Real Time Simulations of Quantum Spin Chains: Density-of-States and Reweighting approaches

Pavel Buividovich[1] and Johann Ostmeyer[1]

[1] *Department of Mathematical Sciences, University of Liverpool, United Kingdom*

(Dated: 11th December 2022)

We put the Density-of-States (DoS) approach to Monte-Carlo (MC) simulations under a stress test by applying it to a physical problem with the worst possible sign problem: the real time evolution of a non-integrable quantum spin chain. Benchmarks against numerical exact diagonalisation and stochastic reweighting are presented. Both MC methods, the DoS approach and reweighting, allow for simulations of spin chains as long as $L = 40$, far beyond exact diagonalisability, though only for short evolution times $t \lesssim 1$. We identify discontinuities of the density of states as one of the key problems in the MC simulations and propose to calculate some of the dominant contributions analytically, increasing the precision of our simulations by several orders of magnitude. Even after these improvements the density of states is found highly non-smooth and therefore the DoS approach cannot outperform reweighting. We prove this implication theoretically and provide numerical evidence, concluding that the DoS approach is not well suited for quantum real time simulations with discrete degrees of freedom.

I. INTRODUCTION

The simulation of a real-time evolution of a many-body quantum system up to some finite physical time t is a problem of utmost importance across many areas of physics. However, it is an NP-hard problem which, in general, can only be solved on a classical computer in a time

$$T_{\text{sim}} \sim \mathrm{e}^{\alpha(t)\, n_{\text{dof}}} \tag{1}$$

that grows exponentially with the number of degrees of freedom n_{dof} in a system. Polynomial-time solutions are only possible on a quantum computer [1, 2]. With the current state of technology, only relatively small quantum systems can be reliably simulated on quantum computers. Therefore, at least for some time we will have to rely on classical computers to simulate the real-time evolution of quantum systems. While we certainly cannot simulate the real-time evolution in polynomial time on a classical computer, we can try to reduce the coefficient $\alpha(t)$ in the exponent in (1).

Conceptually, the simplest approach to simulate real-time evolutions is to find all eigenstates of the Hamiltonian using numerical exact diagonalization. This approach works well for systems with finite-dimensional Hilbert spaces, such as quantum spin chains, and allows to calculate any real-time evolution[1] in a straightforward way. For a chain of L spin-1/2 degrees of freedom, such as Ising spins, the Hilbert space dimension is $N = 2^L$. Correspondingly, the computational cost of the "plain vanilla" exact diagonalization (ED) scales as $\mathcal{O}\left(N^3\right) = \mathcal{O}\left(2^{3L}\right)$. Physical results found in most of the literature scale up to $L \leq 18$, see e.g. [3–6]. More advanced methods, such as the shift-invert method of exact diagonalization (SIMED) [7] and polynomially filtered exact diagonalization (POLFED) [8] allow for extremely efficient ED. In particular the POLFED has only $\mathcal{O}\left(N^2\right) = \mathcal{O}\left(2^{2L}\right)$ runtime and $\mathcal{O}(N) = \mathcal{O}\left(2^L\right)$ memory requirement, both up to polynomial corrections in L and for a fixed number of eigenvalues. The largest systems simulated to date (to our knowledge) have $L = 26$ and the largest ones used for physics beyond mere algorithmic proof of principle have $L = 24$. Unfortunately the polynomial corrections are not specified in the literature but they appear to be quite significant if these are the largest sizes possible. In addition, to our knowledge it is only claimed that the limited number of eigenvalues from the bulk yield the correct results. The truncation in eigenspectrum might well cause relevant artifacts.

Another possible approach is to approximate the quantum evolution $\exp\left(i\hat{H}t\right)$ as a product of many exactly treatable factors using a Suzuki-Trotter decomposition. Instead of exact diagonalization, in this case we have to perform multiple matrix-matrix multiplications, with a computational cost that scales as $\mathcal{O}\left(N^2\right) = \mathcal{O}\left(2^{2L}\right)$ for local Hamiltonians. This allows to simulate spin chains with $L \leq 24$ [9, 10].

An alternative to methods based on calculations in the entire Hilbert space is provided by Monte-Carlo methods, which replace exact summation over all states with importance sampling. For finite-temperature equilibrium partition functions, Monte-Carlo methods typically reduce the exponential scaling in the system size L down to a polynomial one. Monte-Carlo methods can also be used for bosonic systems in a straightforward way, in contrast to exact diagonalization which necessarily relies on truncations of an infinite-dimensional Hilbert space.

In the absolute majority of cases, real-time evolution problems have path integral representations with oscillatory integrands, either real- or complex-valued. The most straightforward approach to deal with non-

[1] Up to floating-point round-off errors, that might become important at very late times.

positive-definite integrands in Monte-Carlo simulations is reweighting, whereby one performs importance sampling with the weight proportional to the absolute value of the integrand, and re-weights the contribution of any field configuration with the corresponding complex phase of the integrand. However, in this case the NP-hardness of the real time evolution problem reveals itself in the exponentially quick deterioration of statistical power with the number of degrees of freedom n_{dof} and the evolution time t. Many contributions with different complex phases cancel each other, and exponentially many Monte-Carlo samples are needed to calculate the path integral with desired precision. This exponential growth of computational complexity is referred to as the "sign problem".

In the past decade, the Density of States (DoS) approach has been advocated as an efficient way to make the sign problem milder [11–14]. Starting from a generic oscillatory path integral representation of a partition function Z of the form

$$Z = \int \mathcal{D}\phi \exp\left(-S_R\left[\phi\right] + \mathrm{i}\, S_I\left[\phi\right]\right) , \tag{2}$$

the basic idea of the DoS approach is to construct a numerical approximation to the Density of States, or probability distribution, of the complex phase $S_I\left[\phi\right]$:

$$\rho\left(E\right) = \int \mathcal{D}\phi \exp\left(-S_R\left[\phi\right]\right) \delta\left(E - S_I\left[\phi\right]\right) . \tag{3}$$

Here $\int \mathcal{D}\phi$ denotes integration or summation over any continuous or discrete variables ϕ. Knowing $\rho\left(E\right)$, which can be determined without encountering a sign problem, we can calculate the full partition function Z as a simple one-dimensional integral over E

$$Z = \int \mathrm{d}E \rho\left(E\right) \mathrm{e}^{\mathrm{i}\, E} . \tag{4}$$

An important development was the *Logarithmic Linear Relaxation* (LLR)~~LLR~~ method [15–17] based on Robbins-Monro iterations [18], which allowed to obtain high-precision results for the Density of States $\rho\left(E\right)$ for lattice models with continuous [19, 20] as well as discrete [21] degrees of freedom. Applications include the compact $U\left(1\right)$ lattice gauge theory [19], heavy-dense QCD at finite quark density [11, 12], finite-density Bose gas [13], finite-density Hubbard model [20] and the Potts model [21]. It is fair to say that in all cases the computational complexity of evaluating the original multi-dimensional oscillatory path integral (2) translates into the complexity of evaluating the one-dimensional oscillatory integral (4), which is very sensitive to statistical errors in the density of states $\rho\left(E\right)$. This problem was tackled in [11–13, 20, 21] by constructing various analytic approximations to $\rho\left(E\right)$ and using them to calculate the integral. The approximations included polynomial fits and splines. Resummations based on Fourier transforms of $\rho\left(E\right)$ were considered in [20]. This allowed to obtain results of practical significance for lattice sizes that are intractable with other methods, e.g. the straightforward reweighting.

In this work, we put the DoS/LLR approach under a stress test by using it to simulate the real-time evolution, for which the partition function in (2) is strongly dominated by the imaginary part of the action S_I. This is hence a physical problem for which the sign problem is expected to be maximally strong. As a model of choice, we use a non-integrable quantum Ising chain with quenched disorder [3]. We explicitly quantify the computational cost required to obtain a fixed error of our measurements, and compare it with that of the reweighting and exact diagonalization methods. We find that without further analytic approximations, the computational cost of evaluating the one-dimensional integral (4) of the Density of States is not better than that of simple reweighting. We also critically examine the smoothness of the DoS, and find an efficient way to remove some of the dominant discontinuities, which improves the performance of both the DoS and the reweighting approaches by orders of magnitude. This also allows us to construct reasonably good polynomial approximations for the DoS $\rho\left(E\right)$, however they turn out to hardly improve the quality of the results because of the high number of parameters required. Therefore, for our particular system, the DoS approach only outperforms reweighting in very rare cases. Our overall conclusion is that both Monte-Carlo approaches, LLR and reweighting, are advantageous compared to exact diagonalization when simulating short-time evolutions for large spatial system sizes, even allowing for sizes completely inaccessible to ED. On the other hand, exact diagonalization is clearly better for long times.

We chose the non-integrable quantum Ising chain with quenched disorder [3] as our test model because it is one of the simplest non-integrable models which exhibit ergodic and many-body localization regimes. The Hamiltonian of the model reads, in its most general form

$$H \ = -\sigma^z \cdot \mathbb{J} \cdot \sigma^z - \mathbb{h} \cdot \sigma^x, \tag{5}$$

where $\sigma^{x,z} = (\sigma_1^{x,z}, \ldots, \sigma_L^{x,z})$ is the spatial vector of Pauli x- and z-spin matrices, respectively, $\mathbb{J} \in \mathbb{R}^{L\times L}$ is some coupling matrix and $\mathbb{h} \in \mathbb{R}^L$ is a vector of local magnetic fields $\mathbb{h}_i$.

We work in the ergodic regime, where no symmetry can accidentally make the sign problem milder. Furthermore, the model has a straightforward path integral representation. Its energy spectrum can be easily obtained using exact numerical diagonalization for chain lengths of order $L \lesssim 20$, which allows to obtain benchmark results for any real-time observables without any statistical errors. In this work we focus on one of the simplest real-time quantities, the infinite-temperature spectral form-factor

$$\mathcal{K}(t) = \langle K(t) \rangle_{\mathbb{J}} = \left\langle \left| \mathrm{Tr}\, \mathrm{e}^{\mathrm{i}\, H t} \right|^2 \right\rangle_{\mathbb{J}} , \tag{6}$$

where $\langle \cdot \rangle_{\mathbb{J}}$ denotes the average over different disorder realisations and $U(t)$ is the real time evolution operator for given disorder $\mathbb{J}$. Universal late-time features of spectral form-factors, such as the "ramp" [22–24], have proven useful as probes of quantum chaos in strongly-interacting

many-body systems. In particular, spectral form-factors are useful for distinguishing ergodic and many-body localized regimes [25].

Originally, our hope was to apply the DoS/LLR approach to resolve the ongoing debate about the existence and the properties of the many-body localization phase in quantum spin chains [3–6, 9, 10, 26–32]. The debate is strongly focused on extrapolations towards the thermodynamic limit, therefore it is useful to devise methods that could work for lattice sizes that are currently inaccessible for numerical exact diagonalization. While our improvements to the DoS/LLR approach allow to simulate spin chains of lengths as large as $L = 40$, well beyond the reach of any exact diagonalization methods, such simulations are bound to early times $t \lesssim 1$. This is by far insufficient to resolve the late-time behavior of spectral form-factors, which starts distinguishing the ergodic and the many-body localized regimes at the Heisenberg times scale $t \gtrsim 2^L$. Let us note in passing that the necessity to simulate up to times of order of 2^L is very likely to make the calculation of the spectral form-factor an NP-hard problem even for a quantum computer. Indeed, the "no fast-forwarding theorem" [33] is likely to hold for our Hamiltonian. Therefore, quantum computers will need an exponentially large computational time $T_{\text{sim}} \gtrsim 2^L$ to simulate physical evolution up to physical times $t \gtrsim 2^L$.

The rest of this work is structured as follows. We provide the details of the model (5) and its classical $1+1$-dimensional counterpart in Section II. Next, we introduce the DoS/LLR algorithm in Section III and we derive how the algorithm can be optimised in Section IV. Analytic estimates for the runtimes of all the different approaches are discussed in Section V and, finally, the results of numerical experiments are presented in Section VI.

II. THE MODEL

In this work, we follow [3] and consider the Hamiltonian (5) with both the nearest-neighbor and next-to-nearest-neighbor couplings:

$$\mathbb{J}_{ij} = (\mathbb{J}_0 + \Delta\mathbb{J}_i)\, \delta_{i,j+1} + \mathbb{J}_2 \delta_{i,j+2}, \tag{7}$$

where the nearest neighbour coupling contains quenched disorder $\Delta\mathbb{J}_i$ that is sampled uniformly from $\Delta\mathbb{J}_i \in [-\Delta\mathbb{J}, \Delta\mathbb{J}]$. Periodic boundary conditions are applied. The magnetic field term $\mathbb{h}_i = \mathbb{h}_0$ is constant. Throughout this work we choose $\mathbb{h}_0 = 0.6$, $\mathbb{J}_2 = 0.3$, $\mathbb{J}_0 = 1$, and $\Delta\mathbb{J} = 1$ deeply in the chaotic regime.

