An Inequality for Gaussians on Lattices

Oded Regev∗†  Noah Stephens-Davidowitz∗
noahsd@cs.nyu.edu

Abstract
We show that for any lattice $L \subseteq \mathbb{R}^n$ and vectors $x, y \in \mathbb{R}^n$,
$$\rho(L + x)^2 \rho(L + y)^2 \leq \rho(L)^2 \rho(L + x + y) \rho(L + x - y),$$
where $\rho$ is the Gaussian measure $\rho(A) = \sum_{w \in A} e^{-\pi \|w\|^2}$. We show a number of applications, including bounds on the moments of the discrete Gaussian distribution, various monotonicity properties, and a positive correlation inequality for Gaussian measures on lattices.

1 Introduction
A lattice $L \subset \mathbb{R}^n$ is the set of all integer linear combinations of $n$ linearly independent vectors $B = (b_1, \ldots, b_n)$. For a lattice $L \subset \mathbb{R}^n$, the dual lattice, denoted $L^*$, is defined as the set of all vectors that have integer inner products with all lattice points,
$$L^* = \{w \in \mathbb{R}^n : \forall y \in L, \langle w, y \rangle \in \mathbb{Z} \}.$$
It is easy to show that $L^*$ is itself a lattice.

For any $s > 0$, we define the function $\rho_s : \mathbb{R}^n \to \mathbb{R}$ as $\rho_s(x) = \exp(-\pi \|x\|^2 / s^2)$. For a discrete set $A \subset \mathbb{R}^n$ we define $\rho_s(A) = \sum_{w \in A} \rho_s(w)$. The discrete Gaussian distribution over $L + x$ with parameter $s$, $D_{L + x,s}$, is the probability distribution over $L + x$ that assigns probability proportional to $\rho_s(w)$ to each vector $w \in L + x$. (See Figure 1a) The periodic Gaussian function over $L$ with parameter $s$ is
$$f_{L,s}(x) := \rho_s(L + x) / \rho_s(L).$$
(See Figure 1b) When $s = 1$, we write $\rho(x), D_L$, and $f_L(x)$. Using the Poisson summation formula, one can write
$$f_L(x) = \mathbb{E}_{w \sim D_{L^*}} \left[ \cos(2\pi \langle w, x \rangle) \right].$$

These objects appear in several guises in mathematics and are well studied. For example, $\rho(L + x)$ is the Riemann theta function in a dual form (see, e.g., [Mum07]) and was studied in connection with the Riemann zeta function [Rie57, BPY01]; it can also be seen as the heat kernel

∗Courant Institute of Mathematical Sciences, New York University.
†Supported by the Simons Collaboration on Algorithms and Geometry and by the National Science Foundation (NSF) under Grant No. CCF-1320188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.
in the flat torus $\mathbb{R}^n/L$; it played an instrumental role in proving tight transference theorems for lattices by Banaszczyk [Ban93]; and it was used to construct bilipschitz embeddings of flat tori into Hilbert space [HR13]. Both $D_L+x$ and $f_L$ have also played an important role in recent years in computer science, especially in cryptographic applications of lattices (e.g., [MR07, GPV08]). Our motivation comes from attempts to improve upon the current fastest known algorithm for the main computational problem on lattices, the Shortest Vector Problem [ADRS14].

2 The main inequality

The following is our main theorem. The proof is essentially a combination of a certain identity related to Riemann’s theta relations (see [Mum07, Chapter 1, Section 5]) and the Cauchy-Schwarz inequality.

**Theorem 2.1.** For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and any two vectors $x, y \in \mathbb{R}^n$, we have

$$\rho(\mathcal{L}+x)^2 \rho(\mathcal{L}+y)^2 \leq \rho(\mathcal{L})^2 \rho(\mathcal{L}+x+y) \rho(\mathcal{L}+x-y).$$

**Proof.** Let $\mathcal{L}^{\oplus 2} := \mathcal{L} \oplus \mathcal{L}$. We can then write $\rho(\mathcal{L}+x)\rho(\mathcal{L}+y) = \rho(\mathcal{L}^{\oplus 2} + (x,y))$. Consider the $2n \times 2n$ matrix

$$T := \begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix},$$

where $I_n$ is the $n \times n$ identity matrix. Note that $T/\sqrt{2}$ is an orthogonal matrix so that $\|Tv\| = \sqrt{2}\|v\|$ for any $v \in \mathbb{R}^{2n}$. We therefore have

$$\rho(\mathcal{L}+x)\rho(\mathcal{L}+y) = \rho_{\sqrt{2}}(T(\mathcal{L}^{\oplus 2} + (x,y))) = \rho_{\sqrt{2}}(T\mathcal{L}^{\oplus 2} + (x+y,x-y)).$$

