The Annals of Probability
1996, Vol. 24, No. 4, 2172–2178

THE WILLS FUNCTIONAL AND GAUSSIAN PROCESSES[1]

BY RICHARD A. VITALE

University of Connecticut

The Wills functional from the theory of lattice point enumeration can be adapted to produce the following exponential inequality for zero-mean Gaussian processes:

$$E \exp\left[\sup_t \left(X_t - (1/2)\sigma_t^2 \right)\right] \le \exp\left(E \sup_t X_t \right).$$

An application is a new proof of the deviation inequality for the supremum of a Gaussian process above its mean:

$$P\left(\sup_t X_t - E \sup_t X_t \ge a \right) \le \exp\left(- \frac{(1/2)\alpha^2}{\sigma^2} \right),$$

where $a > 0$ and $\sigma^2 = \sup_t \sigma_t^2$.

1. Introduction. The use of geometric methods in the study of Gaussian processes is by now well established. Our purpose here is to identify a surprising, and apparently deep, connection in the form of the Wills functional. Originally introduced for bounding lattice point enumeration (see [23]), the Wills functional is built up from the classical quermassintegrals ("projection–measure integrals") of Minkowski, which have effectively been applied before to Gaussian processes under the name mixed volumes (or mixed widths) (e.g., [1], [3], [4], [12], [14], [16], [17], [18], [19] and [21]). Placed in our setting, the Wills functional leads naturally to an exponential moment inequality for Gaussian processes and, as a corollary, a deviation inequality for the supremum of a Gaussian process above its mean. The latter is sharp in the sense of having the best possible constant in the exponent.

In the next section, we discuss two representations for the Wills functional. Then we turn to the exponential inequality and the deviation inequality. Section 4 carries some finite-dimensional complements, leading to a second proof of the deviation inequality. We conclude with some remarks in the last section.

For Gaussian processes and bounds in particular, see [6] and [9].

Received October 1995; revised February 1996.

[1]This work was supported in part by ONR Grant N00014-90-J-1641.

AMS 1991 *subject classifications*. Primary 60G15; secondary 52A20, 60G17.

Key words and phrases. Alexandrov–Fenchel inequality, Gaussian process, deviation inequality, exponential bound, intrinsic volume, mixed volume, quermassintegral, tail bound, Wills functional.

2. The Wills functional. Suppose that K is a convex body in $\mathbb{R}^d$ (compact, convex subset) and that $\delta(x, K)$ is the distance between $x \in \mathbb{R}^d$ and K. The Wills functional can be expressed as

$$W(K) = \int_{\mathbb{R}^d} \exp\left[-\pi\delta^2(x, K)\right] dx, \qquad dx = \text{Lebesgue measure.} \tag{1}$$

It was observed by Hadwiger [8] that (1) coincides with the original definition of Wills [23] in terms of quermassintegrals:

$$W(K) = \sum_{j=0}^{d} \binom{d}{j} \frac{1}{\omega_j} W_j(K). \tag{2}$$

Here $\omega_j = \pi^{j/2}/\Gamma(j/2 + 1)$ is the volume of the unit ball B_j in $\mathbb{R}^j$, and the jth quermassintegral $W_j(K)$ is equal to $(\omega_d/\omega_{d-j})E\ \mathrm{vol}_{d-j}(\Pi_{d-j}K)$, where the expectation is of the $(d-j)$-volume of the projection of K onto a random $(d-j)$-dimensional subspace. It is of interest to consider

$$V_j(K) = \binom{d}{j} \frac{1}{\omega_{d-j}} W_{d-j}(K), \qquad 0 \le j \le d,$$

which are normalized versions of the quermassintegrals that do not depend on the dimension of the ambient space $\mathbb{R}^d$ (e.g., [3] and [10]). Following McMullen [10], they are known in the geometry literature as the *intrinsic volumes* of K [in the probability literature, they have been written $h_j(K)$].

Here is a sketch of the equivalence of (1) and (2) which has a probabilistic flavor. We start with the classical Steiner formula for the volume of a parallel body (which itself can be established using a stochastic argument, [22]; for the general theory, see [15]):

$$\mathrm{vol}(K + \Lambda B_d) = \sum_{j=0}^{d} \binom{d}{j} \Lambda^j W_j(K),$$

where $\Lambda \ge 0$, or, equivalently,

$$\int_{\mathbb{R}^d} 1(\delta(x, K) \le \Lambda)\, dx = \sum_{j=0}^{d} \binom{d}{j} \Lambda^j W_j(K). \tag{3}$$

We then regard Λ as a random variable with density

$$f(\lambda) = 1(\lambda \ge 0) 2\pi\lambda \exp(-\pi\lambda^2) \tag{4}$$

and take expectations on both sides of (3) with the use of Fubini's theorem and the moments $E\Lambda^j = \omega_j^{-1}$, $j = 1, 2, \dots$.

