Integrals of motion in the many-body localized phase

V. Ros [a,b,*], M. Müller [c], A. Scardicchio [b,d,e,f,1]

[a] *SISSA, via Bonomea 265, 34136 Trieste, Italy*
[b] *INFN, Sezione di Trieste, Strada Costiera 11, 34151 Trieste, Italy*
[c] *Abdus Salam ICTP, Strada Costiera 11, 34151 Trieste, Italy*
[d] *Physics Department, Princeton University, Princeton, NJ 08544, USA*
[e] *Physics Department, Columbia University, New York, NY 10027, USA*
[f] *ITS, Graduate College of the City University of New York, New York, NY 10016, USA*

Abstract

We construct a complete set of quasi-local integrals of motion for the many-body localized phase of interacting fermions in a disordered potential. The integrals of motion can be chosen to have binary spectrum $\{0, 1\}$, thus constituting exact quasiparticle occupation number operators for the Fermi insulator. We map the problem onto a non-Hermitian hopping problem on a lattice in operator space. We show how the integrals of motion can be built, under certain approximations, as a convergent series in the interaction strength. An estimate of its radius of convergence is given, which also provides an estimate for the many-body localization–delocalization transition. Finally, we discuss how the properties of the operator expansion for the integrals of motion imply the presence or absence of a finite temperature transition.

1. Introduction

The thermodynamic description of macroscopic bodies, as shown by Boltzmann in his work on the foundations of statistical mechanics, is based on the assumption that the underlying

---

* Corresponding author at: SISSA, via Bonomea 265, 34136 Trieste, Italy. Tel.: +39 040 3787 271.
   E-mail address: vros@sissa.it (V. Ros).
[1] On leave from: Abdus Salam ICTP, Strada Costiera 11, 34151 Trieste, Italy.

microscopic dynamics are ergodic. More precisely, one assumes that the environment of any small subsystem of the macroscopic body acts as a thermal bath, with which the subsystem can exchange particles and energy, and which leads to the eventual thermalization of the subsystems, independently of its initial state. In order for thermalization to occur gradients in particle and energy density must be able to even out, which requires non-vanishing transport over arbitrarily large scales.

However, in the absence of interaction, Anderson [1] has shown that a sufficiently strong quenched disorder can localize quantum particles. This prevents the transport of energy and particles and therefore entails the non-ergodicity of the system. Already in Anderson's first paper, and later in the context of electron–electron interactions [2], it was surmised that this localization might persist in the presence of interactions, despite the widespread belief that any finite interactions would restore transport, ergodicity, and thus standard thermodynamic behavior, in such systems. Later, numerical investigations of the Hubbard model [3] hinted indeed at the possibility of such "many-body localization" (MBL), and more recently the seminal study of disordered electrons with weak short range interactions to all orders of perturbation theory provided important analytical insight into this phenomenon [4], predicting that in an isolated system, decoupled from any external bath, a finite interaction is required to induce delocalization and enable transport. Below this delocalization threshold, truly inelastic decay processes are impossible, as the system ceases to be a heat-bath for itself, and any d.c. transport is strictly absent. In this way many-body localized systems are crucially different from other situations where full ergodicity in phase space breaks down, such as in systems with spontaneously broken symmetries, one-dimensional integrable systems, or spin glasses. In all these examples, the thermal conductivity remains finite.[2] In contrast, a necessary[3] condition for many-body localization is the vanishing of the d.c. transport coefficients at non-zero temperature.

Since the seminal work by Basko, Aleiner and Altshuler [9,4,10], the paradigm of many-body localization has attracted a lot of interest, and the phenomenology of MBL phases and the localization transition have been explored, see for example [11–17]. Many-body localization opens the interesting possibilities of protection of topological order at finite temperature or of phase transitions below the equilibrium lower critical dimension [18–23]. It was even proposed that MBL could survive in the absence of quenched disorder [24–28]. (A different type of non-ergodic behavior, exhibiting, however, ballistic transport, has been conjectured in disorder free 1d systems that are close enough to integrability [29,30].)

Unambiguous experimental evidence of an MBL transition or phase is however still lacking at the time of writing, despite of promising developments [31,32].

An MBL phase can be seen as the prototype of a quantum glass phase, where the dynamics are slowed down indefinitely and where memory of the initial condition is retained in local observables for arbitrarily long times. This latter phenomenon has certain similarities with integrable systems [33,34], in which an extensive number of conserved quantities (integrals of motion) constrain the system to evolve in a much smaller submanifold than the one determined

---

[2] In integrable systems, transport of some quantities is often even more efficient than in non-integrable systems, being ballistic as opposed to diffusive.

[3] The condition is not sufficient, since even in the absence of diffusion thermalization might occur via sub-diffusive processes. This was found empirically in one-dimensional systems close enough to the localization transition [5,6]. Moreover, localization can occur also in time-dependent systems (e.g., periodically driven systems) that may have no conserved local densities and thus no meaningful d.c. transport [7,8]. For these systems, MBL is defined more generally as a phase where any local observable does not thermalize almost certainly.

by the conservation of energy and momenta only. The long-time relaxation then only leads to a restricted (generalized) Gibbs ensemble.

Given this similarity, it was conjectured that, like in the non-interacting limit, an extensive set of (quasi-)local integrals of motion should exist in the MBL phase [33,35,34]. By definition, those do not evolve with time, as they commute with the Hamiltonian. They thus constrain the dynamics to remain very close to the initial condition in which the system was prepared. The existence of such local integrals of motion was recently proven for a particular spin Hamiltonian in [35], under reasonable assumptions bounding potential level attraction. The notion of locality used above refers to the set of degrees of freedom, which the conserved operator affects. Conserved quantities in integrable systems are not local in this sense, as they are sums over all space of certain local terms.[4] The non-locality (in our sense) of those integrals still allows for finite transport in integrable systems. A further important difference between MBL systems and integrable ones is the fact that MBL is robust with respect to any sufficiently small perturbation of the Hamiltonian, while integrability in 1d systems is broken by generic perturbations.

The aim of this paper is to show that quasi-local integrals of motion exist for weakly interacting disordered electrons, under the same set of assumptions that were made in the original work by Basko et al. [4] (henceforth referred to as BAA). We find such integrals by solving equations for conserved operators within perturbation theory. Our approach reduces the problem to the solution of a single-particle-like hopping problem in operator space, for which we present a solution in the strongly localized regime, and determine the radius of convergence of the construction. This furnishes an estimate of the delocalization very similar to that obtained by Basko et al. [4]. We hope that our technique will help to obtain analytic results on many-body localization in the future.

1.1. Outline and summary of this work

Here, we present a short outline of this work, summarizing the main steps, and the problems we address.

We are seeking integrals of motion for disordered electrons with weak short range interactions, as defined in Eqs. (1), (2). In Section 2 we coarse-grain the model, reducing it to an array of coupled quantum dots of size of the order of the single-particle localization length.

The non-interacting model has trivial integrals of motion, namely the occupation numbers of the single-particle eigenstates. We then look for their generalization in the presence of interactions, "dressing" these integrals of motion (Section 3). This leads us to a set of linear equations (Section 4, Eq. (40)) in the space of number conserving operators, which we expand in the basis (28) of products of single particle creation and annihilation operators. For any strength λ of the interaction, these equations define a unique set of conserved operators. The main question to analyze is whether they act locally, or whether they significantly affect a spatially unbounded set of degrees of freedom. We address the question of locality within the so-called forward approximation, introduced in Section 5, where we only determine the leading term in perturbation theory for the expansion coefficients. Since the interaction terms act locally, for the conserved quantities to be non-local increasingly high orders of perturbation theory must contribute to the expansion, i.e., the perturbative expansion diverges (Section 5.2).

---

[4] Quantum mechanics provides quite trivially a large set of mutually commuting, conserved operators in any system, namely the projectors on exact many-body eigenstates. However, those are highly non-local and have minimal rank 1. Such trivial "integrals of motion" are of no interest in the present context.

We represent diagrammatically the particle–hole creation processes, which dominate the forward approximation, in Section 5.4. In order to study the statistics of the diagrams at high orders, we need to solve three main technical problems. One is the estimate of their number, due to the freedom in choosing the interaction vertices and their order. We solve this (Section 5.4) by introducing an integral representation that sums correlated diagrams sharing the same interaction vertices. This reduces the factorially many (in the order N of the perturbation theory) terms to an only sub-exponential number of terms, which are products of N denominators. The second problem concerns their statistical distribution. In the many-body problem the denominators are correlated even within the forward approximation, at variance with one-particle problems. Therefore determining the statistics of large deviations, which dominate the probability of creating excitations at large distance, is a challenge. We solve it using a transfer-matrix technique (Section 7). Finally, we have to count the number of processes leading to a given configuration in operator space, which is a combinatorial problem in the space of diagrams (Section 8). The last two ingredients allow us to determine the decay rate of the largest of these terms, which dominates the expansion. Requiring a positive spatial decay rate determines the range of convergence of the operator expansion in the forward approximation.

After solving these technical problems, we obtain the final result in Section 9, namely the existence of quasi-local integrals of motion for disordered electrons for sufficiently small interaction $\lambda < \lambda_c$. We find λ_c in the forward approximation: in the same spirit as Anderson's "upper bound" approximation, this is expected to yield a lower bound for the actual phase boundary for many-body localized phase of the lattice system at infinite temperature. In a final section, we discuss possible scenarios for a localization transition or crossover at finite temperature (Section 10).

2. Model Hamiltonian and coarse-graining

We consider a Hamiltonian describing weakly interacting, spinless electrons in a disordered background. At variance with the work by BAA, we consider a model on a lattice Λ,

$$H = \sum_{i \in \Lambda} c_i^\dagger \left[-\frac{1}{2m} \Delta^{(\Lambda)} + V_{\text{dis}}(i) \right] c_i + \frac{1}{2} \sum_{i,j \in \Lambda} c_i^\dagger c_j^\dagger U(i-j) c_j c_i, \tag{1}$$

where $\Delta^{(\Lambda)}$ is the lattice Laplacian, V_{dis} is a random disordered potential and U is a short range interaction.

We choose to work with a lattice model, because in a finite volume its Hilbert space is finite, and both spectrum and energy per particle are bounded. This will allow us to take a meaningful limit of infinite temperature, and to make statements about many-body localization in that limit.

It is convenient to write the interaction in the form

$$U(i-j) = \frac{\lambda}{\nu a^d} u(i-j) \tag{2}$$

where ν is the density of states, and $u(i-j)$ is a dimensionless, normalized, short-ranged interaction kernel. The dimensionless parameter λ measures the interaction strength.

We consider a disorder potential such that the single particle part of the Hamiltonian possesses only fully localized wave-functions ϕ^α, $\alpha = 1, ..., |\Lambda|$, with typical localization length ξ. Moreover, we are interested in the disorder regime relatively close to single particle delocalization, where ξ is significantly bigger than the lattice spacing a. Let us denote by $\delta_\xi = 1/\nu\xi^d$ the average level spacing in a localization volume, and by W the band width of the single particle problem. The condition $\xi \gg a$ ensures that a large number

$$N_{\text{loc}} = \frac{W}{\delta_\xi} \tag{3}$$

of single particle wave-functions overlap significantly in space. This will provide a large parameter for our analysis.

It is convenient to switch to the basis of single particle wave-functions ϕ^α, in which the Hamiltonian assumes the form

$$
\begin{aligned}
H &= \sum_\alpha \epsilon_\alpha n_\alpha + \sum_{\alpha < \beta, \gamma < \delta} U_{\alpha\beta,\gamma\delta} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta \\
&= \sum_\alpha \epsilon_\alpha n_\alpha + \frac{\lambda}{v a^d} \sum_{\alpha < \beta, \gamma < \delta} u_{\alpha\beta,\gamma\delta} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta
\end{aligned}
\tag{4}
$$

$$\equiv H_0 + U, \tag{5}$$

where $n_\alpha = c_\alpha^\dagger c_\alpha$, and the Greek indices label single particle eigenstates obtained in the absence of interaction. We also choose a certain ordering relation "$<$" among the indices β, γ.

Our choice of the basis ϕ_α is different from that of BAA, who worked with Hartree–Fock (HF) orbitals. Our choice allows us to work in full generality in the operator space, while HF orbitals depend on the non-interacting occupation numbers, i.e., the many body state around which one analyzes stability with respect to interactions. In Section 8.1 we will argue, however, that in the approximation in which we are working, we can neglect the interaction vertices $U_{\alpha\beta,\gamma\delta}$ with two or more coinciding indices, even without resorting to HF, which resums most of those terms. Thus, the two different choices of basis sets lead essentially to the same combinatoric analysis of diagrams.

To simplify the above model further, we assume the single particle energies ϵ_α to be random and uncorrelated. The interaction term U is antisymmetrized: $U_{\alpha\beta,\gamma\delta} = U_{\beta\alpha,\delta\gamma} = -U_{\beta\alpha,\gamma\delta}$. We further simplify it by taking its matrix elements $U_{\alpha\beta,\gamma\delta}$ to be local in space, i.e., they are assumed to be non-zero only if the corresponding single particle states have localization center within one localization volume. Hereby we define the localization center of a single particle state as

$$\vec{r}_\alpha = \int d^d r \, \phi_\alpha^2(r) \vec{r}. \tag{6}$$

Moreover, it is known that the matrix elements decrease rather rapidly (as a power law) when the energy difference between involved levels exceeds the level spacing in the localization volume δ_ξ. This motivates the use of a simplified interaction in which we take $u_{\alpha\beta,\gamma\delta}$ to be non-zero only if

$$|\epsilon_\alpha - \epsilon_\delta|, |\epsilon_\beta - \epsilon_\gamma| \lesssim \delta_\xi \quad \text{or} \quad |\epsilon_\alpha - \epsilon_\gamma|, |\epsilon_\beta - \epsilon_\delta| \lesssim \delta_\xi. \tag{7}$$

In these cases we assume

$$u_{\alpha\beta,\gamma\delta} = \eta_{\alpha\beta,\gamma\delta} \, v a^d \delta_\xi = \eta_{\alpha\beta,\gamma\delta} \left(\frac{a}{\xi}\right)^d, \tag{8}$$

where $\eta_{\alpha\beta,\gamma\delta}$ is a random variable, box-distributed in $[-1, 1]$.

2.1. Coarse-graining

Let us now coarse-grain the model: we assume that the interaction $U_{\alpha\beta,\gamma\delta}$ connects wave-functions either on the same localization volume or on neighboring localization volumes. For

either vertices we assume the same amplitude λ, as long as the restrictions (7) on the energy levels are respected.

This differs from the coarse-graining by BAA, who divided the sample into d-dimensional regions of linear size ξ, restricted the single particle levels α to those regions, but included a (small) hopping term between localization volumes (elastic processes). In contrast the interaction term, responsible for inelastic processes, was restricted to scattering within a given localization cell.

3. Integrals of motion and absence of transport

In the absence of interactions ($\lambda = 0$), the occupation numbers n_α of single particle levels are mutually commuting, conserved quantities. These operators are *quasi-local* in real space, as follows immediately from their expansion in the basis of lattice operators:

$$n_\alpha = \sum_{i,j} \phi_\alpha^*(i)\phi_\alpha(j)c_i^\dagger c_j, \tag{9}$$

where ϕ_α is the corresponding localized single particle eigenfunction. By *quasi-locality* of the n_α we mean that an operator $c_i^\dagger c_j$ contributes in the expansion with a weight which decays exponentially in the distance between the localization center $\vec{r}_\alpha$ of ϕ_α and the sites it acts on (its *support* – here the sites i, j).

By truncating the sum (9) to terms with support only within a neighborhood of $m\xi$ of $\vec{r}_\alpha$ one obtains an operator, whose commutator with the Hamiltonian vanishes up to exponentially small terms. As $m \to \infty$ the operator rapidly converges (in the operator norm) to the conserved n_α. In the non-interacting case this follows directly from the spatial localization of the single particle wave-functions. Our goal is to find an analogue of these operators in the interacting case.

That such a generalization should exist was proven by Imbrie [35] under certain hypotheses on the spectrum in a 1d spin chain, for which he constructed a quasi-local unitary rotation $\mathcal{U}$ which essentially diagonalizes the Hamiltonian H. More precisely, it brings it to the canonical form

$$\mathcal{U}^\dagger H \mathcal{U} = -\sum_i h_i \sigma_i^z - \sum_{i<j} J_{i,j}\sigma_i^z\sigma_j^z + \sum_{i<j<k} J_{i,j,k}\sigma_i^z\sigma_j^z\sigma_k^z + ..., \tag{10}$$

where the k-spin interactions $J_{i_1,...,i_k}$ decay exponentially with the diameter of their index set. Applying the inverse unitary on the conserved spins σ_i^z provides one with integrals of motion of the original Hamiltonian H, $I_i = \mathcal{U}\sigma_i^z\mathcal{U}^\dagger$. At the same time, Huse and Oganesyan [33], and independently, Serbyn et al. [34] argued for the existence of such local integrals of motion in general MBL systems.

Note that the set of conserved and mutually commuting quantities is by no means unique. For example, any set of independent polynomials of σ_i^z's is conserved as well. A nice property of the set of σ_i^z, however, is the binarity of their spectrum, $\{-1, 1\}$, or the property that $(\sigma_i^z)^2 = 1$. Knowing the eigenvalues of N independent integrals of motion like this allows one to unambiguously label the 2^N eigenstates of the Hilbert space of an N-spin system [33].

An alternative construction of conserved (but non-binary) quantities is discussed in [36] for a random spin chain, where infinite time averages of local operators are considered (such as $n_i(t) = e^{iHt}n_i e^{-iHt}$ in our case). By definition of the time average, it commutes with the Hamiltonian. In

an MBL phase one expects the average to remain non-zero, whereas it vanishes due to diffusion in an ergodic delocalized phase.

In this paper, we make a different choice, which nevertheless defines a unique set of binary integrals of motion. Our construction consists in two steps. We will first prove the existence of local integrals of motion in perturbation theory in λ, not requiring the binarity of their spectrum. This is the most difficult task and will take the largest part of the paper.

3.1. Construction of exact quasiparticles of the Fermi insulator

Since our procedure will leave us some freedom in the choice of integrals, in Appendix A we will show how this freedom can be used to fix the spectrum to be binary, order by order in the interaction λ. Notice that the latter amounts to the construction of *exact* quasiparticle occupation numbers of the interacting Fermi insulator. In contrast to Fermi liquids where such exact quasiparticle operators cannot be constructed, neither in real nor in momentum space, it becomes possible in the MBL phase. Rewritten in terms of these occupation numbers $\tilde{n}_\alpha$, the Hamiltonian

$$H = \sum_\alpha \epsilon_\alpha \tilde{n}_\alpha + \frac{1}{2} \sum_{\alpha \neq \beta} J_{\alpha,\beta} \tilde{n}_\alpha \tilde{n}_\beta + ..., \tag{11}$$

can then be seen as an exact quasiparticle energy functional, which determines the energy $E_\alpha^{(\mathrm{qp})}$ of any quasiparticle as a function of the occupations of all others, as:

$$E_\alpha^{(\mathrm{qp})}\big(\{\tilde{n}_\beta\}\big) \equiv \frac{\partial H}{\partial \tilde{n}_\alpha} = \epsilon_\alpha + \sum_{\beta(\neq \alpha)} J_{\alpha,\beta} \tilde{n}_\beta + ... \tag{12}$$

3.2. Complete set of local integrals implies absence of transport

Before outlining the construction of the integrals of motion, let us first show how their existence implies the absence of any d.c. transport, and hence many-body localization.