The real time dynamics of the system as well as the severity of its sign problem are captured in the spectral form factor (SFF)

$$\mathcal{K}(t) \equiv \langle K(t) \rangle_{\mathbb{J}} \tag{8}$$

$$\equiv \left\langle |\operatorname{Tr} U(t)|^2 \right\rangle_{\mathbb{J}} \tag{9}$$

where $\langle \cdot \rangle_{\mathbb{J}}$ denotes the average over different disorder realisations and $U(t)$ is the real time evolution operator for given disorder $\mathbb{J}$

$$\begin{aligned} U(t) = \mathrm{e}^{-\mathrm{i}\, H t} &= \prod_{j=1}^{N_t} \mathrm{e}^{-\mathrm{i}\, \delta H} \\ &= \prod_{j=1}^{N_t} \mathrm{e}^{\mathrm{i}\, \delta \sigma^z \cdot \mathbb{J} \cdot \sigma^z}\, \mathrm{e}^{\mathrm{i}\, \delta \mathbb{h} \cdot \sigma^x} + \mathcal{O}(\delta) \end{aligned} \tag{10}$$

with the physical evolution time t and the Suzuki-Trotter discretization step size $\delta \equiv t/N_t$. This Suzuki-Trotter decomposition becomes exact in the limit where the number of time steps $N_t \to \infty$.

The transfer-matrix method allows to relate the Trotterised version to a two-dimensional system of classical spins[2]

$$K(t) \propto Z \equiv \sum_{\{s\}} \mathrm{e}^{-S(s)}\,, \tag{11}$$

where we sum over all possible spin $s_{i,k} = \pm 1$ combinations, arriving at the action

$$\begin{aligned} S = &-\mathrm{i} \sum_k \sum_{i,j} s_{i,k} J_{ij} s_{j,k} \\ &- \sum_i \sum_k \left(h_i - \frac{\pi}{4}\mathrm{i} \right) s_{i,k} s_{i,k+1}\,, \end{aligned} \tag{12}$$

$$J_{ij} := \delta \mathbb{J}_{ij}\,, \quad h_i := -\frac{1}{2} \log \tan\left(\delta \mathbb{h}_i\right)\,. \tag{13}$$

J and h_i are real constants. See appendix A for details on the derivation.

Any spin flip results in a change of S by $\lambda\left(h_i - \frac{\pi}{4}\mathrm{i}\right)$ with $\lambda \in \{\pm 4, 0\}$ and some contributions of order δ. Therefore both, real and imaginary part, are significantly changed even for small δ. In particular, $\lambda = \pm 4$ flips the sign of the action.

We split the action into three separate parts

$$S = \mathrm{i}\, S_I + S_R + \mathrm{i}\, \pi S_S\,, \tag{14}$$

$$S_I := -\sum_k \sum_{i,j} s_{i,k} J_{ij} s_{j,k} \tag{15}$$

$$S_R := -\sum_i \sum_k h_i s_{i,k} s_{i,k+1} \tag{16}$$

$$S_S := \frac{1}{4} \sum_i \sum_k s_{i,k} s_{i,k+1}\,. \tag{17}$$

Now each of S_I, S_R and S_S is real. Furthermore S_S is an integer and therefore

$$\mathrm{e}^{\mathrm{i}\, \pi S_S} = \pm 1 \tag{18}$$

contributes only a sign to the total measure.

[2] Such a system can also be interpreted as an anisotropic classical Ising model with complex coupling coefficients.

A. Probability weighted density of states

Monte-Carlo based algorithms can easily implement the probability density induced by S_R whereas the complex phase and real sign contributions by S_I and S_S respectively pose severe issues. As outlined in equations (2)-(4) in the introductory Section, in this work we employ the density of states (DoS) approach, rewriting the partition sum as

$$Z = \sum_{\{s\}} \mathrm{e}^{-S_R} \mathrm{e}^{-\mathrm{i}\, S_I} \mathrm{e}^{-\mathrm{i}\,\pi S_S} \tag{19}$$

$$= \sum_{\{s\}} \mathrm{e}^{-S_R} \int_{\mathbb{R}} \mathrm{d}E\, \mathrm{e}^{-\mathrm{i}\, E}\, \delta(E - S_I) \sum_{z=\pm 1} z\, \delta_{z,\mathrm{e}^{-\mathrm{i}\,\pi S_S}} \tag{20}$$

$$= \sum_{z=\pm 1} z \int_{\mathbb{R}} \mathrm{d}E\, \mathrm{e}^{-\mathrm{i}\, E} \sum_{\{s\}} \mathrm{e}^{-S_R}\, \delta(E - S_I)\, \delta_{z,\mathrm{e}^{-\mathrm{i}\,\pi S_S}} \tag{21}$$

$$= \sum_{z=\pm 1} z \int_{\mathbb{R}} \mathrm{d}E\, \mathrm{e}^{-\mathrm{i}\, E}\, \rho_z(E)\,. \tag{22}$$

In the last step we defined the DoS $\rho_z\,(E)$ as

$$\rho_z(E) \coloneqq \sum_{\{s\}} \mathrm{e}^{-S_R}\, \delta(E - S_I)\, \delta_{z,\mathrm{e}^{-\mathrm{i}\,\pi S_S}}\,. \tag{23}$$

B. Normalisation

Let us define the auxiliary partition sum

$$Z_R = \sum_{\{s\}} \mathrm{e}^{-S_R} \tag{24}$$

$$= \sum_{z=\pm 1} \int_{\mathbb{R}} \mathrm{d}E\, \rho_z(E) \tag{25}$$

which is a sum of manifestly positive contributions and therefore does not suffer from a sign problem. Note that all terms containing the coupling $\mathbb{J}$ have been dropped and Z_R therefore completely decouples into its one-dimensional (time-like) parts, i.e.

$$Z_R = \prod_i Z_{R,i}\,, \tag{26}$$

$$Z_{R,i} = \mathrm{Tr}\left[(T_2)_i^{N_t}\right] \tag{27}$$

$$= 2^{N_t}\left(\cosh^{N_t} h_i + \sinh^{N_t} h_i\right)\,, \tag{28}$$

where the last line is the well known solution of the one-dimensional classical Ising model and can be obtained by diagonalising the 2×2 transfer matrices $(T_2)_i$ from equation (A8).

In the full simulation one would usually normalise the sum over the DoS to one so that the Monte-Carlo estimator (denoted by $\langle\cdot\rangle$) of the partition sum becomes

$$\langle Z \rangle = Z_R \left\langle \frac{Z}{Z_R} \right\rangle\,, \tag{29}$$

where Z_R is known analytically and $\left\langle \frac{Z}{Z_R} \right\rangle$ can be found by evaluating equation (22) with the normalised estimator of $\rho_z(E)$. Combining this result with equation (A14), we finally obtain a formula for the trace of the quantum mechanical time evolution operator

$$\mathrm{Tr}\, U(t) = A Z_R \left\langle \frac{Z}{Z_R} \right\rangle\,. \tag{30}$$

III. THE ALGORITHMS

The canonical choice and baseline algorithm for Monte-Carlo simulations with a sign problem is *reweighting*. In this approach all contributions stemming from the imaginary part of the action are simply considered part of the observable. In this particular case this means that configurations are sampled purely from the probability distribution $\mathbb{p}(S_R)$ induced by S_R (using for instance local Metropolis-Hastings updates) and the partition sum is estimated by

$$\left\langle \frac{Z}{Z_R} \right\rangle = \frac{\sum_{\{s\}} \mathrm{e}^{-S_R} \mathrm{e}^{-\mathrm{i}\, S_I} \mathrm{e}^{-\mathrm{i}\,\pi S_S}}{\sum_{\{s\}} \mathrm{e}^{-S_R}} \tag{31}$$

$$= \sum_{\{s\}} \mathbb{p}(S_R)\, \mathrm{e}^{-\mathrm{i}\, S_I} \mathrm{e}^{-\mathrm{i}\,\pi S_S} \tag{32}$$

$$= \left\langle \mathrm{e}^{-\mathrm{i}\, S_I - \mathrm{i}\,\pi S_S} \right\rangle_{S_R}\,. \tag{33}$$

Another natural approach is to estimate the DoS ρ prior to determining any observables. This can be done efficiently using the Logarithmic Linear Relaxation (LLR) algorithm. The principle idea of LLR was first proposed by Wang and Landau [15] and applied to the classical Ising model with uniform coupling in two dimensions. Later on it has been optimised through the introduction of the $1/t$ algorithm [16, 17] in accordance with the theoretically optimal scheme derived by Robbins and Monro [18].

Let us consider a classical Ising system of $N \equiv LN_t$ spins. The auxiliary variable E (typically identified as an 'energy', see eq. (23)) is divided into M bins E_i, $i = 1, \ldots, M$, so that we discretise the DoS

$$\rho_i \coloneqq \int_{\frac{E_i + E_{i-1}}{2}}^{\frac{E_i + E_{i+1}}{2}} \mathrm{d}E\, \rho(E)\,, \quad i = 2, \ldots, M-1 \tag{34}$$

and for $i = 1$ ($i = M$) the lower (upper) bound of the integral takes the minimal (maximal) value of E. Of course, the arithmetic mean can be exchanged for any other value between E_i and its neighbours without changing the general argument. Discretisation errors are of

order $\mathcal{O}\left(h^2\right)$ and further details are discussed in appendix B.

We write

$$\rho_i \equiv \mathrm{e}^{\alpha_i} \tag{35}$$

and from now on we aim to estimate the α_i as accurately as possible (hence the 'logarithmic' in the method's name). Starting from uniform initial conditions ($\alpha_i = 0$ for all i), we employ an update scheme of Λ steps in total where at the kth step some configuration update $s \mapsto s'$ is proposed and accepted with the probability

$$p_{\mathrm{acc}} = \frac{\mathrm{e}^{-S_R[s']}}{\mathrm{e}^{-S_R[s]}} \cdot \frac{\rho_{i(s)}}{\rho_{i(s')}} \tag{36}$$

$$= \mathrm{e}^{S_R[s]-S_R[s']}\, \mathrm{e}^{\alpha_{i(s)}-\alpha_{i(s')}}\,, \tag{37}$$

where $i\,(s)$ and $i\,(s')$ are the bins that correspond to the imaginary part of the action S_I on configurations s and s', respectively. The DoS of the accepted bin, say i w.l.o.g., has to be updated

$$\alpha_i \mapsto \alpha_i + \beta_{N,M,\Lambda}(k,i)\,, \tag{38}$$

where $\beta \geq 0$. Without prior knowledge of the DoS we drop the explicit dependence on i. Furthermore we know from Ref. [18] that the conditions

$$\sum_{k=0}^{\infty} \beta_{N,M,\Lambda}(k) = \infty\,, \tag{39}$$

$$\sum_{k=0}^{\infty} \beta_{N,M,\Lambda}(k)^2 < \infty \tag{40}$$

have to hold.[3] In particular, they propose a function asymptotically scaling as $1/k$ since this is the fastest decaying function (up to logarithmic factors) of the required class and therefore promises best convergence for large k.

The coefficients in the explicit form

$$\beta_{N,M,\Lambda}(k) = \frac{a_{N,M,\Lambda}}{b_{N,M,\Lambda}+k} \tag{41}$$

are of crucial importance and we derive in appendix C why

$$a_{N,M,\Lambda} = \ln 2\, \frac{MN}{\ln \frac{\Lambda}{b_{N,M,\Lambda}}}\,, \tag{42}$$

$$b_{N,M,\Lambda} = 3M \tag{43}$$

are a particularly good choice. We are going to use them throughout the rest of this work.

LLR iterations do not form a Markov chain because the accept/reject probability depends on the history of the previous updates. Therefore statistical errors cannot be calculated via straight forward bootstrap procedures or similar. Instead every simulation presented hereafter has been repeated $N_{\mathrm{runs}} = 40$ times and the error has been estimated from the resulting distribution. In practice LLR simulations are not always stable but are prone to outliers. To counteract this effect we used the median as best estimator for observables instead of the more commonly used mean. Similarly the errors have been approximated using the 16% and 84% quantiles instead of the standard deviation. For a normal distribution this choice of median and quantiles is fully compatible with the default choice of mean and standard deviation, but for other distributions it is more stable.

IV. SMALL TRANSVERSE COMPONENT EXPANSION

The representation (11) of the real-time partition function in terms of a finite sum over classical Ising spins implies a finite number of states, so that technically speaking the DoS (23) is a collection of delta peaks. However, in the presence of quenched disorder, there are in general $\mathcal{O}\left(2^L\right)$ delta peaks spread randomly over the range of width $\mathcal{O}\left(L\right)$ of possible values of S_I. As a result, the bulk of the DoS approaches the continuous function in the thermodynamic limit $L \to +\infty$. For finite length L, the apparent continuousness is a mere consequence of the finite resolution following from the fixed number M of bins defining the DoS, i.e. the true distribution is convoluted with a function that is 1 within a bin and 0 everywhere else.

Looking at a typical distribution of the DoS (see fig. 1), it appears to be fractured into bands. The top one corresponds to the even sector and consists of a few very narrow peaks. The second (odd) band is closer to a smooth distribution, though it still features some irregularities. An alternating descend of smooth bands with roughly equal step size on a logarithmic scale follows. With the technically available precision we usually cannot resolve any but the four highest bands.

