero, which completes the proof. \square

Proof of Lemma 4.2. Without loss of generality, we assume that A is an increasing set. Let $W = (W_1, \dots, W_n)$ be a standard Gaussian vector on \mathbb{R}^n and define $Y_t = e^{-t}x + \sqrt{1 - e^{-2t}}W$. Then we can write

$$P_t \mathbf{1}_A(x) = \mathbb{E}_\mu[\mathbf{1}_A(Y_t)] = \mathbb{E}_\mu \left[\bar{\Phi} \left(\frac{t_i(A; Y_t^{(-i)}) - e^{-t}x_i}{\sqrt{1 - e^{-2t}}} \right) \right],$$

which, taking partial derivative w.r.t. x_i , yields

$$\partial_i P_t \mathbf{1}_A(x) = \mathbb{E} \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \phi \left(\frac{t_i(A; Y_t^{(-i)}) - e^{-t} x_i}{\sqrt{1-e^{-2t}}} \right) \geq 0.$$

Therefore,

$$\begin{aligned} \mathbb{E}_\mu[\partial_i P_t \mathbf{1}_A] &= \int_{\mathbb{R}^n} \mathbb{E}_W \left[\frac{e^{-t}}{\sqrt{1-e^{-2t}}} \phi \left(\frac{t_i(A; Y_t^{(-i)}) - e^{-t} x_i}{\sqrt{1-e^{-2t}}} \right) \right] \prod_{j=1}^n \phi(x_j) dx \\ &= \mathbb{E}_W \int_{\mathbb{R}^n} \phi(u) \phi \left(\frac{\sqrt{1-e^{-2t}}u + t_i(A; Y_t^{(-i)})}{e^{-t}} \right) du \cdot \prod_{j \neq i} \phi(x_j) dx^{(-i)}, \end{aligned} \tag{4.1}$$

where in the last step we make a change of variable $u = \frac{e^{-t}x_i - t_i(A; Y_t^{(-i)})}{\sqrt{1-e^{-2t}}}$. Note that by Lemma 4.4, we have $t_i(A; Y_t^{(-i)}) \rightarrow t_i(A; x^{(-i)})$ in distribution as $t \rightarrow 0+$. Hence, taking limit as $t \rightarrow 0+$ in (4.1), we obtain, by the Bounded Convergence Theorem,

$$\mathbb{E}_\mu[\partial_i P_t \mathbf{1}_A] = \int_{\mathbb{R}^n} \phi(t_i(A; x^{(-i)})) du \cdot \prod_{j \neq i} \phi(x_j) dx^{(-i)} = I_i^{\mathcal{G}}(A).$$

This completes the proof of the lemma. □

As a consequence of Lemma 4.2, Theorems 1.6 and 1.7 can now be easily derived from their functional counterparts.

Proof of Theorems 1.6 and 1.7. For $t > 0$, define $f_t = P_t \mathbf{1}_A$ and $g_t = P_t \mathbf{1}_B$ for increasing sets A, B of \mathbb{R}^n . Note that f_t and g_t are increasing C^∞ functions which are bounded by 1. Thus we can apply Theorems 3.1 and 2.1 with $f = f_t$ and $g = g_t$ and then let $t \rightarrow 0+$. In view of Lemma 4.2, the right hand sides of the inequalities converge to appropriate quantities involving the geometric influences of the sets A and B . Again by Lemma 4.2, $\mathbf{1}_A$ and $\mathbf{1}_B$ are almost surely continuous, hence $f_t \rightarrow \mathbf{1}_A$ and $g_t \rightarrow \mathbf{1}_B$ in probability. Therefore, $\mathbb{E}_\mu[f_t g_t] - \mathbb{E}_\mu[f_t] \mathbb{E}_\mu[g_t] \rightarrow \mu^{\otimes n}(A \cap B) - \mu^{\otimes n}(A) \mu^{\otimes n}(B)$ by dominated convergence, which completes the proofs of the theorems. □

Note that above proof technique can not be immediately applied to deduce Theorem 1.9 from Theorem 2.2 since Lemma 4.2 does not hold for general non-monotone sets. We overcome this obstacle by establishing a shifting lemma, which implies that it will be sufficient to prove our theorem for increasing sets. This shifting lemma is a Gaussian analogue of Lemma 2.7 in [3].

