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

May 26, 2026

Abstract

This short note presents a dimension-independent subgaussian concentration bound for Gaussian vectors under coordinate-wise nonlinear mappings. Discovered by Gemini 3.5 Flash, this result applies to any bounded function under a well-conditioned covariance. We apply this tool to answer a question of Simone Bombari on sign-quantized linear maps $Y = \operatorname{sgn}(Wx)$.

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 strong positive answer to this question – for any bounded map instead of $\operatorname{sgn}(\cdot)$ – using AI:

- **AI Discovery and Generalization (Theorem 1):** To handle coordinate dependence, 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 for Lipschitz functions. It subsequently generalized the result for the sign function to an arbitrary bounded functions. However, this idea works for well-conditioned matrices W only.
- **Human Design (Section 3):** To apply this general tool to square random matrices W – which are typically ill conditioned – the human authors used a simple row-partitioning trick. This breaks the square matrix into two well-conditioned rectangular blocks, bypassing the singularity issue.

2 Main Result: Subgaussianity under Well-conditioned Covariance

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.

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

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

The following theorem, giving dimension-independent subgaussian bound, was proved by Gemini 3.5 Flash.

Theorem 1. *Let $X \sim \mathcal{N}(0, \Sigma)$ be an n-dimensional normal random vector with a nonsingular covariance matrix Σ. Let $\phi : \mathbb{R} \to \mathbb{R}$ satisfy $\|\phi\|_\infty \le 1$. Define the random vector $Y = \phi(X) \in \mathbb{R}^n$ by applying ϕ to each coordinate of X. Then*

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

where $C > 0$ is an absolute constant.

Proof. By definition, we need to show that for any direction $v \in S^{n-1}$, the random variable

$$\langle v, Y - \mathbb{E}Y \rangle = \sum_{i=1}^{n} v_i (Y_i - \mathbb{E}Y_i)$$

has a uniformly bounded subgaussian norm.

Step 1: Covariance Splitting. Let $a = \lambda_{\min}(\Sigma)$ and $b = \lambda_{\max}(\Sigma)$, so that $\kappa(\Sigma) = b/a$. Defining the residual covariance matrix $\Sigma_G \coloneqq \Sigma - aI_n \ge 0$, we can decompose X into the sum of two independent Gaussian vectors:

$$X = \sqrt{a}Z + G, \quad \text{where } Z \sim \mathcal{N}(0, I_n) \text{ and } G \sim \mathcal{N}(0, \Sigma_G).$$

Step 2: Conditional Hoeffding Bound. Conditioned on G, the coordinates $Y_i = \phi(\sqrt{a}Z_i + G_i)$ are mutually independent. Let

$$\mu_i(x) \coloneqq \mathbb{E}\, \phi(\sqrt{a}Z_i + x), \quad x \in \mathbb{R}.$$

The conditional expectation satisfies $\mathbb{E}[Y_i \mid G] = \mu_i(G_i)$. Since $\|\phi\|_\infty \le 1$, each centered coordinate is deterministically bounded: $|v_i(Y_i - \mu_i(G_i))| \le 2|v_i|$. By Hoeffding's inequality (see e.g. [Ver26, Theorem 2.2.6]), a random variable bounded within an interval of length $4|v_i|$ has MGF bounded by $\exp(\lambda^2(4|v_i|)^2/8) = \exp(2v_i^2\lambda^2)$. Summing over the independent coordinates gives:

$$\mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} v_i(Y_i - \mu_i(G_i))\right) \,\middle|\, G\right] \le \exp\left(\sum_{i=1}^{n} 2v_i^2\lambda^2\right) = \exp\left(2\lambda^2\right). \tag{1}$$

