

2017-06-01

Gaussian ensemble for quantum integrable dynamics

Hyungwon Kim, Anatoli Polkovnikov, Emil A Yuzbashyan. 2017. "Gaussian ensemble for quantum integrable dynamics." *Annals of Physics*, Volume 381, pp. 107 - 120 (14).

<https://hdl.handle.net/2144/28533>

"Downloaded from OpenBU. Boston University's institutional repository."

Gaussian ensemble for quantum integrable dynamics

Hyungwon Kim,¹ Anatoli Polkovnikov,² and Emil A. Yuzbashyan¹

¹*Center for Materials Theory, Department of Physics and Astronomy,
Rutgers University, Piscataway, NJ 08854, USA*

²*Department of Physics, Boston University, Boston, MA 02215, USA*

We propose a Gaussian ensemble as a description of the long-time dynamics of isolated quantum integrable systems. Our approach extends the Generalized Gibbs Ensemble (GGE) by incorporating fluctuations of integrals of motion. It is asymptotically exact in the classical limit irrespective of the system size and, under appropriate conditions, in the thermodynamic limit irrespective of the value of Planck's constant. Moreover, it captures quantum corrections near the classical limit, finite size corrections near the thermodynamic limit, and is valid in the presence of non-local interactions. The Gaussian ensemble bridges the gap between classical integrable systems, where a generalized microcanonical ensemble is exact even for few degrees of freedom and GGE, which requires thermodynamic limit. We illustrate our results with examples of increasing complexity.

The far from equilibrium dynamics of isolated many-body systems with many nontrivial integrals of motion attracted considerable attention as such dynamics have been recently realized in several experiments [1–7]. In particular, it has been conjectured that the infinite time averages of various observables for a system evolving with a time-independent Hamiltonian \hat{H} are described by the Generalized Gibbs Ensemble (GGE) [8]:

$$\hat{\rho}_{\text{GGE}} = C e^{-\sum_i \beta_i \hat{H}_i}, \quad h_i \equiv \langle \hat{H}_i \rangle_0 = \text{tr}(\hat{\rho}_{\text{GGE}} \hat{H}_i), \quad (1)$$

where \hat{H}_i is a complete (in some yet unspecified sense) set of integrals of motion for \hat{H} , the second equation relates β_i to expectation values h_i of the integrals in the initial state, and C is a normalization constant. A key difficulty with quantum GGE stems from the absence of an accepted well-defined notion of quantum integrability. As a result, GGE is strictly speaking unfalsifiable. For example, it was initially shown to fail for the 1D XXZ spin chain [9–11], but later studies [12, 13] cured this by adding new integrals of motion in Eq. (1).

In contrast, classical integrability is well-defined [14]. Moreover, the microcanonical version of GGE – Generalized Microcanonical Ensemble (GME) is exact for a general classical integrable Hamiltonian $H(\mathbf{p}, \mathbf{q})$ [14, 15],

$$\rho(\mathbf{p}, \mathbf{q}) = C \prod_{k=1}^n \delta(H_k(\mathbf{p}, \mathbf{q}) - h_k), \quad (2)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T O(t) dt = \int d\mathbf{p} d\mathbf{q} O(\mathbf{p}, \mathbf{q}) \rho(\mathbf{p}, \mathbf{q}), \quad (3)$$

where $\mathbf{q} = (q_1, \dots, q_n)$ and $\mathbf{p} = (p_1, \dots, p_n)$ are the generalized coordinates and momenta and $H_k(\mathbf{p}, \mathbf{q})$ are the integrals of motion. The time evolution of any dynamical variable (observable) $O(t) \equiv O(\mathbf{p}(t), \mathbf{q}(t))$ is obtained by evolving with $H(\mathbf{p}, \mathbf{q})$ starting at $t = 0$ and h_k is the initial value of $H_k(\mathbf{p}, \mathbf{q})$. Note that unlike the microcanonical distribution for a nonintegrable Hamiltonian or GGE, Eq. (3) holds for any number of degrees of freedom n and arbitrary interactions, i.e. does not require

thermodynamic limit. In some sense, classical integrable dynamics are more ergodic, but in a restricted part of the phase-space cut out by the integrals of motion.

What can play the role or replace the microcanonical ensemble for a quantum integrable system with arbitrary particle number? More precisely, how to quantize Eqs. (2) and (3), i.e. what is a suitable density matrix $\hat{\rho}$ that turns into Eq. (2) in $\hbar \rightarrow 0$ limit? To what extent does it describe the quantum dynamics and how does it compare to GGE? These are the questions we address in this paper. We propose a multivariable Gaussian in \hat{H}_i as an appropriate $\hat{\rho}$ and show that it has several remarkable features. In particular, it provides quantum corrections to Eq. (2) at any particle number, i.e. a semiclassical approximation to the exact density matrix. It further yields finite size correction to GGE and is expected to work well in systems with long-range interactions, see also Fig. 1.

In the case of Gibbs or Generalized Gibbs distributions, one can simply replace $H(\mathbf{p}, \mathbf{q})$ or $H_k(\mathbf{p}, \mathbf{q})$ with the corresponding operators. This does not work for Eq. (2), because the average of a product of \hat{H}_i with so constructed density matrix is equal to the product of averages, which is not the case for a typical quantum state. Therefore, to reproduce various time averages, we need to broaden the delta-functions in Eq. (2).

It is natural to proceed by analogy with the usual quantum microcanonical ensemble and to replace the right hand side of Eq. (3) with an equal weight average over all eigenstates $|n\rangle$ of \hat{H}_i , $\hat{H}_i |n\rangle = E_i^{(n)} |n\rangle$, that have eigenvalues $E_i^{(n)}$ sufficiently close to quantum expectation values $h_i = \langle \hat{H}_i \rangle_0$ of the integrals in the initial state [16, 17]. The problem is that $E_i^{(n)}$ are generally discrete, while h_i can be anywhere in between. For example, integrals of motion for a collection of noninteracting fermions are their occupation numbers taking values 0 and 1, while their expectation values are arbitrary numbers ranging from 0 to 1. As a result, it is not always possible to find even a single eigenstate sufficiently close to the prescribed set of h_i . This problem can perhaps be

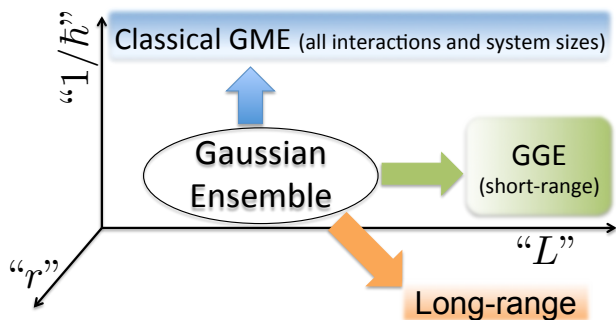


FIG. 1. (color online) Schematic diagram indicating limits where the Gaussian ensemble (4) is exact. By design it converges to the classical GME (2) when $\hbar \rightarrow 0$. The latter is a rigorous theorem in classical mechanics valid for a general integrable system at any number of degrees of freedom. The Gaussian ensemble contains GGE as a particular case and therefore holds whenever GGE does (e.g. for short-range interactions and system size $L \rightarrow \infty$). It is also valid for models where mean-field is exact due to large interaction range r .

resolved for a certain class of models in the thermodynamic limit through coarse-graining or suitable redefinition of the integrals [16–18]. However, in other integrable models there seems to be no simple, well-motivated remedy even at large particle number. We consider one such example below – interaction quenches in the BCS model, where there are no eigenstates reasonably close to h_i (see Fig. 4). In any case, since Eq. (3) is valid for any n , we are looking for a model-independent approach uniformly applicable at any particle number.

