

Free-fermion Page Curve: Canonical Typicality and Dynamical Emergence

Xie-Hang Yu,^{1,2} Zongping Gong,^{1,2} and J. Ignacio Cirac^{1,2}

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany

²Munich Center for Quantum Science and Technology, Schellingstraße 4, 80799 München, Germany

We provide further analytical insights into the newly established noninteracting (free-fermion) Page curve, focusing on both the kinematic and dynamical aspects. First, we unveil the underlying canonical typicality and atypicality for random free-fermion states. The former appears for a small subsystem and is exponentially weaker than the well-known result in the interacting case. The latter explains why the free-fermion Page curve differs remarkably from the interacting one when the subsystem is macroscopically large, i.e., comparable with the entire system. Second, we find that the free-fermion Page curve emerges with unexpectedly high accuracy in some simple tight binding models in long-time quench dynamics. This contributes a rare analytical result concerning quantum thermalization on a macroscopic scale, where conventional paradigms such as the generalized Gibbs ensemble and quasi-particle picture are not applicable.

Introduction.—As a central concept in quantum information science [1], entanglement has been recognized to play vital roles in describing and understanding quantum many-body systems in and out of equilibrium [2–5]. For example, entanglement area laws for ground states of gapped local Hamiltonians enable their efficient descriptions based on tensor networks [6], while their violations may signature quantum phase transitions [7, 8]. The emergence of thermal ensemble from unitary evolution, a process known as quantum thermalization [9], is ultimately attributed to the entanglement generated between a subsystem and the complement [10].

Almost thirty years ago, Page considered the fundamental problem of bipartite entanglement in a fully random many-body system and found a maximal entanglement entropy (EE) up to finite-size corrections [11]. This seminal work was originally motivated by the black-hole information problem [12]. Remarkably, it has been attracted increasing and much broader interest in the past decade, due not only to the new theoretical insights from quantum thermalization [13–20] and quantum chaos [21–23], but also to the practical relevance in light of the rapid experimental development in quantum simulations [24–31]. In particular, the saturation of maximal entropy has been found to be a consequence of canonical typicality [32–34], which means most random states behave locally like the canonical ensemble. This typicality behavior has been argued to emerge in generic interacting many-body systems satisfying the eigenstate thermalization hypothesis [5, 9, 10, 35–37] and can even be rigorously established or ruled out in specific situations [38, 39].

In this Letter, we provide analogous insights into the noninteracting counterpart of Page’s problem. That is, we focus on free fermions or (fermionic) Gaussian states, which are of their own interest in quantum many-body physics, quantum information and computation [40–53]. Somehow surprisingly, in the seemingly simpler noninteracting case, the subsystem-size dependence of averaged EE, which is described by the Page curve pictorially, was not solved until very recently [54, 55]. It turns out to be similar to the interacting case for a small subsystem, but differ significantly otherwise. See Fig. 1(b) for an illustration. With the measure concentration results on compact-group manifolds, we establish the cor-

responding canonical typicality (atypicality) in microscopic (macroscopic) regions for the free-fermion ensemble. Thus we explicitly explain the similarity and difference from the kinematic aspect. In addition, we show that the free-fermion Page curve can be relevant to extremely simple tight-binding models via long-time quench dynamics. By classifying the systems according to their conserved (eigen) mode occupation numbers, we construct two classes of Hamiltonians which can/cannot give rise to a highly similar Page curve. Our finding concerning macroscopic properties which cannot be captured by the generalized Gibbs ensemble or quasi-particle picture and thus goes beyond the conventional paradigm of local thermalization.

Canonical typicality and atypicality.—We start by generalizing the main result in [32] to the random fermionic Gaussian (RFG) ensemble. While [32] already considers possible restrictions, we stress that Gaussianity is inadequate since Gaussian states do not constitute a Hilbert subspace. For simplicity, we consider number-conserving systems with totally N modes occupied by $N/2$ fermions, i.e., the half-filling case. Compared to the fully random case without number conservation, this setting appears to be more physically comprehensible and experimentally relevant, while displaying exactly the same Page curve [54, 55]. More general ensembles are discussed in Supplemental Material [56].

A pictorial illustration of our setup is shown in Fig. 1(a). Due to Wick’s theorem [57], a fermionic Gaussian state ρ is fully captured by its covariance matrix $C_{j,j'} = \text{Tr}(\rho a_j^\dagger a_{j'})$. Here a_j is the annihilation operator for mode j , which may label, e.g., a lattice site. As the covariance matrix for any RFG-pure state can be related to each other by a unitary transformation, the uniform distribution over this ensemble can be generated at the level of the covariance matrix $\{C = UC_0U^\dagger\}$. Here U is taken Haar-randomly over the unitary group $\mathbb{U}(N)$ [54, 55] and C_0 is an arbitrary reference covariance matrix in the ensemble satisfying $C_0^2 = C_0$ and $\text{Tr}C_0 = N/2$.

An important property of Gaussian states is that their subsystems remain Gaussian. We denote C_A as the $N_A \times N_A$ covariance matrix restricted to subsystem A with N_A modes. The EE $S_A = -\text{Tr}(\rho_A \log_2 \rho_A)$ of the reduced state $\rho_A =$

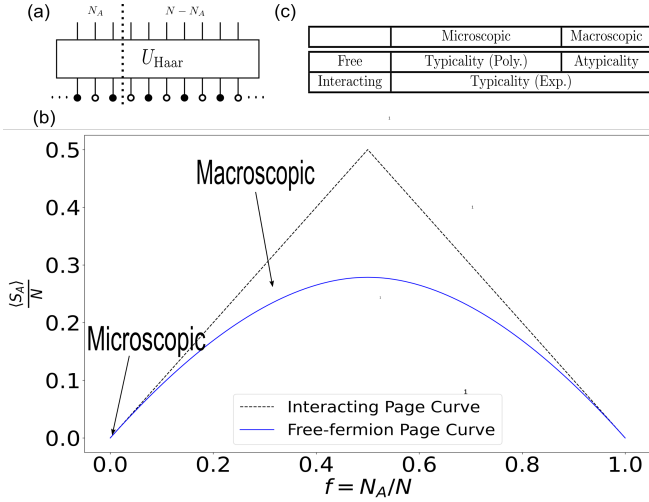


Figure 1. (a) The entire free-fermion system has N sites with half filling. The subsystem of interest has N_A ($N_A \leq N$) sites. The RFG ensemble is generated by Haar-random Gaussian unitaries with number conservation. (b) The Page curves of the RFG and interacting ensemble in the thermodynamic limit $N \rightarrow \infty$. It is obvious that these two Page curves agree with each other in the microscopic region but show a $\mathcal{O}(1)$ deviation in the macroscopic region. The interacting Page curve in the thermodynamic limit is always saturated. (c) The table summarizes the typicality/atypicality property for the RFG and interacting ensembles. Here ‘‘Poly.’’ and ‘‘Exp.’’ indicates polynomial and exponential scalings, respectively.

$\text{Tr}_{\bar{A}} \rho(\bar{A})$: complement of A) then reads:

$$S_A = -\text{Tr}(C_A \log_2 C_A) - \text{Tr}((I_A - C_A) \log_2 (I_A - C_A)), \quad (1)$$

where I_A is the identity matrix with dimension N_A .

Our first result is the measure concentration property of the covariance matrix for RFG ensemble:

Theorem 1 For arbitrary $\epsilon > 0$ and subsystem A , the probability that the reduced covariance matrix of a state in the RFG ensemble deviates from the ensemble average satisfies

$$\mathbb{P}(d_{\text{HS}}(C_A, I_A/2) \geq \eta + 2\epsilon) \leq 2e^{-\frac{\epsilon^2}{\eta}} \quad (2)$$

and

$$\mathbb{P}(d_{\text{HS}}^2(C_A, I_A/2) \leq \eta_a - 2\epsilon) \leq 2e^{-\frac{\epsilon^2}{\eta'_a}} \quad (3)$$

with $\eta = \sqrt{\frac{N_A^2}{2(N-1)}}$, $\eta' = \frac{12}{N}$, $\eta_a = \frac{N_A^2}{4(N+1)}$, $\eta'_a = \frac{12N_A}{N}$ and $d_{\text{HS}}(C, C') = \sqrt{\text{Tr}(C - C')^2}$ being the Hilbert-Schmidt distance.

The proof largely relies on the generalized Levy’s lemma for Riemann manifolds with positive curvature [58–60], which allows us to turn the upper bound on the distance average $\langle d_{\text{HS}}(C_A, I_A/2) \rangle \leq \sqrt{\frac{N_A^2}{2(N-1)}}$ or the lower bound

$\langle d_{\text{HS}}^2(C_A, I_A/2) \rangle \geq \frac{N_A^2}{4(N+1)}$ into a probability inequality [56]. From Eq. (2) we can easily see for infinite environments $N \rightarrow \infty$, the local microscopic system will have maximal entropy $S_A \rightarrow N_A$.

We emphasize that in Eq. (2), η and η' only scales polynomially with the (sub)system size. This contrasts starkly with the exponential scaling canonical typicality for random interacting ensemble [32]. Intuitively, this is because in the interacting case, the Hilbert-space dimension scales exponentially with the (sub)system size, which, however, simply equals to the size of the covariance matrix in the free-fermion case. Physically, the Gaussian constraint makes the ensemble only explore a very limited sub-manifold in the entire Hilbert space. This polynomial scaling means that, for a fixed subsystem size N_A , the reduced state still exhibits canonical typicality, while the atypicality is only polynomially suppressed by the environment size. Accordingly, the averaged EE should achieve the maximal value but with a polynomial finite-size correction. In fact, such an exponentially weaker canonical typicality (2) can also be exploited to explain the qualitatively larger variance of the EE for the RFG ensemble, which is $\mathcal{O}(N^{-2})$ in comparison to $e^{-\mathcal{O}(N)}$ in the interacting case [54–56].

On the other hand, if the subsystem is macroscopically large, meaning that $f = N_A/N$ is $\mathcal{O}(1)$, the concentration inequality (2) becomes meaningless since η is $\mathcal{O}(\sqrt{N})$, the same order as the Hilbert-Schmidt norm of C_A . Instead, we may take $\epsilon = \mathcal{O}(N^\alpha)$ with $\alpha \in (0, 1)$ in Eq. (3), finding that the majority of the reduced covariance matrix differs significantly from the ensemble average. In other words, the RFG ensemble exhibits canonical atypicality in this case. In particular, this result implies an $\mathcal{O}(1)$ deviation in the EE density from the maximal value. We recall that, in stark contrast, the canonical typicality for interacting states is exponentially stronger and persists even on any macroscopic scale with $f < 1/2$.

