Equilibration via Gaussification in fermionic lattice systems

M. Gluza,[1] C. Krumnow,[1] M. Friesdorf,[1] C. Gogolin,[2,3] and J. Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*
[3]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*

(Dated: January 6, 2016)

The perspective of probing quantum many-body systems out of equilibrium under well controlled conditions is attracting enormous attention in recent years, a perspective that extends to the study of fermionic systems. In this work, we present an argument that precisely captures the dynamics causing equilibration and Gaussification under quadratic non-interacting fermionic Hamiltonians. Specifically, based on two basic assumptions — the initial clustering of correlations and the Hamiltonian exhibiting delocalizing transport — we prove that systems become locally indistinguishable from fermionic Gaussian states on precisely controlled time scales. The argument gives rise to rigorous instances of a convergence to a generalized Gibbs ensemble. This argument is general enough to allow for arbitrary pure and mixed initial states, including thermal and ground states of interacting models, and large classes of systems, including high-dimensional lattice and classes of spin systems. Our results allow to develop an intuition of equilibration that is expected to be generally valid, and at the same time relates to current experiments of cold atoms in optical lattices.

Despite the great complexity of quantum many-body systems out-of-equilibrium, local expectation values in such systems show the remarkable tendency to equilibrate to stationary values that do not depend on the microscopic details of the initial state, but rather can be described with few parameters using thermal states or generalized Gibbs ensembles [1–3]. Such behavior has been successfully studied in many settings theoretically and experimentally, most notably in instances of quantum simulations in optical lattices [2, 4, 5].

By now, it is clear that, despite the unitary nature of quantum mechanical evolution, local expectation values equilibrate due to a dephasing between the eigenstates [3, 6–11]. So far it is, however, unclear why this dephasing tends to happen so rapidly. In fact, experiments often observe equilibration on time scales independent of the system size [5, 12], while even the best general theoretical bounds diverge exponentially [2, 11]. This discrepancy is known as the problem of equilibration time scales.

What is more, only little is known about how exactly the equilibrium expectation values emerge. Due to the exponentially many constants of motion present in quantum many-body systems, corresponding to the overlaps with the eigenvectors of the system, there seems to be no obvious reason why equilibrium values often only depend on few macroscopic properties such as temperature or particle number. In short: It is unclear how precisely the memory of the initial conditions is lost during time evolution.

To make progress towards a solution of the two problems introduced above, it is instructive to study the behavior of non-interacting particles captured by so-called quadratic or free models, which provide a guideline for understanding these issues in significantly more depth. Thermal states of these models are fully described by the correlation matrix and are hence instances of so-called Gaussian states, which are thus particularly simple to describe. While studying the time evolution of such states provides valuable insight into the spreading of particles and equilibration, it does not explain how the state ends up appearing Gaussian in the first place.

In this work, we address this question: We show that after a short and well controlled relaxation time very general non-Gaussian initial states become locally indistinguishable from a Gaussian state with the same second-moments. This mechanism is much reminiscent and shares many features of actual thermalisation, in that an initially complex setting appears to be convergent to a state that is captured by astoundingly few parameters only. In this way, we present a significant step forward in the theory of equilibration of quantum many-body systems that have been pushed out of equilibrium.

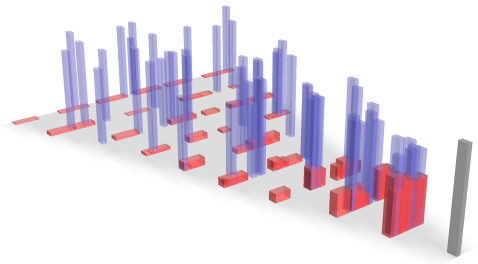

Figure 1. Illustration of the Gaussification process. Expanding a time evolving local operator in the basis of creation and annihilation operators allows us to separate Gaussian (two body, dark blue) and non-Gaussian (multi-body, bright red) contributions to its expectation value in states with exponential clustering of correlations. Delocalizing transport leads to an algebraic suppression of the non-Gaussian contributions in time.

Our results hold for a remarkably large class of initial conditions, including ground states of interacting models, evolving in time under a quadratic fermionic Hamiltonian with finite ranged interactions. This family of Hamiltonians notably includes the case of non-interacting ultra-cold fermions in an optical lattice. By virtue of the Jordan-Wigner transformation our results also apply to certain spin models. We formulate our results in form of a rigorous proof, which at the same time provides an intuitive explanation for this result. In particular, we identify two natural assumptions, namely exponential clustering of correlations in the initial state and what we call delocalizing transport, which provide the basis for our

proof and are immediately connected to the intuition underlying equilibration.

Setting. Let $\mathcal{L}$ be a $d_{\mathcal{L}}$-dimensional cubic lattice with V lattice sites. We restrict to cubic lattices purely for notational convenience, all results can be straight forwardly generalized to Kagomé, honeycomb, or other geometries. Restricting ourself to spin-less fermions, again purely out of notational convenience, each site $r \in \mathcal{L}$ is associated with a fermionic orbital with fermionic creation and annihilation operators $f_r^\dagger$ and f_r. We collect all of them in a vector $c = (f_1, f_1^\dagger, \ldots, f_V, f_V^\dagger)$. The Hamiltonian of a quadratic fermionic system is then of the form

$$H = \sum_{j,k=1}^{2V} c_j^\dagger \, h_{j,k} \, c_k, \tag{1}$$

with $h = h^\dagger$ its hermitian Hamiltonian kernel. The time evolution of the annihilation operator in the Heisenberg picture under such a Hamiltonians is given by (see Appendix A)

$$c_j(t) \coloneqq \mathrm{e}^{\mathrm{i}Ht} \, c_j \, \mathrm{e}^{-\mathrm{i}Ht} = \sum_{k=1}^{2V} W_{j,k}(t) \, c_k \tag{2}$$

with the propagator $W(t) \coloneqq \mathrm{e}^{2th}$.

Next we introduce the concept of Gaussian states and Gaussification. Define the correlation matrix γ of a state ρ as the matrix of its second moments, i.e., $\gamma_{j,k} \coloneqq \mathrm{tr}(\rho \, c_j^\dagger \, c_k)$. A convenient characterization of Gaussian states is the following: They are the states that maximize the von Neumann entropy given the expectation values collected in the correlation matrix (see also Appendix C 1). For every state ρ, we hence define its Gaussified version ρ_G as the Gaussian state with the correlation matrix of ρ, i.e., $\mathrm{tr}(\rho_G \, c_j^\dagger \, c_k) = \mathrm{tr}(\rho \, c_j^\dagger \, c_k)$.

Assumptions. Our main theorem holds for initial states (including non-Gaussian ones) with a certain form of decay of correlations that evolve under quadratic Hamiltonians that exhibit a certain form of transport. We now make these two conditions precise, starting with the correlation decay:

Definition 1 (Exponential clustering of correlations)**.** *We say that a state ρ exhibits exponential clustering of correlations with length scale $\xi > 0$ and constant $C_{\mathrm{Clust}} > 0$ if, for any two operators A, B with $\|A\| = \|B\| = 1$, we have*

$$\begin{aligned} &|\,\mathrm{tr}(\rho \, A \, B) - \mathrm{tr}(A \, \rho) \; \mathrm{tr}(B \, \rho)| \\ &\leq C_{\mathrm{Clust}} \; \mathrm{supp}(A) \; \mathrm{supp}(B) \, \mathrm{e}^{-\,\mathrm{d}(A,B)/\xi}. \end{aligned} \tag{3}$$

Here $\mathrm{d}(A, B)$ is taken to be the natural distance on the lattice between the supports $\mathrm{supp}(A), \mathrm{supp}(B)$ of A and B and $\|\cdot\|$ denotes the operator norm. The above definition is slightly less demanding than the standard definition of exponential clustering of correlations (we allow for a scaling with the volume of the supports rather than just their surfaces) and a wide classes of physically relevant states exhibit this form of correlation decay. It is important to stress that the initial state is not assumed to be a Gaussian state. It may for example perfectly well be a ground state of an interacting gapped model

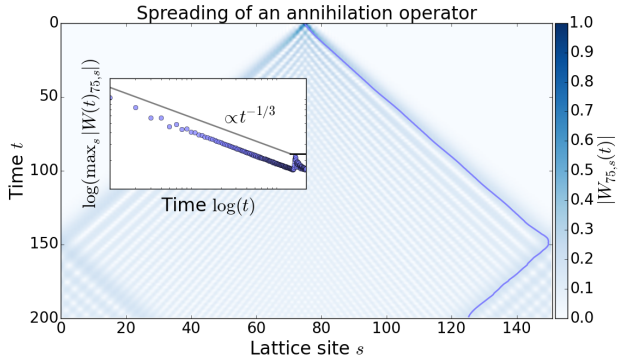

Figure 2. Numerical study of the spreading of the support of a single fermionic annihilation operator described in Eq. (2) under the evolution of the free hopping Hamiltonian $H = -\sum_j (f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j)$ with unit hopping amplitude on a one-dimensional chain of 150 sites with periodic boundary conditions. The blue color denotes the modulus of the evolution matrix elements $W_{75,k}(t)$ of a fermionic operator initially supported on site 75 in the center of the chain. The support expands ballistically creating the Lieb-Robinson cone. Within the cone the support is distributed fairly evenly. The maximal value, indicated by the smoothed blue curve, is encountered at the wavefronts and suppressed by at least $t^{-1/3}$ as shown in the inset. The suppression stabilises, once the wavefronts collide.

[13, 14] or a high temperature state of a locally interacting non-quadratic fermionic system [15].

For our proof of local relaxation towards a Gaussian state we further assume that the Hamiltonian exhibits transport in the following sense:

Definition 2 (Delocalizing transport)**.** *A quadratic Hamiltonian with propagator W on a $d_{\mathcal{L}}$-dimensional cubic lattice of volume V exhibits delocalizing transport with constant $C_{\mathrm{Trans}} > 0$ and recurrence time $t_{\mathrm{Rec}} > 0$ if, for all $t \in (0, t_{\mathrm{Rec}}]$, we have that*

$$\forall j, k: \quad |W_{j,k}(t)| \leq C_{\mathrm{Trans}} \max\{t^{-d_{\mathcal{L}}/3}, V^{-d_{\mathcal{L}}/3}\}. \tag{4}$$

The intuition behind this definition is that an initially localized fermionic operator will spread over a large area, such that its component on a single localized operator is dynamically suppressed. For our main result any exponent smaller than $-d_{\mathcal{L}}/4$ is sufficient. The choice of the exponent $-d_{\mathcal{L}}/3$ in the definition is justified by the fact that this is what can be shown to hold for important classes of models: Quadratic hopping Hamiltonians with constant on-site potential (see Fig. 2 and Appendix B 1), as well as the critical Ising model (see Appendix B 2) show this kind of transport, as well as numerous others.

The recurrence time takes into account that any non-trivial bound of the form (4) is eventually violated due to the recurrent nature of the dynamics of finite dimensional quantum systems. For free hopping Hamiltonians, it can be shown that the recurrence time grows at least like a power-law with the system size V (see Appendix B 2), but it is usually exponentially large.

Main result. With the above definitions our main result can be stated as follows:

Theorem 1 (Gaussification in finite time)**.** *Consider a family of systems on cubic lattices of increasing volume V. Let the initial states exhibit exponential clustering of correlations and let the Hamiltonians be quadratic finite range and have delocalizing transport. If all constants are independent of the system size, except the recurrence time t_{Rec}, which increases as some function of V, then for any local operator A on a fixed finite region and error $\epsilon > 0$ there exists a relaxation time $t_{\mathrm{Relax}} > 0$ independent of the system size such that for all times $t \in [t_{\mathrm{Relax}}, t_{\mathrm{Rec}}]$ it holds that $|\operatorname{tr}[A(t)\,\rho] - \operatorname{tr}[A(t)\,\rho_G]| \leq \epsilon$.*

The theorem states that for all times in $[t_{\mathrm{Relax}},\, t_{\mathrm{Rec}}]$ it will be impossible by any local measurements on S to distinguish the true state ρ from the fermionic Gaussian state ρ_G, which has the same second moments as ρ. Note that since t_{Relax} is independent of the system size, but t_{Rec} increases with its volume, for any ϵ there always exists a system size such that $t_{\mathrm{Rec}} > t_{\mathrm{Relax}}$ and the interval where the theorem applies grows as a function of the system size. We complement this analytical and rigorous approach by a tensor network simulation, the results of which are depicted in Fig. 3, which clearly shows the algebraic suppression of non-Gaussian feature in time.

