Cluster truncated Wigner approximation in strongly interacting systems

Accepted Manuscript

Cluster truncated Wigner approximation in strongly interacting systems

Jonathan Wurtz, Anatoli Polkovnikov, Dries Sels

Please cite this article as: J. Wurtz, A. Polkovnikov, D. Sels, Cluster truncated Wigner approximation in strongly interacting systems, *Annals of Physics* (2018), https://doi.org/10.1016/j.aop.2018.06.001

Cluster Truncated Wigner Approximation in Strongly Interacting Systems

Jonathan Wurtz,[1] Anatoli Polkovnikov,[1] and Dries Sels[1, 2, 3]

[1]*Department of Physics, Boston University, 590 Commonwealth Ave., Boston, MA 02215, USA*
[2]*Department of Physics, Harvard University, 17 Oxford st., Cambridge, MA 02138, USA*
[3]*Theory of quantum and complex systems, Universiteit Antwerpen, B-2610 Antwerpen, Belgium*

We present a general method by which linear quantum Hamiltonian dynamics with exponentially many degrees of freedom is replaced by approximate classical nonlinear dynamics with the number of degrees of freedom (phase space dimensionality) scaling polynomially in the system size. This method is based on generalization of the truncated Wigner approximation (TWA) to a higher dimensional phase space, where phase space variables are associated with a complete set of quantum operators spanning finite size clusters. The method becomes asymptotically exact with the increasing cluster size. The crucial feature of TWA is fluctuating initial conditions, which we approximate by a Gaussian distribution. We show that such fluctuations dramatically increase accuracy of TWA over traditional cluster mean field approximations. In this way we can treat on equal footing quantum and thermal fluctuations as well as compute entanglement and various equal and non-equal time correlation functions. The main limitation of the method is exponential scaling of the phase space dimensionality with the cluster size, which can be significantly reduced by using the language of Schwinger bosons and can likely be further reduced by truncating the local Hilbert space variables. We demonstrate the power of this method analyzing dynamics in various spin chains with and without disorder and show that we can capture such phenomena as long time hydrodynamic relaxation, many-body localization and the ballistic spread of entanglement.

I. INTRODUCTION

Phase space methods recently emerged as very powerful tools for simulating dynamics of quantum systems typically close to some classical limit. Such methods found a wide range of applications in many areas of science such as quantum chemistry, quantum optics, atomic systems, quantum chaos, condensed matter physics and others. Generally these methods replace exponential complexity of simulating quantum systems with the power law complexity of classical systems. This comes at the price of the classical dynamics being nonlinear, leading to potential instabilities. Quantum effects introduce additional stochastic dynamical noise on classical equations of motion, which is usually described by non-positive probability distribution and leads to rapidly increasing complexity of exact simulations with time. A common approximation, where this quantum noise is neglected, is known as the truncated Wigner approximation (TWA)[1–4]. At the same time TWA accounts for the initial noise encoded in the Wigner function describing the initial state. It is this noise which makes TWA distinct from mean field approaches. A standard limitation of TWA is that it is usually applicable for a finite time, which is set by the effective Planck's constant and by the magnitude of nonlinearities (interactions) in the system. TWA becomes exact either in the classical, or the noninteracting limit[4].

In strongly interacting systems far from the classical limit standard phase space methods do not usually lead to simplifications or may even result in uncontrolled approximations. However, as was demonstrated in Refs. 5 and 6 one can partially circumvent these difficulties by increasing dimensionality of phase space by adding extra/(hidden) degrees of freedom representing correlations and treating them as independent variables. Thus in Ref. 5 a set of three level (spin one) systems was efficiently treated by first mapping them to SU(3) representation with 8 classical degrees of freedom per spin and then applying SU(3) TWA instead of the traditional TWA based on the SU(2) spin representation[4,7] with

2 degrees of freedom per spin. This increased phase space dimensionality allows one to map all local on-site interactions to an effective on-site magnetic field, linearizing the local part of the Hamiltonian. Thus the only source of nonlinearity potentially leading to errors due to TWA come from the inter-spin interactions. As was demonstrated in Ref. 5 using the $SU(3)$ representation significantly improves the accuracy of TWA. Similarly, TWA was adopted to fermion bilinears, which can be identified as classical phase space variables[6]. These bilinears (or string variables) form a closed SO(2N) Lie-algebra for N-single particle orbitals, which can be described by a phase space of dimensionality $2N^2$ instead of the naive dimensionality of $2N$. In Refs. 6 and 8 it was shown that fermion TWA accurately describes quantum dynamics in various non-trivial setups like a two-channel model, expansion of interacting fermions and dynamics of the SYK model.