The above discussions can be made more straightforward by considering the measure concentration property of S_A . By bounding S_A using $d_{\text{HS}}(C_A, I_A/2)$ from both sides, we obtain [56]

$$\mathbb{P}(S_A \leq N_A - \epsilon) \leq 2e^{-\frac{(\sqrt{\epsilon} - \xi)^2}{\xi'}}, \quad \forall \epsilon > \xi^2 \quad (4)$$

in the microscopic region and

$$\mathbb{P}(S_A \geq N_A - \xi_a + \epsilon) \leq 2e^{-\frac{\epsilon^2}{\xi'_a}}, \quad \forall \epsilon > 0 \quad (5)$$

in the macroscopic region. Here $\xi = \sqrt{\frac{2N_A^2}{N-1}}$, $\xi' = \frac{192}{N}$, $\xi_a = \frac{N_A^2}{2 \ln 2(N+1)}$ and $\xi'_a = \frac{192N_A}{\ln^2 2N}$. Note that Eq. (4) also becomes meaningless in the macroscopic region since ξ^2 will be comparable with N_A . Choosing $\epsilon = (\xi + \mathcal{O}(N^{-\alpha/2}))^2$ for Eq. (4) and $\epsilon = \mathcal{O}(N^\alpha)$ in Eq. (5) with $\alpha \in (0, 1)$, we fully explain the microscopic similarity and macroscopic difference between the Page curves for the RFG and interacting ensembles. See Fig. 1(b) and (c).

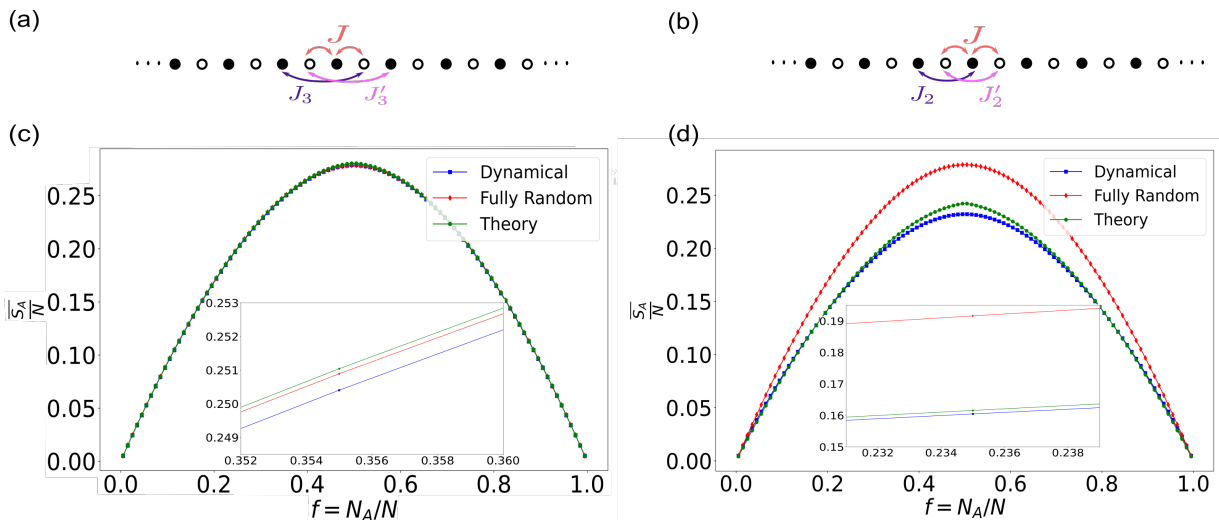


Figure 2. (a) and (b) show the tight binding Hamiltonians $H_0 + H_1$ with period 2. H_1 in (a) only includes the odd-range hopping, while (b) includes even-range hopping. (c) Dynamical Page curve for the minimal model (6) (blue) as a representative of (a) and its comparison with the Page curve for the RFG ensemble (red) as well as our theoretic result Eq. (8) (green). Here $N = 200$. These three lines are very close to each other, with a difference $\sim 10^{-3}$ which agrees with our analysis. A relatively larger deviation of our theoretical result from the other two curves at $f = \frac{1}{2}$ is due to the truncation. This figure can also represent the general dynamical Page curve for Hamiltonians in (a). (d) Dynamical Page curve for Hamiltonian $H = (\sum_{j=1}^N a_j^\dagger a_{j+1} + 0.3 \sum_{j:\text{even}} a_j^\dagger a_{j+2} - 0.3 \sum_{j:\text{odd}} a_j^\dagger a_{j+2}) + \text{H.c.}$ as a representative of (b). Here also $N = 200$. The dynamical Page curve is obviously different from the Page curve for the RFG ensemble. The considerable deviation between the theoretical result and the dynamical Page curve near $f = \frac{1}{2}$, where higher-order terms become least negligible, is because we only calculate up to the third term in Eq. (7) [56].

Dynamically emergent Page curve.—We recall that a particularly intriguing point of the (interacting) Page curve is its emergence in physical many-body systems with local interactions [13–18], which are typically chaotic but yet far from fully random. Indeed, a popular phenomenological theory for describing generic entanglement dynamics on the macroscopic level, the so-called entanglement membrane theory [23], explicitly assumes that the entanglement profile of the thermalized system follows the Page curve. The intuition is that a long-time evolution can generate highly non-local correlations in a state and roughly exhaust the whole Hilbert (sub)space, provided the dynamics is ergodic. It is thus natural to ask whether the free-fermion Page curve could be relevant to thermalization in physical systems without interactions. Note that this question is complementary to the aforementioned (a)typicality results, which are kinematic, i.e., irrelevant to dynamics, as in the interacting case [32].

We try to address the above question by analytically investigating the long-time averaged EE in the quench dynamics governed by some simple local quadratic Hamiltonians with number conservation. Hereafter, we use the term “dynamical Page curve” to refer to this long-time averaged entanglement profile. To simplify the analytic calculations, we assume the Hamiltonian H to be translation-invariant (under the periodic boundary condition) and specify our initial state $|\Psi_0\rangle$ to be a period-2 density wave with half filling. See Fig. 2(a-b) for a schematic illustration. The dynamical Page curve is thus formally given by

$\overline{S(\rho_A(t))}$, where $\rho_A = \text{Tr}_{\bar{A}}[e^{-iHt}|\Psi_0\rangle\langle\Psi_0|e^{iHt}]$ and $\overline{f(t)} = \lim_{T \rightarrow \infty} T^{-1} \int_0^T dt f(t)$ denotes the long-time average. It is worth mentioning that the dynamical Page curve is ensured to be concave by translation invariance, as a result of the strong subadditivity of quantum entropy [61].

We primarily focus on the minimal model, i.e., a one-dimensional lattice with nearest-neighbor hopping:

$$H_0 = \sum_j a_j^\dagger a_{j+1} + \text{H.c.}, \quad (6)$$

whose band dispersion reads $E_k = 2 \cos k$. We believe that the exact results for the large (spatiotemporal) scale dynamical behaviors of this fundamental model are interesting on their own. Moreover, our method and results actually apply to much broader situations, as will soon become clear below.

Surprisingly, despite the additional translation-invariant and energy-conserving constraints compared to the RFG ensemble, this minimal model (6) turns out to give rise to a dynamical Page curve extremely close to that for the RFG ensemble (see blue and red curves in Fig. 2(c)). To gain some analytic insights, we perturbatively expand the entropy expression (1) around $C_A = \frac{I_A}{2}$, obtaining

$$S_A(t) = N_A - \sum_{n=1}^{\infty} \frac{\text{Tr}(2C_A(t) - I_A)^{2n}}{2n(2n-1) \ln 2}. \quad (7)$$

Thanks to the translational invariance, $C_A(t)$ can be related to the block-diagonal momentum-space covariance matrix

$\tilde{C}(t) = \bigoplus_k \tilde{C}_k(t)$ via $C_A(t) = \Pi_A U_F \tilde{C}(t) U_F^\dagger \Pi_A^\dagger$. Here U_F and Π_A are the Fourier transformation matrix and projector to subsystem A , respectively. The off-diagonal elements of a 2×2 block $\tilde{C}_k(t)$ involve a time-dependent phase $e^{i\theta_k(t)}$ with $\theta_k(t) = t(E_k - E_{k+\pi})$. When calculating $\overline{\text{Tr}(2C_A(t) - I_A)^{2n}}$, we will encounter terms like $e^{i\theta_k(t)} e^{i\theta_{k'}(t)}$, which equals to $\delta_{k,k'+\pi}$ in the thermodynamic limit. This contraction allows us to establish a set of Feynman rules for systematically calculating Eq. (7) order by order [56].

Since the bipartite EE is identical for either of the subsystems, the Page curve is reflection-symmetric with respect to $f = \frac{1}{2}$ and thus it suffices to focus on $f = N_A/N \leq \frac{1}{2}$. In the thermodynamic limit, the dynamical Page curve turns out to be [56]

$$\frac{\overline{S_A}}{N} = f - \frac{1}{\ln 2} \left(\frac{1}{2} f^2 + \frac{1}{6} f^3 + \frac{1}{10} f^4 \right) + \mathcal{O}(f^5). \quad (8)$$

On the other hand, the Page curve for RFG ensemble is [54]

$$\langle \frac{S_A}{N} \rangle = f - \frac{1}{\ln 2} \left(\frac{1}{2} f^2 + \frac{1}{6} f^3 + \frac{1}{12} f^4 \right) + \mathcal{O}(f^5). \quad (9)$$

The above two equations differ only by $\frac{1}{60 \ln 2} f^4 + \mathcal{O}(f^5)$, which is as small as about 10^{-3} even for f near $1/2$.

Interestingly, if we add a perturbation H_1 to Eq. (6), as long as H_1 is period-2 and only includes odd-range hopping, as represented by Fig. 2(a), the dynamical Page curve can be analytically demonstrated to be the same as Eq. (8) in the thermodynamic limit, as the same Feynman rules apply [56]. One example is $H_1 = J(\sum_{j:\text{even}} a_j^\dagger a_{j+2m+1} - \sum_{j:\text{odd}} a_j^\dagger a_{j+2m+1}) + \text{H.c.}$ for arbitrary J and integer m . Thus, we have defined another ensemble of fermionic Gaussian states by dynamical evolution, which covers a wide class of Hamiltonians and this ensemble has remarkably similar Page curve as the RFG ensemble.

However, if H_1 includes even-range hopping, as represented by Fig. 2(b) the dynamical Page curve will be very different, as shown in Fig. 2(d). This can be easily explained with the canonical typicality property proved above: for this class of Hamiltonians, their conserved (eigen) mode occupation number n_k deviates from the average value of RFG ensemble, which is $\frac{1}{2}$. Thus, the dynamical ensemble is naturally "atypical" even for microscopic scale because the local conserved observable is constructed from mode occupation numbers [62]. This result implies the reduced state on a small subsystem deviates considerably from being maximally mixed so that the tangent slope of the dynamical Page curve at $f = 0$ is well below that for the RFG ensemble. In contrast, one can show that all the conserved mode occupation number for the class of Hamiltonians mentioned in the last paragraph are $\frac{1}{2}$.