The argument can be understood as a fermionic quantum central limit theorem [16] emerging from a dynamical evolution, in that the second moments are preserved and the higher cumulants can be proven to converge to zero in time. The key steps in the proof, to be presented in the following, can be understood from intuitive physical considerations. They are based on three main ingredients: the finite speed of propagation in lattice systems, the homogeneous suppression of matrix elements of the propagator due to delocalizing transport, and the exponential clustering of correlations in the initial state. For the full proof with all details of the involved combinatorics we refer the reader to Appendix (see Appendix D 5 for an overview). The proof shares some intuition put forth on the equilibration of harmonic classical chains [17].

Proof. We expand the local operator A supported in a fixed finite region S in the basis of fermionic operators. To that end let $\tilde{S} \coloneqq \{s_r\}$ for $r \in [2|S|] \coloneqq \{1, \ldots, 2|S|\}$ be the set of indices of all elements of the vector c with support in S, then

$$A(t) = \sum_{b_1,\ldots,b_{2|S|}=0}^{1} a_{b_1,\cdots,b_{2|S|}}\, c_{s_1}(t)^{b_1} \ldots c_{s_{2|S|}}(t)^{b_{2|S|}} \,. \quad (5)$$

Without loss of generality we can assume normalization $\|A\| = 1$, such that all of the $2^{2|S|}$ coefficients satisfy $|a_{b_1,\cdots,b_{2m}}| \leq 1$. Thus

$$\begin{aligned} &|\operatorname{tr}[A\,\rho(t)] - \operatorname{tr}[A\,\rho_G(t)]| \quad (6)\\ &\leq 2^{2|S|} \max_{J\subset\tilde{S}} \Bigg| \sum_{(k_j)_{j\in J}\in[2V]^{\times|J|}} \operatorname{tr}\Bigg[\prod_{j\in J} W_{j,k_j}(t)\, c_{k_j}\,(\rho-\rho_G)\Bigg]\Bigg| \,. \end{aligned}$$

Here and in the following all products are meant to be preformed in increasing order.

We assumed that the Hamiltonian has finite range interactions, i.e., there exists a fixed length scale l_0, such that whenever $(j,k) > l_0$ it holds that $h_{j,k} = 0$, where we have used the shorthand $\mathrm{d}(j,k) \coloneqq \mathrm{d}(c_j, c_k)$. Such models satisfy Lieb-Robinson bounds [18], which in our free fermionic setting can be stated as follows:

Lemma 2 (Lieb-Robinson bound for quadratic systems [19])**.** *For any quadratic fermionic Hamiltonians H with finite range interactions there exist constants $C_{\mathrm{LR}}, \mu, v > 0$ independent of the system size such that its propagator W fulfills the bound*

$$|W_{j,k}(t)| \leq C_{\mathrm{LR}}\, \mathrm{e}^{\,\mu\,(v|t| - \mathrm{d}(j,k))} \,. \quad (7)$$

The Lieb-Robinson bound tells us that $c_j(t)$ and $c_k(t)$ essentially still have disjoint support as long as t is small enough such that $v|t| \ll \mathrm{d}(j,k)$. We can hence restrict the sum in Eq. (6) to those k_j whose $\min_{s\in\tilde{S}} \mathrm{d}(k_j, s)$ is smaller than $(v + 2\,v_\epsilon)|t|$ for some fixed $v_\epsilon > 0$. The total contribution of the neglected terms can be bounded explicitly (see Appendix D 1 for details) and, importantly, is independent of V and exponentially suppressed in $|v_\epsilon\, t|$.

For each of the remaining summands in Eq. (6) it is now important to keep track of the distribution of the indices k_j inside the cone. For this purpose we define the Δ-partition P_Δ of a subindex set $J \subset \tilde{S}$ and sequence of indices $(k_j)_{j\in J}$ as the unique decomposition of J into subsets (patches) p in the following way: The patches are constructed such that for any two subindices within any given patch p there is a connecting chain of elements from that patch in the sense that the distance between two consecutive c_{k_j} with $j \in p$ along that chain is not greater than Δ and the distance between any two $c_{k_j}, c_{k'_j}$ with j, j' from different patches is larger than Δ. For each patch p in the Δ-partition of a given summand in Eq. (6) we define a corresponding operator

$$\hat{P}_p^{(k_j)_{j\in p}} \coloneqq \prod_{j\in p} W_{j,k_j}(t)\, c_{k_j} \,. \quad (8)$$

We can then reorder the factors in Eq. (6) to write the product as a product over these operators. The exponential clustering of correlations (Definition 1) in the initial state allows us to factor the patches if we scale Δ suitably with $|t|$. Concretely, for $\sigma \in \{\rho, \rho_G\}$ the expectation values appearing in Eq. (6), which we denote by $\langle\cdot\rangle_\sigma$, can be approximated as follows

$$\Bigg\langle \prod_{p\in P_\Delta} \hat{P}_p^{(k_j)_{j\in p}} \Bigg\rangle_\sigma \approx \prod_{p\in P_\Delta} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \,. \quad (9)$$

The error introduced in the factorization is exponentially suppressed with the ratio of patch distance and correlation length Δ/ξ.

It remains to bound the contribution from the factorized patches that are completely inside the Lieb-Robinson cone. Note that the right hand side of Eq. (9) can be non-zero only if all the patches are of even size, as ρ and ρ_G have an even particle number parity. Moreover, as the second moments

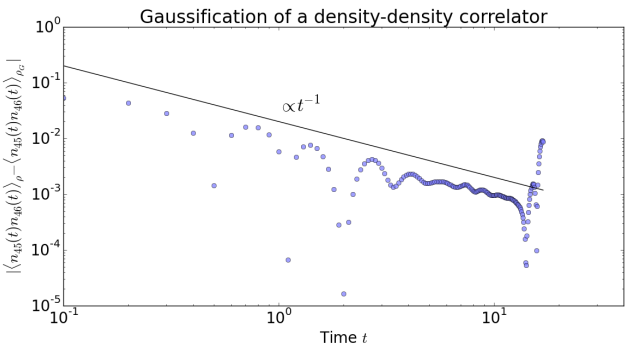

Figure 3. Numerical study of the evolution of the nearest-neighbor density-density correlator for a correlated initial state under the free hopping Hamiltonian with hopping amplitude one. The initial state was obtained by a matrix-product state algorithm ground state search of an interacting spinless Fermi-Hubbard model $H = -\sum_j (f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j) + U \sum_j n_j n_{j+1} + \sum_j \omega_j n_j$ on 64 sites. The system was placed at half filling with periodic boundary conditions and hopping amplitude one, interaction strength $U = 2$, and weak onsite disorder w_j drawn at random independently and identically distributed from a Gaussian with variance $1/4$. The time evolution was performed by an explicit evolution of the operator in the Heisenberg picture using Eq. (2). Shown is the difference of expectation values between the initial state ρ and its Gaussified version ρ_G for a nearest-neighbor density-density correlator on sites 45 and 46 as a function of time. As indicated by the curve scaling as t^{-1}, which serves as a guide to the eye, the difference is suppressed algebraically. At late times, a recurrence due to the finite size of the system occurs, leading to an increase in the difference between the state and its Gaussified version.

of ρ and ρ_G are equal, the difference of the right hand side for $\sigma = \rho$ and $\sigma = \rho_G$ vanishes whenever all patches have size 2. Hence, only partitions that contain at least one patch of size at least 4 can contribute. The delocalizing transport of the Hamiltonian implies that the contribution from such larger patches however is dynamically suppressed. Whenever $|p| \geq 4$ it holds that

$$|\langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma| \leq C^4_{\mathrm{Trans}}\, t^{-4\, d_{\mathcal{L}}/3} \tag{10}$$

as long as V is large enough. The influence of possible patches of size 2 in the same decomposition makes it necessary to bound the overall contribution with an involved recursive and combinatorial argument. However, effectively the dynamical suppression stated in the last inequality allows us to derive a bound that increases with the patch size Δ but is algebraically suppressed in time t. The increase with Δ is a consequence of the fact that allowing for longer distances between the elements of a patch increases the number of possible patches of a given size. Finally, by choosing $\Delta = t^{d_{\mathcal{L}}/24}$ one can obtain an at least algebraic suppression with t of all terms and thereby of the difference $|\operatorname{tr}(A(t)\rho) - \operatorname{tr}(A(t)\rho_G)|$. □

Physical implications and application. The Gaussification result presented above also has profound implications for the study of equilibration of quantum many-body systems. Whenever the second moments equilibrate, which is often observed [5, 20–24], the results imply that the full reduced density matrix becomes stationary. The quadratic models considered here constitute a "theoretical laboratory", in which the mechanisms of Gaussification and equilibration can be very precisely and quantitatively characterized, and all specifics of the processes laid out.

This does not mean that the physics we address is very specific to precisely these quadratic Hamiltonians: Quite to the contrary, we expect the fundamental mechanisms underlying the result — local relaxation due to transport and initial clustering of correlations — to be generic and the reason for relaxation in a wide classes of interacting models [25]. The intuition, reminiscent of a quantum central limit theorem, that incommensurate influences of further and further separated regions lead to mixing and relaxation is then expected to still be valid. It is also important to stress that our main theorem equally applies to mixed initial states, such as thermal states, which are relevant in present day experiments with ultra-cold fermions [26–29].

Returning to the specifics of quadratic Hamiltonians, the result derived here can be interpreted in yet another way: It constitutes a rigorous proof of convergence to a *generalised Gibbs ensemble* (GGE). Fermionic Gaussian states are the maximum entropy states given their second moments so that the resulting state can be captured in terms of a GGE with suitable Lagrange parameters. The same holds true for the integrable spin models that can be mapped to the type of fermionic models considered here. The results complement insights on bosonic systems [6, 30] and are expected to largely carry over to continuous settings as well in which a convergence to a Gaussian generalised Gibbs ensemble has been observed [**?**].

Conclusion and outlook. In this work we have established an understanding of how systems quenched to non-interacting fermionic Hamiltonians locally converge to Gaussian states. Out of equilibrium dynamics is identified as having the tendency to bring systems locally in maximum entropy states given the second moments. This holds even if the initial state was far from being a Gaussian state, e.g., a ground state of a strongly interacting model. This is achieved based on just two natural assumptions: A form of delocalizing transport in the model and exponential clustering of correlations in the initial state. Otherwise the initial state can be completely general. It is the hope that the present work will serve as a stepping stone to gain further insights into the relaxation dynamics of more complex quantum many-body systems and the consequences of the suppression of transport in, for example, localizing systems.

Acknowledgements

We would like to thank the EU (RAQUEL, SIQS, AQuS), the ERC (TAQ, OSYRIS), the BMBF, and the Studienstiftung des Deutschen Volkes, MPQ-ICFO, the Spanish Ministry Project FOQUS (FIS2013-46768-P), the Generalitat de Catalunya (SGR 874 and 875), the Spanish MINECO (Severo Ochoa grant SEV-2015-0522), and the COST Action MP1209

for support. We warmly acknowledge fruitful discussions with M. Cramer, M. M. Wolf and P. Ćwikliński.

[1] Polkovnikov, A., Sengupta, K., Silva, A. & Vengalattore, M. Non-equilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.* **83**, 863 (2011).

[2] Eisert, J., Friesdorf, M. & Gogolin, C. Quantum many-body systems out of equilibrium. *Nature Phys.* **11**, 124 (2015).

[3] Gogolin, C. & Eisert, J. Equilibration, thermalization, and the emergence of statistical mechanics in closed quantum systems. arXiv:1503.0753 (accepted in Rep. Prog. Phys).

[4] Bloch, I., Dalibard, J. & Nascimbene, S. Quantum simulations with ultracold quantum gases. *Nature Phys.* **8**, 267 (2012).

[5] Trotzky, S. *et al.* Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas. *Nature Phys.* **8**, 325 (2012).

[6] Cramer, M., Dawson, C. M., Eisert, J. & Osborne, T. J. Exact relaxation in a class of non-equilibrium quantum lattice systems. *Phys. Rev. Lett.* **100**, 030602 (2008).

[7] Linden, N., Popescu, S., Short, A. J. & Winter, A. Quantum mechanical evolution towards thermal equilibrium. *Phys. Rev. E* **79**, 061103 (2009).

[8] Short, A. J. & Farrelly, T. C. Quantum equilibration in finite time. *New J. Phys.* **14**, 013063 (2012).