Definition 4.5. For $i \in \{1, 2, \dots, n\}$, the i -shift operator M_i acting on subsets of \mathbb{R}^n is defined by:

$$M_i(A) = \{x \in \mathbb{R}^n: x_i \geq \bar{\Phi}^{-1}(\mu(A_i^x))\}.$$

The shifting operator M is defined as $M = M_1 \circ M_2 \circ \dots \circ M_n$.

Lemma 4.6. Let $A \subseteq \mathbb{R}^n$. For any $i \in \{1, 2, \dots, n\}$ and for any $\rho \in (0, 1)$, we have:

- (i) $M(A)$ is increasing.
- (ii) $\mu^{\otimes n}(M(A)) = \mu^{\otimes n}(A)$.
- (iii) $I_i^{\mathcal{G}}(M(A)) \leq I_i^{\mathcal{G}}(A)$.
- (iv) $Z^{\mathcal{G}}(M(A), \rho) \geq Z^{\mathcal{G}}(A, \rho)$.

Proof. The proofs of (i) and (ii) are standard (see [11]).

In order to prove (iii), we recall the notion of h -influences defined in [18] and its relation to geometric influences. For a function $h : [0, 1] \rightarrow [0, 1]$, the h -influence of the i th coordinate on A (in the Gaussian space) is defined as

$$I_i^h(A) := \int h(\mu(A_i^x)) \mu^{\otimes n}(dx),$$

where A_i^x is the restriction of A along the fiber of x in the i th direction. It was shown in previous work that:

- If the function h is concave and continuous, then h -influences of any set can only decrease under the action of the shifting operator M on that set (see Theorem 2.2 of [18]).
- For $h(t) = \phi(\Phi^{-1}(t))$ (which is concave and continuous), we have $I_i^{\mathcal{G}}(A) \geq I_i^h(A)$ for any set A , and $I_i^{\mathcal{G}}(A) = I_i^h(A)$ for monotone increasing sets (see Lemmas 3.5 and 3.7 of [21]).

Combining these two facts, we have

$$I_i^{\mathcal{G}}(M(A)) = I_i^h(M(A)) \leq I_i^h(A) \leq I_i^{\mathcal{G}}(A),$$

as asserted in (iii).

To prove (iv), it is sufficient to show that $Z^{\mathcal{G}}(M_j(A), \rho) \geq Z^{\mathcal{G}}(A, \rho)$ for each $j \in \{1, 2, \dots, n\}$. Let W, W' be two i.i.d. standard Gaussian vectors on \mathbb{R}^n and set $W^\rho = \sqrt{1 - \rho^2}W + \rho W'$ (as defined above). We have

$$\begin{aligned} Z^{\mathcal{G}}(A, \rho) &= \mathbb{E}[\mathbf{1}_{A_j^W}(W_j)\mathbf{1}_{A_j^{W^\rho}}(W_j^\rho)] \\ &= \mathbb{E}_{W^{(-j)}, W'^{(-j)}}[\mathbb{E}[\mathbf{1}_{A_j^W}(W_j)\mathbf{1}_{A_j^{W^\rho}}(W_j^\rho) | W^{(-j)}, W'^{(-j)}]]. \end{aligned} \tag{4.2}$$

By Borell’s isoperimetric inequality [6], amongst all pairs of subsets S, T of the real line such that $\mu(S) = a$ and $\mu(T) = b$, the joint probability $\mathbb{P}[W_j \in S, W_j^\rho \in T]$ is maximized when $S = [\bar{\Phi}^{-1}(a), \infty)$ and $T = [\bar{\Phi}^{-1}(b), \infty)$. This implies that

$$\mathbb{E}[\mathbf{1}_{A_j^x}(W_j)\mathbf{1}_{A_j^y}(W_j^\rho)] \leq \mathbb{E}[\mathbf{1}_{M_j(A)_j^x}(W_j)\mathbf{1}_{M_j(A)_j^y}(W_j^\rho)] \quad \forall x, y \in \mathbb{R}^n. \tag{4.3}$$

Assertion (iv) follows immediately by plugging Eq. (4.3) into Eq. (4.2). □

Proof of Theorem 1.9. By Lemma 4.6, it is sufficient to prove the theorem for increasing sets. Now we can follow the proof of Theorems 1.6 and 1.7 to complete the proof. We omit the details. □

We point out that a Gaussian analogue of the original BKS theorem follows immediately from Theorem 1.9.