All the observations above constitute our second main result:

Theorem 2 *The RFG ensemble-like dynamical Page curve (8) emerges for a period-2 short-range free-fermion Hamiltonian if and only if the conserved mode occupation numbers are $1/2$.*

Discussions.—It is well-known that the generalized Gibbs ensemble (GGE) characterizes the local thermalization of integrable systems including free fermions [62–66]. However, in principle, GGE only predicts the expectation values of observables, which do not include the entropy. Note that the former (latter) is linear (nonlinear) in ρ_A and thus commutes (does not commute) with time average. Moreover, we also study the macroscopic scale, which can not be captured by GGE as well as its recently proposed refined version [67] concerning the purified subsystem by measuring the complement [68]. In this sense, our study goes well beyond the conventional paradigm of quantum thermalization in integrable systems, pointing out especially the highly nontrivial behaviors on the macroscopic level, where typicality may completely break down.

Finally, let us mention the relation between our strategy and the quasi-particle picture, which is widely used to calculate EE growth [66, 69–75]. It turns out this picture fails to reproduce the dynamical Page curve. Under the periodical boundary condition, the quasi-particle picture predicts $\overline{S_A} = N - \frac{N_A^2}{N}$ for the Hamiltonian satisfying the conditions in Theorem 2 [56]. This result is obtained by counting the steady number of entangled pairs shared by A and \bar{A} . On the other hand, noting that $(2C_A - I_A)^{2n} \leq (2C_A - I_A)^2$, if we replace all the higher-order terms of $(2C_A - I_A)$ in Eq. (7) with $(2C_A - I_A)^2$, we will get a lower entropy bound, which coincides with the prediction by the quasi-particle picture: $S_A \geq N_A - \frac{\text{Tr}(2C_A - I_A)^2}{\ln 2} \sum_n \frac{1}{2n(2n-1)} = N_A - \frac{N_A^2}{N}$. It is thus plausible to argue that the quasi-particle picture ignores possible higher-order correlations beyond quasi-particle pairs.

Conclusion and outlook.—We have derived the canonical (a)typicality for the RFG ensemble and pointed out the quantitative scaling difference in atypicality suppression from interacting systems. This explains the very different behaviors of the Page curves. We have also explored the relevance to long-time quench dynamics of free-fermion systems. To our surprise, some simple time-independent Hamiltonians are enough to make the free-fermion Page curve emerge to a very high accuracy. We analytically prove a necessary and sufficient condition about this behavior. The breakdown of the quasi-particle picture was also discussed.

Strictly speaking, we define a new ensemble arising from a wide class of free-fermion Hamiltonians, whose dynamical Page curve resembles a lot but yet differs from the fully random one. The properties of this new ensemble and its corresponding Page curve merit further study. Another interesting question is how the dynamical Page curves will be enriched upon imposing additional symmetries (such as the Altland-Zirnbauer symmetries [76]). Our work proposes a methodology to study this question. Besides, whether or not the fully random Page curve can emerge exactly for a time-independent free Hamiltonian also remains open.

Z.G. is supported by the Max-Planck-Harvard Research Center for Quantum Optics (MPHQ). J.I.C. acknowledges

support by the EU Horizon 2020 program through the ERC Advanced Grant QENOCOBA No. 742102.

Note added.—While finalizing this manuscript, a related work by Isoue *et al.* appeared in Ref. [77], which reported the typicality for random bosonic Gaussian states.

-
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
- [2] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Rev. Mod. Phys.* **80**, 517 (2008).
- [3] J. Eisert, M. Cramer, and M. B. Plenio, *Rev. Mod. Phys.* **82**, 277 (2010).
- [4] J. Eisert, M. Friesdorf, and C. Gogolin, *Nat. Phys.* **11**, 124–130 (2015).
- [5] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [6] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, *Rev. Mod. Phys.* **93**, 045003 (2021).
- [7] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, *Phys. Rev. Lett.* **90**, 227902 (2003).
- [8] P. Calabrese and J. Cardy, *J. Stat. Mech.* **2004**, P06002 (2004).
- [9] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
- [10] R. Nandkishore and D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
- [11] D. N. Page, *Phys. Rev. Lett.* **71**, 1291 (1993).
- [12] D. N. Page, *Phys. Rev. Lett.* **71**, 3743 (1993).
- [13] Z.-C. Yang, C. Chamon, A. Hama, and E. R. Mucciolo, *Phys. Rev. Lett.* **115**, 267206 (2015).
- [14] L. Vidmar and M. Rigol, *Phys. Rev. Lett.* **119**, 220603 (2017).
- [15] Y. O. Nakagawa, M. Watanabe, H. Fujita, and S. Sugiura, *Nat. Commun.* **9**, 1635 (2018).
- [16] J. R. Garrison and T. Grover, *Phys. Rev. X* **8**, 021026 (2018).
- [17] T.-C. Lu and T. Grover, *Phys. Rev. E* **99**, 032111 (2019).
- [18] C. Murthy and M. Srednicki, *Phys. Rev. E* **100**, 022131 (2019).
- [19] J. Pollack, M. Rozali, J. Sully, and D. Wakeham, *Phys. Rev. Lett.* **125**, 021601 (2020).
- [20] K. Kaneko, E. Iyoda, and T. Sagawa, *Phys. Rev. A* **101**, 042126 (2020).
- [21] Y. Sekino and L. Susskind, *J. High Energy Phys.* **2008** (10), 065.
- [22] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, *J. High Energy Phys.* **2016**, 4.
- [23] C. Jonay, D. A. Huse, and A. Nahum (2018), arXiv:1803.00089.
- [24] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, *Nature* **528**, 77 (2015).
- [25] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, *Science* **353**, 794 (2016).
- [26] C. Neill, P. Roushan, M. Fang, Y. Chen, M. Kolodrubetz, Z. Chen, A. Megrant, R. Barends, B. Campbell, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Kelly, J. Mutus, P. J. J. O’Malley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. Polkovnikov, and J. M. Martinis, *Nat. Phys.* **12**, 1037 (2016).
- [27] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samajdar, H. Pichler, W. W. Ho, S. Choi, S. Sachdev, M. Greiner, V. Vuletić, and M. D. Lukin, *Nature* **595**, 227 (2021).
- [28] R. Verresen, M. D. Lukin, and A. Vishwanath, *Phys. Rev. X* **11**, 031005 (2021).
- [29] G. Semeghini, H. Levine, A. Keesling, S. Ebadi, T. T. Wang, D. Bluvstein, R. Verresen, H. Pichler, M. Kalinowski, R. Samajdar, A. Omran, S. Sachdev, A. Vishwanath, M. Greiner, V. Vuletić, and M. D. Lukin, *Science* **374**, 1242 (2021).
- [30] B. Yang, H. Sun, R. Ott, H. Y. Wang, T. V. Zache, J. C. Halimeh, Z. S. Yuan, P. Hauke, and J. W. Pan, *Nature* **587**, 392 (2020).
- [31] T. Liu, S. Liu, H. Li, H. Li, K. Huang, Z. Xiang, X. Song, K. Xu, D. Zheng, and H. Fan (2022), arXiv:2208.13347.
- [32] S. Popescu, A. J. Short, and A. Winter, *Nat. Phys.* **2**, 754 (2006).
- [33] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, *Phys. Rev. Lett.* **96**, 050403 (2006).
- [34] P. Reimann, *Phys. Rev. Lett.* **99**, 160404 (2007).
- [35] M. Srednicki, *J. Phys. A* **32**, 1163 (1999).
- [36] M. Rigol, V. Dunjko, and M. Olshanii, *Nature* **452**, 854 (2008).
- [37] R. Moessner and S. L. Sondhi, *Nat. Phys.* **13**, 424 (2017).
- [38] M. P. Müller, E. Adlam, L. Masanes, and N. Wiebe, *Commun. Math. Phys.* **340**, 499 (2015).
- [39] R. Hamazaki and M. Ueda, *Phys. Rev. Lett.* **120**, 080603 (2018).
- [40] L. G. Valiant, *SIAM Journal on Computing* **31**, 1229 (2002).
- [41] B. M. Terhal and D. P. DiVincenzo, *Phys. Rev. A* **65**, 032325 (2002).
- [42] S. Bravyi, *Quantum Info. Comput.* **5**, 216 (2005).
- [43] M. M. Wolf, *Phys. Rev. Lett.* **96**, 010404 (2006).
- [44] M.-C. Bañuls, J. I. Cirac, and M. M. Wolf, *Phys. Rev. A* **76**, 022311 (2007).
- [45] L. Fidkowski, *Phys. Rev. Lett.* **104**, 130502 (2010).
- [46] J. M. Magán, *Phys. Rev. Lett.* **116**, 030401 (2016).
- [47] T. Shi, E. Demler, and J. I. Cirac, *Ann. Phys.* **390**, 245 (2018).
- [48] M. Gluza, M. Kliesch, J. Eisert, and L. Aolita, *Phys. Rev. Lett.* **120**, 190501 (2018).
- [49] E. c. v. Greplová and G. Giedke, *Phys. Rev. Lett.* **121**, 200501 (2018).
- [50] L. Hackl and R. C. Myers, *J. High Energy Phys.* **2018**, 139.
- [51] A. Zhao, N. C. Rubin, and A. Miyake, *Phys. Rev. Lett.* **127**, 110504 (2021).
- [52] M. Oszmaniec, N. Dangniam, M. E. Morales, and Z. Zimborás, *PRX Quantum* **3**, 020328 (2022).
- [53] G. Matos, C. N. Self, Z. Papić, K. Meichanetzidis, and H. Dreyer, *Characterization of variational quantum algorithms using free fermions* (2022).
- [54] E. Bianchi, L. Hackl, M. Kieburg, M. Rigol, and L. Vidmar, *PRX Quantum* **3**, 030201 (2022).
- [55] E. Bianchi, L. Hackl, and M. Kieburg, *Phys. Rev. B* **103**, L241118 (2021).
- [56] See Supplemental Material for details.
- [57] L. Hackl and E. Bianchi, *SciPost Physics Core* **4**, 1 (2021).
- [58] N. D. Ponti (2017), arXiv:1705.01829.
- [59] M. Ledoux, *The Concentration of Measure Phenomenon*, Vol. 89 (American Mathematical Society, 2005).
- [60] E. S. Meckes, *The Random Matrix Theory of the Classical Compact Groups* (Cambridge University Press, 2019) pp. 1–212.
- [61] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, *Phys. Rev. Lett.* **100**, 070502 (2008).
- [62] T. Ishii and T. Mori, *Phys. Rev. E* **100**, 012139 (2019).
- [63] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, *Phys. Rev. Lett.* **98**, 050405 (2007).
- [64] A. C. Cassidy, C. W. Clark, and M. Rigol, *Phys. Rev. Lett.* **106**, 140405 (2011).
- [65] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, *Science* **348**, 207 (2015).
- [66] F. H. L. Essler and M. Fagotti, *J. Stat. Mech.* **2016**, 064002 (2016).