In order to show the absence of d.c. transport, consider the Kubo formula for the conductivity σ associated with the local current density J_r, associated with a conserved quantity, such as charge or energy. Let

$$J_r(\omega) = \sum_{r'} \sigma\big(r, r'; \omega\big) E(\omega) \tag{13}$$

be the current at frequency ω and position r arising in linear response to a spatially homogeneous field E, and denote by

$$J(\omega) = \frac{1}{V} \sum_r J_r(\omega) \equiv \sigma(\omega) E(\omega) \tag{14}$$

the spatially averaged current density, V being the volume of the system. At finite inverse temperature β, the dissipative part of the conductivity is given by:

$$\mathrm{Re}\big[\sigma(\omega)\big] = -\frac{1}{V} \sum_r \frac{\mathrm{Im}[\Pi(\omega, r)]}{\omega}, \tag{15}$$

where $\Pi(\omega, r)$ is the Fourier transform of the retarded correlation function of the current operator, with Lehmann representation:

$$\Pi(\omega, r) = \frac{1}{\mathcal{Z}} \sum_{m,m'} \sum_{r'} e^{-\beta E_{m'}} \left(1 - e^{-\beta(E_m - E_{m'})}\right) \frac{\langle m'|J_{r'+r}|m\rangle \langle m|J_{r'}|m'\rangle}{\omega + E_{m'} - E_m + i\eta}. \tag{16}$$

Here $\mathcal{Z}$ is the partition function, and the limit $\eta \to 0$ is to be taken after the thermodynamic limit. In the d.c. limit one finds

$$\text{Re}[\sigma(\omega \to 0)] = \frac{\pi\beta}{V} \sum_{r'r} \sum_{m,m'} \frac{e^{-\beta E_{m'}}}{\mathcal{Z}} \langle m'|J_{r'+r}|m\rangle \langle m|J_{r'}|m'\rangle \delta_\eta(E_{m'} - E_m), \tag{17}$$

where we have used

$$\lim_{\omega \to 0} \frac{1 - e^{-\beta\omega}}{\omega} = \beta, \tag{18}$$

and $\delta_\eta(x) = \pi^{-1}\eta/(x^2 + \eta^2)$ is a regularized δ-function.

Let us now show first that for a *complete* set of *strictly local conserved quantities* the conductivity vanishes with probability one in the thermodynamic limit. By strict locality operators we mean that they only act on degrees of freedom that belong to a compact spatial region with finite diameter ζ. We call a set of conserved quantities complete if for any two distinct eigenstates $m \neq m'$ at least one of those integrals of motion takes a different eigenvalue.

For two eigenstates m, m' let $\tilde{I}$ be such a distinguishing integral, with corresponding eigenvalues $\tilde{I}|m\rangle = \tilde{I}_m|m\rangle$ and $\tilde{I}|m'\rangle = \tilde{I}_{m'}|m'\rangle$, with $\tilde{I}_{m'} \neq \tilde{I}_m$. For a strictly local current operator and r sufficiently much bigger than ζ, it follows immediately that one of the two current matrix elements

$$\langle m'|J_{r'}|m\rangle = \frac{\langle m'|[J_{r'}, \tilde{I}]|m\rangle}{(\tilde{I}_m - \tilde{I}_{m'})}, \qquad \langle m'|J_{r'+r}|m\rangle = \frac{\langle m'|[J_{r'+r}, \tilde{I}]|m\rangle}{(\tilde{I}_m - \tilde{I}_{m'})}, \tag{19}$$

vanishes, since one of the two commutators vanishes. Thus, in Eq. (17) the sum over r can be restricted to $r \lesssim \zeta$. Furthermore, for any fixed eigenstate m the sum over eigenstates m' is restricted to a finite set, since $J_{r'}|m\rangle$ can differ only in a finite number ($\leq \exp(c\zeta^d)$, with $c = O(1)$) of integrals of motion from $|m\rangle$. Thus, in the thermodynamic limit, where we have to send $\eta \to 0$, the contribution to the δ-function vanishes with probability one, and thus $\text{Re}[\sigma(\omega = 0)] = 0$. Note that the potentially singular term from $m = m'$ does not contribute because $\langle m|J_r|m\rangle = 0$ by time reversal invariance.

This discussion is of course over-simplified since the actual integrals of motion are only quasi-local, in the sense that there are corrections to a strict locality, which decay exponentially with the diameter of their support on a typical scale ζ. However, the derivation above reflects the essential mechanism by which a complete set of integrals of motion suppresses transport. Consider the matrix elements $\langle m'|J_{r'}|m\rangle$ for eigenstates that differ significantly only in integrals of motion whose support is centered up to a distance $x\zeta$ from r'. These matrix elements are then not exactly zero, but exponentially small in x. There are also exponentially many states m, m' which satisfy these criteria, and thus some energy differences $E_m - E'_m$ in (17) become exponentially small. One might worry that these exponentially small denominators can contribute to the δ-function in the thermodynamic limit, leading to a non-zero conductivity. However, the very construction of the local integrals of motion outlined in the following, and the convergence of that procedure, strongly suggest that with probability tending to one as $\eta \to 0$ the exponential smallness of the

energy denominators is dominated by the decay of the matrix elements in (17), in the sense that at small η the contributions to the δ-function come with weights that are almost surely much smaller than η. If this were not the case, resonant energy denominators would systematically appear in the construction of the conserved quantities and prevent their locality. Therefore, the consistency and convergence of the following construction implies the suppression of d.c. transport in systems admitting a complete set of quasi-local conserved operators.

3.3. Recipe for the construction of integrals of motion

Let us now come back to the actual construction of quasi-local conserved operators. In order to find a generalization of the single particle occupation numbers to the interacting case one should construct an extensive set of $|\Lambda|$ functionally independent operators[5] $\{I_\alpha\}$, which are quasi-local and satisfy

$$[I_\alpha, H] = 0. \tag{20}$$

Since the spectrum of the many-body system is almost surely non-degenerate, it follows that such conserved quantities also satisfy $[I_\alpha, I_\beta] = 0$. Their mutual commutativity implies that they form a commutative algebra. As we discussed above, the choice of a basis spanning this algebra is not at all unique. It is worth mentioning that if the operators I_α commute with H and span the algebra of operators then we can write

$$H = \sum_\alpha \epsilon_\alpha I_\alpha + \sum_{\alpha,\beta} J_{\alpha,\beta} I_\alpha I_\beta + ... \tag{21}$$

as we claimed above. The couplings J's have similar exponential decay as those in (10).

Here we present a specific construction of conserved operators, which fixes the arbitrariness in their definition in a unique way. Our construction starts from the idea that at weak interactions the I_α should be expected to be a perturbed version of the n_α. Thus, we look for a perturbative series in λ,

$$I_\alpha = n_\alpha + \Delta I_\alpha = n_\alpha + \sum_{n\geq 1} \lambda^n \Delta I_\alpha^{(n)}. \tag{22}$$

For the further discussion it is useful to introduce some natural operator subspaces. I_α can be sought as an element of the space C of particle-conserving operators on the Hilbert space, and without loss of generality we may require it to be Hermitian. Since we will require $[H, I_\alpha] = [H_0, I_\alpha] + [U, I_\alpha] = 0$, with H_0 and U as in (5), we consider the kernel K of the linear map $f(X) = [H_0, X]$ defined for $X \in C$, as well as its image, $O = f(C)$. The latter is the orthogonal complement of K with respect to the inner product of operators, $\langle A, B \rangle = \text{Tr}[A^\dagger B]$, $C = K \oplus O$. K is spanned by all possible products of n_α's, while O is spanned by the normally ordered operators

$$\mathcal{O}_{\mathcal{I},\mathcal{J}} = \prod_{\beta \in \mathcal{I}} c_\beta^\dagger \prod_{\gamma \in \mathcal{J}} c_\gamma, \quad \mathcal{I} \neq \mathcal{J}, \tag{23}$$

where the same ordering "$<$" as previously is chosen for the indices β, γ.

---

[5] Functional independence means that no I_α can be expressed as a function of all the other I_β.

At the nth stage of perturbation theory one has to solve the equation

$$[U, \Delta I_\alpha^{(n-1)}] + [H_0, \Delta I_\alpha^{(n)}] = 0. \tag{24}$$

In order for this equation to have a solution one has to make sure that $[U, \Delta I_\alpha^{(n-1)}] \in O$.[6] If this is the case, $\Delta I_\alpha^{(n)}$ is determined up to an element of K. In Appendix A we show how to use this freedom to impose binarity of the spectrum of I_α, $\mathrm{spec}(I_\alpha) = \{0, 1\}$, i.e., $I_\alpha^2 = I_\alpha$. The latter allows these operators to be interpreted as generalized quasiparticle number operators of the interacting Fermi insulator.

Below we describe the construction of conserved I_α based on a simpler choice, however. In particular, we claim that if our Hamiltonian is time-reversal invariant, and thus has real matrix elements in the basis of single particle eigenstates, there is a *unique* solution of (24) with $\Delta I_\alpha \in O$. This choice implies that the only term in the expansion of I_α that commutes with H_0 will be the very first one, n_α. To prove this at the perturbative level, we have to show that one always finds $[U, \Delta I_\alpha^{(n-1)}] \in O$, or equivalently, that, $x(\Psi_0) := \langle \Psi_0 | [U, \Delta I_\alpha^{(n-1)}] | \Psi_0 \rangle = 0$ for every eigenstate Ψ_0 of H_0. One can easily check that at each stage of perturbation theory ΔI_α has real coefficients in the occupation number basis (23). Thus $x(\Psi_0)$ is real. On the other hand, from the anti-Hermiticity of $[U, \Delta I_\alpha^{(n-1)}]$ it follows that $x(\Psi_0)$ is purely imaginary, and thus vanishes indeed.

From the above it follows that we can express the solution of Eq. (24) formally as

$$\Delta I_\alpha^{(n)} = i \lim_{\eta \to 0} \int_0^\infty dt\, e^{-\eta t} e^{iH_0 t} [U, \Delta I_\alpha^{(n-1)}] e^{-iH_0 t}, \tag{25}$$

which determines the successive terms in perturbation theory recursively.

As we show in Appendix A, the recipe to construct a binary operator consists in modifying order by order the terms in the perturbative expansion

$$\Delta I_\alpha^{(n)} \longrightarrow \Delta B_\alpha^{(n)} = \Delta I_\alpha^{(n)} + \Delta K_\alpha^{(n)}, \tag{26}$$

by adding to each $\Delta I_\alpha^{(n)}$ a diagonal operator $\Delta K_\alpha^{(n)} \in K$, which is determined by the previous orders in perturbation theory as:

$$\Delta K_\alpha^{(n)} = (1 - 2n_\alpha) \left[\sum_{m=1}^{n-1} \Delta B_\alpha^{(m)} \Delta B_\alpha^{(n-m)} + \left\{ n_\alpha - \frac{1}{2}, \Delta I_\alpha^{(n)} \right\} \right]. \tag{27}$$

It is plausible that the convergence for binary operators is essentially the same as for the operators constructed below.

Based on the above perturbative argument, we make the following ansatz for the conserved quantities:

$$I_\alpha = n_\alpha + \sum_{\substack{N \geq 1}} \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}| = N = |\mathcal{J}|}} \mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} \left(\mathcal{O}_{\mathcal{I},\mathcal{J}} + \mathcal{O}_{\mathcal{I},\mathcal{J}}^\dagger \right), \tag{28}$$

---

[6] Note that it is not obvious from the outset that this simple perturbative scheme should work and produce a local operator. Indeed we construct perturbation theory for an extensive set of operators which are all null eigenvectors of $[H_0, .]$. In principle one should thus use degenerate perturbation theory for all these operators simultaneously, which could turn out to require a non-local change of basis. The further steps below show, however, that this is not the case.

where the sets $\mathcal{I}, \mathcal{J}$ run over all sets of indices $\{\beta_1 < \cdots < \beta_N\}$ of single particle states. Linear constraints on the coefficients $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ are found by imposing the conservation condition $[I_\alpha, H] = 0$. The coefficients result as λ-dependent functions of the random energies ϵ_α and of the random matrix elements $U_{\alpha\beta,\gamma\delta}$, which vanish in the limit $\lambda = 0$. Since the resulting operators I_α are functionally independent for $\lambda = 0$, we expect the same to hold for any finite λ before the delocalization transition. Indeed it is hard to see how a polynomial of $I_{\beta\neq\alpha}$'s could contain only a single diagonal term n_α.

It is important to note that the expansion (28) should not be seen as an expansion in λ, but rather as an expansion in the support on which the operators $\mathcal{O}_{\mathcal{I},\mathcal{J}}$ act. A formal expansion in λ must always be re-summed locally when rare, but very small denominators are encountered, implying that the naive perturbative series (24), (25) has vanishing radius of convergence in λ [1]. In Appendix B we discuss a simple example where such a re-summation is necessary.

We point out that in any finite system the above ansatz, even though motivated by a perturbative consideration, uniquely determines a conserved operator even if perturbation theory does not converge, despite of re-summations. In that case I_α is defined as the finite (possibly exponentially large) sum (28) whose coefficients satisfy the linear system of Eq. (40) below. In a delocalized regime that operator will have support on the whole system.

3.4. Convergence criterion

We argue that for sufficiently small λ the expansion (28) converges in the operator norm. The convergence holds in probability, that is, for any $\epsilon > 0$:

$$\lim_{R \to \infty} \mathbb{P}\left(\sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=|\mathcal{J}| \\ r(\mathcal{I},\mathcal{J})>R}} \left| \mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} \right| < \epsilon \right) = 1, \tag{29}$$

where $r(\mathcal{I}, \mathcal{J}) = \max_{\beta \in \mathcal{I} \cup \mathcal{J}} |\vec{r}_\alpha - \vec{r}_\beta|$ is the maximal distance between the localization center of the state α and any of the states β that are acted upon by the operator $\mathcal{O}_{\mathcal{I},\mathcal{J}}$. $\mathbb{P}$ is the probability measure over the disorder realizations. This ensures that the series defining the operator I_α converges almost surely, since $\|\mathcal{O}_{\mathcal{I},\mathcal{J}}\| = 1$ for all $\mathcal{I}, \mathcal{J}$.

The resulting operator I_α is quasi-local in the sense defined above. As will become clear below, cf. Section 5.2, one can associate a length scale to the support of these operators like for the non-interacting case: truncating the expansion at that length scale yields operators that are conserved up to exponentially small corrections. This scale is essentially the localization length pertaining to the interacting problem.

The many-body delocalization transition is expected to happen at a sharply defined critical value $\lambda = \lambda_c$ of the interaction strength, at which thermalization and ergodicity are restored. It is natural to expect that this coincides with the delocalization of physically defined conserved quantities, such as the time average of local operators. There is also a sharply defined interaction strength $\lambda = \lambda_c'$ at which our integrals I_α become non-local with probability one. Logically we cannot exclude that λ_c' is slightly smaller than λ_c (since it might be possible to find a prescription for conserved quantities that leads to more local operators than ours); however, we believe that within the approximations we are making, see Section 5, λ_c and λ_c' cannot be distinguished. We therefore use the notation λ_c indistinctly for both critical values.

To discuss the convergence (29), we map the problem of constructing conserved quantities into an equivalent problem of a particle hopping on a disordered lattice whose sites are labeled by

the Fock indices $(\mathcal{I}, \mathcal{J})$. In particular, the exponential decay of the coefficients of I_α corresponds to the localization of the particle on that lattice, in analogy with the non-interacting case (9). In turn, the delocalization of the particle corresponds to the divergence of the operator expansion (28).

4. Explicit construction of the integrals of motion

In this section we present the equations defining $\mathcal{A}_{\mathcal{I},\mathcal{J}}$ in (28) and discuss how to solve them. To illustrate the procedure, we first solve exactly a non-interacting case and then proceed with the interacting problem.

4.1. Non-interacting single-particle example

Consider a non-interacting one-dimensional disordered Hamiltonian:

$$H_{\text{And}} = \sum_i \epsilon_i n_i - t \sum_i \left(c_i^\dagger c_{i+1} + c_{i+1}^\dagger c_i \right), \tag{30}$$

where ϵ_i are random energies and the hopping t is treated perturbatively. In this case, the ansatz

$$I_k = n_k + \sum_{i<j} \mathcal{A}_{ij}^{(k)} \left(c_i^\dagger c_j + c_j^\dagger c_i \right), \tag{31}$$

is consistent. Imposing $[H, I_k] = 0$, we obtain a set of linear equations for the coefficients $\mathcal{A}_{ij}^{(k)}$, one equation for each index k. If for identical indices we define:

$$\mathcal{A}_{ii}^{(k)} \equiv \delta_{k,i} \tag{32}$$

then the equations for $\mathcal{A}_{ij}^{(k)}$ with $i \neq j$ can be compactly written as:

$$(\epsilon_i - \epsilon_j)\mathcal{A}_{ij}^{(k)} - t\left(\mathcal{A}_{i-1\,j}^{(k)} + \mathcal{A}_{i+1\,j}^{(k)} - \mathcal{A}_{ij-1}^{(k)} - \mathcal{A}_{ij+1}^{(k)} \right) = 0. \tag{33}$$

In view of these equations, one may re-interpret $\mathcal{A}_{ij}^{(k)}$ as the wave-function amplitudes of a particle on a square lattice with sites (i, j), and correlated on-site disorder $\mathcal{E}_{i,j} = \epsilon_i - \epsilon_j$, subject to the constraint (32). An explicit expression for them can be given in terms of the eigenfunctions ϕ_α of the Anderson problem (30) as:

$$\mathcal{A}_{ij}^{(k)} = \sum_\alpha \omega_\alpha^k \phi_\alpha(i)\phi_\alpha(j), \tag{34}$$

where the ω_α^k have to be determined from the constraint

$$\sum_\alpha \omega_\alpha^k \left[\phi_\alpha(i) \right]^2 = \delta_{k,i}. \tag{35}$$

The exponential decay of the amplitudes (34) in the distance between the sites i, j follows from the localization in space of the eigenstates ϕ_α. It implies the convergence of the expansion (31). Therefore, the operators I_k are quasi-local conserved operators, similarly to the particle number operators n_α in (9).

However, note that these two sets of operators differ, in particular (31) does not contain any diagonal terms ($i = j \neq k$). Using (34), (35) one can also explicitly check that the operators (31) do not coincide with the time average of the operators $n_k(t)$.

