MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Geometrical bounds for the variance and recentered moments

Tongseok Lim
Institute of Mathematical Sciences, ShanghaiTech University, 393 Middle Huaxia Road, Pudong, Shanghai China
Current address: Krannert School of Management, Purdue University, West Lafayette IN 47907 USA, lim336@purdue.edu

Robert J. McCann
Department of Mathematics, University of Toronto, Toronto ON Canada M5S 2E4, mccann@math.toronto.edu, http://www.math.toronto.edu/mccann/

We bound the variance and other moments of a random vector based on the range of its realizations, thus generalizing inequalities of Popoviciu (1935) and Bhatia and Davis (2000) concerning measures on the line to several dimensions. This is done using convex duality and (infinite-dimensional) linear programming.

The following consequence of our bounds exhibits symmetry breaking, provides a new proof of Jung's theorem (1901), and turns out to have applications to the aggregation dynamics modelling attractive-repulsive interactions: among probability measures on $\mathbf{R}^n$ whose support has diameter at most $\sqrt{2}$, we show that the variance around the mean is maximized precisely by those measures which assign mass $1/(n+1)$ to each vertex of a standard simplex. For $1 \le p < \infty$, the p-th moment — optimally centered — is maximized by the same measures among those satisfying the diameter constraint.

Key words: multidimensional moment bounds, random vectors, convex duality, infinite-dimensional linear programming, variance, Popoviciu, Bhatia, Davis, Jung, Legendre-Fenchel, isodiametric inequality
MSC2000 subject classification: *MSC2010* classification 62H05, Secondary 49N15, 52A40, 60E15, 90C46
OR/MS subject classification: Primary: Probability — distribution comparison; Secondary: Mathematics — convexity and Programming — infinite dimensional
History: Received February 13, 2020; revised August 18, 2020 and September 29, 2020.

---

1. Introduction This article concerns the extension of geometrical variance bounds from one to higher dimensions. Let $K \subset \mathbf{R}^n$ be a compact set and $\mathcal{P}(K)$ denote the Borel probability measures supported on K. Let

$$\bar{x}(\mu) := \int_{\mathbf{R}^n} x d\mu(x) \tag{1}$$

$$\text{and} \quad \mathrm{Var}(\mu) := \int_{\mathbf{R}^n} |x - \bar{x}(\mu)|^2 d\mu(x) \tag{2}$$

denote the barycenter (or mean) and (the trace of the) the variance of $\mu \in \mathcal{P}(K)$. Here $|\cdot|$ is the Euclidean norm, and $\operatorname{spt}\mu$ will denote the smallest closed subset of $\mathbf{R}^n$ containing the full mass of μ. When $K := [\underline{k}, \bar{k}] \subset \mathbf{R}$, an inequality due to Bhatia and Davis [4] asserts

$$\mathrm{Var}(\mu) \le (\bar{k} - \bar{x}(\mu))(\bar{x}(\mu) - \underline{k}), \tag{3}$$

with equality if and only if $\operatorname{spt}\mu \subset \{\underline{\mathrm{k}}, \bar{k}\}$. Optimizing over all possible means yields

$$\mathrm{Var}(\mu) \le \frac{1}{4}(\bar{k} - \underline{\mathrm{k}})^2, \tag{4}$$

with equality if and only if $\mu = \frac{1}{2}(\delta_{\underline{\mathrm{k}}} + \delta_{\bar{k}})$ — a result known since Popoviciu's work [26] on polynomial roots, as explained in Jensen and Styan [18]. We propose to explore higher dimensional, i.e. $n > 1$, generalizations of bounds such as (3)–(4) and their cases of equality.

In higher dimensions, the shape of the set $K \subset \mathbf{R}^n$ plays a non-trivial role in the formulation of such a bound. However, it turns out that the variance maximizing measures must — in each case — be supported on the intersection of K with an enclosing sphere. This is the content of our first result, whose statement requires taking the convex envelope of the function

$$\phi_K(x) := \begin{cases} -|x|^2 & \text{if} \quad x \in K, \\ +\infty & \text{if } x \in \mathbf{R}^n \setminus K. \end{cases} \tag{5}$$

Convex envelopes are conveniently expressed using the Legendre transform.

Given a Banach space Z and its dual Z^*, recall the Legendre-Fenchel transform of a function $f : Z \longrightarrow \mathbf{R} \cup \{+\infty\}$ is defined on Z^* by

$$f^*(z^*) := \sup_{z \in Z} z^*(z) - f(z). \tag{6}$$

where $z^*(z)$ denotes the duality pairing. The double Legendre transform f^{**} is well-known to be the largest lower semicontinuous convex function on Z^{**} whose restriction to Z is dominated by f. Letting $\mathrm{conv}(K)$ denote the smallest closed convex set containing K and $\mathrm{int}(K)$ the interior of K, our multidimensional analogs of the Bhatia, Davis [4] and Popoviciu [26] inequalities (3)–(4) are the following:

Theorem 1 **(Enclosing spheres support variance maximizers)**.
(a) If the measure $\mu \in \mathcal{P}(\mathbf{R}^n)$ has barycenter $\bar{x}(\mu)$ and vanishes outside the compact set $K \subset \mathbf{R}^n$, then

$$\mathrm{Var}(\mu) \le -|\bar{x}(\mu)|^2 - \phi_K^{**}(\bar{x}(\mu)) \tag{7}$$

where ϕ_K^{**} is defined as in (5)–(6). If $\bar{x}(\mu) \in \mathrm{int}(\mathrm{conv}(K))$, then equality holds in (7) if and only if μ vanishes outside the boundary of some closed ball B containing K, i.e. if and only if $\mu[K \cap \partial B] = 1$.

(b) Among measures with all barycenters, μ maximizes variance over $\mathcal{P}(K)$ if and only if $\mu[K \cap \partial B] = 1$ where B is the smallest closed ball containing K, and $\bar{x}(\mu)$ is the center of B. Moreover, in this case $\mathrm{Var}(\mu) = R^2$ where R is the radius of B.

We note that the function $-\phi_K^{**}$ is the concave envelope of $-\phi_K$. Some refinements and examples include:

Remark 1 (Specialization to one-dimension). In the classical context $n = 1$ and $K = [\underline{\mathrm{k}}, \bar{k}]$, we recover (3) from (7) by noting $t \in [0, 1]$ and $x = (1-t)\underline{\mathrm{k}} + t\bar{k}$ imply

$$-\phi_K^{**}((1-t)\underline{\mathrm{k}} + t\bar{k}) = (1-t)\underline{\mathrm{k}}^2 + t\bar{k}^2$$

and hence

$$-\phi_K^{**}(x) = (\bar{k} + \underline{\mathrm{k}})x - \bar{k}\underline{\mathrm{k}}.$$

For our chacterization of equality in (7), the assumption $\bar{x}(\mu) \in \mathrm{int}(\mathrm{conv}(K))$ is in general necessary and cannot be omitted, as the following example indicates.