Instead, we propose broadening δ -functions in Eq. (2) directly. Specifically, we consider a Gaussian ensemble

$$\hat{\rho}_G = C \exp \left[- \sum_{ij} (\hat{H}_i - \mu_i) (\Sigma^{-1})_{ij} (\hat{H}_j - \mu_j) \right], \quad (4)$$

where parameters μ_i and Σ_{ij} are fixed by first and second moments of the initial conditions

$$h_i \equiv \langle \hat{H}_i \rangle_0 = \text{tr}(\hat{\rho}_G \hat{H}_i), \quad (5)$$

$$\langle \hat{H}_i \hat{H}_j \rangle_0 = \text{tr}(\hat{\rho}_G \hat{H}_i \hat{H}_j), \quad (6)$$

i.e. $\hat{\rho}_G$ by design reproduces exact first and second order correlation functions of conserved quantities. Note also that when \hat{H}_i have unbounded and continuous spectra, $\mu_i = \langle H_i \rangle_0$ and Σ is the covariance matrix, $\Sigma_{ij} = \langle \hat{H}_i \hat{H}_j \rangle_0 - \langle \hat{H}_i \rangle_0 \langle \hat{H}_j \rangle_0$.

It turns out that this ensemble has a number of significant advantages over various alternatives. First, it converges to Eq. (2) and therefore is exact in the classical limit for any number of degrees of freedom. Indeed, in this limit $\langle \hat{H}_i \hat{H}_j \rangle_0 = \langle \hat{H}_i \rangle_0 \langle \hat{H}_j \rangle_0$ and spectra of \hat{H}_i are continuous. Therefore, a product of δ -functions solves Eqs. (5) and (6) and, assuming the solution is unique, we arrive at the above statement. Moreover, we find that not only Eq. (4) is exact in $\hbar \rightarrow 0$ limit, but it also captures the leading quantum correction ($\propto \hbar$), at least in all

examples we studied. Higher order corrections, however, do not necessarily agree (see below). For this reason, we also expect the Gaussian ensemble to be exact for models with long-ranged interactions in the thermodynamic limit, such that the mean-field is exact effectively rendering such models classical.

Second, the Gaussian ensemble contains GGE as a particular case (when the coefficients at the bilinear in \hat{H}_i terms in the exponent in Eq. (4) vanish) and therefore always works as well as or better than GGE. Furthermore, we find that averages with $\hat{\rho}_G$ converge to exact infinite time averages faster than for GGE with increasing system size. For example, in a 1D fermion model on L sites we analyze below, the convergence is $1/L^2$ for the Gaussian ensemble and $1/L$ for GGE.

Now suppose the only integral is the total energy and the Eigenstate Thermalization Hypothesis (ETH) holds, i.e. expectation values of observables in an eigenstate are functions of the corresponding eigenenergy only [19]. Then, the Gaussian (unlike Gibbs or microcanonical) ensemble reproduces not only time-averaged expectations but also fluctuations of global and local observables, since these quantities depend on energy and its fluctuation only [20]. Fluctuations in turn enter various Kubo response coefficients, fluctuation-dissipation and other thermodynamic relations. Similarly, the multivariable Gaussian (4) correctly predicts expectations and fluctuations of observables satisfying the generalized ETH [16] in any integrable system. Lastly, Eq. (4) is model-independent and well-defined for any spectra of \hat{H}_i with no need to choose a measure of closeness of eigenstates to the prescribed set of expectation values h_i .

Of course, one can consider other representations of δ -functions in Eq. (2) and include more parameters to match higher order moments of $\{\hat{H}_i\}$. However, $\hat{\rho}_G$ is sufficient for many purposes as outlined above and as we will see in the examples below. Our proposal also does not fully resolve the difficulty with the notion of quantum integrability mentioned in the first paragraph. However, it in part bypasses this issue in models with smooth integrable classical limit, such as e.g. BCS-Gaudin model considered below. In such cases, we can rely on the well-defined concept of classical integrability and Eq. (4) makes good sense due to its connection to the rigorous result (3) for classical integrable systems. In the absence of a sound quantum definition, this route via classical integrability is useful for making unambiguous statements about general properties of quantum systems believed to be integrable. For example, Ref. 21 uses this approach to derive energy level statistics in such systems. In what follows we consider several specific examples to illustrate the above points.

Convergence to classical GME. Perhaps, the simplest quantum system with a transparent classical limit is the 1D harmonic oscillator in a coherent state. In this case, we find that the discrepancy between exact infinite time

averages and the Gaussian ensemble expectation values of observables is of order \hbar^2 [22]. For example,

$$\frac{\overline{\langle \hat{\Lambda}^k \rangle_\infty}}{E_0^k} = \frac{\langle \hat{\Lambda}^k \rangle_G}{E_0^k} = 1 + \frac{k(k-1)}{2} \frac{\hbar\omega}{E_0} + O\left(\frac{\hbar^2\omega^2}{E_0^2}\right), \quad (7)$$

where $\hat{\Lambda} = \hbar\omega\hat{n}$, \hat{n} is the number operator and E_0 is the classical energy.

Next, we look at two-spin Gaudin magnets (see below for larger number of spins) – two interacting spins of arbitrary magnitudes (S_1, S_2) in a magnetic field,

$$\hat{H}_1 = B\hat{S}_1^z + \gamma\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2, \quad (8)$$

$$\hat{H}_2 = B\hat{S}_2^z - \gamma\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2, \quad [\hat{H}_1, \hat{H}_2] = 0. \quad (9)$$

Let us designate \hat{H}_1 as the Hamiltonian to generate quantum dynamics, though both Gaussian ensemble and infinite time averages are the same for a generic linear combination of \hat{H}_1 and \hat{H}_2 . Classically (when spins become angular momenta variables), this is an integrable system with two degrees of freedom. There are five independent parameters in $\hat{\rho}_G$ to be fixed by two first and three second moments of \hat{H}_1 and \hat{H}_2 .

We start the dynamics from $|\psi(0)\rangle = |\sigma_1\rangle \otimes |\sigma_2\rangle$, where $|\sigma_{i=1,2}\rangle$ is a spin coherent state characterized by a direction (θ_i, ϕ_i) in which the projection of $\hat{\mathbf{S}}_i$ is maximal. We choose coherent initial states for the ease of visualizing dynamics in the classical limit, since they correspond to individual points in the phase-space. However, other initial states are equally good. For an observable, we pick \hat{S}_1^z , arbitrarily set $(\theta_1, \phi_1) = (0.5\pi, 0.5\pi)$, $(\theta_2, \phi_2) = (0.3\pi, 0)$, $\gamma = 1$, and choose $B = \gamma S_2$, so that the effects of the magnetic field and interaction on the first spin are comparable.

We consider two cases: $S_1 = S_2$ and $S_1 = 1/2$ at increasing S_2 . The classical limit (for the second spin) is $\hbar \rightarrow 0$, $S_2 \rightarrow \infty$, while keeping $\hbar S_2 = \text{const}$. Therefore, terms of order $1/S_2^k$ are of order \hbar^k . Fig. 2 shows the difference $\langle \hat{S}_1^z \rangle_G - \langle \hat{S}_1^z \rangle_\infty$ between the Gaussian ensemble and infinite time averages as a function of S_2 . In the first case, the system becomes truly classical as $S_1 = S_2 \rightarrow \infty$ and the discrepancy (in $\hbar \langle \hat{S}_1^z \rangle = \text{finite}$) is of order \hbar^2 . We see that the Gaussian ensemble indeed captures the main quantum correction of order \hbar . Somewhat surprisingly, the agreement is even better in the second case, when the first spin stays quantum. Here at large S_2 the difference goes as $1/S_2^3$, i.e. the discrepancy is of order \hbar^3 .

