Work statistics of sudden quantum quenches: A random matrix theory perspective on Gaussianity and its deviations

Miguel Tierz

Shanghai Institute for Mathematics and Interdisciplinary Sciences
Block A, International Innovation Plaza, No. 657 Songhu Road, Yangpu District,
Shanghai, China
October 13, 2026

arXiv:2509.09640v2 [quant-ph] 13 Oct 2026

We show that, for sudden quenches, the work distribution reduces to the statistics of the traces of powers of Haar unitaries, which are random unitary matrices drawn uniformly from the unitary group. For translation–invariant quadratic fermionic chains with interactions extending to m neighbors and periodic boundary conditions, the Loschmidt amplitude admits a unitary matrix model / Toeplitz representation, which yields a work variable of the form $W = \sum_{r\leq m} a_r \,\Re\,\mathrm{Tr}\, U^r$ (and in models with pairing terms -superconducting pairing- additional $b_r \,\Im\,\mathrm{Tr}\, U^r$ terms appear). By invoking multivariate central limit theorems for vectors of traces of unitaries, we obtain a Gaussian distribution for $P(W)$ with variance $\mathrm{Var}(W) = \frac{1}{2}\sum_r r\,(a_r^2 + b_r^2)$ and asymptotic independence across different powers. We also characterise the conditions under which non-Gaussian tails arise, for example from many interaction terms or their slow decay, as well as the appearance of Fisher–Hartwig singularities. We illustrate these mechanisms in the XY chain. Various numerical diagnostics support the analytical results.

I Introduction

Work is a central stochastic quantity in nonequilibrium thermodynamics: it quantifies not only the average energetic cost of a driving protocol but also the size and structure of fluctuations around that cost. In the quantum setting, the definition of work requires care because work is not an observable associated with a single Hermitian operator. A widely used and operationally motivated choice is the two-projective-measurement (TPM) scheme [1–3], which underlies quantum fluctuation relations and has been implemented via interferometric and ancilla protocols in several platforms [4–6].

In this paper we focus on *sudden quenches* in isolated many-body systems, where a control parameter is changed abruptly and the system subsequently evolves unitarily under the post-quench Hamiltonian. Sudden quenches are a standard theoretical and experimental workhorse: they probe far-from-equilibrium physics on the shortest timescales, connect directly to fidelity decay and orthogonality-catastrophe phenomena [7], and play a key role in modern discussions of dynamical quantum phase transitions (DQPTs), where nonanalytic features emerge in real-time Loschmidt amplitudes [8, 9].

Miguel Tierz: tierz@simis.cn

Work statistics of quenches

From the viewpoint of work statistics, quenches in extended systems exhibit a characteristic separation of scales: the *central* part of the work distribution is typically close to Gaussian in generic settings, while *non-Gaussian* features become pronounced in large-deviation regimes and especially near or at quantum critical points. In such regimes, universal behavior can be understood via quantum-to-classical mappings in a film geometry and the associated critical Casimir physics, with the non-Gaussian large-deviation tail governed by the boundary universality class set by the initial state [10–12]. These results hold broadly—well beyond quadratic fermionic models and Gaussian initial states.

More recently, the role of interaction range has come into focus. For long-range systems with power-law decaying couplings $J_r \propto r^{-\alpha}$, non-Gaussian work statistics can arise even away from quantum critical points in the strong long-range regime $\alpha < d$ [13, 14], consistent with the broader classification of long-range quantum systems [15].

The random-matrix perspective

Our results are complementary to this body of work: by recasting the TPM work characteristic function as a unitary matrix model, we obtain (i) a *rigorous, checkable criterion* for the Gaussian core (the strong Szegő condition on the Toeplitz symbol), (ii) an *explicit variance formula* whose physical content is transparent, and (iii) a *taxonomy* of the distinct mechanisms that produce systematic deviations from Gaussianity—slow Fourier decay (long-range couplings), Fisher–Hartwig singularities (criticality, boundary effects), and proximity of Fisher zeros (DQPTs). The random-matrix viewpoint naturally separates the Gaussian core (typical fluctuations, controlled by multivariate CLTs for traces of Haar unitaries) from the tail structure (controlled by Toeplitz asymptotics beyond strong Szegő), providing a mathematically sharp route to both.

TPM work and sudden quenches

We consider a protocol in which external controls $\lambda(t)$ are varied during a time window $t \in [0, \tau]$. The system Hamiltonian is

$$H(t) \equiv H(\lambda(t)), \qquad H_i := H(\lambda(0^-)), \quad H_f := H(\lambda(\tau^+)). \tag{1}$$

The drive generates the unitary

$$U_\tau = \mathcal{T} \exp\Big(- i \int_0^\tau H(\lambda(t))\, dt \Big). \tag{2}$$

In the TPM scheme [3], one measures H_i at $t = 0^-$, evolves with U_τ, and measures H_f at $t = \tau^+$. If $H_i|n^0\rangle = E_n^{(0)}|n^0\rangle$ and $H_f|m^1\rangle = E_m^{(1)}|m^1\rangle$, the work random variable is

$$w = E_m^{(1)} - E_n^{(0)}, \tag{3}$$

with probabilities $p_n^{(0)} = \langle n^0|\rho_0|n^0\rangle$ and $p_{m|n} = |\langle m^1|U_\tau|n^0\rangle|^2$. The characteristic function of work,

$$\chi_W(u) = \int_{\mathbb{R}} e^{iuw} P(w)\, dw, \tag{4}$$

has the standard operator form

$$\chi_W(u) = \mathrm{Tr}\Big(e^{iuH_f}\, U_\tau\, e^{-iuH_i}\, \rho_0\, U_\tau^\dagger \Big). \tag{5}$$

In a *sudden quench* [7] the drive is instantaneous at $t = 0$: we set $\tau = 0^+$ and evolve for a time t under H_f. If the prepared state is an eigenstate, $H_i|\psi_0\rangle = E_0|\psi_0\rangle$ (the usual situation in our applications: a translation-invariant ground state, or a domain-wall preparation), then (5) reduces to

$$\chi_W(u) = e^{-iuE_0}\,\langle\psi_0|e^{iuH_f}|\psi_0\rangle = e^{-iuE_0}\,G(-u), \tag{6}$$

where

$$G(t) := \langle\psi_0|e^{-iH_f t}|\psi_0\rangle \tag{7}$$

is the *Loschmidt amplitude* and $L(t) = |G(t)|^2$ the Loschmidt echo. Thus $P(w)$ is the Fourier transform of a Loschmidt-type object, while $L(t)$ quantifies fidelity decay [1, 2, 7, 8, 16, 17]. This connection has driven much of the modern activity on DQPTs, where nonanalytic features in $G(t)$ govern far-from-equilibrium behavior [8, 9]. On the experimental side, interferometric protocols access $\chi_W(u)$ directly via an ancilla, enabling reconstructions of $P(w)$ and tests of fluctuation relations in a variety of platforms [4–6].

Why random matrices enter: Toeplitz determinants and trace statistics

The key structural observation we develop and exploit is that, for sudden quenches in quadratic (or effectively free) spin/fermion models with periodic boundary conditions (PBC), the Loschmidt amplitude admits a Toeplitz (or block-Toeplitz) determinant representation whose symbol is built from the post-quench single-particle dispersion $\varepsilon(k)$ [18, 19]. Toeplitz determinants are exactly equivalent to partition functions of unitary random-matrix ensembles via the Heine–Szegő identity [20, 21]. In the present context, this equivalence upgrades the work problem to a random-matrix linear-statistics problem and allows us to import sharp probabilistic tools.

In the settings studied here one has a Toeplitz representation $G_N(t) = D_N(e^{-it\,\varepsilon})$, and therefore $\chi_W(u) = e^{-iuE_0}\,D_N(e^{iu\varepsilon})$. The deterministic phase from the zero Fourier mode only shifts the mean and will often be suppressed; the nontrivial factor is a characteristic function of a random-matrix linear statistic (Sec. III). Concretely, expanding the dispersion in Fourier modes,

$$\varepsilon(k) = a_0 + \sum_{m\geq 1}\,(a_m \cos mk + b_m \sin mk), \tag{8}$$

the Heine–Szegő identity recasts the characteristic function as an average over the unitary group[1], For notational simplicity, we absorb the trivial phase from the zero Fourier mode a_0 by writing $D_N(e^{iu\varepsilon}) = e^{iuNa_0}\,\chi(u)$; equivalently, $\chi(u)$ is the characteristic function of the *centered* linear statistic.

$$\chi(u) = \mathbb{E}_{U(N)}\left[\exp\left(iu\sum_{m\geq 1}\left(\alpha_m \mathrm{Tr}\,U^m + \alpha_m^* \mathrm{Tr}\,U^{-m}\right)\right)\right], \tag{9}$$

with $\alpha_m = (a_m - ib_m)/2$ determined by the quench and microscopic couplings. In other words, *the logarithm of the characteristic function is the cumulant generating function of a linear statistic of traces of powers of a Haar-distributed unitary.*

[1]We illustrate with finite range in formulas for brevity; the Gaussian core extends to infinite-range symbols under $\sum_{r\geq 1} r(a_r^2 + b_r^2) < \infty$, or to truncations with $m(N) = o(N^{2/3})$; cf. Sec. IV.

Gaussian core and its breakdown

This random-matrix viewpoint immediately clarifies when and why a Gaussian work distribution emerges. For fixed m and $N \to \infty$, the traces of powers obey a multivariate central limit theorem: for each m, $\operatorname{Tr} U^m$ converges to an independent complex Gaussian with $\mathbb{E}[\operatorname{Tr} U^m] = 0$ and $\operatorname{Var}(\operatorname{Tr} U^m) = m$, and different powers become asymptotically independent [22–24]. Recent multivariate refinements give super-exponential rates in total variation and quantify regimes where the number of included harmonics may grow with N while retaining Gaussianity [25, 26].

Two points are worth emphasizing for a physics audience. First, the normalization differs from the textbook i.i.d. CLT: there is no $1/\sqrt{N}$; instead the intrinsic variance scale of $\operatorname{Tr} U^m$ is m, and Gaussianity emerges in the large-N limit with m fixed (or slowly growing). Second, the physics of microscopic couplings enters only through the weights α_m multiplying the (asymptotically independent) Gaussian modes. As a consequence, to leading order the work distribution $P(w)$ is Gaussian with variance

$$\operatorname{Var}[w] \;=\; 2\sum_{m\geq 1} m\,|\alpha_m|^2 \;=\; \frac{1}{2}\sum_{m\geq 1} m\,(a_m^2 + b_m^2), \tag{10}$$

while controlled, model-specific non-Gaussian corrections arise when at least one hypothesis behind the Szegő/CLT regime is weakened. We identify and illustrate the following mechanisms: (A) *Too many modes / slow decay*—a large number of harmonics or borderline Fourier decay (as in long-range couplings with $J_r \propto r^{-\alpha}$, $\alpha \leq 1$ in 1D) makes higher cumulants visible; (B) *Fisher–Hartwig singularities*—jumps, zeros, or edge exponents in the Toeplitz symbol [20, 27] generate non-quadratic terms in the asymptotics and structured non-Gaussian tails; (C) *Critical quenches / DQPTs*—Fisher zeros of $G(t)$ approach the real-time axis, narrowing the analyticity strip of $\chi_W(u)$ and thickening the tails of $P(w)$ [8, 9]; (D) *Finite-N effects*—at moderate N or with weight on medium/large r, asymptotically vanishing cumulants remain visible.

Summary of main results

(1) Mapping. For sudden quenches in translation-invariant quadratic fermionic chains (PBC), the work variable maps to a CUE linear statistic:

$$W = \sum_{m\geq 1} (a_m\,\Re\,\operatorname{Tr}\,U^m + b_m\,\Im\,\operatorname{Tr}\,U^m), \qquad U \sim \text{Haar on } U(N),$$

where $\{a_m, b_m\}$ are the Fourier coefficients of the post-quench dispersion $\varepsilon(k)$. [Sec. II–III]

(2) Gaussian core. Under the strong Szegő condition $\sum_{r\geq 1} r(a_r^2 + b_r^2) < \infty$, multivariate CLTs for traces of Haar unitaries give $W \Rightarrow \mathcal{N}(0, \sigma^2)$ as $N \to \infty$ with

$$\sigma^2 = \frac{1}{2}\sum_{m\geq 1} m(a_m^2 + b_m^2)$$

and asymptotic independence across different powers. [Sec. IV]

(3) Deviation mechanisms. Non-Gaussian features appear when: (A) too many modes or slow Fourier decay (long-range couplings); (B) Fisher–Hartwig singularities (criticality, boundaries); (C) Fisher zeros of $G(t)$ near the real axis (DQPTs); (D) finite-N effects. [Sec. V]

(4) Long-range threshold. For $J_r \propto r^{-\alpha}$: the Szegő condition holds iff $\alpha > 1$ (1D), identifying the strong long-range regime $\alpha \leq 1$ as the regime where the strong Szegő/finite-variance CLT assumptions break down and pronounced deviations (non-Gaussianity and/or non-standard scaling) are expected, consistent with Refs. [13, 14]. [Sec. V]

Scope

Our focus is the TPM notion of work for sudden quenches, because in this setting the exact reduction (6) and the Toeplitz/unitary-matrix-model mapping (9) provide a mathematically sharp route to Gaussianity criteria and their breakdown. Extensions to finite-time driving [28] or alternative notions of quantum work [1, 2] typically require different structured-determinant objects (generically beyond Toeplitz form) and are discussed in the outlook (Sec. VII).

Organization

In Sec. II we set up the quadratic fermion models and the Toeplitz/matrix-model representation of the Loschmidt amplitude for translation-invariant Gaussian states and domain-wall preparations. In Sec. III we prove the generator identity and the resulting identification of the work distribution with a random-matrix linear statistic. In Sec. IV we derive the Gaussian core using multivariate CLTs for vectors of traces of Haar unitaries, including quantitative regimes where the number of active harmonics may grow slowly with system size. In Sec. V we specialize to quadratic fermionic chains (including pairing models) and discuss in detail when the Gaussian core is guaranteed and how non-Gaussian tails arise, including a long-range corollary and connections to entanglement via Fisher–Hartwig singularities; we illustrate the mechanisms in the XY chain. Sec. VI presents numerical diagnostics (histograms, Q–Q plots, variance and kurtosis scaling). We conclude in Sec. VII with scope and outlook, and we also include a brief discussion of physical implications of non-Gaussianity at the end of Sec. V.