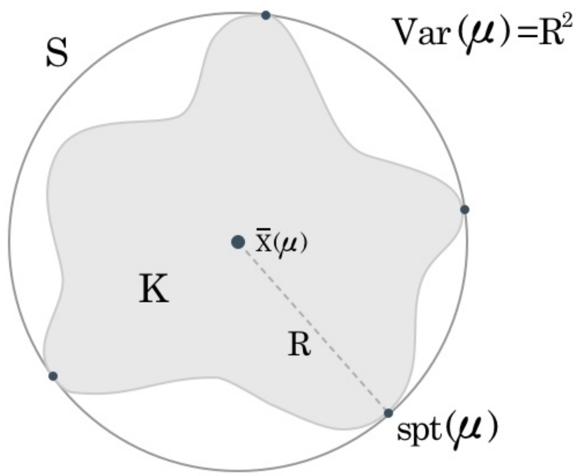


FIGURE 1. Variance maximizer μ over $\mathcal{P}(K)$ in Theorem 1.

EXAMPLE 1 (STADIUM). Taking $K \subset \mathbf{R}^2$ to be the convex hull of two (say unit) balls in $\mathbf{R}^2$ and constraining the barycenter $\bar{x}$ to be (say) the midpoint of one of the flat sides of K shows the conclusion of Theorem 1(a) need not remain true for all $\bar{x}$ in the boundary of $\mathrm{conv}(K)$; the putative enclosing sphere degenerates to a halfspace $H \supset K$ in this example, with $\mu[K \cap \partial H] = 1$ being necessary but not sufficient for equality in (7). See the next remark concerning lower dimensional spheres, however.

REMARK 2 (CASES OF EQUALITY FOR BOUNDARY BARYCENTERS). Let $L := conv(K)$ denote the convex hull of $K \subset \mathbf{R}^n$, i.e. the smallest closed convex set containing K. Theorem 2.1.2 of Schneider's book [29] asserts that each point $x \in L$ belongs to the relative interior of a uniquely determined face F_x of L, where a face $F \subset L$ refers to a set containing the endpoints of every segment in L whose midpoint lies in F. When $\bar{x}(\mu) \in \partial L$ in Theorem 1, let j denote the dimension of $F_{\bar{x}(\mu)}$. When $j > 1$, applying the same theorem to $F_{\bar{x}(\mu)} \subset \mathbf{R}^j$ instead of $K \subset \mathbf{R}^n$ shows equality holds in (7) if and only if μ is concentrated on a round sphere $\mathbf{S}^{j-1} \subset \mathbf{R}^j$ enclosing $F_{\bar{x}(\mu)}$. When $j = 0$ then $\bar{x}(\mu)$ is an extreme point of K, μ is a Dirac measure, and (7) becomes an equality.

EXAMPLE 2 (APPLICATIONS TO SAMPLE GEOMETRIES). Theorem 1(a) and Remark 2 imply:

(a) (Ball) If $K = \overline{B_R(0)}$ then $\mathrm{Var}(\mu) \le R^2 - |\bar{x}(\mu)|^2$, and equality holds if and only if μ is supported on ∂K.

(b) (Ellipse) If $K = \{(x_1, x_2) \in \mathbf{R}^2 \mid (\frac{x_1}{a})^2 + (\frac{x_2}{b})^2 \le 1\}$ with $a > b > 0$ and $\mathrm{Var}(\mu) = -|\bar{x}(\mu)|^2 - \phi_K^{**}(\bar{x}(\mu))$ then $\mathrm{spt}\,\mu$ consists of at most two points.

(c) (Rectangular parallelopiped) If $K = \prod_{i=1}^n [-a_i, a_i]$ is non-empty, then $\mathrm{Var}(\mu) \le -|\bar{x}(\mu)|^2 + \sum_{i=1}^n a_i^2$, and equality holds if and only if μ is concentrated on the vertices of K.

(d) (Diamond) If $a_1 > a_2 > 0$ and $K = \{(x_1, x_2) \in \mathbf{R}^2 \mid |\frac{x_1}{a_1}| + |\frac{x_2}{a_2}| \le 1\}$, then $\mathrm{Var}(\mu) \le a_1^2 - \frac{a_1^2 - a_2^2}{a_2}|\bar{x}_2(\mu)| - |\bar{x}(\mu)|^2$ and equality holds if and only if μ concentrates at the two vertices of K farthest from the origin, plus at most one of its other two vertices.

In cases (a) and (c) it is easy to check ϕ_K^{**} is constant on K: its negative coincides with the square distance from the origin to the farthest point in K (which follows from the fact that any affine perturbation of the concave function $-|x|^2$ from (5) can only be minimized at extreme points of K, and is constant on a sphere through the minimizing points). The conclusions of (b) and (d) follow since an enclosing circle cannot intersect an ellipse in more than two points, nor a diamond in more than the three mentioned points.

Theorem 1(b) also has analogs for other, possibly anisotropic, measures of the extent to which the mass of μ is concentrated or dispersed. To illustrate, we give the following definition, which can be contrasted with other generalizations of the variance from the literature, such as those of Pronzato et al. [27] and its references.

Let $V : \mathbf{R}^n \to [0,\infty)$ be convex. Define

$$\mathrm{Var}_V(\mu) := \inf_{z\in\mathbf{R}^n} \int_{\mathbf{R}^n} V(x-z)d\mu(x). \tag{8}$$

We say that $V : \mathbf{R}^n \to [0,\infty)$ is *coercive* if its sublevel sets $V^{-1}([0,\lambda])$ are compact for each $\lambda \geq 0$.

Remark 3 (Generalized variances and centered p-th moments). If V is coercive the infimum (8) is attained. If V is also strictly convex and $\mathrm{Var}_V(\mu) < \infty$, then the point $\bar{x}_V(\mu)$ attaining it is unique, by the displacement convexity of McCann [23]. We can think of $\mathrm{Var}_V(\mu)$ and $\bar{x}_V(\mu)$ as generalizations of the variance and mean, which reduce to the classical variance and mean in case $V(x) = |x|^2$. When $V(x) = |x|^p$ they reduce to p-th moments, but centered on $\bar{x}_V(\mu)$ rather than the classical mean.

We then generalize Theorem 1(b) as follows:

Theorem 2 **(Maximizing generalized variances)**. *Let $K \subset \mathbf{R}^n$ be compact and $V : \mathbf{R}^n \to [0,\infty)$ be convex and coercive. Let $\lambda \geq 0$ be the smallest value for which there exists $z \in \mathbf{R}^n$ with $K + z \subset V^{-1}([0,\lambda])$. Then*

$$\lambda = \sup_{\mu\in\mathcal{P}(K)} \mathrm{Var}_V(\mu). \tag{9}$$

Moreover $\mu \in \mathcal{P}(K)$ attains this supremum if and only if there exists

$$z_* \in \arg\min_{z\in\mathbf{R}^n} \int V(x-z)d\mu(x) \tag{10}$$

such that $\mathrm{spt}\,\mu \subset V^{-1}(\lambda) - z_*$.