In this work we extend the ideas of Ref. 5 by developing a cluster TWA (CTWA) approach, where we treat all operator degrees of freedom confined to a cluster as independent phase space variables. In this way we first map quantum Hamiltonians to large $\mathcal{N}$-models, where $\mathcal{N} = D^2$ and D is the Hilbert space size of the cluster. Then dynamics of the system is governed by the classical equations of motion as in standard large-$\mathcal{N}$ theories (Ref. 9), with an additional ingredient that the initial conditions are sampled according to a Gaussian distribution set by the initial density matrix of the system. By construction, CTWA becomes asymptotically exact as we increase the cluster size and thus gives a controllable expansion for simulating dynamics in quantum systems, at least in principle. As it will be clear from our discussion, mathematically clusters can be defined in an arbitrary way by splitting the Hilbert space into a sum of orthogonal subspaces (clusters) and mapping the operator basis spanning these subspaces (which form a closed SU(D) algebra) into classical degrees of freedom. In this work we will focus into local spatial clusters.

One of the advantages of using the phase space formalism

FIG. 2. **Example of clustering of spins**. Blue circles are physical spins $\frac{1}{2}$, while red rectangles are the clusters. Hamiltonian dynamics within each cluster is exact, while interactions between clusters are treated approximately.
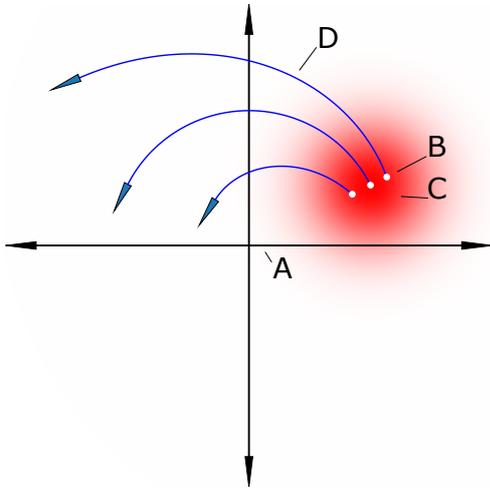

FIG. 1. Visual example of CTWA evolution in phase space. The coordinate system (A) (Black lines) parameterizes a phase space for points $\{x\}$ and functions $O(\{x\})$ mapped from operators $\hat{O}$. Points in phase space (B) (white dots) are selected from a probability distribution (C) (red blob) and evolved independently according to mean field classical equations of motion (D) (blue traces). As evolution is nonlinear, nearby points diverge in time. Observables are found by averaging over all sampled points.

is that it allows one to discuss equations of motion, observables, time correlators, and initial conditions which include both quantum and thermal fluctuations. On the formal level, TWA can be derived by projecting exact time evolution of the density matrix into an appropriate operator product subspace. In this sense TWA is very similar to mean field or the time-dependent variational principle (TDVP)[10–12]. However, because the initial Wigner function can never be represented as a single point in this space (due to e.g. the uncertainty principle) TWA necessarily involves a statistical mixture of different classical trajectories (see Fig. 1) over a range of initial conditions in phase space, which generally cannot be expressed as a single mean field evolution.

This statistical mixture in phase space allows one to simulate build up of quantum correlations and entanglement from initial product states, which is encoded in the classical mutual information induced by nonlinear evolution in phase space. Let us point that the main limitation of the cluster TWA is rapid scaling of the phase space dimensionality with the cluster size. E.g. for spin one-half systems the phase-space dimensionality of a cluster of size N scales as $D^2 = 4^N$. We show that this number can be reduced to $D = 2^N$ by employing multiple conservation laws analogous to spin conservation in the standard $SU(2)$ case. Equivalently one can understand this dimensional reduction using the classical analogues of the Schwinger boson representation of the phase space variables. On a formal level these dimensional reductions map classical (mean field) equations for density functions to classical (mean field) equations for the wave functions. At the end of the paper, following the ideas developed in Ref. 11 for

mean field dynamics, we comment how phase space dimensionality can be reduced further by using incomplete local Hilbert space constructed from e.g. the matrix product state operator basis[13–15].

The rest of this paper will be structured as follows. Sections II A and II B cover the outline of the CTWA for the two methods: Operator, and Schwinger Boson CTWA. Section II C gives a simple example. Sections III and IV, will give various nontrivial demonstrations of the method. Section V outlines details of the method, including rationale for the Wigner function and dimensional reduction of the operator CTWA.

II. CLUSTER TWA: OPERATIONAL MANUAL

Postponing a detailed discussion of the cluster TWA until the end of the paper, let us start with a brief summary of the method and demonstrating how it can be applied to a simple example. For illustration of the method we will use a system of interacting spins $\frac{1}{2}$ described by the Hamiltonian:

$$\hat{H} = \sum_{ij} J_{ab}^{ij} \hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(j)} + \sum_j B_a^j \hat{\sigma}_a^{(j)}, \tag{1}$$

where $a, b \in x, y, z$ denote the indices of Pauli matrices and i, j denote different spins on the lattice. The couplings J_{ab}^{ij} can be either short range of long range, and both J_{ab}^{ij} and B_a^j can generally be time-dependent.