[9] Rigol, M., Dunjko, V., Yurovsky, V. & Olshanii, M. Relaxation in a completely integrable many-body quantum system: An ab initio study of the dynamics of the highly excited states of 1D lattice hard-core bosons. *Phys. Rev. Lett.* **98**, 050405 (2007).

[10] Reimann, P. & Kastner, M. Equilibration of isolated macroscopic quantum systems. *New J. Phys.* **14**, 043020 (2012).

[11] Malabarba, A. S. L., Garcia-Pintos, L. P., Linden, N., Farrelly, T. C. & Short, A. J. Quantum systems equilibrate rapidly for most observables. *Phys. Rev. E* **90**, 012121 (2014).

[12] Pertot, D. *et al.* Relaxation dynamics of a Fermi gas in an optical superlattice. *Phys. Rev. Lett.* **113**, 170403 (2014).

[13] Hastings, M. B. & Koma, T. Spectral gap and exponential decay of correlations. *Commun. Math. Phys.* **265**, 781–804 (2006).

[14] Nachtergaele, B. & Sims, R. Locality estimates for quantum spin systems. *Selected contributions of the XVth International Congress on Mathematical Physics* (2007).

[15] Kliesch, M., Gogolin, C., Kastoryano, M. J., Riera, A. & Eisert, J. Locality of temperature. *Phys. Rev. X* **4**, 031019 (2014).

[16] Hudson, R. L. A quantum-mechanical central limit theorem for anti-commuting observables. *J. Appl. Prob.* **10**, 502–509 (1973).

[17] Dudnikova, T. V., Komech, A. & Spohn, H. On the convergence to statistical equilibrium for harmonic crystals. *J. Math. Phys.* **44**, 2596 (2003).

[18] Lieb, E. H. & Robinson, D. W. The finite group velocity of quantum spin systems. *Commun. Math. Phys.* **28**, 251–257 (1972).

[19] Hastings, M. B. Decay of correlations in Fermi systems at nonzero temperature. *Phys. Rev. Lett.* **93**, 126402 (2004).

[20] Flesch, A., Cramer, M., McCulloch, I. P., Schollwöck, U. & Eisert, J. Probing local relaxation of cold atoms in optical superlattices. *Phys. Rev. A* **78**, 033608 (2008).

[21] Calabrese, P. & Cardy, J. Time dependence of correlation functions following a quantum quench. *Phys. Rev. Lett.* **96**, 136801 (2006).

[22] Calabrese, P., Essler, F. H. & Fagotti, M. Quantum quench in the transverse field Ising chain II: Stationary state properties. *J. Stat. Mech.* P07022 (2012).

[23] Peschel, I. & Eisler, V. Reduced density matrices and entanglement entropy in free lattice models. *J. Phys. A* **42**, 504003 (2009).

[24] Fagotti, M. & Essler, F. H. L. Reduced density matrix after a quantum quench. *Phys. Rev. B* **87**, 245107 (2013).

[25] Manmana, S. R., Wessel, S., Noack, R. M. & Muramatsu, A. Strongly correlated fermions after a quantum quench. *Phys. Rev. Lett.* **98**, 210405 (2007).

[26] Schneider, U. *et al.* Fermionic transport and out-of-equilibrium dynamics in a homogeneous Hubbard model with ultracold atoms. *Nature Phys.* **8**, 213 (2012).

[27] Parsons, M. F. *et al.* Site-resolved imaging of fermionic lithium-6 in an optical lattice. *Phys. Rev. Lett.* **114**, 213002 (2015).

[28] Strohmaier, N. *et al.* Interaction-controlled transport of an ultracold Fermi gas. *Phys. Rev. Lett.* **99**, 220601 (2007).

[29] Haller, E. *et al.* Single-atom imaging of fermions in a quantum-gas microscope. *Nature Phys.* **11**, 738 (2015).

[30] Cramer, M. & Eisert, J. A quantum central limit theorem for non-equilibrium systems: Exact local relaxation of correlated states. *New J. Phys.* **12** (2010).

[31] Kitaev, A. Anyons in an exactly solved model and beyond. *Ann. Phys.* **321**, 2 (2006).

[32] Lieb, E., Schultz, T. & Mattis, D. Two soluble models of an anti-ferromagnetic chain. *Ann. Phys.* **16**, 407–466 (1961).

[33] Kraus, C. V., Wolf, M. M., Cirac, J. I. & Giedke, G. Pairing in fermionic systems: A quantum-information perspective. *Phys. Rev. A* **79**, 012306 (2009).

[34] van Lint, J. H. & Wilson, R. M. *A course in combinatorics* (Cambridge University Press, 1992).

Appendix A: Majorana operators and time evolution in free fermionic systems

In this appendix, we formulate the Majorana operator description that allows to conveniently derive the operator governing transport in the system. We introduce the Majorana operators as

$$m_{2j-1} \coloneqq (f_j^\dagger + f_j)/\sqrt{2}, \tag{A1}$$

$$m_{2j} \coloneqq \mathrm{i}\,(f_j^\dagger - f_j)/\sqrt{2}, \tag{A2}$$

which are collected in a vector $m = (m_1, \ldots, m_{2V})$ [31]. The vector c of creation and annihilation operators used in the main text and m are related by the unitary transformation

$$\Omega \coloneqq \frac{1}{\sqrt{2}} \bigoplus_{j=1}^{V} \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix}, \tag{A3}$$

as $m = \Omega\, c$. The Majorana operators are Hermitian and satisfy the anti-commutation relations $\{m_j, m_k\} = \delta_{j,k}$ for $j, k \in [2V] \coloneqq \{1, \ldots, 2V\}$. The algebra generated by those operators constitutes a Clifford algebra. Linear transformations of the form

$$m'_j = \sum_{j,k=1}^{2V} O_{j,k}\, m_k, \quad O \in SO(2V) \tag{A4}$$

transform a vector of legitimate Majorana operators to a new such vector.

The most general form of a Hamiltonian considered in this work can be written in terms of the Majorana operators as follows

$$H = \mathrm{i} \sum_{j,k=1}^{2V} m_j\, K_{j,k}\, m_k\ , \tag{A5}$$

where $K = -K^T$ is real and anti-symmetric. It is straightforward to relate such Hamiltonians to the ones expressed in the form of Eq. (1). The kernel K can be obtained form h via $K = -\mathrm{i}\,\Omega\, h\, \Omega^\dagger$.

Time evolution can be captured conveniently in the Majorana operator formulation. Using the Baker-Campbell-Hausdorff formula, that K is anti-symmetric, and the algebraic structure of the Majorana fermions, one arrives at the following expression for their time evolution in the Heisenberg picture

$$m_j(t) \coloneqq \mathrm{e}^{\mathrm{i}Ht} m_j \mathrm{e}^{-\mathrm{i}Ht} = \sum_{k=1}^{2V} (\mathrm{e}^{2tK})_{j,k}\, m_k = \sum_{k=1}^{2V} L_{j,k}(t)\, m_k, \tag{A6}$$

where $L(t) \coloneqq \mathrm{e}^{2tK}$. Now notice that as the propagator defined in the main text is related to $L(t)$ via

$$W(t) = \Omega^\dagger\, L(t)\, \Omega \tag{A7}$$

and hence

$$c_j(t) = \sum_{k=1}^{2V} W_{j,k}(t)\, c_k \tag{A8}$$

as claimed in the main text.

Further, we introduce some general notation which we will use in the following for any given operator A that is supported on a region S. For this set, let $\tilde{S} \coloneqq \{s_1, \cdots, s_{2|S|}\}$ with $s_1 < s_2 < \ldots < s_{2|S|}$ be the set of indices of fermionic basis operators in S. Similarly as in Eq. (5) in the main text we can then expand A as

$$A = \sum_{b_1, \ldots, b_{2|S|}=0}^{1} a(\{s_r : b_r = 1\})\, c_{s_1}^{b_1} \ldots c_{s_{2|S|}}^{b_{2|S|}}\ , \tag{A9}$$

with $a(\{s_r : b_r = 1\}) = a_{b_1, \ldots, b_{2|S|}}$. The sum in Eq. (A9) goes over all possible configurations of fermionic basis operators on the region S. We can hence group the summands according to the subset J of indices from $\tilde{S}$ for which a given term actually contains a fermionic basis operator and write it as a sum

$$A = \sum_{J \subset \tilde{S}} a(J)\, A_J\ , \tag{A10}$$

with

$$A_J \coloneqq \prod_{j\in J} c_j \ . \tag{A11}$$

Here, as well as everywhere where such expressions appear, we assume that the product over j is ordered according to the lattice ordering. The time evolution of A is then given by $A(t) = \sum_{J\subset\tilde{S}} a(J)\, A_J(t)$ with

$$A_J(t) = \sum_{(k_j)_{j\in J}\in[2V]^{\times|J|}} \left(\prod_{j\in J} W_{j,k_j}(t)\, c_{k_j}\right) . \tag{A12}$$

Appendix B: Transport

In the following we show that our definition of delocalizing transport (Def. 2) is matched by two prototypical examples: the fermionic nearest neighbor hopping model as well as the fermionisation of the Ising model. Our proofs rely on the following Lemma [30].

Lemma 3 (Spreading for hopping models)**.** *Given the real $V \times V$ matrix*

$$M = \begin{pmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ & & 1 & 0 & \\ 1 & & & \ddots & \end{pmatrix}, \tag{B1}$$

we find for all $0 < t \leq V^{6/7}$

$$|(e^{-\mathrm{i}tM})_{j,k}| \leq \frac{37}{|t|^{1/3}}. \tag{B2}$$

We begin with free hopping Hamiltonians.

1. Transport in fermionic nearest neighbor hopping models in square lattices

Consider fermionic hopping models, i.e., systems whose Hamiltonian is a linear combination of terms of the form $f_j^\dagger f_k$. Instead of the general quadratic form (1), the Hamiltonian can then be written as

$$H = \sum_{j,k=1}^{V} f_j^\dagger M_{j,k} f_k, \tag{B3}$$

with M a real and symmetric matrix, i.e., $M = M^T$. The time evolution of fermionic annihilation operators in the Heisenberg picture is then given by

$$f_j(t) = \mathrm{e}^{\mathrm{i}Ht}\, f_j\, \mathrm{e}^{-\mathrm{i}Ht} = \sum_{k=1}^{V} N_{j,k}(t)\, f_k, \tag{B4}$$

where $N(t) \coloneqq \mathrm{e}^{-\mathrm{i}Mt}$. To connect this to the notation used in the main text, note that with P the permutation matrix that acts as

$$(f_1, f_1^\dagger, \ldots, f_V, f_V^\dagger) = P(f_1, \ldots, f_V, f_1^\dagger, \ldots, f_V^\dagger). \tag{B5}$$

the Hamiltonian from Eq. (B3) can be written in the form of Eq. (1) by adding an appropriate constant and setting

$$h = \frac{1}{2} P\,(M \oplus -M)\, P^\dagger. \tag{B6}$$

$N(t)$ is then related to $W(t)$, via

$$W(t) = P(N(t) \oplus N(t)^\dagger) P^\dagger. \tag{B7}$$

Now consider the particularly important case of nearest neighbor hopping on a square lattice of spacial dimension $d_{\mathcal{L}}$ with V sites, periodic boundary conditions and unite hopping strength. Writing the Hamiltonian as in (B3) the coupling matrix M of this model can be decomposed into a sum over the different $d_{\mathcal{L}}$ spatial directions as follows

$$M = \sum_{k=0}^{d_{\mathcal{L}}-1} \mathbb{1}^{\otimes k} \otimes M^{(1)} \otimes \mathbb{1}^{\otimes (d_{\mathcal{L}}-k-1)}, \tag{B8}$$

with

$$M^{(1)} \coloneqq -\begin{pmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ & & 1 & 0 & \\ 1 & & & & \ddots \end{pmatrix}. \tag{B9}$$

For such models we can bound the spreading as follows:

Lemma 4 (Spreading for fermionic hopping models)**.** *For the fermionic nearest neighbor hopping model on a square lattice of spacial dimension $d_{\mathcal{L}}$ with periodic boundary conditions there is a constant $C_{\mathrm{Trans}} > 0$ independent of volume V of the lattice and the recurrence time $t_{\mathrm{Rec}} = V^{6/7d_{\mathcal{L}}}$ such that for all $t \in]0, t_{\mathrm{Rec}}]$ it holds that*

$$\forall j,k \qquad |W_{j,k}(t)| \leq C_{\mathrm{Trans}}\, t^{-d_{\mathcal{L}}/3}. \tag{B10}$$