Taking $V(x) = |x|^2$ so that $\mathrm{Var}_V = \mathrm{Var}$, we recognize $\mathrm{spt}\,\mu \subset V^{-1}(\lambda) - z_*$ as the sphericity condition from Theorem 1 — and (10) as the barycenter condition from the same theorem. More generally, view (9) as the value to player 1 of a two-player zero-sum game — in which the first player chooses $\mu \in \mathcal{P}(K)$ and the second player, knowing μ, chooses $z \in \mathbf{R}^n$. We can then interpret (10) as player 2's best response to μ, and $\mathrm{spt}\,\mu \subset V^{-1}(\lambda) - z_*$ as characterizing player 1's best response to z_*; together they form the conditions for a saddle-point in the payoff function or equivalently, for a Nash equilibrium; c.f. McCann and Guillen [24].

1.1. Regular simplices maximize moments, given diameter

For fixed barycenter $\bar{x}(\mu)$, the variance (2) is a linear function on the convex set $\mathcal{P}(K)$. It is thus not surprising that our proof of Theorem 1 relies on linear programming duality (and convex-concave minimax theory in the case of Theorem 2). A more challenging question is to give sharp bounds on the variance and moments of all measures μ in a *non-convex* set, to which many of the standard techniques in the calculus of variations, see e.g. Kawohl [20], McCann [23], and Borwein and Zhu [6], no longer apply.

The example which motivated our interest in this problem concerns the measures satisfying a diameter bound $\mathrm{diam}[\mathrm{spt}\,\mu] \leq 1$. Here $\mathrm{spt}\,\mu$ refers to the smallest closed set containing the full mass of μ. This question arises as an important special case in our work on attractive-repulsive interactions, which addresses the patterns formed by a large collection of particles or organisms all preferring to be at distance one from each other, Lim and McCann [21]. We resolve this question below, by showing among measures $\mu \in \mathcal{P}(\mathbf{R}^n)$ with $\mathrm{diam}[\mathrm{spt}\,\mu] \leq 1$, the variance and other moments are maximized precisely when the mass of μ is evenly distributed over the $n+1$ vertices of a regular, unit diameter simplex, i.e. an equilateral triangle if $n=2$ and a regular tetrahedron if $n=3$.

While it may seem surprising to find this solution breaks rotational symmetry, such symmetry breakings undoubtedly bear some responsibility for the zoo of patterns which emerge from the flocking and swarming models discussed in Albi et al. [1], Balagué et al. [2], Bertozzi et al. [3], Burchard et al. [7], Carrillo et al. [9] [10] [11] [8], Choksi et al. [12], Fellner and Raol [16], Frank and Lieb [17], Lopes [22] and the references there, of which the present problem represents a limiting case, Lim and McCann [21]. Indeed, for a particular family of such models, Sun et al. [31] discovered that equidistribution over the vertices of a regular unit diameter simplex is a stationary state around which the linearized dynamics undergoes a bifurcation from stability to instability; in our companion work [ibid], Theorem 3 below is used to confirm their conjecture that nonlinear stability of the full dynamics holds throughout the linearly stable regime. We were also reminded of the role linear programming duality plays in confirming the optimality of sphere packings in certain dimensions, as in Odlyzko and Sloane [25], Cohn and Elkies [13] and Viazovska [32].

Definition 1 (Simplices). (a) A set $K \subset \mathbf{R}^n$ is called a *top-dimensional simplex* if K has non-empty interior and is the convex hull of $n+1$ points $\{x_0, x_1, ..., x_n\}$ in $\mathbf{R}^n$.

(b) A set $K \subset \mathbf{R}^n$ is called a regular k-simplex if it is the convex hull of $k+1$ points $\{x_0, x_1, ..., x_k\}$ in $\mathbf{R}^n$ satisfying $|x_i - x_j| = d$ for some $d > 0$ and all $0 \le i < j \le k$. The points $\{x_0, x_1, ..., x_k\}$ are called *vertices* of the simplex.

(c) In particular, it is called a *unit k-simplex* if $d = 1$.

Remark 4 (Regular n-simplices $K \subset \mathbf{R}^n$ are top-dimensional). A regular n-simplex with sidelength $d = \sqrt{2}$ is linearly isometric to the following *standard simplex* in $\mathbf{R}^{n+1}$

$$\Delta^n := \{a = \{a_1, ..., a_{n+1}\} \in [0,1]^{n+1} \mid \sum_{i=1}^{n+1} a_i = 1\}, \tag{11}$$

which can be verified by simple induction on dimension.

We can now state the following:

Theorem 3 **(Isodiametric variance bounds and cases of equality)**. *Let $V(x) = v(|x|)$ with $v : [0,\infty) \longrightarrow [0,\infty)$ convex and increasing. If the support of a Borel probability measure μ on $\mathbf{R}^n$ has diameter no greater than d, then $\mathrm{Var}_V(\mu) \le v(r_n d)$ where $r_n = \sqrt{\frac{n}{2n+2}}$. Equality holds if and only if μ assigns mass $1/(n+1)$ to each vertex of a regular n-simplex having diameter d.*

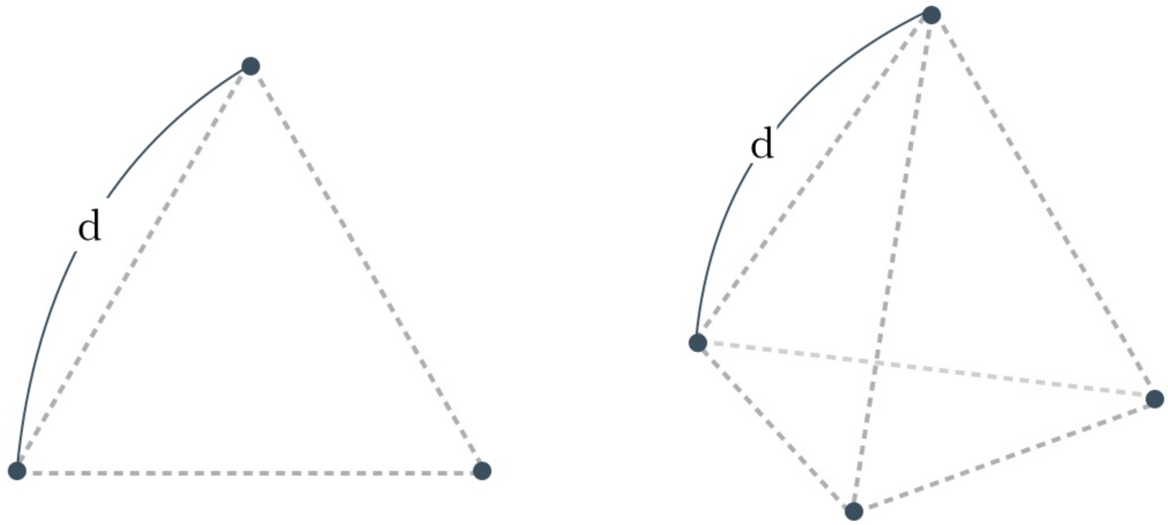
Example 3 (Isodiametric bounds on recentered p-th moments). Take $V(x) = |x|^p$ with $p \ge 1$ in Theorem 3.

This theorem gives a variational characterization of the unit n-simplex. It can also be viewed as another generalization of Popoviciu's inequality (4) from $n = 1$ to higher dimensions $n > 1$.

