On the Subgaussianity of Quantized Linear Maps: An AI-Assisted Note

Guangyi Zou[1] and Roman Vershynin[1]

[1]Department of Mathematics, University of California, Irvine
gzou3@uci.edu, rvershyn@uci.edu

August 17, 2026

Abstract

We prove an elementary bounded-differences inequality for functions of non-isotropic Gaussian vectors. Specifically, if f has bounded coordinate differences and $X \sim \mathcal{N}(\mu, \Sigma)$, then the resulting concentration bound depends on the condition number $\kappa(\Sigma)$. As an application, we answer a question of Simone Bombari concerning the subgaussianity of sign-quantized linear maps $Y = \operatorname{sgn}(Wx)$.

In the special case where f is the coordinatewise sign function, an argument was initially suggested to us by Gemini 3.5 Flash without attribution. We subsequently discovered that it closely resembles an earlier argument of Barber and Kolar [BK18, Lemma 4.5]. This revision corrects the attribution and documents the episode as an instance of AI-assisted mathematical discovery.

1 Introduction: An AI-for-Math Perspective

Simone Bombari asked us whether the 1-bit quantized random vector $Y = \operatorname{sgn}(Wx)$ has subgaussian norm bounded by a universal constant. Here W is an $n \times n$ random Gaussian matrix, and x is an independent standard normal random vector in $\mathbb{R}^n$. The question is nontrivial since the coordinates of Y are not independent. We give a positive answer and, more generally, prove a bounded-differences inequality for a non-isotropic Gaussian input.

- **AI (re)discovery (Theorem 2):** During our initial exploration, Gemini 3.5 Flash[1] proposed decomposing the Gaussian vector into independent parts, using one part to smooth the sign function, and then applying Gaussian concentration. We later learned this idea had already appeared in Barber and Kolar [BK18, Lemma 4.5]. Gemini later returned essentially their same argument without citing the source. The present revision corrects this attribution. The argument extends to general bounded-differences functions, as stated below.
- **Human design (Section 3):** To apply this result to square random matrices W, which are typically ill conditioned, the authors used a row-partitioning trick. This breaks the square matrix into two well-conditioned rectangular blocks.

2 Main Result: Gaussian Bounded Differences

The subgaussian norm of a random vector is defined as the largest subgaussian norm of its 1-dimensional marginals:

$$\|Y\|_{\psi_2} = \sup_{v \in S^{n-1}} \|\langle v, Y \rangle\|_{\psi_2},$$

[1]The original Chinese dialogue is archived at https://gemini.google.com/share/daca36d5cfbc; see https://github.com/The8ravo/conversation-history-with-Gemini/blob/main/conversation_history_with_gemini.pdf for an AI-generated English translation with a timeline. The reproduced interactive dialogue and full derivations are archived at https://gemini.google.com/share/cefbb71da6d2.

see [Ver26, Definition 3.4.1]. The condition number of a positive-definite matrix Σ is the ratio of its largest eigenvalue to its smallest eigenvalue:

$$\kappa(\Sigma) := \frac{\lambda_{\max}(\Sigma)}{\lambda_{\min}(\Sigma)}.$$

Definition 1 (Bounded differences)**.** A measurable function $f : \mathbb{R}^n \to \mathbb{R}$ has bounded-differences constants $b_1, \dots, b_n \geq 0$ if, for every $i = 1, \dots, n$, changing only the i-th coordinate changes f by at most b_i:

$$|f(x) - f(x')| \leq b_i \quad \text{whenever} \quad x_j = x'_j \text{ for all } j \neq i.$$

Write $b = (b_1, \dots, b_n)$.

Theorem 2 (Gaussian bounded-differences inequality)**.** *Let $X \sim \mathcal{N}(\mu, \Sigma)$ be an n-dimensional Gaussian vector with non-singular covariance matrix Σ. Let $f : \mathbb{R}^n \to \mathbb{R}$ have bounded-differences constants $b = (b_1, \dots, b_n)$. Then,*

$$\|f(X) - \mathbb{E}f(X)\|_{\psi_2} \leq C\sqrt{\kappa(\Sigma)}\|b\|_2,$$

where $C > 0$ is an absolute constant.