4.2. Interacting case

We now return to the interacting case. Since the operators I_α will contain strings of $c^\dagger$'s and c's of arbitrary length, we need a way to deal with large index sets. We introduce the following notation: for any index set $\mathfrak{X} = (x_1 \cdots x_N)$, we define diagonal coefficients as zero, except if $\mathfrak{X} = \{\alpha\}$:

$$\mathcal{A}^{(\alpha)}_{\mathfrak{X},\mathfrak{X}} \equiv \delta_{\mathfrak{X},\{\alpha\}}. \tag{36}$$

Moreover, for any l, m (with $l < m$) and any single particle labels γ, δ (with $\gamma < \delta$), define the index sets:

$$\begin{aligned}
\mathfrak{X}_l &\equiv (x_1 \cdots \cancel{x}_l \cdots x_N), \\
\mathfrak{X}^\gamma_{lm} &\equiv (\gamma x_1 \cdots \cancel{x}_l \cdots \cancel{x}_m \cdots x_N), \\
\mathfrak{X}^{\gamma\delta}_{lm} &\equiv (\gamma \delta x_1 \cdots \cancel{x}_l \cdots \cancel{x}_m \cdots x_N).
\end{aligned} \tag{37}$$

In general, the set $\mathfrak{X}^{\cdots}_{\cdots}$ is obtained from $\mathfrak{X}$ by eliminating the indices in the subscript and appending the ones in the superscript on the left. Note that the resulting sets are thus not ordered. Let $\sigma[\cdot]$ denote the sign of the permutation which orders the set, and define:

$$\begin{aligned}
s[\mathfrak{X}_l] &\equiv l, \\
s[\mathfrak{X}^\gamma_{lm}] &\equiv l + m + \sigma[\mathfrak{X}^\gamma_{lm}], \\
s[\mathfrak{X}^{\gamma\delta}_{lm}] &\equiv l + m + \sigma[\mathfrak{X}^{\gamma\delta}_{lm}].
\end{aligned} \tag{38}$$

Finally, for index sets with $|\mathcal{Y}| = |\mathcal{Z}|$, define the modified amplitudes:

$$\tilde{\mathcal{A}}^{(\alpha)}_{\mathcal{Y},\mathcal{Z}} \equiv (-1)^{s[\mathcal{Y}]+s[\mathcal{Z}]} \mathcal{A}^{(\alpha)}_{\mathcal{Y},\mathcal{Z}}. \tag{39}$$

With this notation, the condition $[H, I_\alpha] = 0$ is equivalent to the following set of linear equations for $\mathcal{A}^{(\alpha)}_{\mathfrak{I},\mathfrak{J}}$:

$$\begin{aligned}
0 = {}& \left(\sum_{n=1}^{N} \frac{\epsilon_{\alpha_n} - \epsilon_{\beta_n}}{\delta_\xi} \right) \mathcal{A}^{(\alpha)}_{\mathfrak{I},\mathfrak{J}} \\
& + \lambda \sum_{\substack{l,m=1 \\ l<m}}^{N} \left[\sum_{\gamma<\delta} \left(\eta_{\alpha_l \alpha_m, \gamma\delta} \tilde{\mathcal{A}}^{(\alpha)}_{\mathfrak{I}^{\gamma\delta}_{lm},\mathfrak{J}} - \eta_{\gamma\delta, \beta_l\beta_m} \tilde{\mathcal{A}}^{(\alpha)}_{\mathfrak{I},\mathfrak{J}^{\gamma\delta}_{lm}} \right) \right] \\
& + \lambda \sum_{\substack{l,m=1 \\ l<m}}^{N} \sum_{n=1}^{N} (-1)^{N+1} \left[\sum_{\gamma} \left(\eta_{\alpha_l \alpha_m, \gamma\beta_n} \tilde{\mathcal{A}}^{(\alpha)}_{\mathfrak{I}^\gamma_{lm},\mathfrak{J}_n} - \eta_{\gamma\alpha_n, \beta_l\beta_m} \tilde{\mathcal{A}}^{(\alpha)}_{\mathfrak{I}_n,\mathfrak{J}^\gamma_{lm}} \right) \right],
\end{aligned} \tag{40}$$

where $(\mathfrak{I}, \mathfrak{J}) = (\alpha_1 \cdots \alpha_N, \beta_1 \cdots \beta_N)$ and $\mathfrak{I} \neq \mathfrak{J}$. The diagonal coefficients appearing on the right-hand side are defined in (36).

4.2.1. Topology of the operator lattice

Similarly as in the previous single-particle example, Eq. (40) can be thought of as a hopping problem for a single particle on a lattice with sites given by the Fock indices $(\mathfrak{I}, \mathfrak{J})$ and local, correlated disorder $\mathcal{E}_{\mathfrak{I},\mathfrak{J}} = \sum_{n=1}^{N} (\epsilon_{\alpha_n} - \epsilon_{\beta_n})$. The hopping is provided by the interaction U, see
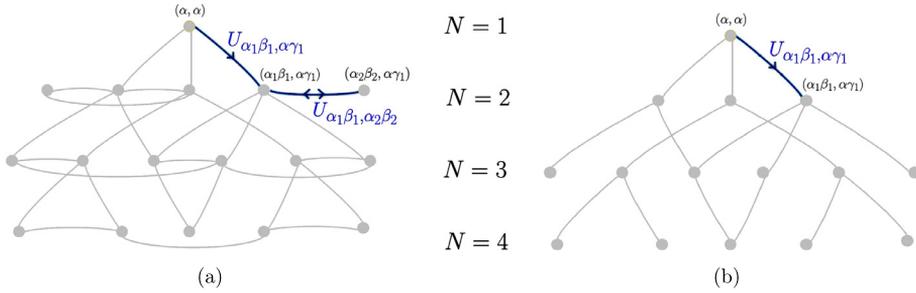
Fig. 1. Structure of the operator lattice before (a) and after (b) making the forward approximation. Vertices correspond to Fock indices $(\mathcal{I}, \mathcal{J})$; links are drawn between index pairs, which are connected by the interaction U, that is, if the pairs appear simultaneously in at least one of Eq. (40).

Fig. 1(a). The non-interacting limit corresponds to the wave-function $\mathcal{A}^{(\alpha)}$ being completely localized on the site $(\mathcal{I}, \mathcal{J}) = (\alpha, \alpha)$.

The lattice topology, as determined by the interactions, is rather complicated. However, Eq. (40) has a clear hierarchical structure: the equation for index sets $\mathcal{I}, \mathcal{J}$ of length N are coupled only to amplitudes with index sets of equal or shorter length. Therefore, the sites can be organized into generations, according to the length of their index sets. Hopping is possible only within the same generation (second term in Eq. (40)) or between consecutive ones (third term in Eq. (40)). In the latter case, the hopping is unidirectional, and thus the hopping problem is non-Hermitian.

The connectivity of the lattice is determined by the restrictions in energy, Eq. (7), and space (particles need to be in the same or in an adjacent localization volume) of the matrix elements $U_{\alpha\beta,\gamma\delta}$. Hoppings from a site $(\mathcal{I}, \mathcal{J})$ in generation N to a site $(\mathcal{I}', \mathcal{J}')$ in generation $N+1$ requires a particle (or hole) in a state α to scatter to the closest energy level γ above or below α, while another particle–hole pair of adjacent levels (β, δ) is created. The particle β can be chosen in N_{loc} ways with N_{loc} given in (3), and there are two choices for γ and δ, respectively. Therefore, the number of Fock states $(\mathcal{I}', \mathcal{J}')$ accessible from $(\mathcal{I}, \mathcal{J})$ via the decay of a given quasiparticle α is:

$$\mathcal{K} = 4\frac{W}{\delta_\xi} = 4N_{\text{loc}}. \tag{41}$$

In contrast, hoppings from $(\mathcal{I}, \mathcal{J})$ to a site of the same generation correspond to processes where each member of a pair of particles (or holes) scatter to one of the two closest energy levels: there are 4 possible final states to which a given pair can decay.

At this point we emphasize that we are not restricting ourselves to a specific many-body state or energy sector. Thus no assumption about the occupation of the levels or about the position of the Fermi level E_F is made. This gives the largest possible connectivity $\mathcal{K}$. It will be reduced to an effective connectivity once we consider the restriction of the integrals I_a to subspaces of a definite energy by means of a projector over many-body states, $\tilde{I}_a = P I_a P$, where

$$P = \sum_{E_a \in [E-\delta E/2, E+\delta E/2]} |E_a\rangle\langle E_a|. \tag{42}$$

This projection will alter the connectivity $\mathcal{K}$, so as to reflect the higher probability for some processes to be Fermi-blocked, since the involved levels might already be occupied. This

yields an effective connectivity $\mathcal{K}_{\text{eff}}$, whose typical value depends both on the average energy density of the states E_a and the average filling fraction of the band. It is not difficult to see that if we use typical values for occupation numbers as given by the Fermi distribution (without assuming the underlying states to be thermal), repeating the above considerations at finite temperature $T \ll E_F$ we obtain $\mathcal{K}_{\text{eff}} \sim T/\delta_\xi$, in analogy to the analysis in [4].

5. The forward approximation

5.1. Simplifications due to large connectivity, $\xi \gg a$

The requirement of convergence of the operator expansion, Eq. (29), can be interpreted as a localization condition for the hopping problem on the disordered lattice of Fock indices. In order to investigate under which conditions localization occurs, we introduce the main approximation of this work: we neglect the second term of Eq. (40), that accounts for the hopping between sites in the same generation.

This approximation is motivated by the following consideration, assuming that the number of single particle levels per localization volume, and thus $\mathcal{K}$, is large: for operator sites with a density of Fock indices per localization volume much smaller than the maximally possible $\sim \mathcal{K}/\xi^d$, the connectivity within the same generation is much smaller than the connectivity $\mathcal{K}$ among sites in different generations (41). Note, however, that transitions from a given state $(\mathcal{I}, \mathcal{J})$ due to the second term of Eq. (40) can involve any pair of particles or holes in the same localization volume. Therefore, for operators with a high density of indices per localization volume those transitions are as numerous as the third class of terms in Eq. (40). Our approximation of dropping the second term is therefore not fully controlled at sufficiently high orders in perturbation theory where operators with a high density of indices per localization volume appear. We postpone further discussions of the subtleties related to this approximation to Section 10.

Once the second term in (40) is dropped, the equations reduce to recursive equations for increasing generations, with the initial condition $\mathcal{A}^{(\alpha)}_{\alpha_1,\beta_1} = \delta_{\alpha_1,\beta_1}\delta_{\alpha_1,\alpha}$. However, only some of the amplitudes $\mathcal{A}^{(\alpha)}_{\mathcal{I},\mathcal{J}}$ in (28) are determined through the recursion, while we approximate all other amplitudes to be zero: in generation N, the non-zero amplitudes correspond to sites $(\mathcal{I}, \mathcal{J})$ that can be reached from (α, α) via *directed* paths of length $N - 1$. Retaining only these sites simplifies the structure of the lattice of Fock indices very substantially, see Fig. 1(b).

The amplitudes on these sites $(\mathcal{I}, \mathcal{J})$ can be written as the sum over all directed paths that connect them to the root (α, α) in Fig. 1(b):

$$\mathcal{A}^{(\alpha)}_{\mathcal{I},\mathcal{J}} = \sum_{\substack{\text{directed paths:} \\ (\alpha,\alpha) \to (\mathcal{I},\mathcal{J})}} \omega_{\text{path}}. \tag{43}$$

The path weights ω_{path} are of the form:

$$\omega_{\text{path}} \equiv (-1)^{\sigma_{\text{path}}} \prod_{i=1}^{N-1} \frac{\lambda \eta_{\alpha_i \beta_i, \gamma_i \delta_i} \delta_\xi}{\sum_{k=1}^{i} \mathcal{E}_{\alpha_i \beta_i, \gamma_i \delta_i}} \tag{44}$$

in close analogy to forward approximations in single particle problems [1,37–40].

The factor $(-1)^{\sigma_{\text{path}}}$ takes into account the global fermionic sign associated with the path, arising from the sign factors in Eq. (40). However, we will see below that these signs are immaterial at the level of our approximation.

Note that the resulting expression for $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ is of order λ^{N-1}, that is, the lowest possible order in λ for amplitudes of operators involving $2N$ particle–hole indices. Indeed, at least $N-1$ interactions are needed to create the corresponding excitations.

5.2. Probability of resonances on the operator lattice

Let us discuss the configuration in real space of the indices $(\mathcal{I}, \mathcal{J})$ with $|\mathcal{I}| = N$, which are retained within the forward approximation, cf. Fig. 1(b). Since the amplitudes $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ are of order λ^{N-1} and the interaction is local, the indices satisfy $r(\mathcal{I},\mathcal{J}) \leq N\xi$: amplitudes involving single particle states sufficiently far away from the localization center α must belong to sufficiently high generations. Within the approximations made, the convergence criterion (29) can then be restated in terms of the generation number N as:

$$\lim_{N^* \to \infty} \mathbb{P}\left(\sum_{\substack{N > N^*}} \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N=|\mathcal{J}|}} \left| A_{\mathcal{I},\mathcal{J}}^{(\alpha)} \right| < \epsilon \right) = 1 \tag{45}$$

for arbitrary $\epsilon > 0$.

A sufficient condition for Eq. (45) to hold is that for some $z < 1$ and for N^* sufficiently big:

$$\mathbb{P}\left(\forall N > N^*, \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N=|\mathcal{J}|}} \left| A_{\mathcal{I},\mathcal{J}}^{(\alpha)} \right| < z^{N-1} \right) = 1 - \zeta\left(N^*\right) \tag{46}$$

with

$$\lim_{N^* \to \infty} \zeta\left(N^*\right) = 0. \tag{47}$$

The left hand side of Eq. (46) can be interpreted as the probability that no resonance[7] occurs at large distance from the unperturbed localization center (α, α). Whenever it holds, it implies the quasi-locality of the operators I_α within the forward approximation: indeed, Eq. (46) implies that the first appearance of operators $c_\beta, c_\beta^\dagger$'s in I_α, with $|\vec{r}_\beta - \vec{r}_\alpha| \approx N\xi$ and $N \gg 1$ is with high probability exponentially small in N.

In the following we will show that Eq. (46) holds in a regime of small couplings λ; the critical value λ_c at which (46) ceases to hold gives an estimate for the radius of convergence of the operator series, and thus for the boundary of the many-body localized phase.

5.3. Similarities and differences with localization problems on trees

The similarity to a one-particle problem allows us to revisit analogies and differences between many-body localization and single particle problems on lattices which have some features of a Cayley tree [41] (see also [42,43] and references therein). Indeed, in the simplified lattice

---

[7] A resonance is said to occur at a site $(\mathcal{I}, \mathcal{J})$ if $\mathcal{A}_{\mathcal{I},\mathcal{J}}$ is comparable with the amplitude at the origin (α, α), i.e., if it is of order $O(1)$.
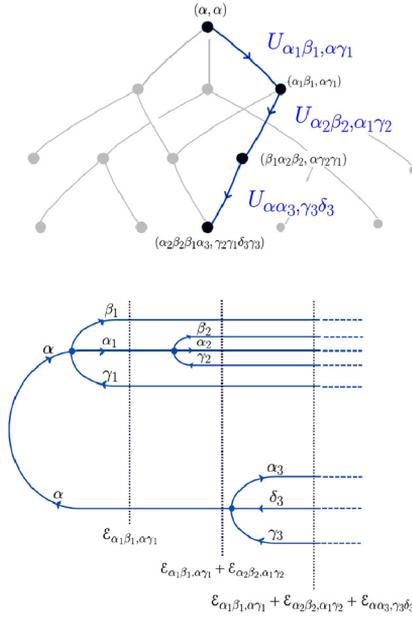
Fig. 2. Directed path in the operator lattice and associated ordered scattering graph. The sites $(\mathfrak{I}, \mathfrak{J})$ along the path correspond to the intermediate states of the graph, indicated by dashed lines. Hoppings on the lattice correspond to vertices $U_{\alpha_1\alpha_2,\beta_1\beta_2}$ in the graph. The energy $\mathcal{E}_{\mathfrak{I},\mathfrak{J}}$ of an intermediate state is the sum of the energy differences $\mathcal{E}_{\alpha_1\alpha_2,\beta_1\beta_2} = \epsilon_{\alpha_1} + \epsilon_{\alpha_2} - \epsilon_{\beta_1} - \epsilon_{\beta_2}$ associated with all preceding scatterings. The three excitations emanating from a vertex are associated to the outgoing legs as follows: the excitation with energy level adjacent to the incoming one is associated with the central leg. The upper and lower leg correspond to the particle and the hole, respectively, of the additionally created pair. The condition (7) requires them to have an energy difference of the order of δ_ξ.

of Fig. 1(b), the number of sites at distance N from the localization center (α, α) grows as $\mathcal{K}^N$ with $\mathcal{K}$ given in (41). This exponential growth is analogous to the growth on trees and other hierarchical lattices, see e.g. [44]. However, we caution the reader that, despite superficial similarities, the calculation we will perform does not reduce to studying an equivalent single particle problem on a Cayley tree as in [45]. Indeed, in the latter problem there is a unique path leading from the root to a given site and thus there are no loops. In contrast, in the operator lattice, there are typically exponentially many diagrams (or effective paths) leading to a given site, and thus plenty of loops, similarly as in finite dimensional lattices. Nevertheless, it is usually the case that among those many paths only very few dominate the sum over all paths – an observation we will heavily rely on in the sequel.

Our present problem also differs from the study of the decay of excitations in a zero-dimensional quantum dot, as considered in [41]. There, no genuine delocalization can take place due to the finite available phase space. Instead, it is essential that our operator expansion leave the localization volume of the initial state α, for delocalization to be possible beyond a critical interaction strength λ_c.

5.4. Connection with many-body diagrammatic perturbation theory

Insight into the meaning of the forward approximation at the level of the many-body system is given by a diagrammatic representation of the paths, as shown in Fig. 2.
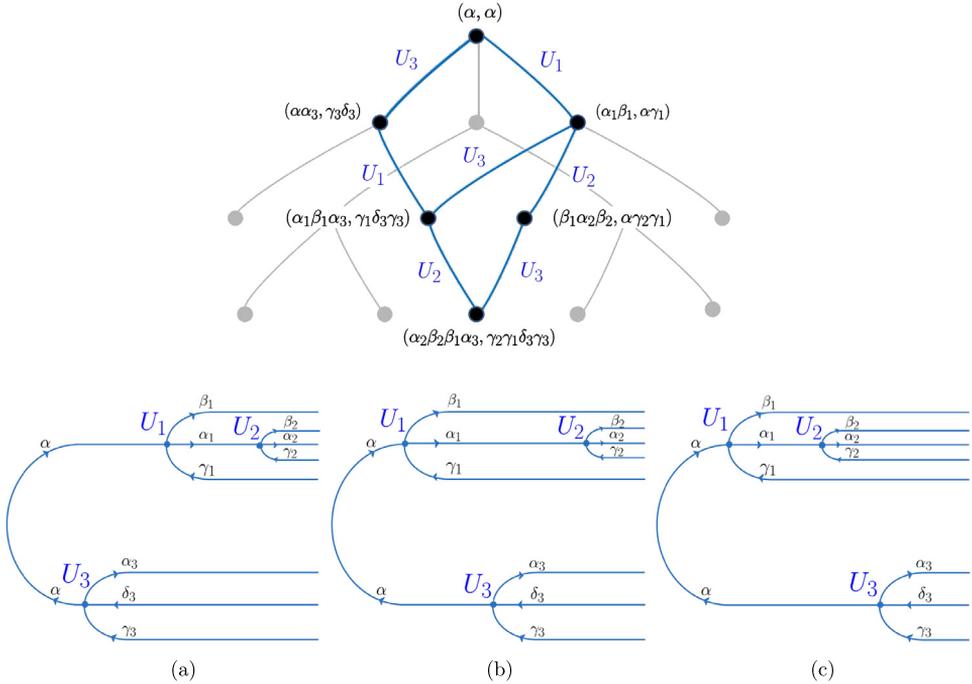
Fig. 3. Loops in the many-body lattice corresponding to different processes with the same final state, and the corresponding ordered graphs. The graphs differ only in the order in which the interactions U_1, U_2, U_3 act. The weights of such paths are strongly correlated: they are all proportional to the same product of matrix elements, $U_1U_2U_3$, and have highly correlated denominators. The sum over all these ordered graphs constitutes a *diagram*.

To any path of length N in the operator lattice we uniquely associate an ordered graph with N vertices. These graphs have two main branches representing the decay of the operators c_α and $c_\alpha^\dagger$ of the initial operator n_α. Directed paths of length N on the lattice translate into graphs having the geometry of a tree, with a root and N nodes corresponding to the creation of particle–hole pairs. The intermediate states of the graph correspond to the sites $(\mathcal{I}, \mathcal{J})$ along the path in the operator lattice, their energy being $\mathcal{E}_{\mathcal{I},\mathcal{J}}$. Note that the order of the sites along the path fixes the order of the interaction vertices in the graph.

Such graphs can be grouped into *diagrams*: members of the same diagram only differ in the ordering of vertices, while sharing the same geometry and labeling of the legs; they are obviously highly correlated among each other. An example is shown in Fig. 3, where all three paths connect the state $(\mathcal{I}, \mathcal{J}) = (\alpha_2\beta_2\beta_1\alpha_3, \gamma_2\gamma_1\delta_3\gamma_3)$ to the root (α, α), and involve the same interaction matrix elements.

Such correlated paths exist for all diagrams with *branchings* (i.e., vertices where more than one of the outgoing excitations undergo further scattering). The order of the subsequent interactions on different branches can be permuted. This corresponds to different paths on the lattice and different ordered graphs, respectively.