1.2. Epilog

After Theorem 3 was announced on the arXiv [21] (in the special case $V(x) = |x|^2$), we learned of an isodiametric inequality due to Jung [19] in which regular simplices also play a crucial role; a modern treatment is given in Danzer et al. [14]:

Theorem 4 **(Jung)**. *Let $K \subset \mathbf{R}^n$ be compact with $\mathrm{diam}(K) = 1$. Then K is contained in a closed ball of radius $r_n = \sqrt{\frac{n}{2n+2}}$. Moreover, unless it lies in some smaller ball, K contains the vertices of a unit n-simplex.*



(A) (B)

Figure 2. Support spt(μ) of the optimizer μ in Theorem 3: (A) in $\mathbf{R}^2$; (B) in $\mathbf{R}^3$.

The constant r_n which appears in these theorems also relates spherical Hausdorff measure to Hausdorff measure, as in Federer [15]. Below we shall show how Theorem 4 follows from our isodiametric variance bound, thus yielding a new proof of Jung's theorem. In an appendix to Lim and McCann [21] we show the converse is also true: Theorem 3 can be derived from Jung's theorem using elementary geometry. Thus the two theorems are in some sense equivalent. We are grateful to Tomasz Tkocz and an anonymous seminar participant at Seoul National University, for drawing our attention to Jung's theorem.

1.3. Plan of the paper:

The next section develops the linear programming and convex duality based proof of Theorems 1 and 2. Section 3 addresses the non-convex problem of maximizing moments under a diameter constraint. It uses induction on dimension to prove a geometric lemma which allows us to deduce Theorem 3, before closing with a new proof of Jung's theorem.

2. A geometric family of ∞-dimensional linear programs This section uses linear programming and convex analysis to extend the one-dimensional inequalities (3)–(4) of Bhatia, Davis and Popoviciu to higher dimensions, i.e. $n > 1$. Translation invariance allows us to center our measures so that $\bar{x}(\mu) = 0$ without loss of generality. For each compact $K \subset \mathbf{R}^n$ let

$$\mathcal{P}_0(K) := \{\mu \in \mathcal{P}(K) \mid \bar{\mathrm{x}}(\mu) = 0\} \tag{12}$$

denote the set of probability measures on K having vanishing mean.

Our first goal is to establish the following duality result of Fenchel and Rockafellar [28] type:

Proposition 1 **(A strong duality with attainment)**. *If $K \subset \mathbf{R}^n$ is compact then*

$$\sup_{\mu \in \mathcal{P}_0(K)} \int_K |x|^2 d\mu(x) = \inf_{q \in \mathbf{R}^n} \phi_K^*(-2q) = -\phi_K^{**}(0) \tag{13}$$

where ϕ_K^* and ϕ_K^{**} denote the Legendre transforms (6) of (5). The supremum is attained if $0 \in L$ and the infimum if $0 \in \mathrm{int}(L)$, where $L := \mathrm{conv}(K)$. A measure $\mu \in \mathcal{P}_0(K)$ and point $q \in \mathbf{R}^n$ optimize (13) if and only if μ vanishes outside $K \cap \partial B_R(q)$ for the smallest sphere $\partial B_R(q)$ centered at q and enclosing K.

Identity (13) can be motivated heuristically as in, e.g. McCann and Guillen [24]. Introducing Lagrange multipliers h and q for the mass and barycenter constraints,

$$\begin{aligned}
&\sup_{\mu\in\mathcal{P}_0(K)} \int_K |x|^2 d\mu(x) \\
&= \sup_{\mu\in\mathcal{M}_+(K)} \inf_{h\in\mathbf{R},q\in\mathbf{R}^n} h\big(1-\mu(K)\big) + \int_K \big(|x|^2 - 2q\cdot x\big) d\mu(x) \\
&\le \inf_{q\in\mathbf{R}^n,h\in\mathbf{R}} \sup_{\mu\in\mathcal{M}_+(K)} \left[h + \int_K \big(|x|^2 - 2q\cdot x - h\big) d\mu(x)\right] \\
&= \inf_{q\in\mathbf{R}^n} \inf_{h\ge|x|^2-2q\cdot x \ \forall x\in K} h \\
&= \inf_{q\in\mathbf{R}^n} \sup_{x\in K} |x|^2 - 2q\cdot x, \\
&= -\phi_K^{**}(0)
\end{aligned}$$

where $\mathcal{M}_+(K)$ denotes the set of non-negative Borel measures of finite total mass on $K \subset \mathbf{R}^n$. This inequality can be interpreted as asserting that foreknowledge of one's opponent's strategy cannot be a disadvantage in a two-player zero-sum game; it may or may not confer an advantage, depending on the structure of the game. Statement (13) is basically the assertion that the inequality can be replaced with an equality in our case, which is a consequence of the payoff expression in square brackets having a saddle point or equivalently, of the game having a Nash equilibrium. Since the payoff is bilinear in the variables μ and (h,q), this may not be surprising. Due to lack of compactness however, a rigorous proof along standard lines requires some machinery. Therefore, recall Theorem 4.4.3 from the book of Borwein and Zhu [5]:

THEOREM 5 **(Fenchel-Rockafellar duality)**. *Let $A: Z \longrightarrow Y$ be a bounded linear transformation of Banach spaces Z and Y, equipped with functions $f: Z \longrightarrow \mathbf{R} \cup \{+\infty\}$ and $g: Y \longrightarrow \mathbf{R} \cup \{+\infty\}$. If g is continuous at some point in $A(\mathrm{dom}\, f)$, then*

$$\sup_{y^*\in Y^*} -f^*(A^*y^*) - g^*(-y^*) = \inf_{z\in Z} f(z) + g(Az),$$

where Y^* denotes the Banach space dual to Y and $\mathrm{dom}\, f := f^{-1}(\mathbf{R})$. Moreover, the supremum is attained if finite.

Proof of Proposition 1: Let $Z := \mathbf{R}^{n+1}$ be Euclidean and equip the continuous functions $Y := C(K)$ on K with the supremum norm, so that $Z^* = \mathbf{R}^{n+1}$ and $Y^* = \mathcal{M}(K)$, the space of signed measures on K normed by total variation. Take $A(z) = z_0 + \sum_{i=1}^n z_i x_i =: \xi(x) \in Y$ so that $A^*\mu = \int_K (1,x) d\mu(x)$ gives the mass and barycenter of $\mu \in \mathcal{M}(K)$. Set $f(z_0,\dots,z_n) := z_0$ so that

$$f^*(z^*) = \begin{cases} 0 & \text{if } z^* = (1,0,\dots,0), \\ \infty & \text{else.} \end{cases}$$

Also set

$$g(\xi) := \begin{cases} 0 & \text{if } \xi(x) \ge |x|^2 \quad \forall x \in K, \\ \infty & \text{else,} \end{cases}$$

so that

$$g^*(\mu) = \begin{cases} \int_K |x|^2 d\mu(x) & \text{if } \mu \le 0, \\ \infty & \text{else.} \end{cases}$$

Inserting these choices into Theorem 5 yields (13), noting the definitions (5)–(6) of ϕ_K^*. If $0 \in L := \mathrm{conv}(K)$ then $\mathcal{P}_0(K)$ is non-empty and the supremum is bounded above and below (by the infimum and zero) hence attained (also by Theorem 5).