(2)

For any $z := (z_1, z_2) \in \mathcal{L}^{\oplus 2}$, we have $Tz = (w_1, w_2)$ where $w_1 = z_1 + z_2$ and $w_2 = w_1 - 2z_2$. It
follows that
\[ TL^2 = \{(w_1, w_2) \in L^2 : w_1 \equiv w_2 \mod 2L\} \]
\[ = \bigcup_{c \in L/(2L)} (2L + c)^2, \]
where the union is disjoint. Plugging in to Eq. (2), we have
\[ \rho(\mathcal{L} + x)\rho(\mathcal{L} + y) = \sum_{c \in L/(2L)} \rho_{\sqrt{2}}(2L + c + x + y) \cdot \rho_{\sqrt{2}}(2L + c + x - y). \] (3)

Call this \( g(x, y) \). Note that, by the right-hand side of (3), we can view \( g(x, y) \) as the inner product of two vectors,
\[ g(x, y) = \langle h(x + y), h(x - y) \rangle. \]
Then, by Cauchy-Schwarz, we have
\[ g(x, y)^2 \leq \|h(x + y)\|^2\|h(x - y)\|^2 = g(x + y, 0)g(x - y, 0), \]
as needed.

We remark that using the same proof with other transformations \( T \) might lead to other such inequalities. We leave this for future work and proceed to list a few immediate corollaries of Theorem 2.1.

**Corollary 2.2.** For any lattice \( L \subset \mathbb{R}^n \) and any two vectors \( x, y \in \mathbb{R}^n \), we have

\[ f_L(x)^2f_L(y)^2 \leq f_L(x + y)f_L(x - y) \] (4a)
\[ f_L(x)^4 \leq f_L(2x) \] (4b)
\[ f_L(x)f_L(y) \leq (f_L(x + y) + f_L(x - y))/2 \] (4c)
\[ \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, x \rangle)]^2 \leq \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, y \rangle)]^2 \] (4d)
\[ \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, x \rangle)] \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, y \rangle)] \leq \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, x \rangle) \cos(2\pi \langle w, y \rangle)]. \] (4e)

**Proof.** Eq. (4a) follows from the definition of \( f_L \). Eq. (4b) follows from plugging in \( y = x \) to Eq. (4a). Eq. (4c) follows from the fact that \( \sqrt{ab} \leq (a + b)/2 \) for all \( a, b \geq 0 \). Eq. (4d) follows from writing \( f_L(x) \) in its dual form (Eq. (1)) in Eq. (4a) and then applying the identity \( \cos(a + b) = \cos(a) \cos(b) - \sin(a) \sin(b) \). Finally, Eq. (4e) follows from applying the same analysis to (4c).

## 3 Moments of the discrete Gaussian distribution

We next show an inequality on the Hessians of \( f_L \). In particular, this inequality constrains the shape of the local maxima of \( f_L \). (As observed in [DRST14], \( f_L \) can in fact have local maxima at non-lattice points.)
Proposition 3.1. For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and any vector $x \in \mathbb{R}^n$, we have the positive semidefinite inequality
\[
\frac{H f_\mathcal{L}(x)}{f_\mathcal{L}(x)} \geq H f_\mathcal{L}(0) + \frac{\nabla f_\mathcal{L}(x)(\nabla f_\mathcal{L}(x))^T}{f_\mathcal{L}(x)^2}.
\]

Proof. By Eq. (4a), we have
\[
f_\mathcal{L}(x + y) f_\mathcal{L}(x) - f_\mathcal{L}(x)^2 f_\mathcal{L}(y)^2 \geq 0.
\]
Note that this is tight when $y = 0$. It follows that, for any $x$, the left-hand side has a local minimum at $y = 0$, and therefore the Hessian with respect to $y$ at 0 must be positive semidefinite. The result follows by taking the Hessian and rearranging.

As a corollary, we obtain that the covariance matrix of $D_{\mathcal{L}+x}$ is minimized at $x = 0$.