Corollary 4.7. *Let $A_\ell \subseteq \mathbb{R}^{n_\ell}$ be a sequence of sets and suppose that $\sum_{i=1}^{n_\ell} I_i^{\mathcal{G}}(B)^2 \rightarrow 0$ as $\ell \rightarrow \infty$. Then $\{A_\ell\}$ is asymptotically Gaussian noise-sensitive.*

Proof of Corollary 3.7. Again, we can follow the proof of Theorems 1.6 and 1.7 to show that, for any increasing set $A \subseteq \mathbb{R}^n$,

$$\text{VAR}^{\mathcal{G}}(A, \rho) \geq (1 - \rho^2) \sum_{i=1}^n I_i^{\mathcal{G}}(A)^2.$$

The assertion of the corollary follows immediately. □

4.1. Comparison between Theorems 1.6 and 1.7

Let us compare the performances of Theorems 1.6 and 1.7 in two important special cases.

- *Threshold sets in \mathbb{R}^n .* Let $A = \{x \in \mathbb{R}^n: n^{-1/2} \sum_{i=1}^n x_i > -t\}$ and $B = \{x \in \mathbb{R}^n: n^{-1/2} \sum_{i=1}^n x_i > t\}$. In this case, $\mu(A) = \varepsilon$ and $\mu(B) = 1 - \varepsilon$ where $\varepsilon = \Phi^{-1}(-t)$, and hence, $\mu^{\otimes n}(A \cap B) - \mu^{\otimes n}(A)\mu^{\otimes n}(B) = \varepsilon^2$. It is easy to show that $I_i^{\mathcal{G}}(A) = I_i^{\mathcal{G}}(B) \asymp n^{-1/2} \varepsilon \sqrt{\log(1/\varepsilon)}$ for each i . Thus, Theorem 1.6 gives a lower bound of order ε^2 whereas Theorem 1.7 yields a lower bound of order $\varepsilon^2/\log n$. Therefore, in this example, Theorem 1.6 is tight as $t \rightarrow \infty$ ($\varepsilon \rightarrow 0$) up to a constant factor for any n , while Theorem 1.7 is off by a factor of $\log n$.
- *Sets that depend on a single coordinate.* Let $n = 1$ (which is equivalent to the case when both sets depend on a single coordinate). In this case, Theorem 1.7 is strictly stronger than Theorem 1.6. Indeed, for $n = 1$, the bounds given by the theorems are (up to a constant):

$$\frac{I^{\mathcal{G}}(A)I^{\mathcal{G}}(B)}{\sqrt{\log(e/I^{\mathcal{G}}(A))\log(e/I^{\mathcal{G}}(B))}} \quad \text{and} \quad \frac{I^{\mathcal{G}}(A)I^{\mathcal{G}}(B)}{\log(e/I^{\mathcal{G}}(A)I^{\mathcal{G}}(B))}.$$

Since

$$\log(e/I^{\mathcal{G}}(A)I^{\mathcal{G}}(B)) \geq \frac{1}{2}(\log(e/I^{\mathcal{G}}(A)) + \log(e/I^{\mathcal{G}}(B))),$$

the left bound is always greater than the right one by the inequality between the arithmetic and geometric means. Moreover, it can be shown that the bound of Theorem 1.7 is asymptotically tight for any choice of the sets A, B , while Theorem 1.6 is not tight for $A = (-t, \infty), B = (e^t, \infty)$ as $t \rightarrow \infty$.

5. Other probability spaces

In this section, we show how one can use the Gaussian Talagrand bounds obtained in the previous sections to prove analogous bounds for other product spaces, including all discrete product spaces, the space $[0, 1]^n$ endowed with the Lebesgue measure, etc. Next we will deduce a BKS theorem for the product biased measure on the discrete cube $\{-1, 1\}^n$ from its Gaussian counterpart. We should mention here that it is not clear if it is possible to find a reduction from the Gaussian BKS theorem to an analogous BKS theorem for a general discrete product space $([q]^n, \gamma^{\otimes n})$. Indeed, while the Ornstein–Uhlenbeck semigroup action is same as adding “small” amount of noise to every coordinate, the standard noise operator (on a discrete product space) amounts to adding “big” noise to a small number of coordinates. That they are equivalent is far from obvious.