We also recall the following bound (see [11]), which is a consequence of the deep Alexandrov–Fenchel inequality (see [15]):

$$W(K) \le \exp V_1(K). \tag{5}$$

3. Bounds. Our main result is as follows.

THEOREM 1. *Suppose that* $\{X_t,\ t \in T\}$ *is a bounded, zero-mean Gaussian process. Then*

$$E \exp\left[\sup_t \{X_t - (1/2)\sigma_t^2\}\right] \le \exp\left[E \sup_t X_t\right], \tag{6}$$

where $\sigma_t^2 = EX_t^2$.

NOTE. Since the process is bounded, it is continuous in probability with respect to the pseudometric $d_X(t_1, t_2) = \sqrt{E(X_{t_1} - X_{t_2})^2}$. We regard $\sup_t X_t$ as over a countable, dense (under d_X) subset of T. Any other countable, dense subset of T gives the same value for $\sup_t X_t$ almost surely.

PROOF OF THEOREM 1. It is enough to show (6) for $T = \{1, 2, \dots, n\}$, a finite set, since the general case follows by approximation. By the Gram–Schmidt procedure, there is a collection $Z_1, Z_2, \dots, Z_d$, $d \le n$, of independent, standard Gaussian variables, such that, for $1 \le k \le n$ and appropriate vectors $a_1, a_2, \dots, a_n \in \mathbb{R}^d$,

$$X_k = \langle a_k, Z \rangle, \tag{7}$$

where $Z = (Z_1, Z_2, \dots, Z_d)^T \in \mathbb{R}^d$ and $\langle \cdot, \cdot \rangle$ signifies inner product. Note that $EX_k^2 = \|a_k\|^2$ and $EX_k X_l = \langle a_k, a_l \rangle$. Let $K \subset \mathbb{R}^d$ be the convex hull of $A = \{a_k/\sqrt{2\pi}\}_1^n$. Employing the natural definition of $W(A)$ (A nonconvex) and making a change of variables at one point, we have

$$\begin{aligned}
W(A) &= \int_{\mathbb{R}^d} \exp\left[-\pi\delta^2(x, A)\right] dx \\
&= \int_{\mathbb{R}^d} \exp\left[-\pi \inf_k \left\| x - a_k/\sqrt{2\pi} \right\|^2\right] dx \\
&= \int_{\mathbb{R}^d} \exp\left[\sup_k \left(\sqrt{2\pi}\langle a_k, x\rangle - (1/2)\|a_k\|^2\right)\right] \exp\left[-\pi\|x\|^2\right] dx \\
&= E \exp\left[\sup_k \left(\langle a_k, Z\rangle - (1/2)\|a_k\|^2\right)\right] \\
&= E \exp\left[\sup_k \left(X_k - (1/2)\sigma_k^2\right)\right].
\end{aligned}$$

Now $A \subseteq K$ implies $W(A) \le W(K)$, and, together with (5), we only have to recall that $V_1(K) = E \sup_k X_k$ [e.g., [16], Proposition 14, where it is written $h_1(K)$]. □

As a consequence, we have the sharp deviation inequality.

COROLLARY 1. *Under the conditions of Theorem* 1, *for any* $a > 0$,

$$P\left(\sup_t X_t - E \sup_t X_t \ge a\right) \le \exp\left[-(1/2)(a^2/\sigma^2)\right], \tag{8}$$

where $\sigma^2 = \sup_t \sigma_t^2$.

The literature contains a variety of applications and results related to (8) (see, e.g., [6] and [9]), often associated with Maurey and Pisier (13). See also the antecedent [4].

PROOF. We use the inhomogeneity of (6) in a variational argument. Consider the process $\{rX_t\}$ for fixed $r > 0$. Then (6) provides

$$E\exp\left[\sup_t \left(rX_t - (1/2)r^2\sigma_t^2\right)\right] \le \exp\left[E\sup_t rX_t\right].$$

Since $\sigma_t^2 \le \sigma^2$ for all t, some rearrangement gives

$$E\exp\left[r\sup_t\left(X_t - E\sup_t X_t\right)\right] \le \exp\left[(1/2)r^2\sigma^2\right].$$

Using Markov's inequality, we have

$$\begin{aligned} P\left(\sup_t X_t - E\sup_t X_t \ge a\right) &= P\left(r\left[\sup_t X_t - E\sup_t X_t\right] \ge ra\right) \\ &= P\left(\exp\left[\sup_t rX_t - E\sup_t rX_t\right] \ge \exp(ra)\right) \\ &\le \exp\left[(1/2)r^2\sigma^2 - ra\right]. \end{aligned}$$

Minimizing the last expression over r then yields $\exp[-a^2(2\sigma^2)^{-1}]$ at $r = a/\sigma^2$. $\square$

4. Finite-dimensional bounds. In general, bounds which are independent of dimension are the most useful. Still, we elaborate a dimensional one here since it leads to a second proof of the deviation inequality. In addition, it illustrates a different use of the Wills functional, here allied with Urysohn's inequality (e.g., [2]; in a probabilistic format, [20]), which we state as follows.