Comparison with GGE. To compare with GGE at increasing system size, we analyze quenches in a free-fermion system, where GGE is exact in the thermody-

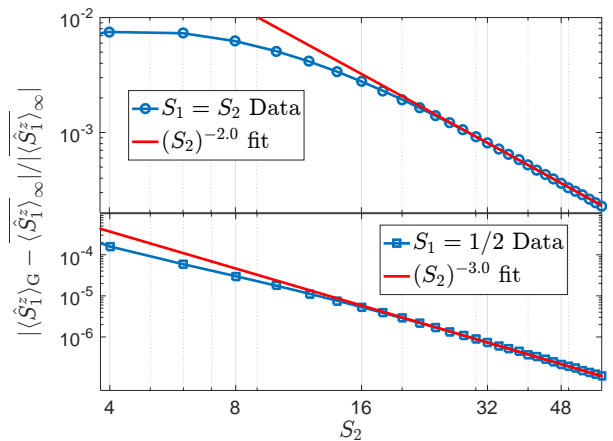


FIG. 2. (color online) Normalized difference between Gaussian ensemble and infinite time averages of S_1^z for the two-spin Hamiltonian (8) as a function of the magnitude S_2 of the second spin. Upper panel: $S_1 = S_2$, lower panel: $S_1 = 1/2$. The two averages converge in the limit $\hbar \rightarrow 0$, $S_2 \rightarrow \infty$, $\hbar S_2 = \text{fixed}$. Corrections to the Gaussian ensemble are of order \hbar^2 in the first case and \hbar^3 in the second.

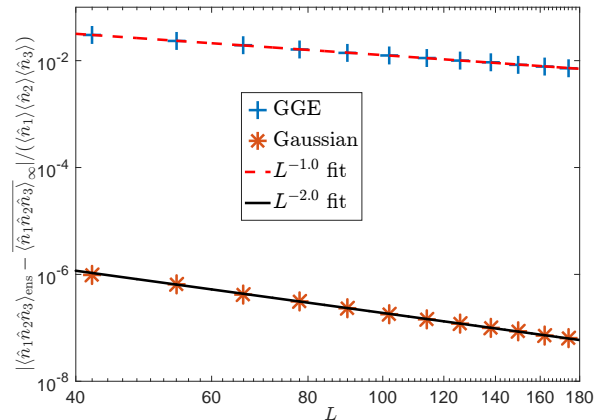


FIG. 3. (color online) Mismatch between ensemble (Gaussian and GGE) and infinite time averages in a three-point correlation function of the lattice site occupation number \hat{n}_i for the free-fermion model (10) with $Q = \pi/3$ as a function of the chain length L . We normalize by the first cumulant, which is the same for both ensembles. The time evolution is due to a quench from $\phi, V_1, V_2 = 0$ to $\phi = 0.3, V_1 = 1.5, V_2 = 1.0$. The mismatch vanishes as L^{-1} for GGE and as L^{-2} for the Gaussian ensemble. The latter thus captures the leading finite size correction ($\propto L^{-1}$) to the thermodynamic limit.

dynamic limit,

$$\hat{H} = - \sum_{j=1}^L (e^{i\phi} \hat{c}_{j+1}^\dagger \hat{c}_j + e^{-i\phi} \hat{c}_j^\dagger \hat{c}_{j+1}) + \sum_{j=1}^L [V_1 \cos(Qj) + V_2 \cos(2Qj)] \hat{n}_j, \quad (10)$$

where \hat{c}_j annihilates a fermion at site j , $\hat{n}_j = \hat{c}_j^\dagger \hat{c}_j$, and $Q = 2\pi/M$ is a commensurate modulation. We impose

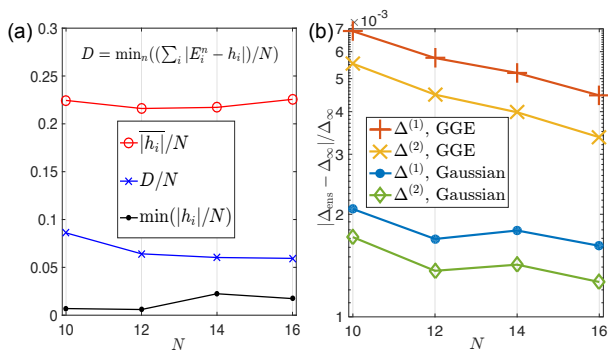


FIG. 4. (color online) Interaction quench from $\lambda_{\text{in}} = 0.5$ to $\lambda_{\text{fn}} = 2.0$ in the BCS model (11) with N single-particle levels (spins). (a) Distance D between the eigenvalues $E_i^{(n)}$ of the integrals of motion and their expectation values h_i in the initial state compared to the arithmetic mean of $|h_i|$ and smallest $|h_i|$. (b) Mismatch between ensemble (Gaussian and GGE) and infinite time (Δ_{∞}) averages for two alternative definitions, $\Delta^{(1)}$ and $\Delta^{(2)}$, of the BCS order parameter (14).

a periodic boundary condition and choose $M = 6$ (L is a multiple of M). We prepare the system in the ground state of \hat{H} with $\phi = V_1 = V_2 = 0$ at half filling and quench to nonzero ϕ , V_1 , and V_2 . This mixes M single-particle eigenstates in the pre-quench Hamiltonian. The parameter ϕ breaks the time-reversal symmetry and V_1, V_2 break the particle-hole symmetry thus removing all symmetry protected degeneracies in the single-particle spectrum. Natural integrals of motion are mode occupation numbers of the post-quench Hamiltonian.

Since GGE by construction captures averages of all single-particle occupation numbers and Gaussian ensemble – those of all their linear and bilinear combinations, they exactly reproduce the time average of single-body and two-body observables, respectively, for any L [22]. Therefore, we study a three-body correlation function $\langle \hat{n}_1 \hat{n}_2 \hat{n}_3 \rangle$. In Fig. 3, we plot the difference between the infinite time and ensemble averages normalized by the first cumulant $\langle \hat{n}_1 \rangle \langle \hat{n}_2 \rangle \langle \hat{n}_3 \rangle$. We see that the GGE approaches the thermodynamic limit $L \rightarrow \infty$ as $1/L$ as expected [17, 23, 24] while the Gaussian ensemble – as $1/L^2$ as anticipated above [22]. In addition to faster approach, the Gaussian ensemble result agrees with the time average better at any given L by orders of magnitude.

Long-range interactions. Our last example is the BCS model [25] on N single-particle levels ϵ_i . In terms of Anderson pseudospin-1/2 operators [26],

$$\hat{H}_{\text{BCS}} = \sum_{i=1}^N 2\epsilon_i \hat{S}_i^z - g \sum_{i,j=1}^N \hat{S}_i^+ \hat{S}_j^-. \quad (11)$$

As usual, we write the BCS coupling constant as $g = \lambda \delta$ [27, 28], where δ is the mean spacing between ϵ_i 's and λ is the dimensionless BCS coupling strength.

The integrals of motion of the BCS Hamiltonian are

Gaudin magnets [29, 30] (central spin models),

$$\hat{H}_i = -\frac{1}{g} \hat{S}_i^z + \sum_{j \neq i}^N \frac{\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j}{\epsilon_i - \epsilon_j}, \quad [\hat{H}_i, \hat{H}_j] = 0, \quad (12)$$

It is straightforward to verify that the total z -projection $\hat{S}^z = -g \sum_i \hat{H}_i$ and $\hat{H}_{\text{BCS}} = \sum_{i=1}^N 2\epsilon_i \hat{H}_i + \text{const.}$ Therefore, $[\hat{H}_{\text{BCS}}, \hat{H}_i] = 0$ and \hat{S}^z is conserved.

This model has several interesting features in addition to being the celebrated BCS model of superconductivity. First, \hat{H}_i are conserved for any (even unequal) spin magnitudes S_i , not just $S_i = 1/2$, i.e. the model is integrable (whatever this means in the quantum case) for arbitrary S_i . It is therefore one of the few models where one can gradually go from extreme quantum to purely classical while maintaining integrability. In the classical limit, when spins become angular momenta variables and commutators turn into Poisson brackets, H_{BCS} is integrable in the strict classical sense [31, 32] and Eq. (3) holds. Thus, the Gaussian ensemble (4) has a good foundation. On the other hand, the range of interactions in Eq. (11) is infinite and \hat{H}_i are not additive. So, there is no obvious justification [33] for a factorizable exponential form of the density matrix in Eq. (1).