Since there is no single natural definition of influences for such spaces, we formulate the results in terms of the h -influences defined in [18] (which turns out to be the most natural way to state them), and then mention the formulation with respect to more common definitions of influences. First we recall the definition of h -influences.

Definition 5.1. Let Ω be a probability space endowed with a probability measure γ . For a function $h: [0, 1] \rightarrow [0, 1]$, the h -influence of the i th coordinate on a set A in the product space $(\Omega^n, \gamma^{\otimes n})$ is defined as

$$I_i^h(A) := \mathbb{E}_{\gamma} [h(\gamma(A_i^x))],$$

where A_i^x is the restriction of A along the fiber of x in the i th direction and \mathbb{E}_{γ} , as always, denotes the expectation w.r.t. the product measure $\gamma^{\otimes n}$.

Throughout this section, we consider h -influences with respect to the function $h(t) = \phi(\Phi^{-1}(t))$. For sake of simplicity, we formulate the results for discrete probability spaces. The results for other spaces, such as the space $[0, 1]^n$ endowed with the Lebesgue measure, can be derived similarly.

For $q > 1$, let $[q] = \{1, 2, \dots, q\}$, and let γ be a probability measure on $[q]$. Without loss of generality, we assume that $\gamma(i) > 0$ for all $i \in [q]$ and denote the smallest atom in $([q], \gamma)$ by $\alpha = \min_{i \in [q]} \gamma(i)$. In order to obtain the reduction from $([q]^n, \gamma^{\otimes n})$ to $(\mathbb{R}^n, \mu^{\otimes n})$, we define $\psi: \mathbb{R} \rightarrow [q]$ to be an increasing function such that the push forward $\mu \circ \psi^{-1}$ has law γ . For example, $\psi(u) = \min\{i \in [q]: F(i) > \Phi(u)\}$, where F is the distribution function of γ . Define $\psi^{\otimes n}: \mathbb{R}^n \rightarrow [q]^n$ by $\psi^{\otimes n}(u_1, \dots, u_n) = (\psi(u_1), \dots, \psi(u_n))$, and set $A_G := (\psi^{\otimes n})^{-1}(A)$.

Obviously, $\mu^{\otimes n}(A_G) = \gamma^{\otimes n}(A)$ for any $A \subseteq [q]^n$. Moreover, a similar equality holds with respect to the restriction along fibers: If $u \in \mathbb{R}^n$ such that $\psi^{\otimes n}(u) = J \in [q]^n$, then the fibers $(A_G)_i^u$ and A_i^J satisfy:

$$(A_G)_i^u = \psi^{-1}(A_i^J).$$

Consequently, $\mu((A_G)_i^u) = \mu(\psi^{-1}(A_i^J)) = \gamma(A_i^J)$. This allows us to relate the geometric influences of A_G to the h -influences of A . Indeed, it was shown in [21] that for $h(t) = \phi(\Phi^{-1}(t))$, we have $I_i^{\mathcal{G}}(B) \geq I_i^h(B)$ for any set $B \subseteq \mathbb{R}^n$, and $I_i^{\mathcal{G}}(B) = I_i^h(B)$ for monotone increasing sets (see Lemmas 3.5 and 3.7 of [21]). Hence, for any $A \subseteq [q]^n$ and for any $1 \leq i \leq n$,

$$I_i^{\mathcal{G}}(A_G) \geq I_i^h(A). \tag{5.1}$$

This allows us to obtain analogues of Gaussian correlation bounds for the product space $([q]^n, \gamma^{\otimes n})$.

Theorem 5.2. *Let A, B be two increasing subsets of $[q]^n$. Then,*

$$\begin{aligned} & \gamma^{\otimes n}(A \cap B) - \gamma^{\otimes n}(A)\gamma^{\otimes n}(B) \\ & \geq c \max \left(\sum_{i=1}^n \frac{I_i^h(A)I_i^h(B)}{\sqrt{\log(1/I_i^h(A))\log(1/I_i^h(B))}}, \varphi \left(\sum_{i=1}^n I_i^h(A)I_i^h(B) \right) \right), \end{aligned}$$

where $c > 0$ is a universal constant, $h(t) = \phi(\Phi^{-1}(t))$, and $\varphi(x) = x/\log(e/x)$.