If $0 \in \mathrm{int}(L)$ then $rB \subset L$ for $r > 0$ sufficiently small, where $B := B_1(0)$ is the centered unit ball. Then $\phi_L \le \phi_{rB}$ hence $\phi_L^*(q) \ge \phi_{rB}^*(q) = r|q| + r^2$ grows without bound as $|q| \to \infty$. Being lower semicontinuous, ϕ_L^* then attains its minimum. On the other hand, the concavity of $x \mapsto -|x|^2$ implies $\phi_L^* = \phi_K^*$, as we now argue. Indeed $\phi_L^* \ge \phi_K^*$ follows directly from $K \subset L$ and $\phi_K \ge \phi_L$. Conversely, given any affine function a on $\mathbf{R}^n$ dominated by ϕ_K, we find $a \le \phi_L$ also, since $\phi_K = \phi_L$ outside $L \setminus K$, and each $x \in L \setminus K$ can be approximated by convex combinations $x^j = \sum_{i=1}^j t_i^j k_i^j$ of points $k_i^j \in K$ with $t_i^j \ge 0$ and $\sum_{i=1}^j t_i^j = 1$, so

$$\begin{aligned}
\phi_L(x^j) &= -|x^j|^2 \\
&\ge -\sum_{i=1}^j t_i^j |k_i^j|^2 \\
&= \sum_{i=1}^j t_i^j \phi_K(k_i^j) \\
&\ge a(x^j),
\end{aligned}$$

and the limit $x = \lim_{j\to\infty} x^j$ yields $\phi_L \ge a$ as desired. Since ϕ_K^{**} is the supremum of such affine functions a, we conclude $\phi_L^{**} \ge \phi_K^{**}$, which implies $\phi_K^* \ge \phi_L^*$ hence $\phi_K^* = \phi_L^*$.

To characterize the optimizers, let $\mu \in \mathcal{P}_0(K)$ and $q \in \mathbf{R}^n$. Then

$$\begin{aligned}
\int_K |x|^2 d\mu(x) &= \int_K (|x-q|^2 - |q|^2) d\mu(x) \\
&\le \max_{x\in K} |x-q|^2 - |q|^2 \\
&= \phi_K^*(-2q) \\
&=: R^2 - |q|^2
\end{aligned} \tag{14}$$

and equality holds if and only if μ vanishes outside the set

$$\arg\max_{x\in K} |x-q|^2 = K \cap \partial B_R(q);$$

here R is the smallest radius for which $K \subset \overline{B_R(q)}$. On the other hand, $\mu \in \mathcal{P}_0(K)$ and $q \in \mathbf{R}^n$ optimize (13) if and only if equality holds in (14), so the proposition is established. $\square$

Expression (13) is particularly convenient for selecting the translation of K which maximizes the value of the linear program using the following lemma:

Lemma 1 **(Optimal translation of a domain relative to the origin)**. *For compact $K \subset \mathbf{R}^n$, we have $\phi_{K-w}^{**}(x) = (|x+w|^2 - |x|^2) + \phi_K^{**}(x+w)$. In particular, $\phi_K^{**}(0) \le \phi_{K-w}^{**}(0)$ for all translations $w \in \mathbf{R}^n$ if and only if ϕ_K^{**} attains its minimum at the origin.*

Proof. The Legendre-Fenchel transform (6), applied to ϕ_K, yields

$$\begin{aligned}
\phi_{K-w}^*(y) &= |w|^2 - w \cdot y + \phi_K^*(y - 2w) \qquad \text{and} \\
\phi_{K-w}^{**}(x) &= |w|^2 + 2w \cdot x + \phi_K^{**}(x + w),
\end{aligned}$$

hence

$$\phi_{K-w}^{**}(0) = |w|^2 + \phi_K^{**}(w). \tag{15}$$

Recall that a convex function f on a Banach space Z attains its minimum at x if and only if $0 \in \partial f(x)$, where

$$\partial f(x) := \{z^* \in Z^* \mid f(z) \geq f(x) + z^*(z-x) \quad \forall z \in Z\}. \tag{16}$$

The formula above shows $f(w) := \phi^{**}_{K-w}(0)$ to be a strictly convex function of w with $\partial f(0) = \partial \phi^{**}_K(0)$, so $\phi^{**}_{K-w}(0)$ attains its minimum at $w = 0$ if and only if $\phi^{**}_K(w)$ does as well. $\square$

Proof of Theorem 1: (a) For a compact set $K \subset \mathbf{R}^n$ with $w \in \mathbf{R}^n$, Lemma 1 yields $\phi^{**}_{K-w}(0) = |w|^2 + \phi^{**}_K(w)$. In (13) this gives

$$\sup_{\nu \in \mathcal{P}_0(K-w)} \int |x|^2 d\nu = -\phi^{**}_{K-w}(0) = -|w|^2 - \phi^{**}_K(w). \tag{17}$$

Letting μ denote the translation of ν by $w := \bar{x}(\mu)$ yields (7). If $w \in \operatorname{int}(\operatorname{conv}(K))$, Proposition 1 states that $\nu \in \mathcal{P}_0(K-w)$ attains the supremum if and only if ν vanishes outside $K \cap \partial B_R(q)$ with $K \subset \overline{B_R(q)}$ for some $q \in \mathbf{R}^n$ and $R > 0$. In other words, if $\bar{x}(\mu) \in \operatorname{int}(\operatorname{conv}(K))$, then (7) becomes an equality if and only if μ is supported on the boundary of a closed ball containing K.

(b) Can be proved using Lemma 1 as in Lim and McCann [21]. Alternately (b) follows from the choice $V(x) = |x|^2$ in Theorem 2, whose proof appears just below. $\square$

Proof of Theorem 2. Recall from e.g. McCann and Guillen [24, §2.2] that

$$\begin{aligned}
\sup_{\mu \in \mathcal{P}(K)} \operatorname{Var}_V(\mu) &= \sup_{\mu \in \mathcal{P}(K)} \inf_{z \in \mathbf{R}^n} \int V(x-z) d\mu(x) \\
&\leq \inf_{z \in \mathbf{R}^n} \sup_{\mu \in \mathcal{P}(K)} \int V(x-z) d\mu(x) \\
&= \min_{z \in \mathbf{R}^n} \max_{x \in K} V(x-z) \\
&= \lambda.
\end{aligned}$$

Combining compactness of K with coercivity and continuity of the convex function V allows us to replace $\mathbf{R}^n$ with a sufficiently large closed ball $\overline{B_R(0)}$ without affecting the values of either infimum; the infima are therefore attained, and the inequality above becomes an equality according to convex-concave minimax theory, as in, e.g. Strasser [30, Theorem 45.8].