We perform a sudden interaction quench $\lambda_{\text{in}} \rightarrow \lambda_{\text{fn}}$, i.e. we evolve with $\hat{H}_{\text{BCS}}(\lambda_{\text{fn}})$ starting from the ground state at λ_{in} . In numerics, we take $N = 2K$, $S^z = 0$, $\lambda_{\text{in}} = 0.5$, and $\lambda_{\text{fn}} = 2.0$. The single-particle spectrum is $\epsilon_i = 2i/(N-1) + \eta_i$, where $\eta_K = \eta_{K+1} = 0.1$ and $\eta_i = 0$ for $i \neq K, K+1$ [34]. Precise choice of the parameters is unimportant as long as the initial state is not dominated by just a few eigenstates of $\hat{H}_{\text{BCS}}(\lambda_{\text{in}})$. Otherwise, Gaussian ensemble becomes effectively exact as it has enough parameters to match all $|c_n|^2$ in time averages $\langle \hat{O} \rangle_{\infty} = \sum_n |c_n|^2 \hat{O}_{nn}$, where c_n are the coefficients in the decomposition of the initial state into the eigenstates of $\hat{H}_{\text{BCS}}(\lambda_{\text{in}})$. This is not a problem for large N , but for $N \leq 12$ we need to choose λ_{in} and λ_{fn} carefully to satisfy this condition. We use exact numerical diagonalization to construct the Gaussian ensemble, GGE, and determine $|c_n|^2$ for $N \leq 16$. The main challenge turns out to be solving for the Gaussian ensemble and GGE parameters, not obtaining the eigenstates.

Our attempts to construct an equal weight ensemble failed as we were unable to find even a single eigenstates with eigenvalues $E_i^{(n)}$ of \hat{H}_i sufficiently close to their expectation values $h_i = \langle \hat{H}_i \rangle_0$ in the initial state. For example, in Fig. 4(a) we show the minimum distance between $E_i^{(n)}$ and h_i ,

$$D = \frac{1}{N} \min_n \left(\sum_{i=1}^N |E_i^{(n)} - h_i| \right), \quad (13)$$

where N^{-1} ensures $D \propto N$ for large N , since $E_i^{(n)}$ and h_i grow as N . We see that D is comparable to the average $\overline{h_i} = N^{-1} \sum_i h_i$ and by far exceeds the smallest

$|h_i|$. Moreover, D/N does not appreciably decrease with increasing N between $N = 12$ and 16 , even though available N are too small for a reliable conclusion.

Now let us compare the Gaussian ensemble and GGE to the quench dynamics. We look at two versions of the BCS order parameter suitable for a system with fixed particle number (fixed S^z) [28],

$$\begin{aligned}\Delta^{(1)} &= g\sqrt{\sum_{i,j}\langle S_i^+ S_j^- \rangle - N_\uparrow} \\ \Delta^{(2)} &= g\sum_i \sqrt{\frac{1}{4} - \langle S_i^z \rangle^2},\end{aligned}\tag{14}$$

where N_\uparrow is the number of up spins, $N_\uparrow = N/2$ for $S^z = 0$. Fig. 4(b) shows the normalized difference between ensemble averaged $\Delta^{(i=1,2)}$ for both ensembles and the infinite time average [performed before taking square roots in Eq. (14)] as a function of N . The difference is of order $10^{-2} - 10^{-3}$ with an overall decreasing trend with increasing N for either ensemble for both definitions of the order parameter. The mismatch is significantly smaller for the Gaussian ensemble for all N .

It is well-known that mean-field becomes exact for the BCS model in $N \rightarrow \infty$ limit due to infinite interaction range [26, 35, 36]. Mean-field is equivalent to replacing quantum spin-1/2s with classical spins [26, 31, 32] in the Hamiltonian (11) and integrals of motion (12). On the other hand, we know that Gaussian ensemble (4) is exact in the classical limit. Thus, we expect it to become exact for the BCS dynamics in the thermodynamic limit. The parameter controlling finite size corrections to mean-field is δ/Δ_0 , where Δ_0 is the ground state gap. This parameter therefore plays the role of \hbar . If the Gaussian ensemble captures the leading quantum correction of order \hbar to the classical limit as in previous examples, the discrepancy between ensemble and infinite time averages should go as $(\delta/\Delta_0)^2 \propto N^{-2}$.

In conclusion, we proposed the multivariable Gaussian ensemble (4) as a quantum extension of exact classical GME (2). Our proposal stems from the classical definition of integrability thus largely bypassing difficulties associated with the absence of a sound widely accepted quantum notion. It is well and uniformly defined for any quantum system. It is exact in the classical limit and provides corrections of order \hbar to this limit at any particle number as well as finite size corrections to GGE whenever GGE holds. Further, we expect the Gaussian ensemble to become exact for models with long-range interactions in thermodynamic limit (as long as mean-field becomes exact). In this paper, we also analyzed two simple one- and two-body models and two many-body models to support these statements.

We thank M. Rigol for fruitful discussions, especially about GGE. H.K. acknowledges technical help from T. Oh. This work was financially supported in part by NSF

DMR-1308141 (H.K.); AFOSR FA9550-13-1-0039, NSF DMR-1506340, and ARO W911NF1410540 (A.P.); and David and Lucile Packard Foundation (E.A.Y.).

-
- [1] T. Kinoshita, T. Wenger, and D. S. Weiss, “A quantum Newton’s cradle”, *Nature* 440, 900 (2006).
 - [2] S. Hofferberth, I. Lesanovsky, B. Fischer, T. Schumm, and J. Schmiedmayer, “Non-equilibrium coherence dynamics in one-dimensional Bose gases”, *Nature* 449 324 (2007)
 - [3] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. Adu Smith, E. Demler, and J. Schmiedmayer, “Relaxation and Prethermalization in an Isolated Quantum System”, *Science* 337, 1318 (2012).
 - [4] S. Trotzky, Y-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, “Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas”, *Nat. Phys.* 8 324 (2012).
 - [5] R. Matsunaga, Y. I. Hamada, K. Makise, Y. Uzawa, H. Terai, Z. Wang, R. Shimano, “Higgs Amplitude Mode in BCS Superconductors $Nb_{1-x}Ti_xN$ induced by Terahertz Pulse Excitation”, *Phys. Rev. Lett.* 111, 057002 (2013).
 - [6] T. Langen, R. Geiger, M. Kuhnert, B. Rauer, and J. Schmiedmayer, “Local emergence of thermal correlations in an isolated quantum many-body system”, *Nature Physics* 9, 640 (2013).
 - [7] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, “Experimental observation of a generalized Gibbs ensemble”, *Science* 348, 207 (2015).
 - [8] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, “Relaxation in a completely integrable many-body quantum system: an ab initio study of the dynamics of the highly excited states of lattice hard-core bosons”, *Phys. Rev. Lett.* 98, 050405 (2007).
 - [9] B. Wouters, J. De Nardis, M. Brockmann, D. Fioretto, M. Rigol, and J.-S. Caux, “Quenching the anisotropic Heisenberg chain: Exact solution and generalized Gibbs ensemble predictions”, *Phys. Rev. Lett.* 113, 117202 (2014).
 - [10] B. Pozsgay, M. Mestyán, M. A. Werner, M. Kormos, G. Zaránd, and G. Takács, “Correlations after quantum quenches in the XXZ spin chain: Failure of the Generalized Gibbs Ensemble”, *Phys. Rev. Lett.* 113, 117203 (2014).
 - [11] G. Goldstein and N. Andrei, “Failure of the local Generalized Gibbs Ensemble for integrable models with bound states”, *Phys. Rev. A* 90, 043625 (2014).
 - [12] E. Ilievski, M. Medenjak, and T. Prosen, “Quasilocal conserved operators in isotropic Heisenberg spin 1/2 chain”, *Phys. Rev. Lett.* 115, 120601 (2015).
 - [13] E. Ilievski, J. De Nardis, B. Wouters, J.-S. Caux, F. H. L. Essler, and T. Prosen, “Complete Generalized Gibbs Ensemble in an interacting theory”, *Phys. Rev. Lett.* 115, 157201 (2015).
 - [14] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1978).
 - [15] E. A. Yuzbashyan, “Generalized microcanonical and Gibbs ensembles in classical and quantum integrable dynamics”, *Ann. Phys.* 367, 288 (2016).
 - [16] A. C. Cassidy, C. W. Clark, and M. Rigol, “Generalized thermalization in an integrable lattice system”, *Phys. Rev.*