II Setup

Purpose.— This section fixes conventions for translation-invariant quadratic chains and the Gaussian-state technology used throughout. These ingredients are the starting point for the Toeplitz/matrix-model representations derived later, which ultimately enable the random-matrix formulation of the TPM work characteristic function for sudden quenches.
Model and conventions.— We work with spinless fermions on a ring of length L with periodic boundary conditions (PBC). The lattice operators c_j and $c_j^\dagger$ annihilate/create a fermion at site j and obey the canonical anticommutation relations (CAR)

$$\{c_i, c_j\} = 0, \qquad \{c_i^\dagger, c_j^\dagger\} = 0, \qquad \{c_i, c_j^\dagger\} = \delta_{ij}. \tag{11}$$

With the discrete Brillouin set $\mathcal{B} = \{2\pi n/L : n = 0, 1, \ldots, L-1\}$ we use

$$c_k := \frac{1}{\sqrt{L}} \sum_{j=1}^{L} e^{-ikj}\, c_j, \qquad c_j = \frac{1}{\sqrt{L}} \sum_{k\in\mathcal{B}} e^{ikj}\, c_k, \tag{12}$$

so that $\{c_k, c_{k'}^\dagger\} = \delta_{k,k'}$ and other anticommutators vanish. We set $\hbar = 1$ throughout. We consider finite–range, translation–invariant quadratic fermion chains on a ring of length L with periodic boundary conditions (PBC). In real space,

$$\begin{aligned} H = &\sum_{j=1}^{L}\sum_{r=1}^{m} \Big(J_r\, c_{j+r}^\dagger c_j + \overline{J_r}\, c_j^\dagger c_{j+r} \Big) \\ &+ \sum_{j=1}^{L}\sum_{r=1}^{m} \Big(\Delta_r\, c_{j+r}^\dagger c_j^\dagger + \overline{\Delta_r}\, c_j c_{j+r} \Big) \;+\; \mu \sum_{j=1}^{L} c_j^\dagger c_j. \end{aligned} \tag{13}$$

We call the number–conserving case $\Delta_r \equiv 0$ *XX-type* (unpaired), and the case $\Delta_r \neq 0$ *paired/ Bogoliubov–de Gennes (BdG)*. For background on the free-fermion/spin-chain correspondence and diagonalization, see [29–31]. For the BdG/Nambu formalism and Gaussian-state technology, see [32–34].

Momentum-space form. After a Fourier transform (to $k = \frac{2\pi}{L}\mathbb{Z}$) and, when pairing is present, writing operators in the Nambu basis $\Psi_k = (c_k, c_{-k}^\dagger)^T$, the Hamiltonian decouples into 2×2 blocks with Pauli matrices $(\sigma_x, \sigma_y, \sigma_z)$

$$H = \sum_k \Psi_k^\dagger H(k) \Psi_k, \qquad H(k) = \xi(k)\,\sigma_z + \Delta(k)\,\sigma_y, \tag{14}$$

with single–particle dispersion $\varepsilon(k) = 2\sqrt{\xi(k)^2 + \Delta(k)^2}$. For finite range m,

$$\xi(k) = 2\sum_{r=1}^{m} \Re(J_r)\cos(rk) - \mu, \tag{15}$$

$$\Delta(k) = 2\sum_{r=1}^{m} \Re(\Delta_r)\sin(rk). \tag{16}$$

The unpaired case corresponds to $\Delta \equiv 0$; pairing simply activates the sine harmonics through $\Delta(k)$.

Initial states. We use two standard Gaussian preparations: (i) a translation-invariant Slater/BCS eigenstate of H_i (typically its ground state); and (ii) a *domain-wall* Slater state occupying a contiguous block $A \subset \{1, \ldots, L\}$,

$$|\mathrm{DW}_A\rangle = \prod_{j \in A} c_j^\dagger\,|0\rangle, \tag{17}$$

$$C_{ij} = \langle \mathrm{DW}_A|\, c_i^\dagger c_j\, |\mathrm{DW}_A\rangle \;=\; \mathbf{1}_{\{i=j\in A\}}. \tag{18}$$

Under PBC, case (i) yields for the Loschmidt amplitude a (block-)Toeplitz determinant with symbol built from $e^{-it\varepsilon(k)}$ (and the Bogoliubov angles in the paired case); see the Heine–Szegő framework and its block extensions in [20, 35, 36]. For the domain-wall state, translation symmetry is broken and one obtains Toeplitz *minors* [21] formed from the restricted propagator; see [18, 37, 38]. These reductions are the starting point for our matrix-model/Toeplitz analysis of $\chi_W(u)$ developed below.

For a sudden quench from H_i to H_f with $|\psi_0\rangle$ an H_i eigenstate, the work characteristic function equals the Loschmidt amplitude, $\chi_W(u) = \langle e^{iuW}\rangle = G(-u)$ and $P(W) = (2\pi)^{-1}\int e^{-iuW} G(u)\,du$. For the number–conserving (XX–type) case under PBC and a translation–invariant Slater/BCS initial state, $G_N(t)$ is a Toeplitz determinant with symbol $f_t(e^{i\theta}) = e^{-it\varepsilon(\theta)}$; for a domain–wall Slater state, $G_N(t)$ becomes a Toeplitz *minor* built from the restricted propagator (see [18, 37, 38]). For range-m hoppings J_r the one-body dispersion is $\varepsilon(\theta) = 2\sum_{r=1}^{m} J_r \cos(r\theta)$ and the symbol is $f_t(\theta) = \exp(-it\,\varepsilon(\theta)) = \exp(-i\sum_{r=1}^{m} a_r \cos(r\theta))$ with

$$a_r = 2J_r \qquad \text{(PBC, XX normalization)}. \tag{19}$$

Consequently the work random variable can be written as the linear statistic (that is, a weighted sum of $\cos r\theta_j$ and $\sin r\theta_j$ modes)

$$W = \sum_{r=1}^{m} a_r\, \Re\,\mathrm{Tr}\, U^r, \qquad U \sim \text{Haar on } U(N). \tag{20}$$

With pairing terms, as in the XY chain, the symbol has both $\cos(r\theta)$ and $\sin(r\theta)$ contributions and (20) generalizes to $W = \sum_{r\le m}\big(a_r\,\Re\,\mathrm{Tr}\,U^r + b_r\,\Im\,\mathrm{Tr}\,U^r\big)$. *Pairing* here means quadratic Bogoliubov–de Gennes terms that do not conserve particle number—i.e., anomalous terms $c_k c_{-k}$ and $c^\dagger_{-k}c^\dagger_k$—equivalently, a nonzero off–diagonal gap $\Delta(k)$ in the 2×2 BdG block. This setting will be studied further in Section V.

Exact identification of the generator. Let $U \sim$ Haar on $U(N)$ with eigenangles $\{\theta_j\}_{j=1}^N$. For real vectors $(x_1,\dots,x_m)$ and $(y_1,\dots,y_m)$ define

$$V(\theta) = i\sum_{r=1}^{m}\big(x_r\cos r\theta + y_r\sin r\theta\big), \qquad f(\theta) = e^{V(\theta)}. \tag{21}$$

Lemma 1 (Exact generator identity (Heine–Szegő))**.** *For any fixed m and $N\in\mathbb{N}$,*

$$D_N(e^V) = \mathbb{E}_{U(N)}\exp\Big(i\sum_{r=1}^{m}\big[x_r\,\Re\,\mathrm{Tr}\,U^r + y_r\,\Im\,\mathrm{Tr}\,U^r\big]\Big). \tag{22}$$

Proof. The Heine–Szegő identity [20] states that for any symbol f on the unit circle, $D_N(f) = \mathbb{E}_{U(N)}\prod_{j=1}^N f(e^{i\theta_j})$, where $\{\theta_j\}$ are the eigenangles of the Haar-distributed U. Applied to $f = e^V$, this gives $D_N(e^V) = \mathbb{E}\exp(\sum_{j=1}^N V(\theta_j))$. The result follows from the identities $\sum_j \cos r\theta_j = \Re\,\mathrm{Tr}\,U^r$ and $\sum_j \sin r\theta_j = \Im\,\mathrm{Tr}\,U^r$. □

Corollary 1 (Work/Loschmidt generator)**.** *With $\varepsilon(\theta) = 2\sum_{r\le m} J_r\cos r\theta$ (and optional pairing $b_r\sin r\theta$), the Loschmidt amplitude is $G_N(t) = D_N(e^{-it\varepsilon})$. Up to the conventional sign in $\chi_W(u) = G(-u)$, the work characteristic function is the generator along the ray $(x_r, y_r) = (u\,a_r,\, u\,b_r)$ with $a_r = 2J_r$:*

$$\chi_W(u) = \mathbb{E}\exp\Big(iu\sum_{r\le m}\big[a_r\,\Re\,\mathrm{Tr}\,U^r + b_r\,\Im\,\mathrm{Tr}\,U^r\big]\Big). \tag{23}$$

Thus the Fourier transform of the Toeplitz model for G_N is exactly the joint moment generator of $(\Re\,\mathrm{Tr}\,U^r, \Im\,\mathrm{Tr}\,U^r)_{r\le m}$ evaluated on a one-dimensional ray. In other words, $\chi_W(u)$ equals the moment-generating function of the multivariate random vector $(\Re\,\mathrm{Tr}\,U^r, \Im\,\mathrm{Tr}\,U^r)_{r\le m}$ evaluated along the specific line $(x_r, y_r) = (ua_r, ub_r)$ in its parameter space.

The physical couplings J_r map to mode coefficients via (19). Therefore,

$$\mathrm{Var}(W) = \frac{1}{2}\sum_{r=1}^{m} r\,a_r^2 = 2\sum_{r=1}^{m} r\,J_r^2, \tag{24}$$

and, with pairing coefficients b_r, $\mathrm{Var}(W) = \frac{1}{2}\sum_r r\,(a_r^2 + b_r^2)$. Different neighbor ranges contribute additively.

III Fourier transform of Loschmidt amplitude: $P(W)$ as linear statistics of CUE eigenangles

Physically, this section establishes that the work distribution $P(W)$ is the probability law of a specific random-matrix observable: a weighted sum of traces of powers of a Haar-distributed unitary, shifted by the initial energy. This is the identification that makes the entire random-matrix toolkit available for analyzing work statistics. The mathematical content is a precise statement (Theorem 1) that the Toeplitz determinant $D_N(e^{iu\varepsilon})$ is the characteristic function of this observable, followed by Fourier inversion.

Setup. Let $U \sim$ Haar on $U(N)$ with eigenangles $\{\theta_j\}_{j=1}^N$. Fix a real 2π-periodic ε with Fourier expansion

$$\varepsilon(\theta) = \varepsilon_0 + \sum_{r\geq 1} (\alpha_r \cos r\theta + \beta_r \sin r\theta), \qquad \alpha_r, \beta_r \in \mathbb{R}. \tag{25}$$

Define the (real) linear statistic

$$S_N(U) := \sum_{j=1}^N \varepsilon(\theta_j) = N\varepsilon_0 + \sum_{r\geq 1} \Big(\alpha_r \,\Re \operatorname{Tr} U^r + \beta_r \,\Im \operatorname{Tr} U^r\Big). \tag{26}$$

Using Lemma 1, we are now going to state and prove Theorem 1, which establishes that the work distribution $P(w)$ is the distribution of the random variable $S_N(U) - E_0$, i.e. a linear combination of random matrix traces offset by the initial energy.

Theorem 1 (Work characteristic function and law)**.**
(i) Characteristic function. *Let $|\psi_0\rangle$ be an eigenstate of H_i with $H_i|\psi_0\rangle = E_0|\psi_0\rangle$ (so E_0 is the initial energy). For a sudden quench to H_f whose Loschmidt amplitude admits the Toeplitz representation $G_N(t) = D_N(e^{-it\,\varepsilon})$ (this holds for translation-invariant free fermion models with PBC; see [18, 19] and Section V below), one has*

$$\begin{aligned} \chi_W(u) &= \langle\psi_0|e^{iuH_f} e^{-iuH_i}|\psi_0\rangle \\ &= e^{-iuE_0}\, D_N(e^{iu\varepsilon}) \\ &= e^{-iuE_0}\, \mathbb{E}\, e^{iu\, S_N(U)}\,. \end{aligned} \tag{27}$$

So, using Lemma 1 with $(x_r, y_r) = (ua_r, ub_r)$ (Corollary 1), we identify $D_N(e^{iu\varepsilon}) = \mathbb{E}[\exp(iu\, S_N(U))]$, hence $\chi_W(u) = e^{-iuE_0}\mathbb{E}[e^{iu\, S_N(U)}]$, proving (i).
(ii) Fourier inversion and pushforward law. *Let P_N denote the work law. As a tempered distribution on $\mathbb{R}$,*

$$P_N(w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iuw}\, \chi_W(u)\, du. \tag{28}$$

Equivalently, for every Schwartz test function $\varphi \in \mathcal{S}(\mathbb{R})$,

$$\int_{\mathbb{R}} \varphi(w)\, P_N(dw) = \mathbb{E}\big[\varphi(S_N(U) - E_0)\big]\,. \tag{29}$$

In particular, in Dirac notation, $P_N = \mathbb{E}\big[\delta(\,\cdot - (S_N(U) - E_0))\big]$ (distributionally). This means that $P_N(w)$ is the probability law of $S_N(U) - E_0$ when U is Haar-random (this is the meaning of (29), interpreting the result as a (tempered) distribution).
Sketch of proof. Let $\widehat{\varphi}(u) := \int_{\mathbb{R}} e^{-iuw}\, \varphi(w)\, dw$. By part (i), $\chi_W(u) = e^{-iuE_0}\, \mathbb{E}\Big[e^{iuS_N(U)}\Big]$. Since $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, Tonelli/Fubini applies and

$$\begin{aligned} \int \varphi(w)\, P_N(dw) &= \frac{1}{2\pi}\int \widehat{\varphi}(u)\, \chi_W(u)\, du \\ &= \frac{1}{2\pi}\int \widehat{\varphi}(u)\, e^{-iuE_0}\, \mathbb{E}\Big[e^{iuS_N(U)}\Big]\, du \\ &= \mathbb{E}\Big[\frac{1}{2\pi}\int \widehat{\varphi}(u)\, e^{iu(S_N(U)-E_0)} du\Big] \\ &= \mathbb{E}\big[\varphi(S_N(U) - E_0)\big]\,, \end{aligned} \tag{30}$$

which proves (29).