For comparison, if X is any random vector with independent coordinates, the classical bounded-differences inequality gives

$$\|f(X) - \mathbb{E}f(X)\|_{\psi_2} \leq C\|b\|_2,$$

where $C > 0$ is an absolute constant. Thus, Theorem 2 extends this conclusion to dependent Gaussian coordinates at the cost of the factor $\sqrt{\kappa(\Sigma)}$.

Proof. The proof extends the argument of Barber and Kolar [BK18, Lemma 4.5]. Replacing f by $f - f(0)$ does not change either its bounded-differences constants or the centered random variable $f(X) - \mathbb{E}f(X)$. After this replacement, a coordinate-by-coordinate telescoping argument gives $|f(x)| \leq \sum_i b_i$ for every $x \in \mathbb{R}^n$. Thus f is bounded, which justifies the Gaussian convolution and differentiation used below.

Step 1: Covariance splitting. Let $\alpha = \lambda_{\min}(\Sigma)$ and $\beta = \lambda_{\max}(\Sigma)$. Define $\Sigma_G = \Sigma - \alpha I_n \geq 0$. We may write

$$X = \mu + \sqrt{\alpha}\, Z + G,$$

where $Z \sim \mathcal{N}(0, I_n)$ and $G \sim \mathcal{N}(0, \Sigma_G)$ are independent.

Step 2: Conditional bounded-differences estimate. Define the function

$$H(g) = \mathbb{E}_Z f(\mu + \sqrt{\alpha}\, Z + g), \quad g \in \mathbb{R}^n.$$

Conditionally on G, the coordinates of Z are independent, and changing Z_i changes only the i-th coordinate of the input of f. Thus the map $z \mapsto f(\mu + \sqrt{\alpha}\, z + G)$ has bounded-differences constants $b_1, \dots, b_n$. By the classical bounded-differences inequality (see, e.g., [Ver26, Theorem 5.7.1]), conditionally on G,

$$\big\|f(\mu + \sqrt{\alpha}\, Z + G) - H(G)\big\|_{\psi_2 | G} \leq C\|b\|_2,$$

where $C > 0$ is an absolute constant. More precisely, its exponential-moment form gives, for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}\big[\exp\big(\lambda\big(f(\mu + \sqrt{\alpha}\, Z + G) - H(G)\big)\big) \,\big|\, G\big] \leq \exp\left(\frac{\lambda^2\|b\|_2^2}{8}\right). \tag{1}$$

Step 3: Concentration of the conditional mean. Gaussian convolution makes H smooth: Gaussian integration by parts (see, e.g., [Ver26, Lemma 7.2.3]) gives

$$\partial_i H(g) = \frac{1}{\sqrt{\alpha}}\mathbb{E}_Z\big[Z_i f(\mu + \sqrt{\alpha}\, Z + g)\big]. \tag{2}$$

Conditionally on Z_j for all $j \neq i$, the range of $f(\mu + \sqrt{\alpha}\, Z + g)$ as a function of Z_i has length at most b_i. Subtracting the midpoint of this range does not change the expectation in (2). This gives

$$|\partial_i H(g)| \leq \frac{b_i}{2\sqrt{\alpha}} \mathbb{E}|Z_i| = \frac{b_i}{\sqrt{2\pi\alpha}}.$$

Thus $\|\nabla H(g)\|_2 \leq \|b\|_2/\sqrt{2\pi\alpha}$, so H is a $\|b\|_2/\sqrt{2\pi\alpha}$-Lipschitz function.

Let $U \sim \mathcal{N}(0, I_n)$. Since $G \overset{d}{=} \Sigma_G^{1/2} U$, define

$$F(u) = H(\Sigma_G^{1/2} u).$$

The Lipschitz constant of F satisfies

$$\mathrm{Lip}(F) \leq \|\Sigma_G^{1/2}\|_{\mathrm{op}} \,\mathrm{Lip}(H) \leq \sqrt{\|\Sigma_G\|_{\mathrm{op}}}\, \frac{\|b\|_2}{\sqrt{2\pi\alpha}} = \sqrt{\beta - \alpha}\, \frac{\|b\|_2}{\sqrt{2\pi\alpha}}.$$