A. Operator Cluster TWA

Here we describe the necessary steps for implementing the operator CTWA.

1. Split the system into clusters labeled by indices i', j', e.g. like shown in Fig. 2. Define a complete operator basis $\{\hat{X}_\alpha^{i'}\}$, $\alpha = 0, \ldots D^2 - 1$ spanning the Hilbert space of each cluster i', where D is the Hilbert space dimension of that cluster: $D = 2^N$ for a cluster of N spins $\frac{1}{2}$. Two examples of a complete basis are products of Pauli matrices (section II C 1) and all rank-1 Hermitian matrices (section V C). We will use the convention that $\hat{X}_0^{i'} = \hat{\mathbf{I}}$ is the identity operator and all the remaining operators $\hat{X}_\alpha^{i'}$ are traceless. Furthermore, we require that the basis operators are trace-orthogonal:

$$\text{Tr}\left[\hat{X}_\alpha^{i'} \hat{X}_\beta^{j'}\right] = D\delta_{\alpha\beta}\delta_{i'j'}. \tag{2}$$

In this way any operator $\hat{O}^{i'}$ within a cluster i' can be written as a linear combination of the basis operators:

$$\hat{O}^{i'} = \sum_\alpha o_\alpha \hat{X}_\alpha^{i'}. \tag{3}$$

Note that the choice of clusters is not unique. Moreover a particular cluster choice can break some underlying symmetries like translational or (discrete) rotational invariance. In this case one should simply average the results over different clustering choices, which effectively restores these symmetries.

2. Define structure constants $f_{\alpha\beta\gamma}$:

$$[\hat{X}_\alpha^{i'}, \hat{X}_\beta^{j'}] = i f_{\alpha\beta\gamma} \delta^{i'j'} \hat{X}_\gamma^{i'}$$
$$\Leftrightarrow i f_{\alpha\beta\gamma} = \frac{1}{D} \text{Tr}[\hat{X}_\gamma^{i'} [\hat{X}_\alpha^{i'}, \hat{X}_\beta^{i'}]]. \tag{4}$$

These two definitions are obviously equivalent. The second definition has an advantage that it can even be applied to basis operators which do not form a complete basis. Then structure constants define projectors of the commutators to the basis spanned by $\{\hat{X}_\alpha^{i'}\}$.

3. Rewrite the Hamiltonian $\hat{H}$ in terms of cluster operators

$$\hat{H} = \sum_{i'j'} \tilde{J}_{\alpha\beta}^{i'j'} \hat{X}_\alpha^{i'} \hat{X}_\beta^{j'} + \sum_{i'} \tilde{B}_\alpha^{i'} \hat{X}_\alpha^{i'}. \tag{5}$$

The new couplings $\tilde{J}$ and fields $\tilde{B}$ may be different then the base couplings; for example, local fields now include couplings between spins local to a cluster, as an operator $\hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(j)}$ is linear in $\hat{X}_\alpha^{i'}$ if both (i) and (j) belong to the same cluster i'.

4. Identify basis operators $\hat{X}_\alpha^{i'}$ to classical phase space variables $x_\alpha^{i'}$ satisfying canonical Poisson bracket relations defined through the structure constants $f_{\alpha\beta\gamma}$:

$$\hat{X}_\alpha^{i'} \to x_\alpha^{i'} - \frac{i}{2} x_\beta^{i'} f_{\alpha\beta\gamma} \frac{\partial}{\partial x_\gamma^{i'}}; \quad \{x_\alpha^{i'}, x_\beta^{j'}\} = -\delta_{i'j'} f_{\alpha\beta\gamma} x_\gamma^{i'}. \tag{6}$$

These Poisson brackets are obtained from a standard rule $i[\hat{A}, \hat{B}] \to \{A(\{x\}), B(\{x\})\}$, where A, B are functions in on phase space variables $\{x\}$.

5. Using the operator identification, define the Hamiltonian and any observables of interest in as functions of the classical phase space variables $\{x\}$, E.g. for an observable local to a particular cluster i'

$$\hat{O} = \sum_\alpha o_\alpha \hat{X}_\alpha^{i'} \to O_W(\{x\}) = \sum_\alpha o_\alpha x_\alpha^{i'}; \quad o_\alpha = \frac{1}{D} \text{Tr}[\hat{O} \hat{X}_\alpha^{(i')}],$$
$$\hat{H} \to H_W(\{x\}) = \sum_{i'j'} \tilde{J}_{\alpha\beta}^{i'j'} x_\alpha^{i'} x_\beta^{j'} + \sum_{i'} \tilde{B}_\alpha^{i'} x_\alpha^{i'}, \tag{7}$$

where index W indicates that this is the Weyl symbol corresponding to the symmetric ordering. For operators linear in the cluster variables Weyl ordering does

not play a role, but for operators nonlinear in $\hat{X}_\alpha^{(i')}$, which e.g. show up in various response functions it is important. Correct ordering always follows if one uses the Bopp representation (6).