Obviously we should sum over all possible vertex order permutations of branched diagrams with fixed geometry and labeling of legs.
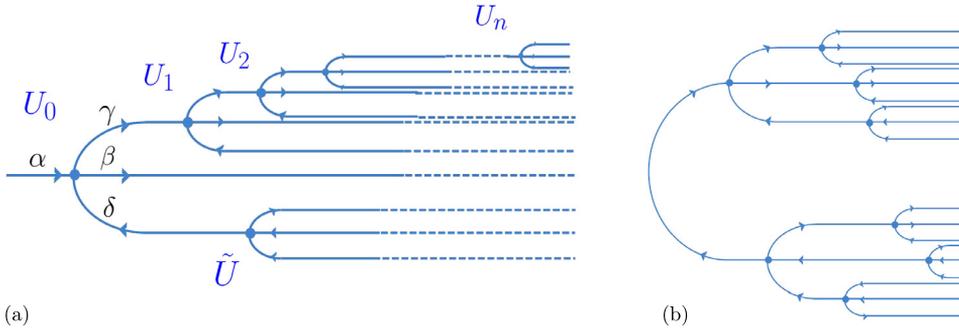
Fig. 4. (a) Branched decay of a single particle. (b) Geometry of a diagram with the maximal possible number of branchings for a fixed number of interactions.

5.4.1. Singly branched diagrams

Consider the sum of the energy denominators[8] of the three path weights in the example of Fig. 3. It is immediate to check that the following holds:

$$\Sigma \equiv \frac{1}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_2)(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3)} + \frac{1}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_3)(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3)}$$
$$+ \frac{1}{\mathcal{E}_3(\mathcal{E}_3 + \mathcal{E}_1)(\mathcal{E}_3 + \mathcal{E}_1 + \mathcal{E}_2)} = \frac{1}{\mathcal{E}_3} \frac{1}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_2)}, \tag{48}$$

where $\mathcal{E}_i$ is the energy difference between out- and in-going states at the vertex i. Thus, the sum over the three paths weights in Fig. 3 can be written as a single term $\tilde{\omega}_\Gamma$:

$$\frac{\tilde{\omega}_\Gamma}{(\lambda \delta_\xi)^3} \equiv \frac{\eta_3}{\mathcal{E}_3} \frac{\eta_1 \eta_2}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_2)}, \tag{49}$$

where η_i is the random variable associated the vertex i. More precisely, $\tilde{\omega}_\Gamma$ is the product of two weights of the form (44), describing the independent decay of the particle $c_\alpha^\dagger$ and the hole c_α, respectively. It can easily be checked by induction that this factorization generalizes to an arbitrary number of interactions in such singly branched diagrams: for any of them, a weight of the form (49) is obtained by summing over all the path weights. We refer to $\tilde{\omega}_\Gamma$ as the weight of the *effective path* associated to the diagram, and denote the latter by Γ.

5.4.2. Multiply branched diagrams

Let us now discuss further branchings in the sub-diagrams describing the independent decays of the particle $c_\alpha^\dagger$ and the hole c_α. Consider a multi-branched decay of the single particle $c_\alpha^\dagger$, as shown in Fig. 4(a). There the particles γ and δ, which are produced in the first scattering, decay further through n vertices $U_{i=1,\dots,n}$, and the vertex $\tilde{U}$, respectively. The possible orderings of this diagram correspond to $n + 1$ correlated paths, which differ by the relative position of the vertex $\tilde{U}$ with respect to the U_i. Their sum,

$$\Sigma' = \frac{1}{\mathcal{E}_0(\mathcal{E}_0 + \tilde{\mathcal{E}})(\mathcal{E}_0 + \tilde{\mathcal{E}} + \mathcal{E}_1) \cdots (\mathcal{E}_0 + \tilde{\mathcal{E}} + \cdots + \mathcal{E}_n)}$$

---

[8] The global sign of amplitudes of tree-like diagrams without loops does not depend on the order in which the interactions act. This is because the associated four-fermion interaction terms mutually commute, which implies that the signs arising from eventually bringing the operators into the normal order are the same for all vertex orders.

$$+ \frac{1}{\mathcal{E}_0(\mathcal{E}_0 + \mathcal{E}_1)(\mathcal{E}_0 + \mathcal{E}_1 + \tilde{\mathcal{E}}) \cdots (\mathcal{E}_0 + \mathcal{E}_1 \cdots + \mathcal{E}_n)} + \cdots$$

$$+ \frac{1}{\mathcal{E}_0(\mathcal{E}_0 + \mathcal{E}_1)(\mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2) \cdots (\mathcal{E}_0 + \mathcal{E}_1 \cdots + \tilde{\mathcal{E}})}, \tag{50}$$

does not simply factorize, but it can nevertheless be written in compact form through an integral representation,

$$\Sigma' = \lim_{\epsilon \to 0} \int \frac{d\omega_1 d\omega_2 \delta(\omega_1 + \omega_2 - \mathcal{E}_0)}{\omega_1^-(\omega_1^- + \tilde{\mathcal{E}}) \cdot \omega_2^-(\omega_2^- + \mathcal{E}_1) \cdots (\omega_2^- + \mathcal{E}_1 + \cdots + \mathcal{E}_n)}, \tag{51}$$

where $\omega_i^- = \omega_i - i\epsilon$. Indeed, the sum Σ' (multiplied by the matrix elements of the correspondent vertices) must be equal to the retarded Green function associated to the independent, parallel decay of the particle γ and the hole δ, computed in the forward scattering approximation and at energy $\mathcal{E}_0$. For loop-free graphs like the one of Fig. 4(a), the decay processes of the particle γ and the hole δ are independent. In the time domain, the Green function of their joint decay is the product of the individual Green functions, which leads to the convolution (51) in frequency space.

The above formula is rather natural when relating with standard many-body perturbation theory. Indeed, after the summation over orderings of vertices, the diagrams of a fixed geometry are in direct correspondence with the diagrams obtained by BAA in the perturbative expansion of the Keldysh self energy in the imaginary self consistent Born approximation. The latter neglects the renormalization of the real part of the self energy and retains only processes where at each vertex an additional particle–hole pair is created. In our formalism, this corresponds to the directed paths jumping from generation to generation, see also the discussion in Appendix B. Not surprisingly, the statistical analysis of this class of diagrams will give an estimate of the radius of convergence for the operator expansion (28) which is similar to the criterion for the breakdown of stability of the localized phase found by BAA, or to its extension to infinite temperature [46]. Our further analysis is also very similar to the calculation in Ref. [47], but differs in some points, which will be indicated.

The expression (51) for a branched diagram is a random variable, whose probability distribution is hard to analyze. However, the analytic structure of the integrand can be exploited to rewrite Σ' as a sum over a much smaller number of terms than the number of orderings in Eq. (50). After performing the integral over ω_2 in Eq. (51), we find a number of poles in the complex plane of ω_1. Using the residue theorem, we can write (51) as the sum over residues of the poles in the half plane, which contains less poles. In the particular example considered, closing the contour on the upper half plane yields the algebraic identity:

$$\Sigma' = \frac{1}{\tilde{\mathcal{E}}} \frac{1}{\mathcal{E}_0(\mathcal{E}_0 + \mathcal{E}_1)(\mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2) \cdots (\mathcal{E}_0 + \mathcal{E}_1 + \cdots + \mathcal{E}_n)}$$

$$- \frac{1}{\tilde{\mathcal{E}}} \frac{1}{(\mathcal{E}_0 + \tilde{\mathcal{E}})(\mathcal{E}_0 + \tilde{\mathcal{E}} + \mathcal{E}_1)(\mathcal{E}_0 + \tilde{\mathcal{E}} + \mathcal{E}_1 + \mathcal{E}_2) \cdots (\mathcal{E}_0 + \tilde{\mathcal{E}} + \mathcal{E}_1 + \cdots + \mathcal{E}_n)}. \tag{52}$$

The two terms in (52) have a similar structure as the denominators in the original path weight (44). For the considered sub-diagram, the sum over all the $n + 1$ orderings of vertices could thus be reduced to the sum of only two "effective path" weights.

5.4.3. General branched diagrams

A convolution formula analogous to Eq. (51) can be written for any branched diagram: to each branching one associates an integral of the form (51) with one auxiliary frequency per decaying branch, as well as an energy conserving δ-function for the vertex (see Appendix C for an example). Then one eliminates the δ-functions by integrating over the frequency variable, that occurs most often in the denominators. Using the residue theorem, the remaining integrals can be carried out, and the sum over all orderings of a diagram with fixed geometry can be expressed as a much smaller sum of weights of effective paths, as in the example above. The number of such terms is given by the product of the number of residues obtained for each auxiliary frequency.

The number of effective paths associated to a general diagram depends on its structure; to obtain an upper bound on this number, consider the diagram with the maximal number of branchings at fixed order N, see Fig. 4(b). In Appendix C, we show that in this case the number of effective paths scales as $\exp[\log 3 (\log N)^2 + O(\log N \log(\log N))]$. This upper bound implies that the number of effective paths associated to an arbitrary diagram is always *sub-exponential* in N.

6. Summing diagrams

In this section we show that in the localized region, at a any given order of the expansion, a few terms dominate the operator sum. The term with the largest coefficient in turn is dominated by the maximal diagram contributing to it.

6.1. Summing over diagrams and their effective paths

Let $\mathcal{D}_{\mathcal{I},\mathcal{J}}$ denote the set of all diagrams with final state $\mathcal{I}, \mathcal{J}$, each diagram being characterized by its geometry and the labeling of its segments. For any diagram $d \in \mathcal{D}_{\mathcal{I},\mathcal{J}}$, let $\mathcal{P}(d)$ be the set of effective path weights $\tilde{\omega}_\Gamma$ associated to it, following the procedure described in the previous section. The corresponding amplitude on the operator lattice can then be written as

$$\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} = \sum_{d \in \mathcal{D}_{\mathcal{I},\mathcal{J}}} \left(\sum_{\Gamma \in \mathcal{P}(d)} \tilde{\omega}_\Gamma \right) \equiv \sum_{d \in \mathcal{D}_{\mathcal{I},\mathcal{J}}} S(d). \tag{53}$$

As we shall prove in the following section, the $\tilde{\omega}_\Gamma$ are random variables with fat-tailed distributions. The effective paths associated to a diagram $d \in \mathcal{D}_{\mathcal{I},\mathcal{J}}$ all involve the same set of energies in their denominators and are thus correlated. Nevertheless, we argue that the tail of the distribution of their sum, $S(d)$, is still very similar to the tail distribution of a single effective path, since in the case of a large deviation, $S(d)$ is very likely to be dominated by the effective path with the biggest weight. Indeed, consider a rare set of energies $\mathcal{E}_i$, which produces an atypically large value of $S(d)$. There is typically one single effective path for which all denominators become simultaneously small, while the combination of energies in the denominators of other effective paths are very likely to be suboptimal for a fraction of the denominators. Therefore, with high probability, $S(d)$ will approximately be equal to the maximum over all effective paths weights: $S(d) \approx \max_{\Gamma \in \mathcal{P}(d)} \tilde{\omega}_\Gamma$.

The set of energies $\mathcal{E}_i$ that optimize distinct effective paths are typically different, and thus these rare events can be approximated as being independent from each other. Hence, the tail of the distribution of $S(d)$ is enhanced with respect to the tail of a single path weight by a factor $|\mathcal{P}(d)|$. We shall see, however, that due to the sub-exponential scaling of the number of effective

paths, this enhancement is immaterial for the estimate of the radius of convergence of the operator series.

Inspecting the explicit examples of Eq. (52) or Eq. (C.3), one can see that there exist energy realizations for which cancellations occur between effective paths with significant weight. This happens when the single path weights are individually big, but $\tilde{\mathcal{E}}$ is much smaller than all the other energy variables $\mathcal{E}_i$, which leads to a cancellation between effective paths. However, such configurations require an atypically small $\tilde{\mathcal{E}}$ and do not occur with significant probability. Therefore the suppression of the tail distribution due to such effects is hardly relevant.

Correlations between effective path weights of different diagrams are even weaker than those above, since they share at most a fraction of all $\mathcal{E}_i$. Therefore we may approximate rare deviations of $S(d)$ and $S(d')$ as independent if $d \neq d'$. Given that the $S(d)$ are themselves fat-tailed random variables, the sum over diagrams is dominated by the largest term. Therefore, the full operator amplitude $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ is likely to be dominated by one single effective path:

$$\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} \approx \max_{d \in \mathcal{D}_{\mathcal{I},\mathcal{J}}} \left(\max_{\Gamma \in \mathcal{P}(d)} \tilde{\omega}_\Gamma \right) \approx \max_{\Gamma : (\alpha,\alpha) \to (\mathcal{I},\mathcal{J})} \tilde{\omega}_\Gamma \tag{54}$$

where on the right hand side the maximum is taken over all effective paths from (α, α) to $(\mathcal{I}, \mathcal{J})$. As a consequence, for the tail of the probability distribution we obtain the approximation

$$P\left(\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} = a\right) \approx |\mathcal{D}_{\mathcal{I},\mathcal{J}}| \overline{\mathcal{P}(d)} P(\tilde{\omega}_\Gamma = a), \tag{55}$$

where $\overline{\mathcal{P}(d)}$ is an average number of effective paths contributing to a diagram.

6.2. Summing over amplitudes: probability of resonances

Similarly to the effective path weights of different diagrams, also the amplitudes $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ associated to different sites $\mathcal{I}, \mathcal{J}$ are weakly correlated, and we treat them as independent random variables. Let us now consider the probability in (46):

$$\mathbb{P}\left(\forall N > N^*, \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| < z^N\right) \approx \prod_{N > N^*} \mathbb{P}\left(\sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| < z^N\right). \tag{56}$$

Here we approximated the probability to satisfy the condition at each generation to be independent from the previous generations. As follows from (55) and from the fact that the effective paths $\tilde{\omega}_\Gamma$ have fat tails, the amplitudes $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ have themselves a fat-tailed distribution. Their sum is therefore dominated by the maximal amplitude, and each factor on the right hand side (56) can be computed as:

$$\mathbb{P}\left(\operatorname*{Max}_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| < z^N\right) = \prod_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \left(1 - \mathbb{P}\left(\left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| > z^N\right)\right)$$

$$\approx \exp\left(- \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \mathbb{P}\left(\left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| > z^N\right)\right). \tag{57}$$

Using (55), the exponent in (57) is re-written as:

$$\sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \mathbb{P}\left(\left|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}\right| > z^N\right) = \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} |\mathcal{D}_{\mathcal{I},\mathcal{J}}| \overline{\mathcal{P}(d)} \mathbb{P}\left(|\tilde{\omega}_\Gamma| > z^N\right). \tag{58}$$

The probability in (58) is a large deviation probability: indeed, the weights $\tilde{\omega}_\Gamma$ of effective paths are of order $O(\lambda^N)$: in order for $\tilde{\omega}_\Gamma$ to be bigger than z^N (with z arbitrarily close to 1), this decay factor must be compensated by an atypical smallness of the energy denominators. We devote the following section to the computation of the probability of these large deviation events. The calculation will reveal that, for λ sufficiently small, the probability decays exponentially with N. This decay competes with the exponential growth of the total number of effective paths of length N:

$$\mathcal{N}_N \equiv \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} |\mathcal{D}_{\mathcal{I},\mathcal{J}}| \overline{\mathcal{P}(d)}, \tag{59}$$

which we estimate in Section 8 below. The competition between these two terms leads to a transition at a given critical value of λ, which we determine in Section 9.

7. Large deviations of paths with correlated denominators

In the previous section we argued that the large deviations of operator amplitudes are essentially determined by the large deviations of effective path weights. The weight of any effective path is the products of two terms, describing the decay of $c_\alpha^\dagger$ and c_α, respectively. In each of those terms (cf. (52) e.g.), the functional dependence on the $\mathcal{E}_i$ is similar to that in the original path weights (44). We will first discuss the latter and then show that general effective paths behave essentially identically.

Because of the energy restrictions (7) the energy differences $\mathcal{E}_{\alpha\beta,\gamma\delta}/\delta_\xi$ are random variables of order $O(1)$. For simplicity, we take them as independent Gaussian random variables with zero mean and unit variance. The denominators in (44) are partial sums of such energies, and we may write:

$$|\omega_{\text{path}}| = \prod_{i=1}^{N-1} \frac{\lambda |\eta_{\alpha_i \beta_i, \gamma_i \delta_i}|}{|s_i|}, \tag{60}$$

where $s_i = (\mathcal{E}_1 + \cdots + \mathcal{E}_i)/\delta_\xi$, with $\mathcal{E}_i \equiv \mathcal{E}_{\alpha_i \beta_i, \gamma_i \delta_i}$.

In path weights of the form (60) we are mostly interested in characterizing the distribution of the product of denominators. The numerator behaves as $\sim (\lambda \eta_{\text{typ}})^{N-1}$, with $\eta_{\text{typ}} = \exp[\langle \log |\eta| \rangle] = 1/e$, and we neglect the Gaussian fluctuations of its logarithm.

The fact that the denominators in (60) are correlated distinguishes the many-body problem from single particle localization. These correlations are a feature that any perturbative treatment of MBL has to deal with, and it is thus important to develop a method to calculate the large deviations in this case.

The distribution function $P_N(y)$ of the logarithm of the product of denominators,

$$Y_N \equiv -\sum_{i=1}^{N} \log |s_i|, \tag{61}$$

can be obtained from its generating function,

$$G_N(k) \equiv \mathbb{E}\big[e^{-kY_N}\big], \tag{62}$$

by inverse Laplace transform,

$$P_N(y) = \frac{1}{2\pi i} \int_{\mathcal{B}} e^{yk} G_N(k) dk, \tag{63}$$

where $\mathcal{B}$ is the Bromwich path in the complex k-plane.

In the present case, the relevant y scales linearly with N, and thus we define $\tilde{y} = y/N$, and

$$P_N(N\tilde{y}) = \frac{1}{2\pi i} \int_{\mathcal{B}} e^{N\phi_N} dk, \tag{64}$$

where the function

$$\phi_N(\tilde{y}, k) = \tilde{y}k + \frac{\log G_N(k)}{N} \stackrel{N\to\infty}{\to} \phi(\tilde{y}, k) \tag{65}$$

has a well-defined limit, $\phi(\tilde{y}, k)$, for large N. In that limit, the integral over k can be done by a saddle point approximation. The contour has to be deformed to pass parallel to the imaginary axis through $k^* = k^*(\tilde{y})$, which satisfies:

$$\tilde{y} = -\frac{d}{dk}\left[\lim_{N\to\infty} \frac{\log G_N(k)}{N}\right]_{k=k^*(\tilde{y})}. \tag{66}$$

Large deviations correspond to $\tilde{y} = O(1)$. In the case of parametrically small interaction strength λ (which is relevant in the case of large connectivity $\mathcal{K}$) we will see that we can restrict our attention to $\tilde{y} \gg 1$, see Section 9. For large values of $\tilde{y}$, we will see that the saddle point tends to $k^* \to -1$.

The computation of the generating function G_N is given in Appendix D. Here it suffices to say that the recursive structure of the denominators s_i lends itself naturally to a transfer matrix expression for G_N, which grows as the Nth power of the largest eigenvalue.

The final result for the exponent at the saddle point is

$$\phi(\tilde{y}, k^*(\tilde{y})) = -\tilde{y} + \log\left(\frac{2e\tilde{y}}{\sqrt{2\pi}}\right) + \frac{\gamma}{2\tilde{y}} + O\left(\frac{1}{\tilde{y}^2}\right), \tag{67}$$

for $\tilde{y} \gg 1$. From this we obtain the large deviation probability:

$$P_N\left(N\tilde{y} = \log\left[\prod_{i=1}^{N} \frac{1}{|s_i|}\right]\right) = C(\tilde{y}, N)\left(\frac{2e}{\sqrt{2\pi}}\right)^N \tilde{y}^N e^{-N\mathcal{F}(\tilde{y})}, \tag{68}$$

where C contains only negligible logarithmic corrections to the exponent, and

$$\mathcal{F}(\tilde{y}) = \tilde{y} - \frac{\gamma}{2\tilde{y}} + O\left(\frac{1}{\tilde{y}^2}\right). \tag{69}$$

7.1. Comparison between correlated and uncorrelated denominators

It is interesting to compare the large deviation distribution (68) with the tails of the distribution of the random variable:

$$Y'_N \equiv -\sum_{i=1}^{N} \log|X_i| \tag{70}$$

where X_i are i.i.d. Gaussian random variables with zero mean and unit variance. As derived in Appendix D, at leading order in N, up to a correction $\mathcal{F} \to \mathcal{F} - \log 2/(2\tilde{y}) + O(1/\tilde{y}^2)$, both have the same form (68).