Proof. Since the functions $x \mapsto x/\sqrt{\log(1/x)}$ and $x \mapsto x/\log(e/x)$ are increasing in $(0, 1)$, the assertion follows by applying Theorem 1.6 and Theorem 1.7 to the increasing sets $A_G, B_G \subseteq \mathbb{R}^n$ coupled with the observation (5.1). \square

An interesting special case is the discrete cube $\{-1, 1\}^n$ endowed with the product biased measure $\nu_\alpha^{\otimes n}$, where $\nu_\alpha = \alpha\delta_1 + (1 - \alpha)\delta_{-1}$ (w.l.o.g. for $0 < \alpha < 1/2$). In this case, the h -influence with $h(t) = \phi(\Phi^{-1}(t))$ satisfies

$$I_i^h(A) = \alpha\sqrt{\log(1/\alpha)}I_i(A),$$

where $I_i(A)$ is defined similarly to (1.1) (but instead of taking the expectation w.r.t. the uniform measure $\nu^{\otimes n}$, we use the product biased measure $\nu_\alpha^{\otimes n}$). Hence, Theorem 5.2 gives the bound

$$\nu_\alpha^{\otimes n}(A \cap B) - \nu_\alpha^{\otimes n}(A)\nu_\alpha^{\otimes n}(B) \geq c\varphi \left(\alpha^2 \log(1/\alpha) \sum_{i=1}^n I_i(A)I_i(B) \right),$$

which was already shown in [19], Proposition 3.12. We note that unlike the result of [19], in Theorem 5.2 the h -influences in the RHS appear without a ‘‘scaling factor’’ depending on α . This shows that in some sense, this h -influence, which is the discrete variant of the geometric influence, is more natural than the definition of influence used in [19] for the biased measure.

In order to obtain an analogue of Theorem 1.9 for the biased cube $(\{-1, 1\}^n, \nu_\alpha^{\otimes n})$, we need to find the exact relation between Gaussian noise sensitivity and discrete noise sensitivity (as defined in the introduction but now both X and X^η are distributed (marginally) as $\nu_\alpha^{\otimes n}$).

Lemma 5.3. *Consider the probability space $(\{-1, 1\}^n, \nu_\alpha^{\otimes n})$. Let A be a subset of $\{-1, +1\}^n$ and let A_G be as defined above. Then for any $\rho \in (0, 1)$*

$$\text{VAR}^{\mathcal{G}}(A_G, \rho) = \text{VAR}(A, \eta),$$

for $\eta = \eta(\rho, \alpha) = \frac{\mathbb{P}[W_1 < \Phi^{-1}(\alpha), W_1^\rho > \Phi^{-1}(\alpha)]}{\alpha(1-\alpha)}$, where (W_1, W_1^ρ) is a bivariate normal random vector with mean zero, unit variance and correlation $\sqrt{1 - \rho^2}$.

Proof. Let X and X^η be two $(1 - \eta)$ -correlated vectors on $(\{-1, +1\}^n, \nu_\alpha^{\otimes n})$ and let (W, W^ρ) be Gaussian vectors on \mathbb{R}^n as defined in Definition 1.8. Clearly, $\mu^{\otimes n}(A_G) = \nu_\alpha^{\otimes n}(A)$. To equate $Z(A, \eta)$ to $Z^G(A_G, \rho)$, we want to choose $\rho > 0$ such that the random vectors (X, X^η) and $(\psi^{\otimes n}(W), \psi^{\otimes n}(W^\rho))$ have the same distributions on $\{-1, 1\}^n \times \{-1, 1\}^n$. Note that this is equivalent to the condition

$$\mathbb{P}[X_1 = -1, X_1^\eta = -1] = \mathbb{P}[W_1 < \Phi^{-1}(\alpha), W_1^\rho < \Phi^{-1}(\alpha)],$$

which is same as

$$\alpha - \alpha(1 - \alpha)\eta = \mathbb{P}[W_1 < \Phi^{-1}(\alpha), W_1^\rho < \Phi^{-1}(\alpha)].$$

The lemma now follows immediately. \square

Theorem 5.4. Consider the product space $(\{-1, 1\}^n, \nu_\alpha^{\otimes n})$. For any n , for any set $A \subset \{-1, 1\}^n$, and for any $\eta \in (0, 1)$,

$$\text{VAR}(A, \eta) \leq C_1 \cdot \left(\sum_{i=1}^n I_i^h(A)^2 \right)^{C_2 \rho^2},$$

where $h(t) = \phi(\Phi^{-1}(t))$, ρ is as defined in Lemma 5.3, and $C_1, C_2 > 0$ are universal constants.