- Lett. 106, 140405 (2011).
- [17] Kai He, Lea F. Santos, Tod M. Wright, and Marcos Rigol, “Single-particle and many-body analyses of a quasiperiodic integrable system after a quench”, Phys. Rev. A 87, 063637 (2013).
- [18] J.-S. Caux and F. H. L. Essler, “Time evolution of local observables after quenching to an integrable model”, Phys. Rev. Lett. 110, 257203 (2013).
- [19] M. Srednicki, “The approach to thermal equilibrium in quantized chaotic systems”, J. Phys. A 32, 1163 (1999).
- [20] L. D’Alessio, Y. Kafri, A. Polkovnikov, M. Rigol, “From Quantum Chaos and Eigenstate Thermalization to Statistical Mechanics and Thermodynamics”, arXiv:1509.06411, see Sect. 4.2.1.
- [21] M. V. Berry and M. Tabor, “Level Clustering in the Regular Spectrum”, Proc. R. Soc. Lond. A 356, 375 (1977).
- [22] See supplementary material.
- [23] C. Gramsch and M. Rigol, “Quenches in a quasisordered integrable lattice system: Dynamics and statistical description of observables after relaxation”, Phys. Rev. A 86, 053615 (2012).
- [24] T. M. Wright, M. Rigol, M. J. Davis, Karén V. Kheruntsyan, “Nonequilibrium dynamics of one-dimensional hard-core anyons following a quench: Complete relaxation of one-body observables”, Phys. Rev. Lett. 113, 050601 (2014).
- [25] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, “Theory of superconductivity”, Phys. Rev. 108, 1175 (1957).
- [26] P. W. Anderson, “Random-phase approximation in the theory of superconductivity”, Phys. Rev. 112, 1900 (1958).
- [27] I. L. Kurland, I. L. Aleiner, B. L. Altshuler, “Mesoscopic magnetization fluctuations for metallic grains close to the Stoner instability”, Phys. Rev. B 62, 14886 (2000).
- [28] J. von Delft, D.C. Ralph, “Spectroscopy of discrete energy levels in ultrasmall metallic grains”, Physics Reports, 345, 61 (2001); see footnote 11 on p. 96 and references therein.
- [29] M. Gaudin, *The Bethe Wavefunction* (Cambridge University Press, 2014).
- [30] M. C. Cambiaggio, A. M. F. Rivas, and M. Saraceno, “Integrability of the pairing Hamiltonian”, Nucl. Phys. A 424, 157 (1997).
- [31] E. A. Yuzbashyan, B. L. Altshuler, V. B. Kuznetsov, V. Z. Enolskii, “Solution for the dynamics of the BCS and central spin problems”, J. Phys. A 38, 7831, (2005).
- [32] E. A. Yuzbashyan, M. Dzero, V. Gurarie, M. S. Foster, “Quantum quench phase diagrams of an s-wave BCS-BEC condensate”, Phys. Rev. A 91, 033628 (2015).
- [33] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1*, 3rd ed., Vol. 5 (Butterworth-Heinemann, 1980).
- [34] We make the spectrum asymmetric with respect to the Fermi energy to avoid degeneracies in the $S^z = 0$ sector due to particle-hole symmetry.
- [35] R.W. Richardson, “Pairing in the limit of a large number of particles”, J. Math. Phys. 18,1802 (1977).
- [36] E. A. Yuzbashyan, A. A. Baytin, B. L. Altshuler, “Finite Size Corrections for the Pairing Hamiltonian”, Phys. Rev. B 71, 094505 (2005).

Supplementary Material: Gaussian ensemble for quantum integrable dynamics

CONVERGENCE TO THE CLASSICAL LIMIT FOR 1D HARMONIC OSCILLATOR

Here we show that the Gaussian ensemble converges to the classical microcanonical ensemble in $\hbar \rightarrow 0$ limit and reproduces the leading quantum correction to it for 1D harmonic oscillator in a coherent state. In this example, there is only one integral of motion (the Hamiltonian itself) in Eqs. (2) and (4). Suppose the initial state is a coherent state with eigenvalue z : $\hat{a}|z\rangle = z|z\rangle$, where \hat{a} is the annihilation operator. This corresponds to taking a particle in the ground state of $\hat{H}_0 = (\hat{p} - p_0)^2/(2m) + m\omega^2(\hat{q} - q_0)^2/2$ and time-evolving with $\hat{H} = \hat{p}^2/(2m) + m\omega^2\hat{q}^2/2$, where

$$z = \sqrt{\frac{m\omega}{2\hbar}}q_0 + i\frac{p_0}{\sqrt{2m\hbar\omega}},$$

i.e. to a quantum quench from $p_0, q_0 \neq 0$ to $p_0, q_0 = 0$

To construct ρ_G , we need to fix two parameters, μ and $\Sigma^{-1} \equiv 2\sigma^2$, with the help of Eqs. (5) and (6) for $\langle \hat{H} \rangle_0$ and $\langle \hat{H}^2 \rangle_0$. Decomposing $|z\rangle$ into number operator eigenstates $|n\rangle$, we obtain

$$C \sum_{n=0}^{\infty} \hbar\omega n e^{-\frac{(\hbar\omega n - \mu)^2}{2\sigma^2}} = \hbar\omega |z|^2, \quad (\hat{\rho}_G)_{nn} = C e^{-\frac{(\hbar\omega n - \mu)^2}{2\sigma^2}}, \quad (\text{S15})$$

$$C \sum_{n=0}^{\infty} (\hbar\omega n)^2 e^{-\frac{(\hbar\omega n - \mu)^2}{2\sigma^2}} = (\hbar\omega |z|^2)^2 (|z|^2 + 1), \quad (\text{S16})$$

where $C^{-1} = \sum_{n=0}^{\infty} \exp(-(\hbar\omega n - \mu)^2/(2\sigma^2))$ and we include the zero point energy $\hbar\omega/2$ into μ . The classical limit is

$$\hbar \rightarrow 0, \quad |z|^2 \rightarrow \infty, \quad \hbar\omega |z|^2 = \frac{p_0^2}{2m} + \frac{m\omega^2 q_0^2}{2} \equiv E_0 = \text{fixed}. \quad (\text{S17})$$

In this limit, sums in Eqs. (S15) and (S16) turn into integrals resulting in $\mu = E_0$ and $\sigma = \hbar\omega |z|$. We see that $\sigma \rightarrow 0$, $\hbar\omega n \rightarrow E$ and $\hat{\rho}_G \rightarrow C\delta(E - E_0)$, i.e. we recover the classical microcanonical ensemble.

Now let us demonstrate that the Gaussian ensemble captures the leading quantum correction to infinite time averages. We restrict ourselves to powers of the number operator, $\hat{n}^k = (\hat{a}^\dagger \hat{a})^k$, where k is a nonnegative integer. Note that expansion in \hbar is equivalent to the expansion in $1/|z|^2$, see Eq. (S17). The infinite time average is

$$\overline{\langle \hat{n}^k \rangle}_\infty = \sum_{n=0}^{\infty} n^k |\langle z|n \rangle|^2 = |z|^{2k} \left[1 + \frac{k(k-1)}{2|z|^2} + O\left(\frac{1}{|z|^4}\right) \right] = |z|^{2k} \left[1 + \frac{k(k-1)\hbar\omega}{2E_0} + O\left(\frac{\hbar^2\omega^2}{E_0^2}\right) \right], \quad (\text{S18})$$

where we used the fact that $|\langle z|n \rangle|^2 = |z|^{2n} e^{-|z|^2}/n!$ is a Poisson distribution with parameter $|z|^2$ whose k^{th} factorial moment, i.e. the expectation value of $\hat{n}(\hat{n}-1)\dots(\hat{n}-k+1)$, is $|z|^{2k}$, which is straightforward to verify directly.