Since $\beta/\alpha = \kappa(\Sigma)$, it follows that

$$\mathrm{Lip}(F)^2 \leq \frac{\kappa(\Sigma) - 1}{2\pi} \|b\|_2^2.$$

Standard Gaussian concentration applied to $F(U)$ (see, e.g., [Ver26, Theorem 5.2.3] or [BLM13, Theorem 5.5]) therefore gives

$$\begin{aligned} \mathbb{E}\exp(\lambda(H(G) - \mathbb{E}H(G))) &= \mathbb{E}\exp(\lambda(F(U) - \mathbb{E}F(U))) \\ &\leq \exp\left(\frac{\lambda^2 \,\mathrm{Lip}(F)^2}{2}\right) \\ &\leq \exp\left(\frac{\lambda^2 \|b\|_2^2}{4\pi}(\kappa(\Sigma) - 1)\right). \end{aligned} \tag{3}$$

Step 4: Combining the bounds. Conditioning on G and using (1)–(3), we obtain

$$\log \mathbb{E}\exp(\lambda(f(X) - \mathbb{E}f(X))) \leq \lambda^2 \|b\|_2^2 \left(\frac{1}{8} + \frac{\kappa(\Sigma) - 1}{4\pi}\right) \leq \frac{\lambda^2 \kappa(\Sigma) \|b\|_2^2}{8}.$$

By the standard equivalence of subgaussian properties (see, e.g., [Ver26, Proposition 2.6.1]), this implies

$$\|f(X) - \mathbb{E}f(X)\|_{\psi_2} \leq C\sqrt{\kappa(\Sigma)} \|b\|_2,$$

where $C > 0$ is an absolute constant. □

The coordinatewise result from the first version is an immediate consequence.

Corollary 3. *Let $X \sim \mathcal{N}(\mu, \Sigma)$ be an n-dimensional Gaussian vector with non-singular covariance matrix Σ. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a measurable function with $\|\phi\|_\infty \leq 1$, and define $Y = \phi(X)$ coordinatewise. Then*

$$\|Y - \mathbb{E}Y\|_{\psi_2} \leq C\sqrt{\kappa(\Sigma)},$$

where $C > 0$ is an absolute constant.

This follows immediately: for $v \in S^{n-1}$, the function $f_v(x) = \sum_i v_i \phi(x_i)$ has bounded-differences constants $b_i \leq 2|v_i|$, and hence $\|b\|_2 \leq 2$. Theorem 2 and the definition of the subgaussian norm of a random vector give the claim after taking the supremum over v and absorbing the factor 2 into C.

Remark 4 (Extension to Continuous Functions)**.** If ϕ is γ-Hölder continuous instead of globally bounded, the same decomposing trick still works, and the subgaussian norm becomes bounded by

$$O\left(\sqrt{\lambda_{\max}(\Sigma)}\, \lambda_{\min}(\Sigma)^{\frac{\gamma-1}{2}}\right).$$

See the conversation history for the derivation.

Furthermore, the well-conditioning of Σ is necessary for a uniform bound. A natural counterexample is the rank-one matrix $\Sigma = \mathbf{1}\mathbf{1}^\top$, where $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbb{R}^n$, which yields $\|\operatorname{sgn}(X)\|_{\psi_2} \geq c\sqrt{n}$ for an absolute constant $c > 0$.

Remark 5 (Simultaneous work of Roth)**.** Simultaneously and independently of our work, Roth [Rot26] obtained a bounded-differences inequality under dependence by a different method. His result applies to distributions more general than Gaussian distributions, including strongly log-concave and log-smooth measures. Our argument is restricted to Gaussian inputs, but it is elementary and short.

3 Application: Sign-Quantized Square Linear Maps

We now apply Corollary 3 to resolve the question posed by Simone Bombari.

Corollary 6. *Let $W \in \mathbb{R}^{n\times n}$ be a random matrix with i.i.d. standard normal entries $W_{ij} \sim \mathcal{N}(0,1)$. Let $x \sim \mathcal{N}(0, I_n)$ be an n-dimensional standard normal random vector. Define $Y = \operatorname{sgn}(Wx) \in \{-1,1\}^n$ by applying the sign function to each coordinate of Wx. Then*

$$\mathbb{E}\|Y\|_{\psi_2|W} \le C,$$

where $C > 0$ is an absolute constant. Here, the conditional norm is with respect to x for fixed W, and the expectation is with respect to W.