Physically, this result can be understood as follows. By restricting to $\tilde{y} \gg 1$, we are concentrating on very rare realizations of Y_N. Those are insensitive to the details in the structure of the denominators. Indeed, atypically big values of objects like $(\prod_{i=1}^{N} s_i)^{-1}$ arise from restraining the random walk $(s_1, \cdots, s_N)$ to the vicinity of the origin. This boils down to computing the probability that s_i is small *conditioned* on the fact that s_{i-1} was small. To leading order in the typical smallness of such denominators, one obtains the same result as by minimizing N denominators independently. The leading correction with respect to the case of i.i.d. denominators consists in a small suppression of the tail, since it is slightly less probable to encounter small denominators, when they are correlated.

The above reasoning can be extended to more general weights $\tilde{\omega}_\Gamma$, associated with effective paths. Indeed, the corresponding denominators are still products of single energies or partial sums (see Eq. (52) or Eq. (C.3)). In the limit of very large deviations ($\tilde{y} \gg 1$) they all share the same tail distribution (68), the only relevant parameter being the total number N of denominators. Therefore, approximating the numerator in $\tilde{\omega}_\Gamma$ with its typical value $(\lambda \eta_{\text{typ}})^N$ and using (68), we finally obtain:

$$\mathbb{P}\left(\frac{\log |\tilde{\omega}_\Gamma|}{N} = \tilde{x} + \log \lambda \eta_{\text{typ}}\right) \approx C(\tilde{x}, N) \left(\frac{2e}{\sqrt{2\pi}}\right)^N \tilde{x}^N e^{-N\mathcal{F}(\tilde{x})}, \tag{71}$$

with $\mathcal{F}$ given in (69).

8. Counting diagrams

8.1. Justification of neglecting interaction vertices with equal indices

We recall that we have neglected interaction terms $U_{\alpha\beta,\gamma\delta}$ where two or more indices are identical. This will significantly simplify the combinatorics of counting diagrams. Let us now give a justification a posteriori for this approximation, by showing that such terms would make contributions which are down by factors of $1/\mathcal{K}$. Consider the various scattering processes with one pair of equal indices among the four legs of a vertex, whereby we restrict to one ingoing and three out-going particles. Consider first the scattering $\alpha \to \beta$ with the simultaneous creation of a pair (γ, α). The constraints $|\epsilon_\alpha - \epsilon_\gamma| < \delta_\xi, |\epsilon_\alpha - \epsilon_\beta| < \delta_\xi$ imply that all levels have to lie within δ_ξ from each other. The phase space for such events is smaller by a factor of $1/\mathcal{K}$ with respect to generic scattering processes where γ is unrestricted.

The second case is more subtle. It consists in a scattering $\alpha \to \beta$ from a particle γ, which remains in place. If this is to be a resonant contribution one needs the energy increment $\Delta\mathcal{E}$ of the vertex to be $|\Delta\mathcal{E}| = |\epsilon_\alpha - \epsilon_\beta| \lesssim \delta_\xi/\mathcal{K}$. In a scattering where γ switches to a neighboring state δ, with $|\epsilon_\gamma - \epsilon_\delta| \sim \delta_\xi$, one can optimize α, β among the $\mathcal{K}$ different choices, such as to make $\Delta\mathcal{E}$ of order $\delta_\xi/\mathcal{K}$. However, if γ remains in place, the optimum over the $\mathcal{K}$ choices for α, β will yield a parametrically bigger $\Delta\mathcal{E} = \epsilon_\alpha - \epsilon_\beta$, because of the repulsion between the neighboring levels α, β. Therefore such processes are systematically much less resonant than processes involving four distinct levels.[9]

---

[9] BAA dropped such terms for a different reason, working directly with Hartree–Fock orbitals, which implicitly depend on the interaction strength and the initial state to be studied. The latter introduces a slight dependence of the localization

8.2. Combinatorics of diagrams

We now estimate the total number of diagrams $\mathcal{N}_N$ at a given order N, cf. Eq. (59). For simplicity, we restrict here to the case of spatial dimension $d = 1$.

Consider any amplitude $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$ with index set $(\mathcal{I}, \mathcal{J}) = (\alpha_1 \cdots \alpha_N, \beta_1 \cdots \beta_N)$. The localization centers $r_{\alpha_i}, r_{\beta_i}$, cf. Eq. (6), of the single particle indices are distributed over a certain number of localization volumina of length ξ around r_α, with a given number of single particle indices per localization length. Due to the energy restrictions imposed on the interactions, particles and holes belonging to the same localization volume are organized in pairs: members of a pair are produced in the same scattering process, and have an energy difference of order δ_ξ.

Due to the fact that the interaction is local, only particle–hole pairs in nearby localization volumina can be involved in the same interaction vertex: this imposes some constraints on the geometry of the diagrams representing the scattering processes with $(\mathcal{I}, \mathcal{J})$ as final state. For example, states $(\mathcal{I}, \mathcal{J})$ having only one particle–hole pair per localization length must be associated to diagrams with no branchings in the decays of $c_\alpha^\dagger$ and c_α, since the particle–hole pairs must be created in a fixed order dictated by their spatial sequence, and thus no permutation is possible. In contrast, final states with several pairs per localization length can be reached by a variety of diagrams.

In the following, we construct the subset of diagrams corresponding to scattering processes with a "necklace structure", in which the particle–hole pairs are created in a sequence of n groups of $m_{i=1,\dots,n}$ pairs, each group belonging to a single localization volume. This furnishes a lower bound on the number of all diagrams. Note that m_i is bounded by the maximal number of particle–hole pairs per localization volume ($N_{\text{loc}} = \mathcal{K}/4$), and $\sum_{i=1}^{n} m_i = N$. Due to locality, pairs belonging to the ith and $(i + 1)$th group belong to neighboring localization volumina in real space; pairs belonging to different groups $i, j \neq \{i - 1, i + 1\}$ might belong to the same localization volume.

This construction is done in two steps: first, for every group i we build all possible sub-diagrams with final indices corresponding to the indices of the m_i pairs, as illustrated in Fig. 5. In a second step, we connect sub-diagrams of neighboring groups by a single scattering vertex. We thus obtain a global necklace diagram, and count how many different diagrams with this structure there are. The counting is similar to Ref. [47], but here we include diagrams corresponding to final states with a non-uniform density of particle and hole indices per localization length, since these have a larger abundance.

A central ingredient for the combinatorics is the number of all possible geometries of diagrams with m interactions in a given localization volume, see Fig. 5. We denote this number by $\mathcal{T}_m$. It equals the number of trees with one root (of connectivity 2) and m nodes (of connectivity 4). As we derive in Appendix E:

$$\mathcal{T}_m = \frac{3^{\frac{3}{2}+3m}}{\pi} \frac{\Gamma\left(m + \frac{2}{3}\right)\Gamma\left(m + \frac{4}{3}\right)}{\Gamma(2m + 3)} \sim \frac{3}{4}\sqrt{\frac{3}{\pi}} \frac{1}{m^{\frac{3}{2}}} \left(\frac{27}{4}\right)^m. \tag{72}$$

---

length ξ, and hence of the level spacing δ_ξ, on λ. This in turn might induce a small shift of λ_c. However, since their subsequent analysis boils down to dropping the same terms as we have argued above, this shift is expected to be a $1/\mathcal{K}$ correction.
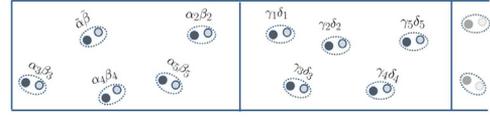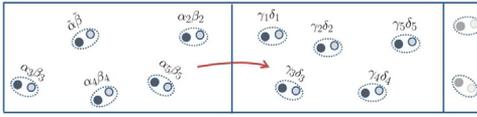
(a) The m_1 pairs of the first group belong to the localization volume containing the localization center r_α, with one pair ($\tilde{\alpha}\tilde{\beta}$ in the Figure) with energies close to ϵ_α. The remaining m_1 pairs ($m_1 - 1$ in the same volume and one in the adjacent volume) are produced in m_1 scatterings organized in diagrams with all possible geometries ($\mathcal{T}_{m_1}$ of them). For each geometry, a factor m_1 comes from the choice of the vertex (red in the Figure) that produces a pair in the subsequent localization volume.

(b) The indices of the remaining $m_1 - 1$ pairs in the localization volume are assigned to the legs. This fixes all labels up to the internal legs. For the first localization volume, the possible internal indices satisfying the energy restrictions are 2^{m_1-2}, α being fixed.
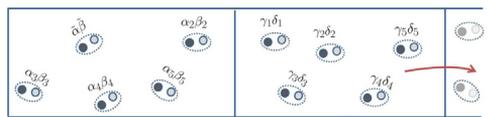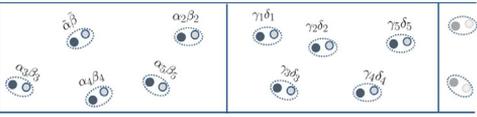
(c) For a fixed geometry and a set of final labels, the permutation of assignments of labels to the legs gives rise to an independent diagram with the same final state, since the matrix elements of the interactions change. There are $(m_1 - 1)!$ such permutations (the pair $\tilde{\alpha}\tilde{\beta}$ is fixed). When legs are permuted, the corresponding internal indices must change as well in order to satisfy the energy restriction of the interactions (in the Figure, $\tilde{\alpha}_3^{(i)} \to \tilde{\alpha}_5^{(i)}$ and $\tilde{\beta}_3^{(i)} \to \tilde{\beta}_5^{(i)}$).

(d) Diagrams corresponding to the decay processes in the adjacent localization volume are attached to a pair of legs selected in (a). Again, m_2 interactions occur on two branches, in a total of $\mathcal{T}_{m_2}$ possible distinct geometries. There are m_2 choices to select the pair of legs to which to attach the next subdiagram. For any of the $m_2!$ labelings of the remaining external legs, there are two choices for each internal index, corresponding to whether the incoming particle scatters up or down in energy. In total there are $m_2 2^{m_2} m_2! \mathcal{T}_{m_2}$ different diagrams associated to this group of pairs. The same counting holds true for the subsequent groups.

Fig. 5. Construction of the diagrams representing the decay of groups of m_i particle–hole pairs, where members of the same group belong to the same localization volume. The diagrams are constructed by connecting sub-diagrams describing the decay of each single group of pairs. We restrict the combinatorics to only one scattering vertex connecting the sub-diagrams of different groups. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Following the reasonings explained in Fig. 5, we find the number of necklace diagrams associated with fixed groups of m_i pairs to be

$$n_{\text{neck}} = \prod_{i=1}^{n} \left[m_i 2^{m_i} m_i! \mathcal{T}_{m_i} \right]. \tag{73}$$

The origin of the various factors is explained in detail in Fig. 5: the factor m_i counts the number of pairs which are created subsequently to the first pair entering the volume associated to the group i. One of those m_i pairs belongs to the adjacent localization volume and creates the subsequent cascade of pair creations there. The second factor describes the choice of two levels (the level closest in energy above or below) to which an incoming quasiparticle may scatter at a vertex. The factorial term comes from the choice of assigning the m_i pairs to the final legs of a given tree diagram in the localization volume of group i.

Consider first the case in which only a single group i of pairs occupies a given localization volume. The number of choices of $\{m_i\}$ particle–hole pairs is then given by

$$n_s\left(\{m_i\}, \mathcal{K}\right) \equiv \prod_{i=1}^{n} 2^{m_i} \binom{N_{\text{loc}} - m_i}{m_i} = \prod_{i=1}^{n} \left[2^{m_i} \binom{\mathcal{K}/4 - m_i}{m_i} \right]. \tag{74}$$

Indeed, a configuration of m_i pairs of (disjoint) adjacent levels, and the remaining $N_{\text{loc}} - 2m_i$ untouched levels in the same localization volume form a set of $N_{\text{loc}} - m_i$ objects, out of which m_i are pairs. This explains the binomial factor. For each pair, one can choose how to assign the two levels to particle and hole, respectively. This yields the factor 2^{m_i}.

As we will see below, the relevant m_i are of order $O(1) \ll \mathcal{K}$. We therefore approximate:

$$\binom{\mathcal{K}/4 - m_i}{m_i} \approx \frac{(\mathcal{K}/4)^{m_i}}{m_i!}. \tag{75}$$

Note that the necklace structure will in general fold back and forth in real space, such that several groups will get to lie in the same volume. Nevertheless, the above approximation remains good as long as the total number of pairs created in a given localization volume is significantly smaller than $\mathcal{K}$.

Combining Eqs. (73)–(75), the total number of necklace diagrams is:

$$\mathcal{N}_N \approx \overline{\mathcal{P}(d)} \sum_{\{m_i\} | \sum_i m_i = N} \frac{1}{2} \prod_{i=1}^{n} \left[2\mathcal{K}^{m_i} m_i \mathcal{T}_{m_i} \right], \tag{76}$$

where the average number of effective paths per diagram, $\overline{\mathcal{P}(d)}$, scales sub-exponentially with N. The factors of 2 arise due to freedom of each group to scatter to the left or the right of the preceding group as long as there is still significant phase space in the corresponding localization volumina. The correction due to the finiteness of $\mathcal{K} \gg 1$ is small and was thus neglected.

We now determine the distribution of group sizes $\{m_i\}$ which dominates the sum (76), writing

$$\mathcal{N}_N = \frac{1}{2} \sum_{\{m_i\} | \sum_i m_i = N} \prod_i 2\mathcal{K}^{m_i} m_i \mathcal{T}_{m_i} = \frac{\mathcal{K}^N}{2} \sum_{\{m_i\} | \sum_i m_i = N} \prod_i 2 m_i \mathcal{T}_{m_i}$$

$$= \frac{\mathcal{K}^N}{2} \sum_{\{n_m\} | \sum_m m n_m = N} \binom{\sum_m n_m}{n_1, n_2, ..., n_m} \prod_m (2m \mathcal{T}_m)^{n_m}, \tag{77}$$
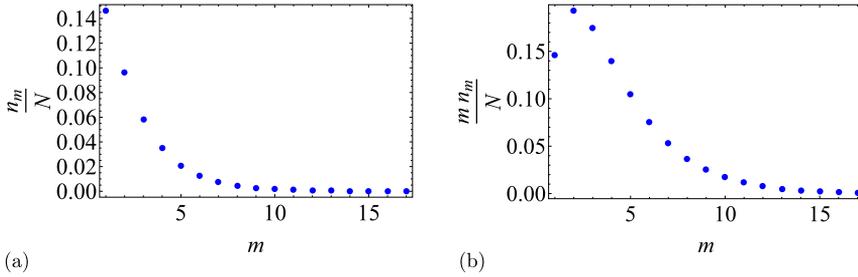
Fig. 6. The plot (a) shows the distribution of the number n_m/N of groups of m particle–hole pairs in necklace diagrams dominating $\mathcal{N}_N$. The plot (b) shows the probability $m n_m/N$ that a given pair belongs to a group containing m pairs.

where $n_m = \sum_i \delta_{m,m_i}$ is the number of groups i with m pairs. For the relevant m's, $n_m \sim N \gg 1$; therefore, at large N the sum (77) is dominated by the saddle point over the n_m. Imposing the constraint $\sum_m m n_m = N$ with a Lagrange multiplier μ yields the saddle point equations:

$$\mu m = -\log(n_m) + \log\left(\sum_m n_m\right) + \log(2m\mathcal{T}_m), \tag{78}$$

and thus

$$\frac{n_m}{\sum_{m'} n_{m'}} = 2m\mathcal{T}_m e^{-\mu m}. \tag{79}$$

The Lagrange multiplier μ is fixed by the constraint:

$$1 = \sum_m 2m\mathcal{T}_m e^{-\mu m} = -2\frac{d}{d\mu}\left[\mathcal{T}\left(x = e^{-\mu}\right)\right], \tag{80}$$

with $\mathcal{T}(x) = \sum_m \mathcal{T}_m x^m$. As discussed in Appendix E, $\mathcal{T}(x) = [T(x)]^2$, where $T(x)$ is the generating function of 3-branched trees satisfying $T(x) = 1 + xT^3(x)$. The solution of Eq. (80) is:

$$e^{-\mu} = 0.0941. \tag{81}$$

The saddle point solution can thus be written as

$$\frac{n_m}{N} = Am\mathcal{T}_m e^{-\mu m}, \tag{82}$$

where $1/A = d^2/d\mu^2[T(x = e^{-\mu})^2] = 1/0.778$, as follows from the constraint $\sum_m m n_m = N$. The resulting values for n_m/N are shown in Fig. 6(a). The probability that a given pair is created in a scattering process involving a total of m pairs in the same localization volume is plotted in Fig. 6(b). We see that most pairs are created together with a few more pairs within the same localization volume.

Plugging (82) into the saddle point for $\mathcal{N}_N$, we find the number of diagrams to grow like (dropping pre-exponential factors)

$$\mathcal{N}_N \approx \left(\mathcal{K}e^{\mu}\right)^N \approx (10.6\mathcal{K})^N. \tag{83}$$

This result is based on the approximation that we only allow for diagrams with a necklace structure, where groups of m_i pairs are connected by a single scattering between subsequent localization volumina. Performing the calculation without this restriction is difficult since it is

less easy to control the spatial constraints. However, we can easily obtain an upper bound by realizing that all possible diagrams consist in all geometrically possible labellings of trees of size N. The number of trees grows as $(27/4)^N$. For each label one has roughly $3\mathcal{K}$ possibilities, as the pair must lie in a localization volume adjacent to or identical with the one of the pair preceding it on the tree. This yields the simple upper bound

$$\mathcal{N}_N < \left(3 \cdot \frac{27}{4} \mathcal{K} \right)^N \approx (20.25\mathcal{K})^N, \tag{84}$$

which yields a growth factor which is only about a factor of 2 bigger than the much more conservative estimate (83). Let us thus write

$$\mathcal{N}_N \approx (C\mathcal{K})^N, \tag{85}$$

with

$$10.6 < C < 20.25. \tag{86}$$

8.3. Effect of Fermi blocking

The above counting is still not entirely complete. Indeed, eventually the operators we have constructed should act on some many body states, and get annihilated when attempting to create particles on occupied states or holes on already empty states. In an infinite temperature state, and at a filling fraction ν each particle–hole creation operator has a chance to annihilate the state with probability $1 - \nu(1 - \nu)$, or, in other words, only a fraction of $[\nu(1 - \nu)]^N$ of all operators will not annihilate a typical infinite temperature state. One should thus modify the number of relevant diagrams to

$$\mathcal{N}_N \to \mathcal{N}_N \simeq \big(C\nu(1 - \nu)\mathcal{K} \big)^N. \tag{87}$$

In the next section we use this result to determine the radius of convergence of the operator series. Similar considerations apply to finite temperature as we will discuss below.

8.4. Structure of the dominant operator terms

Our result differs from the similar analysis in Ref. [47]. The main difference consists in our assumption that the sum of diagrams that add up to the amplitude of a given operator $O_{\mathcal{I},\mathcal{J}}$ is dominated by the biggest term (provided the considered amplitude is among the largest ones at that order). In contrast, the authors of [47] assumed that the exponentially many diagrams have comparable amplitudes, but random signs, and applied the central limit theorem to the sum. Moreover, we allow for fluctuations of the number of pairs generated in each localization volume instead of imposing a homogeneous spatial density. We find that in the restricted set of necklace diagrams the optimal distribution of group sizes m_is is peaked at values of order $O(1)$, but still clearly larger than one. Upon folding of the necklace, the number of pairs per localization volume will become even more significantly larger than 1. Thus we see that multiple scattering processes within a localization volume significantly enhance the delocalization tendency. This shows that the many-body problem is genuinely different from an effective one-body problem, in which a simple excitation would propagate nearly ballistically, by shedding one particle–hole excitation in every localization volume.