- [67] M. Lucas, L. Piroli, J. D. Nardis, and A. D. Luca (2022), arXiv:2207.13628.
- [68] W. W. Ho and S. Choi, *Phys. Rev. Lett.* **128**, 060601 (2022).
- [69] C.-A. Chen, S. Khlebnikov, and C.-L. Hung, *Phys. Rev. Lett.* **127**, 060404 (2021).
- [70] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, *Nature* **511**, 202 (2014).
- [71] O. A. Castro-Alvaredo, B. Doyon, and T. Yoshimura, *Phys. Rev. X* **6**, 041065 (2016).
- [72] P. Calabrese and J. Cardy, *J. Stat. Mech.* **2005**, P04010 (2005).
- [73] M. Fagotti and P. Calabrese, *Phys. Rev. A* **78**, 010306 (2008).
- [74] B. Bertini, M. Fagotti, L. Piroli, and P. Calabrese, *J. Phys. A* **51**, 39LT01 (2018).
- [75] B. Bertini, E. Tartaglia, and P. Calabrese, *J. Stat. Mech.* **2018**, 063104 (2018).
- [76] A. Altland and M. R. Zirnbauer, *Phys. Rev. B* **55**, 1142 (1997).
- [77] J. T. Iosue, A. Ehrenberg, D. Hangleiter, A. Deshpande, and A. V. Gorshkov, arXiv:2209.06838.
- [78] S. Lloyd, *Phys. Rev. A* **55**, 1613 (1997).
- [79] M. Hayashi, *Group Representation for Quantum Theory* (Springer International Publishing, 2017).
- [80] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, Cambridge Mathematical Library (Cambridge University Press, 1988).
- [81] Another proof for general cases can be obtained with the techniques in Subsec. III C.
- [82] E. Stein and R. Shakarchi, *Fourier Analysis: An Introduction* (Princeton University Press, 2003).
- [83] R. Kress, *Numerical Analysis*, Vol. 181 (Springer New York, 1998).
- [84] Noting that in higher n -expansion of $\overline{\text{Tr} X_A^{2n}}$, there will also be contributions to the entropy density $\frac{S_A}{N}$ at the order of $\mathcal{O}(f)$, like the term $\frac{N_A}{N} \frac{\sum_k n_k^{2n}}{N}$. Nonetheless, the convergence of expansion is guaranteed by $X_A^{2n} \leq X_A^{2n-2}$.
- [85] V. Alba and P. Calabrese, *SciPost Physics* **4**, 017 (2018).

SUPPLEMENTARY INFORMATION

CONTENTS

References	5
References	5
I. Proof of Canonical Typicality/Atypicality for the RFG ensemble	S1
II. Some Applications of Measure Concentration Typicality	S3
A. Measure concentration property for entropy	S3
B. Upper bound on the variance of entropy	S3
III. Detailed Calculation of the Dynamical Page Curves	S4
A. Calculation for the minimal model	S4
Second order in X_A	S5
Third order in X_A	S5
Forth Order in X_A	S6
B. General Feynman Rules	S6
C. Proof of Theorem 2	S8
D. Generalization to the atypical Page curves	S9
IV. Calculation of entanglement Entropy in the Quasi-Particle Picture	S11
<p>In this Supplemental Mateiral, we provide detailed proof of Theorem 1 and 2 in the main text. We also provide the calculations of other results and conclusions in the main text and discuss their generalization.</p>	

I. PROOF OF CANONICAL TYPICALITY/ATYPICALITY FOR THE RFG ENSEMBLE

In this section, we consider number conserving fermionic Gaussian ensemble with N modes occupied by m fermions. The notation follows the main text. In particular, $\langle \cdot \cdot \cdot \rangle$ is used to denote the average value over the ensemble. The covariance matrix of the subsystem A for a particular random Gaussian state is

$$C_A = \Pi_A U C_0 U^\dagger \Pi_A^\dagger, \quad (\text{S.1})$$

where Π_A is the projection operator on the the subsystem with size $N_A \times N$, U is taken Haar randomly over $\mathbb{U}(N)$ and C_0 satisfies $C_0^2 = C_0$ and $\text{Tr} C_0 = m$. In the following, the distance between two matrices is measured by Hilbert-Schmidt distance d_{HS} . We define a function $f : \mathbb{U}(N) \rightarrow \mathbb{R}$ as

$$f(U) = d_{\text{HS}}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger, \langle C_A \rangle). \quad (\text{S.2})$$

It is easy to check that f is Lipschitz continuous with constant 2:

$$\begin{aligned} |f(U_1) - f(U_2)| &\leq d_{\text{HS}}(U_1 C_0 U_1^\dagger, U_2 C_0 U_2^\dagger) \\ &\leq d_{\text{HS}}(U_1 C_0 U_1^\dagger, U_1 C_0 U_2^\dagger) + d_{\text{HS}}(U_1 C_0 U_2^\dagger, U_2 C_0 U_2^\dagger) \\ &= \|C_0(U_1^\dagger - U_2^\dagger)\|_{\text{HS}} + \|(U_1 - U_2)C_0\|_{\text{HS}} \\ &\leq 2d_{\text{HS}}(U_1, U_2). \end{aligned}$$

The generalized Levy's lemma [58–60] states that for any Lipschitz continuous function over some Riemann manifolds with positive curvature, its values are concentrated around the mean one. For the unitary group, we have

$$\mathbb{P}(|f(U) - \langle f(U) \rangle| \geq l\epsilon) \leq 2e^{-\frac{N\epsilon^2}{12}}, \quad (\text{S.3})$$

where l is the Lipschitz constant. In what follows, we will bound $\langle f(U) \rangle = \langle d_{\text{HS}}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger, \langle C_A \rangle) \rangle$. First, Since $\langle C_A \rangle = \Pi_A \int d_H(U) U C_0 U^\dagger \Pi_A^\dagger$ is invariant under any unitary on $\mathbb{U}(N_A)$, according to Schur's lemma, $\langle C_A \rangle = \frac{m}{N} I_A$ and

$$\begin{aligned} \langle \|\Pi_A U C_0 U^\dagger \Pi_A^\dagger - \langle C_A \rangle\|_{\text{HS}} \rangle &\leq \sqrt{\langle \|\Pi_A U C_0 U^\dagger \Pi_A^\dagger - \langle C_A \rangle\|_{\text{HS}}^2 \rangle} \\ &= \sqrt{\langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger - \langle C_A \rangle)^2 \rangle} \\ &= \sqrt{\langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger)^2 \rangle - \frac{m^2}{N^2} N_A}. \end{aligned}$$

Next, we need to calculate $\langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger)^2 \rangle$. The idea is similar as in [32] and originally comes from random quantum channel coding [78]: we introduce another reference space R' which has the same dimension as the original total system R . The following equation holds:

$$\langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger)^2 \rangle = \int d_H(U) \text{Tr}[(\Pi_A \otimes \Pi_A)(U C_0 U^\dagger \otimes U C_0 U^\dagger) \text{SWAP}_{RR'}(\Pi_A^\dagger \otimes \Pi_A^\dagger)],$$

where $\text{SWAP}_{RR'}$ is the SWAP operation between the original system R and the reference one R' . From Schur-Weyl duality [79], we obtain

$$\int d_H(U) (U C_0 U^\dagger \otimes U C_0 U^\dagger) \text{SWAP}_{RR'} = \alpha I_{RR'} + \beta \text{SWAP}_{RR'}. \quad (\text{S.4})$$

Now, for simplicity, we can take $C_0 = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$. The following relations hold

$$\begin{aligned} \text{Tr} \text{SWAP}_{RR'} &= N, \\ \text{Tr}[(U \otimes U)(C_0 \otimes C_0)(U^\dagger \otimes U^\dagger) \text{SWAP}_{RR'}] &= \text{Tr}[(U \otimes U)(C_0 \otimes C_0) \text{SWAP}_{RR'}(U^\dagger \otimes U^\dagger)] = \text{Tr} C_0^2 = m, \\ \text{Tr}[(U \otimes U)(C_0 \otimes C_0)(U^\dagger \otimes U^\dagger)] &= m^2. \end{aligned}$$

The trace of Eq. (S.4) gives $N^2 \alpha + N \beta = m$. Multiplying Eq. (S.4) by $\text{SWAP}_{RR'}$ and tracing it, we have $N \alpha + N^2 \beta = m^2$.

Solving the equations leads to $\begin{cases} \alpha = \frac{Nm - m^2}{N(N^2 - 1)} \\ \beta = \frac{Nm^2 - m}{N(N^2 - 1)} \end{cases}$. As a result,

$$\begin{aligned} \langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger)^2 \rangle &= \text{Tr}[(\Pi_A \otimes \Pi_A)(\alpha I_{RR'} + \beta \text{SWAP}_{RR'})(\Pi_A^\dagger \otimes \Pi_A^\dagger)] \\ &= \alpha N_A^2 + \beta N_A. \end{aligned} \quad (\text{S.5})$$

Assuming that in the thermodynamic limit $N \rightarrow \infty$, the density of charge $\frac{m}{N}$ is fixed as $\mathcal{O}(1)$, we obtain

$$\langle f(U) \rangle^2 \leq \frac{m N_A^2}{N(N-1)} \sim \mathcal{O}\left(\frac{N_A^2}{N}\right)$$

and the typicality

$$\mathbb{P}\left(d_{\text{HS}}\left(C_A, \frac{m}{N} I_A\right) \geq 2\epsilon + \sqrt{\frac{m N_A^2}{N(N-1)}}\right) \leq 2e^{-\frac{N\epsilon^2}{12}}. \quad (\text{S.6})$$

For the other direction, we take $f(U) = d_{\text{HS}}^2(C_A, \langle C_A \rangle)$, which is also Lipschitz continuous with constant calculated as

$$\begin{aligned} |f(U_1) - f(U_2)| &\leq 2(d_{\text{HS}}(\Pi_A U_1 C_0 U_1^\dagger \Pi_A^\dagger, \langle C_A \rangle) + d_{\text{HS}}(\Pi_A U_2 C_0 U_2^\dagger \Pi_A^\dagger, \langle C_A \rangle)) d_{\text{HS}}(U_1, U_2) \\ &\leq 4\sqrt{N_A} \left(1 - \frac{m}{N}\right) d_{\text{HS}}(U_1, U_2). \end{aligned} \quad (\text{S.7})$$

Here we assume $m \leq \frac{N}{2}$ due to the particle-hole symmetry (otherwise, we may replace $1 - \frac{m}{N}$ by $\frac{m}{N}$). According to Eq. (S.5), we obtain

$$\langle f(U) \rangle = \langle \text{Tr}(\Pi_A U C_0 U^\dagger \Pi_A^\dagger)^2 \rangle - \frac{m^2}{N^2} N_A \geq \frac{(N-m)m N_A^2}{N^2(N+1)}. \quad (\text{S.8})$$

If $\frac{N_A}{N}$ and $\frac{m}{N}$ are both fixed as $\mathcal{O}(1)$ in the thermodynamic limit, this formula scales linear with N . Applying generalized Levy's lemma leads to

$$\mathbb{P}\left(d_{\text{HS}}^2\left(C_A, \frac{m}{N}I_A\right) \leq \frac{(N-m)mN_A^2}{N^2(N+1)} - 4\sqrt{N_A}\left(1 - \frac{m}{N}\right)\epsilon\right) \leq 2e^{-\frac{N}{48}\epsilon^2}. \quad (\text{S.9})$$

For example, if we choose $\epsilon \sim \mathcal{O}(N^{\frac{1}{3}})$, the above inequality means that C_A will deviate from its ensemble average by an $\mathcal{O}(N)$ factor with almost unit probability. This is the atypicality discussed in the main text.