Remarks. *(1) What is and is not Fourier–transformed.* By Lemma 1 and Theorem 1(i), $D_N(e^{iu\varepsilon})$ is (up to the trivial phase e^{-iuE_0}) the *characteristic function* of the real linear statistic $S_N(U)$, see Eq. (20); hence it is positive–definite and satisfies $\chi_W(0) = 1$. The inverse Fourier transform acts on this characteristic function in the u–variable and, in the tempered sense, yields the pushforward law in Eq. (28) (equivalently Eq. (29)). In particular,

$$P_N \;=\; \mathbb{E}\big[\delta(\,\cdot\,-(S_N(U)-E_0))\big] \quad \text{(as distributions)}.$$

Crucially, nothing is Fourier–transformed pointwise in the angle θ: the symbol $e^{iu\varepsilon(\theta)}$ *enters through* the Toeplitz/matrix–model identity [20, 21] that produces the characteristic function; the Fourier inversion is only in u.
(2) Note that each trace $\operatorname{Tr} U^r$ is complex, so writing S_N in terms of $\Re \operatorname{Tr} U^r$ and $\Im \operatorname{Tr} U^r$ emphasizes that each mode contributes two independent Gaussian degrees of freedom. In other words, each trace is like a 2D Gaussian (pair of quadratures), ensuring S_N is a real linear combination of those 2D components
(3) *Centering and variance.* Since $\mathbb{E}[\operatorname{Tr} U^r] = 0$, we have $\mathbb{E}[S_N] = N\varepsilon_0$ and

$$\operatorname{Var}(S_N - N\varepsilon_0) = \frac{1}{2}\sum_{r\geq 1} r(\alpha_r^2+\beta_r^2) \quad \text{(exact for } N \geq \max r), \tag{31}$$

using $\mathbb{E}|\operatorname{Tr} U^r|^2 = \min\{r, N\}$ [22] and orthogonality across different r (see next Section for more on orthogonality). Notice that the real and imaginary parts of each T_r are independent Gaussians with variance $r/2$ and this is why the variance (31) is a simple sum of contributions.

Corollary 2 (Gaussian core with a growing number of modes)**.** *Let $U \sim$ Haar on $U(N)$ and set*

$$W_N = \sum_{r\geq 1}\big(a_r\,\Re\,\operatorname{Tr} U^r + b_r\,\Im\,\operatorname{Tr} U^r\big).$$

Assume the Szegő regularity $\sum_{r\geq 1} r(a_r^2+b_r^2) < \infty$. For any sequence $m(N) \to \infty$ with $m(N) = o(N^{2/3})$, write

$$W_N = W_N^{(\leq m)} + R_N^{(>m)}, \tag{32}$$

with the truncation

$$W_N^{(\leq m)} := \sum_{r=1}^{m(N)} (a_r X_r + b_r Y_r), \qquad \text{with } X_r := \Re\,\operatorname{Tr} U^r, \quad Y_r := \Im\,\operatorname{Tr} U^r. \tag{33}$$

Then, as $N \to \infty$,

$$W_N \;\Rightarrow\; \mathcal{N}\Big(0,\; \tfrac{1}{2}\sum_{r\geq 1} r(a_r^2+b_r^2)\Big). \tag{34}$$

Moreover, for each compact $K \subset \mathbb{R}$,

$$\sup_{u\in K}\big|\chi_{W_N}(u) - e^{-u^2\sigma^2/2}\big| \;\longrightarrow\; 0. \tag{35}$$

$$\sigma^2 = \tfrac{1}{2}\sum_{r\geq 1} r\,(a_r^2+b_r^2).$$

In addition, if $m(N) \ll \sqrt{N}$ then the total-variation error for $W_N^{(\leq m)}$ is super-exponentially small in $N/m(N)$ [25]. Sketch. *By the multivariate CLT for traces [25], the $2m(N)$–vector*

$(X_1, Y_1, \ldots, X_{m(N)}, Y_{m(N)})$, *scaled by* $\sqrt{r/2}$ *on each coordinate, converges (in total variation) to a centered Gaussian in* $\mathbb{R}^{2m(N)}$ *whenever* $m(N) = o(N^{2/3})$, *with a super-exponential rate if* $m \ll \sqrt{N}$. *Hence* $W_N^{(\leq m)}$ *is asymptotically* $\mathcal{N}(0, \frac{1}{2}\sum_{r\leq m(N)} r(a_r^2 + b_r^2))$. *The tail has vanishing variance,* $\operatorname{Var} R_N^{(>m)} = \frac{1}{2}\sum_{r>m(N)} r(a_r^2 + b_r^2) \to 0$, *so* (35) *follows by Slutsky's lemma.*

IV Gaussian core: why traces of powers are asymptotically independent Gaussians

This section answers the central physical question: *why is the work distribution generically Gaussian for large systems?* The answer is that the individual trace variables $\operatorname{Tr} U^r$ entering the work (Eq. (20)) become independent Gaussians in the large-N limit, with a variance that depends on r but *not* on N. The work, being a linear combination of these, inherits Gaussianity with the explicit variance $\frac{1}{2}\sum_r r(a_r^2 + b_r^2)$. We also explain why this CLT is qualitatively different from the textbook i.i.d. version.

Let $U \sim$ Haar on $U(N)$ and write $T_r := \operatorname{Tr} U^r = \sum_{j=1}^N e^{ir\theta_j}$. Already at finite N one has [22, 23]

$$\mathbb{E}\,|T_r|^2 = \min\{r, N\}, \qquad \mathbb{E}\,T_r = 0, \qquad \mathbb{E}\,T_r T_s = 0 \ (r \neq s), \tag{36}$$

This fixes the natural fluctuation scale: the variance of T_r is of order r (rather than N). For any fixed m, we have convergence in distribution to i.i.d. complex Gaussians, namely

$$\left(\tfrac{T_1}{\sqrt{1}}, \tfrac{T_2}{\sqrt{2}}, \ldots, \tfrac{T_m}{\sqrt{m}}\right) \xrightarrow[N\to\infty]{d} (Z_1, \ldots, Z_m), \tag{37}$$

with Z_r i.i.d. standard complex Gaussians. Equivalently,

$$\mathbb{E}[T_r] = 0, \quad \mathbb{E}[T_r \overline{T_s}] = \delta_{rs}\, r, \quad \mathbb{E}[T_r T_s] = 0. \tag{38}$$

So, different trace powers become independent Gaussians in the large-N limit and it is the orthogonality $\mathbb{E}[T_r\overline{T_s}] = 0$ for $r \neq s$ (and $\mathbb{E}[T_r T_s] = 0$) that makes the joint distribution asymptotically factorize. There are two complementary derivations which are physically illuminating.

(i) Toeplitz/cumulants (strong Szegő). Heine–Szegő identifies $D_N(e^{iu\varepsilon})$ with the characteristic function of $S_N = \sum_j \varepsilon(\theta_j) = \sum_r(\alpha_r \Re T_r + \beta_r \Im T_r) + N\varepsilon_0$. The strong Szegő limit theorem expands $\log D_N$ into cumulants and shows that as $N \to \infty$ [24]

$$\log D_N(e^{iu\varepsilon}) = iu\, N\varepsilon_0 - \frac{u^2}{4}\sum_{r\geq 1} r(\alpha_r^2 + \beta_r^2) + o(1), \tag{39}$$

so $\kappa_1(S_N) = N\varepsilon_0$, $\kappa_2(S_N) = \frac{1}{2}\sum_r r(\alpha_r^2 + \beta_r^2)$, and $\kappa_p(S_N) \to 0$ for $p \geq 3$. Recall that $o(1)$ denotes a term that vanishes as $N \to \infty$.

(ii) Direct multivariate CLT for traces. K. Johansson [24] established the CLT for linear statistics, Rains [39] studied the distribution of T_r when r grows with N, and Johansson and Lambert [25] gave quantitative multivariate bounds (TV metric) and growth windows for $m(N)$: the multivariate CLT remains valid uniformly for $m = o(N^{2/3})$ and when $m(N) = o(\sqrt{N})$ the total-variation error is super-exponentially small in N/m [25]. This is convenient for controlling $\chi_W(u)$ uniformly on compact u-intervals.

Under any of these $m(N)$ growth values, for any fixed m and coefficients $\{a_r, b_r\}_{r\le m}$,

$$W = \sum_{r=1}^{m} \left(a_r \,\Re T_r + b_r \,\Im T_r\right) \implies \mathcal{N}\Big(0,\ \tfrac{1}{2}\sum_{r=1}^{m} r(a_r^2 + b_r^2)\Big). \tag{40}$$

This is our *Gaussian core* for $P(W)$, verified below numerically. The Gaussian core is visible in the straight quantile–quantile plots, e.g. Fig. 1(b,e) and Fig. 2(b,e).

Finite-N corrections and growth of m. At fixed m, the leading non-Gaussian features are governed by the third and fourth cumulants. When m grows with N, the multivariate bounds of Johansson–Lambert quantify how fast m may increase while $(T_1, \dots, T_m)$ stays close (in total variation) to its Gaussian limit. Finite-N deviations appear only in the extremes (detrended Q–Q): see Fig. 1(c,f) and Fig. 2(c,f). The surrogate and exact Haar data are indistinguishable in the bulk (compare Fig. 3, rows 2–3).

Different powers $r \neq s$ are uncorrelated (indeed asymptotically independent), which justifies the variance additivity in (31) and later formulas.

A Products of traces: selection rules, Wick limits, and large-N structure

Many observables with couplings to 1st, 2nd, …, m-th neighbors lead to correlators of products of traces such as $\mathbb{E}\big[\prod_{r\le m} T_r^{n_r}\,\overline{T_r}^{\bar{n}_r}\big]$. Two structural facts guide computations [22, 23].

Selection rule (exact for all N). Invariance under the global phase $U \mapsto e^{i\phi}U$ implies

$$\mathbb{E}\Big[\prod_{a=1}^{p} T_{k_a} \prod_{b=1}^{q} \overline{T_{\ell_b}}\Big] = 0 \quad \text{unless} \quad \sum_{a=1}^{p} k_a = \sum_{b=1}^{q} \ell_b. \tag{41}$$

So, this invariance is why any moment with an unmatched number of T factors vanishes. In particular $\mathbb{E}[T_1 T_2 \cdots T_m] = 0$ for all $m \ge 1$.

Wick factorization (asymptotic). Since (37) holds with independent complex Gaussian limits,

$$\mathbb{E}\Big[\prod_{r=1}^{m} T_r^{n_r}\,\overline{T_r}^{\bar{n}_r}\Big] = \prod_{r=1}^{m} \left(\delta_{n_r,\bar{n}_r}\, n_r!\, r^{n_r}\right) + o(1). \tag{42}$$

This means that mixed moment factorizes into second moments (pairings) plus $o(1)$. Thus, at leading order, only pairings T_r with $\overline{T_r}$ survive; cross-pairings between different powers are subleading. Two basic consequences are, with r, s fixed, $N \to \infty$

$$\mathbb{E}\,|T_r|^{2q} = q!\, r^{\,q} + o(1), \qquad \mathbb{E}\big[T_r\,\overline{T_s}\big] = \delta_{rs}\, r. \tag{43}$$

The lack of cross-covariance and the $\sqrt{r/2}$ scaling are borne out by the scatter panels with 95% ellipses; see Fig. 5(a,b) and the standardized versions in Fig. 5(c,d).

V Quadratic fermionic chains, Gaussian core, and non-Gaussian tails

Purpose.— In this section we translate the abstract CUE/trace-CLT framework of Secs. III–IV into concrete criteria on the physical couplings of quadratic fermionic chains. We also catalogue the main mechanisms that produce deviations from a finite-variance Gaussian work core (slow Fourier decay and long-range couplings, Fisher–Hartwig singularities, and critical/DQPT effects), using the XY chain as a representative benchmark.

Notation. In the abstract framework of Secs. III–IV, the Fourier coefficients of the dispersion were denoted α_r (cosine) and β_r (sine). From this section onward we write a_r and b_r for the same quantities, absorbing any conventional factors of 2 (cf. Eq. (19)). In particular, $\mathrm{Var}(W) = \frac{1}{2}\sum_r r(a_r^2 + b_r^2)$ and all formulas below use a_r, b_r exclusively; the correspondence is $a_r \leftrightarrow \alpha_r$, $b_r \leftrightarrow \beta_r$.

Before turning to concrete models, it is worth emphasizing that the strong Szegő condition (48) is not just a technical regularity assumption. In the present setting it acts as a sharp *physical diagnostic*: through the Fourier tail of the single-particle data it detects (i) long-range couplings (slow decay of a_r, b_r), (ii) critical or boundary-induced singular structure (Fisher–Hartwig data) that produces nonstandard fluctuation scaling, and (iii) parameter regimes where entanglement scaling is anomalous in free-fermion chains. In this sense, the Szegő/Fisher–Hartwig classification provides a single language—and, in favorable cases, exact thresholds and variance formulas—for predicting when the work distribution has a finite-variance Gaussian core and when it does not.

A Starting point: a general quadratic chain (with pairing)

Consider a translation-invariant quadratic fermionic Hamiltonian (with possible pairing) that, after Fourier and Nambu (BdG) lifting [29, 32], decouples as

$$H = \sum_{k\in\mathcal{B}} \Psi_k^\dagger\, \mathcal{H}(k)\, \Psi_k, \qquad \mathcal{H}(k) = \xi(k)\,\sigma_z + \Delta(k)\,\sigma_y, \tag{44}$$

with dispersion $\varepsilon(k) = 2\sqrt{\xi(k)^2 + \Delta(k)^2}$ and a Brillouin set $\mathcal{B}$. For a sudden quench from H_i to H_f, the work characteristic function

$$\chi_W(u) = \langle e^{iuH_f} e^{-iuH_i}\rangle_{\psi_0} = e^{-iuE_0}\,\langle\psi_0|e^{iuH_f}|\psi_0\rangle \tag{45}$$

is representation-invariant; it factorizes over k and admits a compact determinant/Pfaffian form [18, 34]. Under PBC and translation invariance, $\log\chi_W(u)$ is a Riemann sum over k that, upon Fourier expansion in θ, reduces to the Toeplitz model used above:

$$\chi_W(u) = D_N\left(e^{i\sum_{r\geq 1}[x_r\cos(r\theta)+y_r\sin(r\theta)]}\right), \tag{46}$$

with coefficients (x_r, y_r) determined by the Fourier content of ε together with the Bogoliubov angles that encode the initial BCS vacuum. Consequently the work variable is the linear statistic

$$W = \sum_{r\geq 1}\left(a_r\,\Re\,\mathrm{Tr}\,U^r + b_r\,\Im\,\mathrm{Tr}\,U^r\right), \tag{47}$$

where $a_r \propto \hat{\varepsilon}_r^{(\cos)}$ and $b_r \propto \hat{\varepsilon}_r^{(\sin)}$ (pairing produces the sine-harmonics). We stress that in (46), the coefficients x_r and y_r are directly related to the Fourier expansion of the single-particle energy $\varepsilon(k)$. Writing $\varepsilon(\theta) = \varepsilon_0 + \sum_{r\geq 1}[\varepsilon_r^{(c)}\cos(r\theta) + \varepsilon_r^{(s)}\sin(r\theta)]$, and noting that pairing (nonzero $\Delta(k)$) contributes sine terms via the Bogoliubov rotation, one finds $x_r = u\,\varepsilon_r^{(c)}$ and $y_r = u\,\varepsilon_r^{(s)}$. Equivalently, $a_r = \varepsilon_r^{(c)}$ and $b_r = \varepsilon_r^{(s)}$ in the notation of (47) (up to the factor of 2 coming from the Jordan-Wigner convention, cf. Corollary 1).
The fact that the initial state is a *Bogoliubov* (BCS) vacuum rather than the Jordan–Wigner vacuum merely changes the values of $\{a_r, b_r\}$; the identities $\chi_W(u) = G(-u)$ and the Toeplitz representation are unaffected.