6. Define an approximate Gaussian Wigner function $W(\{x\})$ describing initial conditions. For simplicity we only explicitly consider here pure initial states which factorize between clusters such that $W(\{x\}) = \prod_{i'} W_{i'}(\{x_\alpha^{i'}\})$

$$W_{i'}(\{x\}) = \frac{1}{Z} \exp\left[(x_\alpha - \rho_\alpha^{i'}) \Sigma_{\alpha\beta}^{i'} (x_\beta - \rho_\beta^{i'})\right]. \tag{8}$$

Determine the coefficients $\rho_\alpha^{i'}$ and $\Sigma_{\alpha\beta}^{i'}$ using the initial density matrix on the i'th cluster $\hat{\rho}^{i'}$ by fixing the expectation values and fluctuations of the basis operators:

$$\text{Tr}[\hat{\rho}^{i'} \hat{X}_\beta^{i'}] = \int \prod_\alpha dx_\alpha x_\beta W_{i'}(\{x\}),$$
$$\text{Tr}[\hat{\rho}^{i'} (\hat{X}_\beta^{i'} \hat{X}_\gamma^{i'} + \hat{X}_\gamma^{i'} \hat{X}_\beta^{i'})] = 2 \int \prod_\alpha dx_\alpha x_\beta x_\gamma W_{i'}(\{x\}) \tag{9}$$

Further discussion on this choice of Wigner function is given in section V A; fixing the terms in this way corresponds to matching quantum expectation values and fluctuations with their associated values in phase space. This choice guarantees asymptotic accuracy of CTWA at short times.

7. Solve the classical equations of motion for phase space variables:

$$\frac{dx_\alpha^{i'}(t)}{dt} = -\{x_\alpha^{i'}, H_W\} = f_{\alpha\beta\gamma} \frac{\partial H_W}{\partial x_\beta^{i'}} x_\gamma^{i'}. \tag{10}$$

Initial conditions $x_\alpha^{i'}(t=0)$ are randomly sampled from the Gaussian Wigner function and independently evolved. Note that these equations are identical to the Dirac mean-field equations if we assume that the density matrix is always factorized. Formally mean-field is recovered from the cluster TWA by setting the inverse of the fluctuation matrix $\Sigma_{\alpha\beta}$ in Eq. (8) to zero.

8. To find an expectation value of some observable $\hat{O}(t)$ in Heisenberg representation at some time $t > 0$, average the corresponding classical function for each point $\{x(t)\}_n$ in phase space, where n is the n-th sample from the Wigner function.

$$\langle \hat{O}(t) \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} O_W(\{x(t)\}_n) = \overline{O_W(\{x(t)\})} \tag{11}$$

The over-line represents averaging with respect to the initial conditions sampled by the Gaussian Wigner function.

4

9. If one is interested in the non-equal time correlation functions then instead of the previous step use

$$\langle \hat{X}_\alpha^{i'}(t_1)\hat{X}_\beta^{j'}(t_2) + \hat{X}_\beta^{j'}(t_2)\hat{X}_\alpha^{i'}(t_1)\rangle \rightarrow 2\overline{x_\alpha^{i'}(t_1)x_\beta^{j'}(t_2)}$$

$$i\langle[\hat{X}_\alpha^{i'}(t_1),\hat{X}_\beta^{j'}(t_2)]\rangle = f_{\alpha\gamma\nu}\, x_\gamma^{i'}(t_1)\overline{\frac{\partial x_\beta^{j'}(t_2)}{\partial x_\nu^{i'}(t_1)}}. \tag{12}$$

Here, the derivative stands for the response of $x_\beta^{j'}$ at time t_2 due to an infinitesimal perturbation of $x_\nu^{i'}(t_1)$ at time t_1: $x_\nu^{i'}(t_1) \rightarrow x_\nu^{i'}(t_1) + \varepsilon$. To preserve causality we assume that $t_1 \le t_2$, otherwise one should swap the operators $\hat{X}_\alpha^{i'}(t_1)$ and $\hat{X}_\beta^{j'}(t_2)$ and the corresponding phase space variables (see Ref. 4 for details). Using the Bopp operators, one may compute out of time order correlators in this fashion as well.