It is straight forward to interpret these bands, as they directly correspond to a particular parity sector each. Every spin-flip pair in temporal direction (see fig.2) changes this sector. It comes with a sign flip and a Boltzmann suppression, i.e.

$$S_S \mapsto S_S - 1\,, \tag{44}$$

$$S_R \mapsto S_R + 4h\,, \tag{45}$$

$$\Rightarrow p \mapsto -p\left(\tan\frac{t\hbar}{N_t}\right)^2\,, \tag{46}$$

where $p = \mathrm{e}^{-S_R - \mathrm{i}\,\pi S_S}$ is the sign-full 'probability' weight and $S_{S,R}$ as in equation (14). This provides a canonical way of expanding the DoS $\rho_z(E)$ and the partition sum Z in powers of $\tan\frac{t\hbar}{N_t}$, that is around small times t or transverse fields $\hbar$.

[3] Actually these are not exactly the conditions from [18], but they are equivalent to this simpler version.

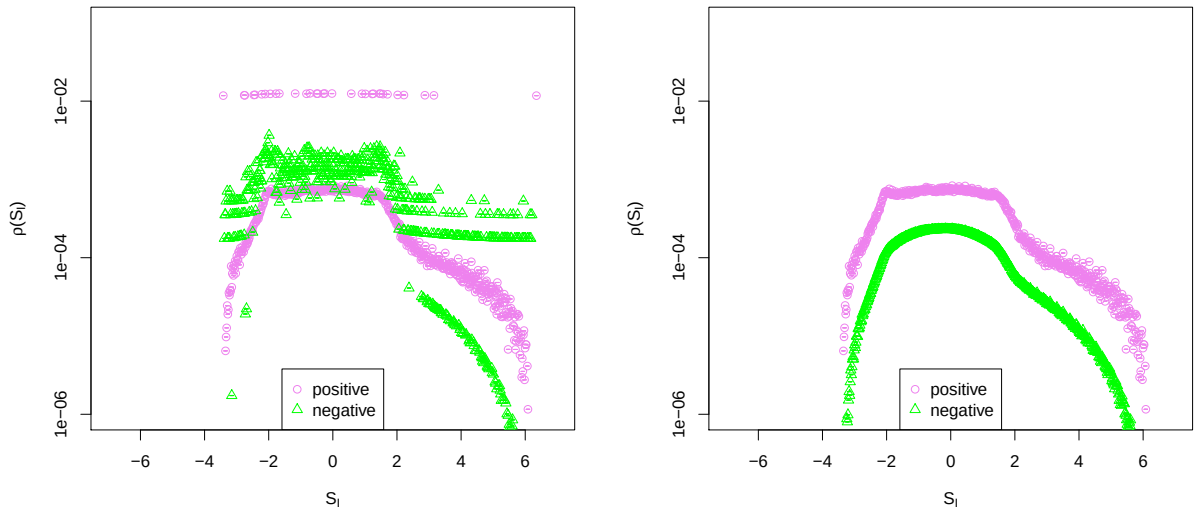

Figure 1. Densities of states in the even (positive) and odd (negative) parity sectors respectively, calculated with LLR_0 (left) and LLR_2 (right). The spin chain length is $L = 6$, the evolution time is $t = 1$, the number of Monte-Carlo sweeps is $N_s = 10^8$.

From now on we will denote simulations restricted to parity sectors with at least n spin-flip pairs by LLR_n (or REW_n for reweighting). In practice this means that the acceptance probability in equation (34) is set to zero for any proposed configuration with less than n spin-flip pairs. Otherwise the sampling algorithm proceeds unchanged.

Thus, the "plain vanilla" algorithm allowing to visit the full phase space will be denoted LLR_0 and a lot of our studies will be focused on LLR_2 as a good trade off between precision and performance. The remaining contributions from all the configurations with less than n spin-flip pairs are calculated analytically as follows.

A. Leading order: $n = 0$ spin-flip pairs

Without any spin flips in temporal direction the phase space is confined to one of the 2^L spatial configurations repeated N_t times. Therefore, the DoS and the classical partition sum reduce to

$$\rho_z(E)|_{n=0} \propto \delta_{z,1} \sum_{\{s\}} \delta\left(E - t\sum_{i,j} s_i \mathbb{J}_{ij} s_j\right) , \tag{47}$$

$$Z|_{n=0} = \sum_{\{s\}} \mathrm{e}^{\mathrm{i}\, N_t \sum_{i,j} s_i J_{ij} s_j + L N_t h} \tag{48}$$

$$= \left(\tan\frac{t\hbar}{N_t}\right)^{-\frac{1}{2}LN_t} \sum_{\{s\}} \mathrm{e}^{\mathrm{i}\, t \sum_{i,j} s_i \mathbb{J}_{ij} s_j} \tag{49}$$

and thus

$$\mathrm{Tr}\, U(t)|_{n=0} = \left(\frac{\sin\delta\hbar\cos\delta\hbar}{\tan\delta\hbar}\right)^{\frac{1}{2}LN_t} \sum_{\{s\}} \mathrm{e}^{\mathrm{i}\, t \sum_{i,j} s_i \mathbb{J}_{ij} s_j} \tag{50}$$

$$= (\cos\delta\hbar)^{LN_t} \sum_{\{s\}} \mathrm{e}^{\mathrm{i}\, t \sum_{i,j} s_i \mathbb{J}_{ij} s_j} . \tag{51}$$

The prefactor $Z_0 \equiv (\cos\delta\hbar)^{LN_t} \xrightarrow{\delta\to 0} 1$ is clearly an artefact from Trotterisation, but it cannot be dropped in realistic simulation since it crucially counteracts finite δ effects in the LLR_n simulation itself.

Here we assume that the couplings $\mathbb{J}_{ij}$ are non-uniform and hence the full sum over all configurations $\{s\}$ has to be calculated. In particular simpler cases, e.g. $\mathbb{J}_{ij} = \mathbb{J}\,\delta_{i+1,j}$, the sum can be simplified and therefore the computational complexity reduced dramatically.

Note that the continuum limit of this result corresponds to the $\hbar = 0$ approximation of the quantum mechanical time evolution operator

$$\mathrm{Tr}\, U(t)|_{n=0} = \mathrm{Tr}\left[\mathrm{e}^{\mathrm{i}\, t\sigma^z\cdot\mathbb{J}\cdot\sigma^z}\, \mathrm{e}^{\sigma^x\, \mathcal{O}(ht)}\right] \tag{52}$$

$$= \sum_{\{s\}} \mathrm{e}^{\mathrm{i}\, t \sum_{i,j} s_i \mathbb{J}_{ij} s_j} + \mathcal{O}\left((ht)^2\right) , \tag{53}$$

where the linear order in ht can be dropped because $\mathrm{Tr}\,\sigma^x = 0$.

B. Next-to-leading odd order: $n = 1$ spin-flip pair

Let us now consider the case of a single spin-flip pair, that is exactly one of the quantum spins has exactly two

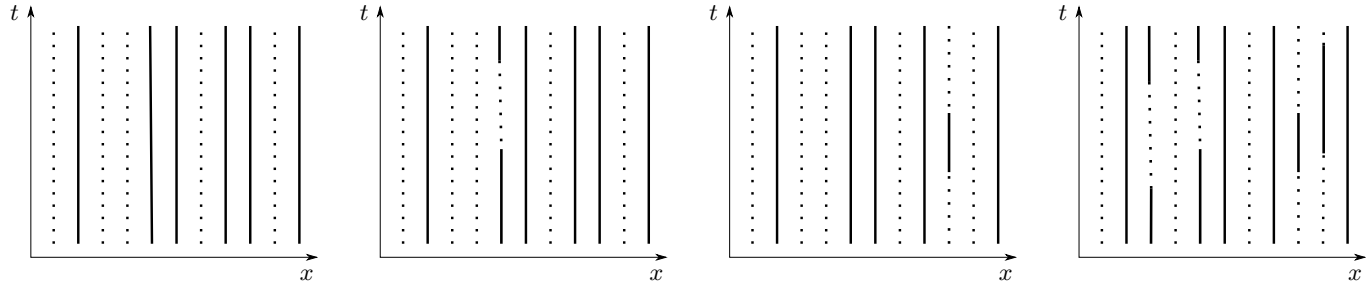

Figure 2. Visualisation of configurations with different numbers n of spin-flip pairs in the time continuum. Periodic boundary conditions are applied, so flips come in pairs only. W.l.o.g. the solid lines represent spin-up and the dotted lines spin-down. From left to right: $n = 0$, $n = 1$, $n = 1$, $n = 4$.

different values over time in the z-basis. Some visualisations can be found in figure 2. We absorb the probability weight $-(\tan\delta\mathbb{h})^2$ and the combinatorial factor $N_t/2$ (temporal translational invariance yields a factor N_t, exchangeability of start and end point of the flipped region a factor 1/2) into a new prefactor

$$Z_1 \equiv -\frac{N_t}{2}\left(\tan\delta\mathbb{h}\right)^2 Z_0\,. \tag{54}$$

Now for every configuration s without flips there are L different candidates l for a flip with $N_t - 1$ different possible lengths τ each. We can write the resulting phase as

$$\begin{aligned} S_I|_{n=1} &= -\left(N_t-\tau\right)\sum_{i,j} s_i J_{ij} s_j - \\ &\quad -\tau\sum_{i,j} s_i\left(1-2\delta_{il}\right)J_{ij}\left(1-2\delta_{lj}\right)s_j \\ &= -\left(t\sum_{i,j} s_i \mathbb{J}_{ij} s_j - 4\tau\sum_j s_l J_{lj} s_j\right)\,. \end{aligned} \tag{55}$$

Combining the pre-factor and the phase leads to

$$\rho_z(E)|_{n=1} = Z_1\,\delta_{z,-1}\sum_{\{s\}}\sum_l\sum_{\tau=1}^{N_t-1}\delta\left(E - t\sum_{i,j} s_i\mathbb{J}_{ij}s_j + 4\tau\sum_j s_l J_{lj} s_j\right)\,, \tag{56}$$

$$\operatorname{Tr} U(t)|_{n=1} = Z_1\sum_{\{s\}} \mathrm{e}^{\mathrm{i}\,t\sum_{i,j} s_i\mathbb{J}_{ij}s_j}\sum_l\sum_{\tau=1}^{N_t-1}\mathrm{e}^{-4\,\mathrm{i}\,\tau\sum_j s_l J_{lj} s_j} \tag{57}$$

$$= Z_1\sum_{\{s\}} \mathrm{e}^{\mathrm{i}\,t\sum_{i,j} s_i\mathbb{J}_{ij}s_j}\sum_l \mathrm{e}^{-2\,\mathrm{i}\,t\sum_j s_l\mathbb{J}_{l_j}s_j}\,\frac{\sin\left(2\left(N_t-1\right)\sum_j s_l J_{lj}s_j\right)}{\sin\left(2\sum_j s_l J_{lj} s_j\right)} \tag{58}$$

with the well-defined continuum limit

$$\rho_z(E)|_{n=1} = \frac{(ht)^2}{2}\,\delta_{z,-1}\sum_{\{s\}}\sum_l\theta\left(-\left|E - t\sum_{i,j} s_i\mathbb{J}_{ij}s_j\right| + \left|4\tau\sum_j s_l J_{lj} s_j\right|\right)\,, \tag{59}$$

$$\lim_{\delta\to 0}\operatorname{Tr} U(t)|_{n=1} = -\frac{(ht)^2}{2}\sum_{\{s\}} \mathrm{e}^{\mathrm{i}\,t\sum_{i,j} s_i\mathbb{J}_{ij}s_j}\sum_l \mathrm{e}^{-2\,\mathrm{i}\,t\sum_j s_l\mathbb{J}_{l_j}s_j}\,\operatorname{sinc}\left(2t\sum_j s_l\mathbb{J}_{lj}s_j\right)\,, \tag{60}$$

where θ denotes the Heaviside step function.

Note that the temporal sum in (55) could be calculated exactly since it reduced to a geometric series. This allows an evaluation of the remaining sum in $\mathcal{O}\left(L2^L\right)$, i.e. the same runtime as the contribution without flips.

Again we could have obtained the continuum result up to leading order directly from the time evolution operator using the Zassenhaus formula

$$\operatorname{Tr} U(t)|_{n=1} = \operatorname{Tr}\left[\mathrm{e}^{\mathrm{i}\,t\sigma^z\cdot\mathbb{J}\cdot\sigma^z}\mathrm{e}^{\mathrm{i}\,t\mathbb{h}\sum_i\sigma_i^x}\right] + \mathcal{O}\left((ht)^4\right) \tag{61}$$

$$= \mathrm{Tr}\left[\mathrm{e}^{\mathrm{i}\, t\sigma^z\cdot\mathbb{J}\cdot\sigma^z}\prod_i\left(\mathbb{1} + \mathrm{i}\, ht\sigma_i^x - \frac{1}{2}(ht)^2\mathbb{1}\right)\right] + \mathcal{O}\left((ht)^4\right) \tag{62}$$

$$= \left(1 - \frac{1}{2}L(ht)^2\right)\mathrm{Tr}\left[\mathrm{e}^{\mathrm{i}\, t\sigma^z\cdot\mathbb{J}\cdot\sigma^z}\right] + \mathcal{O}\left((ht)^4\right). \tag{63}$$

Here the cyclic property of the trace ensures that the simple first-order decomposition into σ^z and σ^x parts is equivalent to a symmetric decomposition with an error of order $\mathcal{O}\left((ht)^3\right)$. Additionally $\mathrm{Tr}(\sigma^x)^3 = 0$ reduces the residual error to $\mathcal{O}\left((ht)^4\right)$.

