Bull. Aust. Math. Soc. (First published online 2026), page 1 of 9*
doi:10.1017/S0004972726101713
*Provisional—final page numbers to be inserted when paper edition is published

EXTREME LEAST SINGULAR VALUES OF GAUSSIAN ROW SUBMATRICES AND A PHASE RETRIEVAL STABILITY PROBLEM

YITZCHAK SHMALO

(Received 24 May 2026 UTC; accepted 16 June 2026 UTC)

Abstract

Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$, and put $d_{\mathbb{F}} = \dim_{\mathbb{R}} \mathbb{F}$. Let $A_m \in \mathbb{F}^{N_m \times m}$ have independent standard Gaussian entries and assume $N_m/m \to \gamma > 1$. We determine the exponential scale of the smallest least singular value among all square row submatrices:

$$M_m^{\mathbb{F}} = \min_{T \subset [N_m],\, |T|=m} \sigma_{\min}(A_{m,T}).$$

With $c_\gamma = \gamma^\gamma/(\gamma-1)^{\gamma-1}$, we prove $m^{-1} \log M_m^{\mathbb{F}} \xrightarrow{P} -d_{\mathbb{F}}^{-1} \log c_\gamma$ and, when $N_m = \gamma m + O(1)$, the convergence holds at rate $O(m^{-1})$. As an application, for real phase retrieval at the critical number $N = 2m-1$ of measurements, the Balan–Wang stability parameter satisfies $\omega(A_m) = 4^{-m+o_P(m)}$, so the sharp exponential base in the Gaussian Balan–Wang problem is $1/4$. The lower estimate extends to all bounded-density ensembles and continuity is necessary.

2020 *Mathematics subject classification*: primary 60B20; secondary 15A18, 15B52, 42C15, 94A12.

Keywords and phrases: phase retrieval, random matrices, Gaussian matrices, least singular value, hard edge, complement property, second moment method.

1. Introduction

The least singular value of a random matrix is a basic hard-edge observable in random matrix theory. This paper studies a related extremal quantity: the minimum of the least singular value over all square row submatrices of a tall Gaussian matrix. The number of such submatrices is exponential in the dimension and the main point is to show that their strong dependence does not change the exponential scale predicted by the one-matrix hard-edge small-ball exponent.

The research presented in this paper was supported by the European Research Council (ERC) under the European Union's Horizon 2022 research and innovation programme (grant agreement No. 101041711), the Simons Foundation, Heights Labs, the Israel Science Foundation (grant number 2258/19), and the Israel Science Foundation (ISF Grant 4101/25).

© The Author(s), 2026. Published by Cambridge University Press on behalf of Australian Mathematical Publishing Association Inc. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $d_{\mathbb{F}} = \dim_{\mathbb{R}} \mathbb{F}$, so that $d_{\mathbb{R}} = 1$ and $d_{\mathbb{C}} = 2$. A standard Gaussian variable over $\mathbb{R}$ is $N(0,1)$; a standard Gaussian variable over $\mathbb{C}$ is $(\xi + i\eta)/\sqrt{2}$, where ξ, η are independent $N(0,1)$ variables. For $A \in \mathbb{F}^{N\times m}$ and $T \subset [N]$, A_T denotes the submatrix formed from the rows indexed by T. Define

$$M_m^{\mathbb{F}}(A) = \min_{T\subset[N],\,|T|=m} \sigma_{\min}(A_T). \tag{1.1}$$

Our main result identifies the limiting exponential base for (1.1) for every fixed row aspect ratio $\gamma > 1$.

THEOREM 1.1 (Extremal square submatrices). *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $d_{\mathbb{F}} = \dim_{\mathbb{R}} \mathbb{F}$. Let $A_m \in \mathbb{F}^{N_m\times m}$ have independent standard Gaussian entries and assume $N_m/m \to \gamma > 1$. Put*

$$h(\gamma) = \gamma \log \gamma - (\gamma - 1)\log(\gamma - 1), \quad c_\gamma = e^{h(\gamma)} = \frac{\gamma^\gamma}{(\gamma-1)^{\gamma-1}}. \tag{1.2}$$

Then,

$$\frac{1}{m}\log M_m^{\mathbb{F}}(A_m) \xrightarrow{P} -\frac{h(\gamma)}{d_{\mathbb{F}}}.$$