Next, we evaluate the Gaussian ensemble averages. Let $n_0 = \mu/(\hbar\omega)$ and $s = \sigma/(\hbar\omega)$. We begin by showing that corrections to the classical answer $n_0 = s^2 = |z|^2$ as $|z|^2 \rightarrow \infty$ are exponentially small ($\propto e^{-|z|^2} = e^{-E_0/\hbar\omega}$) and therefore can be neglected when calculating corrections of order \hbar . Eq. (S15) becomes

$$\frac{\sum_{n=0}^{\infty} n \exp\left(-\frac{(n-n_0)^2}{2s^2}\right)}{\sum_{n=0}^{\infty} \exp\left(-\frac{(n-n_0)^2}{2s^2}\right)} = |z|^2. \quad (\text{S19})$$

Observe the following relation for any nonnegative integer k as n_0, s , and n_0/s tend to infinity,

$$\left| \sum_{n=0}^{\infty} n^k \exp\left(-\frac{(n-n_0)^2}{2s^2}\right) - \sum_{n=-\infty}^{\infty} n^k \exp\left(-\frac{(n-n_0)^2}{2s^2}\right) \right| < \sum_{n=2n_0+1}^{\infty} n^k \exp\left(-\frac{(n-n_0)^2}{2s^2}\right) < \int_{2n_0}^{\infty} x^k \exp\left(-\frac{(x-n_0)^2}{2s^2}\right) dx = O\left(s^2 n_0^{k-1} e^{-n_0^2/s^2}\right), \quad (\text{S20})$$

where we obtained the last relation by integrating by parts. Therefore, we can extend summations in Eq. (S19) from $[0, \infty)$ to $(-\infty, \infty)$ with an exponentially small error, i.e.

$$|z|^2 = \frac{\sum_{n=0}^{\infty} n \exp\left(-\frac{(n-n_0)^2}{2s^2}\right)}{\sum_{n=0}^{\infty} \exp\left(-\frac{(n-n_0)^2}{2s^2}\right)} = \frac{\sum_{n=-\infty}^{\infty} (n+n_0) \exp\left(-\frac{n^2}{2s^2}\right)}{\sum_{n=-\infty}^{\infty} \exp\left(-\frac{n^2}{2s^2}\right)} + O\left(|z|^2 e^{-|z|^2}\right) = n_0 + O\left(|z|^2 e^{-|z|^2}\right). \quad (\text{S21})$$

Similarly, we derive $s^2 = |z|^2$ up to an exponentially small error.

Therefore, the Gaussian ensemble expectation value is

$$\langle \hat{n}^k \rangle_{\text{G}} = \frac{\sum_{n=0}^{\infty} n^k \exp\left[-(n-|z|^2)^2/(2|z|^2)\right]}{\sum_{n=0}^{\infty} \exp\left[-(n-|z|^2)^2/(2|z|^2)\right]}. \quad (\text{S22})$$

We evaluate the sums involved using the Poisson summation formula,

$$A_k \equiv \sum_{n=0}^{\infty} n^k \exp\left[\frac{-(n-|z|^2)^2}{2|z|^2}\right] = |z|^{2k+2} \sum_{p=-\infty}^{\infty} \int_0^{\infty} e^{-|z|^2[(x-1)^2/2-2i\pi px]} x^k dx. \quad (\text{S23})$$

The saddle-point analysis of the integrals on the r.h.s. shows that the contribution of $p \neq 0$ terms is suppressed by a factor $e^{-2\pi^2|z|^2}$, i.e. it is sufficient to keep only the $p = 0$ term,

$$A_k = |z|^{2k+2} \left[\int_{-1}^{\infty} (1+y)^k e^{-|z|^2 y^2/2} dy + O\left(e^{-2\pi^2|z|^2}\right) \right] = |z|^{2k} \sqrt{2\pi} |z| \left[1 + \frac{k(k-1)}{2} \frac{1}{|z|^2} + O\left(\frac{1}{|z|^4}\right) \right], \quad (\text{S24})$$

where we changed the variable $x = y + 1$ in the integral and evaluated it with a simple version of the saddle-point method. Thus,

$$\langle \hat{n}^k \rangle_{\text{G}} = \frac{A_k}{A_0} = |z|^{2k} \left[1 + \frac{k(k-1)}{2|z|^2} + O\left(\frac{1}{|z|^4}\right) \right] = |z|^{2k} \left[1 + \frac{k(k-1)\hbar\omega}{2E_0} + O\left(\frac{\hbar^2\omega^2}{E_0^2}\right) \right]. \quad (\text{S25})$$

Comparing Eqs. (S18) and (S25), we see that they agree up to terms proportional to \hbar . In other words, Gaussian ensemble reproduces the leading term and the first quantum correction.

This agreement, however, does not extend to terms of order \hbar^2 and higher. Consider, for example, \hat{n}^3 . The exact infinite time average follows from the first three factorial moments mentioned above

$$\overline{\langle \hat{n}^3 \rangle}_\infty = |z|^6 + 3|z|^4 + |z|^2 = |z|^6 \left[1 + 3\frac{\hbar\omega}{E_0} + \frac{\hbar^2\omega^2}{E_0^2} \right]. \quad (\text{S26})$$

To obtain the Gaussian ensemble, we evaluate the integral in the first equation in (S24) for $k = 3$ and $k = 0$,

$$\langle \hat{n}^3 \rangle_G = |z|^6 + 3|z|^4 + O(e^{-|z|^2/2}) = |z|^6 \left[1 + 3\frac{\hbar\omega}{E_0} + O(e^{-E_0/\hbar\omega}) \right], \quad (\text{S27})$$

where the error arises from $|z|^{2k} \int_{-\infty}^{-1} (1+y)^k e^{-|z|^2 y^2/2} = O(e^{-|z|^2/2})$, which we estimated via repeated integration by parts. Therefore, there is a discrepancy of order \hbar^2 . For instance, since $\langle \hat{q}^6 \rangle = (5/2)(\hbar/m\omega)^3 \langle \hat{n}^3 \rangle +$ expectation values of lower powers of \hat{n} that are captured exactly by construction, the discrepancy in the third order cumulant of \hat{q}^2 between the Gaussian ensemble and exact infinite time averages is $2.5(\hbar/m\omega)^3 |z|^2 = 2.5\hbar^2 E_0/m^3 \omega^4$.

CALCULATION OF THREE-BODY OBSERVABLES IN THE FREE-FERMION MODEL

In the main text, we compared GGE and the Gaussian ensemble for quenches in a free-fermion model

$$\hat{H} = - \sum_{j=1}^L (e^{i\phi} \hat{c}_{j+1}^\dagger \hat{c}_j + e^{-i\phi} \hat{c}_j^\dagger \hat{c}_{j+1}) + \sum_{j=1}^L [V_1 \cos(Qj) + V_2 \cos(2Qj)] \hat{n}_j, \quad (\text{S28})$$

with periodic boundary conditions and $Q = 2\pi/M$ with integer M , such that L is a multiple of M . The pre-quench Hamiltonian \hat{H}_{in} has $\phi = V_1 = V_2 = 0$ and is therefore diagonal in the momentum basis,

$$\hat{H}_{\text{in}} = - \sum_k 2 \cos(k) \hat{c}_k^\dagger \hat{c}_k, \quad (\text{S29})$$

where $\hat{c}_k = L^{-1/2} \sum_{j=1}^L \hat{c}_j e^{-ikj}$ and $k = 2\pi\nu/L$ with $\nu = 1, 2, \dots, L$.