II. SOME APPLICATIONS OF MEASURE CONCENTRATION TYPICALITY

A. Measure concentration property for entropy

In this subsection, we will use Eq. (S.6) and Eq. (S.9) to derive the measure concentration typicality/atypicality for subsystem entropy. For simplicity the half filling condition is assumed. The eigenvalues of C_A are denoted as $\{\frac{1}{2} + \lambda_i\}$, $i \in \{1, \dots, N_A\}$ with $\sum_{i=1}^{N_A} \lambda_i^2 = d_{\text{HS}}^2(C_A, \frac{I_A}{2})$ and $\lambda_i \in [-\frac{1}{2}, \frac{1}{2}]$.

If the subsystem is microscopically small, we know the typicality of entropy follows by noting that

$$S_A = \sum_{i=1}^{N_A} H\left(\frac{1}{2} + \lambda_i, \frac{1}{2} - \lambda_i\right) = \sum_{i=1}^{N_A} \left[1 - \sum_{n=1}^{\infty} \frac{(2\lambda_i)^{2n}}{2n(2n-1)\ln 2}\right] \geq N_A - 4d_{\text{HS}}^2\left(C_A, \frac{I_A}{2}\right), \quad (\text{S.10})$$

where we replace $(2\lambda_i)^{2n}$ by $(2\lambda_i)^2$ in the last inequality since $(2\lambda_i)^2 \leq 1$. Here $H(p_0, p_1) = -p_0 \log_2 p_0 - p_1 \log_2 p_1$ is the Shannon entropy. Combined with Eq. (S.6) we obtain

$$\begin{aligned} \mathbb{P}(N_A - S_A \geq x) &\leq \mathbb{P}\left(d_{\text{HS}}\left(C_A, \frac{I_A}{2}\right) \geq \frac{\sqrt{x}}{2}\right) \\ &\leq \begin{cases} 2 \exp\left[-\frac{N}{48}\left(\frac{\sqrt{x}}{2} - \sqrt{\frac{N_A^2}{2(N-1)}}\right)^2\right], & x > \frac{2N_A^2}{N-1}; \\ 1, & x \leq \frac{2N_A^2}{N-1}. \end{cases} \end{aligned} \quad (\text{S.11})$$

As long as $N \gg N_A^2$, we conclude the subsystem entropy will be nearly maximal.

For the other direction, if N_A is macroscopically large, we can upper bound the lhs of Eq. (S.10) by

$$S_A \leq N_A - \frac{2}{\ln 2} d_{\text{HS}}^2\left(C_A, \frac{I_A}{2}\right). \quad (\text{S.12})$$

Following the atypicality of $d_{\text{HS}}^2(C_A, \frac{I_A}{2})$ in Eq. (S.9), the subsystem entropy density will show an $\mathcal{O}(1)$ deviation from the maximal value:

$$\begin{aligned} \mathbb{P}\left(S_A \geq N_A - \frac{N_A^2}{2\ln 2(N+1)} + \frac{2}{\ln 2}\epsilon\right) &\leq \mathbb{P}\left(d_{\text{HS}}^2\left(C_A, \frac{I_A}{2}\right) \leq \frac{N_A^2}{4(N+1)} - \epsilon\right) \\ &\leq 2e^{-\frac{N}{48N_A}\epsilon^2}. \end{aligned} \quad (\text{S.13})$$

We may take $\epsilon \sim \mathcal{O}(N^\alpha)$ for arbitrary $\alpha \in (0, 1)$, finding that the majority of subsystem entropy will be comparable or smaller than $N_A - \frac{N_A^2}{2\ln 2(N+1)}$. This clearly illustrates the difference of the Page curve for the RFG ensemble from the interacting one.

B. Upper bound on the variance of entropy

At the end of this section, we will discuss the variance of entropy in microscopic region. Here the half filling condition is also assumed. From Eq. (S.11), we obtain

$$\mathbb{P}((S_A - N_A)^2 \geq x) \leq \begin{cases} 1, & x \leq \frac{4N_A^4}{(N-1)^2}; \\ 2 \exp\left[-\frac{N}{48}\left(\frac{x^{\frac{1}{4}}}{2} - \sqrt{\frac{N_A^2}{2(N-1)}}\right)^2\right], & x > \frac{4N_A^4}{(N-1)^2}. \end{cases}$$

Therefore

$$\begin{aligned} \langle (S_A - N_A)^2 \rangle &= \int \mathbb{P}((S_A - N_A)^2 \geq x) dx \\ &\leq \frac{4N_A^4}{(N-1)^2} + 2 \int_{\frac{4N_A^4}{(N-1)^2}}^{\infty} dx \exp \left[-\frac{N}{48} \left(\frac{x^{\frac{1}{4}}}{2} - \sqrt{\frac{N_A^2}{2(N-1)}} \right)^2 \right]. \end{aligned}$$

For the last line, we can change the integral variable into $t = \sqrt{N} \left(\frac{x^{\frac{1}{4}}}{2} - \sqrt{\frac{N_A^2}{2(N-1)}} \right)$, obtaining

$$\begin{aligned} 2 \int_{\frac{4N_A^4}{(N-1)^2}}^{\infty} dx \exp \left[-\frac{N}{48} \left(\frac{x^{\frac{1}{4}}}{2} - \sqrt{\frac{N_A^2}{2(N-1)}} \right)^2 \right] &= \frac{128}{\sqrt{N}} \int_0^{\infty} dt \left(\frac{t}{\sqrt{N}} + \sqrt{\frac{N_A^2}{2(N-1)}} \right)^3 e^{-\frac{t^2}{48}} \\ &= \frac{128}{N^2} \int_0^{\infty} dt \left(t + \sqrt{\frac{NN_A^2}{2(N-1)}} \right)^3 e^{-\frac{t^2}{48}} \sim \mathcal{O} \left(\frac{1}{N^2} \right), \end{aligned}$$

provided that N_A is fixed as $\mathcal{O}(1)$. In conclusion, we obtain

$$\text{Var}(S_A) \leq \langle (S_A - N_A)^2 \rangle \sim \mathcal{O}(N^{-2}),$$

which agrees with [54, 55].

III. DETAILED CALCULATION OF THE DYNAMICAL PAGE CURVES

As mentioned in the main text, for all the models in this section, we assume the initial state is a period-2 density wave with half filling. Following the same notation in the main text, we further define $X_A(t) = 2C_A(t) - I_A$. Thus

$$S_A(t) = N_A - \sum_{n=1}^{\infty} \frac{\text{Tr} X_A^{2n}(t)}{2n(2n-1) \ln 2}. \quad (\text{S.14})$$

A. Calculation for the minimal model

We first consider the minimal model. Remember that the minimal model means only nearest neighbor hopping is included. After introducing the Fourier transformed mode $a_k^\dagger = \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{-ikj} a_j^\dagger$, we can easily obtain the correlation function in momentum space

$$\text{Tr}[\rho a_k^\dagger(t) a_{k'}(t)] = \frac{1}{2} \delta_{k,k'} + \frac{1}{2} \delta_{k,k'+\pi} e^{i\theta_k(t)}, \quad (\text{S.15})$$

where $a_k(t)$ ($a_k^\dagger(t)$) is the annihilation (creation) operator in the Heisenberg picture, $\theta_k(t) = t(E_k - E_{k+\pi})$ and $\rho = |\Psi_0\rangle\langle\Psi_0|$ corresponds to the initial density matrix. In the following, we may omit the index t if there is no ambiguity.

After inverse Fourier transformation back to position space, the covariance matrix for subsystem A reads $[C_A]_{m_1 m_2} = \frac{\delta_{m_1, m_2}}{2} + \frac{1}{2N} \sum_k e^{i\theta_k} e^{ik(m_1 - m_2)} e^{i\pi m_2}$ and thus

$$[X_A]_{m_1 m_2} = \frac{1}{N} \sum_k e^{i\theta_k} e^{ik(m_1 - m_2)} e^{i\pi m_2}. \quad (\text{S.16})$$

Second order in X_A

With Eq. (S.16), we can calculate $\overline{\text{Tr}X_A^2}$ as

$$\begin{aligned}\overline{\text{Tr}X_A^2} &= \frac{1}{N^2} \sum_{k_1, k_2, m_1, m_2} \overline{e^{i\theta_{k_1}} e^{ik_1(m_1-m_2)} e^{i\pi m_2} e^{i\theta_{k_2}} e^{ik_2(m_2-m_1)} e^{i\pi m_1}} \\ &= \frac{1}{N^2} \sum_{k_1, k_2, m_1, m_2} (\delta_{k_1, k_2 + \pi} e^{ik_1(m_1-m_2)} e^{ik_1(m_2-m_1)} e^{-i\pi(m_2-m_1)} e^{i\pi(m_1+m_2)} \\ &\quad + \delta_{k_1+k_2, \pi} e^{i2k_1(m_1-m_2)} e^{i\pi(m_1+m_2)} e^{i\pi(m_2-m_1)}) \\ &= \frac{N_A^2}{N} + \frac{1}{N} \sum_{k, m_1, m_2} e^{i2k(m_1-m_2)} = \frac{N_A^2}{N} + \frac{1}{N} \sum_{m_1, m_2} (\delta_{m_1-m_2, 0} + \delta_{m_1-m_2, \frac{N}{2}} + \delta_{m_1-m_2, -\frac{N}{2}}).\end{aligned}$$

Noting that the first term in the middle step comes from $\theta_{k+\pi} = -\theta_k$, which holds for general Hamiltonians according to the definition. However, the second term is due to the reflection symmetry (which implies $E_k = E_{-k}$) of the minimal model and is thus not universal. Fortunately, this model dependent term will vanish in the thermodynamic limit.

This calculation can be diagrammatically represented as shown in Fig. S1(a), where we draw an arrow from m_1 to m_2 (m_2 to m_1) because there is a factor $e^{ik_1(m_1-m_2)}$ ($e^{ik_2(m_2-m_1)}$). As will become clear below, such a diagrammatic representation provides a convenient and systematic way for dealing with higher order terms.