Equivalently,

$$M_m^{\mathbb{F}}(A_m) = c_\gamma^{-m/d_{\mathbb{F}} + o_P(m)}.$$

If in addition $N_m = \gamma m + O(1)$, for instance, $N_m = \lfloor \gamma m \rfloor$, then for every fixed $\varepsilon > 0$,

$$\mathbb{P}\left\{\left|\frac{1}{m}\log M_m^{\mathbb{F}}(A_m) + \frac{h(\gamma)}{d_{\mathbb{F}}}\right| > \varepsilon\right\} = O_{\gamma,\varepsilon,\mathbb{F}}(m^{-1}). \tag{1.3}$$

The constant c_γ is the exponential growth rate of the relevant binomial coefficient: $\binom{N_m}{m} = \exp\{h(\gamma)m + o(m)\}$. For one $m \times m$ Gaussian matrix over $\mathbb{F}$, the probability that $\sigma_{\min}$ is at most t is of order $t^{d_{\mathbb{F}}}$ at the origin, up to powers of m. If the square submatrices behaved independently, the minimum over $\exp\{h(\gamma)m + o(m)\}$ of them would therefore be expected at scale $\exp\{-h(\gamma)m/d_{\mathbb{F}} + o(m)\}$. Theorem 1.1 proves that this heuristic gives the correct exponential scale.

A closely related question was studied by Rademacher and Shu [11] in the smoothed analysis of Frank–Wolfe methods. In our row notation, their results show that, when the number of Gaussian rows is a fixed factor larger than the dimension, some square submatrix has exponentially small least singular value, and all square submatrices have least singular value bounded below at a possibly different exponential scale. The present paper refines that picture by identifying the exact exponential rate, by treating both real and complex Gaussian matrices, and by extracting the sharp Gaussian base in the Balan–Wang phase retrieval stability question.

The motivating application is the stability of real finite-dimensional phase retrieval; this is the setting of Randomstrasse Open Problem 21. If $A \in \mathbb{R}^{N\times m}$ has rows $a_1, \dots, a_N$, the real phaseless measurement map is $x \bmod \{\pm 1\} \mapsto |Ax| = (|\langle a_i, x\rangle|)_{i=1}^N$. The finite-dimensional frame formulation goes back to Balan, Casazza and Edidin [2];

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

see also [5, 7, 10] for injectivity and stability background. In the real case, injectivity is characterised by the complement property: for every partition of the rows, one side must span $\mathbb{R}^m$. Thus, the sharp generic threshold is $N = 2m - 1$.

Balan and Wang introduced quantitative stability parameters measuring the conditioning of spanning subfamilies and related these parameters to Lipschitz bounds for phaseless inversion [3] (see also [1, 4, 9]). At the real critical threshold, for a full-spark $A \in \mathbb{R}^{(2m-1)\times m}$, define

$$\omega(A) = \min_{J\subset[2m-1],\, \operatorname{rank}(A_J)<m} \sigma_m(A_{J^c}). \tag{1.4}$$

Balan and Wang conjectured an exponential upper bound

$$\omega(A) \le C \max_i \|a_i\|\, b^m \tag{1.5}$$

with universal constants $C > 0$ and $0 < b < 1$. Randomstrasse Open Problem 21 asks for the value of $\omega(A)$ for independent and identically distributed Gaussian A and for the corresponding value of the exponential base in (1.5) [6]. Thus, the following real critical case is not just an example of Theorem 1.1; it is the stated Gaussian Balan–Wang open problem.

For full-spark $A \in \mathbb{R}^{(2m-1)\times m}$, (1.4) reduces to the minimum over square row submatrices:

$$\omega(A) = \min_{T\subset[2m-1],\, |T|=m} \sigma_{\min}(A_T). \tag{1.6}$$

Indeed, the rank-deficient sets J are exactly those with $|J| \le m - 1$ and adding rows can only increase the Gram matrix. Applying Theorem 1.1 with $\mathbb{F} = \mathbb{R}$ and $N_m = 2m - 1$ gives the following answer.