B When the Gaussian core is guaranteed

Let $V(\theta) = iu\,\varepsilon(\theta) = iu\big(\varepsilon_0 + \sum_{r\geq 1}(\alpha_r \cos r\theta + \beta_r \sin r\theta)\big)$. The strong Szegő condition

$$\sum_{r\geq 1} r\,|\widehat{V}_r|^2 = u^2 \sum_{r\geq 1} r(\alpha_r^2 + \beta_r^2) < \infty \qquad (V \in H^{1/2}) \tag{48}$$

is satisfied whenever ε has Fourier modes decaying a bit faster than $1/r$ [24]. Under (48): (i) the cumulant expansion gives $\kappa_1(S_N) = N\varepsilon_0$, $\kappa_2(S_N) = \frac{1}{2}\sum_{r\geq 1} r(\alpha_r^2 + \beta_r^2)$ and $\kappa_p(S_N) \to 0$ for $p \geq 3$; (ii) the multivariate CLT for $(T_1, \ldots, T_m)$ after $\sqrt{r}$ rescaling holds for fixed m, with quantitative extensions when $m = m(N)$ grows slowly. Therefore, for $W = \sum_{r\leq m}(a_r\,\Re T_r + b_r\,\Im T_r)$ one has

$$W \;\Rightarrow\; \mathcal{N}\Big(0,\; \tfrac{1}{2}\sum_{r\leq m} r(a_r^2 + b_r^2)\Big) \qquad (N \to \infty). \tag{49}$$

It is worth stressing that the strong Szegő condition (48) is what guarantees the Gaussian core rigorously: W converges in distribution to $\mathcal{N}(0, \frac{1}{2}\sum r(a_r^2 + b_r^2))$. Two useful corollaries: (1) if ε is real-analytic (gapped models, generic BdG), then (a_r, b_r) decay exponentially and the Gaussian core is extremely sharp; (2) even with mild cusps (e.g., linear dispersion near a gap closing), the sum in (48) typically converges, so the core remains Gaussian while finite-N corrections increase. The exponential vs. power-law Fourier decay manifests exactly as expected: the exponential case shows a virtually perfect Q–Q line, while the power-law case exhibits mild tail curvature (Fig. 2(e) vs. Fig. 2(b,c)).

Long-range couplings and the Szegő threshold. The strong Szegő condition (48) acquires a transparent physical interpretation when the couplings decay as a power law [15]. Consider a one-dimensional chain (13) with hopping and pairing amplitudes

$$J_r = \frac{J}{r^\alpha}, \qquad \Delta_r = \frac{\Delta}{r^\alpha}, \qquad r = 1, 2, 3, \ldots \tag{50}$$

(with J, Δ setting the overall scales and $\alpha > 0$ the decay exponent). From Eqs. (15)–(16), the dispersion inherits Fourier coefficients that scale as $a_r \sim r^{-\alpha}$ and $b_r \sim r^{-\alpha}$ (up to model-dependent constants and the Bogoliubov rotation). Substituting into (48),

$$\sum_{r\geq 1} r\,(a_r^2 + b_r^2) \;\sim\; \sum_{r\geq 1} r^{1-2\alpha}, \tag{51}$$

which converges if and only if $1 - 2\alpha < -1$, i.e. $\alpha > 1$. This identifies $\alpha = 1$ (in one spatial dimension) as a sharp threshold within the Toeplitz/RMT framework:

- *Weak long-range* ($\alpha > 1$, equivalently $\alpha > d$ in d dimensions): the strong Szegő condition is generically satisfied, and the Gaussian core with variance $\frac{1}{2}\sum_r r(a_r^2 + b_r^2) < \infty$ is robust. Physical intuition: the effective number of strongly contributing harmonics is finite, and the multivariate CLT for traces applies.

- *Strong long-range* ($\alpha \leq 1$): the Szegő condition fails; the variance associated with the Gaussian-core formula is no longer finite in the fixed-symbol setting, and the strong Szegő/CLT regime used above does not apply. One expects pronounced deviations from the finite-variance Gaussian core (non-Gaussianity and/or different scaling) even away from quantum critical points; see Fig. 4c and Fig. 4d for numerical illustrations (a pedagogical long-range CLT proxy and a direct CUE variance check).

The threshold $\alpha = 1$ in one dimension coincides exactly with the boundary $\alpha = d$ between the weak and strong long-range regimes in the standard classification of long-range quantum systems [15]. It is noteworthy that this boundary emerges here from a purely *mathematical* criterion (Szegő regularity of the Toeplitz symbol), rather than from a physical scaling or renormalization-group argument. Yet the conclusion is fully consistent with recent independent results obtained by very different methods: universal non-Gaussian work statistics have been found in long-range systems at criticality for $\alpha > d$ [14], while non-Gaussianity has been shown to persist even away from criticality in the strong long-range regime $\alpha < d$ [13].

At the borderline $\alpha = 1$, the sum in (51) diverges logarithmically: $\sum_{r=1}^{R} r^{-1} \sim \log R$. Correspondingly, the variance of the work variable grows as $\mathrm{Var}(W) \sim \log N$ when the number of active harmonics extends to $r \sim N$, and the correct scaling for the centered work distribution changes. This borderline regime connects to the Fisher–Hartwig discussion in the next subsection: the Fourier coefficients $c_m \sim m^{-1}$ that signal the onset of logarithmic variance growth are precisely those associated with jump-type FH singularities in the Toeplitz symbol (cf. the variance $\sim \sum_j \beta_j^2 \log N$ computed below for FH symbols with jump exponents β_j).

Kac normalization and extensivity. The strong long-range regime $\alpha \leq d$ is also the regime most relevant to current experimental platforms for long-range quantum matter (e.g. trapped ions, Rydberg arrays, cavity-mediated systems), where an N-dependent Kac normalization is routinely introduced to keep the energy extensive; see Ref. [15]. In this setting a natural question is whether Kac normalization restores the *finite-variance Gaussian-core* regime predicted by the strong Szegő condition.

Implementing the Kac prescription amounts to replacing $J_r \to J_r/\mathcal{N}_\alpha(N)$ (and similarly for pairing) with $\mathcal{N}_\alpha(N) = \sum_{r=1}^{N} r^{-\alpha}$. This makes the effective Fourier weights N-dependent, $a_r \to a_r(N)$, and the Szegő condition must be reconsidered in the joint limit $N \to \infty$ with an evolving symbol. For $\alpha < 1$, $\mathcal{N}_\alpha(N) \sim N^{1-\alpha}/(1-\alpha)$, so that (heuristically) $a_r(N) \sim r^{-\alpha}/N^{1-\alpha}$ and hence

$$\sum_{r=1}^{N} r\, a_r(N)^2 \;\sim\; \frac{1}{N^{2(1-\alpha)}} \sum_{r=1}^{N} r^{1-2\alpha} \;=\; O(1). \tag{52}$$

This estimate suggests that Kac normalization can restore the *finiteness* of the Szegő sum. We therefore formulate the following working conjecture.

Conjecture (Gaussian core under Kac normalization). For $\alpha < 1$ with Kac normalization, the work distribution recovers a Gaussian core with a finite limiting variance

$$\sigma_\infty^2 = \frac{1}{2} \lim_{N\to\infty} \sum_{r=1}^{N} r\, a_r(N)^2 < \infty,$$

and higher cumulants vanish in the thermodynamic limit (after centering by the mean work). A rigorous proof would require controlling Toeplitz determinants with an N-dependent symbol and establishing CLT error bounds uniform in N; this goes beyond the standard strong Szegő theorem [24] and its fixed-symbol extensions [25].

At the borderline $\alpha = 1$, $\mathcal{N}_1(N) \sim \log N$, giving $a_r(N) \sim r^{-1}/\log N$ and

$$\sum_{r=1}^{N} r\, a_r(N)^2 \sim (\log N)^{-2} \sum_{r=1}^{N} r^{-1} \sim \frac{1}{\log N} \to 0,$$

suggesting that Kac normalization may "over-correct" the fluctuations at $\alpha = 1$ and that a different rescaling is required to obtain a nontrivial limiting work distribution in this borderline case.

C Where non-Gaussian tails come from

In several physically relevant mappings, such as for the case of open boundary conditions presented in [18] the symbol picks up genuine Fisher–Hartwig (FH) singularities such as algebraic endpoints and/or jump phases, unlike the PBC case studied here, where $\varepsilon(\theta)$ is smooth unless crossing a critical point. A canonical FH symbol has the form, with $\alpha_j > -1/2$ [20]

$$f_{\mathrm{FH}}(\theta) = e^{V(\theta)} \prod_{j=1}^{J} |2 - 2\cos(\theta - \theta_j)|^{\alpha_j} \, e^{i\beta_j(\theta - \theta_j - \pi)}. \tag{53}$$

For such symbols, the Toeplitz asymptotics depart from the pure strong Szegő quadratic law: $D_N(f_{\mathrm{FH}})$ exhibits explicit N-dependent power laws, oscillatory factors, and constant terms determined by the FH data $\{\alpha_j, \beta_j, \theta_j\}$.
Under PBC, Loschmidt symbols naturally enter the Fisher–Hartwig (FH) class whenever the quench creates discontinuities at Fermi points or algebraic endpoints. For a *flux (twist) quench* without pairing, one has a scalar Toeplitz symbol $f_t(e^{i\theta}) = \exp(it\,\varepsilon(\theta))\, g(\theta)$ in which g carries *jump* FH singularities at the (pre/post) Fermi points $\{\theta_j\}$ with exponents $\beta_j \neq 0$; expanding $\log f_t(\theta) = \sum_{m\neq 0} c_m e^{im\theta}$ gives $c_m^{(j)} = \frac{i\beta_j}{m} e^{-im\theta_j}$ and, by Heine–Szegő plus the covariance $\mathbb{E}|\mathrm{Tr}\, U^m|^2 = m$ [22, 24], the linear statistic variance scales as

$$\mathrm{Var}\big(\log G_N\big) = \sum_{m\leq N} m|c_m|^2 = \Big(\sum_j \beta_j^2\Big) \log N + O(1),$$

i.e. FH enhances fluctuations to a $\sqrt{\log N}$ scale while the *Gaussian core* from the smooth part (Szegő) remains.

For quenches with pairing (XY or Ising case, to be further discussed below), the object is a *block Toeplitz* determinant with 2×2 symbol $\Phi(k,t) = (\mathbb{I} - \mathcal{C}_0(k)) + \mathcal{C}_0(k) e^{-i\mathcal{H}_1(k)t}$; quenches across criticality generate *root* FH exponents at $k = 0, \pi$ ($\alpha = \pm\frac{1}{2}$) and possibly jumps ($\beta \neq 0$), producing deterministic power-law/oscillatory factors on top of the same CLT-driven Gaussian core for the finite Fourier harmonics of $\log\det\Phi$. Thus, in PBC as in the OBC case that will be discussed elsewhere, the trace-statistics mechanism gives the Gaussian center, while the FH data dictate the normalization and non-Gaussian tails.

Non-Gaussian features arise when at least one hypothesis behind Szegő/CLT degrades. In particular:

(A) Too many modes / slow decay. If the effective number of harmonics m grows with N and (a_r, b_r) decay slowly, higher cumulants need not vanish. A sufficient condition for a stable Gaussian core is $\sum_{r\geq 1} r(a_r^2 + b_r^2) < \infty$ together with admissible growth $m(N)$ (cf. Johansson–Lambert). At the borderline $a_r, b_r \sim r^{-1}$ the variance diverges logarithmically and the correct scaling changes (non-Gaussian-looking histograms at fixed normalization). That is, a very large number of Fourier modes or a dispersion whose Fourier coefficients a_r, b_r decay too slowly (borderline $1/r$ behavior or weaker) can invalidate the central limit approximation. See the tail curvature in detrended Q–Q, Fig. 2(c), the kurtosis-scaling diagnostic in Fig. 4c, and the direct CUE variance check in Fig. 4d.

(B) Fisher–Hartwig singularities. In variants where the symbol develops genuine FH factors, which are nonanalytic features in the symbol such as jump discontinuities or zeros [27], $\log D_N$ contains non-quadratic terms (power-laws/oscillations), and persistent non-Gaussian edges appear. This includes certain OBC reductions and matrix-model weights related to Painlevé V [18, 40, 41] (to be discussed elsewhere).

(C) Critical quenches / DQPTs. Crossing a quantum critical manifold brings Fisher zeros of $G(t)$ close to the real axis [8, 9]. Indeed, if the quench crosses a phase transition, the Loschmidt amplitude develops Fisher zeros arbitrarily close to the real axis, which translates into $\chi_W(u)$ acquiring singularities for real u (non-analytic behavior in time). This will thicken the tails of $P(W)$. Persistent edge features (cf. discussion here; qualitative signature matches the residuals in Fig. 1(c,f)).

(D) Finite-N effects. For moderate N or weight on medium/large r, the third/fourth cumulants—which are $o(1)$ asymptotically—are visible. Edgeworth corrections from the cumulant method would capture these tails. Notice how small residual deviations are confined to extremes, Fig. 1(c,f).

1 Connection to entanglement: Fisher–Hartwig singularities as a common origin

The Toeplitz/Fisher–Hartwig (FH) mechanisms behind deviations from the strong Szegő regime also underlie a familiar phenomenon in free-fermion many-body physics: *anomalous scaling of entanglement entropy.* Beyond a formal analogy, this yields a simple and experimentally testable qualitative implication in quadratic chains: *parameter regimes that display anomalous entanglement scaling (e.g. area-law violations or logarithmic growth) are natural candidates for persistent deviations from a finite-variance Gaussian work core under sudden quenches, and vice versa.* The quantitative map between entanglement measures and work cumulants is model dependent, but the *trigger*—the regularity class of the underlying Toeplitz symbols—is shared.