Due to our choice of the modulation wavenumber $Q = 2\pi/M$, the quenched Hamiltonian \hat{H}_{fn} mixes momenta $k, k+Q, k+2Q, \dots, k+(M-1)Q$ only among themselves. It is convenient to introduce a two-index momentum notation that reflects this property: $\hat{c}_k \rightarrow \hat{c}_{q,\alpha}$, where $q = 2\pi/L, 4\pi/L, \dots, 2\pi/M$, $\alpha = 0, 1, \dots, M-1$, and $k = q + \alpha Q$. The quenched Hamiltonian splits into L/M independent sub-Hamiltonians (sectors)

$$\hat{H}_{\text{fn}} = \sum_q \sum_{\alpha, \beta=0}^{M-1} h_{\alpha, \beta}^q \hat{c}_{q,\alpha}^\dagger \hat{c}_{q,\beta} = \sum_q \sum_{\gamma=0}^{M-1} \epsilon(q, \gamma) \hat{N}_{q,\gamma}, \quad \hat{N}_{q,\gamma} = \hat{b}_{q,\gamma}^\dagger \hat{b}_{q,\gamma}, \quad (\text{S30})$$

where $h_{\alpha, \beta}^q$ is an $M \times M$ matrix,

$$\mathbf{h}^q = \begin{pmatrix} -2 \cos(q + \phi) & V_1/2 & V_2/2 & \dots & V_1/2 \\ V_1/2 & -2 \cos(q + \phi + Q) & V_1/2 & \dots & V_2/2 \\ V_2/2 & V_1/2 & -2 \cos(q + \phi + 2Q) & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & V_1/2 \\ V_1/2 & \dots & V_2/2 & V_1/2 & -2 \cos(q + \phi + (M-1)Q) \end{pmatrix}. \quad (\text{S31})$$

Diagonalizing \mathbf{h}^q , we obtain single-particle energies $\epsilon(q) = (\mathbf{U}^q)^{-1} \mathbf{h}^q \mathbf{U}^q$ and new fermion operators $\hat{b}_{q,\gamma} = \sum_{\beta} (U^q)_{\gamma, \beta}^{-1} \hat{c}_{q,\beta}$.

The conservation laws are mode occupation numbers $\hat{N}_{q,\gamma}$. The GGE and Gaussian ensemble density matrices are

$$\hat{\rho}_{\text{GGE}} = \exp \left[- \sum_{k,\alpha} \lambda_{k,\alpha} \hat{N}_{k,\alpha} \right], \quad \langle \hat{N}_{k,\alpha} \rangle_{\text{GGE}} \equiv \frac{\text{tr}(\hat{\rho}_{\text{GGE}} \hat{N}_{k,\alpha})}{\text{tr} \hat{\rho}_{\text{GGE}}} = \langle \hat{N}_{k,\alpha} \rangle_0, \quad (\text{S32})$$

$$\hat{\rho}_G = \exp \left[- \sum_{p,q,\alpha,\beta} \sigma_{p\alpha,q\beta} \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \right], \quad \langle \hat{N}_{p,\alpha} \rangle_G \equiv \frac{\text{tr}(\hat{\rho}_G \hat{N}_{p,\alpha})}{\text{tr} \hat{\rho}_G} = \langle \hat{N}_{p,\alpha} \rangle_0, \quad \langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_G = \langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_0, \quad (\text{S33})$$

where we used $\hat{N}_{q,\beta} = \hat{N}_{q,\beta}^2$ to absorb the linear in $\hat{N}_{q,\beta}$ terms in the definition of $\hat{\rho}_G$ into the quadratic part.

We are interested in the three-point correlation function of the lattice site occupation number

$$\langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle = \langle \hat{c}_j^\dagger \hat{c}_j \hat{c}_\ell^\dagger \hat{c}_\ell \hat{c}_m^\dagger \hat{c}_m \rangle \quad (\text{S34})$$

$$= \frac{1}{L^3} \sum_{k,q,u,v,p,s} \sum_{\alpha,\beta,\gamma,\delta,\zeta,\eta} e^{i[k-q+(\beta-\alpha)Q]j+i[u-v+(\delta-\gamma)Q]\ell+i[s-p+(\eta-\zeta)Q]m} \langle \hat{c}_{q,\alpha}^\dagger \hat{c}_{k,\beta} \hat{c}_{v,\gamma}^\dagger \hat{c}_{u,\delta} \hat{c}_{p,\zeta}^\dagger \hat{c}_{s,\eta} \rangle \quad (\text{S35})$$

$$= \frac{1}{L^3} \sum_{k,q,u,v,p,s} \sum_{\alpha,\beta,\gamma,\delta,\zeta,\eta} e^{i[k-q+(\beta-\alpha)Q]j+i[u-v+(\delta-\gamma)Q]\ell+i[s-p+(\eta-\zeta)Q]m} \times \sum_{\alpha',\beta',\gamma',\delta',\zeta',\eta'} (U^q)_{\alpha,\alpha'}^* (U^k)_{\beta,\beta'} (U^v)_{\gamma,\gamma'}^* (U^u)_{\delta,\delta'} (U^p)_{\zeta,\zeta'}^* (U^s)_{\eta,\eta'} \langle \hat{b}_{q,\alpha}^\dagger \hat{b}_{k,\beta'} \hat{b}_{v,\gamma'}^\dagger \hat{b}_{u,\delta'} \hat{b}_{p,\zeta'}^\dagger \hat{b}_{s,\eta'} \rangle. \quad (\text{S36})$$

Therefore, we need to evaluate the following average:

$$\langle \hat{b}_{q,\alpha}^\dagger \hat{b}_{k,\beta'} \hat{b}_{v,\gamma'}^\dagger \hat{b}_{u,\delta'} \hat{b}_{p,\zeta'}^\dagger \hat{b}_{s,\eta'} \rangle, \quad (\text{S37})$$

with respect to the time-evolved state of the system as well as Gaussian ensemble and GGE.

The time-evolution is

$$\langle \hat{b}_{q,\alpha}^\dagger \hat{b}_{k,\beta'} \hat{b}_{v,\gamma'}^\dagger \hat{b}_{u,\delta'} \hat{b}_{p,\zeta'}^\dagger \hat{b}_{s,\eta'} \rangle_t = e^{i[\epsilon(q,\alpha') - \epsilon(k,\beta') + \epsilon(v,\gamma') - \epsilon(u,\delta') + \epsilon(p,\zeta') - \epsilon(s,\eta')]t} \langle \hat{b}_{q,\alpha}^\dagger \hat{b}_{k,\beta'} \hat{b}_{v,\gamma'}^\dagger \hat{b}_{u,\delta'} \hat{b}_{p,\zeta'}^\dagger \hat{b}_{s,\eta'} \rangle_0, \quad (\text{S38})$$

where $\langle \dots \rangle_t$ is the expectation value at time t . The infinite time average is nonzero only for terms with zero phase factor, i.e. when $\epsilon(q,\alpha') - \epsilon(k,\beta') + \epsilon(v,\gamma') - \epsilon(u,\delta') + \epsilon(p,\zeta') - \epsilon(s,\eta') = 0$. Since we removed time-reversal and particle-hole symmetries, there are no degeneracies in the single-particle spectrum as well as no two-particle and three-particle resonances. Consequently, the time average is zero unless the double indices on creation and annihilation operators are pairwise equal, i.e. the set $\{(q,\alpha'), (v,\gamma'), (p,\zeta')\}$ is a permutation of $\{(k,\beta'), (u,\delta'), (s,\eta')\}$. Therefore, only terms that can be cast into the form $\langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_t = \langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_0$ survive. The same holds for expectation value (S37) evaluated in any eigenstate of \hat{H}_{fn} and hence for both ensemble averages. Thus,

$$\overline{\langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle} - \langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle_{\text{ens}} = \frac{1}{L^3} \sum_{q,r,s;\alpha,\beta,\gamma} R_{\alpha\beta\gamma}^{qrs} \left(\langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_0 - \langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_{\text{ens}} \right), \quad (\text{S39})$$

where $R_{\alpha\beta\gamma}^{qrs}$ are coefficients of order one and $\langle \dots \rangle_{\text{ens}}$ stands for the average with respect to either ensemble.