Proof. Consider the set A_G defined as above, and the corresponding ‘‘monotonized’’ set $M(A_G)$ (see Lemma 4.6 above). By Lemma 4.6(iv) and Lemma 5.3,

$$\text{VAR}^G(M(A_G), \rho) \geq \text{VAR}^G(A_G, \rho) = \text{VAR}(A, \eta). \quad (5.2)$$

On the other hand, by properties of the monotonization operator M , we have:

$$I_i^G(M(A_G)) = I_i^h(M(A_G)) \leq I_i^h(A_G) = I_i^h(A) \quad (5.3)$$

(see the proof of Lemma 4.6(iii) above). Applying Corollary 1.9 to the set $M(A_G)$, we get:

$$\text{VAR}^G(M(A_G), \rho) \leq C_1 \cdot \left(\sum_{i=1}^n I_i^G(M(A_G))^2 \right)^{C_2 \rho^2}. \quad (5.4)$$

Combination of (5.4) with (5.2) and (5.3) yields the assertion. \square

Let’s compare the above bound to the following bound obtained in [20], Theorem 7, in the regime when $\eta > 0$ is small but fixed and $\alpha \rightarrow 0$:

$$\text{VAR}(A, \eta) \leq c'_1 \cdot \left(\alpha(1 - \alpha) \sum_{i=1}^n I_i(A)^2 \right)^{\beta(\eta, \alpha) \cdot \eta},$$

where $\beta(\eta, \alpha) \cdot \eta \asymp_\eta 1/\log(1/\alpha)$. Note that after switching back to ordinary influences, Theorem 5.4 reads:

$$\text{VAR}(A, \eta) \leq C_1 \cdot \left(\alpha^2 \log(1/\alpha) \sum_{i=1}^n I_i(A)^2 \right)^{C_2 \cdot \rho(\eta, \alpha)^2}. \quad (5.5)$$

We are interested in finding a reasonable lower bound (up to a constant that may depend on η) on $\rho(\eta, \alpha)$. Set $t = \bar{\Phi}^{-1}(\alpha) \asymp \sqrt{\log(1/\alpha)}$. Note that,

$$\begin{aligned} \mathbb{P}[W_1 > t, W_1^\rho < t] &\leq \mathbb{P}[W_1 > t] \mathbb{P}[t\sqrt{1 - \rho^2} + \rho W_1' < t] \\ &= \alpha \mathbb{P}\left[W_1' < \frac{t(1 - \sqrt{1 - \rho^2})}{\rho} \right] \leq \alpha \mathbb{P}[W_1' < t\rho]. \end{aligned}$$

Since both $\mathbb{P}[W_1 > t, W_1^\rho < t]$ and $\alpha\mathbb{P}[W'_1 < t\rho]$ are increasing functions of ρ , a lower bound on $\rho(\eta, \alpha)$ can be achieved by solving $\eta\alpha(1 - \alpha) = \alpha\mathbb{P}[W'_1 < t\rho]$, which yields $t\rho \asymp_\eta 1$, or, $\rho^2 \asymp_\eta 1/\log(1/\alpha)$. So, the asymptotic performance of Theorem 5.4 matches with that of [20] (which was shown in [20] to be essentially tight).

We now relate our results to more common definitions of influences in the product spaces $([q]^n, \gamma^{\otimes n})$.

Variance influence. This notion, used e.g. in [14,25], is defined as:

$$I_i^{\text{Var}}(A) := \mathbb{E}_\gamma[\text{Var}(\mathbf{1}_{A_i^x})].$$

It is clear that the variance influence coincides with the h -influence for $h(t) = t(1 - t)$, and hence, it is always smaller (up to a constant factor) than the h -influence with $h(t) = \phi(\Phi^{-1}(t))$. Hence, Theorem 5.2 holds without change for the variance influences. In order to find a lower bound of variance influence in terms of h -influence, we consider the contribution of a single fiber to $I_i^h(A)$ and to $I_i^{\text{Var}}(A)$. If $\mu(A_i^x) = t \leq 1/2$, then these contributions are $t\sqrt{\log(1/t)}$ and $t(1 - t)$, respectively. Note that if α is the size of the smallest atom in $([q], \gamma)$, then either $t \in \{0, 1\}$ or $t \in [\alpha, 1 - \alpha]$. In both cases,