C. Higher orders: $n \geq 2$ spin-flip pairs

In principle, we can follow the steps of the previous section in order to write down the contribution of the nth order (n spin flips) for arbitrary n. The prefactor (without combinatorics) reads

$$Z_n \equiv \left(-\tan^2\delta\hbar\right)^n Z_0 \tag{64}$$

and the phase (including combinatorics) generalises to

$$S_I = \sum_k\sum_{i,j} s_i\left(\prod_l\left(1 - 2\delta_{il}\theta(k-\tau_1^l)\theta(\tau_2^l - k)\right)\right) \times J_{ij}\left(\prod_l\left(1 - 2\delta_{lj}\theta(k-\tau_1^l)\theta(\tau_2^l - k)\right)\right)s_j \tag{65}$$

where θ denotes the Heaviside step function and $\tau_{1,2}^l$ are successive flipping times of spins at spatial position l. The partition sum has to include all possible combinations of $\{l\}$ and corresponding $\tau_{1,2}^l$.

The evaluation of the partition sum for any $n \geq 2$ however is prohibitively expensive because the temporal sum can no longer be evaluated analytically. Two different regions flipped at overlapping times introduce non-trivial entanglement (see right panel of fig. 2). Another way to see this is the fact that the first-order Trotter decomposition of the time evolution operator results in an error of order $\mathcal{O}\left((ht)^4\right)$. A higher order treatment therefore necessitates more complicated decompositions with non-zero commutators taken into account.

D. Spline fitting

In LLR_2 simulations all the leading peaks have been removed from the DoS so that the remaining function is relatively smooth and can be approximated by a fit. Since we do not know anything about the global structure of the DoS in general, a smooth local approximation like a cubic spline[4] is the canonical choice. Such a spline with a fixed number of free parameters or degrees of freedom (d.o.f.), where d.o.f. smaller or equal the number of data points, is fitted to the log-DoS as depicted in figure 3. The fit can then be numerically integrated to high precision.

A smooth fit comes with two advantages. First, discretisation errors in S_I due to a finite bin size in (32) are reduced. In practice this effect is usually negligible and if it is not, a finer discretisation should be chosen in the first place. More relevantly, fluctuations of neighbouring bins with large uncertainties are averaged out. This allows decent estimations of the spectral form factor (SFF) ~~SFF~~ at relatively high noise levels.

The single but crucial disadvantage of fitting the data is that a priori the optimal number k_{dof} of d.o.f. is completely unclear. In figure 3 we show the same data fitted with different order splines to the effect that the leftmost is under-fitted and will yield completely wrong results, whereas the rightmost is over-fitted and the desired smoothing effect is absent. We use the Bayesian information criterion (BIC) [35]

$$\mathrm{BIC} = M\ln\left(\chi^2/M\right) + k_{\mathrm{dof}}\ln(M) \tag{66}$$

to find the optimal number of d.o.f. Here M is again the number of bins of the DoS and χ^2 is the usual chi-squared value. The BIC is supposed to be minimal for the best choice of k_{dof}.

Overall the fitting allows to extract results from otherwise too noisy data, but they come at the cost of additional uncertainty and possible bias so that high-precision data is better evaluated without any fitting. This becomes clear from the comparison of the two top plots on Fig. 5, where the errors obtained from spline-fitted results are only considerably smaller for small numbers of Monte-Carlo updates.

E. Origin of the non-smooth DoS

The highly irregular shape of the DoS is the crucial difference to the DoS of models with continuous (and compact) variables where the LLR approach has been shown to outperform reweighting significantly [11]. It is therefore important to understand the origin of this shape. Naïvely one might think that it has to do with the random coupling and the chaotic nature of the system, but it turns out that this is not the case. Even the simplest transverse Ising model with constant coupling and no next-to-nearest neighbour interaction features a non-smooth DoS.

Qualitatively the origin of the DoS's strange form can be understood starting with the spikes in LLR_0 forming the highest 'band'. In the case when the nearest-neighbor

[4] We use the `smooth.spline` implementation in `R` [34].

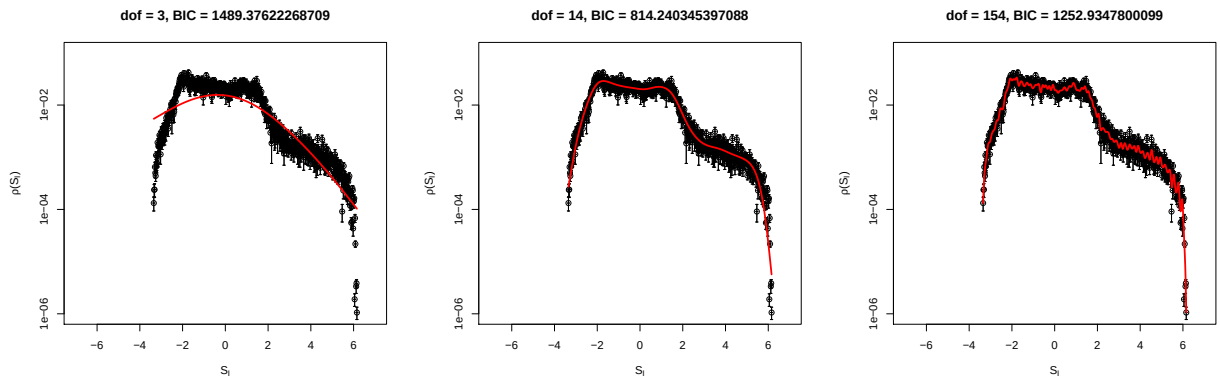

Figure 3. Spline fits to $\log\left(\rho_+(S_I) - \rho_-(S_I)\right)$ with different degrees of freedom (d.o.f.) for 538 data points and respective Bayesian information criterion (BIC), eq. (64). The spin chain length is $L = 6$, the evolution time is $t = 1$, the number of Monte-Carlo sweeps is $N_s = 10^4$.

couplings contain quenched disorder $\Delta\mathbb{J}_i$, there are 2^{L-1} spikes of equal height. In the case of constant couplings, for example, for the conventional transverse field quantum Ising model without quenched disorder, there are $\lfloor L/2 \rfloor + 1$ spikes with binomially distributed heights. Now the second band, i.e. the single flip-pair expansion, can be seen (up to a proportionality factor) as a convolution of the first band with a rectangular function, taking the value 1 for every S_I that can be reached by a single flip-pair and the value 0 everywhere else. Thus every delta-peak smears out to a superposition of rectangle-shaped functions at the same position, as can be seen in equation (57). Its width relative to the total range of S_I scales as $\mathcal{O}(1/L)$. The generalisation is straight forward. With every additional flip-pair the previous distribution is convoluted with a rectangular function. For instance the DoS of LLR_2 looks like the original spikes of LLR_0 convoluted with a triangle function, yielding the 'batman' profile in figure 4 for the simple transverse field Ising model with constant coupling $\mathbb{J}_{ij} = \delta_{i,j+1}$ and $L = 6$.

Note that the notion of successive convolutions is only exact up to a single flip-pair. Beyond, it still allows to understand the shape of the DoS qualitatively but crucial contributions from interferences between two or more flip pairs are neglected (e.g. two flip pairs might be on the same spatial site) so that the DoS in not reproduced correctly.

With larger L the DoS smooths out in the disordered case. The overall smoothing stems from an exponentially large number of peaks constituting LLR_0, not from a broader smearing of the lower bands, since the convolution always acts on scales of $\mathcal{O}(1/L)$.

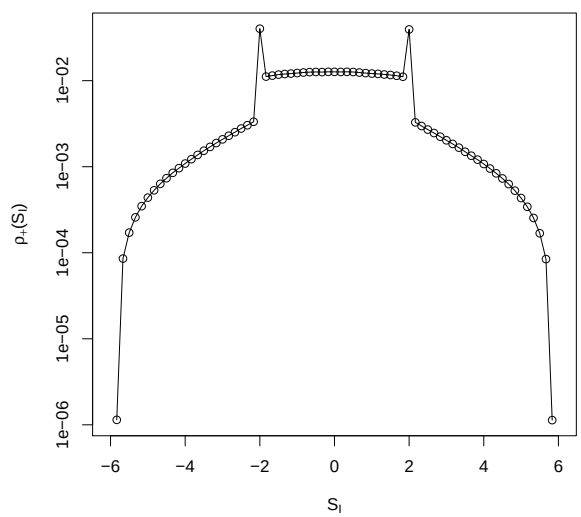

Figure 4. Density of states in the even (positive) parity sector, calculated with LLR_2 for the pure transverse Ising model with constant coupling, i.e. $\mathbb{J}_2 = \Delta\mathbb{J} = 0$. Chain length $L = 6$; time, field strength and coupling $t = \hbar = \mathbb{J}_0 = 1$; number of Monte-Carlo sweeps $N_s = 10^7$.

V. RUNTIME COMPARISON

A. Exact diagonalisation

The straight forward solution for arbitrary times is the exact diagonalisation of the full 2^L-dimensional system. Since the diagonalisation of a matrix has cubic scaling, we end up with a runtime in $\mathcal{O}\left(2^{3L}\right)$.

In our benchmarks we used the `LAPACKE_dsyev` routine provided by the `lapacke` package in `C` to calculate the complete spectrum of the respective Hamiltonian to ma-

chine precision.

For completeness we also mention that the brute force approach of simply trying all the combinations in the classical system to calculate the exact partition sum has a runtime of $\mathcal{O}\left(2^{LN_t}\right)$.

B. Reweighting

Stochastic methods like reweighting cannot be compared to exact methods like ED directly since the computational effort crucially depends on the desired precision ε. It is, however, reasonable to expect square-root convergence in the amount of statistics due to the central limit theorem and to set a sweep over the full lattice of $N = L\,N_t$ sites as the smallest unit of data. With these assumptions we get a runtime scaling of $\mathcal{O}\left(LN_t\varepsilon^{-2}\right)$ for a system without a sign problem. It is important to keep in mind that this is the best case scenario and phenomena like critical slowing down are completely neglected.

For the sake of simplicity we are going to assume a small enough and constant Trotter step size δ, so that $N_t \propto t$. Note that this implies exponential runtime even without a sign problem for physically interesting long times $t \sim 2^L$.

The crucial change we have to introduce for systems with a sign problem is a factor of the statistical power

$$\Sigma = \left|\frac{Z}{Z_R}\right| , \tag{67}$$

that is the modulus of the expectation value of the complex phase. This modification leads to a runtime scaling of

$$T_{\text{sim}} \sim \mathcal{O}\left(Lt\varepsilon^{-2}\Sigma^{-2}\right) , \tag{68}$$

so in the following we have to estimate the magnitude of Σ as best we can. It is clear that we cannot calculate it exactly since that would allow to solve the complete initial quantum mechanical problem.

We start with the observation that

$$|\mathrm{Tr}\,U(t)| = |AZ_R|\;\Sigma\,, \tag{69}$$

where the left hand side can be estimated from universal properties of the SFF~~spectral form factor (SFF)~~ and the product AZ_R is known analytically. Let us expand the latter product near the continuum limit

$$\begin{aligned}
|AZ_R| &= \prod_i \left(\sin\delta\mathbb{h}_i\cos\delta\mathbb{h}_i\right)^{N_t/2}\left(\cosh\left(-\frac{1}{2}\ln\tan\delta\mathbb{h}_i\right)^{N_t} + \sinh\left(-\frac{1}{2}\ln\tan\delta\mathbb{h}_i\right)^{N_t}\right) && (70)\\
&= \prod_i \left(\delta\mathbb{h}_i\right)^{N_t/2}\left(\left(\frac{1}{\sqrt{\delta\mathbb{h}_i}} + \sqrt{\delta\mathbb{h}_i}\right)^{N_t} + \left(\frac{1}{\sqrt{\delta\mathbb{h}_i}} - \sqrt{\delta\mathbb{h}_i}\right)^{N_t}\right)\left(1+\mathcal{O}\left(\delta\mathbb{h}_i\right)\right) && (71)\\
&= \prod_i \left(\left(1+\delta\mathbb{h}_i\right)^{N_t} + \left(1-\delta\mathbb{h}_i\right)^{N_t}\right)\left(1+\mathcal{O}\left(\delta\mathbb{h}_i\right)\right) && (72)\\
&= \prod_i \left(\left(1+\frac{t\mathbb{h}_i}{N_t}\right)^{N_t} + \left(1-\frac{t\mathbb{h}_i}{N_t}\right)^{N_t}\right)\left(1+\mathcal{O}\left(\delta\mathbb{h}_i\right)\right) && (73)\\
&= \prod_i 2\cosh\left(t\mathbb{h}_i\right)\left(1+\mathcal{O}\left(\delta\mathbb{h}_i\right)\right) . && (74)
\end{aligned}$$

In the common case of all fields equal $\mathbb{h}_i \equiv \mathbb{h}$ we thus obtain the formula

$$\Sigma = \frac{|\mathrm{Tr}\,U(t)|}{2^L\cosh\left(t\mathbb{h}\right)^L} , \tag{75}$$

which is exact as $\delta \to 0$.