Corollary 3.2. For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and vector $x \in \mathbb{R}^n$, we have the positive semidefinite inequality
\[
\mathbb{E}_{w \sim D_{\mathcal{L}+x}} [ww^T] - \mathbb{E}_{w \sim D_{\mathcal{L}+x}} [w] \mathbb{E}_{w \sim D_{\mathcal{L}+x}} [w^T] \succeq \mathbb{E}_{w \sim D_{\mathcal{L}}} [ww^T] - \mathbb{E}_{w \sim D_{\mathcal{L}}} [w] \mathbb{E}_{w \sim D_{\mathcal{L}}} [w^T].
\]

In particular,
\[
\mathbb{E}_{w \sim D_{\mathcal{L}+x}} [\|w\|^2] - \mathbb{E}_{w \sim D_{\mathcal{L}+x}} [\|w\|^2] \succeq \mathbb{E}_{w \sim D_{\mathcal{L}}} [\|w\|^2].
\]

Proof. A straightforward computation shows that
\[
\frac{\nabla f_\mathcal{L}(x)}{f_\mathcal{L}(x)} = -2\pi \mathbb{E}_{w \sim D_{\mathcal{L}+x}} [w],
\]
\[
\frac{H f_\mathcal{L}(x)}{f_\mathcal{L}(x)} = 4\pi^2 \mathbb{E}_{w \sim D_{\mathcal{L}+x}} [ww^T] - 2\pi I_n.
\]
The result then follows from Proposition 3.1.

The following proposition (with $x = y$) implies that the one-dimensional projections of the discrete Gaussian distribution are “leptokurtic,” i.e., have kurtosis greater than 3, the kurtosis of a normal variable. We remark that the case $n = 1$ follows from a known inequality related to the Riemann zeta function $\zeta(3)$ (see also [BPY01, Section 2.2]).

Proposition 3.3. For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and vectors $x, y \in \mathbb{R}^n$,
\[
\mathbb{E}_{w \sim D_{\mathcal{L}}} [(w, x)^2 (w, y)^2] \geq \mathbb{E}_{w \sim D_{\mathcal{L}}} [(w, x)^2] \mathbb{E}_{w \sim D_{\mathcal{L}}} [(w, y)^2] + 2 \mathbb{E}_{w \sim D_{\mathcal{L}}} [(w, x) (w, y)]^2.
\]

Proof. From Eq. (4d), for any $u, v \in \mathbb{R}^n$, we have
\[
\mathbb{E}_{w \sim D_{\mathcal{L}}} [\cos(2\pi (w, u))]^2 \mathbb{E}_{w \sim D_{\mathcal{L}}} [\cos(2\pi (w, v))]^2 \leq \mathbb{E}_{w \sim D_{\mathcal{L}}} [\cos(2\pi (w, u)) \cos(2\pi (w, v))]^2 - \mathbb{E}_{w \sim D_{\mathcal{L}}} [\sin(2\pi (w, u)) \sin(2\pi (w, v))]^2.
\]
As in the proof of Proposition 3.1, we note that this inequality is tight when \( u = 0 \). So, by taking the second derivative in the \( x \) direction, we have

\[
\mathbb{E}_{w \sim D_L} [(\langle w, x \rangle)^2] \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, v \rangle)]^2 \geq \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle)^2 \cos(2\pi \langle w, v \rangle)] \mathbb{E}_{w \sim D_L} [\cos(2\pi \langle w, v \rangle)] + \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle \sin(2\pi \langle w, v \rangle))^2].
\]

This new inequality is tight when \( v = 0 \). So, by taking the derivative twice at \( v = 0 \) in the \( y \) direction, we have

\[
2 \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle)^2] \mathbb{E}_{w \sim D_L} [(\langle w, y \rangle)^2] \leq \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle)^2 (\langle w, y \rangle)^2] + \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle)^2] \mathbb{E}_{w \sim D_L} [(\langle w, y \rangle)^2] - 2 \mathbb{E}_{w \sim D_L} [(\langle w, x \rangle \langle w, y \rangle)^2],
\]

and the result follows. \( \Box \)

## 4 Monotonicity of the periodic Gaussian function

**Proposition 4.1.** For any lattice \( L \subset \mathbb{R}^n \) and vector \( x \in \mathbb{R}^n \),

\[
\frac{d}{ds} f_{L,s}(x) = \frac{2\pi f_{L,s}(x)}{s^3} \cdot \mathbb{E}_{w \sim D_{L \times s}} [||w||^2] - \frac{2\pi f_{L,s}(x)}{s^3} \cdot \mathbb{E}_{w \sim D_{L,s}} [||w||^2].
\]

In particular, \( f_{L,s}(x) \) is non-decreasing as a function of \( s \).