Entanglement and Toeplitz determinants. For the ground state of a translation-invariant free-fermion chain, the Rényi entanglement entropy of a contiguous block of ℓ sites is determined by the eigenvalues of the $\ell \times \ell$ restricted correlation matrix $[C_\ell]_{jk} = \langle c_j^\dagger c_k \rangle$, $j, k \in A$. Under PBC, C_ℓ is a Toeplitz matrix with symbol given by the momentum-space occupation function $n(k)$. For paired/BdG chains the reduced covariance matrix is block-Toeplitz, but the same Toeplitz asymptotic machinery applies. The Rényi entropy $S_q(\ell)$ can be expressed in terms of a Toeplitz determinant $D_\ell(f_q)$ with a symbol f_q built from $n(k)$ and the Rényi index q [36]. When $n(k)$ has jump discontinuities (as at Fermi points of a gapless chain), f_q acquires FH singularities. FH asymptotics then give the celebrated logarithmic scaling $S_q(\ell) \sim c_q \log \ell$, with a prefactor fixed by the FH exponents [36, 42]. In gapped short-range chains, $n(k)$ is smooth (Szegő-regular) and entanglement obeys an area law.

Work statistics side: the same diagnostic. In our work problem, the determinant is different—it is the Toeplitz/block-Toeplitz determinant for the Loschmidt amplitude (or χ_W), with symbol built from the post-quench single-particle data $\varepsilon_f(k)$ together with the Bogoliubov angles encoding the initial Gaussian state [18, 19]. Nevertheless, the same classification governs the fluctuation regime:

- If the symbol is Szegő-regular (Fourier coefficients decay fast enough), the multivariate trace CLT yields a sharp Gaussian work core with finite variance $\mathrm{Var}(W) = \frac{1}{2}\sum_r r(a_r^2 + b_r^2)$.
- If the symbol develops FH structure (jumps/roots/edge exponents) or the Fourier tail is too slow to satisfy (48), then the finite-variance CLT regime is lost and FH and finite-N corrections produce nonstandard scaling and visible deviations from a purely Gaussian description.

Thus the Szegő/FH regularity class functions as a unified diagnostic for "generic" versus "anomalous" behavior in both entanglement and work statistics.

Concrete anchor: the XY chain (gapped vs critical). This parallel can be made vivid within the same model already used below. For quenches within a gapped phase of the XY chain, the ground-state symbol entering C_ℓ is smooth (area-law entanglement), while the post-quench dispersion and Bogoliubov data are analytic, leading to exponentially decaying Fourier modes and an extremely sharp Gaussian work core (Sec. 2). By contrast, at (or across) a quantum critical line the correlation-matrix symbol acquires non-analytic structure at $k = 0, \pi$ that yields logarithmic entanglement [36, 42], and the Loschmidt block symbol develops FH-type root/jump singularities at the same special points, producing logarithmic/power-law factors and enhanced non-Gaussian features in $P(W)$ (see the FH discussion in Sec. V and the XY analysis below).

Long-range chains: a bidirectional experimental prediction. The overlap is particularly concrete for power-law couplings $J_r \propto r^{-\alpha}$ in one dimension. In this setting, the strong long-range regime $\alpha \leq 1$ is known to generate algebraic singularities in the correlation-matrix symbol and to produce area-law violations—including logarithmic, fractal, and even volume-law scaling—even away from criticality [43–45]. In our framework, the *same regime* $\alpha \leq 1$ is precisely where the Szegő condition for the work/Loschmidt symbol fails (Eq. (51)), i.e. where the finite-variance Gaussian-core theorem no longer applies. We therefore expect that in free-fermion realizations of long-range chains, an observation of anomalous entanglement scaling is accompanied by persistent deviations from a simple Gaussian work distribution under sudden quenches, and conversely that pronounced non-Gaussian work features signal an entanglement anomaly in the same parameter regime.

Scope and caveats. The symbols governing entanglement ($f_q[n(k)]$) and work ($e^{iu\varepsilon_f(k)}$ dressed by Bogoliubov angles) are not identical, so FH exponents and prefactors need not match numerically. What our framework does provide is a unified, sharp diagnostic: the Toeplitz regularity class of the underlying single-particle data controls both entanglement and work statistics, and Toeplitz/FH asymptotics give a single language in which quantitative relations (inequalities or functional links between $S_q(\ell)$ and work cumulants) can be formulated and tested.

Model and diagonalization. After the Jordan–Wigner map, the translation-invariant XY chain in a transverse field is a quadratic (BdG) fermion system [29–31] with

$$\mathcal{H}(k) = \mathbf{d}(k)\cdot\boldsymbol{\sigma}, \tag{54}$$

$$\mathbf{d}(k) = (\gamma \sin k,\, 0,\, h - \cos k), \tag{55}$$

$$\varepsilon(k) = 2|\mathbf{d}(k)| = 2\sqrt{(h-\cos k)^2 + (\gamma \sin k)^2}. \tag{56}$$

The *Bogoliubov angle* is

$$\tan(2\theta_k) = \frac{\gamma \sin k}{h - \cos k}, \tag{57}$$

so that a rotation by $2\theta_k$ in Nambu space diagonalizes $\mathcal{H}(k)$.

Loschmidt amplitude and work characteristic function. For a sudden quench $(\gamma_i, h_i) \to (\gamma_f, h_f)$ from the ground state of H_i (a BCS vacuum), the Loschmidt amplitude and TPM work characteristic function are

$$G(t) := \langle \psi_0 | e^{-iH_f t} | \psi_0 \rangle, \qquad \chi_W(u) = e^{-iuE_0}\, G(-u), \tag{58}$$

with $H_i|\psi_0\rangle = E_0|\psi_0\rangle$ [17]. The overlap factorizes over $(k, -k)$ and admits the standard product form [8]

$$G(t) = \prod_{k>0} \Big[\, \cos(\varepsilon_f(k)\, t) - \, i\, \sin(\varepsilon_f(k)\, t)\, \cos(2\Delta\theta_k) \,\Big], \tag{59}$$

where $\Delta\theta_k := \theta_k(\gamma_f, h_f) - \theta_k(\gamma_i, h_i)$.

From block Toeplitz to a scalar symbol. Equivalently, using the Klich reduction for Gaussian states [34] (and the HFB/Pfaffian viewpoint [32, 33]),

$$\Phi(k,t) = (\mathbb{I} - \mathcal{C}_0(k)) + \mathcal{C}_0(k)\, e^{-i\mathcal{H}_f(k)t}, \tag{60}$$

with $\mathcal{C}_0(k) = \frac{1}{2}(\mathbb{I} - \hat{\mathbf{d}}_i(k)\cdot\boldsymbol{\sigma})$.
For translation-invariant chains this is a *block Toeplitz* operator in real space; see e.g. the block-Toeplitz structure for XY correlation-matrix functionals in [36] and strong Szegő–type expansions for matrix symbols in [35]. A short 2×2 calculus gives

$$\det \Phi(k,t) = \Big[\, \cos(\varepsilon_f(k)t) - \, i\, \sin(\varepsilon_f(k)t)\, \cos(2\Delta\theta_k) \,\Big]^2. \tag{61}$$

Thus the *scalar* symbol

$$f_t(e^{ik}) = \cos(\varepsilon_f(k)t) - i\, \sin(\varepsilon_f(k)t)\, \cos(2\Delta\theta_k) \tag{62}$$

controls $G(t)$ via $G(t) = \sqrt{\det T_N(\Phi)} = D_N(f_t)^{1/2}$ and (59).

Small-u reduction to traces. Setting $u = t$ in $\chi_W(u) = e^{-iuE_0}G(-u)$ and expanding for small u,

$$f_u(e^{ik}) = 1 - i\,u\,\tilde{\varepsilon}(k) - \tfrac{1}{2}u^2\,\varepsilon_f(k)^2 + O(u^3), \tag{63}$$

$$\tilde{\varepsilon}(k) := \varepsilon_f(k)\,\cos(2\Delta\theta_k), \tag{64}$$

$$\log f_u(e^{ik}) = -i\,u\,\tilde{\varepsilon}(k) + O(u^2). \tag{65}$$

With the Fourier series

$$\tilde{\varepsilon}(k) = \sum_{m\in\mathbb{Z}} \tilde{\alpha}_m\,e^{imk}, \qquad \tilde{\alpha}_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{\varepsilon}(k)\,e^{-imk}\,dk, \tag{66}$$

Heine–Szegő [20] gives the linear-statistics form

$$\log\chi_W(u) = iu\sum_{m\geq 1}\Big(\tilde{\alpha}_m\,\operatorname{Tr}U^m + \overline{\tilde{\alpha}_m}\,\operatorname{Tr}U^{-m}\Big) + O(u^2), \tag{67}$$

so the selection rule and Wick factorization of Sec. IV apply verbatim with $\varepsilon(k) \mapsto \tilde{\varepsilon}(k)$, and the a_r, b_r of our general framework are now the real and imaginary parts of $2\tilde{\alpha}_r$. In particular,

$$\sigma^2 \;=\; 2\sum_{m\geq 1} m\,|\tilde{\alpha}_m|^2. \tag{68}$$

When non-Gaussian tails appear. If the quench does *not* cross a critical line, then $k \mapsto f_t(e^{ik})$ is smooth for fixed t; $\tilde{\varepsilon}(k)$ is analytic and the Fourier coefficients $\tilde{\alpha}_m$ decay exponentially, so the central $P(w)$ is Gaussian with $O(1)$ width. If the quench *does* cross a critical line, $\Delta\theta_k$ acquires a endpoint jump at $k = 0$ or π, f_t becomes piecewise smooth (FH type), and Toeplitz FH asymptotics yield deterministic power-law/oscillatory corrections [46] (visible non-Gaussian tails).

XX limit For $\gamma_i = \gamma_f = 0$ one has $\theta_k \equiv 0$, hence $\Delta\theta_k = 0$ and (59) collapses to

$$G(t) = \prod_{k>0} e^{-i\,\varepsilon_f(k)\,t}, \tag{69}$$

the translation-invariant, number-conserving XX limit (cf. [29, 30]). This should not be confused with the *domain-wall XX* protocol analyzed earlier, where $G_N(t) = D_N(e^{-it\,\varepsilon})$ is a Toeplitz determinant built from the restricted propagator on a spatial block. Here $\varepsilon_f(k) = 2\,|h_f - \cos k|$ is the BdG (positive) branch. If one prefers the Jordan–Wigner convention $\xi_f(k) = 2(\cos k - h_f)$, the same reduction is obtained by keeping the piecewise sign in $\cos(2\Delta\theta_k)$, which amounts to writing $e^{-i\,|\xi_f|t}$ together with jump phases at the Fermi points (FH language). This change of convention does not affect the Gaussian-core variance nor our diagnostics.

3 Quick summary of Gaussianity and tail behavior

- *Gapped BdG (analytic ε):* sharp Gaussian core; pairing just introduces b_r in the variance $\frac{1}{2}\sum r(a_r^2 + b_r^2)$.
- *Critical (no FH):* core Gaussian; tails show visible but shrinking non-Gaussian corrections.

- *FH/jumps/zeros or borderline decay:* persistent non-Gaussian structures (edges/oscillations); the correct asymptotics are beyond strong Szegő.
- *Large m or slow decay of (a_r, b_r):* the fixed-normalization histogram looks non-Gaussian; after the right rescaling a Gaussian limit often returns.

4 Physical implications of non-Gaussianity

The preceding analysis identifies when the work distribution is Gaussian and when it is not, but one may ask: *what are the physical consequences of non-Gaussian work statistics?* We briefly discuss three aspects that connect the Gaussian/non-Gaussian distinction to observable thermodynamic behavior.

Fluctuation theorems and rare-event sampling. The Jarzynski equality and the Crooks fluctuation theorem [1, 2] relate the work distribution to equilibrium free-energy differences and hold regardless of whether $P(W)$ is Gaussian. However, the *practical* use of these relations—estimating ΔF from finitely many realizations—depends sensitively on the tail structure of $P(W)$. The Jarzynski estimator $\langle e^{-\beta W}\rangle$ is dominated by rare trajectories in the left tail (negative-work events), and its convergence is controlled by the weight in those tails. When $P(W)$ is Gaussian, the tail probabilities are exponentially suppressed and the estimator converges at a rate determined by the mean and variance alone. Non-Gaussian tails—whether heavier (e.g., algebraic, as in FH regimes) or lighter (e.g., bounded support)—change this convergence behavior: heavy tails can make the estimator converge more slowly (requiring exponentially more samples), while truncated tails can improve it. The mechanisms identified in this work—slow Fourier decay, FH singularities, Fisher zeros—provide a concrete, model-level understanding of when and why such sampling difficulties arise in sudden-quench protocols.

Irreversible work and higher cumulants. For a sudden quench from an eigenstate, the average work is $\langle W\rangle = \langle\psi_0|H_f|\psi_0\rangle - E_0$. More generally, when the initial state is thermal at inverse temperature β, the irreversible (excess) work $W_{\rm irr} = \langle W\rangle - \Delta F$ is related to the full work distribution via the Jarzynski equality [1, 2]. When $P(W)$ is Gaussian, $W_{\rm irr}$ is fully determined by the first two cumulants (mean and variance); specifically, in the Gaussian thermal-state case, $W_{\rm irr} = \beta\ \mathrm{Var}(W)/2$ [2]. When $P(W)$ develops non-Gaussian features, higher cumulants contribute and this simple relation breaks down: the cumulant expansion of $W_{\rm irr}$ receives corrections from the third cumulant (skewness), fourth cumulant (kurtosis), and beyond. Explicitly, expanding the Jarzynski equality in cumulants gives $W_{\rm irr} = \frac{\beta}{2}\kappa_2 - \frac{\beta^2}{6}\kappa_3 + \frac{\beta^3}{24}\kappa_4 - \cdots$, so the direction of the effect is set by the persistent cumulants: positive skewness *reduces* the excess work relative to the Gaussian value, while positive excess kurtosis *enhances* it. In the present framework, the cumulant structure of $P(W)$ is directly controlled by the Toeplitz/RMT machinery: the strong Szegő theorem guarantees that $\kappa_p \to 0$ for $p \geq 3$ in the Gaussian regime, while Fisher–Hartwig corrections or slow Fourier decay generate *persistent* higher cumulants. Thus, the same mechanisms that produce non-Gaussian tails also modify the relation between work fluctuations and irreversibility, regardless of whether the initial state is an eigenstate or a thermal mixture.