Since the SFF is bound from above $|\mathrm{Tr}\,U(t)| \leq 2^L$, the best possible runtime scaling amounts to

$$T_{\text{sim}} \sim \mathcal{O}\left(Lt\varepsilon^{-2}\cosh\left(t\mathbb{h}\right)^{2L}\right) , \tag{76}$$

so that reweighting can never outperform exact diagonalisation when $t\mathbb{h} \gtrsim \operatorname{acosh} 2 \approx 1.3$.

Empirically we find that for small times $|\mathrm{Tr}\,U(t)| \sim 2^L\mathrm{e}^{-Lt^2}$ until it reaches the plateau at $2^{L/2}$. Moreover, in order to maintain a high precision, usually we have to scale the Trotter step with the system size $\delta \sim 1/L$, so that a more realistic runtime estimate is given by

$$T_{\text{sim}} \sim \mathcal{O}\left(\frac{1}{2^{-L}+\mathrm{e}^{-Lt^2}}\,L^2t\varepsilon^{-2}\cosh\left(t\mathbb{h}\right)^{2L}\right) . \tag{77}$$

C. LLR

In what follows we show that for both reweighting and LLR the error is strongly dominated by the ‘bulk’, that

is the region with highest DoS. Since LLR is only advantageous in sampling the 'tails' of the DoS distribution more accurately, it cannot significantly outperform reweighting. This conclusion is also confirmed by our Monte-Carlo data.

For a more formal comparison we have to introduce several assumptions. First of all we demand that all the LLR parameters have been tuned properly. In particular the bin size used for LLR is small enough to have negligible discretisation errors. Such a choice is always possible though the need to tune the parameters is a serious disadvantage of LLR. Secondly, we bin the data accumulated for reweighting in the same fashion, again without significant loss of precision. Now the normalised DoS obtained by either method can be interpreted as a probability distribution ρ_i on the same discrete and finite set $i \in \{1, \dots, M\}$. Same as in equation (32), M denotes the number of bins the DoS has been split into.

It is easy to see that with reweighting every bin i is visited $N_i = \rho_i \Lambda$ times on average, where Λ is the total number of samples. Therefore, in the large Λ limit, according to the central limit theorem (or naturally assuming a Poisson distribution of bin counts), the absolute uncertainty of the number of counts in said bin is $\sqrt{N_i}$ yielding an error of

$$\Delta\rho_i \equiv \frac{\rho_i}{\sqrt{N_i}} \tag{78}$$

$$= \sqrt{\frac{\rho_i}{\Lambda}} \,. \tag{79}$$

Since LLR is a special case of the Robbins-Monro (RM) algorithm with a non-differentiable target function, the coefficients in the exponent converge strictly as $\mathcal{O}\left(1/\sqrt{\Lambda}\right)$ [36]. Thus, the relative error of ρ_i lies in $\mathcal{O}\left(\mathrm{e}^{1/\sqrt{\Lambda}} - 1\right) = \mathcal{O}\left(1/\sqrt{\Lambda}\right)$. We can obtain some bounds on the proportionality factor as well, though these results are by no means as precise as for reweighting.

In the early stages of the RM iteration the probability to update the i-th bin is simply ρ_i (same as for reweighting). As Robins-Monro iterations proceed and the coefficients α_i are updated, at asymptotically large Λ all bins are visited with equal probabilities $\sim 1/M$. Therefore in the *bulk* of the DoS defined as all the regions with $\rho_i \gg 1/M$ the probability to visit bin i decreases with the number of iterations Λ and eventually approaches $1/M$ from above. We can thus provide a bounds on N_i in the bulk region:

$$\rho_i \Lambda \geq N_i \geq \frac{\Lambda}{M} \,. \tag{80}$$

On the other hand, in the *tail* of the distribution where $\rho_i \ll 1/M$, the probability to visit the bin i increases with the number of RM iterations, and eventually approaches $1/M$ from below. The expectation value of the iteration count N_i in this region is therefore bounded by

$$\rho_i \Lambda \leq N_i \leq \frac{\Lambda}{M} \,. \tag{81}$$

In the following we assume the best possible scenario, using the upper bounds on N_i, in which case the statistical error is given by

$$\Delta\rho_i = \begin{cases} \sqrt{\frac{\rho_i}{\Lambda}} & \text{if } i \text{ in bulk,} \\ \rho_i \sqrt{\frac{M}{\Lambda}} & \text{if } i \text{ in tail,} \end{cases} \tag{82}$$

and show that LLR cannot outperform reweighting significantly even in that case.

Let us consider some probability ρ_0 from the bulk and another $\rho_\infty = \zeta \rho_0$ from the tail, where ζ is some small parameter. For reweighting we need $\Lambda \sim \zeta^{-2}$ samples to resolve ρ_∞ at all. For LLR on the other hand the error in the tail is much smaller. But even if we knew ρ_∞ exactly, this knowledge is useless as long as $\Delta\rho_0 > \rho_\infty$, i.e. the error is bulk-dominated. From this follows the condition

$$\Delta\rho_0 < \overset{!}{\leq} \rho_\infty \tag{83}$$

$$\Rightarrow \qquad \sqrt{\frac{\rho_0}{\Lambda}} < \zeta \rho_0 \tag{84}$$

$$\Leftrightarrow \qquad \Lambda > \zeta^{-2}/\rho_0 \,. \tag{85}$$

Thus, not only does LLR follow the same scaling as for reweighting in Eq. (66), it even has very similar errors quantitatively since the statistical uncertainties in the bulk are almost identical for reweighting and LLR, and these uncertainties dominate in both cases.

We remark that this argument is applicable independently of the specific form of the DoS and the conclusion therefore generalisable beyond its application to spin chains. Generally, complexity of a given problem is encoded in the magnitude of the parameter ζ. For instance in our case of real time spin chain simulations $\zeta \sim \exp(-Lt)$ is exponentially suppressed with the system size providing an alternative perspective on the origin of the NP-hardness of the sign problem.

D. Augmented LLR_2 or REW_2

Removing the two leading order contributions from the stochastic calculations and evaluating them exactly, comes with an additional cost of

$$\mathcal{O}\left(L 2^L\right) \tag{86}$$

as compared to "plain vanilla" LLR_0 or REW_0. However, it allows to achieve the same precision with a significantly lower accuracy goal since the stochastic part contributes with a reduced weight. More specifically,

$$\varepsilon^{-1} \mapsto \varepsilon^{-1} \left(1 - \cosh(t\hbar)^{-L} \left(1 + \frac{1}{2} L (t\hbar)^2\right)\right) \tag{87}$$

results in an overall runtime of the approximate order

$$T_{\text{sim}} \sim \mathcal{O}\left(L\,2^L + Lt\varepsilon^{-2}\left(\cosh\left(t\mathbb{h}\right)^L - 1 - \frac{L(t\mathbb{h})^2}{2}\right)^2\right). \tag{88}$$

This estimate should be compared with (74). We see that summation over the two leading-order contributions allows to obtain equally precise results at short times with significantly smaller statistics (i.e. computational effort) and to extend the stochastically calculable time region towards somewhat larger values.

VI. RESULTS

All the simulations presented in this work have been performed using `R` as a front- and `C` as a back-end. The complete code and most of the data is publicly available at [**?**]. Some of the data is rather large and has therefore not been published in the same way, but we are happy to provide it in case of interest.

We tested our algorithm on chains of up to $L = 50$ with $\text{LLR}_0/\text{REW}_0$ and up to $L = 40$ with $\text{LLR}_2/\text{REW}_2$. Small lengths~~length~~ $L \leq 16$ have been benchmarked against results from exact diagonalisation. We emphasise that the larger lattices with $L \gtrsim 24$ are far out of the reach of any exact calculations. For instance $L = 40$ would require 16 TB of memory to even store a single complex state vector in double precision. We also stress that while $\text{LLR}_2/\text{REW}_2$ has a runtime scaling as $\mathcal{O}\left(2^L\right)$ and therefore $L > 40$ quickly becomes unfeasible, there are no such limitations for $\text{LLR}_0/\text{REW}_0$ and L can be chosen practically arbitrarily large at the cost of somewhat reducing the evolution time t.

In each simulation we chose a single random realisation of the coupling J. The disorder average over J is technically straight forward and would only add a layer of uncertainties to the results obscuring the quality of the algorithms.

Some typical examples of the scaling of the relative error between the Monte-Carlo and exact diagonalization results with the number of Monte-Carlo iterations are shown on Fig. 5. The error is plotted as a function of the total number of sweeps $N_s = \Lambda/N$ for different evolution times t and correspondingly different strengths of the sign problem. Respective statistical powers Σ are shown as lines. In an ideal stochastic simulation the error should scale as $\Sigma^{-1}/\sqrt{N_s}$ and it does so usually. There are some exceptions in the decrease with N_s though. The shortest evolution times just plateau out at a constant error level. For LLR_0 the simulations would require a much finer discretisation of S_I in order to resolve the spiked DoS properly. The LLR_2 and REW_2 calculations on the other hand are so precise that the Trotter error for the particular choice of N_t is resolved.

From the two top plots in Fig. 5 we also conclude that the results obtained with spline fitting tend to be better than the ones from simple integration for small numbers $N_s \lesssim 10^5$ of Monte-Carlo updates, and for large evolution times $t \sim 1$ where the sign problem is more severe (compare the top left corners of both plots). More statistics undermines the usefulness of fits as they cannot extract any additional information.

The true reason for the favourable spline results in the case of small statistics appears to be that the best fitting functions are very smooth in this case. Therefore the highly oscillatory integral (22) required to obtain the spectral form factor yields a value close to zero. Since the analytic approximation used in LLR_2 is not too bad by itself, the overall result is better than the noise in the unfitted case. Thus, using the stochastic data does not help to improve the analytic approximation for too little statistics. But in the spline-fitted case it does not add any harm either.

It is evident that the augmentation defining LLR_2 allows to reduce the error at short times by several orders of magnitude compared to "plain vanilla" LLR_0. It is also clear, however, that it does not significantly mitigate the sign problem at long times confirming the predictions from equation (86).

A direct comparison of LLR_2 and REW_2 shows very little differences. If anything, reweighting appears to be slightly better. This is exactly what we expect from section V C where we showed that LLR and reweighting have principally the same sign problem. The minor difference in favour of reweighting most likely comes from the binning required for LLR as well as a suboptimal parameter choice (see appendix C).

Figure 6 shows the spectral form factors corresponding to the errors in figure 5. We added the approximations for continuous times derived in section IV. The strength of the LLR_2 method becomes clear when the small differences between the approximations and the exact results are observed. Only these differences have to be calculated stochastically. Note, however, that even at very short times $t \lesssim 0.4$ the stochastic contribution cannot be neglected and the Monte-Carlo results are not compatible with the approximations.

We further remark that the LLR results at large times and small statistics are systematically incompatible with the correct values. This can be explained by a required minimal number of iterations to 'fill up' the DoS. We expect that a different choice of coefficients $a_{L,M,\Lambda}$ and b_M than in equations (40) and (41) might reduce this problem, but we did not conduct additional extensive parameter tuning. This is certainly a great disadvantage of LLR as opposed to reweighting.

The results obtained for longer chains $L \geq 20$ and depicted in figure 7 are qualitatively similar to those discussed above for $L = 16$. Again LLR_2 and REW_2 yield extremely precise results at short times and reweighting generally outperforms LLR. Notably these results demonstrate the scalability of our stochastic algorithms towards