Because $\hat{\rho}_{\text{GGE}}$ in Eq. (S32) is a (tensor) product of functions of individual occupation numbers [S1],

$$\langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_{\text{GGE}} = \langle \hat{N}_{q,\alpha} \rangle_{\text{GGE}} \langle \hat{N}_{r,\beta} \rangle_{\text{GGE}} \langle \hat{N}_{s,\gamma} \rangle_{\text{GGE}} = \langle \hat{N}_{q,\alpha} \rangle_0 \langle \hat{N}_{r,\beta} \rangle_0 \langle \hat{N}_{s,\gamma} \rangle_0, \quad (\text{S40})$$

as long as no two occupation number operators, i.e. *pairs* of indices coincide, $\{q,\alpha\} \neq \{r,\beta\} \neq \{s,\gamma\}$ and $\{q,\alpha\} \neq \{s,\gamma\}$. Since occupation numbers in different sectors are uncorrelated in the initial state, the same factorization holds for $\langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_0$ if q,r,s are distinct. For most remaining terms, when $q=r$ (but $\alpha \neq \beta$) or $q=s$ (but $\alpha \neq \gamma$) or $r=s$ (but $\beta \neq \gamma$), the GGE and initial state averages do not necessarily agree. In other words, GGE fails to capture the correlations between different occupation numbers *within* the same sector. Since the number of such terms (\propto number of pairs of sectors) is of order $(L/M)^2 \propto L^2$, $\overline{\langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle} - \langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle_{\text{GGE}} \propto 1/L$ at large L (M is fixed). Similarly, the Gaussian ensemble automatically matches $\langle \hat{N}_{q,\alpha} \hat{N}_{r,\beta} \hat{N}_{s,\gamma} \rangle_0$ for distinct q,r,s and for $q=r \neq s$, $q \neq r = s$, $q = s \neq r$, but not for $q=r=s$ (except when two of the indices α,β,γ are equal). In other words, it reproduces two-body, but not three-body correlations between occupation numbers in the initial state. The number of $q=r=s$ terms is proportional to L , so $\overline{\langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle} - \langle \hat{n}_j \hat{n}_\ell \hat{n}_m \rangle_{\text{G}} \propto 1/L^2$. This behavior with the system size L is in agreement with Fig. 3 in the main text for both ensembles. The fact that the Gaussian ensemble is exact for one- and two-point correlation functions of lattice site occupation numbers forced us to consider three-point functions.

Computing initial state and ensemble averages

Another implication of Eq. (S40) is that we need not determine the Lagrange multipliers $\lambda_{q,\alpha}$ for GGE in Eq. (S32), since GGE averages reduce to expectation values of single mode occupation numbers $\langle \hat{N}_{k,\alpha} \rangle_0$ in the initial state.

We do however need $\langle \hat{N}_{k,\alpha} \rangle_0$ and we also need $\langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_0$ in Eqs. (S33) to construct the Gaussian ensemble,

$$\langle \hat{N}_{k,\alpha} \rangle_0 = \langle 0 | \prod_q \hat{c}_q (\hat{b}_{k,\alpha}^\dagger \hat{b}_{k,\alpha}) \prod_p \hat{c}_p^\dagger | 0 \rangle = \sum_\beta |(U^k)_{\alpha,\beta}^{-1}|^2 \langle 0 | \prod_q \hat{c}_q (\hat{c}_{k,\beta}^\dagger \hat{c}_{k,\beta}) \prod_p \hat{c}_p^\dagger | 0 \rangle = \sum_{\beta: k+\beta Q \in \{\text{gr}\}} |(U^k)_{\alpha,\beta}^{-1}|^2, \quad (\text{S41})$$

$$\begin{aligned}
\langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_0 &= \langle 0 | \prod_u \hat{c}_u (\hat{b}_{p,\alpha}^\dagger \hat{b}_{p,\alpha} \hat{b}_{q,\beta}^\dagger \hat{b}_{q,\beta}) \prod_v \hat{c}_v^\dagger | 0 \rangle = \\
&= \sum_{\gamma, \gamma', \delta, \delta'} (U^p)_{\alpha, \gamma}^{-1} (U^{p*})_{\alpha, \gamma'}^{-1} (U^q)_{\beta, \delta}^{-1} (U^{q*})_{\beta, \delta'}^{-1} \langle 0 | \prod_u \hat{c}_u (\hat{c}_{p, \gamma}^\dagger \hat{c}_{p, \gamma} \hat{c}_{q, \delta}^\dagger \hat{c}_{q, \delta}) \prod_v \hat{c}_v^\dagger | 0 \rangle = \\
&= \sum_{q+\delta Q, p+\gamma Q \in \{\text{gr}\}} |(U^p)_{\alpha, \gamma}^{-1}|^2 |(U^q)_{\beta, \delta}^{-1}|^2 + \sum_{\substack{q+\delta Q \in \{\text{gr}\} \\ q+\gamma Q \notin \{\text{gr}\}}} (U^q)_{\alpha, \gamma}^{-1} (U^{q*})_{\beta, \gamma}^{-1} (U^q)_{\alpha, \delta}^{-1} (U^{q*})_{\beta, \delta}^{-1}, \tag{S42}
\end{aligned}$$

where $\{\text{gr}\}$ is the set of momenta occupied in the ground state of the pre-quench Hamiltonian.

To construct the Gaussian ensemble, we have to solve the last two equations in (S33) for $\sigma_{p\alpha, q\beta}$. Since $\langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_0 = \langle \hat{N}_{p,\alpha} \rangle_0 \langle \hat{N}_{q,\beta} \rangle_0$ for $p \neq q$ in the initial state, we set $\sigma_{p\alpha, q\beta} = 0$ for $p \neq q$. Then, $\hat{\rho}_G$ is a tensor product over different sectors labeled by p , which ensures [S1] $\langle \hat{N}_{p,\alpha} \hat{N}_{q,\beta} \rangle_G = \langle \hat{N}_{p,\alpha} \rangle_G \langle \hat{N}_{q,\beta} \rangle_G$. Thus, we are left with

$$\frac{\text{tr}(\hat{\rho}_G \hat{N}_{p,\alpha})}{\text{tr} \hat{\rho}_G} = \langle \hat{N}_{p,\alpha} \rangle_0, \quad \frac{\text{tr}(\hat{\rho}_G \hat{N}_{p,\alpha} \hat{N}_{p,\beta})}{\text{tr} \hat{\rho}_G} = \langle \hat{N}_{p,\alpha} \hat{N}_{p,\beta} \rangle_0, \quad \hat{\rho}_G = \exp \left[- \sum_{p, \alpha, \beta} \sigma_{p\alpha, p\beta} \hat{N}_{p,\alpha} \hat{N}_{p,\beta} \right]. \tag{S43}$$

There are $M(M+1)/2$ nonlinear equations for $M(M+1)/2$ unknown $\sigma_{p\alpha, p\beta}$ for each of L/M sectors labeled by p , a total of $L(M+1)/2$ equations, to be solved numerically. This is feasible for moderate $M (\leq 12)$.

Note also that the number of particles in each sector [each sub-Hamiltonian in Eq. (S30)] is conserved. Therefore, the number of eigenstates $|n\rangle$ of H_{fm} involved in the time-evolution in each sector is C_m^M . The infinite time average of an observable \hat{O} in the absence of degeneracies is $\langle \hat{O} \rangle_\infty = \sum_n |c_n|^2 \langle n | \hat{O} | n \rangle$, where c_n are the coefficients in the decomposition of the initial state into the eigenstates (diagonal ensemble). We need $M(M+1)/2 > C_m^M$ or else the Gaussian ensemble becomes exact as it has enough parameters to match all $|c_n|^2$. The smallest M that satisfies this criterion is $M = 6$ at $m = 3$. This dictates our choice of $M = 6$ and filling fraction = $1/2$.

Finally, we comment on a technical aspect of the computation. Even though we are dealing with a free-fermion model, a direct computation of three-point correlation functions (unlike two- and one-point ones) is prohibitive at large $L (\geq 120)$ due to a large number ($\propto L^3$) of nonzero terms in Eq. (S36). However, only a small fraction of these terms contribute to the difference between ensemble and infinite time averages in Eq. (S39). As discussed above, there is a factor of $(L/M)^2$ reduction in the number of terms for the Gaussian ensemble and a factor of L/M for GGE. This allows us to go to much larger L in numerical evaluation of the difference.

[S1] Note that $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$ and $\hat{O} = \hat{O}_1 \otimes \hat{O}_2$ implies $\text{tr}(\hat{\rho} \hat{O}) = \text{tr}(\hat{\rho}_1 \hat{O}_1) \text{tr}(\hat{\rho}_2 \hat{O}_2)$.