$$\frac{t\sqrt{\log(1/t)}}{t(1 - t)} \leq 2\sqrt{\log(1/\alpha)}.$$

Hence, for any A and i ,

$$I_i^h(A) \leq 2\sqrt{\log(1/\alpha)}I_i^{\text{Var}}(A),$$

and thus, Theorem 5.4 holds (for $q = 2$) with $4\log(1/\alpha)\sum_i I_i^{\text{Var}}(A)^2$ in place of $\sum_i I_i^h(A)^2$.

BKKKL influence. This influence, used in [7,14] is given by:

$$I_i^{\text{BKKKL}}(A) := \mathbb{E}_\gamma[h(\mu(A_i^x))],$$

where $h(t) = 1$ if $t \in (0, 1)$, and $h(t) = 0$ if $t \in \{0, 1\}$. This definition coincides with $I_i(A)$ for $q = 2$ and we have already seen in (5.5) how Theorem 5.4 should look in this case. As for Theorem 5.2, since the contribution of each fiber to $I_i^h(A)$ is either zero or at least $\alpha\sqrt{\log(1/\alpha)}$, it follows that the theorem holds with $\alpha^2\log(1/\alpha)I_i^{\text{BKKKL}}(A)I_i^{\text{BKKKL}}(B)$ instead of $I_i^h(A)I_i^h(B)$.

6. Open problems

We conclude the paper with a few directions for further research suggested by our results and by recent related work.

- (1) The first issue left open in this paper is to prove a quantitative BKS theorem for all other discrete product spaces. In fact, we weren't able to deduce it by a reduction from the Gaussian version even for the simplest case $([q]^n, \lambda^{\otimes n})$ where $q > 2$ and λ is the uniform measure on $[q]$. We note that we have a direct proof of quantitative BKS for all discrete spaces, using a generalization of the techniques used in [20], along with hypercontractive estimates for general discrete measures obtained by Wolff [30]. However, the proof is cumbersome and the result is not tight, and hence, a reduction from the Gaussian case is more desirable.
- (2) It would be interesting to find alternative "direct" proofs of Theorems 1.6 and 1.9, which do not rely on their counterparts on the discrete cube. In particular, we wonder whether one can combine the reverse hypercontractivity technique used in the proof of Theorem 1.7 with the classical hypercontractivity used in the proof of Theorem 1.6 to obtain a new lower bound that will enjoy the benefits of both theorems.
- (3) Probably the most interesting direction is to find applications of the results. Both Talagrand's lower bound and the BKS theorem have various applications, and even the recent generalization of the BKS theorem to biased measures [20] was already applied to percolation theory [1]. On the other hand, Gaussian noise sensitivity was recently studied by Kindler and o'Donnell [22] and used to obtain applications to isoperimetric inequalities and to hardness of approximation. Hence, it will be interesting to find also applications of Talagrand's lower bound or of the BKS theorem in the Gaussian setting.

- (4) Finally, our understanding of influences in product spaces is still very far from complete. In particular, only a very few is known about influences with respect to non-product measures, and it is even unclear what should be the natural definition of influences in such a general setting (see [12]).