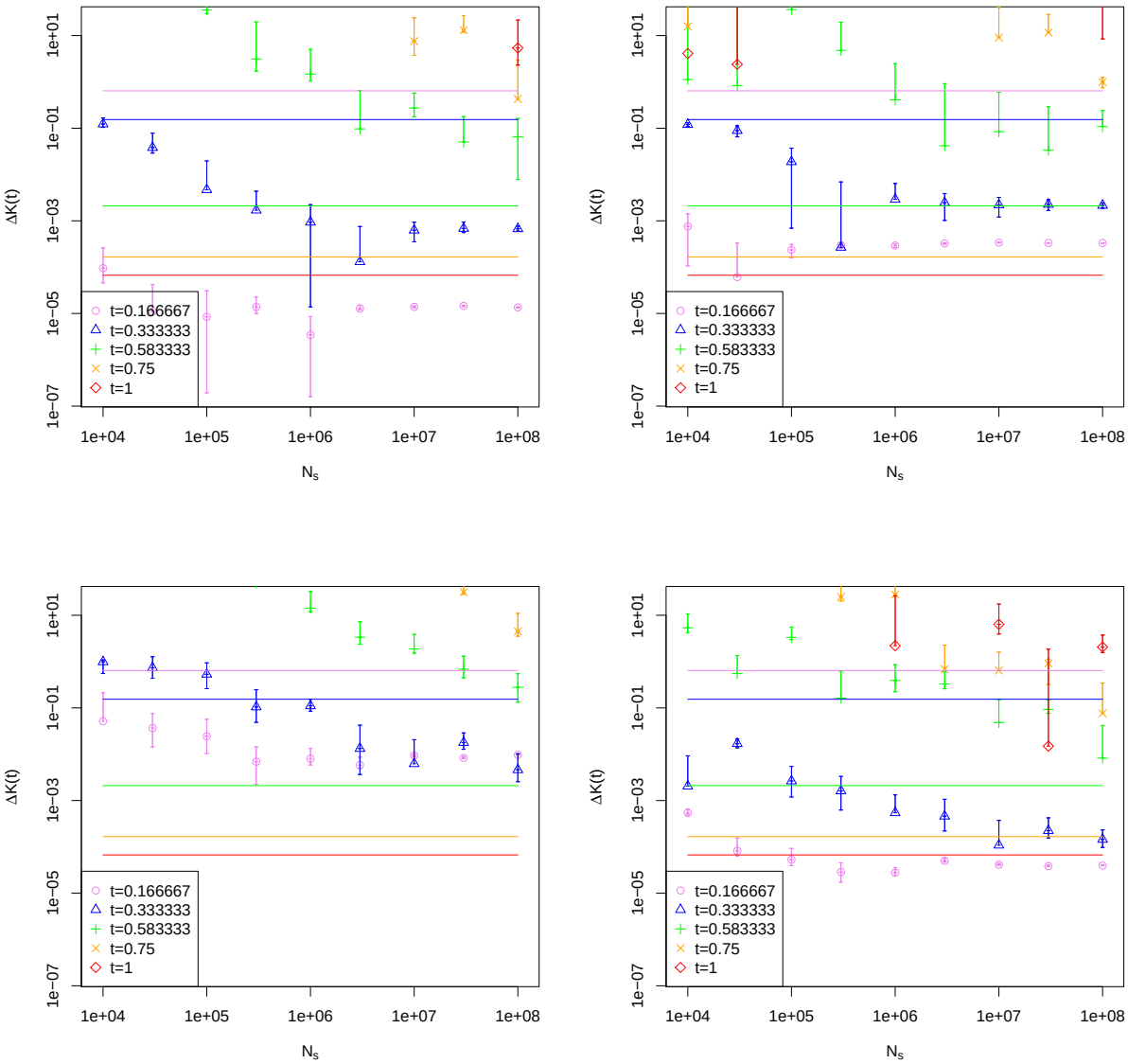

Figure 5. Relative errors of the spectral form factor $\Delta K(t) = |K(t)/K_{\mathrm{ED}}(t) - 1|$, where $K_{\mathrm{ED}}(t)$ has been calculated with exact diagonalisation, as a function of the number of sweeps N_s. For small evolution times $t \lesssim 0.3$, $\Delta K\,(t)$ reaches a plateau value that is entirely saturated by non-statistical Trotterization errors. Error bars in negative direction that would extend below 0 are omitted. Lines show the respective statistical powers. Top left to bottom right: LLR_2, LLR_2 with spline fitting, LLR_0, REW_2. The chain length is $L = 16$ for all plots.

very long chains provided $t \lesssim 1$. Especially for REW_0 we can obtain reliable results for virtually arbitrarily large values of L.

VII. DISCUSSION AND OUTLOOK

In this work, we investigated the possibility to evaluate the spectral form-factor (real-time partition sum) of a non-integrable spin chain (5) using Monte-Carlo simulations. The partition sum contains a multitude of mutually cancelling complex-valued contributions, which turns its direct evaluation into an NP-hard problem. We investigated the infinite-temperature spectral function because it poses the most challenging problem and is also often considered in the literature. We remark however that the method would be applicable with only minor modifications (a larger, in principle unlimited, number of parity sectors) to finite temperature systems as well. In fact, finite temperatures might severely reduce the sign problem.

We compared the most straightforward approach based

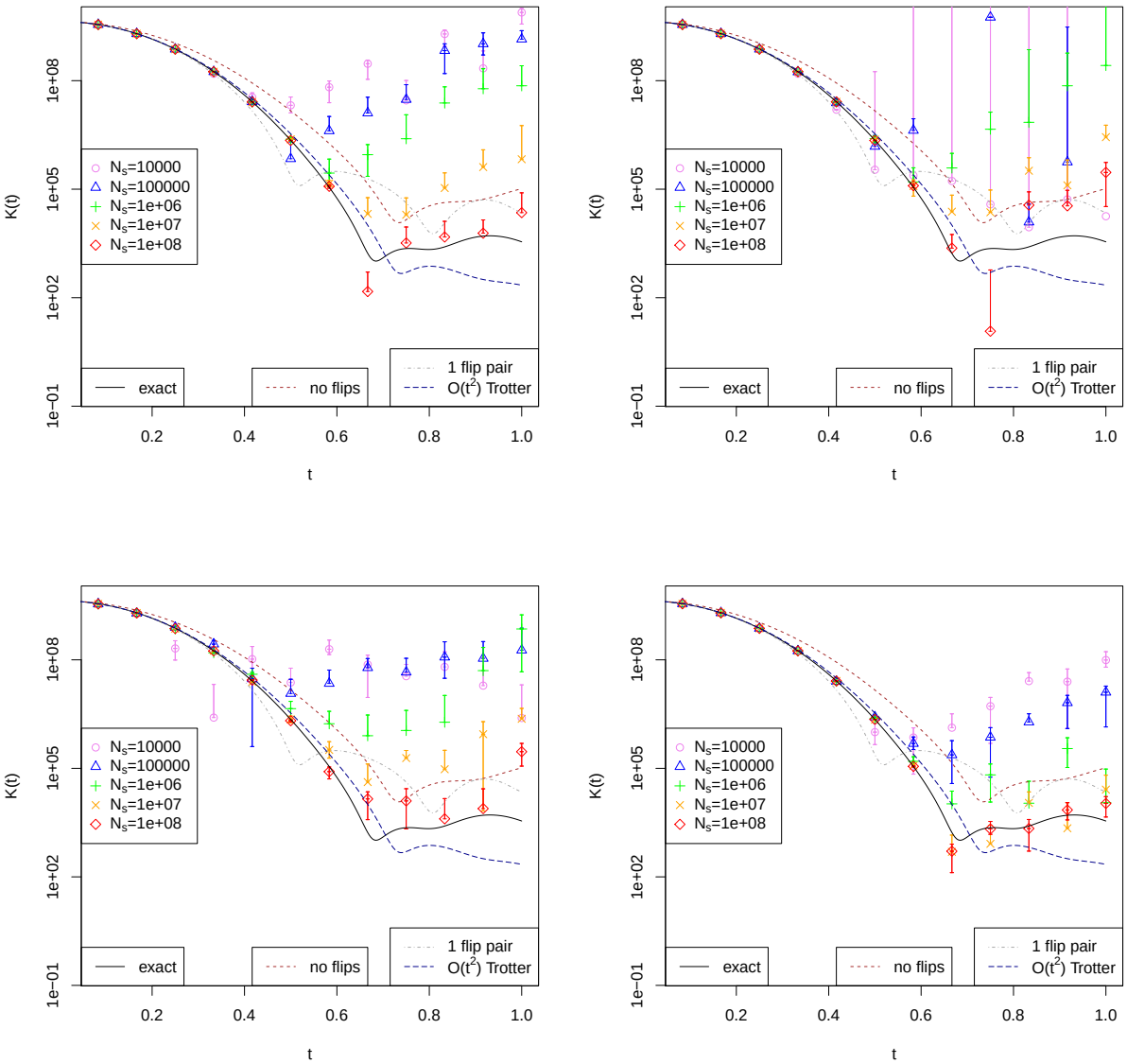

Figure 6. The spectral form factors $K(t)$ compared with the solution from exact diagonalisation ("exact") and the approximations "no flips" (49), "1 flip pair" (58), and "$\mathcal{O}\left(t^2\right)$ Trotter" (61) for different numbers of Monte-Carlo sweeps N_s. Error bars in negative direction that would extend below 0 are omitted. Top left to bottom right: LLR_2, LLR_2 with spline fitting, LLR_0, REW_2. The chain length is $L = 16$ for all plots.

on importance sampling with subsequent reweighting with a more advanced Density-of-States (DoS) approach, in which the statistical distribution of the complex phase of the action is calculated with high precision.

One of our main results is the improved simulation strategy, in which the contribution of configurations with zero and one spin flips into the real-time partition function are summed over exactly. These configurations contribute to the leading and the next-to-leading orders of the expansion of the real-time partition function $\mathrm{Tr}\, U(t)$ in (6) in powers of the magnetic field $\mathbb{h}$. Summation over all other configurations is performed stochastically, using either the standard Monte-Carlo updates (for reweighting), or the LLR algorithm (for the DoS approach). As we discuss below, both approaches work nearly equally well.

As illustrated in Fig. 5, the improved simulation strategy significantly accelerates convergence to exact results, which we obtained using numerical diagonalization of the Hamiltonian. It allows to study real-time evolution of spin chains with lengths up to $L = 40$ and up to physical times $t \lesssim 1$. Our method thus clearly

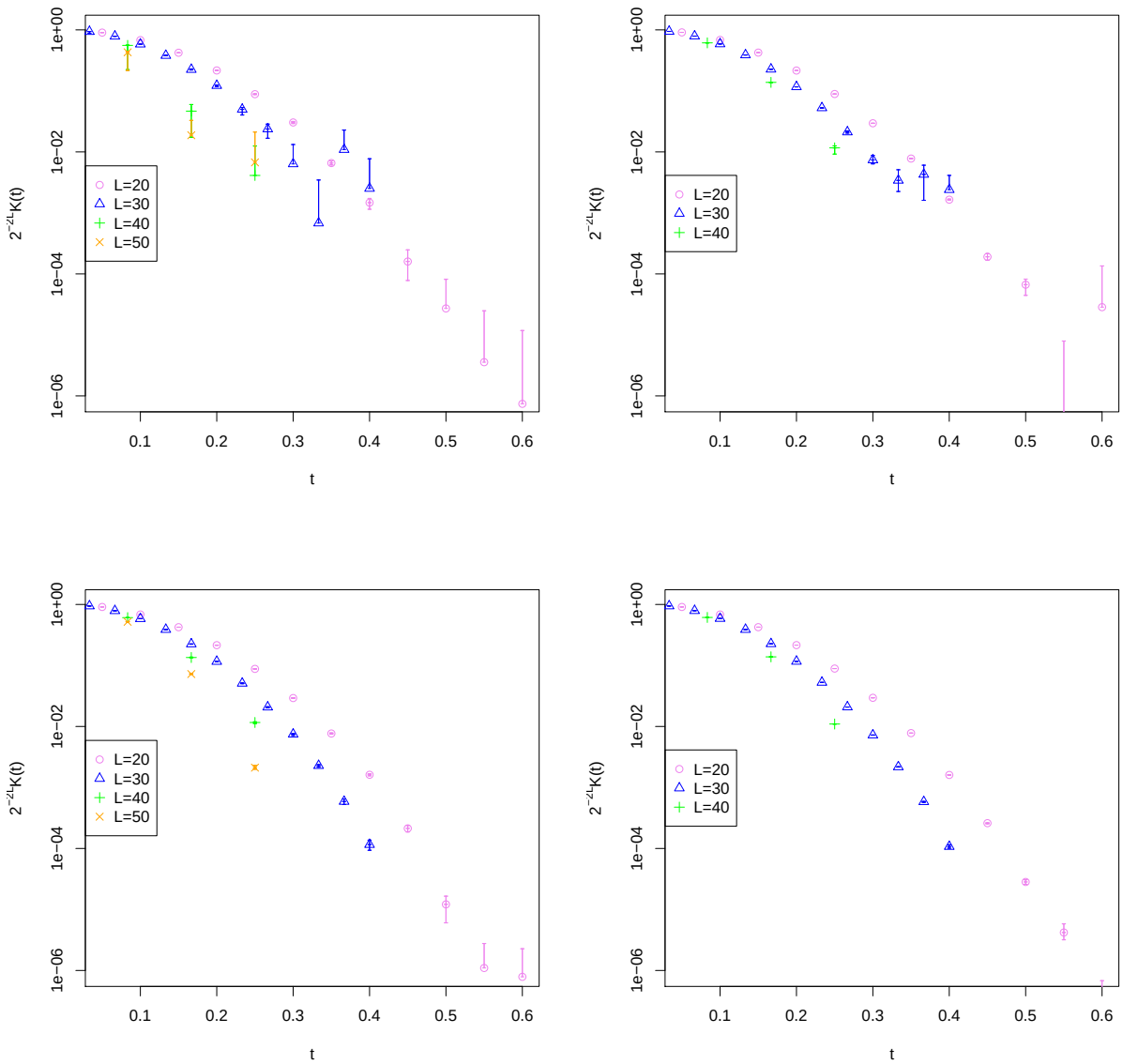

Figure 7. Normalised spectral form factors $2^{-2L}K(t)$ for system sizes hardly ($L = 20$) or not at all ($L \geq 30$) treatable by exact diagonalisation. Error bars in negative direction that would extend below 0 are omitted. Top left to bottom right: LLR_0, LLR_2, REW_0, REW_2. The number of Monte-Carlo sweeps is $N_s = 10^7$ for all plots.

outperforms numerical exact diagonalization methods at short evolution times $t \lesssim 1$. Since for our spin chain the numerical cost of exact diagonalization does not depend on the evolution times (up to possible issues with numerical precision of the result due to round-off errors), exact diagonalization becomes more advantageous for $t \gtrsim 1$. The algorithm could also be translated to a continuous time formulation along the lines of [37], which removes trotterization errors.