B. Wave function (Schwinger boson) CTWA

Instead of using D^2 phase space variables per cluster x_α, one can use the Schwinger boson representation of the operators $\hat{X}_\alpha$ and the associated phase space, which has dimensionality D. In section V D we show how a similar dimensional reduction can also be done for the operator CTWA. The essential steps for the associated TWA are very similar except that we operate with standard bosonic field Poisson brackets. Instead of the Schwinger boson representation one can use e.g. an angular momentum or any other representation of the basis operators.

1. Define the basis operators $\hat{X}_\alpha^{i'}$ in the same way as in the operator TWA (see item 1. in Sec. II A).

2. Define Schwinger boson representation of the basis operators:

$$\hat{X}_\alpha^{i'} = \sum_{a,b=0}^{D-1} \hat{b}_{a,i'}^\dagger T_\alpha^{ab}\hat{b}_{b,i'}, \tag{13}$$

where the matrices $\{T_\alpha\}$ form the fundamental representation of the algebra spanned by the operators $\hat{X}_\alpha^{i'}$. Here $\hat{b}_{b,i'}^\dagger$ is the raising operator for cluster i' and site b, corresponding to the b-th basis state $|b\rangle$. A simple way to define these matrices is through first fixing some basis within each cluster like

$$|0\rangle = |\uparrow,\uparrow,\dots\uparrow\rangle \tag{14}$$
$$|1\rangle = |\downarrow,\uparrow,\dots\uparrow\rangle$$
$$|D-1\rangle = |\downarrow,\downarrow,\dots\downarrow\rangle$$

and then defining

$$T_\alpha^{ab} = \langle a|\hat{X}_\alpha|b\rangle.$$

3. Represent the Hamiltonian (and all other observables of interest) through the Schwinger bosons

$$\hat{H} = \sum_{i'j'} \tilde{J}_{\alpha\beta}^{i'j'} T_\alpha^{ab} T_\beta^{cd} \hat{b}_{a,i'}^\dagger \hat{b}_{b,i'} \hat{b}_{c,j'}^\dagger \hat{b}_{d,j'} + \sum_{i'} \tilde{B}_\alpha^{i'} T_\alpha^{ab} \hat{b}_{a,i'}^\dagger \hat{b}_{b,i'}. \tag{15}$$

4. Associate Schwinger boson creation and annihilation operators with complex phase space amplitudes $\hat{b}_{a,i'}^\dagger \rightarrow b_{a,i'}^*$ and $\hat{b}_{a,i'} \rightarrow b_{a,i'}$ satisfying canonical Poisson bracket relations:

$$\{b_{a,i'}^*, b_{b,j'}\} = -i\delta_{ab}\delta_{i',j'}. \tag{16}$$

In this way, operators get mapped to functions in phase space $\hat{O} \rightarrow O_W(\vec{b},\vec{b}^*)$. For example the Hamiltonian (1) reads

$$H(\vec{b},\vec{b}^*) =$$
$$= \sum_{i'j'} J_{\alpha\beta}^{i'j'} T_\alpha^{ab} T_\beta^{cd} b_{a,i'}^* b_{b,i'} b_{c,j'}^* b_{d,j'} + \sum_{i'} B_\alpha^{i'} T_\alpha^{ab} b_{a,i'}^* b_{b,i'}. \tag{17}$$

5. Define an approximate Gaussian Wigner function in terms of Schwinger bosons. For the pure state $|0\rangle$ the Wigner function reads

$$W(\vec{b},\vec{b}^*) = \delta(|b_0|^2 - 1 - r_D)\prod_{a=1}^{D-1} e^{-\frac{b_a^* b_a}{r_D}},$$
$$r_D = \frac{\sqrt{1+D}-1}{D}. \tag{18}$$

Note that this Wigner function is not a standard Wigner function for Schwinger bosons, which would have a noise $1/2$ per each empty bosonic mode[4]. The phase space variables $b_{a,i'}$ can be interpreted as wave function amplitudes in the cluster basis $|0\rangle,\dots|D-1\rangle$, for a wave function which factorizes between clusters. For any other pure state related to $|0\rangle$ by some unitary rotation $|\psi_0\rangle = U|0\rangle$ the associated Wigner function can be obtained from the one in Eq. (18) by applying the unitary rotation to Schwinger bosons within a cluster: $\vec{b}_{i'} \rightarrow U\vec{b}_{i'}$. Because any unitary transformation is canonical (it preserves the Poisson brackets) the Wigner function and the equations of motion are invariant under basis transformations. More details on this rotation can be found in section V C. As in the operator case, formally mean-field can be obtained from TWA by setting the fluctuations in the Wigner function (18) to zero. Note that for the fluctuating initial state $\sum_a |b_{a,i'}|^2 > 1$ so the identification of $b_{a,i'}$ with the wave function amplitudes is only qualitative but the equations of motion are not affected by the normalization. More details are found in section V E.