Fluctuation bounds and signal-to-noise. Variance-based precision–dissipation bounds provide constraints on how large work fluctuations can be for a given amount of irreversibility. In particular, thermodynamic uncertainty relations (TURs) bound the signal-to-noise ratio of currents in terms of entropy production in classical and quantum Markovian

settings [47, 48]. For closed, unitary protocols and TPM work statistics the status of TUR-type bounds is more subtle, but an important general message remains: second-moment information (mean/variance) captures the full fluctuation structure only when $P(W)$ is close to Gaussian. When the work distribution has persistent non-Gaussian features, higher cumulants carry independent information not captured by variance-only bounds. The taxonomy developed here—distinguishing slow-decay, FH, and DQPT mechanisms—thus provides a systematic way to identify quench protocols where rare events and higher cumulants are expected to matter, and motivates cumulant-adapted bounds for such regimes.

Non-Gaussianity as a diagnostic. Finally, it is worth noting that the physical regimes producing non-Gaussian work statistics—long-range interactions, proximity to quantum critical points, Fisher–Hartwig singularities—are precisely those where quantum correlations and many-body effects are most pronounced. This suggests that the degree of non-Gaussianity in $P(W)$ (quantified, for instance, by the ratio of higher cumulants to appropriate powers of the variance, or by the skewness proxy $T_m/S_m^{3/2}$ defined in Sec. VI) could serve as a practical, experimentally accessible diagnostic for the complexity or "quantumness" of a quench protocol. Interferometric reconstructions of $\chi_W(u)$ [4, 6] already provide access to the full characteristic function, and therefore in principle to all cumulants; our results identify the microscopic parameters (interaction range, symbol regularity) that control which cumulants are visible.

VI Numerical diagnostics

The preceding sections make concrete predictions: a Gaussian work distribution with a specific variance formula, asymptotic independence of trace components, and identifiable non-Gaussian corrections in specific regimes. This section tests those predictions against Monte Carlo sampling of the CUE linear statistic, using four complementary diagnostics: (i) histogram with Gaussian overlay; (ii) normal quantile-quantile plots (Q–Q plots; defined below) (and, when informative, detrended Q–Q); (iii) scatter plots of trace components with theory ellipses, including standardized versions; and (iv) variance and excess kurtosis versus N with error bars. The full set of panels is shown after the references.

Notation and basic definitions. Throughout this section, N denotes the matrix dimension of $U(N)$ in the unitary matrix model; m is the number of active Fourier harmonics in the dispersion (equivalently, the neighbor range retained in the chain); and n is the number of Monte–Carlo realizations used per panel (sample size behind each histogram or Q–Q plot). A *Q–Q (quantile–quantile) plot* compares the empirical quantiles of the (centered) work variable on the vertical axis to the corresponding theoretical normal quantiles on the horizontal axis; alignment with the line $y = x$ indicates a Gaussian core, while systematic curvature near the extremes reveals tail deviations. Histograms use the Freedman–Diaconis bin width [49]

$$h \;=\; \frac{2\,\mathrm{IQR}}{n^{1/3}}, \quad \mathrm{IQR} := Q_{0.75} - Q_{0.25},$$

where IQR is the *interquartile range* (the difference between the 75th and 25th percentiles).

Sampling and rendering choices. To display the Gaussian core with low sampling noise, most panels are generated with a surrogate.

We use a Gaussian surrogate that preserves the exact first-/second-moment structure (Sec. IV). To generate surrogate data, we replace the Haar matrix U by independent complex Gaussian variables: $\operatorname{Tr} U^r$ is sampled from $\mathcal{CN}(0, r)$ (a complex normal with mean 0 and variance r) for each r, with different r samples taken independent. This preserves the exact first- and second-moment structure but assumes asymptotic independence outright. This means that we draw:

$$\operatorname{Tr} U^r \overset{d}{\approx} \mathcal{CN}(0, r), \tag{70}$$

independently over the finite set of harmonics and form $W = \sum_r (a_r \Re \operatorname{Tr} U^r + b_r \Im \operatorname{Tr} U^r)$ (cf. (47)). This reproduces the central limit covariance exactly for fixed r and large N, at a fraction of the cost of explicit Haar draws. As a cross-check, we include "Haar vs surrogate" comparisons at $N = 80$, $n = 300$, showing near-identical bulk behavior. See Fig. 3 for a Haar vs surrogate comparison. Histograms use the Freedman–Diaconis bin width $h = 2\,\mathrm{IQR}/n^{1/3}$ [49] and overlay the fitted Gaussian $\mathcal{N}(0, \hat{\sigma}^2)$. Q–Q plots use the classic "points + $y = x$" style; we also provide *detrended* Q–Q (residuals vs theoretical quantiles), which makes small tail deviations easier to see.

What the panels show and how to read them.

1. **Histogram/Q–Q pairs.** High-n panels (histogram with Gaussian overlay; Q–Q with $y = x$ line) display the Gaussian core and isolate any tail curvature.

2. **Scatter plots.** Clouds of $(\Re \operatorname{Tr} U^r, \Re \operatorname{Tr} U^s)$ include the *theory 95% ellipse*; standardized versions divide each axis by $\sqrt{m/2}$ so the contour is a circle, making different (r, s) pairs comparable.

3. **Variance vs N.** With fixed harmonics, the CLT variance $\mathrm{Var}(W) = \frac{1}{2}\sum_r r(a_r^2 + b_r^2)$ is N-independent; the plot plateaus at this constant (dashed line) with SE bars $\hat{\sigma}^2\sqrt{2/(n-1)}$.

4. **Excess kurtosis vs N.** The trend toward 0 (with error bars $\sqrt{24/n}$) matches the CLT/Wick picture; residual scatter reflects the estimator's finite-n noise and small negative bias $\approx 6/(n+1)$. Panels (c)–(d) in Fig. 4 illustrate the long-range threshold: panel (c) uses a long-range correlated non-Gaussian proxy field with $\langle y_0 y_r \rangle_c \sim r^{-\alpha}$, where the excess kurtosis decays toward 0 for $\alpha > 1$ but remains finite for $\alpha < 1$; panel (d) then checks directly in the CUE linear statistic that the variance follows the Szegő sum and is finite for $\alpha > 1$ but grows with N for $\alpha < 1$.

Compact diagnostics that tie figures to theory. For a given choice of weights, define

$$S_m = \sum_{r \le m} r\,(a_r^2 + b_r^2), \qquad T_m = \sum_{r \le m} r^{3/2}\,(|a_r|^3 + |b_r|^3).$$

Then $\mathrm{Var}(W) = \frac{1}{2} S_m$ (prediction for the dashed line in Fig. 4a), while the dimensionless skewness proxy $T_m/S_m^{3/2}$ tracks the relative size of the cubic cumulant: it is small (and shrinks with m) when the Fourier weights decay fast (analytic symbols, within-phase quenches), and grows in borderline/FH regimes. If all weight is on a single mode, this ratio corresponds exactly to the skewness of that mode's distribution. This explains why the Q–Q bulk is straight and why any curvature, when present, is confined to the extremes.

Figure 4(b) shows how the kurtosis trends to the Gaussian value as N increases, confirming our CLT-based prediction that $P(W)$ becomes Gaussian in the large-N limit.

Summary. Across the histogram/Q–Q pairs (Figures 1,2 and 3), and variance/kurtosis trends (Figure 4) and scatter panels (Figure 5) the data agree with the random-matrix CLT for traces: a *Gaussian bulk* with the variance set by the quadratic form $\sum_r r(a_r^2 + b_r^2)$ and *asymptotic independence* across different powers. Non-Gaussian features, when visible, appear where Szegő hypotheses are weakened (many harmonics, slow decay) or in settings with genuine FH structure (jumps/zeros/edge singularities), in line with the discussion in Sec. V.

VII Conclusions and outlook

We mapped the work distribution in a sudden quench onto a random matrix problem (the traces of Haar unitaries). The multivariate central limit theorem then predicts a Gaussian core to the distribution for a large system size, N, with a variance fully determined by the quench parameters (10). More precisely, this is determined by the Fourier components of the post-quench dispersion. The coefficients a_m and b_m come from expanding the single-particle dispersion $\varepsilon(k)$ into cosine and sine harmonics, as given in equation (8). Note that, in contrast to summing N independent random variables, the large-N limit of the unitary group plays the role of the thermodynamic limit here, yielding an intrinsic variance of order m rather than vanishing with N.

This is a direct physical consequence of the fact that the central–limit behavior we exploit here is of a *random–matrix* type rather than the textbook i.i.d. kind. For Haar $U(N)$, the random variables $\operatorname{Tr} U^r$ have $O(1)$ variances (no $1/\sqrt{N}$ normalization), different powers become asymptotically independent by unitary orthogonality, and the relevant regularity condition is of Szegő type [24] (summability of $r(a_r^2 + b_r^2)$), not Lindeberg's. These features lead to error controls and growth windows (e.g. total-variation CLTs up to $m(N) = o(N^{2/3})$ and super-exponential accuracy if $m \ll \sqrt{N}$ [25]) that are quite unlike the Berry–Esseen picture for sums of i.i.d. variables; see also [22] for finite-N trace moments and orthogonality, [23] for the linear statistics CLTs, and the Gibbs lecture by Diaconis [50] for context.

We identify how deviations (non-Gaussian tails) arise when the conditions of the CLT are weakened (multiple slow-decaying modes, symbol singularities, etc.), and we confirm these predictions with both analytic results (Toeplitz determinant asymptotics) and numerical simulations. This Gaussian 'core' of the work distribution, with model-specific non-Gaussian tails, suggests clear experimental signatures – e.g. the prevalence of Gaussian statistics in sudden quenches unless the system is tuned to special conditions (long-range interactions or critical points) that produce visible deviations.

Therefore, a natural next step is to drive the system across quantum critical manifolds (or to open boundaries) and ask how the *tails* of the work distribution reorganize. In our mapping, the work characteristic function is the Loschmidt amplitude $\chi_W(u) = G(-u)$. When $G(t)$ develops *Fisher zeros* that approach the real time axis, the strip of analyticity of $\chi_W(u)$ narrows, and by standard Fourier arguments this slows the decay of $P(W)$—revealing structured, non-Gaussian tails. In parallel, when the (block)-Toeplitz symbol acquires *Fisher–Hartwig* features (jumps/zeros) at criticality or at hard edges (OBC), one expects robust algebraic/oscillatory fingerprints in the tails while the *bulk* remains close to Gaussian. Charting these effects side-by-side—within-phase vs across-critical quenches—would provide an experimentally testable prediction from our Gaussian core with FH/DQPT-controlled tails picture. In particular, we already have computations on the open boundary case with the large u (tails) asymptotics of the Toeplitz/Hankel determinant evaluated through Painlevé V: the expansion contains explicit power-laws and oscillations that will give

non-Gaussian tails controlled by the nearest PV singularity to the real u axis. This will be discussed elsewhere.

Likewise, by tuning the $\{a_r\}$ and probing the large-deviation regime of the matrix model, one encounters the Gross–Witten–Wadia (GWW) transition [51, 52], a third-order phase transition in unitary matrix models whose realization in spin-chain Loschmidt amplitudes was identified in [18], together with its multicritical extensions [53]. These transitions reorganize the *tails* of $P(W)$ in universal ways without changing the near-Gaussian bulk. More broadly, the random-matrix perspective naturally suggests that the extreme-value statistics of the work distribution—i.e., the rare large-$|w|$ events that dominate Jarzynski sampling—should fall into Tracy–Widom universality classes [20, 54], with the specific class (soft edge, hard edge, or GWW-type) determined by whether the Toeplitz symbol has smooth endpoints, algebraic vanishing, or a critical saturation. Establishing this connection quantitatively would link quench thermodynamics to the well-charted landscape of random-matrix edge universality.

Beyond sudden quenches and TPM work. The exact CUE/Toeplitz mapping developed here exploits the sudden-quench identity $\chi_W(u) = e^{-iuE_0}G(-u)$ together with translation invariance, which yields a (block-)Toeplitz determinant controlled by single-particle data. For more general *finite-time* protocols in quadratic systems, $\chi_W(u)$ can still be written in closed form as a determinant/Pfaffian of a one-body operator built from the single-particle evolution (Klich-type formulas) [34], and recent work has shown that finite-time driving can lead to non-Gaussian work statistics in its own right [28]. Alternative (non-TPM) notions of quantum work correspond to different generating functions [1, 2] and are not addressed here. For homogeneous drives the kernels in these determinant representations remain Toeplitz/block-Toeplitz in real space, suggesting that Szegő/Fisher–Hartwig regularity and trace-CLT ideas may continue to control the emergence of a Gaussian core whenever the effective symbol stays sufficiently regular, while protocols that generate non-analytic structure should produce persistent higher-cumulant/tail effects.

Author contributions

Miguel Tierz carried out the original research and wrote the manuscript. Large language models (Claude and ChatGPT) were used during the revision to edit parts of the text, check the bibliography and prepare the LaTeX files. The author takes responsibility for the content of the work.

Acknowledgements

We thank Gregory Schehr for correspondence.

References

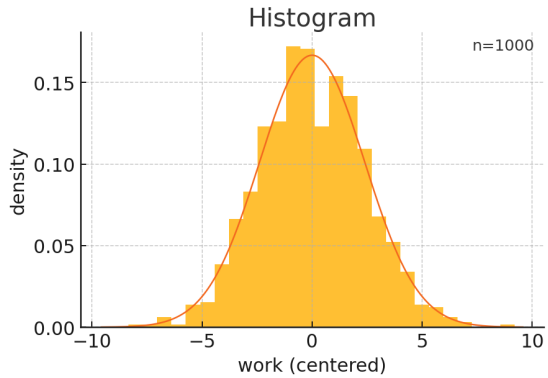
[1] P. Talkner, E. Lutz, and P. Hänggi, Fluctuation theorems: Work is not an observable, Physical Review E **75**, 050102(R) (2007).

[2] M. Campisi, P. Hänggi, and P. Talkner, Colloquium: Quantum fluctuation relations: Foundations and applications, Reviews of Modern Physics **83**, 771 (2011).

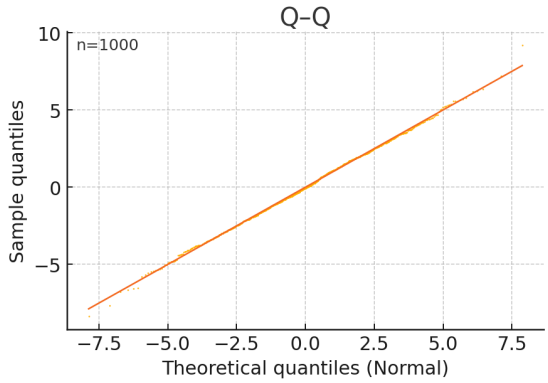
[3] S. Deffner and S. Campbell, *Quantum Thermodynamics: An introduction to the thermodynamics of quantum information* (Morgan & Claypool Publishers, 2019).