An advantage of the stochastic approach is that it can be easily extended to bosonic systems with infinite-dimensional Hilbert spaces, for which exact diagonalization typically becomes prohibitively expensive even for $O(10)$ degrees of freedom. Indeed, to obtain precise results for bosonic degrees of freedom, the size of the local Hilbert space N_{loc} associated with each bosonic degree of freedom should usually be much larger than 2, which results in a significantly faster growth of computational cost $T_{\mathrm{sim}} \sim (N_{\mathrm{loc}})^{n_{\mathrm{dof}}}$ with the number of degrees of freedom than for quantum spin chains. In addition, we need to extrapolate the result to the limit $N_{\mathrm{loc}} \to +\infty$. The possibility to simulate short periods of real-time evolution for systems with a large number of bosonic degrees of freedom is particularly attractive for numerical stud-

ies of the early-time evolution of quark-gluon plasma in heavy-ion collisions before the onset of hydrodynamic behavior.

It is less clear how our results would translate to fermionic systems in general because the performance would crucially depend on a particular realisation. Broadly speaking, we can distinguish two classes; Hubbard-Stratonovich transformations resulting in continuous degrees of freedom and those resulting in discrete degrees of freedom (e.g. Hirsch transformations). In the first case we expect LLR to perform better than reweighting because of the implied smoothness of the DoS. On the other hand the latter case enables improvements of the $\mathrm{LLR}_n/\mathrm{REW}_n$ type.

Another important and somewhat counter-intuitive conclusion is that, at least for our particular Hamiltonian and the trotterization scheme, the DoS/LLR approach does not appear to offer significant advantage over the simple reweighting. In fact, reweighting often outperforms the DoS/LLR approach, presumably, because of the sub-optimal choice of parameters in the LLR algorithm. Reweighting, on the other hand, does not depend on any tunable parameters.

In general, the main technical advantage of the DoS/LLR approach is that it allows to resolve the tails of the distribution of the complex phase with much higher precision than reweighting. In our case, the contribution of these tails to the real-time partition sum is subdominant in comparison to the bulk contribution (central region in the distributions on Fig. 1). Therefore, sampling the tails with high precision is not an advantage in itself. As stressed in [11, 38? , 39], high-precision data for the tails of the density of states becomes an advantage if the density of states $\rho(E)$ can be well approximated by some function with a small number of parameters, e.g. a low-degree polynomial fit or a spline. In this case, precise knowledge on the distribution tails translates into a better choice of fit parameters. This allows for more precise numerical estimates of the oscillatory integrals of the form $Z = \int \mathrm{d}E \rho(E)\, \mathrm{e}^{\mathrm{i}\, E}$ in the representation (4) of the real-time partition function.

As discussed in Section IV E, in our case the DoS in Eq. (23) is a sum of δ-functions and piecewise constant terms for any finite L. The DoS only looks like a continuous function in the thermodynamic limit $L \to +\infty$. While separating out the leading and the next-to-leading contributions to the DoS, as described in Subsections IV A and IV B, leaves us with a much smoother DoS, we were still not able to approximate it by a function (polynomial or spline) with a sufficiently small number of parameters. As a result, spline or polynomial approximations improve the convergence to exact results only for small numbers of Monte-Carlo samples in Robins-Monro iterations (36). Consequently, the LLR approach did not offer significant improvement over the reweighting approach.

In this respect, the LLR approach might work better for systems with continuous degrees of freedom, where the Density of States $\rho(E)$ in (3) is a continuous function even for a finite number of degrees of freedom (away from the thermodynamic limit). See the fermionic system considered in Ref. [20] for an example. In this case, the DoS might allow for good-quality approximations in terms of functions with a small number of parameters, which allows to effectively use the information about the tails of the DoS to improve the precision of oscillatory integrals of the form (4). However, other complications might arise when applying the DoS/LLR method to continuous degrees of freedom. To provide the most obvious example, the DoS will be clearly a non-normalizable and unbounded function for the real-time partition function of even the simplest one-dimensional quantum harmonic oscillator. For the harmonic oscillator, this unboundedness can be easily solved by complexification of the path integral, which makes the path integral completely real-valued. The complexification, however, becomes increasingly complicated for anharmonic potentials. As a result, we will be inevitably forced to adapt the Lefschetz thimble/holomorphic flow approaches, which have their own complications [40–44]. A further problem with the DoS/LLR approach for continuous variables might be the loss of ergodicity for simulations with constrained values of the imaginary part of the action S_I, especially if one uses the Hybrid Monte-Carlo algorithm based on nearly continuous updates of field variables. Finally, the polynomial approximations to the density of states, which plays a crucial role in reducing statistical errors, might fail in the vicinity of phase transitions [14].

We should note that the DoS/LLR approach might also be useful for exploratory studies, for example, to guide the construction of analytic approximations. In fact, our work proceeded in exactly this way - we first measured the "unsubtracted" DoS function with LLR_0, and only after looking at the DoS plots in Fig. 1 we realized how to subtract the leading and next-to-leading contributions specified in eqs. (49) and (58).

ACKNOWLEDGEMENTS

This work was funded in part by the STFC Consolidated Grant ST/T000988/1. Numerical simulations were undertaken on Barkla, part of the High Performance Computing facilities at the University of Liverpool, UK as well as on the Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility (`www.dirac.ac.uk`). The DiRAC component of CSD3 was funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1. DiRAC is part of the National e-Infrastructure. We used the `R` library `hadron` [45] for data analysis. JO would like to thank the other participants of the Pollica Summer Workshop on Effective Field Theories 2022, in particular Michael Flynn, for

many an insightful discussion. We also thank Kurt Langfeld, Paul Rakow and David Schaich for their valuable comments.

Appendix A: Derivation of the classical Ising model in $d+1$ dimensions and further details

The exponentials in equation (10) can be calculated exactly in the canonical z-basis because the first term is diagonal and the second one can be decomposed into local blocks of size 2×2 yielding

$$\mathrm{e}^{\mathrm{i}\,\delta\mathbb{h}_i\sigma_i^x} = \begin{pmatrix} \cos\delta\mathbb{h}_i & \mathrm{i}\sin\delta\mathbb{h}_i \\ \mathrm{i}\sin\delta\mathbb{h}_i & \cos\delta\mathbb{h}_i \end{pmatrix} . \tag{A1}$$

Let us now compare these matrices with the local building blocks of the anisotropic classical Ising model with an arbitrary coupling matrix J' within the first dimension and a nearest-neighbour coupling h' within the second dimension defined by the action

$$S = S^1 + S^2 , \tag{A2}$$

$$S^1 = -\sum_k \sum_{i,j} s_{i,k} J'_{ij} s_{j,k} , \tag{A3}$$

$$S^2 = -\sum_i \sum_k h'_i s_{i,k} s_{i,k+1} . \tag{A4}$$

The physics of such a classical statistical system are governed by the Boltzmann weighted partition sum (over all possible spin configurations)

$$Z = \sum_{\{s\}} \mathrm{e}^{-S(s)} \tag{A5}$$

inducing the transfer matrices

$$T_1 = \left(\mathrm{e}^{-S^1}\right)_{k,k+1} \tag{A6}$$

$$= \mathrm{diag}\left(\left\{\sum_{i,j} s_i J'_{ij} s_j \mid s\in\{\pm 1\}^L\right\}\right) \tag{A7}$$

and

$$(T_2)_i = \left(\mathrm{e}^{-(S^2)_i}\right)_{k,k+1} = \begin{pmatrix} \mathrm{e}^{h'_i} & \mathrm{e}^{-h'_i} \\ \mathrm{e}^{-h'_i} & \mathrm{e}^{h'_i} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{h'_i}+\mathrm{e}^{-h'_i} & 0 \\ 0 & \mathrm{e}^{h'_i}-\mathrm{e}^{-h'_i} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} . \tag{A8}$$

We immediately observe the similarity between T_1 and the interacting part of the time evolution operator. Therefore we set

$$J' \equiv \mathrm{i}\,\delta\mathbb{J} . \tag{A9}$$

Note again that this $\mathbb{J}$ corresponds to the purely spatial coupling matrix. The coupling in temporal direction turns out to be more challenging as we have to exploit the similarity of T_2 with the 2×2-matrix (A1):

$$\begin{pmatrix} \mathrm{e}^{h'_i} & \mathrm{e}^{-h'_i} \\ \mathrm{e}^{-h'_i} & \mathrm{e}^{h'_i} \end{pmatrix} \overset{!}{\propto} \begin{pmatrix} \cos\delta\mathbb{h}_i & \mathrm{i}\sin\delta\mathbb{h}_i \\ \mathrm{i}\sin\delta\mathbb{h}_i & \cos\delta\mathbb{h}_i \end{pmatrix} \tag{A10}$$

$$\Rightarrow \mathrm{e}^{-2h'_i} = \mathrm{i}\tan\delta\mathbb{h}_i \tag{A11}$$

Together with the proportionality factor

$$A = \prod_i A_i^{N_t} , \tag{A12}$$

$$A_i = \sqrt{\mathrm{i}\sin\delta\mathbb{h}_i\cos\delta\mathbb{h}_i} \tag{A13}$$

this allows an exact identification between the trace of the quantum mechanical time evolution operator and the classical partition sum

$$\mathrm{Tr}\,U(t) = AZ . \tag{A14}$$

We can define the purely real constants

$$J := \delta\mathbb{J} , \tag{A15}$$

$$h_i := -\frac{1}{2}\log\tan\delta\mathbb{h}_i , \tag{A16}$$

so that the action reads

$$S = -\mathrm{i}\sum_k\sum_{i,j} s_{i,k} J_{ij} s_{j,k} - \\ -\sum_i\sum_k \left(h_i - \frac{\pi}{4}\mathrm{i}\right) s_{i,k} s_{i,k+1} . \tag{A17}$$

1. Distribution of the imaginary part

It is instructive to investigate the distribution of S_I before we go on. For this we define the auxiliary variable

$$\theta_k := \sum_{i,j} s_{i,k} J_{ij} s_{j,k} \tag{A18}$$

with the upper bound

$$|\theta_k| \le ||J||_1 \tag{A19}$$

$$\le nL\,\delta\,||\mathbb{J}||_\infty , \tag{A20}$$

where n is the number of neighbours a single site can couple to ($n = 2$ for pure nearest neighbour coupling in 1 dimension) and $||\cdot||_p$ denotes the p-norm. Since the considered systemsystems is translationally invariant in the time direction, all θ_k are distributed identically, though not independently in general. The $\mathbb{Z}_2$ symmetry of spin reflections is unbroken, so the expectation value of θ_k is $\bar{\theta}_k = 0$.

Suppose the θ_k were uncorrelated. Then according to the central limit theorem S_I would follow a normal distribution

$$S_I \sim \mathcal{N}(0, v) \tag{A21}$$

with a variance

$$v \leq N_t \left(nL\,\delta\,||\mathcal{J}||_\infty\right)^2 \tag{A22}$$
$$= \frac{1}{N_t}\left(nL\,t\,||\mathcal{J}||_\infty\right)^2 \tag{A23}$$

approaching zero in the continuum limit.

In the opposite limit of maximal correlation on the other hand S_I simplifies to

$$S_I = N_t\theta_k \tag{A24}$$
$$\Rightarrow |S_I| \leq N_t\, nL\delta\, ||\mathcal{J}||_\infty \tag{A25}$$
$$= nL\,t\,||\mathcal{J}||_\infty\,, \tag{A26}$$

that is a constant in the time steps size. We do not take the case of anti-correlation into account since it is highly unphysical. Therefore we find that in any relevant case S_I is bounded by the physical time extent and the variance is going to approach a constant

$$\lim_{N_t\to\infty} v \leq t\,\xi\left(nL\,||\mathcal{J}||_\infty\right)^2 \tag{A27}$$

in the continuum limit, where ξ is the correlation length in time direction. ξ is proportional to the standard deviation of S_I.

2. Alternative boundary conditions

We remark here that the periodic boundary conditions (pbc) silently assumed above are the most usual but not the only possible choice. Physically they correspond to same-to-same scattering and are best suited for infinite time approximations. In contrast open boundary conditions (obc) correspond to all-to-all scattering and might be relevant in comparisons with matrix product state (MPS) or similar calculations where they are much easier to realise than pbc. Closed boundary conditions (cbc) allow for particular choices of initial and final states.

In all the non-periodic cases there is one less time step than time slices, so the proportionality factor between quantum and classical partition sums reduces to $\prod_i A_i^{N_t-1}$. Furthermore the trace in equation (27) has to be replaced by projections to corresponding vectors.

For *obc* one obtains

$$Z_{R,i} = \frac{1}{2}\,(1,1)\,(T_2)_i^{N_t}\begin{pmatrix}1\\1\end{pmatrix} \tag{A28}$$
$$= 2^{N_t-1}\cosh^{N_t-1} h_i'\underline{\tilde{h}_i'} \tag{A29}$$

and for *cbc*

$$Z_{R,i} = (1,0)\,(T_2)_i^{N_t}\begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix} \tag{A30}$$
$$= 2^{N_t-1}\Big((\cos\varphi+\sin\varphi)\cosh^{N_t-1} h_i'\underline{\tilde{h}_i'} + (\cos\varphi-\sin\varphi)\cosh^{N_t-1} h_i'\underline{\tilde{h}_i'}\Big)\,, \tag{A31}$$

where initial and final states are rotated against each other by the angle φ. We observe that, as expected, all the boundary conditions lead to the same result at infinite times $t\to\infty$, $N_t\to\infty$ and moreover the cbc case of $\varphi = 0$ (same-to-same scattering) yields the same result as pbc up to an irrelevant difference of N_t by one.