Third order in X_A

We move on to calculate $\overline{\text{Tr}X_A^3}$ and will see it vanishes in thermodynamic limit. The expression is

$$\overline{\text{Tr}X_A^3} = \frac{1}{N^3} \sum_{k_1, 2, 3, m_1, 2, 3} \overline{e^{i\theta_{k_1} + i\theta_{k_2} + i\theta_{k_3}} e^{ik_1(m_1-m_2) + ik_2(m_2-m_3) + ik_3(m_3-m_1)} e^{i\pi(m_1+m_2+m_3)}}.$$

The non-zero contribution of $\overline{e^{i\theta_{k_1} + i\theta_{k_2} + i\theta_{k_3}}}$ in the minimal model only comes from two cases:

1. $k_2 = k_1 + \pi, k_3 = \pm \frac{\pi}{2}$ and cyclic permutations. The contribution is proportional to

$$\frac{1}{N^3} \sum_{k_1, m_1, m_2, m_3} e^{ik_1(m_1-m_3)} e^{i\frac{\pi}{2}(m_1+m_3)} = \frac{N_A}{N^2} \sum_{m_1, m_3} \delta_{m_1, m_3} e^{i\pi m_1} \sim \mathcal{O}\left(\frac{1}{N}\right).$$

2. k_1, k_2 satisfy $|\cos k_1 + \cos k_2| \leq 1$ and $k_3 = \arccos(-\cos k_1 - \cos k_2)$. The contribution is proportional to

$$\sim \frac{1}{N^3} \sum_{k_1, k_2} \sum_{m_1} e^{im_1(k_1-k_3+\pi)} \sum_{m_2} e^{im_2(k_2-k_1+\pi)} \sum_{m_3} e^{im_3(k_3-k_2+\pi)}.$$

With Hölder inequality [80]: $\sum_i |a_i| |b_i| |c_i| \leq [(\sum_i |a_i|^3)(\sum_i |b_i|^3)(\sum_i |c_i|^3)]^{\frac{1}{3}}$, we can upper bound the above contribution as

$$\leq \frac{1}{N^3} \left\{ \left[\sum_{k_1, k_2} \left| \sum_{m_1} e^{im_1(k_1-k_3+\pi)} \right|^3 \right] \left[\sum_{k_1, k_2} \left| \sum_{m_2} e^{im_2(k_2-k_1+\pi)} \right|^3 \right] \left[\sum_{k_3, k_2} \left| \sum_{m_3} e^{im_3(k_3-k_2+\pi)} \right|^3 \right] \right\}^{\frac{1}{3}}.$$

Since here, no pair of two k differ by π (as it is the case already considered in the first case), the sum of m_1, m_2, m_3 only contributes to $\mathcal{O}(1)$, so the total contribution will be upper bounded by $\mathcal{O}\left(\frac{1}{N}\right)$.

In conclusion

$$\overline{\text{Tr}X_A^3} \sim \mathcal{O}\left(\frac{1}{N}\right)$$

and thus vanishes in the thermodynamic limit.

The above discussion can be generalized to higher orders: as long as the degeneracy point of a Hamiltonian is not dense, we can safely ignore the model dependent contribution and put $\overline{e^{i\theta_{k_1}} e^{i\theta_{k_2}}} = \delta_{k_1, k_2 + \pi}$, which we call a contraction. In the following, we will directly use this contraction rule.[81]

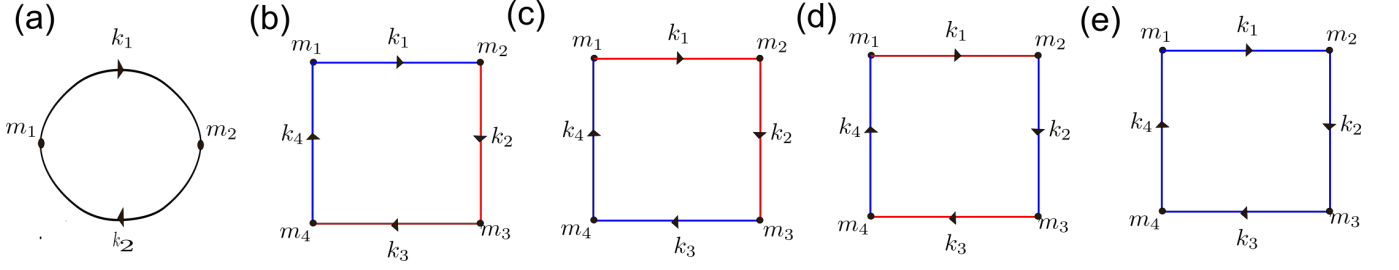


Figure S1. Feynman diagrams for calculating the entanglement entropy order by order. Here each vertex represents a position index and each leg represents a momentum index. Each leg is associated with $e^{i\theta_k} e^{ik(m-m')} e^{i\pi m'}$. In these diagrams, the legs with same color need to be contracted. Each color corresponds to one contraction.

Forth Order in X_A

Now we are moving to calculate $\overline{\text{Tr}X_A^4}$:

$$\overline{\text{Tr}X_A^4} = \frac{1}{N^4} \sum_{k_1, 2, 3, 4, m_1, 2, 3, 4} \prod_j^4 e^{i\theta_{k_j}} e^{ik_j(m_j - m_{j+1})} e^{im_j\pi}, \quad (\text{S.17})$$

where $m_5 = m_1$.

The first contracting class for $\overline{e^{i\theta_{k_1}} e^{i\theta_{k_2}} e^{i\theta_{k_3}} e^{i\theta_{k_4}}}$ is to contract the dynamical phase factors in pairs. Two patterns in this class are shown in Fig. S1(b) and Fig. S1(c). The legs with same colors mean that they are contracted together. These patterns correspond to $\overline{e^{i\theta_{k_1}} e^{i\theta_{k_4}}} \overline{e^{i\theta_{k_2}} e^{i\theta_{k_3}}} = \delta_{k_1, k_4 + \pi} \delta_{k_2, k_3 + \pi}$ and $\overline{e^{i\theta_{k_1}} e^{i\theta_{k_2}}} \overline{e^{i\theta_{k_4}} e^{i\theta_{k_3}}} = \delta_{k_1, k_2 + \pi} \delta_{k_4, k_3 + \pi}$, respectively. Substituting these delta functions into Eq. (S.17), we obtain

$$2 \times \frac{N_A^2}{N^4} \sum_{k_1, 3, m_1, 3} e^{ik_1(m_1 - m_3)} e^{ik_3(m_3 - m_1)} = \frac{2N_A^3}{N^2}, \quad (\text{S.18})$$

where the factor 2 comes from the equal contribution of these two diagrams.

Another pattern from this contracting class is shown in Fig. S1(d) which gives $\overline{e^{i\theta_{k_1}} e^{i\theta_{k_3}}} \overline{e^{i\theta_{k_2}} e^{i\theta_{k_4}}} = \delta_{k_1, k_3 + \pi} \delta_{k_2, k_4 + \pi}$. However, substituting this expression into Eq. (S.17) leads to

$$\frac{1}{N^4} \sum_{k_1, 2, m_1, 2, 3, 4} e^{i(k_1 - k_2)(m_1 + m_3 - m_2 - m_4)} e^{i\pi(m_2 + m_4)} = \frac{1}{N^2} \sum_{m_1, 2, 3, 4} \delta_{m_1 + m_3, m_2 + m_4} e^{i\pi(m_2 + m_4)} \sim \mathcal{O}\left(\frac{1}{N}\right). \quad (\text{S.19})$$

This diagram thus vanishes in the thermodynamic limit.

It should be emphasized that, in the above discussion, some terms are calculated multiple times. These form the other contracting class: we contract the four legs all together, as shown in Fig. S1(e). The contribution from this diagram needs to be subtracted due to the multiple calculation in Fig. S1(b) and (c):

$$-\frac{1}{N^4} \sum_{k_1, 2, 3, 4, m_1, 2, 3, 4} \delta_{k_2, k_1 + \pi} \delta_{k_3, k_1} \delta_{k_4, k_1 + \pi} \prod_j^4 e^{ik_j(m_j - m_{j+1})} e^{im_j\pi} = -\frac{N_A^4}{N^3}. \quad (\text{S.20})$$

Combining all the contributions together, we obtain

$$\overline{\text{Tr}X_A^4} = \frac{2N_A^3}{N^2} - \frac{N_A^4}{N^3}. \quad (\text{S.21})$$

From the above discussion, we can see that the contraction rules here are obviously different from Wick's theorem.

B. General Feynman Rules

The method presented in the previous subsection allows us to calculate Eq. (S.14) to arbitrary orders. Here we summarize our Feynman rules for contraction. A general Feynman diagram for calculating $\overline{\text{Tr}X_A^{2n}}$ is shown in Fig. S2(a). We will use the integer i to label the legs associated with momentum k_i .

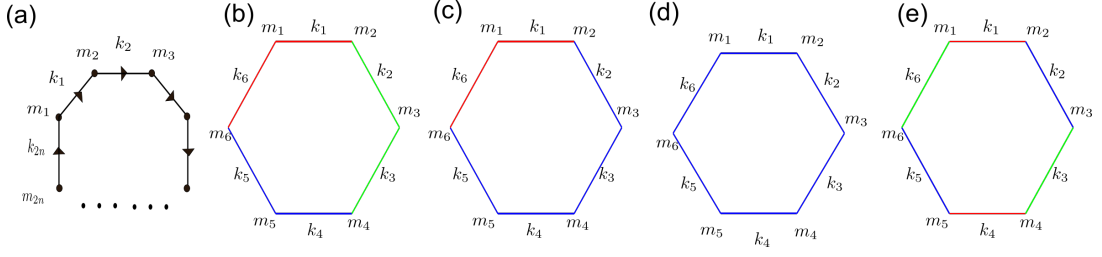


Figure S2. In this figure, more complicated Feynman diagrams are shown compared to Fig. S1. (a) shows a general Feynman diagram, (b)-(e) are the Feynman diagrams for calculating $\overline{\text{Tr} X_A^6}$.

1. All the legs in Fig. S2(a) must be contracted. Each contraction leads to a delta function of momenta k s and has to include even number of legs, where half of the legs should be labeled as even and the other half as odd. This last requirement arises from the phase factor $e^{i\pi m}$ in each leg and is to ensure the diagram not to vanish in the thermodynamic limit (we recall that Fig. S1(d) does not contribute as this requirement is not satisfied). For totally $2n$ legs with l contractions, the N, N_A dependence for this diagram is $\frac{N^{2n-l+1}}{N^{2n-l}}$.
2. Each contraction with $2j$ legs should also be assigned a multiple factor a_{2j} , accounting for the multiple calculations. a_2 and a_4 are obtained in previous discussion while higher a_{2j} can be obtained iteratively, as shown below.
3. For each diagram, multiply the term $\frac{N^{2n-l+1}}{N^{2n-l}}$ and factors a_{2j} obtained in Rule 1 and Rule 2 together. Some diagrams also need to multiply by subsystem correction factor β (see details below). Sum over all possible diagrams leads to the desired result.