Proof. From the structure of the Hamiltonian it follows that $N(t) = (\mathrm{e}^{i(d_{\mathcal{L}}-1)t}\mathrm{e}^{\mathrm{i}M^{(1)}t})^{\otimes d_{\mathcal{L}}}$ where $M^{(1)}$ agrees up to a sign with the matrix in Lemma 3. From Lemma 3 we obtain that for all $0 < t < t_{\mathrm{Rec}} = V^{6/7d_{\mathcal{L}}}$ and j,k

$$|N_{j,k}(t)| \leq \frac{37^{d_{\mathcal{L}}}}{t^{d_{\mathcal{L}}/3}}. \tag{B11}$$

This bound is inherited by W as N and W are related by a permutation of rows and columns. □

2. Spreading in the Ising model

In this appendix, we discuss the 1D Ising model and specifically show that the required conditions of transport are satisfied at criticality. Its Hamiltonian for V sites is

$$H_{\mathrm{IS}} = -\sum_{j=1}^{V} X_j X_{j+1} - g\sum_{j=1}^{V} Z_j\ , \tag{B12}$$

where X_j, Z_j are the Pauli matrices supported on site, j and g are a real parameter, and we adopt periodic boundary conditions. For convenience we restrict the discussion to V even. Invoking the Jordan-Wigner transformation [32], this spin system can be mapped to fermions, using the substitutions

$$Z_j \mapsto f_j f_j^\dagger - f_j^\dagger f_j = 1 - 2n_j\ , \tag{B13}$$

$$S_j = \frac{1}{2}(X_j - \mathrm{i}Y_j) \mapsto \prod_{l<j}(1-2n_l) f_j\ , \tag{B14}$$

$$S_j^\dagger = \frac{1}{2}(X_j + \mathrm{i}Y_j) \mapsto \prod_{l<j}(1-2n_l) f_j^\dagger\ , \tag{B15}$$

where S_j is the spin annihilation operator associated with site j and $n_j = f_j^\dagger f_j$ denotes the usual fermionic number operator. After this transformation, the Ising Hamiltonian takes the form,

$$H_{\mathrm{IS}} = -\sum_{j=1}^{V}(f_j^\dagger + f_j)(1-2n_j)(f_{j+1}^\dagger + f_{j+1}) - g\sum_{j=1}^{V}(1-2n_j) \tag{B16}$$

$$= -\sum_{j=1}^{V}(f_j^\dagger - f_j)(f_{j+1}^\dagger + f_{j+1}) - g\sum_{j=1}^{V}(1-2n_j)\ . \tag{B17}$$

Using the Majorana operators introduced in Appendix A, we can rewrite the Hamiltonian as

$$H_{\mathrm{IS}} = \mathrm{i}\sum_{j=1}^{V-1}(m_{2j}m_{2j+1} - m_{2j+1}m_{2j}) - \mathrm{i}g\sum_{j=1}^{V}(m_{2j}m_{2j-1} - m_{2j-1}m_{2j}). \tag{B18}$$

That is, it is of the form

$$H = \mathrm{i}\sum_{j,k=1}^{2V} m_j K_{j,k} m_k \tag{B19}$$

with

$$K = \begin{pmatrix} 0 & g & & & & \\ -g & 0 & 1 & & & \\ & -1 & 0 & g & & \\ & & -g & 0 & 1 & \\ & & & -1 & 0 & \\ & & & & & \ddots \end{pmatrix}. \tag{B20}$$

In what follows, we will consider a special case. First, we will set $g = 1$, corresponding to the critical Ising model. Second, we will make use of periodic boundary conditions, to render the discussion simpler. The model actually considered, therefore, is the Hamiltonian (B19), with the circulant kernel

$$K = \begin{pmatrix} 0 & 1 & & & & & -1 \\ -1 & 0 & 1 & & & & \\ & -1 & 0 & 1 & & & \\ & & -1 & 0 & 1 & & \\ & & & -1 & 0 & & \\ & & & & & \ddots & \\ 1 & & & & & & \end{pmatrix}. \tag{B21}$$

For this model, essentially the same statement as above holds true.

Lemma 5 (Spreading for the critical Ising model)**.** *For the one-dimensional fermionic model in Eq.* (B19) *and* (B21) *corresponding to the critical Ising model with V even, there is a constant $C_{\mathrm{Trans}} > 0$ such that for all $t \in]0, t_{\mathrm{Rec}}]$ with the recurrence time $t_{\mathrm{Rec}} = (2V)^{6/7}/2$ we have*

$$|W_{j,k}(t)| \leq C_{\mathrm{Trans}} t^{-1/3} \,\forall j, k. \tag{B22}$$

Proof. The proof follows again from Lemma 3. For V even we can relate the coupling matrix K in (B21) to M by

$$\mathrm{i}M = Q^\dagger K Q \tag{B23}$$

with $Q = \mathrm{diag}(1, \mathrm{i}, -1, -\mathrm{i}, 1, \ldots, -\mathrm{i})$. We then find for all $0 < t \leq (2V)^{6/7}/2$

$$|L_{j,k}(t)| = |(e^{2tK})_{j,k}| = |(Q^\dagger \mathrm{e}^{2tK} Q)_{j,k}| = |(e^{2\mathrm{i}tM})_{j,k}| \leq \frac{37}{|2t|^{1/3}}. \tag{B24}$$

From the block structure of Ω in Eq. (A7) follows directly

$$|W_{j,k}(t)| \leq 2\frac{37}{|2t|^{1/3}}. \tag{B25}$$

□

Appendix C: Fermionic Gaussian states and clustering of correlations

In this appendix, we consider Gaussian fermionic states. We first demonstrate that they can be thought of as maximum entropy states given a correlation matrix. Moreover, we investigate the Gaussified version of states with exponential clustering of correlations. By using the exponential decay of second moments, we show that also the Gaussified version of the state shows such a correlation decay.

1. Gaussian states as maximum entropy states

In this subsection, we show that fermionic Gaussian states are the maximum entropy states given the second moments of fermionic operators, which we state here for completeness of the argument. That is to say, the state the system locally converges to can be seen as a proper generalised Gibbs ensemble (GGE), with the constants of motion being provided by the second moments.

Lemma 6 (Gaussian states as maximum entropy states)**.** *For a given correlation matrix $\gamma \in \mathbb{C}^{2V\times 2V}$,*

$$\rho_G = \arg\max_{\rho} \left\{ S(\rho) : \gamma(\rho) = \gamma \right\}, \tag{C1}$$

where ρ is a quantum state and $\gamma(\rho)$ its correlation matrix.

Proof. This statement follows immediately from the positivity of the quantum relative entropy. For an arbitrary state ρ and the Gaussian state ρ_G with the same correlation matrix, we have

$$0 \leq S(\rho\|\rho_G) = -S(\rho) - \operatorname{tr}(\rho \log(\rho_G)). \tag{C2}$$

Since ρ_G is a Gaussian state, it can be written as $\rho_G = e^{ih}$ for a suitable Hermitian h that is quadratic in the fermionic operators, which means that

$$\operatorname{tr}(\rho \log(\rho_G)) = \operatorname{tr}(\rho_G \log(\rho_G)) = -S(\rho_G), \tag{C3}$$

from which the assertion follows. □

2. Clustering of correlations of Gaussified states

In this subsection we show that given a state ρ that exhibits exponential clustering of correlation as defined in Def. 1, its Gaussified version ρ_G inherits the exponential clustering of correlations with a changed scaling of the pre-factor with the support of the considered operators. We proof this statement by the aid of Wick's theorem which connects higher to second moments for a general Gaussian state. Precisely, Wick's theorem can be stated as follows.

Lemma 7 (Wick's theorem [33])**.** *A Gaussian state ρ_G fulfills*

$$\operatorname{tr}[\prod_{k=1}^{n} c_{i_k} \rho_G] = \operatorname{Pf}(\gamma^c[i_1, \ldots, i_n]), \tag{C4}$$

where

$$\gamma^c[i_1, \ldots, i_n]_{a,b} = \begin{cases} \operatorname{tr}(c_{i_a} c_{i_b} \rho_G) & \text{for } a < b, \\ -\operatorname{tr}(c_{i_b} c_{i_a} \rho_G) & \text{for } b < a, \\ 0 & \text{else.} \end{cases} \tag{C5}$$

Given a state with exponential clustering of correlations also its Gaussified version will show clustering of correlations in following sense:

Lemma 8 (Weak clustering of correlations for Gaussified states)**.** *Let ρ be a state that exhibits exponential clustering of correlations according to Definition 1 with constants $C_{\mathrm{Clust}}, \xi > 0$, then for all operators A, B with $\|A\| = \|B\| = 1$ its Gaussified version ρ_G satisfies*

$$|\operatorname{tr}(\rho_G\, A\, B) - \operatorname{tr}(A\, \rho_G)\, \operatorname{tr}(B\, \rho_G)| \leq C_{\mathrm{Clust}}\, 4^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|} (|\operatorname{supp}(A)| + |\operatorname{supp}(B)|)^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|} \mathrm{e}^{-\operatorname{d}(A,B)/\xi}. \tag{C6}$$

Proof. Note that we can assume without loss generality $C_{\mathrm{Clust}} \mathrm{e}^{-\operatorname{d}(A,B)/\xi} \leq 1$ as otherwise the trivial bound $|\operatorname{tr}(\rho_G\, A\, B) - \operatorname{tr}(A\, \rho_G)\, \operatorname{tr}(B\, \rho_G)| \leq 2$ concludes proof.

We decompose a general operator supported on $\operatorname{supp}(A)$ and $\operatorname{supp}(B)$ as in Eq. (A10) into the fermionic operator-basis

$$A = \sum_{K \subset \operatorname{supp}(A)} a(K) \prod_{k \in K} c_k \tag{C7}$$

and

$$B = \sum_{J \subset \mathrm{supp}(B)} b(J) \prod_{j \in J} c_j \tag{C8}$$

correspondingly. From $\|A\| = 1 = \|B\|$ follows $|b(J)| \leq 1$ and $|a(K)| \leq 1$ for all corresponding J and K. Using the triangle inequality, we can therefore write

$$|\operatorname{tr}(\rho_G\, A\, B) - \operatorname{tr}(A\,\rho_G)\,\operatorname{tr}(B\,\rho_G)| \leq 2^{|\operatorname{supp}(A)\cup\operatorname{supp}(B)|} \max_{\substack{K\subset\operatorname{supp}(A),\\ J\subset\operatorname{supp}(B)}} \left|\operatorname{tr}(\rho_G \prod_{k\in K} c_k \prod_{j\in J} c_j) - \operatorname{tr}(\prod_{k\in K} c_k\,\rho_G)\,\operatorname{tr}(\prod_{j\in J} c_j\,\rho_G)\right|. \tag{C9}$$

Assuming that the maximum is attained for the sets J' and K', Wick's theorem allows us to write the expectation values in terms of second moments:

$$|\operatorname{tr}(\rho_G\, A\, B) - \operatorname{tr}(A\,\rho_G)\,\operatorname{tr}(B\,\rho_G)| \leq 2^{|\operatorname{supp}(A)\cup\operatorname{supp}(B)|} |\operatorname{Pf}\gamma^c[(k)_{k\in K'},(j)_{j\in J'}] - \operatorname{Pf}\gamma^c[(k)_{k\in K'}]\,\operatorname{Pf}\gamma^c[(j)_{j\in J'}]|. \tag{C10}$$

Note that from the definition of γ^c in (C5) follows that $\gamma^c[(k)_{k\in K'},(j)_{j\in J'}]$ decomposes into blocks according to

$$\gamma^c[(k)_{k\in K'},(j)_{j\in J'}] = \gamma^c[(k)_{k\in K'}] \oplus \gamma^c[(j)_{j\in J'}] + \begin{pmatrix} 0 & E \\ -E^T & 0 \end{pmatrix}. \tag{C11}$$

As E contains only second moments of which one operator is supported on $\operatorname{supp}(A)$ and the other on $\operatorname{supp}(B)$ we obtain from the exponential clustering of correlations of ρ that $|E_{a,b}| \leq C_{\mathrm{Clust}} \mathrm{e}^{-\,\mathrm{d}(A,B)/\xi}$. Expanding therefore the Pfaffians in Eq. (C10) yields that each term either appears in both terms of the difference and cancels out or that it contains at least one element of E as a factor. Counting the number of terms in the expansion of the Pfaffians gives that the sum contains

$$(2[|\operatorname{supp}(A)| + |\operatorname{supp}(B)|] - 1)!! \leq 2^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|}(|\operatorname{supp}(A)| + |\operatorname{supp}(B)|)^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|} \tag{C12}$$

many terms which yields the final bound

$$|\operatorname{tr}(\rho_G\, A\, B) - \operatorname{tr}(A\,\rho_G)\,\operatorname{tr}(B\,\rho_G)| \leq C_{\mathrm{Clust}}\, 4^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|}(|\operatorname{supp}(A)| + |\operatorname{supp}(B)|)^{|\operatorname{supp}(A)|+|\operatorname{supp}(B)|}\mathrm{e}^{-\,\mathrm{d}(A,B)/\xi}. \tag{C13}$$

□

Appendix D: Details of the proof of Theorem 1

In this Appendix we provide all the details of the proof of our main result Theorem 1. We proceed as follows: In Section D 1 we bound the error introduced by truncating to the Lieb-Robinson cone. In Section D 2 we introduce the necessary concepts and notation to then in Section D 3 bound the error made by factorizing expectation values into a products of local contributions from different patches. In Section D 4 we use the properties of delocalizing transport to show a bound on the remaining non-Gaussian contributions to the expectation value. Finally, in Section D 5, we assemble all the parts of the proof and state a more technical version of the main theorem.