[4] R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, and V. Vedral, Extracting quantum work statistics and fluctuation theorems by single-qubit interferometry, Physical Review Letters **110**, 230601 (2013).
[5] L. Mazzola, G. De Chiara, and M. Paternostro, Measuring the characteristic function of the work distribution, Physical Review Letters **110**, 230602 (2013).
[6] T. B. Batalhão, A. M. Souza, L. Mazzola, R. Auccaise, R. S. Sarthour, I. S. Oliveira, J. Goold, G. De Chiara, M. Paternostro, and R. M. Serra, Experimental reconstruction of work distribution and study of fluctuation relations in a closed quantum system, Physical Review Letters **113**, 140601 (2014).
[7] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, Dynamics of Loschmidt echoes and fidelity decay, Physics Reports **435**, 33 (2006).
[8] M. Heyl, A. Polkovnikov, and S. Kehrein, Dynamical quantum phase transitions in the transverse-field Ising model, Physical Review Letters **110**, 135704 (2013).
[9] M. Heyl, Dynamical quantum phase transitions: a review, Reports on Progress in Physics **81**, 054001 (2018).
[10] A. Gambassi and A. Silva, Statistics of the work in quantum quenches, universality and the critical Casimir effect (2011), arXiv:1106.2671 [cond-mat.stat-mech].
[11] A. Gambassi and A. Silva, Large deviations and universality in quantum quenches, Phys. Rev. Lett. **109**, 250602 (2012).
[12] J. Goold, F. Plastina, A. Gambassi, and A. Silva, The role of quantum work statistics in many-body physics, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer, Cham, 2019) pp. 317–336, arXiv:1804.02805 [cond-mat.stat-mech].
[13] S. Gherardini, L. Buffoni, and N. Defenu, Universal defects statistics with strong long-range interactions, Phys. Rev. Lett. **133**, 113401 (2024).
[14] A. Solfanelli and N. Defenu, Universal work statistics in long-range interacting quantum systems, Phys. Rev. Lett. **134**, 030402 (2025).
[15] N. Defenu, T. Donner, T. Macrì, G. Pagano, S. Ruffo, and A. Trombettoni, Long-range interacting quantum systems, Rev. Mod. Phys. **95**, 035002 (2023).
[16] M. Esposito, U. Harbola, and S. Mukamel, Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems, Reviews of Modern Physics **81**, 1665 (2009).
[17] A. Silva, Statistics of the work done on a quantum critical system by quenching a control parameter, Physical Review Letters **101**, 120603 (2008).
[18] D. Pérez-García, L. Santilli, and M. Tierz, Dynamical quantum phase transitions from random matrix theory, Quantum **8**, 1271 (2024).
[19] D. Pérez-García, L. Santilli, and M. Tierz, Hawking-Page transition on a spin chain, Physical Review Research **6**, 033007 (2024).
[20] P. J. Forrester, *Log-gases and random matrices (LMS-34)* (Princeton university press, 2010).
[21] D. García-García and M. Tierz, Toeplitz minors and specializations of skew Schur polynomials, Journal of Combinatorial Theory, Series A **172**, 105201 (2020).
[22] P. Diaconis and M. Shahshahani, On the eigenvalues of random matrices, Journal of Applied Probability **31A**, 49 (1994).
[23] P. Diaconis and S. N. Evans, Linear functionals of eigenvalues of random matrices, Transactions of the American Mathematical Society **353**, 2615 (2001).
[24] K. Johansson, On random matrices from the compact classical groups, Annals of Mathematics **145**, 519 (1997).
[25] K. Johansson and G. Lambert, Multivariate normal approximation for traces of random unitary matrices, The Annals of Probability **49**, 2961 (2021).
[26] M. Duits and K. Johansson, Powers of large random unitary matrices and Toeplitz determinants, Transactions of the American Mathematical Society **362**, 1169 (2010).
[27] R. E. Hartwig and M. E. Fisher, Asymptotic behavior of Toeplitz matrices and determinants, Archive for Rational Mechanics and Analysis **32**, 190 (1969).
[28] K. Zawadzki, A. Kiely, G. T. Landi, and S. Campbell, Non-Gaussian work statistics at finite-time driving, Physical Review A **107**, 012209 (2023).
[29] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, Annals of Physics **16**, 407 (1961).

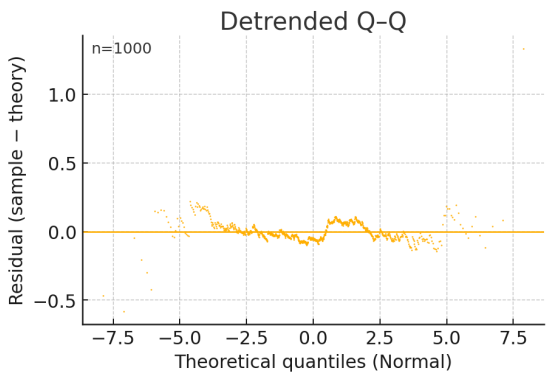
[30] P. Pfeuty, The one-dimensional Ising model with a transverse field, Annals of Physics **57**, 79 (1970).
[31] E. Barouch and B. M. McCoy, Statistical mechanics of the XY model. II. spin-correlation functions, Physical Review A **3**, 786 (1971).
[32] N. Onishi and S. Yoshida, Generator coordinate method applied to nuclei in the transition region, Nuclear Physics **80**, 367 (1966).
[33] L. M. Robledo, Sign of the overlap of Hartree–Fock–Bogoliubov wave functions, Physical Review C **79**, 021302 (2009).
[34] I. Klich, An elementary derivation of Levitov's formula, Quantum Noise in Mesoscopic Physics, 397 (2003).
[35] E. L. Basor and T. Ehrhardt, Asymptotics of determinants of Bessel operators, Communications in Mathematical Physics **234**, 491 (2003).
[36] B.-Q. Jin and V. E. Korepin, Quantum spin chain, Toeplitz determinants and the Fisher–Hartwig conjecture, Journal of Statistical Physics **116**, 79 (2004).
[37] D. Perez-Garcia and M. Tierz, Mapping between the Heisenberg XX spin chain and low-energy QCD, Physical Review X **4**, 021050 (2014).
[38] L. Santilli and M. Tierz, Phase transition in complex-time Loschmidt echo of short and long range spin chain, Journal of Statistical Mechanics: Theory and Experiment **2020**, 063102 (2020).
[39] E. M. Rains, High powers of random elements of compact Lie groups, Probability Theory and Related Fields **107**, 219 (1997).
[40] E. L. Basor, Y. Chen, and T. Ehrhardt, Painlevé V and time-dependent Jacobi polynomials, Journal of Physics A: Mathematical and Theoretical **43**, 015204 (2010).
[41] O. Lisovyy, H. Nagoya, and J. Roussillon, Irregular conformal blocks and connection formulae for Painlevé V functions, Journal of Mathematical Physics **59**, 091409 (2018), arXiv:1806.08344.
[42] F. Ares, J. G. Esteve, and F. Falceto, Rényi entanglement entropy in fermionic chains, Int. J. Geom. Methods Mod. Phys. **12**, 1560002 (2015).
[43] F. Ares, J. G. Esteve, F. Falceto, and A. R. de Queiroz, Entanglement in fermionic chains with finite-range coupling and broken symmetries, Phys. Rev. A **92**, 042334 (2015), arXiv:1506.06665 [quant-ph].
[44] F. Ares, J. G. Esteve, F. Falceto, and A. R. de Queiroz, Entanglement entropy in the long-range Kitaev chain, Phys. Rev. A **97**, 062301 (2018), arXiv:1801.07043 [quant-ph].
[45] A. Solfanelli, S. Ruffo, S. Succi, and N. Defenu, Logarithmic, fractal and volume-law entanglement in a Kitaev chain with long-range hopping and pairing, Journal of High Energy Physics **05**, 066 (2023), arXiv:2301.13231 [quant-ph].
[46] P. Deift, A. Its, and I. Krasovsky, Asymptotics of Toeplitz, Hankel, and Toeplitz+Hankel determinants with Fisher-Hartwig singularities, Annals of mathematics **174**, 1243 (2011).
[47] A. C. Barato and U. Seifert, Thermodynamic uncertainty relation for biomolecular processes, Phys. Rev. Lett. **114**, 158101 (2015), arXiv:1502.05944 [cond-mat.stat-mech].
[48] Y. Hasegawa and T. Van Vu, Fluctuation theorem uncertainty relation, Phys. Rev. Lett. **123**, 110602 (2019), arXiv:1902.06376 [cond-mat.stat-mech].
[49] D. Freedman and P. Diaconis, On the histogram as a density estimator: L_2 theory, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **57**, 453 (1981).
[50] P. Diaconis, Patterns in eigenvalues: the 70th Josiah Willard Gibbs lecture, Bulletin of the American Mathematical Society **40**, 155 (2003).
[51] D. J. Gross and E. Witten, Possible third-order phase transition in the large-N lattice gauge theory, Physical Review D **21**, 446 (1980).
[52] S. R. Wadia, $N = \infty$ phase transition in a class of exactly soluble model lattice gauge theories, Physics Letters B **93**, 403 (1980).
[53] P. Le Doussal, S. N. Majumdar, and G. Schehr, Multicritical edge statistics for the momenta of fermions in nonharmonic traps, Physical review letters **121**, 030603 (2018).
[54] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, Communications in Mathematical Physics **159**, 151 (1994).



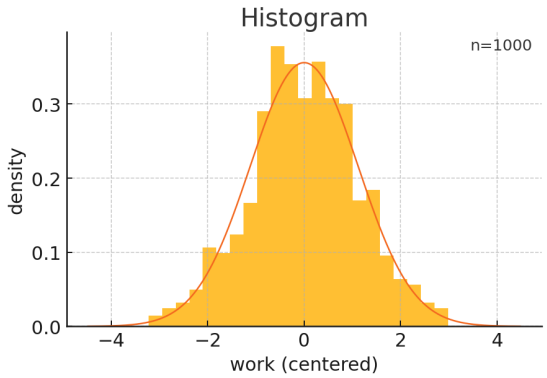
(a) 3-mode: Histogram



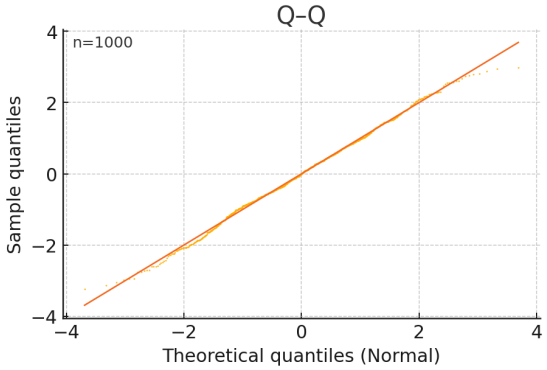
(b) 3-mode: Q–Q



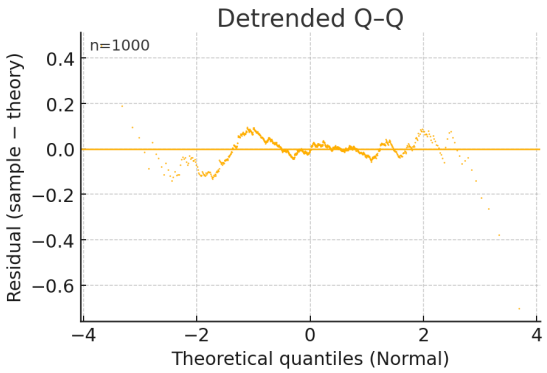
(c) 3-mode: Detrended Q–Q



(d) 5-mode even: Histogram

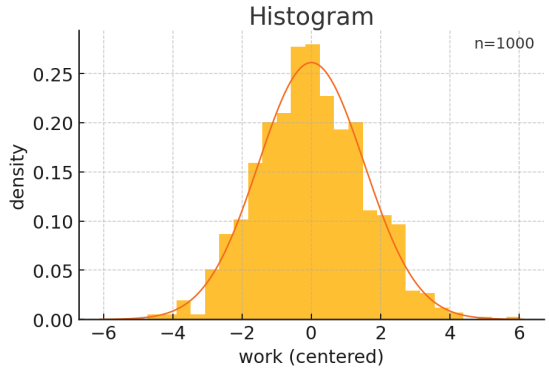


(e) 5-mode even: Q–Q

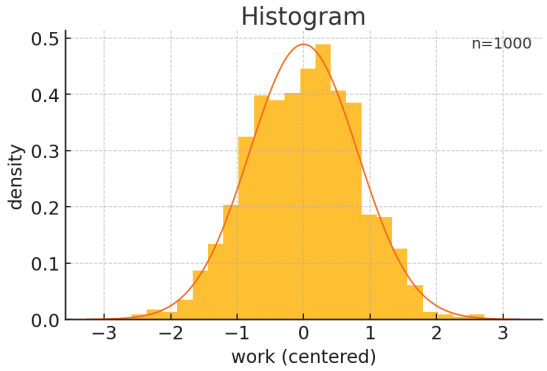


(f) 5-mode even: Detrended Q–Q

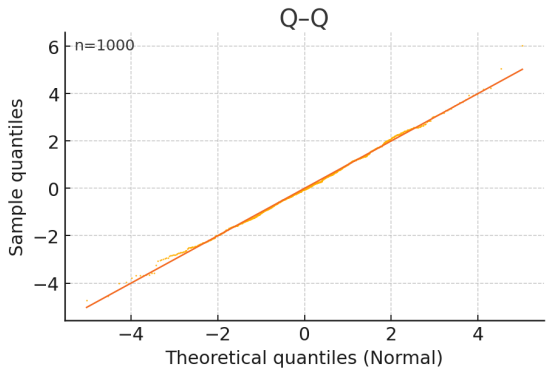
Figure 1: Side-by-side diagnostics for two families at $N = 80$, $n = 1000$ samples. Left: 3-mode model $\mathcal{M} = \{1, 2, 3\}$ with $\alpha_1 = 1.0, \alpha_2 = 0.7, \alpha_3 = 0.5$. Right: 5-mode even model $\mathcal{M} = \{2, 4, 6, 8, 10\}$, $a_m = e^{-0.4m}$. Each column shows the histogram with Gaussian overlay, the Q–Q plot (points + $y = x$), and the detrended Q–Q (residuals vs theory).