COROLLARY 1.2 (Gaussian Balan–Wang base). *Let $A_m \in \mathbb{R}^{(2m-1)\times m}$ have independent standard Gaussian entries. Then,*

$$\frac{1}{m} \log \omega(A_m) \xrightarrow{P} -\log 4, \tag{1.7}$$

with the rate

$$\mathbb{P}\left\{\left|\frac{1}{m} \log \omega(A_m) + \log 4\right| > \varepsilon\right\} = O_\varepsilon(m^{-1}) \tag{1.8}$$

for every fixed $\varepsilon > 0$. Equivalently, $\omega(A_m) = 4^{-m+o_P(m)}$.

Let $R_m = \max_{1\le i\le 2m-1} \|a_i\|$. For every $b > 1/4$, $\mathbb{P}\{\omega(A_m) \le R_m b^m\} \to 1$. For every $b < 1/4$ and every fixed $C < \infty$, $\mathbb{P}\{\omega(A_m) \le C R_m b^m\} \to 0$. Thus, the Gaussian critical base in the Balan–Wang question is $1/4$.

The complex part of Theorem 1.1 is not a statement about the complex phase retrieval threshold. It is the corresponding square-submatrix random matrix theorem over $\mathbb{C}$. For example, when $N_m/m \to 2$, one obtains $M_m^{\mathbb{C}}(A_m) = 2^{-m+o_P(m)}$, reflecting the two-real-dimensional small-ball exponent of a complex Gaussian scalar.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

2. One-matrix lower tails and binomial growth

We first record the elementary one-matrix estimate needed for the lower bound. It is deliberately weaker than the sharp hard-edge asymptotics of Edelman [8] and subsequent work such as [12, 13]; the polynomial loss is harmless at the exponential scale considered here.

LEMMA 2.1 (Polynomial least-singular-value tail). *Let $G \in \mathbb{F}^{m\times m}$ have independent standard Gaussian entries, where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $d_{\mathbb{F}} = \dim_{\mathbb{R}} \mathbb{F}$. There is a constant $C_{\mathbb{F}}$ such that, for all $m \geq 1$ and all $t > 0$,*

$$\mathbb{P}\{\sigma_{\min}(G) \leq t\} \leq C_{\mathbb{F}} m^{1+d_{\mathbb{F}}/2} t^{d_{\mathbb{F}}}. \tag{2.1}$$

PROOF. Let $R_1, \dots, R_m$ be the rows of G, let $H_i = \operatorname{span}_{\mathbb{F}}\{R_j : j \neq i\}$ and set $d_i = \operatorname{dist}(R_i, H_i)$. Almost surely G is invertible. The negative second moment identity gives

$$\sum_{k=1}^{m} \sigma_k(G)^{-2} = \sum_{i=1}^{m} d_i^{-2}. \tag{2.2}$$

For completeness, if $y_i = G^{-1}e_i$, then $y_i \perp H_i$ and $\langle R_i, y_i\rangle = 1$; hence, $\|y_i\| = d_i^{-1}$. Summing $\|y_i\|^2 = e_i^*(GG^*)^{-1}e_i$ over i proves (2.2).

If $\sigma_{\min}(G) \leq t$, then the left-hand side of (2.2) is at least t^{-2}, so some i satisfies $d_i \leq \sqrt{m}\, t$. Conditional on the other rows, d_i has the same distribution as $|g_{\mathbb{F}}|$, where $g_{\mathbb{R}} \sim N(0,1)$ and $g_{\mathbb{C}}$ is a standard complex Gaussian; hence, $\mathbb{P}\{|g_{\mathbb{R}}| \leq u\} \leq \sqrt{2/\pi}\, u$ and $\mathbb{P}\{|g_{\mathbb{C}}| \leq u\} = 1 - e^{-u^2} \leq u^2$. A union bound over i gives (2.1). □

We also use the following elementary asymptotic form of Stirling's formula.

LEMMA 2.2 (Binomial scale). *Let $N_m/m \to \gamma > 1$. Then, $m^{-1}\log\binom{N_m}{m} \to h(\gamma)$, with h as in (1.2). If $N_m = \gamma m + O(1)$, then $\binom{N_m-1}{m-1} = \Theta_\gamma(m^{-1/2}e^{h(\gamma)m})$.*

3. Lower bound: no square submatrix is much smaller

PROPOSITION 3.1 (Lower estimate). *Under the assumptions of Theorem 1.1, for every fixed $\varepsilon > 0$,*

$$\mathbb{P}\left\{M_m^{\mathbb{F}}(A_m) \leq \exp\left[-\left(\frac{h(\gamma)}{d_{\mathbb{F}}} + \varepsilon\right)m\right]\right\} \to 0. \tag{3.1}$$

If $N_m = \gamma m + O(1)$, the probability in (3.1) is bounded by $C_{\gamma,\varepsilon,\mathbb{F}} m^{1+d_{\mathbb{F}}/2} e^{-d_{\mathbb{F}}\varepsilon m}$ for all large m.