1. Truncating to the Lieb-Robinson cone

We decompose a general operator supported in the region S according to Eq. (A10). Without loss of generality we can assume that A is normalized, i.e., $\|A\| \leq 1$, which implies $|a(J)| \leq 1$ and so in the following we concentrate on the individual terms of the form given in Eq. (A12). We will demonstrate that sums over time evolved fermionic operators in Eq. (A12) can be truncated to an enlarged Lieb-Robinson cone up to an error that decays exponentially with time. Rather than summing over all possible index positions $k_j \in [2V]$ it is then sufficient to only sum over positions inside this enlarged Lieb-Robinson cone. We will require the following auxiliary lemma:

Lemma 9 (Norm bound on restricted sums of fermionic operators)**.** *Let $I \subset [2V]$ and $W \in U(2V)$ be unitary. Then for all $j \in [2V]$*

$$\left\| \sum_{k_j \in I} W_{j,k_j} c_{k_j} \right\| \leq 1\,. \tag{D1}$$

Proof. The proof can be carried out with straightforward norm estimates and using the normalisation of the two-point correlator γ. We begin with

$$\left\| \sum_{k_j \in I} W_{j,k_j} c_{k_j} \right\| = \sup_{\substack{|\psi\rangle \\ \||\psi\rangle\|=1}} \langle\psi| \sum_{r_j \in I} \overline{W}_{j,r_j} c_{r_j}^\dagger \sum_{k_j \in I} W_{j,k_j} c_{k_j} |\psi\rangle = \sup_{\substack{|\psi\rangle \\ \||\psi\rangle\|=1}} \sum_{r_j \in I} \sum_{k_j \in I} \overline{W}_{j,r_j} \langle\psi| c_{r_j}^\dagger c_{k_j} |\psi\rangle W_{j,k_j} . \tag{D2}$$

We now rewrite this as a matrix multiplication on the index space

$$\left\| \sum_{k_j \in I} W_{j,k_j} c_{k_j} \right\| \leq \sup_{\gamma} \langle j| \overline{W} P_I \gamma P_I W^T |j\rangle , \tag{D3}$$

where $|j\rangle$ is a vector on the index space, P_I denotes the projector onto the interval I and γ denotes fermionic correlation matrices. A straightforward norm estimate and using that $\|\gamma\| \leq 1$, as every fermionic mode can be occupied by at most one particle, gives

$$\left\| \sum_{k_j \in I} W_{j,k_j} c_{k_j} \right\| \leq \|\overline{W}\| \, \|P_I\| \, \|\gamma\| \, \|P_I\| \, \|W^T\| \leq 1 , \tag{D4}$$

which concludes the proof. □

As introduced in the main text, we then denote by $\mathrm{d}(A, B)$ the shortest distance between the supports $\mathrm{supp}(A), \mathrm{supp}(B)$ of two operators A and B. For $k_1, k_2 \in [2V]$ we then define the distance $\mathrm{d}(k_1, k_2) \coloneqq \mathrm{d}(c_{k_1}, c_{k_2})$. Note that d defines only a pseudometric on $[2V]$ as for $k_1, k_2 \in [2V]$ with $c_{k_1} = f_s$ and $c_{k_2} = f_s^\dagger$ we have $k_1 \neq k_2$ but $\mathrm{d}(k_1, k_2) = 0$.

Given a pseudometric, we define a ball around a set as follows.

Definition 3 (Ball around set)**.** *Given $l > 0$, a set M with pseudometric d and $J \subset M$, we define the l-ball $\mathcal{B}_l(J) \subset M$ around J by*

$$\mathcal{B}_l(J) = \{s \in M : \min_{j \in J} \mathrm{d}(j, s) \leq l\}. \tag{D5}$$

With this, we define an enlarged Lieb-Robinson cone around a set of indices J with radius $(v + 2v_\epsilon)|t|$ for some $v_\epsilon > 0$ and bound the error made by restricting sums of the form given in Eq. (A12) to this widened Lieb-Robinson cone:

Lemma 10 (Error made in restricting to widened Lieb-Robinson cone)**.** *Given a $d_{\mathcal{L}}$-dimensional cubic lattice system with a quadratic Hamiltonian H that satisfies a Lieb-Robinson bound of the form given in Lemma 2 with parameters $C_{\mathrm{LR}}, \mu, v > 0$. Let $v_\epsilon > 0$ and define for any set $J \subset [2V]$ the widened cone $\mathcal{C} \coloneqq \mathcal{B}_{(v+2v_\epsilon)|t|}(J)$, then there exists a constant $\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}})$, such that*

$$\left\| \sum_{(k_j)_{j \in J} \notin \mathcal{C}^{\times |J|}} \prod_{j \in J} W_{j,k_j}(t) \, c_{k_j} \right\| \leq \tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}}) |J|^2 \mathrm{e}^{-\mu v_\epsilon |t|} . \tag{D6}$$

Proof. We begin by splitting the sum according to whether the first index is inside the cone or not. All other indices are free if k_{j_1} is outside the cone, while at least one other index is outside the cone if k_{j_1} lies in it. Using Lemma 9, we obtain

$$\begin{aligned} \left\| \sum_{(k_j)_{j \in J} \notin \mathcal{C}^{\times |J|}} \prod_{j \in J} W_{j,k_j}(t) \, c_{k_j} \right\| &\leq \left\| \sum_{k_{j_1} \notin \mathcal{C}} W_{j_1,k_{j_1}}(t) \, c_{k_{j_1}} \right\| \left\| \prod_{j \in J \setminus \{j_1\}} c_j(t) \right\| \\ &\quad + \left\| \sum_{k_{j_1} \in \mathcal{C}} W_{j_1,k_{j_1}}(t) \, c_{k_{j_1}} \right\| \left\| \sum_{(k_j)_{j \in J \setminus \{j_1\}} \notin \mathcal{C}^{\times |J|-1}} \prod_{j \in J \setminus \{j_1\}} W_{j,k_j}(t) \, c_{k_j} \right\| \\ &\leq \left\| \sum_{k_{j_1} \notin \mathcal{C}} W_{j_1,k_{j_1}}(t) \, c_{k_{j_1}} \right\| + \left\| \sum_{(k_j)_{j \in J \setminus \{j_1\}} \notin \mathcal{C}^{\times |J|-1}} \prod_{j \in J \setminus \{j_1\}} W_{j,k_j}(t) \, c_{k_j} \right\| . \end{aligned} \tag{D7}$$

The first term in the above equation now, due to Lemma 2, satisfies

$$\begin{aligned} \left\| \sum_{k_{j_1} \notin \mathcal{C}} W_{j_1,k_{j_1}}(t) \, c_{k_{j_1}} \right\| &\leq C_{\mathrm{LR}} \sum_{l=(v+2v_\epsilon)\,|t|}^{V} |\mathcal{B}_{l+1}(J) \setminus \mathcal{B}_l(J)| \, \mathrm{e}^{\mu\,(v\,|t|-l)} \\ &\leq 2^{d_{\mathcal{L}}+1} \, d_{\mathcal{L}} \, |J| \, C_{\mathrm{LR}} \, \mathrm{e}^{\mu\, v\, |t|} \sum_{l=(v+2v_\epsilon)\,|t|}^{\infty} l^{d_{\mathcal{L}}-1} \, \mathrm{e}^{-\mu\, l} , \end{aligned} \tag{D8}$$

where we have used that the number $|\mathcal{B}_{l+1}(J)\backslash\mathcal{B}_l(J)|$ of points in the surface of a cone with radius l around J in a cubic lattice is bounded by $4\,d_{\mathcal{L}}\,|J|\,(2l)^{d_{\mathcal{L}}-1}$. Shifting the limits of the sum then yields

$$\left\| \sum_{k_{j_1}\notin\mathcal{C}} W_{j_1,k_{j_1}}(t)\,c_{k_{j_1}} \right\| \leq \mathrm{e}^{-\mu\,v_\epsilon\,|t|}\,|J|\,2^{d_{\mathcal{L}}+1}\,d_{\mathcal{L}}\,C_{\mathrm{LR}}\,\sum_{l=0}^{\infty}(l+(v+2v_\epsilon)|t|)^{d_{\mathcal{L}}-1}\,\mathrm{e}^{-\mu\,(l+v_\epsilon|t|)}\,. \tag{D9}$$

We now define the time independent constant

$$\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}}) \coloneqq \sup_{t\in\mathbb{R}^+} 2^{d_{\mathcal{L}}+1}\,d_{\mathcal{L}}\,C_{\mathrm{LR}}\,\sum_{l=0}^{\infty}(l+(v+2v_\epsilon)|t|)^{d_{\mathcal{L}}-1}\,\mathrm{e}^{-\mu\,(l+v_\epsilon|t|)} \tag{D10}$$

which can be written in terms so of the Hurwitz-Lerch-Phi function Φ (also known as Lerch transcendent)

$$\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}}) \coloneqq \sup_{t\in\mathbb{R}^+} 2^{d_{\mathcal{L}}+1}\,d_{\mathcal{L}}\,C_{\mathrm{LR}}\,\Phi(\mathrm{e}^{-\mu},1-d_{\mathcal{L}},(v+2v_\epsilon)|t|)\mathrm{e}^{-\mu v_\epsilon|t|}\,. \tag{D11}$$

Inserting the estimate into Eq. (D7) and iteratively using the resulting inequality $|J|$-times gives the result as stated.

As argued above, the constant is directly related Hurwitz-Lerch-Phi function and can easily be explicitly evaluated for physical dimensions $d_{\mathcal{L}}=1,2,3$. In one dimension, the constant takes the form

$$\tilde{C}_{\mathrm{LR}}(1) = 4\,C_{\mathrm{LR}}\,\frac{1}{1-\mathrm{e}^{-\mu}}\,. \tag{D12}$$

□

2. Partitions: Tracking indices on the lattice

Using the result of Lemma 10 we can restrict the time evolution in Eq. (A12) to the Lieb-Robinson cone at the cost of an exponentially suppressed error term. In this section we therefore look

$$A_J^{LR}(t) \coloneqq \sum_{(k_j)_{j\in J}\in\mathcal{C}^{\times|J|}} \left(\prod_{j\in J} W_{j,k_j}(t)\,c_{k_j}\right), \tag{D13}$$

the restriction of a term of the form Eq. (A12) to the widened Lieb-Robinson cone $\mathcal{C}=\mathcal{B}_{(v+2v_\epsilon)|t|}(J)$. By grouping summands according to how close the respective indices k_j are on the lattice we will rewrite $A_J^{LR}(t)$ as a sum over partitions of the sub-index set J. This will later allow us to factorize certain expectation values using the exponential decay of correlations in the initial state.

We start by introducing some notation. Given a finite non-empty set J, a partition P of J is a set of non-empty subsets (patches) of J, whose union is J, i.e., $\overline{P}\coloneqq\bigcup_{p\in P}p=J$. We denote by $\pi_m(P)\coloneqq\{p\in P:|p|=m\}$ the subset of all patches in a partition with a given size m and by $\pi_{>m}(P)=\{p\in P:|p|>m\}$ that of all patches with size larger than m. We refer to patches of size two as pairs and patches of size at least four as clusters. Partitions will be called even, if all patches in it have an even size. We further denote by $\mathcal{P}(J)=\{P: P \text{ partition of } J\}$ the set of all partitions of J, and by $\mathcal{P}_m(J)$ and $\mathcal{P}_{>m}(J)$ the sets of all partitions into patches of size exactly equal to, or larger than m, respectively. Given two partitions P,Q we say that Q is a coarsening of P and write $Q<P$ if $\forall q\in Q\,\exists p\in P: q\subset p$ and $Q\neq P$.