6. Solve the classical (Gross-Pitaevskii) equations of motion for the phase space variables $\vec{b}$:

$$i\frac{\partial b_{a,i'}}{\partial t} = \frac{\partial H_w}{\partial b_{a,i'}^*}, \tag{19}$$

satisfying the initial conditions drawn from the initial Gaussian Wigner distribution. As in the operator TWA these equations of motion are identical to the Dirac mean-field equations of motion obtained under the assumption that the wave-function factorizes over clusters: $|\psi(t)\rangle = \prod_{i'} |\psi_{i'}(t)\rangle$.

7. To find the expectation value of some observable $\hat{O}(t)$ at some time $t > 0$ average the corresponding classical function for each point $(\vec{b}_n^*(t), \vec{b}_n(t))$ in phase space, where n is the nth sample from the Wigner function:

$$\langle \hat{O}(t) \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} O_W(\vec{b}_n(t), \vec{b}_n^*(t)) = \overline{O_W(\vec{b}(t), \vec{b}^*(t))}. \tag{20}$$

8. If one is interested in the non-equal time correlation functions of some observables $\hat{O}(t_1)$ and $\hat{Q}(t_2)$ then instead of the previous step use

$$\langle \{\hat{O}(t_1), \hat{Q}(t_2)\} \rangle \to \overline{2 O_W(\vec{b}(t_1), \vec{b}^*(t_1)) Q_W(\vec{b}(t_2), \vec{b}^*(t_2))} \tag{21}$$

$$i\langle [\hat{O}(t_1), \hat{Q}(t_2)] \rangle \to \overline{\{O_W(t_1), Q_W(t_2)\}}. \tag{22}$$

where $\{O_W(t_1), Q_W(t_2)\}$ is the usual non-equal time Poisson bracket (c.f. Eq. (46) in Ref. 16):

$$\{O_W(t_1), Q_W(t_2)\} =$$
$$i \sum_{a,i'} \frac{\partial O_W(t_1)}{\partial b_{a,i'}^*(t_1)} \frac{\partial Q_W(t_2)}{\partial b_{a,i'}(t_1)} - \frac{\partial O_W(t_1)}{\partial b_{a,i'}(t_1)} \frac{\partial Q_W(t_2)}{\partial b_{a,i'}^*(t_1)}. \tag{23}$$

C. Example: cluster TWA for four coupled spins.

Let us illustrate how one can apply cluster TWA to a simple system of four coupled spins shown in Fig. 3, which is described by the following Hamiltonian:

$$\hat{H} = J \sum_{j=1}^{3} \hat{\sigma}_z^{(j)} \hat{\sigma}_z^{(j+1)} + h_x \sum_{j=1}^{4} \hat{\sigma}_x^{(j)}. \tag{24}$$

In this example one can consider clusters of size one, two, and four. The size one cluster leads to standard TWA for spins[4], while the size four cluster leads to exact representation of dynamics by TWA. Size two clusters should be still approximate but lead to an improved accuracy of the method compared to the standard TWA, and will be detailed here, going through the steps highlighting how they are implemented in this example.

1. Operator cluster TWA

1. Our model system is split into two size two clusters, $1' \equiv L$ and $2' \equiv R$. The Hilbert space dimension of the cluster is $D = 4$ and the dimension of the operator basis spanning $SU(4)$ algebra plus identity is $D^2 = 16$. A possible and convenient choice for the operator basis
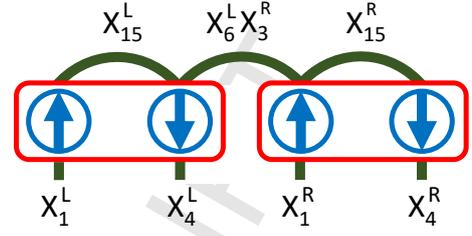

FIG. 3. **Representation of the 4-spin Ising example**. Blue circles and arrows represent the physical spins in the initial Neel state, while red rectangles show the clusters labeled by $i' = L, R$. Green arches represent the nearest-neighbor $\hat{\sigma}_z^{(j)} \hat{\sigma}_z^{(j+1)}$ couplings, with the relevant representation through the cluster operators $\hat{X}_\alpha^{L,R}$. Bottom lines represent the on-site $\hat{\sigma}_x^{(j)}$ fields in terms of the cluster operators. In this way the original Hamiltonian (24) exactly maps to the two-cluster Hamiltonian (26)

in the first cluster (and similarly in the second cluster) is:

$$\hat{X}_0 = \hat{I}^{(1)} \otimes \hat{I}^{(2)}$$
$$\hat{X}_1 = \hat{\sigma}_x^{(1)} \otimes \hat{I}^{(2)}, \quad \hat{X}_2 = \hat{\sigma}_y^{(1)} \otimes \hat{I}^{(2)}, \quad \hat{X}_3 = \hat{\sigma}_z^{(1)} \otimes \hat{I}^{(2)}$$
$$\hat{X}_4 = \hat{I}^{(1)} \otimes \hat{\sigma}_x^{(2)}, \quad \hat{X}_5 = \hat{I}^{(1)} \otimes \hat{\sigma}_y^{(2)}, \quad \hat{X}_6 = \hat{I}^{(1)} \otimes \hat{\sigma}_z^{(2)}$$
$$\hat{X}_7 = \hat{\sigma}_x^{(1)} \otimes \hat{\sigma}_x^{(2)}, \quad \hat{X}_8 = \hat{\sigma}_x^{(1)} \otimes \hat{\sigma}_y^{(2)}, \quad \hat{X}_9 = \hat{\sigma}_x^{(1)} \otimes \hat{\sigma}_z^{(2)}$$
$$\hat{X}_{10} = \hat{\sigma}_y^{(1)} \otimes \hat{\sigma}_x^{(2)}, \quad \hat{X}_{11} = \hat{\sigma}_y^{(1)} \otimes \hat{\sigma}_y^{(2)}, \quad \hat{X}_{12} = \hat{\sigma}_y^{(1)} \otimes \hat{\sigma}_z^{(2)}$$
$$\hat{X}_{13} = \hat{\sigma}_z^{(1)} \otimes \hat{\sigma}_x^{(2)}, \quad \hat{X}_{14} = \hat{\sigma}_z^{(1)} \otimes \hat{\sigma}_y^{(2)}, \quad \hat{X}_{15} = \hat{\sigma}_z^{(1)} \otimes \hat{\sigma}_z^{(2)},$$

where superscripts 1 and 2 refer to the first and the second site of the cluster. It is easy to check that these basis operators satisfy the required normalization conditions (2).

2. The structure constants are easy to get from the commutation relations of the Pauli matrices. There are in total 20 independent structure constants in this operator basis. Let us explicitly show a few of them:

$$[\hat{X}_1, \hat{X}_2] = 2i\hat{X}_3 \Rightarrow f_{1,2,3} = 2, \tag{25}$$
$$[\hat{X}_7, \hat{X}_{10}] = 2i\hat{X}_3 \Rightarrow f_{7,10,3} = 2.$$

3. In terms of the cluster operators the Hamiltonian (24) reads

$$\hat{H} = J \left(\hat{X}_{15}^L + \hat{X}_{15}^R + \hat{X}_6^L \hat{X}_3^R \right) \tag{26}$$
$$+ h_x \left(\hat{X}_1^L + \hat{X}_4^L + \hat{X}_1^R + \hat{X}_4^R \right).$$

Note that this Hamiltonian contains only a single nonlinear coupling between the operators $\hat{X}_6^L$ and $\hat{X}_3^R$. When written in terms of Pauli matrices (single site clusters) the Hamiltonian has three non-linear couplings.

4. Identify fifteen phase space variables per each cluster $x_1^{L,R} \dots x_{15}^{L,R}$ satisfying the Poisson bracket relations with the structure constants determined above.

5. Define the classical Hamiltonian H_W:

$$H_W = J\left(x_{15}^L + x_{15}^R + x_6^L x_3^R\right) \tag{27}$$
$$+ h_x\left(x_1^L + x_4^L + x_1^R + x_4^R\right). \tag{28}$$

6. Let us consider an initial product state

$$|\psi_0\rangle = |\uparrow, \downarrow, \uparrow, \downarrow\rangle,$$

which is an eigenstate of the Hamiltonian with $h_x = 0$. Then our setup can be regarded as a quench from this initial state. The Gaussian Wigner function for this state factorizes into a product of left and right Wigner functions W_L and W_R. For the left cluster (and similar for the right cluster) the nonzero expectation values have only the following three operators:

$$\langle\psi_0|\hat{X}_3^L|\psi_0\rangle = 1, \ \langle\psi_0|\hat{X}_6^L|\psi_0\rangle = -1, \ \langle\psi_0|\hat{X}_{15}^L|\psi_0\rangle = -1. \tag{29}$$