**Proof.** A straightforward computation shows that

\[
\frac{d}{ds} f_{L,s}(x) = \frac{2\pi f_{L,s}(x)}{s^3} \cdot \mathbb{E}_{w \sim D_{L \times s}} [||w||^2] - \frac{2\pi f_{L,s}(x)}{s^3} \cdot \mathbb{E}_{w \sim D_{L,s}} [||w||^2]
\]

\[
\geq \frac{2\pi f_{L,s}(x)}{s^3} \cdot \mathbb{E}_{w \sim D_{L \times s}} [||w||^2]^2,
\]

where we have applied Corollary 3.2. The result then follows from the fact that

\[
\frac{\|\nabla f_{L,s}(x)\|}{f_{L,s}(x)} = \frac{2\pi}{s^2} \cdot \mathbb{E}_{w \sim D_{L \times s}} [||w||^2]. \quad \Box
\]

The next proposition can be viewed as a generalization of Proposition 4.1 in which we replace the scalar parameter \( s > 0 \) by a positive definite matrix \( S \).

**Proposition 4.2.** For any lattice \( L \subset \mathbb{R}^n \), positive definite matrix \( S \in \mathbb{R}^{n \times n} \), and \( x \in \mathbb{R}^n \), let \( f_{L,S}(x) := f_{S^{-1}L}(S^{-1}x) \). Then, for any positive definite matrix \( S' \) satisfying the positive semidefinite inequality \( S' \preceq S \),

\[
f_{L,S'}(x) \leq f_{L,S}(x).
\]

This answers a question of Price [Pri14b], who proved monotonicity for the case \( n = 1 \). He also asked if there are other manifolds for which such a monotonicity property holds, a question that we leave open.
Proof. By replacing $L$ by $S^{-1}L$, $x$ by $S^{-1}x$, and $S$ by $S^{-1}S$, we can assume without loss of generality that $S' = I_n$. Because $S$ is symmetric, there exists an orthonormal basis that diagonalizes $S$. Let $s_1, \ldots, s_n \geq 1$ be the entries along the diagonal of $S$ in this basis (i.e., the eigenvalues of $S$).

The proof now proceeds nearly identically to the proof of Proposition 4.1. Differentiating with respect to $s_i$, we have

$$\frac{d}{ds_i} f_{L,S}(x) = \frac{2\pi f_{L,S}(x)}{s_i^3} \left( \mathbb{E}_{w \sim D_{S^{-1}(L+x)}} [w_i^2] - \mathbb{E}_{w \sim D_{S^{-1}L}} [w_i^2] \right),$$

where $w_i$ is the $i$th coordinate of $w$ (in the basis that diagonalizes $S$). The result follows by noting that Corollary 3.2 implies that this derivative is positive for all $s_i > 0$.

Proposition 4.3. For any lattice $L \subset \mathbb{R}^n$, sublattice $M \subseteq L$, and vector $x \in \mathbb{R}^n$,

$$f_M(x) \leq f_L(x).$$

Proof.

$$\rho(M + x) \rho(L) = \sum_{c \in L/M} \rho(M + x) \rho(M + c)$$

$$\leq \sum_{c \in L/M} \rho(M) \rho(M + x + c) / 2 \quad \text{(Eq. (4c))}$$

$$= \sum_{c \in L/M} \rho(M) \rho(M + x + c)$$

$$= \rho(M) \rho(L + x).$$

The result follows.

As a corollary, we answer a question asked by Price [Pri14a].

Corollary 4.4. For any lattice $L \subset \mathbb{R}^n$, sublattice $M \subseteq L$, and subspace $V \subseteq \mathbb{R}^n$, we have

$$\frac{\rho(M)}{\rho(M \cap V)} \leq \frac{\rho(L)}{\rho(L \cap V)}.$$

Proof.

$$\frac{\rho(L)}{\rho(L \cap V)} = \sum_{c \in L/(L \cap V)} \frac{\rho((L \cap V) + c)}{\rho(L \cap V)}$$

$$\geq \sum_{c \in L/(L \cap V)} \frac{\rho((M \cap V) + c)}{\rho(M \cap V)} \quad \text{(Prop. 4.3)}$$

$$\geq \sum_{c \in M/(M \cap V)} \frac{\rho((M \cap V) + c)}{\rho(M \cap V)}$$

$$= \frac{\rho(M)}{\rho(M \cap V)},$$

where the last inequality follows from the fact that for any two distinct cosets $c, c' \in M/(M \cap V)$, $c - c' \notin V$ and therefore $(L \cap V) + c \neq (L \cap V) + c'$. \qed
5 Positive correlation of Gaussian measure on lattices

The following shows that under the normalized Gaussian measure on a lattice, sublattices are positively correlated. This has superficial resemblance to the celebrated Gaussian correlation conjecture (see [SSZ98] and references therein).