Next we introduce the notion of a Δ-partition of the sub-index set J. Such a partition To each configuration of indices $(k_j)_{j\in J}$, we assign a unique partition P of the subindices j such that all indices $(k_j)_{j\in p}$ with subindices that lie within one set p of the partition are connected by a path of steps with maximal length Δ and all indices k_j corresponding to subindices in two different sets of the partition lie more than a distance Δ apart.

Definition 4 (Δ-partition)**.** *Given a distance $\Delta>0$, a finite set $J\subset\mathbb{N}$, a finite set M equipped with a pseudometric d, and a sequence of elements $(k_j)_{j\in J}\in M^{\times|J|}$. We define the Δ-partition $P_\Delta(J,(k_j)_{j\in J})$ to be the unique partition of J which fulfills (1) Each set in the partition is path connected by hops of length at most Δ in the sense that*

$$\forall p\in P_\Delta(J,(k_j)_{j\in J}):\forall x,y\in p\exists z_1,\ldots,z_N\in p: x=z_1, y=z_N\wedge\forall i\in[N-1]:\mathrm{d}(k_{z_i},k_{z_{i+1}})\leq\Delta. \tag{D14}$$

(2) The different patches in the partition are separated by a distance larger than Δ in the sense that

$$\forall p\neq q\in P_\Delta(J,(k_j)_{j\in J}):\forall x\in p,y\in q:\mathrm{d}(k_x,k_y)>\Delta. \tag{D15}$$

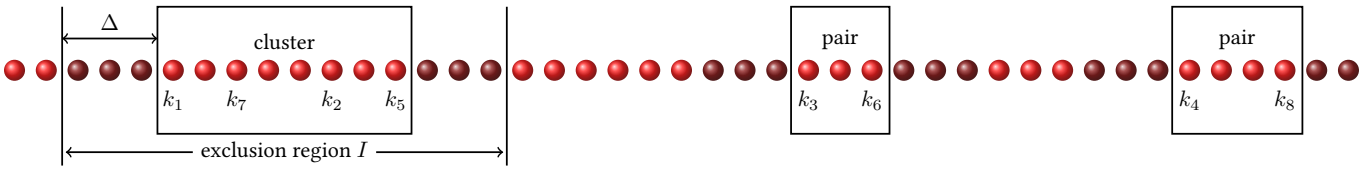

Figure 4. Illustration of a Δ-partition of a given index configuration with $\Delta = 3$ and $J = \{1, \ldots, 8\}$. To the above configuration of the indices $k_1, \ldots k_8$ we associate the Δ-partition $\{\{1, 2, 5, 7\}, \{3, 6\}, \{4, 8\}\}$ such that a patch of size four and two patches of size two (pairs) are formed. Around each patch exists an buffer region (shaded nodes) which separates different patches. For a given Δ-partition P we obtain all possible index configurations in $\mathcal{K}_P^M$ by placing the patches iteratively. A placed patch will hereby create an exclusion region consisting of the patch itself and the buffer region around it in which no further patch can be placed.

In addition, we define a compact notation for index configurations distributed over the lattice such that their Δ-partition agrees with a given partition P.

Definition 5 (Index sets respecting Δ-partitions)**.** *Given a distance $\Delta > 0$, the set $[2V]$ of all sites on the lattice equipped with a pseudometric d, a sub-index set $M \subset [2V]$, and a partition $P \in \mathcal{P}(J)$. We denote the set of sequences contained in M whose Δ partitions is equal to P by*

$$\mathcal{K}_P^M \coloneqq \{(k_j)_{j\in\overline{P}} \in M^{\times|\overline{P}|} : P_\Delta(\overline{P}, (k_j)_{j\in\overline{P}}) = P\}. \tag{D16}$$

The notation introduced above allows us to rewrite the sum over the indices $(k_j)_{j\in J}$ in Eq. (D13) inside the cone by sorting them according to their associated Δ-partition.

$$\begin{aligned} A_J^{LR}(t) = \sum_{(k_j)_{j\in J}\in\mathcal{C}^{\times|J|}} \left(\prod_{j\in J} W_{j,k_j}(t)\, c_{k_j}\right) &= \sum_{P\in\mathcal{P}(J)} \sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}} \prod_{j\in J} W_{j,k_j}(t)\, c_{k_j} \\ &= \sum_{P\in\mathcal{P}(J)} \mathrm{sign}(P) \sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}} \prod_{p\in P} \hat{P}_p^{(k_j)_{j\in p}}, \end{aligned} \tag{D17}$$

where for each $p \in P$ we have introduced a patch operator $\hat{P}_p^{(k_j)_{j\in p}}$, defined by

$$\hat{P}_p^{(k_j)_{j\in p}} = \prod_{j\in p} W_{j,k_j}(t) c_{k_j} \ . \tag{D18}$$

The $\mathrm{sign}(P)$ denotes the sign picked up from reordering the fermionic basis operators into the corresponding patches, where keeping the relative order of the operators inside each patch fix.

3. Factorizing expectation values

In the last section, we have developed the formalism to group indices on the lattice according to their distribution on the lattice and introduced the patch operators $\hat{P}_p^{(k_j)_{j\in p}}$. We now use the exponential clustering of correlations in the initial state to show that expectation values of products of such patch operators can be factorized into a product of expectation values of the individual patch operators up to a small error.

Lemma 11 (Factorizing expectation values in states with exponential clustering of correlations)**.** *Let ρ be a state that exhibits exponential clustering of correlations as defined in Definition 1 with system size independent parameters $C_{\mathrm{Clust}}, \xi > 0$. Let $J \subset [2V]$ and $P \in \mathcal{P}(J)$ be a partition of J. Then for any distance $\Delta > 0$ it holds that*

$$\sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}} \left| \left\langle \prod_{p\in P} \hat{P}_p^{(k_j)_{j\in p}} \right\rangle_\rho - \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\rho \right| \leq |J|^3\, C_{\mathrm{Clust}}\, |\mathcal{C}|^{|J|} \mathrm{e}^{-\Delta/\xi} \ . \tag{D19}$$

Proof. We begin by factorizing out the contribution from the first patch $p_1 \in P$. Using Lemma 9 and the exponential clustering of the initial state, we find for a given $(k_j)_{j\in J} \in \mathcal{K}_P^{\mathcal{C}}$

$$\left| \left\langle \prod_{p\in P} \hat{P}_p^{(k_j)_{j\in p}} \right\rangle_\rho - \langle \hat{P}_{p_1}^{(k_j)_{j\in p_1}} \rangle_\rho \left\langle \prod_{p\in P\setminus\{p_1\}} \hat{P}_p^{(k_j)_{j\in p}} \right\rangle_\rho \right| \leq |p_1|\, |\overline{P} \setminus p_1|\, C_{\mathrm{Clust}}\, \mathrm{e}^{-\Delta/\xi} \ . \tag{D20}$$

Using the trivial bound $|p_1|\,|\overline{P}\setminus p_1| \leq |J|^2$, iterating the step above $|P| \leq |J|$ times and using that $|\mathcal{K}_P^{\mathcal{C}}| \leq |\mathcal{C}|^{|J|}$ yields the result as stated. □

From Lemma 8 follows that the same theorem applies to the Gaussified version of a state exhibiting clustering of correlations if we allow C_{Clust} to scale with the size of the support according to $C_{\mathrm{Clust}} \to C_{\mathrm{Clust}} 4^{|J|}|J|^{|J|}$.

4. Suppression of non-Gaussian contributions

In the last two sections we have bounded the error made in approximating expectation values of terms of the from given in Eq. (A12) by certain sums of products of expectation values of patch operators. This allows us to bound the the difference between the left and right hand side of

$$\mathrm{tr}[A(t)\,\rho] - \mathrm{tr}(A(t)\,\rho_G] \approx \sum_{J\subset\tilde{S}} a(J) \sum_{P\in\mathcal{P}(J)} \mathrm{sign}(P) \sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}} \left(\prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\rho - \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_{\rho_G} \right). \tag{D21}$$

It is obvious that to the right hand side only partitions P in which all patches p are of even size can contribute, as ρ and ρ_G have an even particle number parity. Moreover, as the second moments of ρ and ρ_G are equal by definition, the difference between the products also vanishes whenever all patches have size exactly 2. Hence, every contributing term contains at least one patch of size at least 4. In the remainder of this section we now bound the contribution of such partitions to the right hand side of the above equation. The combinatorial nature of the problem makes this a tedious endeavor. The final result is summarized in the following lemma, which is the last result that we need before we can assemble all the parts of the proof in Section D 5.

Lemma 12 (Bounding contributions from partitions that contain large patches)**.** *Let ρ be a state exhibiting exponential clustering of correlations as defined in Definition 1 with system size independent parameters $C_{\mathrm{Clust}}, \xi > 0$ and ρ_G its Gaussified version. Let $\Delta \geq 1$, $J \subset [2V]$, and $P \in \mathcal{P}(J)$ be an even partition that contains a patch of size at least four (cluster), and $m \coloneqq |\pi_2(P)|$ many patches of size two (pairs). Given that the time evolution of the system is governed by a Hamiltonian showing delocalizing transport as defined in Definition 2 with parameter C_{Trans}, for all $t \in (0, \min(t_{\mathrm{Rec}}, V)]$ and $\sigma \in \{\rho, \rho_G\}$ it holds that*

$$\begin{aligned} R_P(J,m,t) &\coloneqq \left| \sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right| \\ &\leq (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} 2^{d_{\mathcal{L}}+1} |J|^{d_{\mathcal{L}}} \Delta^{d_{\mathcal{L}}})^{|J|-2m} \\ &\quad \times \left[(1 + C_{\mathrm{Clust}}|\mathcal{C}|^2 \mathrm{e}^{-\Delta/\xi}) + 2^{(d_{\mathcal{L}}+1)|J|} |J|^{(d_{\mathcal{L}}+1)|J|+1} \sum_{r=0}^{|J|/2} (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} \Delta^{d_{\mathcal{L}}})^{2r+4} \right]^m . \end{aligned} \tag{D22}$$

Proof. In order to prove the above bound we separate the partition into a part containing only pairs and one containing the rest

$$P = \pi_2(P) \cup \pi_{>2}(P)\,. \tag{D23}$$

For every given fixed position of the indices corresponding to clusters in $\pi_{>2}(P)$, the indices corresponding to the pairs are restricted to the set $K \coloneqq \mathcal{C} \setminus \mathcal{B}_\Delta(\{k_j\}_{j\in\overline{\pi_{>2}(P)}})$, as all patches are separated by a distance larger than Δ. Thus, we can write

$$R_P(J,m,t) = \left| \sum_{(k_j)_{j\in\overline{\pi_{>2}(P)}}\in\mathcal{K}^{\mathcal{C}}_{\pi_{>2}(P)}} \left[\left(\prod_{p\in\pi_{>2}(P)} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right) \sum_{\substack{(k_j)_{j\in\overline{\pi_2(P)}}\in\mathcal{K}^K_{\pi_2(P)},\\ K\coloneqq\mathcal{C}\setminus\mathcal{B}_\Delta(\{k_j\}_{j\in\overline{\pi_{>2}(P)}})}} \prod_{p\in\pi_2(P)} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right] \right| . \tag{D24}$$

The homogeneous suppression due to delocalizing transport for $t \in (0, \min(t_{\mathrm{Rec}}, V)]$ and that $|\overline{\pi_{>2}(P)}| = |J| - 2m$ implies

$$\left\| \prod_{p\in\pi_{>2}(P)} \hat{P}_p^{(k_j)_{j\in p}} \right\| \leq (C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3})^{|J|-2m}\,. \tag{D25}$$