From the definition of λ, there exists z_* such that $K - z_* \subset V^{-1}([0,\lambda])$. Thus $\mu \in \mathcal{P}(K)$ satisfies

$$\inf_{z \in \mathbf{R}^n} \int_K V(x-z) d\mu(z) \leq \int_K V(x - z_*) d\mu(x) \leq \lambda$$

with the first inequality being saturated if and only if (10) holds, and the second inequality being saturated if and only if $V(x - z_*) = \lambda$ on $\operatorname{spt}\mu$. In light of (9), these two conditions are necessary and sufficient to ensure that μ is a maximizer. $\square$

3. Isodiametric variance and p-th moment bounds

This section establishes our isodiametric variance bound and cases of equality: Theorem 3. Let us briefly outline the strategy of our proof. Fix $V(x) = v(|x|)$ radially symmetric, convex and increasing. For each compact set $K \subset \mathbf{R}^n$ of unit diameter, Theorem 2 asserts (i) that the maximizer of $\operatorname{Var}_V(\mu)$ on $\mathcal{P}(K)$ vanishes outside the smallest sphere enclosing K and (ii) the center of this sphere attains the infimum (8) defining $\operatorname{Var}_V(\mu)$. We may, without loss of generality assume that K has been translated so that this sphere is centered on the origin. We shall now show the radius of this sphere cannot exceed the radius $r_n := \sqrt{\frac{n}{2n+2}}$ of the unit n-simplex. To do so we use an induction on dimension, which is based on the idea that if the centered sphere is too large, no measure whose support has unit diameter can have its center of mass at the origin. More precisely, we show the following elementary yet crucial geometric proposition which characterizes the unit simplex.

Proposition 2 **(Tension between diameter and center-of-mass constraints)**.
(a) If $K \subset \partial B_r(0)$ is a subset of the radius $r > r_n := \sqrt{\frac{n}{2n+2}}$ centered sphere in $\mathbf{R}^n$ and $\operatorname{diam}(K) \le 1$, *then* $0 \notin \operatorname{conv}(K)$.

(b) If K is a subset of the centered sphere in $\mathbf{R}^n$ of radius r_n, $\operatorname{diam}(K) \le 1$ *and* $0 \in \operatorname{conv}(K)$, *then K is the set of vertices of a unit n-simplex.*

Proof of Proposition 2. (a) The proposition is trivial to verify when $n = 1$. To derive a contradiction, suppose the proposition holds in $\mathbf{R}^{n-1}$ but fails in $\mathbf{R}^n$. Then there exists a centered sphere S of radius r with $r > r_n$, and $K \subset S$ with $\operatorname{diam}(K) \le 1$ and $0 \in \operatorname{conv}(K)$. We can find $n+1$ points in K, say $X := \{x_0, x_1, ..., x_n\} \subset K$, such that $0 \in \operatorname{conv}(X)$. If the origin lies on the boundary of $\operatorname{conv}(X)$, then after intersecting the problem with a hyperplane supporting $\operatorname{conv}(X)$ at 0, the inductive hypothesis yields the desired contradiction using $r_{n-1} < r_n$. We may therefore assume $0 \in \operatorname{int}\operatorname{conv}(X)$, so that $\operatorname{conv}(X)$ is a top-dimensional simplex in $\mathbf{R}^n$.

Without loss of generality, let $x_0 = r\hat{e}_1 = (r, 0, ..., 0)$. Define

$$U := \{x \in S \mid |x - x_0| \le 1\}.$$

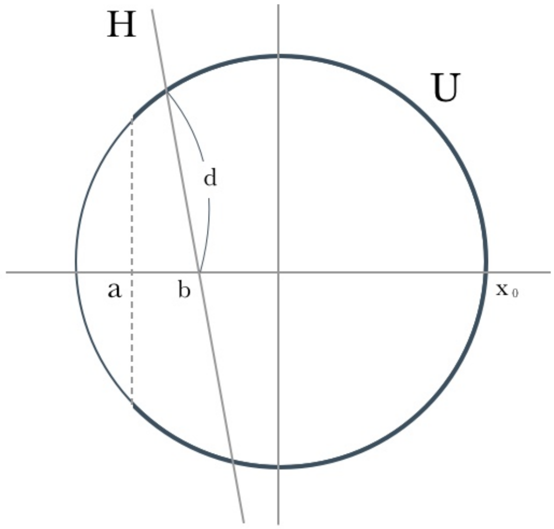
Then $\partial_{rel} U := \{x \in S \mid |x - x_0| = 1\}$ is a $(n-2)$-dimensional sphere of radius r' and center $a = a_1\hat{e}_1$ for some $r' > 0$ and $a_1 \in \mathbf{R}$. Since $0 \in \operatorname{int}\operatorname{conv}(X)$ implies $0 \in \operatorname{int}\operatorname{conv}(U)$, we see that $a_1 < 0$. And $r > r_n$ implies $r' > r_{n-1}$, as $r' = r_{n-1}$ precisely when $r = r_n$. Now consider the unique hyperplane H which contains the $(n-1)$-simplex with vertices $X' = \{x_1, ..., x_n\} \subset X$. Let L be the one-dimensional subspace spanned by $\hat{e}_1$. Then $H \cap L \ne \emptyset$ since $0 \in \operatorname{int}\operatorname{conv}(X)$. Let $b = b_1\hat{e}_1 := H \cap L$. Then $a_1 \le b_1$ since $X' \subset U$, and $b_1 < 0$ since $0 \in \operatorname{int}\operatorname{conv}(X)$. Now define the disk $D := \operatorname{conv}(H \cap S)$ whose (relative) boundary is the $(n-2)$-dimensional sphere $\partial_{rel} D := H \cap S$. Note that $b \in D$ and $X' \subset \partial_{rel} D$. Define

$$d := \operatorname{dist}(b, \partial_{rel} D).$$

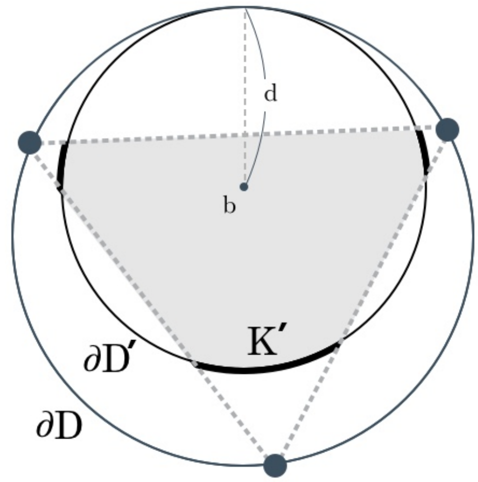
Notice that the facts $a_1 \le b_1 < 0$ and $\partial_{rel} D \subset U$ imply $d \ge r'$, hence $d > r_{n-1}$ (see Figure **??**).