9. Estimate of the radius of convergence

We now have all the ingredients to estimate the probability of resonances at generation N, in order to prove that for λ sufficiently small there are no delocalizing resonances and (46) holds true.

Consider the probability in expression (58). Using (71), we estimate:

$$
\mathbb{P}\big(|\tilde{\omega}_\Gamma| > z^N\big) \approx \left(\frac{2e}{\sqrt{2\pi}}\right)^N \int_{\log\left(\frac{z}{\lambda\eta_{\mathrm{typ}}}\right)}^{\infty} C(\tilde{x}, N)\tilde{x}^N e^{-N\mathcal{F}(\tilde{x})} d\tilde{x}. \tag{88}
$$

Note that the large deviation result applies since $\tilde{x} \geq \log(\frac{z}{\lambda\eta_{\mathrm{typ}}}) \gg 1$. Approximating the integral with the value of the integrand at the extremum, setting $z = 1$ and neglecting sub-exponential terms in N we obtain:

$$
\mathbb{P}\big(|\tilde{\omega}_\Gamma| > 1\big) \approx \left(\frac{2e}{\sqrt{2\pi}}\log\left(\frac{1}{\lambda\eta_{\mathrm{typ}}}\right)\right)^N e^{-N\left[\log\left(\frac{1}{\lambda\eta_{\mathrm{typ}}}\right)+O\left(1/\log\left(\frac{1}{\lambda\eta_{\mathrm{typ}}}\right)\right)\right]}. \tag{89}
$$

Substitution of (89) and (87) into (58) yields:

$$
\sum_{\substack{\mathcal{I}\neq\mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} \mathbb{P}\big(|\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}| > 1\big) \simeq \exp\big[N\log\mathcal{G}(\lambda,\mathcal{K}) + o(N)\big], \tag{90}
$$

with

$$
\mathcal{G}(\lambda,\mathcal{K}) = \nu(1-\nu)\frac{2e\,C\eta_{\mathrm{typ}}}{\sqrt{2\pi}}\lambda\mathcal{K}\log\left(\frac{1}{\lambda\eta_{\mathrm{typ}}}\right). \tag{91}
$$

Taking into account (56) and (57), we finally obtain:

$$
\mathbb{P}\left(\forall N > N^*, \sum_{\substack{\mathcal{I}\neq\mathcal{J} \\ |\mathcal{I}|=N+1=|\mathcal{J}|}} |\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}| < 1\right) = \prod_{N>N^*} \exp\big[-e^{N\log\mathcal{G}(\lambda,\mathcal{K})+o(N)}\big]. \tag{92}
$$

If $\mathcal{G}(\lambda,\mathcal{K}) < 1$, then, for N^* sufficiently big, each of the factors in (92) is arbitrarily close to 1. Therefore, their product converges to 1 in the limit $N^* \to \infty$ (see also [48] for a similar reasoning). This allows us to conclude that, for all values of λ for which $\mathcal{G}(\lambda,\mathcal{K}) < 1$ holds, (46) holds, too, and the series in operator space (28) converges to a quasi-local operator. In this regime, the excitation of the single particle level α, localized at $\vec{r}_\alpha$, is very unlikely to create a distant disturbance at $\vec{r}_\beta$ with large $L = |\vec{r}_\beta - \vec{r}_\alpha|$, its probability tending to zero exponentially as $L \to \infty$: there is no diffusion at small λ.

The critical value for λ is given by $\mathcal{G}(\lambda_c,\mathcal{K}) = 1$. For large $\mathcal{K}$, it equals to:

$$
\lambda_c = \frac{\sqrt{2\pi}}{C\nu(1-\nu)2e}\frac{1}{\mathcal{K}\log\mathcal{K}}, \tag{93}
$$

where we used $\eta_{\mathrm{typ}} = 1/e$.

9.1. Comparison with a single particle on the Bethe lattice

It is interesting to note that the delocalization threshold (93) looks identical to the critical ratio between hopping and disorder strength for a single particle problem on a Bethe lattice (see Eq. (5.8) in [45]) with effective connectivity $\mathcal{K}_{\text{eff}} = \nu(1-\nu)(C/\sqrt{2\pi})\mathcal{K}$, which is a significantly bigger than the connectivity associated with each vertex, $\nu(1-\nu)\mathcal{K}$. This reflects the fact that in the many-body problem the same final state can be reached with many different decay processes. The results are nevertheless similar, because both problems are dominated by very few resonant paths, whereby the large local connectivity in the many-body problem ensures that different resonant paths are likely to be uncorrelated, even if they lead to the same final state.

9.2. Possible implications for delocalization in higher dimensions

According to the above calculation, in the dominating decaying processes only groups of $\mathcal{O}(1)$ particle–hole pairs are created at the same time in a localization volume. This suggests that the necklace-type diagrams are diffusing back and forth a lot. This contrasts with the model of BAA, where the hopping strength between adjacent volumina was assumed to be parametrically smaller than λ, which favored the particle–hole creation cascade to fully explore a localization volume before moving on to the next volume. The latter led them to conjecture a critical exponent for the localization length in higher dimensions by relating the decay processes of single particle excitations to self-avoiding random walks. This scenario hardly holds in our model, as the optimal processes are not of this kind.

10. Finite temperature

So far we have been discussing the convergence of the expansion of integrals of motion in the forward approximation. If the expansion converges, we have succeeded in constructing a complete set of quasi-local conserved quantities which entail the absence of transport in whatever state the system is, in particular at any temperature, including the limit $T \to \infty$. Note again, that the latter limit is meaningful because we work on a lattice on which the energy density is bounded.

An interesting question arises when we ask about transport at finite temperature, and the possibility of a MBL transition as a function of temperature, as predicted by BAA. How would this reflect at the level of integrals of motion? If there is a finite temperature transition, one expects that the localized low T phase is still governed by local conservation laws which inhibit transport, while local integrals of motion do not exist at higher temperature. Clearly the latter rules out the convergence of the conserved operators in the operator norm. Rather one has to invoke that the norm of operators $\mathcal{O}_{\mathcal{I},\mathcal{J}}$, when restricted to typical low temperature states, becomes exponentially small in $N = |\mathcal{I}|$, if the index sets $\mathcal{I}, \mathcal{J}$ contain a finite fraction of hole excitations above $E_F + T$ or particle excitations below $E_F - T$. This effect may enhance the convergence of the series expansion. This is certainly so at the level of the forward approximation where the temperature T essentially replaces the bandwidth in the analytical estimates of our expansion. This will lead to a larger domain of (weak) convergence of the operator expansion, suggesting the possibility of a delocalization transition at finite temperature.

A similar consideration shows that the transition (93) at $T = \infty$ takes place in a regime where the operator expansion is not convergent in the operator norm, but converges only weakly on typical high energy states. This is due to the Fermi blocking discussed in Section 8.3.
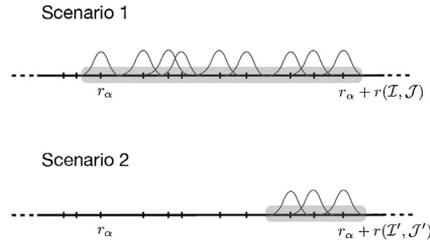
Fig. 7. Pictorial representation of the supports of two different operators $\mathcal{O}_{\mathcal{I},\mathcal{J}}$ and $\mathcal{O}_{\mathcal{I}',\mathcal{J}'}$ contributing to the series expansion (28). In the pictures, the wave-functions are the single particle states contributing to $(\mathcal{I},\mathcal{J})$ and $(\mathcal{I}',\mathcal{J}')$. Both operators involve degrees of freedom whose maximal distance to the localization center r_α is the same: $r(\mathcal{I},\mathcal{J}) = r(\mathcal{I}',\mathcal{J}')$; however, the length of the support N of the operators (shaded in the picture) increases when N grows in the first case, while it remains bounded in the second case.

However, a different scenario is possible as well. The operator series in Eq. (28), or subsequences of it, can diverge for two reasons: (i) Either the amplitude of terms with growing N do not decrease sufficiently fast, and thus the diameter of the support of these terms grows indefinitely. (ii) There can be subsequences of (28) whose terms have bounded index level N, but supports which wander off to infinity. These two possibilities are illustrated in Fig. 7.

Possibility (i) is what is obtained within the forward approximation. The fraction of terms at λ_c, which survive when applied on finite T states, decreases rapidly with N. However, such a projector would not affect the convergence properties of a subsequence of type (ii). Upon restricting to finite T states the norm of the relevant operators is typically reduced by a factor, which remains bounded from below. Therefore the series will continue to diverge despite the projection.

To address the question of whether or not a finite temperature transition is possible one has to consider the interaction strength λ_c at which the infinite T transition takes place, i.e., where the integrals of motion delocalize. If there is a subsequence of type (ii), which diverges at this point, the delocalization transition is a function of λ only, but independent of T. In the delocalized phase $(\lambda > \lambda_c)$ transport would always remain finite, even though it may become very inefficient and strongly activated at low T. If instead there is no subsequence with bounded index cardinality, which diverges at λ_c, a transition in temperature should be expected, as predicted by BAA. Such a transition was recently reported by a numerical study [17].

Physically the scenario (ii) corresponds to transport and delocalization driven by rare, compact, but mobile regions with a local "temperature" above the putative T_c. At first sight one is tempted to rule this out because one would expect such a hot bubble to diffuse and loose its extra energy forever to the environment. However, the environment being in the supposed MBL phase cannot transport the extra energy to infinity, and thus there should be a finite recurrence time until the hot bubble forms again. Whether such a bubble would nevertheless have to remain localized, or whether its internally delocalized state would allow it to move around is a difficult open question. Recently, it was argued that big enough bubbles could undergo resonant delocalization [49]. At the level of integrals of motion these two scenarii translate into the above dichotomy about critical subsequences.

Note that a divergence of type (ii) by a set of operators with bounded support is made less likely by the large parameter $\mathcal{K}$. We in fact invoked this large parameter to neglect these terms, similarly as BAA. However, it is difficult to exclude that there is no such divergent subsequence which contributes with a finite, but with a relative weight which is parametrically small in $\mathcal{K}$. In that case, numerical approaches such as [17,50] would not capture this divergence.

It would be interesting to revisit the question of the finite T transition also as a function of density. In the low density limit, the effective connectivity $\mathcal{K}_{\text{eff}}$ (resulting from projection onto typical states) can be reduced to $K_{\text{eff}} \ll 1$, in which propagation channels of type (ii) become parametrically favorable, and may be the ones to induce delocalization – if interactions can induce a transition at all under such circumstances.

11. Conclusion

In this work we have constructed explicit quasi-local integrals of motion within the weakly interacting regime, which we argued to imply the absence of any d.c. transport. We reduced the problem of constructing such operators to a non-Hermitian hopping problem in operator space, an idea that we hope to have potential for further more rigorous studies. We have also obtained an explicit recipe for constructing generalized occupation numbers of a Fermi insulator order by order in perturbation theory.

We have used the large parameter $\mathcal{K}$ (proportional to the number of sites in a single particle localization volume) to concentrate on processes where one more particle–hole pair is created at every order of perturbation theory. Within this forward approximation, and based on an analysis of rare resonances at large distance, we found an analytical estimate of the radius of convergence of this perturbative construction, yielding a critical value of the reduced interaction strength $\lambda_c = \sqrt{2\pi}/(C\nu(1-\nu)2e\,\mathcal{K}\log\mathcal{K})$ with $10.6 < C < 20.25$, at infinite T and filling fraction ν, similar to the prediction by BAA based on the analysis of the life time of a single particle injection.

We believe that the spatial structure of our integrals of motion provides a good picture for the "quantum avalanche" created by injection of an extra particle. We have found that the optimal way of its propagation is by exciting a necklace of groups of $O(1)$ particle–hole pairs per localization volume. Due to the meandering of the necklace structure, several groups of such pairs may be created in the same localization volume, an effect which is enhanced in low dimensions.

The convergence of our construction for the local integrals of motion implies the absence of transport and equilibration at any temperature and density. Taken as such, it appears to be blind to potential phase transitions upon varying those parameters. However, projecting the operator series onto typical states with thermal single particle occupations, one may discuss the *weak* convergence of the operator expansion. In this vein, we have discussed the question of the existence of a genuine finite temperature transition, depending on the properties of the operator series at its critical point at $T = \infty$. Further investigations of this question would be interesting.

Acknowledgements

Appendix A. Imposing binary spectrum

In this appendix we show how one can modify, order by order in λ, the previously obtained integrals of motion in order to fix their spectrum to be that of occupation numbers, i.e., $\{0, 1\}$. This is equivalent to the condition:

$$I_\alpha^2 = I_\alpha. \tag{A.1}$$

This procedure leads to a modified expansion for I_α:

$$I_\alpha = n_\alpha + \sum_{m \geq 1} \lambda^m \Delta B_\alpha^{(m)}, \tag{A.2}$$

with $\Delta B_\alpha^{(m)}$ given explicitly in Eqs. (26), (27).

In the following, we work by induction on m. We set $\Delta B_\alpha^{(0)} = n_\alpha$ and we omit the index α for simplicity. We define the truncation to mth order of I:

$$I^{\leq m} \equiv n + \sum_{i=1}^{m} \lambda^i \Delta B^{(i)}, \tag{A.3}$$

and assume that the property (A.1) holds to order $O(\lambda^{m-1})$, namely:

$$\left(I^{\leq m-1}\right)^2 = I^{\leq m-1} + o\left(\lambda^{m-1}\right). \tag{A.4}$$

Note that $I^{\leq 0}$ is naturally binary, with $(I^{\leq 0})^2 = I^{\leq 0}$.

We denote with $\Delta \hat{I}^{(m)}$ the solution of the equation:

$$\left[H_0, \Delta \hat{I}^{(m)}\right] + \left[U, \Delta B^{(m-1)}\right] = 0 \tag{A.5}$$

in the subspace O, cf. Eq. (25), and define

$$\hat{I}^{\leq m} \equiv I^{\leq m-1} + \lambda^m \Delta \hat{I}^{(m)}. \tag{A.6}$$

The operator $\hat{I}^{\leq m}$ is not binary to order $O(\lambda^m)$; however, we show that it is possible to add to $\Delta \hat{I}^{(m)}$ a suitably chosen operator $\Delta K^{(m)}$ in the kernel K of the linear map $f(X) = [H_0, X]$, so that

$$I^{\leq m} = \hat{I}^{\leq m} + \lambda^m \Delta K^{(m)} \equiv I^{\leq m} + \lambda^m \Delta B^{(m)} \tag{A.7}$$

is binary to order $O(\lambda^m)$. To show this, it is sufficient to show that the difference $(\hat{I}^{\leq m})^2 - \hat{I}^{\leq m}$, truncated to order $O(\lambda^m)$, is an element of the subspace K, i.e.:

$$\left[H_0, \left(\hat{I}^{\leq m}\right)^2\right] = \left[H_0, \hat{I}^{\leq m}\right] + o\left(\lambda^m\right). \tag{A.8}$$

This holds, since:

$$\left(I^{\leq m}\right)^2 = \left[\left(I^{\leq m-1}\right)^2\right]_{m-1} + \lambda^m \sum_{a=0}^{m} \Delta B^{(a)} \Delta B^{(m-a)} + o\left(\lambda^m\right)$$

$$= I^{\leq m-1} + \lambda^m \sum_{a=0}^{m} \Delta B^{(a)} \Delta B^{(m-a)} + o\left(\lambda^m\right), \tag{A.9}$$

where $[X]_{m-1}$ denotes the restriction of the Taylor series of $X(\lambda)$ to terms up to order λ^{m-1}. Using the inductive step $m - 1$ we have from (A.9)

$$\left[H_0, \left(\hat{I}^{\leq m}\right)^2\right] = \left[H_0, I^{\leq m-1}\right] + \lambda^m \left[H_0, \sum_{a=0}^{m} \Delta B^{(a)} \Delta B^{(m-a)}\right] + o\left(\lambda^m\right), \tag{A.10}$$

where in the terms with $a = 0, m$ we have replaced $\Delta \hat{I}^{(m)}$ with $\Delta B^{(m)}$, since Eq. (A.10) does not depend on the choice of $\Delta K^{(m)}$. Given that

$$\left[H_0, \Delta B^{(a)} \Delta B^{(m-a)}\right] = \Delta B^{(a)}\left[H_0, \Delta B^{(m-a)}\right] + \left[H_0, \Delta B^{(a)}\right]\Delta B^{(m-a)}$$
$$= -\Delta B^{(a)}\left[U, \Delta B^{(m-a-1)}\right] - \left[U, \Delta B^{(a-1)}\right]\Delta B^{(m-a)},$$

summing over a we get

$$\left[H_0, \sum_{a=0}^{m} \Delta B^{(a)} \Delta B^{(m-a)}\right] = -\left[U, \sum_{a=0}^{m-1} \Delta B^{(a)} \Delta B^{(m-a)}\right]. \tag{A.11}$$

Using that (A.1) at the inductive step $m-1$ implies

$$\sum_{a=0}^{m-1} \Delta B^{(a)} \Delta B^{(m-a)} = \Delta B^{(m-1)}, \tag{A.12}$$

and using (A.5), we find

$$\left[H_0, \left(\hat{I}^{\leq m}\right)^2\right] = \left[H_0, I^{\leq m-1}\right] + \lambda^m \left[H_0, \Delta \hat{I}^{(m)}\right] + o\left(\lambda^m\right)$$
$$= \left[H_0, \hat{I}^{\leq m}\right] + o\left(\lambda^m\right), \tag{A.13}$$

which proves (A.8).

A simple computation shows that by choosing:

$$I^{\leq m} \equiv \hat{I}^{\leq m} + \lambda^m \left(1 - 2\Delta B^{(0)}\right)\left[\left(\hat{I}^{\leq m}\right)^2 - \hat{I}^{\leq m}\right]_m \tag{A.14}$$

the condition (A.4) is fulfilled to order $O(\lambda^m)$. Eq. (27) follows from noticing that:

$$\left[\left(\hat{I}^{\leq m}\right)^2 - \hat{I}^{\leq m}\right]_m = \sum_{i=1}^{m-1} \Delta B^{(i)} \Delta B^{(m-i)} + \left\{\Delta B^{(0)} - \frac{1}{2}, \Delta \hat{I}^{(m)}\right\}. \tag{A.15}$$

Appendix B. Local re-summation in the case of small denominators

In the following we present a simple example in which the perturbative expansion in λ, Eq. (22), diverges. Suppose that at order n the series expansion contains the term:

$$J_n \equiv \mathcal{J}_n O c_\alpha, \tag{B.1}$$

where $O = c_{i_1}^\dagger \cdots c_{i_m}^\dagger c_{j_1} \cdots c_{j_{m-1}}$ is a string of operators with $i, j \neq \{\alpha, \beta, \gamma, \delta\}$, and that the amplitude $\mathcal{J}_n = O(\lambda^n)$ therefore contains the energy denominator:

$$\mathcal{J}_n \propto \left(\sum_{k=1}^{m} \epsilon_{i_k} - \sum_{k=1}^{m-1} \epsilon_{j_k} - \epsilon_\alpha\right)^{-1} \equiv (\Delta\mathcal{E})^{-1}. \tag{B.2}$$

Suppose $\Delta\mathcal{E}$ to be atypically small. One then easily finds a subsequence of the series (22), which contains arbitrarily high powers of the small denominator. Indeed, let us restrict the interaction to the term $U_{\alpha\beta,\gamma\delta}(c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta + h.c.)$ in the interaction U; higher order terms in the perturbative expansion are obtained by subsequent application of (25) to J_n; this produces:

$$J_{n+1} \equiv \mathcal{J}_n \frac{U_{\alpha\beta,\gamma\delta}}{\Delta\mathcal{E} + \mathcal{E}_{\alpha\beta,\gamma\delta}} O c_\beta^\dagger c_\gamma c_\delta \equiv \mathcal{J}_{n+1} O c_\beta^\dagger c_\gamma c_\delta,$$

$$J_{n+2} \equiv -\mathcal{J}_{n+1} \frac{U_{\alpha\beta,\gamma\delta}}{\Delta\mathcal{E}} O\left(n_\beta(1 - n_\gamma)(1 - n_\delta) + (1 - n_\beta)n_\gamma n_\delta\right)c_\alpha,$$

$$J_{n+3} \equiv \mathcal{J}_{n+1}\left[\frac{U_{\alpha\beta,\gamma\delta}}{\Delta\mathcal{E}} \frac{U_{\alpha\beta,\gamma\delta}}{\Delta\mathcal{E} + \mathcal{E}_{\alpha\beta,\gamma\delta}}\right] O c_\beta^\dagger c_\gamma c_\delta, \tag{B.3}$$
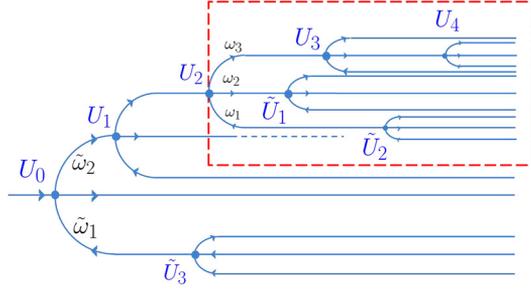
Fig. C.8. A diagram with multiple branchings. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

with $\mathcal{E}_{\alpha\beta,\gamma\delta} = \epsilon_\alpha + \epsilon_\beta - \epsilon_\gamma - \epsilon_\delta$. By iteration of this procedure, a sub-sequence of operators containing arbitrarily high powers of $(\Delta\mathcal{E})^{-1}$ is generated, preventing the convergence of the series if the term in brackets is larger than 1.