Using this and the triangle inequality for the first sum in Eq. (D24) we arrive at

$$R_P(J,m,t) \leq (C_{\text{Trans}} t^{-d_{\mathcal{L}}/3})^{|J|-2m} \sum_{(k_j)_{j\in\overline{\pi_{>2}(P)}}\in\mathcal{K}^{\mathcal{C}}_{\pi_{>2}(P)}} \left| \sum_{\substack{(k_j)_{j\in\overline{\pi_2(P)}}\in\mathcal{K}^{K}_{\pi_2(P)},\\ K=\mathcal{C}\setminus\mathcal{B}_\Delta(\{k_j\}_{j\in\overline{\pi_{>2}(P)}})}} \prod_{p\in\pi_2(P)} \langle \hat{P}_p^{(k_j)_{j\in p}}\rangle_\sigma \right| . \tag{D26}$$

At this point, the inner sum still depends on the position of the clusters, as they create an exclusion region for the pairs (see Fig. 4). Taking the supremum over such exclusion regions decouples the sum. Then bounding the possible number of positions of the $|J|-2m$ indices in the $|\pi_{>2}(P)|$ clusters

$$|(k_j)_{j\in\overline{\pi_{>2}(P)}} \in \mathcal{K}^{\mathcal{C}}_{\pi_{>2}(P)}| \leq |\mathcal{C}|^{|\pi_{>2}P|}(2^{d_{\mathcal{L}}+1}|J|^{d_{\mathcal{L}}}\Delta^{d_{\mathcal{L}}})^{|J|-2m} , \tag{D27}$$

gives

$$\begin{aligned} R_P(J,m,t) &\leq |\mathcal{C}|^{|\pi_{>2}(P)|}(C_{\text{Trans}} t^{-d_{\mathcal{L}}/3} 2^{d_{\mathcal{L}}+1}|J|^{d_{\mathcal{L}}}\Delta^{d_{\mathcal{L}}})^{|J|-2m} \max_{I\subset\mathcal{C}} \left| \sum_{(k_j)_{j\in\overline{\pi_2(P)}}\in\mathcal{K}^{\mathcal{C}\setminus I}_{\pi_2(P)}} \prod_{p\in\pi_2(P)} \langle \hat{P}_p^{(k_j)_{j\in p}}\rangle_\sigma \right| \\ &\leq (|\mathcal{C}|^{1/4} C_{\text{Trans}} t^{-d_{\mathcal{L}}/3} 2^{d_{\mathcal{L}}+1}|J|^{d_{\mathcal{L}}}\Delta^{d_{\mathcal{L}}})^{|J|-2m} \max_{M\subset J:|M|=2m} \max_{\substack{I\subset\mathcal{C},\\ P\in\mathcal{P}_2(M)}} \left| \sum_{(k_j)_{j\in M}\in\mathcal{K}^{\mathcal{C}\setminus I}_{P}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}}\rangle_\sigma \right| . \end{aligned} \tag{D28}$$

We now define

$$f(m,t) \coloneqq \max_{M\subset J:|M|=2m} \max_{\substack{I\subset\mathcal{C},\\ P\in\mathcal{P}_2(M)}} \left| \sum_{(k_j)_{j\in M}\in\mathcal{K}^{\mathcal{C}\setminus I}_{P}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}}\rangle_\sigma \right| \tag{D29}$$

and apply an recursive argument to achieve a bound of the form

$$f(m,t) \leq C(m,t) . \tag{D30}$$

a. Start of recursion: For the case of one pair ($m = 1$), by using Lemma 9, we can bound the appearing maximum as follows

$$\begin{aligned} f(1,t) &= \max_{M\subset J:|M|=2} \max_{I\subset\mathcal{C}} \left| \sum_{(k_j)_{j\in M}\in\mathcal{K}^{\mathcal{C}\setminus I}_{\{M\}}} \langle \hat{P}_M^{(k_j)_{j\in M}}\rangle_\sigma \right| = \max_{M\subset J:|M|=2} \max_{I\subset\mathcal{C}} \left| \sum_{\substack{k_{l_1},k_{l_2}\in\mathcal{C}\setminus I,\\ \{l_1,l_2\}=M:\\ \mathrm{d}(k_{l_1},k_{l_2})\leq\Delta}} \langle \hat{P}_M^{(k_j)_{j\in M}}\rangle_\sigma \right| \\ &\leq \max_{M\subset J:|M|=2} \max_{I\subset\mathcal{C}} \left(\left| \sum_{\substack{k_{l_1},k_{l_2}\in\mathcal{C}\setminus I,\\ \{l_1,l_2\}=M}} \langle \hat{P}_M^{(k_j)_{j\in M}}\rangle_\sigma \right| + \left| \sum_{\substack{k_{l_1},k_{l_2}\in\mathcal{C}\setminus I,\\ \{l_1,l_2\}=M:\\ \mathrm{d}(k_{l_1},k_{l_2})>\Delta}} \langle \hat{P}_M^{(k_j)_{j\in M}}\rangle_\sigma \right| \right) \\ &\leq \max_{M\subset J:|M|=2} \max_{I\subset\mathcal{C}} (1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi}) \\ &= 1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi} . \end{aligned} \tag{D31}$$

With this, we can move to setting up the recursion.

b. Setting up the recursion: To obtain a recursion formula for an upper bound $C(m,t)$ on $f(m,t)$ we now relax the condition that different pairs may not occupy close-by lattice regions. If we drop this constraint, the only remaining constraint is that paired indices k_l and $k_{l'}$ lie close to each other, i.e. $\mathrm{d}(k_l,k_{l'}) \leq \Delta$. For any set $M \subset J$ with $|M| = 2m$, set $I \subset \mathcal{C}$,

$P \in \mathcal{P}_2(M)$ Eq. (D31) yields directly

$$\left| \prod_{p\in P} \sum_{(k_j)_{j\in p}\in\mathcal{K}_{\{p\}}^{\mathcal{C}\setminus I}} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right| \leq (1 + C_{\text{Clust}}|\mathcal{C}|^2 \mathrm{e}^{-\Delta/\xi})^m. \tag{D32}$$

This leaves us with controlling the difference between the constrained and unconstrained pairs and yields

$$f(m,t) \leq \left(1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi}\right)^m + \max_{M\subset J:|M|=2m} \max_{\substack{I\subset\mathcal{C},\\ P\in\mathcal{P}_2(M)}} g(M,P,I), \tag{D33}$$

$$g(M,P,I) \coloneqq \left| \sum_{(k_j)_{j\in M}\in\mathcal{K}_P^{\mathcal{C}\setminus I}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma - \prod_{p\in P} \sum_{(k_j)_{j\in p}\in\mathcal{K}_{\{p\}}^{\mathcal{C}\setminus I}} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right|. \tag{D34}$$

In order to get the error under control, we need the notion of a coarsening of a partition defined above [34]. The key insight here is that the above difference between constrained and unconstrained pairs precisely amounts to all possible coarsening of the partition P

$$g(M,P,I) = \left| \sum_{\substack{Q\in\mathcal{P}(M):\\ Q>P}} \sum_{\substack{(k_j)_{j\in M}\in\mathcal{K}_Q^{\mathcal{C}\setminus I}:\\ \forall\{l,l'\}\in P:\mathrm{d}(k_l,k_{l'})\leq\Delta}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right| \leq \sum_{w=0}^{m-2} \sum_{\substack{Q\in\mathcal{P}(M):\\ Q>P,|\pi_2(Q)|=w}} \left| \sum_{\substack{(k_j)_{j\in M}\in\mathcal{K}_Q^{\mathcal{C}\setminus I}:\\ \forall\{l,l'\}\in P:\mathrm{d}(k_l,k_{l'})\leq\Delta}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right|, \tag{D35}$$

where we have applied the triangle inequality and sorted the coarsenings by the numbers of pairs w they have. This is almost the original expression $R_P(J,m,t)$ which this lemma is trying to bound, with the exception that there is still a signature left of the fact that the clusters were created by joining pairs, such that consecutive indices have to be at most distance Δ apart in the cluster. Following the same steps as above, one can show that

$$\begin{aligned}
g(M,P,I) &\leq \sum_{w=0}^{m-2} \sum_{\substack{Q\in\mathcal{P}(M):\\ Q>P,|\pi_2(Q)|=w}} \left| \sum_{\substack{(k_j)_{j\in M}\in\mathcal{K}_Q^{\mathcal{C}\setminus I}:\\ \forall\{l,l'\}\in P:\mathrm{d}(k_l,k_{l'})\leq\Delta}} \prod_{p\in P} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right| \\
&\leq \sum_{w=0}^{m-2} \sum_{\substack{Q\in\mathcal{P}(M):\\ Q>P,|\pi_2(Q)|=w}} (C_{\text{Trans}} t^{-d_\mathcal{L}/3})^{2m-2w} \left[\sum_{\substack{(k_j)_{j\in\overline{\pi_{>2}(P)}}\in\mathcal{K}^{\mathcal{C}}_{\pi_{>2}(P)}\\ \forall\{l,l'\}\in P:\mathrm{d}(k_l,k_{l'})\leq\Delta}} \left| \sum_{\substack{(k_j)_{j\in\overline{\pi_2(P)}}\in\mathcal{K}^{K}_{\pi_2(P)},\\ K=\mathcal{C}\setminus\mathcal{B}_\Delta(\{k_j\}_{j\in\overline{\pi_{>2}(P)}})}} \prod_{p\in\pi_2(P)} \langle \hat{P}_p^{(k_j)_{j\in p}} \rangle_\sigma \right| \right] \\
&\leq \sum_{w=0}^{m-2} (2m)^{2m} (|\mathcal{C}|^{1/4} C_{\text{Trans}} t^{-d_\mathcal{L}/3} 2^{d_\mathcal{L}+1} |J|^{d_\mathcal{L}} \Delta^{d_\mathcal{L}})^{2m-2w} f(w,t) \,,
\end{aligned} \tag{D36}$$

where we used that for a finite set M, we can upper bound the number of partitions of this set by $|\mathcal{P}(M)| \leq |M|^{|M|}$. Thus we obtain for the function $f(m,t)$ the following upper bound

$$\begin{aligned}
f(m,t) &\leq \left(1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi}\right)^m + \sum_{w=0}^{m-2} |J|^{|J|} (|\mathcal{C}|^{1/4} C_{\text{Trans}} t^{-d_\mathcal{L}/3} 2^{d_\mathcal{L}+1} |J|^{d_\mathcal{L}} \Delta^{d_\mathcal{L}})^{2m-2w} f(w,t) \\
&\leq \left(1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi}\right)^m + 2^{(d_\mathcal{L}+1)|J|} |J|^{(d_\mathcal{L}+1)|J|} \sum_{w=0}^{m-2} (|\mathcal{C}|^{1/4} C_{\text{Trans}} t^{-d_\mathcal{L}/3} \Delta^{d_\mathcal{L}})^{2m-2w} f(w,t) \\
&\leq \left(1 + |\mathcal{C}|^2 C_{\text{Clust}} \mathrm{e}^{-\Delta/\xi}\right)^m + 2^{(d_\mathcal{L}+1)|J|} |J|^{(d_\mathcal{L}+1)|J|} \sum_{r=0}^{|J|/2} (|\mathcal{C}|^{1/4} C_{\text{Trans}} t^{-d_\mathcal{L}/3} \Delta^{d_\mathcal{L}})^{2r+4} \sum_{w=0}^{m-2} f(w,t) \\
&\leq \alpha^m + \delta \sum_{w=0}^{m-1} C(w,t) =: C(m,t) \,,
\end{aligned} \tag{D37}$$

where we overestimated by introducing the sum over r and adding the $m-1$ term to the second sum and have introduced the abbreviations

$$\alpha = (1 + C_{\mathrm{Clust}}|\mathcal{C}|^2 \mathrm{e}^{-\Delta/\xi}), \tag{D38}$$

$$\delta = 2^{(d_{\mathcal{L}}+1)|J|}|J|^{(d_{\mathcal{L}}+1)|J|} \sum_{r=0}^{|J|/2} (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} \Delta^{d_{\mathcal{L}}})^{2r+4} \ . \tag{D39}$$

By now setting

$$C(0,t) \coloneqq f(0,t) = 1, \tag{D40}$$

$$C(1,t) \coloneqq \alpha + \delta \geq \alpha \geq f(1,t), \tag{D41}$$

$$\forall m \geq 1: \quad C(m,t) = \alpha^m + \delta \sum_{w=0}^{m-1} C(w,t), \tag{D42}$$

we now have a recursively defined upper bound $C(m,t)$ on $f(m,t)$.