Appendix B: Error scaling with bin size

The discretisation of the DoS into bins as in equation (32) naturally leads to errors vanishing as the bin width goes to zero, but relevant for finite bin size. The observable we are ultimately interested in amounts to a Fourier transformation of the DoS, so we can judge the quality of the discretisation method by its ability to approximate the integral

$$F_a(x_0,h) := \frac{1}{2h}\int_{x_0-h}^{x_0+h} \mathrm{d}x\,\mathrm{e}^{f(x)+\mathrm{i}\,ax}\,. \tag{B1}$$

The separate estimation of the DoS in bins of length $2h$ succeeded by a multiplication with the phase at the mid-point of the interval as employed in this work results in

$$\frac{\mathrm{e}^{\mathrm{i}\,ax_0}}{2h}\int_{x_0-h}^{x_0+h}\mathrm{d}x\,\mathrm{e}^{f(x)} = = F_a(x_0,h)\,\mathrm{e}^{-\frac{1}{6}\,\mathrm{i}\,a\left(2f'(x_0)+\mathrm{i}\,a\right)h^2+\mathcal{O}\left(h^4\right)}. \tag{B2}$$

Alternative approximations include the classical mid-point formula

$$\mathrm{e}^{f(x_0)+\mathrm{i}\,ax_0} = = F_a(x_0,h)\,\mathrm{e}^{-\frac{1}{6}\left(f''(x_0)+\left(f'(x_0)+\mathrm{i}\,a\right)^2\right)h^2+\mathcal{O}\left(h^4\right)} \tag{B3}$$

as well as a more intricate piecewise linear (trapezoidal) formula

$$\frac{\mathrm{e}^{f(x_0)}}{2h}\int_{x_0-h}^{x_0+h}\mathrm{d}x\,\mathrm{e}^{f'(x_0)x+\mathrm{i}\,ax} = = F_a(x_0,h)\,\mathrm{e}^{-\frac{1}{6}f''(x_0)h^2+\mathcal{O}\left(h^4\right)} \tag{B4}$$

as used e.g. in Ref. [19].

Thus, all the different methods' multiplicative errors are of order $\mathcal{O}\left(h^2\right)$. Empirically we find that the bin size can easily be chosen small enough to completely neglect this error since the uncertainties in the estimation of $f(x)$ itself dominate the total error.

Appendix C: LLR step size

First, we note that condition (37) is not sufficient in practice as the sum does not go up to infinity but only up to Λ. Therefore we require additionally

$$\sum_{k=0}^{\Lambda} \beta_{N,M,\Lambda}(k) \geq \sum_i \hat{\alpha}_i \,, \tag{C1}$$

where we denote the exact logarithmic DoS with the hat $\hat{\alpha}_i$. The left hand side of equation (C1) evaluates to

$$\sum_{k=0}^{\Lambda} \beta_{N,M,\Lambda}(k) = \\ = a_{N,M,\Lambda} \left(\ln \frac{\Lambda}{b_{N,M,\Lambda}} + \mathcal{O}\left(b_{N,M,\Lambda}^{-1}\right) + \mathcal{O}\left(\Lambda^{-1}\right) \right) \tag{C2}$$

while the right hand side can only be approximated

$$\sum_i \hat{\alpha}_i \leq \sum_i N \ln 2 \tag{C3}$$
$$= MN \ln 2 \tag{C4}$$
$$\approx 0.69 MN \,. \tag{C5}$$

Though this is a conservative upper bound, it is not very far from realistic estimations. If we for instance assume a parabolic shape for the α_i with a maximum close to $N \ln 2$, we obtain $\sum_i \hat{\alpha}_i \lesssim \frac{2}{3} \ln 2\, MN \approx 0.46 MN$. Thus, up to a factor of order one, we find that

$$a_{N,M,\Lambda} = \ln 2 \, \frac{MN}{\ln \frac{\Lambda}{b_{N,M,\Lambda}}} \tag{C6}$$

is a good choice.

In order to get a value for the remaining offset $b_{N,M,\Lambda}$ we have to consider the small k behaviour rather than the large k limit as we did before because the offset is going to be irrelevant in the latter case. Consider the likely case that one of the first contributions $\beta_{N,M,\Lambda}(k \ll b_{N,M,\Lambda})$ is added to a state that has a much smaller exact $\hat{\alpha}_i$ relative to the other states. To compensate this 'mistake' after some k_0 steps (k_0 can be large if the initial random walk remains in a distant region of the phase space for a while), we have to spend κ steps adding this contribution to every other of the M states. For this to be possible, the condition

$$\frac{M}{b_{N,M,\Lambda}} \leq \sum_{k=k_0}^{k_0+\kappa} \frac{1}{b_{N,M,\Lambda} + k} \tag{C7}$$
$$= \ln \frac{k_0 + \kappa}{k_0} + \mathcal{O}\left(\frac{b_{N,M,\Lambda}}{k_0}\right) \tag{C8}$$

has to hold. This provides a lower bound for κ

$$\kappa \geq k_0 \left(\mathrm{e}^{\frac{M}{b_{N,M,\Lambda}}} - 1 \right) \tag{C9}$$

which has to be small for the algorithm to be efficient. Thus, it turns out that the offset $b_{N,M,\Lambda} \equiv b_M$ solely depends on the number of states and $b_M \gtrsim M$ for the runtime to not blow up exponentially. On the other hand we cannot choose b_M arbitrarily large either since that would defeat the purpose of $\beta_{N,M,\Lambda}(k)$ decreasing quickly. Numerical test suggest that

$$b_M = 3M \tag{C10}$$

yields a performance close to optimal.

[1] R. P. Feynman, Quantum mechanical computers, Found. Phys. **16**, 507 (1986).

[2] D. Nagaj, Fast universal quantum computation with railroad-switch local hamiltonians, J. Math. Phys. **51**, 062201 (2010), 0908.4219.

[3] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Many-Body Localization in a Disordered Quantum Ising Chain, Phys. Rev. Lett. **113**, 107204 (2014), 1403.1568.

[4] M. Serbyn, Z. Papić, and D. A. Abanin, Criterion for Many-Body Localization-Delocalization Phase Transition, Phys. Rev. X **5**, 041047 (2015), 1507.01635.

[5] D. Abanin, J. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. Parameswaran, F. Pollmann, A. Potter, M. Serbyn, and R. Vasseur, Distinguishing localization from chaos: Challenges in finite-size systems, Ann. Phys. **427**, 168415 (2021), 1911.04501.

[6] D. Sels and A. Polkovnikov, Dynamical obstruction to localization in a disordered spin chain, Phys. Rev. E **104**, 054105 (2021), 2009.04501.

[7] F. Pietracaprina, N. Macé, D. J. Luitz, and F. Alet, Shift-invert diagonalization of large many-body localizing spin chains, SciPost Phys. **5**, 45 (2018), 1803.05395.

[8] P. Sierant, M. Lewenstein, and J. Zakrzewski, Polynomially filtered exact diagonalization approach to many-body localization, Phys. Rev. Lett. **125**, 156601 (2020), 2005.09534.

[9] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Slow delocalization of particles in many-body localized phases, Phys. Rev. B **103**, 024203 (2021), 2010.00565.

[10] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Unlimited growth of particle fluctuations in many-body localized phases, Ann. Phys. **435**, 168481 (2021), 2012.12436.

[11] N. Garron and K. Langfeld, Controlling the sign problem in finite-density quantum field theory, Eur. Phys. J. C **77** (2017), 1703.04649.

[12] N. Garron and K. Langfeld, Tackling the sign problem with a moment expansion and application to heavy dense QCD, PoS **LATTICE2016**, 084 (2017), 1611.01378.

[13] L. Bongiovanni, K. Langfeld, B. Lucini, R. Pellegrini, and A. Rago, The density of states approach at finite chemical potential: a numerical study of the Bose gas, PoS **LATTICE2015**, 192 (2016), 1601.02929.

[14] O. Francesconi, M. Holzmann, B. Lucini, and A. Rago, Free energy of the self-interacting relativistic lattice Bose gas at finite density, Phys. Rev. D **101**, 014504 (2020), 1910.11026.

[15] F. Wang and D. P. Landau, Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States, Phys. Rev. Lett. **86**, 2050 (2001), cond-mat/0011174.

[16] R. E. Belardinelli and V. D. Pereyra, Wang-Landau algorithm: A theoretical analysis of the saturation of the error, J. Chem. Phys. **127**, 184105 (2007), cond-mat/0702414.

[17] R. E. Belardinelli, S. Manzi, and V. D. Pereyra, Analysis of the convergence of the $1/t$ and Wang-Landau algorithms in the calculation of multidimensional integrals, Phys. Rev. E **78**, 067701 (2008), 0806.0268.

[18] H. Robbins and S. Monro, A Stochastic Approximation Method, Ann. Math. Stat. **22**, 400 (1951).

[19] K. Langfeld, B. Lucini, R. Pellegrini, and A. Rago, An efficient algorithm for numerical computations of continuous densities of states, Eur. Phys. J. C **76**, 306 (2016), 1509.08391.

[20] M. Körner, K. Langfeld, D. Smith, and L. von Smekal, Density of states approach to the hexagonal Hubbard model at finite density, Phys. Rev. D **102**, 054502 (2020), 2006.04607.

[21] M. Guagnelli, Sampling the density of states (2012), 1209.4443.

[22] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, JHEP **1711**, 048, 1706.05400.

[23] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, Onset of random matrix behavior in scrambling systems, JHEP **1807**, 124, 1803.08050.

[24] J. Liu, Spectral form factors and late time quantum chaos, Phys. Rev. D **98**, 086026 (2018), 1806.05316.

[25] A. Prakash, J. H. Pixley, and M. Kulkarni, Universal spectral form factor for many-body localization, Phys. Rev. Research **3**, 012019 (2021), 2008.07547.

[26] R. Vosk, D. A. Huse, and E. Altman, Theory of the Many-Body Localization Transition in One-Dimensional Systems, Phys. Rev. X **5**, 031032 (2015), 1412.3117.

[27] D. J. Luitz, N. Laflorencie, and F. Alet, Many-body localization edge in the random-field Heisenberg chain, Phys. Rev. B **91**, 081103 (2015), 1411.0660.

[28] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, Rev. Mod. Phys. **91**, 021001 (2019), 1804.11065.

[29] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Interacting electrons in disordered wires: Anderson localization and low-t transport, Phys. Rev. Lett. **95**, 206603 (2005), cond-mat/0506411.

[30] D. Basko, I. Aleiner, and B. Altshuler, Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states, Ann. Phys. **321**, 1126 (2006), cond-mat/0506617.

[31] W. De Roeck and F. Huveneers, Stability and instability towards delocalization in many-body localization systems, Phys. Rev. B **95**, 155129 (2017), 1608.01815.

[32] P. Sierant and J. Zakrzewski, Challenges to observation of many-body localization, Phys. Rev. B **105**, 224203 (2022), 2109.13608.

[33] Y. Atia and D. Aharonov, Fast-forwarding of Hamiltonians and exponentially precise measurements, Nature Comm. **8**, 1572 (2017), 1610.09619.

[34] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2020).

[35] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde, Bayesian measures of model complexity and fit, J. R. Stat. Soc. B **64**, 583 (2002).

[36] A. Nemirovskii and D. Yudin, *Problem complexity and method efficiency in optimization* (John Wiley & Sons, New York, 1983).

[37] B. B. Beard and U.-J. Wiese, Simulations of Discrete Quantum Systems in Continuous Euclidean Time, Phys. Rev. Lett. **77**, 5130 (1996), cond-mat/9602164.

[38] K. Langfeld and B. Lucini, The density of states approach to dense quantum systems, Phys. Rev. D **90**, 094502 (2014), 1404.7187.

[39] N. Garron and K. Langfeld, Anatomy of the sign-problem in heavy-dense QCD, Eur. Phys. J. C **76**, 569 (2016), 1605.02709.

[40] Z.-G. Mou, P. M. Saffin, A. Tranberg, and S. Woodward, Real-time quantum dynamics, path integrals and the method of thimbles, JHEP **1906**, 094, 1902.09147.

[41] A. Alexandru, G. Basar, P. F. Bedaque, S. Vartak, and N. C. Warrington, Monte Carlo study of real time dynamics, Phys. Rev. Lett. **117**, 081602 (2016), 1605.08040.

[42] A. Alexandru, G. Basar, P. F. Bedaque, and G. W. Ridgway, Schwinger-Keldysh on the lattice: a faster algorithm and its application to field theory, Phys. Rev. D **95**, 114501 (2017), 1704.06404.

[43] J.-L. Wynen, E. Berkowitz, S. Krieg, T. Luu, and J. Ostmeyer, Machine learning to alleviate Hubbard-model sign problems, Phys. Rev. B **103**, 125153 (2021), 2006.11221.

[44] M. Rodekamp, E. Berkowitz, C. Gäntgen, S. Krieg, T. Luu, and J. Ostmeyer, Mitigating the Hubbard Sign Problem with Complex-Valued Neural Networks, Phys. Rev. B **106**, 125139 (2022), 2203.00390.

[45] B. Kostrzewa, J. Ostmeyer, M. Ueding, and C. Urbach, *hadron: Analysis Framework for Monte Carlo Simulation Data in Physics* (2020), r package version 3.1.2.