PROOF. Let $t > 0$. By a union bound and Lemma 2.1,

$$\mathbb{P}\{M_m^{\mathbb{F}}(A_m) \leq t\} \leq \binom{N_m}{m} C_{\mathbb{F}} m^{1+d_{\mathbb{F}}/2} t^{d_{\mathbb{F}}}.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

Take $t = \exp[-(h(\gamma)/d_{\mathbb{F}} + \varepsilon)m]$. Since $m^{-1}\log\binom{N_m}{m} \to h(\gamma)$, the bound tends to zero. The stated quantitative bound follows from Lemma 2.2 when $N_m = \gamma m + O(1)$. □

4. Upper bound: one row and many random hyperplanes

The matching upper estimate is the main point. We fix one row and compare it with all hyperplanes generated by $m-1$ of the remaining rows. Let A_m have rows $a_1, \dots, a_N$, with $N = N_m$. Put $x = a_N$ and $b_i = a_i$ for $1 \le i \le N-1$, and define

$$\mathcal{S} = \{S \subset [N-1] : |S| = m-1\}, \quad L = |\mathcal{S}| = \binom{N-1}{m-1}.$$

For $S \in \mathcal{S}$, let $H_S = \operatorname{span}_{\mathbb{F}}\{b_i : i \in S\}$, almost surely a hyperplane in $\mathbb{F}^m$. Choose a unit normal u_S to H_S (over $\mathbb{C}$ determined up to a phase; all events below are phase-invariant). For $t > 0$, set

$$X_S = \mathbf{1}_{\{|\langle x, u_S\rangle| \le t\}}, \quad Z_t = \sum_{S \in \mathcal{S}} X_S. \tag{4.1}$$

If $Z_t > 0$, then for some S, the row x lies within distance t of H_S; for $T = S \cup \{N\}$ and a unit vector $v \perp H_S$, $\|A_{m,T}v\| = |\langle x, v\rangle| = \operatorname{dist}(x, H_S) \le t$. Thus,

$$Z_t > 0 \implies M_m^{\mathbb{F}}(A_m) \le t. \tag{4.2}$$

4.1. Pair geometry

LEMMA 4.1 (Normals for two overlapping row sets). *Fix distinct $S, T \in \mathcal{S}$. Let $r = |S \cap T|$ and $D = m - r$. Then, $D \ge 2$. Conditional on the rows indexed by $S \cap T$, the two normal lines $\mathbb{F}u_S$ and $\mathbb{F}u_T$ are independent Haar-distributed lines in the D-dimensional space $E^{\perp}$, where $E = \operatorname{span}_{\mathbb{F}}\{b_i : i \in S \cap T\}$. Equivalently, for all phase-invariant events, one may take u_S, u_T to be independent Haar unit vectors in $E^{\perp}$.*

PROOF. Almost surely $\dim_{\mathbb{F}} E = r$, so $\dim_{\mathbb{F}} E^{\perp} = D$. The sets $S \setminus T$ and $T \setminus S$ each contain $D-1$ rows, and the corresponding row families are independent conditional on the rows in $S \cap T$. Projecting the rows in $S \setminus T$ to $E^{\perp}$ gives $D-1$ independent standard Gaussian vectors in $E^{\perp}$; their span is a hyperplane and by rotational invariance, its normal line is Haar-distributed. The same argument applies to $T \setminus S$, independently. □

For $D \ge 2$, define, with u, v independent Haar unit vectors in $\mathbb{F}^D$ and g an independent standard Gaussian vector in $\mathbb{F}^D$,

$$q_D^{\mathbb{F}}(t) = \mathbb{P}\{|\langle g, u\rangle| \le t,\ |\langle g, v\rangle| \le t\}, \quad p_{\mathbb{F}}(t) = \mathbb{P}\{|g_{\mathbb{F}}| \le t\}, \tag{4.3}$$

where $g_{\mathbb{F}}$ is a standard scalar Gaussian over $\mathbb{F}$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