c. Solving the recursion: To resolve the recursion, we first show that

$$\alpha C(m,t) \leq C(m+1,t) \ , \tag{D43}$$

by relying on an induction. To begin with, we have

$$\alpha\, C(0,t) = \alpha \leq \alpha + \delta = C(1,t) \ . \tag{D44}$$

For the induction step, we use

$$\begin{aligned} \alpha\, C(m,t) &= \alpha^{m+1} + \delta \sum_{w=0}^{m-1} \alpha\, C(w,t) \\ &\leq \alpha^{m+1} + \delta \sum_{w=0}^{m-1} C(w+1,t) \\ &\leq \alpha^{m+1} + \delta \sum_{w=0}^{m} C(w,t) = C(m+1,t) \ , \end{aligned} \tag{D45}$$

where we used the induction when moving from the first to the second line by relying on the fact that the sum only goes until $w = m-1$ and we added $\delta C(0,t) > 0$ in the last line. From this, we immediately know that $C(m,t)$ is monotonically increasing as a function of m, since $\alpha \geq 1$. This implies

$$C(m,t) \leq \alpha\, C(m,t) \leq C(m+1,t) \ . \tag{D46}$$

This allows us to resolve the recursion by iteratively using this estimate as follows

$$\begin{aligned} C(m,t) &= \alpha^m + \delta \sum_{w=0}^{m-1} C(w,t) \\ &\leq \alpha^m + (m-1)\delta C(m-1,t) \\ &\leq \sum_{j=0}^{m} \alpha^{m-j} \delta^j \frac{m!}{(m-j)!} \\ &\leq \sum_{j=0}^{m} \alpha^{m-j} \delta^j \frac{m!}{(m-j)!} \frac{|J|^j}{j!} \\ &= (\alpha + |J|\delta)^m. \end{aligned} \tag{D47}$$

For $f(m,t)$, we hence obtain

$$\begin{aligned} f(m,t) &\leq (\alpha + |J|\delta)^m \\ &= \left[(1 + C_{\mathrm{Clust}}|\mathcal{C}|^2 \mathrm{e}^{-\Delta/\xi}) + |J| 2^{(d_{\mathcal{L}}+1)|J|} |J|^{(d_{\mathcal{L}}+1)|J|} \sum_{r=0}^{|J|/2} (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} \Delta^{d_{\mathcal{L}}})^{2r+4} \right]^m \end{aligned} \tag{D48}$$

and for the original quantity considered in this lemma, this yields

$$R_P(J,m,t) \leq (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} 2^{d_{\mathcal{L}}+1} |J|^{d_{\mathcal{L}}} \Delta^{d_{\mathcal{L}}})^{|J|-2m} \tag{D49}$$
$$\times \left[(1 + C_{\mathrm{Clust}} |\mathcal{C}|^2 \mathrm{e}^{-\Delta/\xi}) + |J| 2^{(d_{\mathcal{L}}+1)|J|} |J|^{(d_{\mathcal{L}}+1)|J|} \sum_{r=0}^{|J|/2} (|\mathcal{C}|^{1/4} C_{\mathrm{Trans}} t^{-d_{\mathcal{L}}/3} \Delta^{d_{\mathcal{L}}})^{2r+4} \right]^m ,$$

which concludes the proof. □

5. Overview of the proof of Theorem 1

Collecting all the results of the preceding sub-sections, we can now proof the following result, which directly implies theorem 1 in the main text.

Theorem (Gaussification in finite time)**.** *Let $C_{\mathrm{Clust}}, \xi, C_{\mathrm{Trans}} > 0$. Consider a family of systems on cubic lattices of increasing volume V and let S be some fixed finite region of sites. Let the corresponding initial states exhibit exponential clustering of correlations with constant C_{Clust} and correlation length ξ. Let the Hamiltonians of these systems be quadratic finite range and let them exhibit delocalizing transport with constant C_{Trans} and a recurrence time t_{Rec} increasing as some function of the volume V. Then for any $\epsilon > 0$ there exists a relaxation time $t_{\mathrm{Relax}} > 0$ independent of the system size such that for all $t \in [t_{\mathrm{Relax}}, t_{\mathrm{Rec}}]$ it holds that $\|\rho^S(t) - \rho^S_G(t)\|_1 \leq \epsilon$.*

Proof. To begin with we rewrite the one-norm as

$$\|\rho^S(t) - \rho^S_G(t)\|_1 = \sup_{A \in \mathcal{A}_S} \mathrm{tr}\left(A(t)(\rho - \rho_G)\right) , \tag{D50}$$

and expand the operator A in the basis of fermionic operators

$$A(t) = \sum_{b_1,\ldots,b_{2|S|}=0}^{1} a_{b_1,\cdots,b_{2|S|}} \, c_{s_1}(t)^{b_1} \ldots c_{s_{2|S|}}(t)^{b_{2|S|}} . \tag{D51}$$

Normalization of the operator $\|A\| = 1$ implies that all of the $2^{2|S|}$ coefficients satisfy $|a_{b_1,\cdots,b_{2m}}| \leq 1$, thus

$$\begin{aligned} \|\rho^S(t) - \rho^S_G(t)\|_1 &\leq 2^{2|S|} \max_{J \subset \tilde{S}} \left| \mathrm{tr}\Big(\prod_{j \in J} c_j(t) \, (\rho - \rho_G) \Big) \right| \\ &\leq 2^{2|S|} \max_{J \subset \tilde{S}} \Bigg| \sum_{(k_j)_{j\in J} \in [2V]^{\times |J|}} \mathrm{tr}\Big(\prod_{j \in J} W_{j,k_j}(t) \, c_{k_j} \, (\rho - \rho_G) \Big) \Bigg| . \end{aligned} \tag{D52}$$

Using the Lieb-Robinson bound stated in Lemma 10, we can restrict the sum in the right hand side of the previous inequality to the Lieb-Robinson cone $\mathcal{C}$. This leads to an error term that is exponentially suppressed in time t and we obtain

$$\|\rho^S(t) - \rho^S_G(t)\|_1 \leq 2^{2|S|} \max_{J \subset \tilde{S}} \Bigg(\Bigg| \sum_{(k_j)_{j\in J} \in [\mathcal{C}]^{\times |J|}} \mathrm{tr}\Big(\prod_{j \in J} W_{j,k_j}(t) \, c_{k_j} \, (\rho - \rho_G) \Big) \Bigg| + 2\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}}) |J|^2 \mathrm{e}^{-\mu v_\epsilon |t|} \Bigg) . \tag{D53}$$

We now reorder the terms in the sum according to how the indices k_j are distributed on the lattice. To that end, in Section D 2, we have introduced the concept of a Δ-partition. We turn the sum into a sum over all possible partitions $\mathcal{P}(J)$ and then, for each partition $P \in \mathcal{P}(J)$, sum over all possible ways $\mathcal{K}^{\mathcal{C}}_P$ to distributed the indices over the lattice whose Δ-partition coincides with that given partition P. Partitions consist of patches and we collect the factors $W_{j,k_j}(t) \, c_{k_j}$ from the product over $j \in J$ into patch operators $\hat{P}^{(k_j)_{j\in p}}_p$ for each patch p, as defined in Eq. (D18). Together with the triangle inequality this yields

$$\|\rho^S(t) - \rho^S_G(t)\|_1 \leq 2^{2|S|} \max_{J \subset \tilde{S}} \Bigg(\sum_{P \in \mathcal{P}(J)} \Bigg| \sum_{(k_j)_{j\in J} \in \mathcal{K}^{\mathcal{C}}_P} \Big(\langle \prod_{p \in P} \hat{P}^{(k_j)_{j\in p}}_p \rangle_\rho - \langle \prod_{p \in P} \hat{P}^{(k_j)_{j\in p}}_p \rangle_{\rho_G} \Big) \Bigg| + 2\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}}) |J|^2 \mathrm{e}^{-\mu v_\epsilon |t|} \Bigg) . \tag{D54}$$

Lemma 11 allows us to factor the expectation values with respect to ρ and using Lemma 8 its Gaussified version ρ_G into products of expectation values of the individual patch operators. This leads to an additional error term that grows polynomially with the size of the cone, but is exponentially suppressed in the minimal patch distance Δ, so that we get

$$\begin{aligned}\|\rho^S(t)-\rho_G^S(t)\|_1 \leq 2^{2|S|}\max_{J\subset\tilde{S}}\Bigg(\sum_{P\in\mathcal{P}(J)}\Bigg|\sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}}\Bigg(\prod_{p\in P}\langle\hat{P}_p^{(k_j)_{j\in p}}\rangle_\rho-\prod_{p\in P}\langle\hat{P}_p^{(k_j)_{j\in p}}\rangle_{\rho_G}\Bigg)\Bigg| \\ +(1+2^{2|J|}|J|^{|J|})|J|^{|J|+3}C_{\mathrm{Clust}}|\mathcal{C}|^{|J|}\mathrm{e}^{-\Delta/\xi}+2\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}})|J|^2\mathrm{e}^{-\mu v_\epsilon|t|}\Bigg)\,.\end{aligned} \tag{D55}$$

It is now apparent that partitions that contain at least one patch of odd size do not contribute to the sum as then the corresponding patch operator does not fulfill the parity super-selection rule. Likewise, partitions that contain only patches of size two do not contribute, as the expectation values of their patch operators are the same in ρ and ρ_G. It remains to bound the contribution from the remaining partitions. For these we cannot use cancellations between the parts coming from ρ and those coming from ρ_G, but instead bound them in absolute value. All these partitions contain at least one cluster of size at least four, which allows us to bound the corresponding term from the homogeneous suppression of the elements of the propagator implied by the delocalizing transport (see Definition 2). Doing this explicitly is tedious because of the interplay of contributions from the larger patches and those of size two, and the involved combinatorics of how the smaller patches can be distributed on the lattice. All this is done by first ordering contributions according to the number m of patches of size two they contain and then applying Lemma 12, which internally uses a recursive argument. It yields an upper bound on absolute value of sums of the form $\sum_{(k_j)_{j\in J}\in\mathcal{K}_P^{\mathcal{C}}}\prod_{p\in P}\langle\hat{P}_p^{(k_j)_{j\in p}}\rangle_\rho$ that grows with Δ but is algebraically suppressed with time t. Assuming $\Delta\geq 1$ this yields for $t\leq\min(t_{\mathrm{Rec}},V)$ the following bound

$$\begin{aligned}\|\rho^S(t)-\rho_G^S(t)\|_1 \leq 2^{2|S|}\Bigg(2^{2|S|+1}|S|^{2|S|}\sum_{m=0}^{|S|}(2^{2d_{\mathcal{L}}+1}|S|^{d_{\mathcal{L}}}|\mathcal{C}|^{1/4}C_{\mathrm{Trans}}t^{-d_{\mathcal{L}}/3}\Delta^{d_{\mathcal{L}}})^{2m+4}\Bigg[1+C_{\mathrm{Clust}}|\mathcal{C}|^2\mathrm{e}^{-\Delta/\xi} \\ +2^{4(d_{\mathcal{L}}+1)|S|+1}|S|^{2|S|(d_{\mathcal{L}}+1)+1}\sum_{r=0}^{|S|}(|\mathcal{C}|^{1/4}C_{\mathrm{Trans}}t^{-d_{\mathcal{L}}/3}\Delta^{d_{\mathcal{L}}})^{2r+4}\Bigg]^{|S|} \\ +8\tilde{C}_{\mathrm{LR}}(d_{\mathcal{L}})|S|^2\mathrm{e}^{-\mu v_\epsilon|t|}+2^{8|S|+4}|S|^{4|S|+3}C_{\mathrm{Clust}}|\mathcal{C}|^{2|S|}\mathrm{e}^{-\Delta/\xi}\Bigg)\,.\end{aligned} \tag{D56}$$

Recalling that $|\mathcal{C}|\leq 2|S|[2(v+2v_\epsilon)t+1]^{d_{\mathcal{L}}}$ one realizes that by letting Δ grow in time as $\Delta=\max(1,t^{d_{\mathcal{L}}/24})$ all terms are at least algebraically suppressed in t. Therefore for every ϵ there exists a critical system size from which on the bound above decays below ϵ for a suitable relaxation time $t_{\mathrm{Relax}}\leq\min(t_{\mathrm{Rec}},V)$. As the saturation of the the delocalizing transport once t is of the order of V implies a saturation of the bound above at a value defined by $t=V$ we set for systems smaller than the critical system size $t_{\mathrm{Relax}}>t_{\mathrm{Rec}}$. It follows then that $\|\rho^S(t)-\rho_G^S(t)\|_1\leq\epsilon$ for all $t\in[t_{\mathrm{Relax}},t_{\mathrm{Rec}}]$. □