Step 3: Concentration of the Conditional Mean. Define the function

$$F(x_1, \ldots, x_n) \coloneqq \sum_{i=1}^{n} v_i \mu_i(x_i).$$

Although ϕ is not assumed to be differentiable, the Gaussian density smooths the expectation via convolution. By writing $\mu_i(x) = \int_{\mathbb{R}} \phi(u) \frac{1}{\sqrt{2\pi a}} e^{-\frac{(u-x)^2}{2a}}\, du$ and differentiating the Gaussian kernel directly with respect to x, a simple change of variables yields:

$$|\mu_i'(x)| = \left| \frac{1}{\sqrt{a}} \mathbb{E}\, Z_i\, \phi(\sqrt{a}Z_i + x) \right| \le \frac{1}{\sqrt{a}} \mathbb{E}|Z_i| = \sqrt{\frac{2}{\pi a}}.$$

Thus, the gradient satisfies $\|\nabla F\|_2 \le \max_i |\mu_i'(x_i)| \cdot \|v\|_2 \le \sqrt{\frac{2}{\pi a}}$, meaning F is L_μ-Lipschitz with $L_\mu = \sqrt{\frac{2}{\pi a}}$. Note also that

$$\mathbb{E}F(G) = \sum_{i=1}^{n} v_i \mathbb{E}Y_i.$$

By standard Gaussian concentration for the Lipschitz function of $G \sim \mathcal{N}(0, \Sigma_G)$ (see e.g. [Ver26, Theorem 5.2.11] or [BLM13, Theorem 5.5]), we have:

$$\mathbb{E}[\exp(\lambda(F(G) - \mathbb{E}F(G)))] \leq \exp\left(\frac{\lambda^2 L_\mu^2 \|\Sigma_G\|_{\mathrm{op}}}{2}\right) \leq \exp\left(\frac{\lambda^2}{\pi a}(b-a)\right) = \exp\left(\frac{\lambda^2}{\pi}(\kappa(\Sigma) - 1)\right). \quad (2)$$

Step 4: Total MGF Bound. We decompose the target variable into conditional fluctuations and the variation of the conditional mean:

$$\langle v, Y - \mathbb{E}Y\rangle = \sum_{i=1}^{n} v_i(Y_i - \mu_i(G_i)) + (F(G) - \mathbb{E}F(G)).$$

By the law of total expectation, we combine (1) and (2):

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda\langle v, Y - \mathbb{E}Y\rangle\right)\right] &= \mathbb{E}_G\left[\mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} v_i(Y_i - \mu_i(G_i))\right) \,\middle|\, G\right] \cdot \exp\left(\lambda(F(G) - \mathbb{E}F(G))\right)\right] \\ &\leq \exp(2\lambda^2) \cdot \exp\left(\frac{\lambda^2}{\pi}(\kappa(\Sigma) - 1)\right) = \exp\left(\frac{\sigma^2\lambda^2}{2}\right),\end{aligned}$$

where

$$\sigma^2 = 4 + \frac{2}{\pi}(\kappa(\Sigma) - 1) \leq 4\kappa(\Sigma),$$

since $\kappa(\Sigma) \geq 1$.

By the standard equivalence of subgaussian properties (see e.g. [Ver26, Proposition 2.6.1]), this implies $\|\langle v, Y - \mathbb{E}Y\rangle\|_{\psi_2} \leq C\sqrt{\kappa(\Sigma)}$, where C is an absolute constant. Taking the supremum over $v \in S^{n-1}$ completes the proof. □

Remark 2 (Extension to Continuous Functions)**.** If ϕ is α-Hölder continuous instead of globally bounded, the same decomposing trick still works, and the subgaussian norm becomes bounded by $O(\sqrt{\lambda_{\max}(\Sigma)}\,\lambda_{\min}(\Sigma)^{\frac{\alpha-1}{2}})$. See the conversation history for Gemini 3.5 Flash's proof.

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

3 Application: Sign-Quantized Square Linear Maps

We now apply Theorem 1 to resolve the question posed by Simone Bombari.