The subsystem correction factor β does not appear in $\overline{\text{Tr} X_A^4}$ and $\overline{\text{Tr} X_A^2}$, but will appear in calculating $\overline{\text{Tr} X_A^6}$. For pedagogical purpose, we will now show how to calculate $\overline{\text{Tr} X_A^6}$. Some diagrams are shown in Fig. S2(b-e). In diagram (b), there are three contractions, each contracts two legs. The contribution for this diagram is $\frac{N^4}{N^3} a_2^3$. In diagram (c), there are two contractions, one contracts four legs together and the other contracts two legs. The contribution is $\frac{N^5}{N^4} a_4 a_2$. In diagram (d), all the six legs are contracted together, contributing to $\frac{N^6}{N^5} a_6$. Attentions should be payed to diagram (e). After contracting the three pairs of legs, we obtain

$$\begin{aligned} & \sum_{k_{1,2,3} m_{1,2,3,4,5,6}} e^{ik_1(m_1-m_2+m_4-m_5)} e^{ik_2(m_2-m_3+m_5-m_6)} e^{ik_3(m_3-m_4+m_6-m_1)} \\ &= N^3 \sum_{m_{1,2,3,4,5,6}} \delta_{m_1+m_4=m_2+m_5=m_3+m_6 \bmod N}. \end{aligned} \quad (\text{S.22})$$

Naively, one may conjecture the result of the last sum to be N_A^4 . However, this is only true when $f = \frac{N_A}{N} = 1$. If $f \leq \frac{1}{2}$, the sum will be much smaller. After carefully counting the pairs satisfying the delta function, we obtain for $f \leq \frac{1}{2}$

$$\sum_{k_{1,2,3} m_{1,2,3,4,5,6}} e^{ik_1(m_1-m_2+m_4-m_5)} e^{ik_2(m_2-m_3+m_5-m_6)} e^{ik_3(m_3-m_4+m_6-m_1)} = \frac{1}{2} N^3 (N_A^4 + N_A^2) \rightarrow \beta_1 N^3 N_A^4, \quad (\text{S.23})$$

where β_1 is defined as the subsystem correction factor in thermodynamic limit: $\beta_1 = \begin{cases} \frac{1}{2} & f \leq \frac{1}{2} \\ 1 & f = 1 \end{cases}$. Summing over all possible diagrams, the total contribution is

$$\overline{\text{Tr} X_A^6} = (5 + \beta_1) \frac{N_A^4}{N^3} - (6 + 3\beta_2) \frac{N_A^5}{N^4} + \frac{N_A^6}{N^5} a_6, \quad (\text{S.24})$$

where β_2 is another subsystem correction factor $\beta_2 = \begin{cases} \frac{2}{3} & f \leq \frac{1}{2} \\ 1 & f = 1 \end{cases}$. Here a_6 can be determined by considering the case when $N_A = N$ (i.e. $f = 1$). In this case, $\overline{\text{Tr} X_A^6} = N$, resulting in $a_6 = 4$. Therefore,

$$\overline{\text{Tr} X_A^6} = \frac{11}{2} \frac{N_A^4}{N^3} - 8 \frac{N_A^5}{N^4} + 4 \frac{N_A^6}{N^5} \text{ if } f \leq \frac{1}{2}. \quad (\text{S.25})$$

C. Proof of Theorem 2

In this subsection, we go beyond the minimal model and consider the general Hamiltonians satisfying the condition in Theorem 2 in the main text. We will find the Feynman rule as well as the subsystem entropy is indeed the same as the previous subsection.

Since the Hamiltonian is period-2, we can use a modified Fourier transformation to block diagonalize it:

$$A_k^\dagger = \sqrt{\frac{2}{N}} \sum_{j=1}^{\frac{N}{2}} e^{-ik(2j-1)} a_{2j-1}^\dagger, \quad B_k^\dagger = \sqrt{\frac{2}{N}} \sum_{j=1}^{\frac{N}{2}} e^{-i2kj} a_{2j}^\dagger, \quad k \in \left\{ \frac{2n\pi}{N} \right\}_{n=0}^{\frac{N}{2}-1}. \quad (\text{S.26})$$

Here A_k^\dagger, B_k^\dagger are related to the conserved (eigen) modes P_k^\dagger, Q_k^\dagger via a 2×2 unitary transformation U^k as $A_k^\dagger = U_{11}^k P_k^\dagger + U_{12}^k Q_k^\dagger$ and $B_k^\dagger = U_{21}^k P_k^\dagger + U_{22}^k Q_k^\dagger$. Substituting into Eq. (S.26) leads to

$$a_{2m}^\dagger = \sqrt{\frac{2}{N}} \sum_{k=0}^{\pi} e^{ik2m} (U_{21}^k P_k^\dagger + U_{22}^k Q_k^\dagger), \quad a_{2m+1}^\dagger = \sqrt{\frac{2}{N}} \sum_{k=0}^{\pi} e^{ik(2m+1)} (U_{11}^k P_k^\dagger + U_{12}^k Q_k^\dagger). \quad (\text{S.27})$$

If we define $Q_{k+\pi} = P_k$ and

$$Z_k^m = \begin{cases} \sqrt{2}U_{22}^k, & \text{if } m \text{ is even and } k < \pi; \\ \sqrt{2}U_{21}^{k-\pi}, & \text{if } m \text{ is even and } k \geq \pi; \\ \sqrt{2}U_{12}^k, & \text{if } m \text{ is odd and } k < \pi; \\ -\sqrt{2}U_{11}^{k-\pi}, & \text{if } m \text{ is odd and } k \geq \pi, \end{cases} \quad (\text{S.28})$$

the above inverse Fourier transformation (S.27) can be rewritten as

$$a_m^\dagger = \frac{1}{\sqrt{N}} \sum_{k=0}^{2\pi} Z_k^m Q_k^\dagger e^{ikm}.$$

In the following, we will simplify $\sum_{k=0}^{2\pi}$ as \sum_k and k should be understood as module 2π . By assumption, all the conserved quantity $\text{Tr}(\rho Q_k^\dagger Q_k) = \frac{1}{2}$ for $k \in [0, 2\pi)$, thus

$$\begin{aligned} [C_A]_{ml} &= \frac{1}{2N} \sum_k Z_k^m Z_k^{l*} e^{ik(m-l)} + \frac{1}{2N} \sum_k e^{i\theta_k} Z_k^m Z_{k+\pi}^{l*} e^{ik(m-l)} e^{i\pi l} \\ &= \frac{1}{2} \delta_{m,l} + \frac{1}{2N} \sum_k e^{i\theta_k} Z_k^m Z_{k+\pi}^{l*} e^{ik(m-l)} e^{i\pi l} \end{aligned}$$

where $\theta_{k+\pi} = -\theta_k$ by definition. In the last equality, we have used the unitarity of U^k . Namely, if $m-l$ is odd:

$$\begin{aligned} \frac{1}{2N} \sum_k Z_k^m Z_k^{l*} e^{ik(m-l)} &= \frac{1}{4N} \sum_k [Z_k^m Z_k^{l*} e^{ik(m-l)} + Z_{k+\pi}^m Z_{k+\pi}^{l*} e^{i(k+\pi)(m-l)}] \\ &= \frac{1}{4N} \sum_k (Z_k^m Z_k^{l*} - Z_{k+\pi}^m Z_{k+\pi}^{l*}) e^{ik(m-l)} = 0. \end{aligned}$$

A similar calculation can be carried out for the case in which $m-l$ is even. Therefore, we have

$$[X_A]_{ml} = \frac{1}{N} \sum_k e^{i\theta_k} Z_k^m Z_{k+\pi}^{l*} e^{ik(m-l)} e^{i\pi l}.$$

This expression is very similar to Eq. (S.16) except for the extra factors Z_k . Nonetheless, we will show those extra Z_k 's do not contribute in the thermodynamic limit. As a result, the same Feynman rules and dynamical Page curve follows.

When evaluating a Feynman diagram in the thermodynamic limit with N_A, N both going to infinity, we can first sum over the

position indices. Introducing $s_j = k_j - k_{j-1}$, $k_0 = k_{2n}$ and $m_{2n+1} = m_1$, we obtain

$$\begin{aligned}
& \sum_{m_1, 2, \dots, 2n} \prod_{j=1}^{2n} e^{ik_j(m_j - m_{j+1})} e^{i\pi m_j} Z_{k_j}^{m_j} Z_{k_{j-1} + \pi}^{m_j*} \\
&= \sum_{m_1, 2, \dots, 2n} \prod_{j=1}^{2n} e^{im_j s_j} e^{i\pi m_j} Z_{k_j}^{m_j} Z_{k_{j-1} + \pi}^{m_j*} \\
&= \prod_{j=1}^{2n} \left[\sum_{m_j: \text{even}} e^{im_j(s_j + \pi)} Z_{k_j}^0 Z_{k_{j-1} + \pi}^{0*} + \sum_{m_j: \text{odd}} e^{im_j(s_j + \pi)} Z_{k_j}^1 Z_{k_{j-1} + \pi}^{1*} \right] \\
&= \prod_{j=1}^{2n} \left[\frac{1 - e^{iN_A(s_j + \pi)}}{1 - e^{2i(s_j + \pi)}} \right] \prod_{j=1}^{2n} (Z_{k_j}^0 Z_{k_{j-1} + \pi}^{0*} + e^{i(s_j + \pi)} Z_{k_j}^1 Z_{k_{j-1} + \pi}^{1*}) \\
&= \frac{e^{i\frac{N_A}{2} \sum_{j=1}^{2n} (s_j + \pi)}}{e^{i \sum_{j=1}^{2n} (s_j + \pi)}} \prod_{j=1}^{2n} \left[\frac{\sin \frac{N_A(s_j + \pi)}{2}}{\sin(s_j + \pi)} \right] \prod_{j=1}^{2n} (Z_{k_j}^0 Z_{k_{j-1} + \pi}^{0*} + e^{i(s_j + \pi)} Z_{k_j}^1 Z_{k_{j-1} + \pi}^{1*}) \\
&= \prod_{j=1}^{2n} \left[\frac{\sin \frac{N_A(s_j + \pi)}{2}}{\sin(s_j + \pi)} \right] \prod_{j=1}^{2n} (Z_{k_j}^0 Z_{k_{j-1} + \pi}^{0*} + e^{i(s_j + \pi)} Z_{k_j}^1 Z_{k_{j-1} + \pi}^{1*})
\end{aligned}$$

In the above calculation, we have assumed N_A to be even for simplicity. We also emphasize that $s_1 \cdots s_{2n}$ is not independent since $\sum_{i=1}^{2n} s_i = 0$. The factor $\frac{\sin \frac{N_A(s_j + \pi)}{2}}{\sin(s_j + \pi)}$ will be dominated by the contribution from $s_j = 0$ or $s_j = \pi$ if Z_k is smooth enough. This is due to the convergence of Fourier series, see [82]. In [83] the difference between the discrete sum over momenta and the integration is also upper bounded. However, if $s_j \simeq 0$, it will lead to

$$Z_{k_j}^0 Z_{k_{j-1} + \pi}^{0*} + e^{i(s_j + \pi)} Z_{k_j}^1 Z_{k_{j-1} + \pi}^{1*} \simeq Z_{k_j}^0 Z_{k_j + \pi}^{0*} - Z_{k_j}^1 Z_{k_j + \pi}^{1*} = 0 \quad (\text{S.29})$$

due to the unitarity of U . In the end, we obtain

$$\begin{aligned}
\sum_{m_1, 2, \dots, 2n} \prod_{j=1}^{2n} e^{ik_j(m_j - m_{j+1})} e^{i\pi m_j} Z_{k_j}^{m_j} Z_{k_{j-1} + \pi}^{m_j*} &\simeq (Z_k^0 Z_k^{0*} + Z_k^1 Z_k^{1*})^{2n} \prod_{j=1}^{2n} \left[\frac{\sin \frac{N_A(s_j + \pi)}{2}}{\sin(s_j + \pi)} \right] s_j \neq 0 \\
&= 2^{2n} \prod_{j=1}^{2n} \left[\frac{\sin \frac{N_A(s_j + \pi)}{2}}{\sin(s_j + \pi)} \right] s_j \neq 0
\end{aligned} \quad (\text{S.30})$$

Now we can see in the final expression Eq. (S.30) that the model-dependent factor Z disappears. Therefore, those Hamiltonians satisfying the conditions in Theorem 2 in the main text will have the same Feynman rules and dynamical Page curve as the minimal model.