The desired contradiction (and proposition) will follow if we show that $b \notin \operatorname{conv}(X')$, as this will imply $0 \notin \operatorname{conv}(X)$. To achieve this, suppose on the contrary $b \in \operatorname{conv}(X')$. Let D' be the $(n-1)$-dimensional closed ball in H of center b and radius d, and let $\partial_{rel} D'$ be its boundary sphere. Note that $b \in \operatorname{conv}(X') \cap D'$. Since none of the extreme points of $\operatorname{conv}(X')$ lie in the relative interior of D', it follows the extreme points of $\operatorname{conv}(X') \cap D'$ all lie on the boundary sphere $\partial_{rel} D'$. Setting $K' := \operatorname{conv}(X') \cap \partial_{rel} D'$, the Krein-Milman theorem implies $b \in \operatorname{conv}(K')$. But this contradicts the inductive hypothesis, which asserts that the center b of a sphere $S' := \partial_{rel} D'$ of radius $d > r_{n-1}$ cannot lie in the convex hull of any subset $K' \subset S'$ whose diameter is bounded by one.
(b) We proceed as in part (a). Suppose the proposition holds in $\mathbf{R}^{n-1}$. Let S be the centered sphere of radius r_n in $\mathbf{R}^n$, and let $K \subset S$ be such that $\operatorname{diam}(K) \le 1$ and $0 \in \operatorname{conv}(K)$. As before we can find a subset X of K, the vertices of a n-simplex with $0 \in \operatorname{conv}(X)$, and in fact $0 \in \operatorname{int}\operatorname{conv}(X)$ by part (a). Note that the sphere $\partial_{rel} U$ now has radius r_{n-1}. Again consider the hyperplane H spanned by X', and observe that $b = b_1\hat{e}_1 \in \operatorname{conv}(X')$ since $0 \in \operatorname{conv}(X)$. Now if $a_1 < b_1$, then as before we have $d > r_{n-1}$. This yields a contradiction by part (a) and the last part of its proof. We conclude that $a_1 = b_1$, and this implies that H is the hyperplane containing b and having $x_0 = r_n\hat{e}_1$ as its normal. Then $X' \subset H \cap S = \partial_{rel} U$, and the induction hypothesis implies that X' must form vertices of a unit $(n-1)$-simplex. Hence X forms vertices of a unit n-simplex, inscribed in the sphere $S = \partial B_{r_n}(0)$.

It remains to show that $K = X$. Since $\operatorname{conv}(X)$ is an intersection of $n+1$ closed halfspaces and $X = \operatorname{conv}(X) \cap S$, any point $x' \in K \setminus X$ lies outside at least one of these halfspaces. Without loss of generality, we may suppose it lies in the halfspace $H_a := \{x \in \mathbf{R}^n \mid x \cdot \hat{e}_1 < a_1\}$. But this means $x' \in S \setminus U$, yielding $|x' - x_0| > 1$, which contradicts the assumption $\operatorname{diam}(K) \le 1$. $\square$



(A)



(B)

FIGURE 3. In the proof of Proposition 2: (A), $a, b \in \mathbf{R}^n$ and $d > 0$; (B) $b \in \mathrm{conv}(K')$ yields a contradiction.

We are now in a position to prove Theorem 3 by characterizing variance maximizing measures under a diameter constraint.

Proof of Theorem 3: Set $V(x) = v(|x|) \geq 0$ with v convex and increasing, and fix a compact set $K \subset \mathbf{R}^n$ with diameter no greater than 1, and let $\mu \in \mathcal{P}(K)$ be the probability measure on K which maximizes Var_V. Such a measure exists, by the weak-$*$ compactness of $\mathcal{P}(K)$ in the Banach space $\mathcal{M}(K)$ dual to $(C(K), \|\cdot\|_\infty)$ (or by Proposition 1 in case $V(x) = |x|^2$). We may assume K has been translated so that the origin $z_* = 0$ satisfies (10). In this case we claim $0 \in \mathrm{conv}(\mathrm{spt}\,\mu)$. If not, letting $0 \neq z$ be the point of $\mathrm{conv}(\mathrm{spt}\,\mu)$ closest to the origin, say $z = (r, 0, \ldots, 0)$, we find each point $x \in \mathrm{conv}(\mathrm{spt}\,\mu)$ lies in the halfspace to the right of z, hence is strictly closer to z than to 0, contradicting (10). Theorem 2 asserts μ vanishes outside the smallest sphere $\overline{B_R(0)}$ enclosing K, so that $\mathrm{Var}_V(\mu) = v(R)$. On the other hand, $\mathrm{spt}\,\mu \subset \partial B_R(0)$ has diameter at most one and contains the origin in its convex hull. Proposition 2 therefore asserts that $R \leq r_n$ and that when equality holds $\mathrm{spt}\,\mu$ coincides with the vertices of a unit n-simplex. Note that the uniform measure $\hat{\mu}$ on the vertices of this simplex has center of mass at the origin and $\mathrm{Var}_V(\hat{\mu}) = v(r_n)$. Remark 5 below shows no other measure on the vertices of the simplex has center of mass at the origin. If $R < r_n$ we conclude $\mathrm{Var}_V(\mu) < \mathrm{Var}_V(\hat{\mu})$, while if $R = r_n$ we conclude $\mu = \hat{\mu}$. Thus for the given diameter $d = 1$ of support, we have identified the maximum of $\mathrm{Var}_V(\cdot)$ and the measures which attain it uniquely (up to translations and rotations). □

REMARK 5 (EQUIDISTRIBUTION OVER THE SIMPLEX VERTICES). Since the vertices of the standard simplex (11) form a basis for $\mathbf{R}^{n+1}$, each point inside the simplex can be uniquely expressed as a convex combination of its vertices. Thus among measures on the vertices of the simplex, only the uniform measure has its barycenter at the point $\frac{1}{n+1}(1, \ldots, 1)$.

3.1. A new proof of Jung's theorem

Let conclude by showing how Jung's theorem [19] follows from the results just derived:

Proof of Theorem 4 using Theorems 3 and 1(b). Let $K \subset \mathbf{R}^n$ be compact with $\mathrm{diam}(K) \le 1$. Theorem 3 asserts that any $\mu \in \mathcal{P}(K)$ satisfies $\mathrm{Var}(\mu) \le r_n^2$, and Theorem 1(b) then implies that K can be contained in a closed ball of radius at most r_n. Now suppose K does not lie in a ball with radius strictly smaller than r_n. Then Theorem 1(b) provides $\mu \in \mathcal{P}(K)$ with $\mathrm{Var}(\mu) = r_n^2$ and Theorem 3 then implies that $\mathrm{spt}(\mu)$ contains the vertices of a unit n-simplex. □

Conversely, an appendix to our companion work, Lim and McCann [21], shows how Jung's theorem can be used to prove Theorem 3 — at least for $V(x) = |x|^2$, but the proof there adapts easily to other radially symmetric, convex increasing choices of $V(x) = v(|x|)$.

Acknowledgments. The present manuscript is partially based on material which appeared in an early draft of Lim and McCann [21] and which has been excised from subsequent versions of that preprint. TL is grateful for the support of ShanghaiTech University, and in addition, to the University of Toronto and its Fields Institute for the Mathematical Sciences, where parts of this work were performed. RM acknowledges partial support of his research by Natural Sciences and Engineering Research Council of Canada Grants 2015-04383 and 2020-04162. The authors are grateful to Guido de Philippis, Greg Kuperberg, Tomasz Tkocz, and an anonymous seminar participant at Seoul National University for stimulating interactions, to Cameron Davies for his careful reading of the manuscript, to Hyejung Choi for drawing the figures.

References

[1] Albi G, Balagué D, Carrillo JA, von Brecht J (2014) Stability analysis of flock and mill rings for second order models in swarming. *SIAM J. Appl. Math.* 74:794–818, URL http://dx.doi.org/10.1137/13091779X.