Corollary 3. *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 the random vector $Y = \operatorname{sgn}(Wx) \in \{-1, 1\}^n$ by applying the sign function to each coordinate of x. Then*

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

where $C > 0$ is an absolute constant. Here, the conditional norm $\|Y\|_{\psi_2|W}$ is with respect to the random vector x for a fixed realization of W, and the expectation is with respect to the random matrix W.

Proof. Directly applying Theorem 1 to W fails because the Wishart matrix $\Sigma = WW^\top$ typically has large condition number.

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

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}. \quad (3)$$

This induces a corresponding split on the output vector into

$$Y^{(1)} = \operatorname{sgn}(W_1 x) \in \{-1, 1\}^m \quad \text{and} \quad Y^{(2)} = \operatorname{sgn}(W_2 x) \in \{-1, 1\}^{n-m}.$$

Applying the triangle inequality yields $\|Y\|_{\psi_2|W} \le \|Y^{(1)}\|_{\psi_2|W_1} + \|Y^{(2)}\|_{\psi_2|W_2}$. By symmetry, it suffices to show that $\mathbb{E}\|Y^{(1)}\|_{\psi_2|W_1}$ is bounded by a universal constant.

Since the aspect ratio satisfies $m/n \le 1/2$, the rectangular block W_1 is well-conditioned. Specifically, its sample covariance $\Sigma_1 = W_1 W_1^\top$ satisfies $nc_1 I \preceq \Sigma_1 \preceq nc_2 I$, and therefore

$$\kappa(\Sigma_1) \le c_2/c_1,$$

with high probability $1 - c_3 e^{-c_4 n}$ for universal constants $c_1, c_2, c_3, c_4 > 0$ (see e.g., [Ver26, Corollary 7.3.2 and Exercise 7.13]). Let $\mathcal{E} := \{\kappa(\Sigma_1) \le c_2/c_1\}$ denote this well-conditioned event.

Since $W_1 x$ is a centered Gaussian vector, its distribution is symmetric about the origin, which implies $\mathbb{E}[Y^{(1)} \mid W_1] = 0$. Applying Theorem 1 directly to $Y^{(1)}$ yields:

$$\|Y^{(1)}\|_{\psi_2|W_1} \le C_0 \sqrt{\kappa(\Sigma_1)} \le C_0 \sqrt{\frac{c_2}{c_1}} =: C_1. \tag{4}$$

To bound the expectation over W, we decompose

$$\mathbb{E}\|Y^{(1)}\|_{\psi_2|W_1} = \mathbb{E}\|Y^{(1)}\|_{\psi_2|W_1} \mathbb{1}_{\mathcal{E}} + \mathbb{E}\|Y^{(1)}\|_{\psi_2|W_1} \mathbb{1}_{\mathcal{E}^c}.$$

On the event $\mathcal{E}$, we use (4). On the rare complement $\mathcal{E}^c$, we use the trivial deterministic bound $|\langle v, Y^{(1)}\rangle| \le \sqrt{m} \le \sqrt{n}$ for all $v \in S^{m-1}$, which implies $\|Y^{(1)}\|_{\psi_2|W_1} \le c_5\sqrt{n}$. Plugging in the bounds for each case and using the tail probability $\mathbb{P}(\mathcal{E}^c) \le c_3 e^{-c_4 n}$ gives:

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

Summing the bounds for both blocks via the triangle inequality completes the proof. □

Remark 4 (What AI could do, and what it couldn't)**.** It is instructive to contrast the AI model's capabilities across these results. While Gemini 3.5 Flash autonomously proved the initial query for the sign function and generalized the argument to arbitrary bounded mappings (yielding Theorem 1), it could not prove Corollary 3 outright, via direct prompting. Transitioning from well-conditioned covariances to square random matrices required a human-designed row-partitioning trick (3) to bypass the singularity issues.

Acknowledgments

The authors thank Simone Bombari for posing the open question that motivated this work. We also acknowledge the mathematical collaboration with Google Gemini 3.5 Flash. 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

[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.

[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.