Divergences of this kind are of the same nature as local resonances encountered in single particle localization [1]. They have to be properly re-summed for the series expansion to make sense. For example, all terms multiplying $Oc^\dagger_\beta c_\gamma c_\delta$ re-sum into a self-energy correction of the denominator in the first line of (B.3):

$$J \equiv \mathcal{J}_n \left[\frac{U_{\alpha\beta,\gamma\delta}}{(\Delta\mathcal{E} + \mathcal{E}_{\alpha\beta,\gamma\delta}) - \frac{U^2_{\alpha\beta,\gamma\delta}}{\Delta\mathcal{E}}} \right] Oc^\dagger_\beta c_\gamma c_\delta \equiv \mathcal{J} Oc^\dagger_\beta c_\gamma c_\delta. \tag{B.4}$$

The term in square brackets in (B.4) contains a very large self energy correction $U^2_{\alpha\beta,\gamma\delta}/\Delta\mathcal{E}$, which compensates the divergence in $\mathcal{J}_n$ when $\Delta\mathcal{E} \to 0$.

Self-energy corrections like this are neglected in the forward approximation. Their main effect is to weaken the role of small denominators: As noticed by Anderson, small denominators essentially neutralize themselves by introducing enormous self-energies for the neighboring sites which then appear as very large denominators [1]. The resummation thus increases the convergence as compared to the naive perturbative expansion in forward approximation in [1]. In single particle localization problems with large connectivity, the critical hopping is increased by a factor $e/2$ [45], and a similar effect is expected here [4].

Appendix C. Evaluating diagrams as sums over effective paths: a more involved example

As an additional example for the evaluation of diagrams as sums over effective paths, we give the explicit expression for effective path weights associated to diagrams with the geometry of Fig. C.8.

For fixed indices on all segments, there are 105 different orderings of the interactions.[10] Their sum has the integral representation:

$$I_0(\{U\}) = \lim_{\epsilon \to 0} \int \frac{d\tilde{\omega}_1 d\tilde{\omega}_2 \delta(\tilde{\omega}_1 + \tilde{\omega}_2 - \mathcal{E}_0)}{\tilde{\omega}_1^-(\tilde{\omega}_1^- + \tilde{\mathcal{E}}_3)\tilde{\omega}_2^-(\tilde{\omega}_2^- + \mathcal{E}_1)} I_1(\omega = \tilde{\omega}_2 + \mathcal{E}_1 + \mathcal{E}_2), \tag{C.1}$$

---

[10] The interactions in the red dashed frame can be ordered in 15 different ways, for each of which the interaction $\tilde{U}_3$ can be placed in 7 different positions.
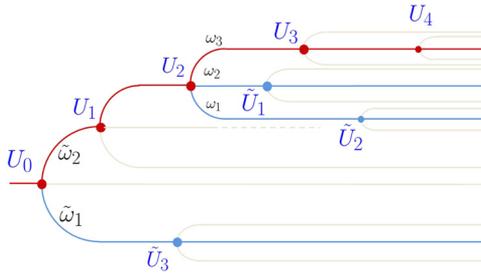
Fig. C.9. Diagram with colored branches. The branches with maximal and minimal (equal to zero) number of interactions along them are colored in red and gray, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where $\omega_i^- \equiv \omega_i - i\epsilon$, and $I_1(\omega)$ is the integral representation of the sum of all the weights of the subdiagram in the dashed frame, with incoming energy ω:

$$I_1(\omega) = \int \frac{d\omega_1 d\omega_2 d\omega_3 \delta(\omega_1 + \omega_2 + \omega_3 - \omega)}{\omega_1^-(\omega_1^- + \tilde{\mathcal{E}}_2)\omega_2^-(\omega_2^- + \tilde{\mathcal{E}}_1)\omega_3^-(\omega_3^- + \mathcal{E}_3)(\omega_3^- + \mathcal{E}_3 + \mathcal{E}_4)}. \tag{C.2}$$

By means of the residue theorem, I_0 can be rewritten as the sum over only 8 effective path weights:

$$I_0(\{U\}) = \frac{1}{\tilde{\mathcal{E}}_3} \frac{1}{\mathcal{E}_0(\mathcal{E}_0 + \mathcal{E}_1)} I_1(\mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_1)$$
$$- \frac{1}{\tilde{\mathcal{E}}_3} \frac{1}{(\tilde{\mathcal{E}}_3 + \mathcal{E}_0)(\tilde{\mathcal{E}}_3 + \mathcal{E}_0 + \mathcal{E}_1)} I_1(\tilde{\mathcal{E}}_3 + \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_1) \tag{C.3}$$

with

$$I_1(\omega) = \frac{1}{\tilde{\mathcal{E}}_1 \tilde{\mathcal{E}}_2} \left[f(\omega) - f(\omega + \tilde{\mathcal{E}}_1) - f(\omega + \tilde{\mathcal{E}}_2) + f(\omega + \tilde{\mathcal{E}}_1 + \tilde{\mathcal{E}}_2) \right] \tag{C.4}$$

and

$$f(X) = \frac{1}{X(X + \mathcal{E}_3)(X + \mathcal{E}_3 + \mathcal{E}_4)}. \tag{C.5}$$

Note that as a function of the $\mathcal{E}_i$ and $\tilde{\mathcal{E}}_i$, I_0 has poles only due to denominators which involve the incoming energy $\mathcal{E}_0$, while I_0 remains regular as any of the $\tilde{\mathcal{E}}_i \to 0$, due to cancellations among different terms.

The minimal number of effective paths associated to a diagram equals to the product of the number of residua of any of the performed integrals. This number can be determined from the structure of the diagram using the following rules: First, one eliminates the final leaves which are not associated to auxiliary frequencies, since they do not contribute with poles in the integral representation (Fig. C.9 represents the diagram of Fig. C.8, with these eliminated branches colored in gray). Then, one determines the directed path (branch) with the maximal number of interactions along it (red one in Fig. C.9). The auxiliary frequencies along this path are eliminated integrating the corresponding δ-functions. All remaining branches contribute one more residua than interactions along the branch. In the example of Fig. C.9, the three branches that remain after eliminating the red one contribute 2 residua each. The total number of effective paths is obtained by multiplying these numbers, which gives $2^3 = 8$ in the present case.
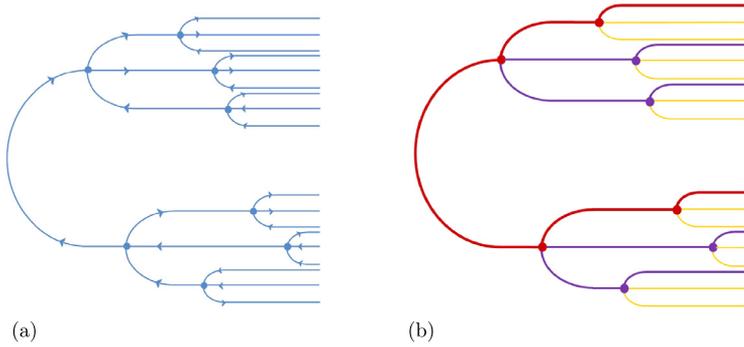
Fig. C.10. (a) Diagram with the maximal possible number of branchings. (b) Branches with the same number of interactions are drawn with the same color.

With the help of these rules, we count the minimal number of effective paths associated to the maximally branched diagram with N interactions, shown in Fig. C.10(a). We denote this number by $|\mathcal{P}|$.

The maximally branched diagram consists of two regular rooted trees with $L(N) \equiv \log(N+1)/\log 3$ generations. Since the weights of the two sub-diagrams factorize, we need to count only the effective paths associated to one of them, and square their number. We therefore consider one sub-diagram, and organize its branches according to the number of interactions along it (in Fig. C.10(b), branches with the same number of interactions have the same color). The number l of interactions along a branch ranges from 1 to $L(N) - 1$. There are $2 \cdot 3^{L(N)-1-l}$ branches with l interactions; each of them contributes with $(l+1)$ residua, yielding a total number of

$$\prod_{l=1}^{L(N)-1} 2(l+1)3^{L(N)-1-l} = L(N)!3^{\frac{(L(N)-1)^2}{2}}\left(\frac{2}{\sqrt{3}}\right)^{L(N)-1} \tag{C.6}$$

terms. The total number of effective paths associated to the diagram of Fig. C.10(a) is the square of this number, which grows as:

$$|\mathcal{P}| = \exp\left[(\log N)^2 \log 3 + O\left(\log N \log(\log N)\right)\right]. \tag{C.7}$$

As claimed in the main text, this number is sub-exponential in N.

Appendix D. Probability of large deviations in products of correlated denominators

Here we derive the probability of large deviations of effective path weights, i.e., the product of correlated denominators, as they occur in perturbation theory in the forward approximation.

We denote by $s_k = x_1 + \cdots + x_k$ the partial sums of i.i.d. random variables $x_i \equiv \mathcal{E}_i/\delta_\xi$. Let us assume the x_i to be unit Gaussian variables with probability density

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}. \tag{D.1}$$

Consider the distribution function $P_N(y)$ of the random variable

$$Y_N \equiv \log\left(\prod_{i=1}^{N} \frac{1}{|s_i|}\right) = -\sum_{i=1}^{N} \log|s_i|, \tag{D.2}$$

and its generating function $G_N(k)$

$$G_N(k) \equiv \mathbb{E}\big[e^{-kY_N}\big]. \tag{D.3}$$

Let us compute G_N for $N \gg 1$. We start by taking the expectation value over the joint distribution of $x_i = s_i - s_{i-1}$ ($s_0 \equiv 0$):

$$G_N(k) = \int \prod_{i=1}^N f(s_i - s_{i-1})e^{k \log|s_i|}ds_i = \int \mathcal{O}_k^{N-1}[f](s_N)|s_N|^k ds_N, \tag{D.4}$$

where the integral operator $\mathcal{O}_k[\cdot]$ acting on a function g is given by

$$\mathcal{O}_k[g](s) = \int f(s-x)|x|^k g(x)dx. \tag{D.5}$$

Consider now the basis of even functions:

$$g_n(x) = \frac{e^{-\frac{x^2}{2}}x^{2n}}{\sqrt{2\pi(2n)!}}, \quad n = 0, 1, \dots. \tag{D.6}$$

In this basis the linear action of $\mathcal{O}_k$ is given by:

$$\mathcal{O}_k[g_n](x) = \frac{1}{2\pi}\int e^{-\frac{1}{2}(x-y)^2}e^{-\frac{y^2}{2}}|y|^{k+2n}dy = \sum_{m\geq 0} O_{mn}(k)g_m(x), \tag{D.7}$$

with the matrix

$$O_{mn}(k) = \frac{1}{\sqrt{2\pi}}\frac{\Gamma\big(\frac{1+k}{2}+n+m\big)}{\sqrt{(2m)!(2n)!}}. \tag{D.8}$$

From (D.4) we thus readily obtain the following expression for $G_N(k)$:

$$G_N(k) = \sum_{m=0}^{\infty}\big(O(k)^{N-1}\big)_{m0}a_m, \tag{D.9}$$

with $a_m(k) = \int_{-\infty}^{\infty} g_m(s_N)|s_N|^k\, ds_N = \frac{2^{\frac{k+1}{2}+m}}{\sqrt{2\pi(2m)!}}\Gamma(\frac{k+1}{2}+m)$.

The matrix $O_{mn}(k)$ can be interpreted as a k-dependent Hamiltonian describing a particle hopping on a semi-infinite open chain with sites labeled by integers $m = 0, 1, 2, \dots$.

The large N behavior of $\log G_N$ is dominated by the largest eigenvalue $\lambda_{\max}(k)$ of $\mathcal{O}$. Since for any $k > -1$, $O(k)$ is symmetric and positive definite, the Perron–Frobenius theorem ensures that $\lambda_{\max}(k)$ is positive and unique, and

$$G_N(k) \approx c(k)\big[\lambda_{\max}(k)\big]^{N-1}, \tag{D.10}$$

where $c(k) = \phi_{\max,0} \cdot \sum_{m\geq 0} a_m \phi_{\max,m}$, and $\phi_{\max}$ is the normalized eigenvector corresponding to $\lambda_{\max}$.

Numerical results for the maximal eigenvalue are shown in Fig. D.11. They are obtained by truncating O to an increasing set of basis states (or chain of sites) $m \leq L$. For k close to the singularity $k = -1$ the results rapidly converge with increasing size L. In this region, we can extract information on the limiting curve $\lambda_{\max}(k)$ from the truncated chain. In particular, we see from the plot that both the function $\log \lambda_{\max}(k)$ and its negative slope diverge at $k = -1$, which
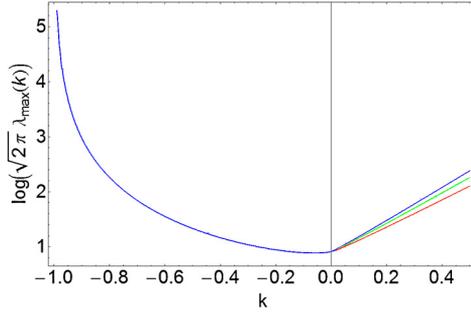
Fig. D.11. Maximal eigenvalue $\log \lambda_{max}(k)$ computed for truncated matrices $O(k)$ with basis sets of size $L = 120$ (red), $L = 200$ (green), $L = 300$ (blue). Close to the singularity $k = -1$, $\lambda_{max}(k)$ converges rapidly with L. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
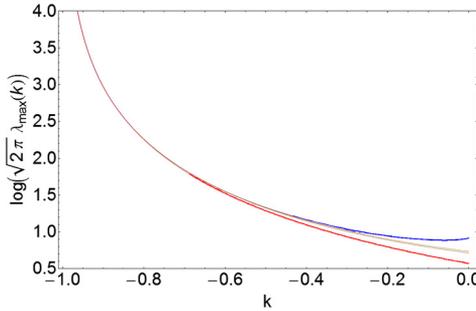

Fig. D.12. Comparison between $\log \lambda_{max}(k)$ obtained numerically for the truncated matrix (with $L = 300$ basis functions) and the analytic expression $\log[\Gamma(\frac{k+1}{2}) + \sqrt{2\pi}\delta\lambda(k)]$ with $\sqrt{2\pi}\delta\lambda(k)$ expanded at zeroth (red), first (brown) and second (orange) order in $(k + 1)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

will also follow form the analysis below. Hence, $k \gtrsim -1$ is the relevant region for the saddle point approximation of Eq. (65), if very large deviations $\tilde{y} \gg 1$ are considered.

Due to the proximity to a logarithmic divergence at $k = -1$, to order $O(1 + k)$ the eigenstate ϕ_{max} for $k \sim -1$ is localized on the first site ($n = 0$) of the corresponding hopping chain:

$$|\phi_{max}\rangle \simeq |0\rangle, \tag{D.11}$$

with an eigenvalue

$$\lambda_{max}(k) \simeq O(k)_{00} = \frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{1+k}{2}\right). \tag{D.12}$$

Corrections to the maximal eigenvalue (D.12) can be evaluated perturbatively in the matrix elements $O_{ik\neq00}$ (D.8), which yields

$$\lambda_{max}(k) = \frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{1+k}{2}\right) + \lambda_{max}^{(2)}(k) + \lambda_{max}^{(3)}(k) + \dots$$

$$\equiv \frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{1+k}{2}\right) + \delta\lambda(k). \tag{D.13}$$

One can show that $\delta\lambda(k)$ is analytic around $k = -1$ and satisfies $\delta\lambda(k \to -1) \to 0$. This is due to the fact that in nth order perturbation theory $\lambda_{\max}^{(n)}$ is proportional to denominators of the form $1/O_{00}^{n-1} \sim (k+1)^{n-1}$. The leading term in $\delta\lambda(k)$ results from:

$$\lambda_{\max}^{(2)}(k) = \sum_{m=1}^{\infty} \frac{(2\pi)^{-\frac{1}{2}}\left[\Gamma\left(\frac{1+k}{2}+m\right)\right]^2}{\Gamma\left(\frac{1+k}{2}\right)(2m)! - \Gamma\left(\frac{1+k}{2}+2m\right)} = \frac{1}{\sqrt{2\pi}}\frac{\pi^2}{36}(k+1) + O(k+1)^2. \quad \text{(D.14)}$$

A plot of the corrections to the maximal eigenvalue (D.12) is given in Fig. D.12.

The inverse Laplace transform of the generating function is governed by

$$\begin{aligned}
\phi(\tilde{y}, k) &:= \tilde{y}k + \lim_{N\to\infty} \frac{G_N(k)}{N} \\
&= \tilde{y}k + \log\left[\Gamma\left(\frac{k+1}{2}\right) + \sqrt{2\pi}\delta\lambda(k)\right] - \frac{1}{2}\log 2\pi.
\end{aligned} \quad \text{(D.15)}$$

It has a saddle point at $k = k^*(\tilde{y})$, determined by

$$\begin{aligned}
\tilde{y} &= -\frac{d}{dk}\left\{\log\left[\Gamma\left(\frac{1+k}{2}\right) + \sqrt{2\pi}\delta\lambda(k)\right]\right\}_{k=k^*(\tilde{y})} \\
&= -\left\{\frac{1}{2}\psi^{(0)}\left(\frac{1+k}{2}\right)\left[1 + \frac{\sqrt{2\pi}\delta\lambda(k)}{\Gamma\left(\frac{1+k}{2}\right)}\right]^{-1} + \frac{\sqrt{2\pi}\delta\lambda'(k)}{\Gamma\left(\frac{1+k}{2}\right) + \sqrt{2\pi}\delta\lambda(k)}\right\}_{k=k^*(\tilde{y})},
\end{aligned} \quad \text{(D.16)}$$

where $\psi^{(0)}(x) \equiv d\log[\Gamma(x)]/dx$.