LEMMA 4.2 (Two-slab and two-disk estimate). *For each* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *there is a constant* $C_{\mathbb{F}}$ *such that, for* $0 < t \le 1$,

$$q_D^{\mathbb{F}}(t) \le (1 + C_{\mathbb{F}} t^2) p_{\mathbb{F}}(t)^2 \Big(1 + \frac{C_{\mathbb{F}}}{D}\Big), \quad D \ge 3, \tag{4.4}$$

and $q_2^{\mathbb{F}}(t) \le p_{\mathbb{F}}(t)$.

PROOF. The case $D = 2$ is immediate, since the joint event is contained in $\{|\langle g, u\rangle| \le t\}$. Assume $D \ge 3$. For $\mathbb{F} = \mathbb{R}$, conditional on u, v, write $\rho = \langle u, v\rangle$ and $s = (1 - \rho^2)^{1/2}$; the pair $(\langle g, u\rangle, \langle g, v\rangle)$ is a centred real Gaussian vector with covariance determinant s^2 and density at the origin $(2\pi s)^{-1}$, so

$$\mathbb{P}\{|\langle g, u\rangle| \le t, |\langle g, v\rangle| \le t \mid u, v\} \le \frac{2t^2}{\pi s}. \tag{4.5}$$

For independent Haar unit vectors in $\mathbb{R}^D$,

$$\mathbb{E}(1 - \rho^2)^{-1/2} = \frac{\Gamma(D/2)\Gamma((D-2)/2)}{\Gamma((D-1)/2)^2} \le 1 + C/D$$

by the gamma-ratio asymptotics. Averaging (4.5) and using $p_{\mathbb{R}}(t) = \sqrt{2/\pi}\, t + O(t^3)$ gives (4.4) over $\mathbb{R}$. For $\mathbb{F} = \mathbb{C}$, conditional on u, v, write $s^2 = 1 - |\rho|^2$ with $\rho = \langle u, v\rangle$; the pair is a centred complex Gaussian vector in $\mathbb{C}^2$ with covariance determinant s^2 and density at the origin $(\pi^2 s^2)^{-1}$, so the probability of the product of the two disks of radius t is at most $t^4/(1 - |\rho|^2)$. For independent Haar unit vectors in $\mathbb{C}^D$, $|\rho|^2$ has the Beta$(1, D-1)$ distribution; hence, $\mathbb{E}(1 - |\rho|^2)^{-1} = (D-1)/(D-2) \le 1 + C/D$ for $D \ge 3$. Since $p_{\mathbb{C}}(t) = 1 - e^{-t^2} = t^2 + O(t^4)$, averaging gives (4.4) over $\mathbb{C}$. □

LEMMA 4.3 (Overlap combinatorics). *Let* S, T *be independent uniform elements of* $\mathcal{S} = \{S \subset [N-1] : |S| = m - 1\}$ *and set* $D = m - |S \cap T|$. *Then, for* $d \ge 1$,

$$\mathbb{P}\{D = d\} = \frac{\binom{m-1}{d-1}\binom{N-m}{d-1}}{\binom{N-1}{m-1}}, \tag{4.6}$$

with the convention that binomial coefficients outside their natural range vanish. Moreover,

$$\mathbb{E}\frac{1}{D} = \frac{N}{m(N-m+1)}, \quad \mathbb{P}\{D = 2\} = \frac{(m-1)(N-m)}{\binom{N-1}{m-1}}.$$