References

- [1] D. Ahlberg, E. I. Broman, S. Griffith and R. Morris. Noise sensitivity in continuum percolation. *Israel J. Math.* To appear, 2014. Available at <http://arxiv.org/abs/1108.0310>.
- [2] W. Beckner. Inequalities in Fourier analysis. *Ann. Math. (2)* **102** (1975) 159–182. [MR0385456](#)
- [3] I. Benjamini, G. Kalai and O. Schramm. Noise sensitivity of boolean functions and applications to percolation. *Publ. Math. Inst. Hautes Études Sci.* **90** (1999) 5–43. [MR1813223](#)
- [4] A. Bonami. Etude des coefficients Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier* **20** (1970) 335–402. [MR0283496](#)
- [5] C. Borell. Positivity improving operators and hypercontractivity. *Math. Z.* **180** (2) (1982) 225–234. [MR0661699](#)
- [6] C. Borell. Geometric bounds on the Ornstein–Uhlenbeck velocity process. *Probab. Theory Related Fields* **70** (1) (1985) 1–13. [MR0795785](#)
- [7] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson and N. Linial. The influence of variables in product spaces. *Israel J. Math.* **77** (1992) 55–64. [MR1194785](#)
- [8] S. Chatterjee. Chaos, concentration, and multiple valleys, 2008. Available at <http://arxiv.org/abs/0810.4221>.
- [9] D. Cordero-Erausquin and M. Ledoux. Hypercontractive measures, Talagrand’s inequality, and influences. Preprint, 2011. Available at <http://arxiv.org/abs/1105.4533>.
- [10] C. M. Fortuin, P. W. Kasteleyn and J. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** (1971) 89–103. [MR0309498](#)
- [11] P. Frankl. The shifting technique in extremal set theory. In *Surveys in Combinatorics* 81–110. C. W. Whitehead (Ed). Cambridge Univ. Press, Cambridge, 1987. [MR0905277](#)
- [12] G. R. Grimmett and B. Graham. Influence and sharp-threshold theorems for monotonic measures. *Ann. Probab.* **34** (2006) 1726–1745. [MR2271479](#)
- [13] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Math. Proc. Cambridge Philos. Soc.* **56** (1960) 13–20. [MR0115221](#)
- [14] H. Hatami. Decision trees and influence of variables over product probability spaces. *Combin. Probab. Comput.* **18** (2009) 357–369. [MR2501432](#)
- [15] J. Kahn, G. Kalai and N. Linial. The influence of variables on boolean functions. In *Proc. 29th Ann. Symp. on Foundations of Comp. Sci.* 68–80. Computer Society Press, 1988.
- [16] G. Kalai and M. Safra. Threshold phenomena and influence. In *Computational Complexity and Statistical Physics* 25–60. A. G. Percus, G. Istrate and C. Moore (Eds). Oxford Univ. Press, New York, 2006.
- [17] N. Keller. Influences of variables on boolean functions. Ph.D. thesis, Hebrew Univ. Jerusalem, 2009.
- [18] N. Keller. On the influences of variables on boolean functions in product spaces. *Combin. Probab. Comput.* **20** (1) (2011) 83–102. [MR2745679](#)
- [19] N. Keller. A simple reduction from the biased measure on the discrete cube to the uniform measure. *European J. Combin.* **33** 1943–1957. Available at <http://arxiv.org/abs/1001.1167>. [MR2950492](#)
- [20] N. Keller and G. Kindler. A quantitative relation between influences and noise sensitivity. *Combinatorica* **33** 45–71. Available at <http://arxiv.org/abs/1003.1839>. [MR3070086](#)
- [21] N. Keller, E. Mossel and A. Sen. Geometric influences. *Ann. Probab.* **40** (3) (2012) 1135–1166. [MR2962089](#)
- [22] G. Kindler and R. O’Donnell. Gaussian noise sensitivity and Fourier tails. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity* 137–147. IEEE Computer Society, Washington, DC, 2012. Available at <http://www.cs.cmu.edu/~odonnell/papers/gaussian-noise-sensitivity.pdf>.
- [23] D. J. Kleitman. Families of non-disjoint subsets. *J. Combin. Theory* **1** (1966) 153–155. [MR0193020](#)
- [24] M. Ledoux. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math. (6)* **9** (2) (2000) 305–366. [MR1813804](#)
- [25] E. Mossel, R. O’Donnell and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math. (2)* **171** (1) (2010) 295–341. [MR2630040](#)
- [26] E. Mossel, R. O’Donnell, O. Regev, J. E. Steif and B. Sudakov. Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami–Beckner inequality. *Israel J. Math.* **154** (2006) 299–336. [MR2254545](#)
- [27] R. O’Donnell. Some topics in analysis of boolean functions. In *Proceedings of the 40th Annual ACM Symposium on the Theory of Computing* 569–578. ACM, New York, 2008. [MR2582688](#)
- [28] M. Talagrand. On Russo’s approximate zero–one law. *Ann. Probab.* **22** (1994) 1576–1587. [MR1303654](#)
- [29] M. Talagrand. How much are increasing sets positively correlated? *Combinatorica* **16** (2) (1996) 243–258. [MR1401897](#)
- [30] P. Wolff. Hypercontractivity of simple random variables. *Studia Math.* **180** (3) (2007) 219–236. [MR2314078](#)