To isolate the singularity in $k = -1$ we use the Laurent expansion of $\psi^{(0)}(x)$ around $x = 0$:

$$\psi^{(0)}\left(\frac{1+k}{2}\right) = -\frac{2}{k+1} - \gamma + \frac{\pi^2}{12}(k+1) + O\left((k+1)^3\right), \quad \text{(D.17)}$$

where γ is the Euler constant. This allows us to recast (D.16) in the following form:

$$\tilde{y} = \frac{1}{k^*+1} + Q\left(k^*+1\right). \quad \text{(D.18)}$$

Here, $Q(\cdot)$ is an analytic function with expansion:

$$Q(x) = \frac{\gamma}{2} - \frac{\pi^2}{18}x + O\left(x^2\right). \quad \text{(D.19)}$$

This yields the equation

$$1 + k^* = \frac{1}{\tilde{y}}\left(1 - \frac{Q(k^*+1)}{\tilde{y}}\right)^{-1}, \quad \text{(D.20)}$$

which can be solved by iteration as an expansion in $1/\tilde{y}$:

$$1 + k^*(\tilde{y}) = \frac{1}{\tilde{y}} + \frac{\gamma}{2}\frac{1}{\tilde{y}^2} + \left(\frac{\gamma^2}{4} - \frac{\pi^2}{18}\right)\frac{1}{\tilde{y}^3} + O\left(\frac{1}{\tilde{y}^4}\right). \quad \text{(D.21)}$$

Expanding (D.15) in powers of $k+1$ and substituting (D.21) we find:

$$\phi\left(\tilde{y}, k^*(\tilde{y})\right) = -\tilde{y} + \log\tilde{y} - \frac{1}{2}\log\left(\frac{\pi}{2e^2}\right) - \frac{\gamma}{2\tilde{y}} + \frac{1}{8}\left(\frac{5\pi^2}{18} - \gamma^2\right)\frac{1}{\tilde{y}^2} + O\left(\frac{1}{\tilde{y}^3}\right). \quad \text{(D.22)}$$

Finally, within the saddle point approximation to Eq. (63), for $\tilde{y} \gg 1$ we find the large deviation probability

$$P_N\left(-\log\left[\prod_{i=1}^{N}\frac{1}{|s_i|}\right] = N\tilde{y}\right) = C(\tilde{y}, N)\left(\frac{2e}{\sqrt{2\pi}}\right)^N \tilde{y}^N e^{-N\mathcal{F}(\tilde{y})}\left[1 + \frac{1}{N}\right], \tag{D.23}$$

where

$$\mathcal{F}(\tilde{y}) = \tilde{y} + \frac{\gamma}{2\tilde{y}} - \frac{1}{8}\left(\frac{5\pi^2}{18} - \gamma^2\right)\frac{1}{\tilde{y}^2} + O\left(\frac{1}{\tilde{y}^3}\right). \tag{D.24}$$

The prefactor

$$C(\tilde{y}, N) = \left(\frac{1}{2\pi N\phi_N''(k^*(\tilde{y}))}\right)^{\frac{1}{2}} \frac{c(k^*(\tilde{y}))}{\lambda_{\max}(k^*(\tilde{y}))} \tag{D.25}$$

yields only logarithmic corrections to the exponent.

As commented in the main text, when restricting to the linear term in (D.24), the large deviation statistics for the correlated denominators coincides with that of *independent* identically distributed energy denominators. Indeed, from Eqs. (D.10) and (D.12) it follows that to leading order in $k + 1$ the exponential growth of G_N is almost equal to that of the generating function $g_N(k) = [2^{\frac{k+1}{2}}\Gamma(\frac{k+1}{2})/\sqrt{2\pi}]^N$ associated with products of N independent Gaussian denominators with unit variance. For $\tilde{y} \gg 1$, the tail of the distribution is determined by the residue of the pole of the generation function at $k = -1$, which is identical in the two cases. Repeating the above derivation of large deviations for independent denominators with generating function $g_N(k)$, one finds that it differs from (D.23) at order $O(\frac{1}{\tilde{y}})$: the tails for correlated denominators are suppressed by a factor $\exp(-N\frac{\log 2}{2}\frac{1}{\tilde{y}})$. The correction $\delta\lambda(k)$ in (D.13) contributes to (D.24) only at order $O(1/\tilde{y}^2)$.

Appendix E. Some useful combinatoric results for diagrams

Let $\mathcal{T}_n$ be the number of tree-like diagrams with a root of connectivity 2, and n vertices with connectivity 4. These trees are obtained by merging two trees of branching ratio 3 at the root, and therefore

$$\mathcal{T}_n = \sum_{\substack{n_1,n_2\geq 0 \\ n_1+n_2=n}} T^{(n_1)}T^{(n_2)}, \tag{E.1}$$

where $T^{(m)}$ is the number of trees with m vertices (including the root) and branching ratio 3. This number satisfies the recursion equations

$$T^{(0)} = 1, \tag{E.2}$$

$$T^{(n)} = \sum_{n_1+n_2+n_3=n-1} T^{(n_1)}T^{(n_2)}T^{(n_3)}. \tag{E.3}$$

We can define the generating function

$$T(x) = \sum_{n\geq 0} x^n T^{(n)}, \tag{E.4}$$

which in virtue of (E.3) satisfies the polynomial equation

$$T(x) = 1 + xT(x)^3.$$ (E.5)

Notice that the first singularity of $T(x)$ is a branch-cut at $x = 4/27$, which implies the large-n behavior $T^{(n)} \sim (27/4)^n$. However, we can find the nth order of the expansion for small x using Lagrange's inversion theorem for the inverse function of

$$x(T) = \frac{T-1}{T^3},$$ (E.6)

expanding around $T = 1$ ($x = 0$).

This yields

$$T^{(n)} = \frac{1}{n} \lim_{T \to 1} \left[\frac{1}{(n-1)!} \frac{d^{n-1}}{dT^{n-1}} \left(\frac{T-1}{x(T)} \right)^n \right] = \frac{1}{2n+1} \binom{3n}{n}.$$ (E.7)

In general, for k-body interactions we have $T^{(n)} = \binom{(k-1)n}{n} / ((k-2)n + 1)$ diagrams. For $k = 3$ these are the numbers of binary trees with n vertices, or Catalan numbers.

There are two ways to solve Eq. (E.1) and find $\mathcal{T}_n$. The first one is to notice that its generating function $\mathcal{T}(x)$ satisfies $\mathcal{T}(x) = T(x)^2$, write Eq. (E.6) in terms of $\mathcal{T}$ and use Lagrange's inversion theorem again. Alternatively, one can use the explicit form of $T^{(n)}$ and apply a summation formula for the ratio of four Γ-functions to obtain:

$$\mathcal{T}_n = \frac{3^{\frac{3}{2}+3n}}{\pi} \frac{\Gamma\left(n+\frac{2}{3}\right)\Gamma\left(n+\frac{4}{3}\right)}{\Gamma(2n+3)} \sim \frac{3}{4}\sqrt{\frac{3}{\pi}} \frac{1}{n^{\frac{3}{2}}} \left(\frac{27}{4}\right)^n$$ (E.8)

which is the result quoted in the text.

References

[1] P.W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. 109 (1958) 1492.
[2] L. Fleishman, P.W. Anderson, Interactions and the Anderson transition, Phys. Rev. B 21 (1980) 2366.
[3] R. Berkovits, B.I. Shklovskii, Statistics of energy spectra of a strongly disordered system of interacting electrons, J. Phys. Condens. Matter 11 (1999) 779.
[4] D.M. Basko, I.L. Aleiner, B.L. Altshuler, Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states, Ann. Phys. 321 (2006) 1126.
[5] Y. Bar Lev, G. Cohen, D.R. Reichman, Absence of diffusion in an interacting system of spinless fermions on a one-dimensional disordered lattice, arXiv:1407.7535 [cond-mat.dis-nn].
[6] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Mueller, E. Demler, Anomalous diffusion and Griffiths effects near the many-body localization transition, arXiv:1408.3413 [cond-mat.dis-nn].
[7] L. D'Alessio, A. Polkovnikov, Many-body energy localization transition in periodically driven systems, Ann. Phys. 333 (2013) 19.
[8] P. Ponte, A. Chandran, Z. Papić, D.A. Abanin, Periodically driven ergodic and many-body localized quantum systems, arXiv:1403.6480 [cond-mat.dis-nn].
[9] D.M. Basko, I.L. Aleiner, B.L. Altshuler, On the problem of many-body localization, eprint, arXiv:cond-mat/0602510.
[10] D. Basko, I. Aleiner, B. Altshuler, Possible experimental manifestations of the many-body localization, Phys. Rev. B 76 (2007) 052203.
[11] V. Oganesyan, D.A. Huse, Localization of interacting fermions at high temperature, Phys. Rev. B 75 (2007) 155111.
[12] M. Žnidarič, T. Prosen, P. Prelovšek, Many-body localization in the Heisenberg XXZ magnet in a random field, Phys. Rev. B 77 (2008) 064426.
[13] A. Pal, D.A. Huse, Many-body localization phase transition, Phys. Rev. B 82 (2010) 174411.
[14] J.H. Bardarson, F. Pollmann, J.E. Moore, Unbounded growth of entanglement in models of many-body localization, Phys. Rev. Lett. 109 (2012) 017202.

[15] R. Vosk, E. Altman, Dynamical quantum phase transitions in random spin chains, arXiv:1307.3256 [cond-mat.dis-nn].

[16] A. De Luca, A. Scardicchio, Ergodicity breaking in a model showing many-body localization, Europhys. Lett. 101 (2013) 37003.

[17] J.A. Kjäll, J.H. Bardarson, F. Pollmann, Many-body localization in a disordered quantum Ising chain, arXiv:1403.1568 [cond-mat.str-el].

[18] I. Aleiner, B. Altshuler, G. Shlyapnikov, A finite-temperature phase transition for disordered weakly interacting bosons in one dimension, Nat. Phys. 6 (2010) 900.

[19] D.A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, S. Sondhi, Localization-protected quantum order, Phys. Rev. B 88 (2013) 014206.

[20] A. Chandran, V. Khemani, C.R. Laumann, S.L. Sondhi, Many-body localization and symmetry-protected topological order, Phys. Rev. B 89 (2014) 144201.

[21] D. Pekker, G. Refael, E. Altman, E. Demler, V. Oganesyan, Hilbert-glass transition: new universality of temperature-tuned many-body dynamical quantum criticality, Phys. Rev. X 4 (2014) 011052.

[22] Y. Bahri, R. Vosk, E. Altman, A. Vishwanath, Localization and topology protected quantum coherence at the edge of 'hot' matter, arXiv:1307.4092 [cond-mat.dis-nn].

[23] N.Y. Yao, C.R. Laumann, S. Gopalakrishnan, M. Knap, M. Mueller, E.A. Demler, M.D. Lukin, Many-body localization with dipoles, arXiv:1311.7151 [cond-mat.stat-mech].

[24] Y. Kagan, L. Maksimov, Localization in a system of interacting particles diffusing in a regular crystal, Sov. Phys. JETP 60 (1984) 1.

[25] M. Schiulaz, M. Müller, Ideal quantum glass transitions: many-body localization without quenched disorder, arXiv:1309.1082 [cond-mat.dis-nn].

[26] M. Schiulaz, A. Silva, M. Müller, Dynamics in many-body localized quantum systems without disorder, arXiv:1410.4690 [cond-mat.dis-nn].

[27] W. De Roeck, F. Huveneers, Asymptotic quantum many-body localization from thermal disorder, arXiv:1308.6263 [cond-mat.stat-mech].

[28] N.Y. Yao, C.R. Laumann, J.I. Cirac, M.D. Lukin, J.E. Moore, Quasi many-body localization in translation invariant systems, arXiv:1410.7407 [cond-mat.dis-nn].

[29] T. Prosen, Time evolution of a quantum many-body system: transition from integrability to ergodicity in the thermodynamic limit, Phys. Rev. Lett. 80 (1998) 1808–1811.

[30] T. Prosen, Ergodic properties of a generic nonintegrable quantum many-body system in the thermodynamic limit, Phys. Rev. E 60 (1999) 3949–3968.

[31] C.D. Errico, E. Lucioni, L. Tanzi, L. Gori, G. Roux, I.P. McCulloch, T. Giamarchi, M. Inguscio, G. Modugno, Observation of a disordered bosonic insulator from weak to strong interactions, arXiv:1405.1210 [cond-mat.dis-nn].

[32] Z. Ovadyahu, Suppression of inelastic electron–electron scattering in Anderson insulators, Phys. Rev. Lett. 108 (2012) 156602.

[33] D.A. Huse, V. Oganesyan, A phenomenology of certain many-body-localized systems, arXiv:1305.4915 [cond-mat.dis-nn].

[34] M. Serbyn, Z. Papić, D.A. Abanin, Local conservation laws and the structure of the many-body localized states, Phys. Rev. Lett. 111 (2013) 127201.

[35] J.Z. Imbrie, On many-body localization for quantum spin chains, arXiv:1403.7837 [math-ph].

[36] A. Chandran, I.H. Kim, G. Vidal, D.A. Abanin, Constructing local integrals of motion in the many-body localized phase, arXiv:1407.8480 [cond-mat.dis-nn].

[37] V. Nguyen, B. Spivak, B. Shklovskii, Tunnel hopping in disordered systems, Sov. Phys. JETP 62 (1985) 1021.

[38] V. Nguyen, B. Spivak, B. Shklovskii, Aharonov–Bohm oscillations with normal and superconductive flux-quanta in hopping conduction, Sov. Phys. JETP Lett. 41 (1985) 42.

[39] M. Kardar, Statistical Physics of Particles, Cambridge University Press, 2007.

[40] M. Müller, Magnetoresistance and localization in bosonic insulators, Europhys. Lett. 102 (2013) 67008.

[41] B.L. Altshuler, Y. Gefen, A. Kamenev, L.S. Levitov, Quasiparticle lifetime in a finite system: a nonperturbative approach, Phys. Rev. Lett. 78 (1997) 2803.

[42] A. De Luca, A. Scardicchio, V.E. Kravtsov, B.L. Altshuler, Support set of random wave-functions on the Bethe lattice, arXiv:1401.0019 [cond-mat.stat-mech].

[43] A. De Luca, B.L. Altshuler, V.E. Kravtsov, A. Scardicchio, Anderson localization on the Bethe lattice: nonergodicity of extended states, Phys. Rev. Lett. 113 (2014) 046806.

[44] C.R. Laumann, A. Pal, A. Scardicchio, Many-body mobility edge in a mean-field quantum spin glass, arXiv:1404.2276 [cond-mat.stat-mech].

[45] R. Abou-Chacra, D. Thouless, P. Anderson, A selfconsistent theory of localization, J. Phys. C 6 (1973) 1734.
[46] Y. Bar Lev, D.R. Reichman, Dynamics of many-body localization, arXiv:1402.0502 [cond-mat.dis-nn].
[47] I. Gornyi, A. Mirlin, D. Polyakov, Interacting electrons in disordered wires: Anderson localization and low-T transport, Phys. Rev. Lett. 95 (2005) 206603.
[48] D. Thouless, Electrons in disordered systems and the theory of localization, Phys. Rep. 13 (1974) 93.
[49] W. De Roeck, F. Huveneers, Scenario for resonant delocalization in translation invariant systems, arXiv:1405.3279 [cond-mat.stat-mech].
[50] B. Bauer, C. Nayak, Area laws in a many-body localized state and its implications for topological order, J. Stat. Mech. 2013 (2013) P09005.

Corrigendum

Corrigendum to "Integrals of motion in the many-body localized phase" [Nucl. Phys. B 891 (2015) 420–465]

V. Ros [a,b,*], M. Müller [c,d], A. Scardicchio [c,b]

[a] *SISSA, via Bonomea 265, 34136 Trieste, Italy*
[b] *INFN, Sezione di Trieste, Strada Costiera 11, 34151 Trieste, Italy*
[c] *Abdus Salam ICTP, Strada Costiera 11, 34151 Trieste, Italy*
[d] *Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland*

Abstract

We correct a small error in our article *Integrals of motion in the many body localized phase* [1]. The correction does not alter the main result regarding the convergence of the perturbative expansion for integrals of motion in forward approximation, but reduces the estimate of the radius of convergence by a numerical factor of roughly $\simeq 1.79$.

© 2015 The Authors. Published by Elsevier B.V.

In Ref. [1] we associated to any diagram with N interaction vertices a minimal number of 'effective paths' that result from the integration procedure described in Section 5. We stated that this minimal number grows *sub-exponentially* with N, which relied on computing this number for diagrams d with a maximally branched geometry. However, the latter was erroneously determined to be sub-exponential in N in Eq. (C.6). The corrected version of Eq. (C.6) (counting the number of effective paths associated to either of the two rooted trees of the diagram) instead reads:

$$\prod_{l=1}^{L(N)-1} (l+1)^{2 \cdot 3^{L(N)-1-l}} = \exp\left(3^{L(N)} \sum_{k=1}^{\infty} \frac{2 \log k}{3^k} - \frac{2}{3} \sum_{k=0}^{\infty} \frac{\log(k+L(N)+1)}{3^k}\right), \tag{1}$$

with $L(N) = \log(N+1)/\log 3$. Using that $2\sum_{k=1}^{\infty} 3^{-k} \log(k) = 0.29$, one finds that the minimal number of effective paths $|\mathcal{P}|$ for diagrams with this geometry, which is the square of the above, scales as:

$$|\mathcal{P}| = \exp\left[0.58 \, N + o(N)\right], \tag{2}$$

which should replace the estimate in Eq. (C.7).

Among all the possible geometries of diagrams with a fixed number of interactions N, the maximally branched geometry is the one that maximizes the number of effective paths. Thus, Eq. (2) is an upper bound for the average $\overline{|\mathcal{P}(d)|}$ introduced in Eq. (55), which we also expect to have an exponential scaling in N:

$$\overline{\mathcal{P}(d)} \sim \exp\left[\alpha \, N + o(N)\right] \tag{3}$$

with $0 < \alpha < 0.58$. Accounting for this correction, the total number of effective paths of length N, $\mathcal{N}_N$ in Eqs. (59), is modified accordingly:

$$\mathcal{N}_N \to e^{\alpha N} \mathcal{N}_N. \tag{4}$$

Since the additional factor is only exponential in N, the conclusion about the convergence of the construction of integrals of motion for small enough interactions is unaltered. The effect of the correction is to slightly diminish the radius of convergence of the construction.

The precise effect of this correction depends on the relative weight of the effective paths and their mutual interference.

If we make the simplifying assumption that the effective paths associated to the same diagram can be treated as independent random variables, the sum $S(d)$ in Eq. (53) is dominated by the largest term, and the factor $\overline{|\mathcal{P}(d)|}$ enhances the tail of the distribution of $\mathcal{A}_{\mathcal{J},\mathcal{J}}^{(\alpha)}$ as compared to the tail corresponding to a single path weight, see Eq. (58). The above discussed correction would thus modify by a factor e^{α} the numerical constant C in Eq. (93). (Note, however, that this constant C was already subject to uncertainty, see Eq. (85), due to the approximations going into the estimate of $\mathcal{N}_N$.) Approximating $e^{\alpha} \approx e^{0.58} = 1.79$ we find that the result of Eq. (93) holds with the following uncertainty on C:

$$18.97 < C < 36.25. \tag{5}$$

The above assumption neglects, however, that the effective paths associated to a given diagram are not independent. Indeed, they involve the same energy variables in the denominators, but in different combinations. These correlations might be relevant when computing the large deviations for $S(d)$. In fact there could be disorder realizations in which all the energy variables are simultaneously small, in such a way that there is no dominant effective path. In an extreme case, all $\tilde{\omega}_\Gamma$ contributing to $S(d)$ might happen to be of the same order of magnitude and atypically large. These contributions will come with different signs and partially cancel, which counteracts the enhancement of the total amplitude. To estimate an upper bound for the effect of the exponential number of effective paths on the constant C we neglect those partial cancellations, and assume that the diagrams dominating the tails of $S(d)$ are such that essentially all effective paths add up constructively with comparable weights. Under this extreme scenario, the large deviations of $S(d)$ would be given in terms of those of a single path weight $\tilde{\omega}$ by setting $S(d) \sim \overline{|\mathcal{P}(d)|}\tilde{\omega}$. In

this approximation, Eq. (93) is recovered with the substitution $\lambda \to \lambda e^{\alpha}$. This shifts the estimated interval for C in Eq. (5) only by a logarithmic factor.

Acknowledgement

References

[1] V. Ros, M. Müller, A. Scardicchio, Nucl. Phys. B 891 (2015) 420.