**Theorem 5.1.** For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and sublattices $\mathcal{M}$ and $\mathcal{N}$,

$$\frac{\rho(\mathcal{M})}{\rho(\mathcal{L})} \cdot \frac{\rho(\mathcal{N})}{\rho(\mathcal{L})} \leq \frac{\rho(\mathcal{M} \cap \mathcal{N})}{\rho(\mathcal{L})} .$$

**Proof.** We claim that for any $w \in \mathcal{L}$,

$$\sum_{c \in M^*/L^*} \cos(2\pi \langle w, c \rangle) = |M^*/L^*| \cdot 1_M(w). \quad (5)$$

If $w \in \mathcal{M}$ this is obvious, so suppose that $w \in \mathcal{L} \setminus \mathcal{M}$. Then, there exists $c' \in M^*/L^*$ such that $\langle w, c' \rangle \not\in \mathbb{Z}$. Since the cosets form a group, we have

$$\sum_{c \in M^*/L^*} \cos(2\pi \langle w, c \rangle) = \sum_{c \in M^*/L^*} \cos(2\pi \langle w, c + c' \rangle)$$

$$= \cos(2\pi \langle w, c' \rangle) \sum_{c \in M^*/L^*} \cos(2\pi \langle w, c \rangle) - \sin(2\pi \langle w, c' \rangle) \sum_{c \in M^*/L^*} \sin(2\pi \langle w, c \rangle),$$

where we have used the identity $\cos(a + b) = \cos(a) \cos(b) - \sin(a) \sin(b)$. Eq. (5) follows after noting that the second term on the right-hand side must be zero by symmetry. Similarly,

$$\sum_{d \in N^*/L^*} \cos(2\pi \langle w, d \rangle) = |N^*/L^*| \cdot 1_N(w).$$

By (4e), we have that for any $c \in M^*/L^*$ and $d \in N^*/L^*$,

$$\mathbb{E}_{w \sim D_c} \left[ \cos(2\pi \langle w, c \rangle) \right] \mathbb{E}_{w \sim D_c} \left[ \cos(2\pi \langle w, d \rangle) \right] \leq \mathbb{E}_{w \sim D_c} \left[ \cos(2\pi \langle w, c \rangle) \cos(2\pi \langle w, d \rangle) \right].$$

The result then follows by summing both sides over all cosets $c$ and $d$. \qed

**Acknowledgements**

We thank Tom Price for helpful discussion.

**References**

[ADRS14] D. Aggarwal, D. Dadush, O. Regev, and N. Stephens-Davidowitz. Solving the shortest vector problem in $2^n$ time via discrete Gaussian sampling, 2014. Submitted. Available at [http://arxiv.org/abs/1412.7994](http://arxiv.org/abs/1412.7994).

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.
[BPY01] P. Biane, J. Pitman, and M. Yor. Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc. (N.S.)*, 38(4):435–465, 2001.

[Chu76] K. L. Chung. Excursions in Brownian motion. *Ark. Mat.*, 14(2):155–177, 1976.

[DRS14] D. Dadush, O. Regev, and N. Stephens-Davidowitz. On the closest vector problem with a distance guarantee. In *IEEE 29th Conference on Computational Complexity*, pages 98–109. 2014. Full version available at [http://arxiv.org/abs/1409.8063](http://arxiv.org/abs/1409.8063).

[GPV08] C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206. 2008.

[HR13] I. Haviv and O. Regev. The Euclidean distortion of flat tori. *J. Topol. Anal.*, 5(2):205–223, 2013.

[MR07] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302 (electronic), 2007.

[Mum07] D. Mumford. *Tata lectures on theta. I*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007. With the collaboration of C. Musili, M. Nori, E. Previato and M. Stillman, Reprint of the 1983 edition.

[New76] C. M. Newman. Fourier transforms with only real zeros. *Proc. Amer. Math. Soc.*, 61(2):245–251 (1977), 1976.

[Pri14a] T. Price. Inequality regarding sum of gaussian on lattices. MathOverflow, 2014. [http://mathoverflow.net/q/160507](http://mathoverflow.net/q/160507) (version: 2014-12-01).

[Pri14b] T. Price. Is the heat kernel more spread out with a smaller metric? MathOverflow, 2014. [http://mathoverflow.net/q/186428](http://mathoverflow.net/q/186428) (version: 2014-12-11).

[Rie57] B. Riemann. Theorie der Abel’schen Functionen. *Journal für die reine und angewandte Mathematik*, 54:101–155, 1857.

[SSZ98] G. Schechtman, T. Schlumprecht, and J. Zinn. On the Gaussian measure of the intersection. *Ann. Probab.*, 26(1):346–357, 1998.