[2] Balagué D, Carrillo JA, Laurent T, Raoul G (2013) Dimensionality of local minimizers of the interaction energy. *Arch. Ration. Mech. Anal.* 209:1055–1088, URL http://dx.doi.org/10.1007/s00205-013-0644-6.

[3] Bertozzi AL, Kolokolnikov T, Sun H, Uminsky D, von Brecht J (2015) Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Commun. Math. Sci.* 13:955–985, URL http://dx.doi.org/10.4310/CMS.2015.v13.n4.a6.

[4] Bhatia R, Davis C (2000) A better bound on the variance. *Amer. Math. Monthly* 107(4):353–357, ISSN 0002-9890, URL http://dx.doi.org/10.2307/2589180.

[5] Borwein JM, Zhu QJ (2005) *Techniques of variational analysis.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 20 (New York: Springer-Verlag), ISBN 0-387-24298-8.

[6] Borwein JM, Zhu QJ (2013) Variational methods in the presence of symmetry. *Adv. Nonlinear Anal.* 2(3):271–307, ISSN 2191-9496, URL http://dx.doi.org/10.1515/anona-2013-1001.

[7] Burchard A, Choksi R, Topaloglu I (2018) Nonlocal shape optimization via interactions of attractive and repulsive potentials. *Indiana Univ. Math. J.* 67:375–395, URL http://dx.doi.org/10.1512/iumj.2018.67.6234.

[8] Carrillo JA, Choi YP, Hauray M (2014) The derivation of swarming models: mean-field limit and Wasserstein distances. *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, 1–46 (Springer, Vienna), URL http://dx.doi.org/10.1007/978-3-7091-1785-9_1.

[9] Carrillo JA, Figalli A, Patacchini FS (2017) Geometry of minimizers for the interaction energy with mildly repulsive potentials. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 34:1299–1308, URL http://dx.doi.org/10.1016/j.anihpc.2016.10.004.

[10] Carrillo JA, Hittmeir S, Volzone B, Yao Y (2019) Nonlinear aggregation-diffusion equations: radial symmetry and long time asymptotics. *Invent. Math.* 218(3):889–977, ISSN 0020-9910, URL http://dx.doi.org/10.1007/s00222-019-00898-x.

[11] Carrillo JA, Huang Y (2017) Explicit equilibrium solutions for the aggregation equation with power-law potentials. *Kinet. Relat. Models* 10:171–192, URL http://dx.doi.org/10.3934/krm.2017007.

[12] Choksi R, Fetecau RC, Topaloglu I (2015) On minimizers of interaction functionals with competing attractive and repulsive potentials. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 32:1283–1305, URL http://dx.doi.org/10.1016/j.anihpc.2014.09.004.

[13] Cohn H, Elkies N (2003) New upper bounds on sphere packings. I. *Ann. of Math. (2)* 157:689–714, URL http://dx.doi.org/10.4007/annals.2003.157.689.

[14] Danzer L, Grünbaum B, Klee V (1963) Helly's theorem and its relatives. *Proc. Sympos. Pure Math., Vol. VII*, 101–180 (Amer. Math. Soc., Providence, R.I.).

[15] Federer H (1969) *Geometric Measure Theory* (New York: Springer-Verlag).

[16] Fellner K, Raoul G (2010) Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.* 20:2267–2291, URL http://dx.doi.org/10.1142/S0218202510004921.

[17] Frank RL, Lieb EH (2018) A "liquid-solid" phase transition in a simple model for swarming, based on the "no flat-spots" theorem for subharmonic functions. *Indiana Univ. Math. J.* 67:1547–1569, URL http://dx.doi.org/10.1512/iumj.2018.67.7398.

[18] Jensen ST, Styan GPH (1999) Some comments and a bibliography on the Laguerre-Samuelson inequality with extensions and applications in statistics and matrix theory. *Analytic and geometric inequalities and applications*, volume 478 of *Math. Appl.*, 151–181 (Kluwer Acad. Publ., Dordrecht).

[19] Jung H (1901) Über die kleinste Kugel, die eine räumliche Figur einschliesst. *J. Reine Angew. Math.* 123:241–257, ISSN 0075-4102, URL http://dx.doi.org/10.1515/crll.1901.123.241.

[20] Kawohl B (1985) *Rearrangements and convexity of level sets in PDE*, volume 1150 of *Lecture Notes in Mathematics* (Springer-Verlag, Berlin), URL http://dx.doi.org/10.1007/BFb0075060.

[21] Lim T, McCann RJ (2019+) Isodiametry, variance, and regular simplices from particle interactions. Preprint at *arXiv:1907.13593* .

[22] Lopes O (2019) Uniqueness and radial symmetry of minimizers for a nonlocal variational problem. *Comm. Pure Appl. Anal.* 18:2265–2282, URL http://dx.doi.org/10.3934/cpaa.2019102.

[23] McCann RJ (1997) A convexity principle for interacting gases. *Adv. Math.* 128(1):153–179, ISSN 0001-8708, URL http://dx.doi.org/10.1006/aima.1997.1634.

[24] McCann RJ, Guillen N (2013) Five lectures on optimal transportation: geometry, regularity and applications. *Analysis and geometry of metric measure spaces*, volume 56 of *CRM Proc. Lecture Notes*, 145–180 (Amer. Math. Soc., Providence, RI), URL http://dx.doi.org/10.1090/crmp/056/06.

[25] Odlyzko AM, Sloane NJA (1979) New bounds on the number of unit spheres that can touch a unit sphere in n dimensions. *J. Combin. Theory Ser. A* 26:210–214, URL http://dx.doi.org/10.1016/0097-3165(79)90074-8.

[26] Popoviciu T (1935) Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica (Cluj)* 9:129–145.

[27] Pronzato L, Wynn HP, Zhigljavsky AA (2017) Extended generalised variances, with applications. *Bernoulli* 23(4A):2617–2642, ISSN 1350-7265, URL http://dx.doi.org/10.3150/16-BEJ821.

[28] Rockafellar RT (1970) *Convex analysis*. Princeton Mathematical Series, No. 28 (Princeton University Press, Princeton, N.J.).

[29] Schneider R (1993) *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications* (Cambridge University Press, Cambridge), ISBN 0-521-35220-7, URL http://dx.doi.org/10.1017/CBO9780511526282.

[30] Strasser H (1985) *Mathematical theory of statistics*, volume 7 of *De Gruyter Studies in Mathematics* (Walter de Gruyter & Co., Berlin), ISBN 3-11-010258-7, URL http://dx.doi.org/10.1515/9783110850826, statistical experiments and asymptotic decision theory.

[31] Sun H, Uminsky D, Bertozzi AL (2012) Stability and clustering of self-similar solutions of aggregation equations. *J. Math. Phys.* 53(11):115610, 18, ISSN 0022-2488, URL http://dx.doi.org/10.1063/1.4745180.

[32] Viazovska MS (2017) The sphere packing problem in dimension 8. *Ann. of Math. (2)* 185:991–1015, URL http://dx.doi.org/10.4007/annals.2017.185.3.7.