PROOF. Fix S. To have $D = d$, the set T omits $d - 1$ elements of S and includes $d - 1$ elements from $[N-1] \setminus S$ (of size $N - m$); this gives (4.6). With $n = N - 1$ and $k = m - 1$,

$$\mathbb{E}\frac{1}{D} = \frac{1}{\binom{n}{k}} \sum_{j=0}^{k} \frac{1}{j+1}\binom{k}{j}\binom{n-k}{j} = \frac{1}{\binom{n}{k}} \cdot \frac{1}{k+1} \sum_{j=0}^{k} \binom{k+1}{j+1}\binom{n-k}{j} = \frac{1}{\binom{n}{k}} \cdot \frac{1}{k+1}\binom{n+1}{k},$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

using $(j+1)^{-1}\binom{k}{j} = (k+1)^{-1}\binom{k+1}{j+1}$ and Vandermonde's identity. Hence,

$$\mathbb{E}\frac{1}{D} = \frac{n+1}{(k+1)(n+1-k)} = \frac{N}{m(N-m+1)}.$$

The value of $\mathbb{P}\{D=2\}$ is (4.6) at $d=2$. □

4.2. The second moment

PROPOSITION 4.4 (Upper estimate). *Under the assumptions of Theorem 1.1, for every fixed $\varepsilon > 0$,*

$$\mathbb{P}\Big\{M_m^{\mathbb{F}}(A_m) > \exp\Big[-\Big(\frac{h(\gamma)}{d_{\mathbb{F}}} - \varepsilon\Big)m\Big]\Big\} \to 0. \tag{4.7}$$

If $N_m = \gamma m + O(1)$, the probability in (4.7) is $O_{\gamma,\varepsilon,\mathbb{F}}(m^{-1})$.

PROOF. Take $0 < \varepsilon < h(\gamma)/d_{\mathbb{F}}$. First, assume $N_m = \gamma m + O(1)$ and write $h = h(\gamma)$, $d = d_{\mathbb{F}}$, $t = \exp[-(h/d-\varepsilon)m] \in (0,1]$ for large m. With Z_t as in (4.1) and $L = \binom{N-1}{m-1}$, $\mu := \mathbb{E}Z_t = L\,p_{\mathbb{F}}(t)$. Since $p_{\mathbb{F}}(t) \ge c_{\mathbb{F}}t^d$ for $0 < t \le 1$ and $L \ge c_\gamma m^{-1/2}e^{hm}$ (Lemma 2.2),

$$\mu \ge c_{\gamma,\mathbb{F}}\, m^{-1/2} e^{d\varepsilon m} \to \infty. \tag{4.8}$$

For $S \ne T$, Lemma 4.1 and the independence of x from all b_i give $\mathbb{E}[X_S X_T] = q^{\mathbb{F}}_{D(S,T)}(t)$ with $D(S,T) = m - |S \cap T|$ (projecting x to $E^\perp$ makes it a standard Gaussian vector there, independent of u_S, u_T). It follows that, for S, T independent uniform in $\mathcal{S}$ and $D = m - |S \cap T|$,

$$\frac{\mathbb{E}Z_t^2}{\mu^2} \le \frac{1}{L\,p_{\mathbb{F}}(t)} + \mathbb{E}\Big[\frac{q_D^{\mathbb{F}}(t)}{p_{\mathbb{F}}(t)^2}\mathbf{1}_{\{D\ge 2\}}\Big].$$

By Lemmas 4.2 and 4.3, the $D \ge 3$ part is at most $(1 + C_{\mathbb{F}}t^2)(1 + C_{\mathbb{F}}\,\mathbb{E}[1/D]) \le 1 + C_{\mathbb{F}}t^2 + C_{\gamma,\mathbb{F}}/m$ (using $\mathbb{E}[1/D] = O(1/m)$). Further, the $D=2$ part is at most $\mathbb{P}\{D=2\}/p_{\mathbb{F}}(t) \le C_{\gamma,\mathbb{F}}m^{5/2}e^{-d\varepsilon m}$. With (4.8),

$$\frac{\operatorname{Var} Z_t}{\mu^2} = \frac{\mathbb{E}Z_t^2}{\mu^2} - 1 \le \frac{C_{\gamma,\mathbb{F}}}{m} + C_{\gamma,\varepsilon,\mathbb{F}}m^{5/2}e^{-d\varepsilon m} = O_{\gamma,\varepsilon,\mathbb{F}}(m^{-1}),$$

so by Chebyshev, $\mathbb{P}\{Z_t = 0\} \le \operatorname{Var}(Z_t)/\mu^2 = O(m^{-1})$. With (4.2), this proves (4.7) and its rate when $N_m = \gamma m + O(1)$. For a general sequence $N_m/m \to \gamma$, repeat the proof with $h(N_m/m)$ in place of $h(\gamma)$ and use $h(N_m/m) \to h(\gamma)$. □