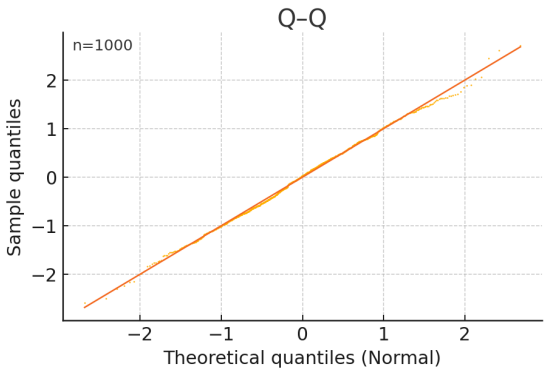
(a) Power-law: Histogram



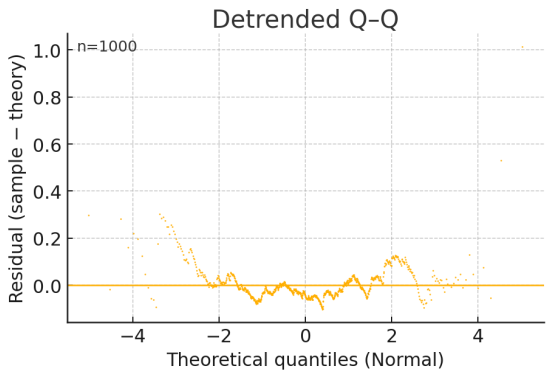
(d) Exponential even: Histogram



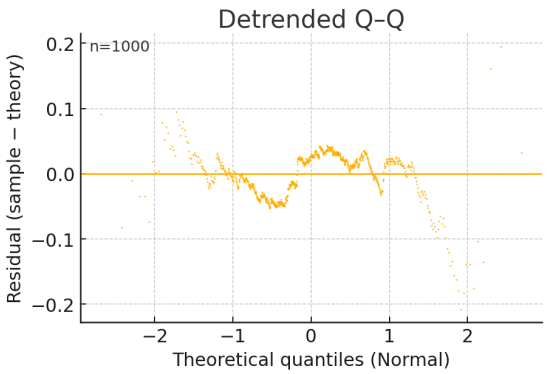
(b) Power-law: Q–Q



(e) Exponential even: Q–Q

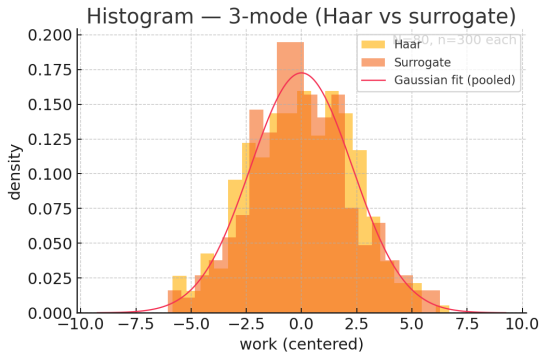


(c) Power-law: Detrended Q–Q

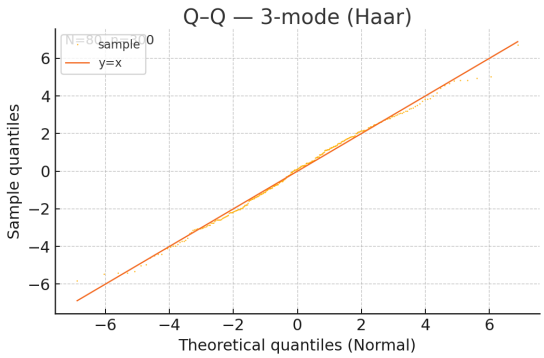


(f) Exponential even: Detrended Q–Q

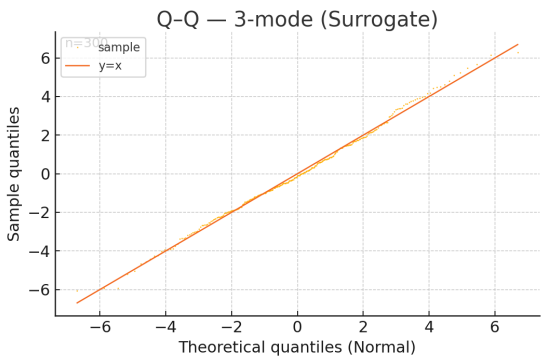
Figure 2: Diagnostics for power-law (left; $a_m = m^{-2}$) and exponential even (right; $a_m = e^{-0.5m}$) families at $N = 80$, $n = 1000$ samples.



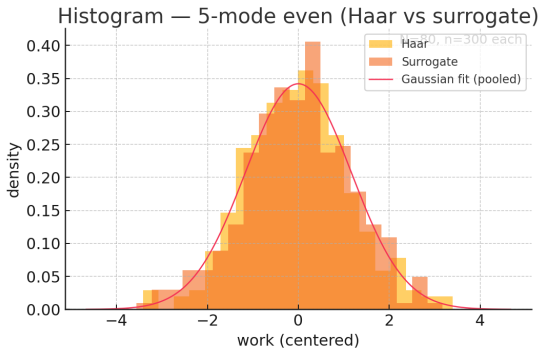
(a) Hist. (Haar vs sur.)



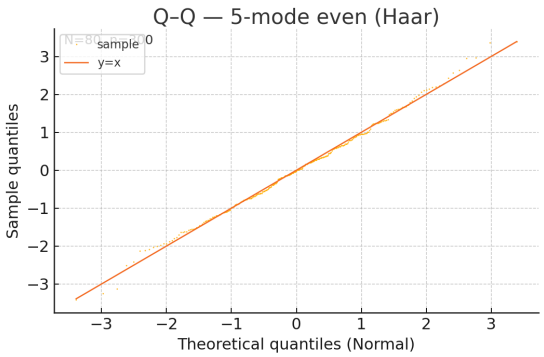
(b) Q–Q (Haar)



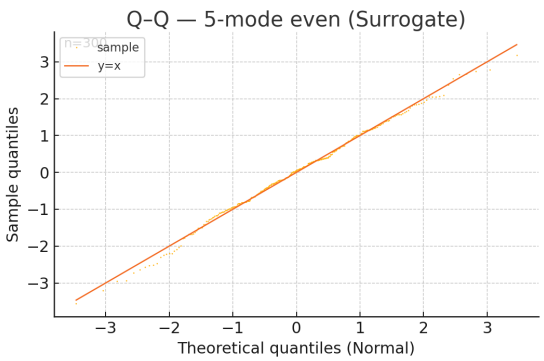
(c) Q–Q (Sur.)



(d) Hist. (Haar vs sur.)

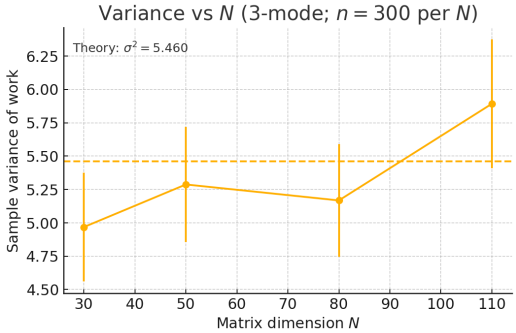


(e) Q–Q (Haar)

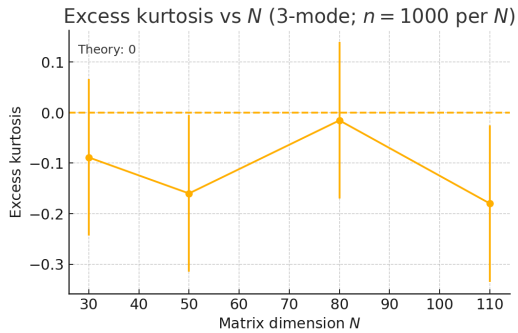


(f) Q–Q (Sur.)

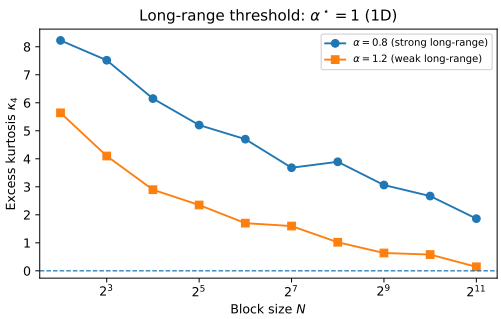
Figure 3: Surrogate vs exact Haar at $N = 80$, $n = 300$ per sample. Left: 3-mode $\mathcal{M} = \{1, 2, 3\}$, $\alpha_1 = 1.0, \alpha_2 = 0.7, \alpha_3 = 0.5$. Right: 5-mode even $\mathcal{M} = \{2, 4, 6, 8, 10\}$, $a_m = e^{-0.4m}$. Rows: (top) histogram overlays (Freedman–Diaconis bins; pooled Gaussian fit), (middle) Q–Q (Haar), (bottom) Q–Q (surrogate).



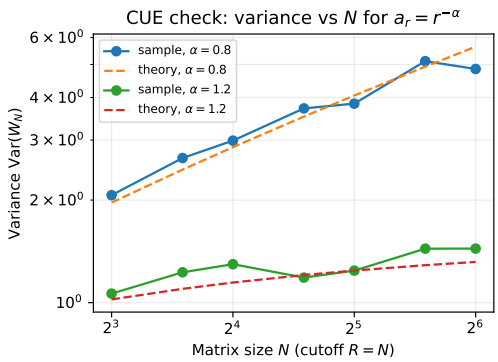
(a) Variance vs N.



(b) Excess kurtosis vs N.

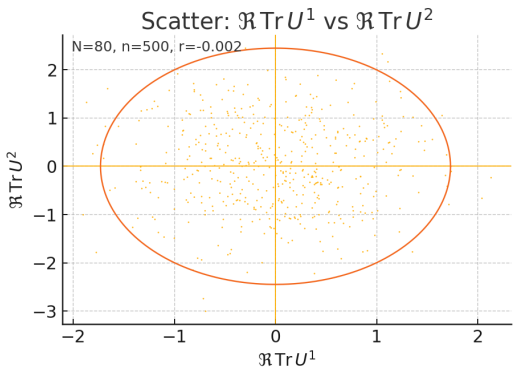


(c) Long-range proxy: kurtosis of block means.

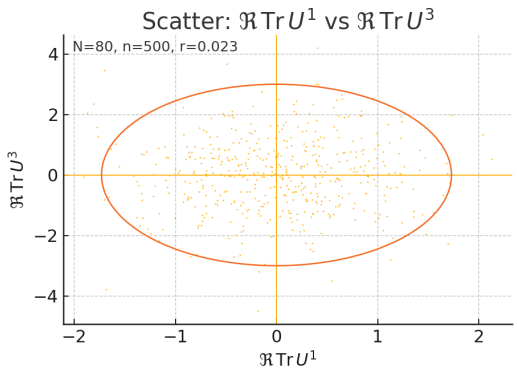


(d) CUE check: variance scaling for $a_r = r^{-\alpha}$.

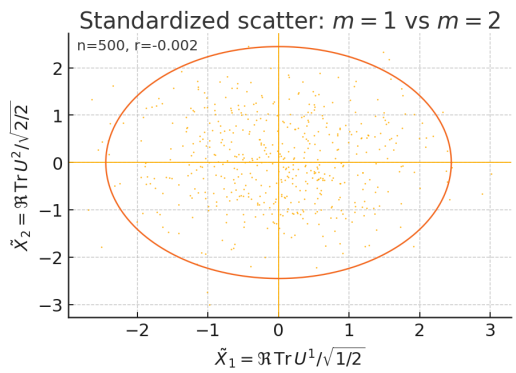
Figure 4: Scaling diagnostics. (a) Sample variance vs N for the 3-mode model $\mathcal{M} = \{1, 2, 3\}$, $\alpha_1 = 1.0$, $\alpha_2 = 0.7$, $\alpha_3 = 0.5$, with analytic prediction (dashed) and SE bars $\hat{\sigma}^2\sqrt{2/(n-1)}$ ($n = 300$ per N). (b) Excess kurtosis vs N for the same model, with error bars $\sqrt{24/n}$ ($n = 1000$ per N); dashed line marks the Gaussian limit. (c) Pedagogical CLT diagnostic for the long-range threshold $\alpha^* = d = 1$: excess kurtosis of standardized block means of a non-Gaussian field $y = x^2 - 1$ built from fractional Gaussian noise with connected correlations $\langle y_0 y_r \rangle_c \sim r^{-\alpha}$. For $\alpha = 0.8 < 1$ (strong long-range, $H = 0.8 > 3/4$) the kurtosis remains elevated at all block sizes, signalling persistent non-Gaussianity (Dobrushin–Major regime); for $\alpha = 1.2 > 1$ (weak long-range, $H = 0.7 < 3/4$) it decays toward zero, confirming CLT convergence. (d) Direct CUE check of the Szegő sum mechanism: for Haar $U(N)$, define the linear statistic $W_N = \sum_{r=1}^{N} a_r \, \Re \operatorname{Tr} U^r$ with $a_r = r^{-\alpha}$ (cutoff $R = N$). Points show the Monte–Carlo sample variance (here $n = 400$ per N); dashed curves show the exact prediction $\mathrm{Var}(W_N) = \frac{1}{2}\sum_{r=1}^{N} r\, a_r^2 = \frac{1}{2}\sum_{r=1}^{N} r^{1-2\alpha}$ from unitary orthogonality [22]. For $\alpha > 1$ the variance saturates (finite Szegő sum), while for $\alpha < 1$ it grows with N (divergent Szegő sum), matching the threshold discussed in Sec. V.



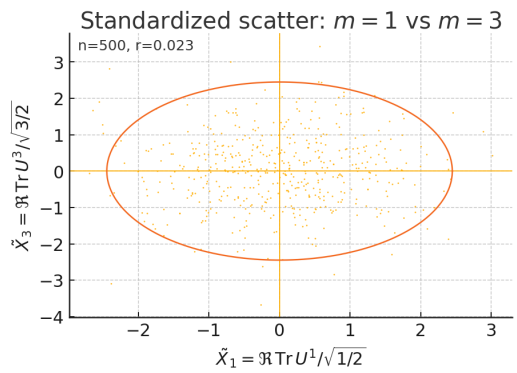
(a) $\Re\,\mathrm{Tr}\,U$ vs $\Re\,\mathrm{Tr}\,U^2$.



(b) $\Re\,\mathrm{Tr}\,U$ vs $\Re\,\mathrm{Tr}\,U^3$.



(c) Standardized: $m = 1$ vs $m = 2$.



(d) Standardized: $m = 1$ vs $m = 3$.

Figure 5: Scatter diagnostics with theory 95% contours (ellipse/circle). Top: raw coordinates show the $(\sqrt{m/2})$ scale on each axis; bottom: standardized coordinates divide by $\sqrt{m/2}$ so the 95% contour is a circle, making different (m, n) pairs directly comparable. Annotations in the panels report the sample correlation r (theory: $r = 0$); here $n = 500$ samples.