Proof. The case $n = 1$ is immediate, so assume $n \ge 2$. Directly applying Corollary 3 to W fails because $\Sigma = WW^\top$ typically has large condition number.

To fix this, we split W into two row blocks $W_1 \in \mathbb{R}^{m\times n}$ and $W_2 \in \mathbb{R}^{(n-m)\times n}$, where $m = \lfloor n/2 \rfloor$:

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}. \tag{4}$$

This induces the split

$$Y^{(1)} = \operatorname{sgn}(W_1 x), \qquad Y^{(2)} = \operatorname{sgn}(W_2 x).$$

The triangle inequality gives

$$\|Y\|_{\psi_2|W} \le \|Y^{(1)}\|_{\psi_2|W_1} + \|Y^{(2)}\|_{\psi_2|W_2}.$$

Both blocks have at most $2n/3$ rows. The same argument applies to both, so we treat W_1.

Since $m/n \le 1/2$, the rectangular block W_1 is well conditioned with high probability. Specifically, its covariance $\Sigma_1 = W_1 W_1^\top$ satisfies

$$nc_1 I \preceq \Sigma_1 \preceq nc_2 I$$

with probability at least $1 - c_3 e^{-c_4 n}$ for universal positive constants c_1, c_2, c_3, c_4; see [Ver26, Corollary 7.3.2 and Exercise 7.13]. Let $\mathcal{E} = \{\kappa(\Sigma_1) \le c_2/c_1\}$.

Since $W_1 x$ is centered Gaussian, $\mathbb{E}[Y^{(1)} \mid W_1] = 0$. On $\mathcal{E}$, Corollary 3 gives

$$\|Y^{(1)}\|_{\psi_2|W_1} \le C_0 \sqrt{\kappa(\Sigma_1)} \le C_0 \sqrt{\frac{c_2}{c_1}} =: C_1. \tag{5}$$

On $\mathcal{E}^c$, the deterministic bound $|\langle v, Y^{(1)}\rangle| \le \sqrt{m} \le \sqrt{n}$ gives $\|Y^{(1)}\|_{\psi_2|W_1} \le c_5\sqrt{n}$ for an absolute constant $c_5 > 0$. Therefore

$$\mathbb{E}\|Y^{(1)}\|_{\psi_2|W_1} \le C_1 + c_5 c_3 \sqrt{n}\, e^{-c_4 n} \le C_2.$$

Here $C_2 > 0$ is an absolute constant. The same estimate for W_2 and the triangle inequality complete the proof. □

Remark 7 (Role of the AI system)**.** Gemini initially returned the covariance-splitting argument for the sign function and extended it to bounded coordinate maps. We later learned that the sign case and the same proof strategy appeared in Barber and Kolar [BK18, Lemma 4.5]; Gemini did not identify that source. The bounded-differences formulation above follows the same argument. The row-partitioning step in (4) was added by the authors.

Acknowledgments

The authors thank Simone Bombari for posing the question that motivated this work and for pointing out its connection with the bounded-differences inequality. We thank Valentin Roth for suggesting that we state our argument in bounded-differences form and for sharing his simultaneous and independent work on bounded-differences inequalities under dependence. Google Gemini 3.5 Flash was used during the initial exploration, as described above. A reproduced interactive dialogue and derivation transcripts are archived at https://gemini.google.com/share/cefbb71da6d2. The authors are supported by NSF Grant DMS 2451011 and U.S. Air Force Grant FA9550-25-1-0294.

References

[BK18] Rina Foygel Barber and Mladen Kolar. ROCKET: robust confidence intervals via Kendall's tau for transelliptical graphical models. *Ann. Statist.*, 46(6B):3422–3450, 2018.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[Rot26] Valentin Roth. On McDiarmid's inequality under dependence via approximate tensorization of entropy. *arXiv preprint arXiv:2606.12720*, 2026.

[Ver26] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2 edition, 2026.