5. Proofs of the main theorem and the phase retrieval corollary

PROOF OF THEOREM 1.1. By Propositions 3.1 and 4.4, for every $\varepsilon > 0$, $m^{-1}\log M_m^{\mathbb{F}}(A_m)$ is below $-h(\gamma)/d_{\mathbb{F}} - \varepsilon$ and above $-h(\gamma)/d_{\mathbb{F}} + \varepsilon$, each with probability $\to 0$; this is convergence in probability. When $N_m = \gamma m + O(1)$, the lower probability is exponentially small and the upper probability is $O(m^{-1})$, giving (1.3). □

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

PROOF OF COROLLARY 1.2. For Gaussian $A_m \in \mathbb{R}^{(2m-1)\times m}$, every m rows are independent almost surely, so the reduction (1.6) holds almost surely. Theorem 1.1 with $\mathbb{F} = \mathbb{R}$, $N_m = 2m-1$ and $h(2) = \log 4$ gives (1.7)–(1.8). For the base in (1.5), note $m^{-1}\log R_m \xrightarrow{P} 0$: the upper tail follows from a union bound and $\mathbb{P}\{\|g\| \geq e^{\eta m}\} \leq \exp(-ce^{2\eta m})$ for fixed $\eta > 0$, and $R_m \geq 1$ with high probability. If $b > 1/4$, pick ε with $e^{-(\log 4-\varepsilon)} < b$; then with high probability, $\omega(A_m) \leq e^{-(\log 4-\varepsilon)m} \leq b^m \leq R_m b^m$. If $b < 1/4$ and $C < \infty$, pick ε, η with $e^{\eta} b < e^{-(\log 4+\varepsilon)}$; then with high probability, $\omega(A_m) \geq e^{-(\log 4+\varepsilon)m}$ and $R_m \leq e^{\eta m}$, so $CR_m b^m \leq C(e^{\eta} b)^m < e^{-(\log 4+\varepsilon)m}$ for large m, whence $\mathbb{P}\{\omega(A_m) \leq CR_m b^m\} \to 0$. □

REMARK 5.1 (What drives the constants). The theorem separates the two inputs that determine the exponential base. The combinatorial input is the entropy $\binom{N_m}{m} = \exp\{h(\gamma)m + o(m)\}$. The small-ball input is the real dimension of the scalar field: $\mathbb{P}\{|g_{\mathbb{R}}| \leq t\} \asymp t$, while $\mathbb{P}\{|g_{\mathbb{C}}| \leq t\} \asymp t^2$. Thus, the same binomial entropy is divided by $d_{\mathbb{F}} = 1$ over $\mathbb{R}$ and by $d_{\mathbb{F}} = 2$ over $\mathbb{C}$.

6. Universality and the necessity of continuity

The lower estimate is insensitive to the entry law, provided it has a bounded density.

PROPOSITION 6.1 (Universal lower estimate). *Let $A_m \in \mathbb{R}^{N_m\times m}$ have independent and identically distributed entries with density bounded by K, and $N_m/m \to \gamma > 1$. Then, for every fixed $\varepsilon > 0$, $\mathbb{P}\{M_m^{\mathbb{R}}(A_m) \leq e^{-(h(\gamma)+\varepsilon)m}\} \to 0$.*

PROOF. The identity (2.2) is purely algebraic. For $G \in \mathbb{R}^{m\times m}$ with rows R_i and n_i the unit normal to $\operatorname{span}\{R_j : j \neq i\}$, the variable $\langle R_i, n_i\rangle = \sum_j (R_i)_j (n_i)_j$ is, conditionally on the other rows, a sum of independent terms; the term with $|(n_i)_{j^*}| = \max_j |(n_i)_j| \geq 1/\sqrt{m}$ has density at most $K/|(n_i)_{j^*}| \leq K\sqrt{m}$, so the convolution has density at most $K\sqrt{m}$ and $\mathbb{P}\{d_i \leq s\} \leq 2K\sqrt{m}\, s$. Hence, $\mathbb{P}\{\sigma_{\min}(G) \leq t\} \leq 2Km^2 t$ and a union bound over $\binom{N_m}{m} = e^{h(\gamma)m+o(m)}$ blocks tends to 0 at $t = e^{-(h(\gamma)+\varepsilon)m}$. □