In Fig. S3, we plotted the dynamical Page curve for the Hamiltonian

$$H = \sum_i a_i^\dagger a_{i+1} + 0.3 \sum_{i: \text{even}} a_i^\dagger a_{i+3} - 0.3 \sum_{i: \text{odd}} a_i^\dagger a_{i+3} + \text{H.C.} \quad (\text{S.31})$$

This dynamical Page curve is nearly the same as the one of minimal model in the main text.

D. Generalization to the atypical Page curves

In this subsection, we further consider the case beyond the condition in Theorem 2 in the main text, namely $\text{Tr}(\rho Q_k^\dagger Q_k) \neq \frac{1}{2}$ (we recall that $Q_{k+\pi} = P_k$). We denote $n_k = \text{Tr}(\rho Q_k^\dagger Q_k)$ and $\eta_k = \sqrt{n_k(1 - n_k)}$. Following the half filling condition, we have

$$n_{k+\pi} = 1 - n_k, \quad \eta_{k+\pi} = \eta_k$$

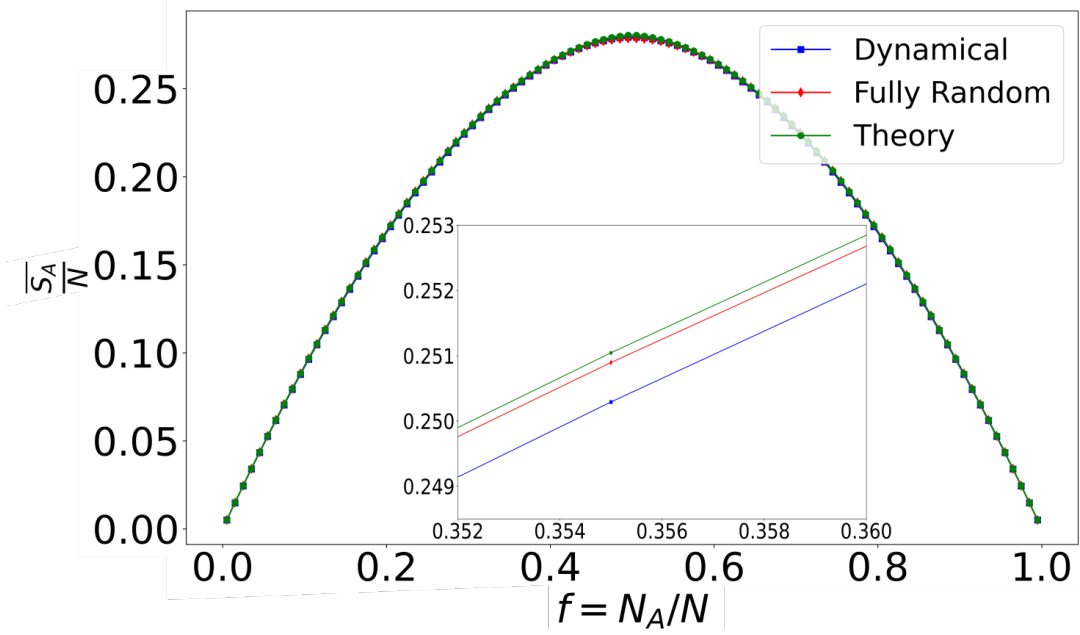


Figure S3. The dynamical Page curve for Hamiltonian (S.31) (blue curve) and its comparison with the one for the RFG ensemble (red curve) and the theoretical result (green curve). Here $N = 200$. The theoretical result is truncated up to order $\mathcal{O}(f^4)$, the same as in the main text.

and still $a_m^\dagger = \frac{1}{\sqrt{N}} \sum_{k=0}^{2\pi} Z_k^m Q_k^\dagger e^{ikm}$ with Z_k^m defined in Eq. (S.28). The covariance matrix can be calculated as

$$\begin{aligned} [C_A]_{ml} &= \frac{1}{N} \sum_{k_1, k_2} Z_{k_1}^m Z_{k_2}^{l*} e^{ik_1 m} e^{-ik_2 l} \text{Tr}(\rho Q_{k_1}^\dagger Q_{k_2}) \\ &= \frac{1}{N} \sum_k Z_k^m Z_k^{l*} e^{ik(m-l)} n_k + \frac{1}{N} \sum_k e^{i\theta_k} Z_k^m Z_{k+\pi}^{l*} e^{ik(m-l)} e^{i\pi l} \eta_k. \end{aligned}$$

Since $n_k \neq \frac{1}{2}$, in general there is no simple expression for X_A .

Using the same techniques as in the previous subsection, we can still establish the Feynman rules for this case. However, we have to distinguish two kinds of legs, one like $Z_{k_j}^{m_j} Z_{k_j}^{m_{j+1}*} e^{ik_j(m_j - m_{j+1})} n_{k_j}$ and the other like $e^{i\theta_{k_j}} Z_{k_j}^{m_j} Z_{k_j+\pi}^{m_{j+1}*} e^{ik_j(m_j - m_{j+1})} e^{i\pi m_{j+1}} \eta_{k_j}$. There is no dynamical phase $e^{i\theta_k}$ in the former, namely no delta functions associated with contraction. Also, there is no extra $e^{i\pi m_{j+1}}$ phase term in the former. Due to the difference between these two kinds of legs, the rule is more complicated than the previous case.

As an example, we can calculate the first three non-trivial terms to obtain:

$$\overline{\text{Tr} C_A^2} = \frac{N_A}{N} \sum_k n_k^2 + \frac{N_A^2}{N^2} \sum_k \eta_k^2,$$

$$\overline{\text{Tr} C_A^3} = \frac{N_A}{N} \sum_k n_k^3 + \frac{3N_A^2}{N^2} \sum_k n_k \eta_k^2,$$

$$\overline{\text{Tr} C_A^4} = \frac{N_A}{N} \sum_k n_k^4 + 4 \frac{N_A^2}{N^2} \sum_k n_k^2 \eta_k^2 + 2 \frac{N_A^2}{N^2} \sum_k \eta_k^4 + \frac{2N_A^3}{N^3} \sum_k \eta_k^4 - \frac{N_A^4}{N^4} \sum_k \eta_k^4.$$

Therefore, Up to $\overline{\text{Tr} X_A^4}$, [84]

$$\overline{S_A} \simeq \frac{N_A (\ln 2 + \frac{3}{4}) - \frac{N_A}{N} \sum_k (4n_k^2 - \frac{8}{3}n_k^3 + \frac{4}{3}n_k^4) - \frac{N_A^2}{N^2} \sum_k (4\eta_k^2 - 8n_k \eta_k^2 + \frac{16}{3}n_k^2 \eta_k^2 + \frac{8}{3}\eta_k^4) - \frac{N_A^3}{N^3} \frac{8}{3} \sum_k \eta_k^4}{\ln 2}.$$

IV. CALCULATION OF ENTANGLEMENT ENTROPY IN THE QUASI-PARTICLE PICTURE

In the quasi-particle picture, a nonequilibrium initial state is a source for generating quasi-particles with opposite momenta, which travel ballistically through the system. Here the main assumption is those quasi-particle pairs generated at different locations and times are incoherent. Therefore, the entanglement entropy of subsystem A is proportional to the number of pairs shared between A and its complement. Without loss of generality, we can assume the subsystem A is located in $[0, N_A)$, $N_A \leq \frac{N}{2}$. For a certain type of pairs with velocity $\pm v(k)$ ($v(k) > 0$), if the right-end of the pair is at position $0 \leq x < N_A$, i.e. within subsystem A , only when x satisfies

$$N_A - N \leq x - 2v(k)t < 0$$

can this pair contribute to the entanglement entropy of A . Here the periodic boundary condition is taken into account and $2v(k)t$ should be understood as modulo N . The solutions of this inequality is a continuous range $x \in [x_{\min}, x_{\max})$, where $x_{\min} = \max\{0, 2v(k)t + N_A - N\}$ and $x_{\max} = \min\{N_A, 2v(k)t\}$. Accordingly, we obtain

$$\Delta x = x_{\max} - x_{\min} = \begin{cases} 2v(k)t, & 2v(k)t \leq N_A; \\ N_A, & N_A < 2v(k)t < N - N_A; \\ N - 2v(k)t, & 2v(k)t \geq N - N_A. \end{cases}$$

A similar argument holds if the left-end is in subsystem A . We assume that after a sufficiently long time, the quasi-particle pairs will distribute uniformly among the system. Hence, the contribution of quasi-particle pairs with momentum k to the entanglement entropy upon the long-time average is given by

$$\frac{S_A(k)}{N^2} \int_0^N d(2v(k)t) \Delta x = S_A(k) \left[\frac{N_A}{N} - \left(\frac{N_A}{N} \right)^2 \right],$$

where the coefficient $S_A(k)$ is to be determined. Summing over all types of pairs, we obtain

$$S_A^{\text{qp}} = \left(\frac{N_A}{N} - \frac{N_A^2}{N^2} \right) \sum_k S_A(k).$$

If $N_A \rightarrow 0$, the limit $S_A^{\text{qp}} \rightarrow N_A \sum_k \frac{H(n_k)}{N}$ should hold [85]. Therefore, $S_A(k) = H(n_k)$. If $n_k = \frac{1}{2}$ for all k 's, the entanglement entropy for subsystem A is

$$S_A^{\text{qp}} = N_A - \frac{N_A^2}{N},$$

which deviates considerably from the dynamical Page curve discussed in the main text.