PROPOSITION 1. *Suppose that K is a convex body in $\mathbb{R}^d$. Then*

$$\text{vol}(K) \le \omega_d\left(\frac{V_1(K)}{V_1(B_d)}\right)^d. \tag{9}$$

As before, B_d is the unit ball in $\mathbb{R}^d$ with volume ω_d. There is equality in (9) *if and only if K is a ball.*

The bound is as follows.

THEOREM 2. *Suppose that $\{X_t, t \in T\}$ is a bounded Gaussian process that can be identified with $A \subset \mathbb{R}^d$ via* (7). *That is, for each t, $X_t = \langle a_t, Z\rangle$ for some $a_t \in \sqrt{2\pi}A$. Then*

$$E\exp\left[\sup_t\left\{X_t - (1/2)\sigma_t^2\right\}\right] \le \omega_d E\left(\frac{E\sup_t X_t}{V_1(B_d)} + \Lambda\right)^d. \tag{10}$$

Here Λ *has the density* (4), *and there is equality if and only if the closed convex hull of* A *is a ball.*

PROOF. Let K be the closed convex hull of A. By Urysohn's inequality.

$$\operatorname{vol}(K + \Lambda B_d) \le \omega_d \left(\frac{V_1(K)}{V_1(B_d)} + \Lambda \right)^d .$$

We then take expectations with respect to Λ and express the left-hand side as in the proof of Theorem 1. □

COROLLARY 2. *For each* $a > 0$ *and* $c \ge 1$,

$$P\left[\sup_t X_t - E \sup_t X_t \ge a \right] \le c^d \exp\left[-\frac{1}{2\sigma^2} \left\{ a + \left(1 - \frac{1}{c} \right) E \sup_t X_t \right\}^2 \right]. \tag{11}$$

For the proof, we use the following estimate. It bounds a polynomial with an exponential, but that is sharp enough for our purpose.

LEMMA 1.

$$\psi_d(\theta) = \omega_d E \left(\frac{\theta}{V_1(B_d)} + \Lambda \right)^d \le c^d e^{\theta/c}$$

for all $\theta \ge 0$, $c \ge 1$ *and* $d = 1, 2, \dots$.

PROOF. Fix $c \ge 1$. We verify the property for each $\psi_d(\cdot)$ by induction on d. For $d = 1$, $\psi_1(\theta) = 1 + \theta \le ce^{\theta/c}$. Given that the claim has been shown for ψ_{d-1} and noting that $\psi_d(0) \le ce^{0/c} = c$, it is enough to compare derivatives. Using the facts that $V_1(B_{d-1}) \le V_1(B_d)$ and $\omega_d d = V_1(B_d)\omega_{d-1}$, we have

$$\begin{aligned} \psi_d'(\theta) &= \frac{\omega_d d}{V_1(B_d)} E \left(\frac{\theta}{V_1(B_d)} + \Lambda \right)^{d-1} \\ &\le \omega_{d-1} E \left(\frac{\theta}{V_1(B_{d-1})} + \Lambda \right)^{d-1} \\ &\le c^{d-1} e^{\theta/c} = \frac{d}{d\theta} c^d e^{\theta/c}, \end{aligned}$$

which completes the proof. □

PROOF OF COROLLARY 2. Consider the process $\{rX_t\}$. Using the lemma and a variational argument similar to that for Corollary 1 leads to the assertion. □

Setting $c = 1$ gives a second proof of the deviation inequality (Corollary 1). Finally, we can tailor (11) a bit as follows: let $c = 1 + 1/d$ and using the (possibly more accessible) quantity $\operatorname{diam}(X) = \sup_{t,t'} d_X(t,t') \le \sqrt{2\pi} E \sup_t X_t$ leads to the variant

$$P\Big[\sup_t X_t - E\sup_t X_t \ge a\Big] \le \exp\left[1 - \frac{1}{2\sigma^2}\left\{a + \left(\frac{1}{d+1}\right)\frac{1}{\sqrt{2\pi}}\operatorname{diam}(X)\right\}^2\right].$$

5. Remarks.

REMARK 1. The connection we have drawn between Gaussian processes and the Wills functional, which can be thought of as a measure of size for convex bodies, is in the spirit of [5], which has motivated much work. For further references on the Wills functional, see the survey [7].