REMARK 6.2 (Continuity is necessary). For atomic entries, the lower estimate fails. If the entries are ± 1 Bernoulli and $N = 2m-1$, every $m \times m$ submatrix is singular with probability at least $\mathbb{P}\{R_1 = R_2\} = 2^{-m}$, so the expected number of singular $m \times m$ submatrices is at least $\binom{2m-1}{m} 2^{-m} \geq c\, 2^m/\sqrt{m} \to \infty$. Thus, the first-moment heuristic behind Proposition 6.1 breaks down and a bounded-density assumption (in particular, Gaussianity) cannot be removed.

REMARK 6.3 (Towards a universal upper estimate). The upper estimate is expected to persist for independent and identically distributed sub-Gaussian entries with bounded density. The only use of Gaussianity is Lemma 4.1, where rotational invariance makes the hyperplane normals Haar-distributed. For bounded-density entries, the normals are delocalised with high probability, so a density-level local central limit theorem yields $\mathbb{P}\{|\langle x, u_S\rangle| \leq t\} = \Theta(t)$ together with the matching pair estimate; making this uniform

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713

over the $\binom{N-1}{m-1}$ hyperplanes requires a delocalisation bound we do not pursue here. A Berry–Esseen estimate does not suffice at the scale $t = e^{-\Theta(m)}$, since its $O(m^{-1/2})$ error dwarfs the signal; a density-level statement is needed.

Acknowledgement

During the preparation of this work, the author used generative AI tools (large language model assistants) to assist with drafting, revision and bibliographic organisation. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the manuscript.

References

[1] R. Balan, 'Stability of frames which give phase retrieval', *Houston J. Math.* **43**(3) (2017), 905–918.
[2] R. Balan, P. G. Casazza and D. Edidin, 'On signal reconstruction without phase', *Appl. Comput. Harmon. Anal.* **20**(3) (2006), 345–356.
[3] R. Balan and Y. Wang, 'Invertibility and robustness of phaseless reconstruction', *Appl. Comput. Harmon. Anal.* **38**(3) (2015), 469–488.
[4] R. Balan and D. Zou, 'On Lipschitz analysis and Lipschitz synthesis for the phase retrieval problem', *Linear Algebra Appl.* **496** (2016), 152–181.
[5] A. S. Bandeira, J. Cahill, D. G. Mixon and A. A. Nelson, 'Saving phase: injectivity and stability for phase retrieval', *Appl. Comput. Harmon. Anal.* **37**(1) (2014), 106–125.
[6] A. S. Bandeira, D. Dmitriev, K. Lucca, P. Nizić-Nikolac and A. Rödder, 'Randomstrasse101: Open Problems of 2025', Preprint, 2026, arXiv:2603.29571.
[7] A. Conca, D. Edidin, M. Hering and C. Vinzant, 'An algebraic characterization of injectivity in phase retrieval', *Appl. Comput. Harmon. Anal.* **38**(2) (2015), 346–356.
[8] A. Edelman, 'Eigenvalues and condition numbers of random matrices', *SIAM J. Matrix Anal. Appl.* **9**(4) (1988), 543–560.
[9] Y. C. Eldar and S. Mendelson, 'Phase retrieval: stability and recovery guarantees', *Appl. Comput. Harmon. Anal.* **36**(3) (2014), 473–494.
[10] P. Grohs, S. Koppensteiner and M. Rathmair, 'Phase retrieval: uniqueness and stability', *SIAM Rev.* **62**(2) (2020), 301–350.
[11] L. Rademacher and C. Shu, 'The smoothed complexity of Frank–Wolfe methods via conditioning of random matrices and polytopes', *Math. Stat. Learn.* **5**(3/4) (2022), 273–310.
[12] M. Rudelson and R. Vershynin, 'The smallest singular value of a random rectangular matrix', *Comm. Pure Appl. Math.* **62**(12) (2009), 1707–1739.
[13] T. Tao and V. Vu, 'Random matrices: the distribution of the smallest singular values', *Geom. Funct. Anal.* **20**(1) (2010), 260–297.

YITZCHAK SHMALO, Einstein Institute of Mathematics,
The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel
e-mail: yitzchak.shmalo@gmail.com

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 04:18:47, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972726101713