REMARK 2. Derived independently, the bound (6) is formally equivalent to Corollary 1 of Tsirel'son [18], which is proved by other means. The context there is different, an infinite-dimensional maximum likelihood problem, and it is presented with no reference to the bounds for Gaussian processes. In any case, the interested reader should consult this important paper together with the others in its series, [17] and [19].

REMARK 3. The corresponding left-tail bound to (8),

$$P\Big(\sup_t X_t - E\sup_t X_t \le a\Big) \le \exp\left(L - \frac{(1/2)a^2}{\sigma^2}\right),$$

does not seem to be accessible by the approach described here.

Acknowledgments. I am grateful to P. McMullen and J. M. Wills for an introduction to the Wills functional and to M. Talagrand for the reference to [4]. The referee made several helpful comments.

REFERENCES

[1] BADRIKIAN, A. and CHEVET, S. (1974). Mesures cylindriques, espaces de Wiener et fonctions aléatoires gaussiennes. *Lecture Notes in Math.* **379**. Springer, New York.

[2] BURAGO, YU. D. and ZALGALLER, V. A. (1988). *Geometric Inequalities*. Springer, New York.

[3] CHEVET, S. (1976). Processus Gaussiens et volumes mixtes. *Z. Wahrsch. Verw. Gebiete* **36** 47–65.

[4] CIREL'SON [TSIREL'SON], B. S., IBRAGIMOV, I. A. and SUDAKOV, V. N. (1976). Norms of Gaussian sample functions. *Proceedings of the Third Japan–USSR Symposium on Probability. Lecture Notes in Math.* **550** 20–41. Springer, New York.

[5] DUDLEY R. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330.

[6] FERNIQUE, X. (1995). Fonctions aléatoires gaussiennes, vecteurs aléatoires gaussiens (troisième version). Unpublished manuscript.

[7] GRITZMANN, P. and WILLS, J. M. (1993). Lattice points. In *Handbook of Convex Geometry* (P. M. Gruber and J. M. Wills, eds.) **B** 765–797. North-Holland, Amsterdam.

[8] HADWIGER H. (1975). Das Wills'sche Funktional. *Monatsh. Math.* **79** 213–221.

[9] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, New York.

[10] McMullen, P. (1975). Non-linear angle–sum relations for polyhedral cones and polytopes. *Math. Proc. Cambridge Philos.* **78** 247–260.
[11] McMullen, P. (1991). Inequalities between intrinsic volumes. *Monatsh. Math.* **111** 47–53.
[12] Milman, V. D. and Pisier, G. (1987). Gaussian processes and mixed volumes. *Ann. Probab.* **15** 292–304.
[13] Pisier, G. (1986). Probabilistic methods in the geometry of Banach spaces. *Probability and Analysis. Lecture Notes in Math.* **1206** 167–241. Springer, New York.
[14] Pisier, G. (1989). *The Volume of Convex Bodies and Banach Space Geometry*. Springer, New York.
[15] Schneider, R. (1993). *Convex Bodies: The Brunn–Minkowski Theory*. Cambridge Univ. Press.
[16] Sudakov, V. N. (1979). *Geometric Problems in the Theory of Infinite-Dimensional Probability Distributions. Proceedings of the Steklov Institute of Mathematics* No. 2. Amer. Math. Soc., Providence, RI.
[17] Tsirel'son, B. S. (1982). A geometric approach to maximum likelihood estimation for infinite-dimensional Gaussian location. *Theory Probab. Appl.* **27** 411–418.
[18] Tsirel'son, B. S. (1985). A geometric approach to maximum likelihood estimation for infinite-dimensional Gaussian location. II. *Theory Probab. Appl.* **30** 820–828.
[19] Tsirel'son, B. S. (1986). A geometric approach to maximum likelihood estimation for infinite-dimensional location. III. *Theory Probab. Appl.* **31** 470–483.
[20] Vitale, R. A. (1988). An alternate formulation of mean value for random geometric figures. *J. Microscopy* **151** 197–204.
[21] Vitale, R. A. (1993). A class of bounds for convex bodies in Hilbert space. *Set-Valued Anal.* **1** 89–96.
[22] Vitale, R. A. (1995). On the volume of parallel bodies: a probabilistic derivation of the Steiner formula. *Adv. in Appl. Probab.* **27** 97–101.
[23] Wills, J. M. (1973). Zur Gitterpunktanzahl konvexer Mengen. *Elem. Math.* **28** 57–63.

Department of Statistics
Box U-120
University of Connecticut
Storrs, Connecticut 06269
E-mail: rvitale@uconnum.uconn.edu