There are 24 non-zero symmetric correlation functions excluding the identity for the left cluster (and similar for the right cluster):

$$\left\langle\psi_0\right|\left(\hat{X}_\alpha^L\right)^2\left|\psi_0\right\rangle = 1, \quad \alpha \in \{1,\ldots,15\} \tag{30}$$

$$\frac{1}{2}\langle\psi_0|\{\hat{X}_\alpha^L, \hat{X}_\beta^L\}_+|\psi_0\rangle = 1$$
$$(\alpha, \beta) \in \left[(4,13)\ (5,14)\ (6,15)\ (7,11)\right]$$
$$\frac{1}{2}\langle\psi_0|\{\hat{X}_\alpha^L, \hat{X}_\beta^L\}_+|\psi_0\rangle = -1$$
$$(\alpha, \beta) \in \left[(1,9)\ (2,12)\ (3,6)\ (3,15)\ (8,10)\right],$$

where $\{\hat{A}, \hat{B}\}_+$ stands for anti-commutator. The fact that

$$\langle\psi_0|\hat{X}_8^2|\psi_0\rangle = \langle\psi_0|\hat{X}_{10}^2|\psi_0\rangle = -\frac{1}{2}\langle\psi_0|\{\hat{X}_8, \hat{X}_{10}\}_+|\psi_0\rangle = 1$$

implies that the phase space variables x_8 and x_{10} should be perfectly anti-correlated. After finding the connected correlators, the associated Gaussian Wigner functions W_L and W_R read

$$W_L = \frac{1}{Z}\delta(x_3-1)\delta(x_6+1)\delta(x_{15}+1)\delta(x_4-x_{13})\delta(x_5-x_{14})$$
$$\times \delta(x_1+x_9)\delta(x_2+x_{12})\delta(x_7-x_{11})\delta(x_8+x_{10})$$
$$\times \exp\left[-\frac{\sum_{\alpha\in\{1,2,4,5,7,8\}}x_\alpha^2}{2}\right], \tag{31}$$

where Z is the normalization constant. The δ-functions can be understood as limiting cases of Gaussians and thus are allowed in our ansatz. We see that each initial condition must draw 12 uncorrelated numbers (6 per each cluster).

7. Now we need to solve the system of classical equations of motion (10) for our phase space variables with the initial conditions $x_\alpha^{L,R}(t=0)$ randomly sampled

from the left and right Wigner functions (31). There are overall thirty equations: fifteen for left variables and fifteen for right variables (the variables $x_0^{L,R}$ corresponding to identity operators are obviously conserved in time). Let us explicitly show a few of them:

$$\partial_t x_4^L = J x_3^R x_5^L + J x_{14}^L \tag{32}$$
$$\partial_t x_3^L = -h x_2^L$$
$$\partial_t x_{12}^L = -h x_{11}^L + h x_{15}^L - J x_1^L.$$

8. We can now compute the expectation values of observables by averaging the corresponding Weyl symbols computed on time-dependent phase space points over the initial conditions. For example:

$$\langle\hat{\sigma}_z^{(1)}(t)\hat{\sigma}_z^{(2)}(t)\rangle = \langle\hat{X}_{15}^L(t)\rangle \approx \overline{x_{15}^L(t)}, \tag{33}$$
$$\langle\hat{\sigma}_z^{(2)}(t)\hat{\sigma}_z^{(3)}(t)\rangle = \langle\hat{X}_6^L(t)\hat{X}_3^R(t)\rangle \approx \overline{x_6^L(t)x_3^R(t)},$$
$$\langle\hat{\sigma}_z^{(1)}(t)\rangle = \langle\hat{X}_3^L(t)\rangle \approx \overline{x_3^L(t)}.$$

9. In order to compute a non-equal time correlation function, for example of the z-magnetization on the first site, we can use Eq. (12) identifying $\hat{\sigma}_z^{(1)} \to x_3^L$:

$$\langle\{\hat{\sigma}_z^{(1)}(t), \hat{\sigma}_z^{(1)}(0)\}_+\rangle = 2\overline{x_3^L(t)x_3^L(0)} \tag{34}$$
$$i\langle[\hat{\sigma}_z^{(1)}(0), \hat{\sigma}_z^{(1)}(t)]\rangle = f_{3\mu\nu}\overline{\left(x_\mu^L(0)\frac{\partial x_3^L(t)}{\partial x_\nu^L(0)}\right)}.$$

In Fig. 4 we show the results of simulations using the cluster TWA. For comparison we also show the single cluster (conventional) TWA and exact results. For simulations we choose specific parameters: $J = \frac{1}{8}$, $h_x = 1$.

2. Wave function CTWA

This same example system can be done under wavefunction (Schwinger Boson) CTWA.

1. First, define the 15+1 basis operators for each of the two clusters in the same fashion as the Operator cluster TWA example. Define these operators as 15 4x4 matrices $\hat{X}_\alpha \to T_\alpha^{ab}\hat{a}_a^\dagger\hat{a}_b$ with complex coordinates for (L/R) of $((a,a^*)/(b,b^*))$. For example, the operators:

$$\hat{X}_4^L = \hat{I}^{(1)} \otimes \hat{\sigma}_x^{(2)} \to \begin{bmatrix}\hat{a}_0^\dagger \\ \hat{a}_1^\dagger \\ \hat{a}_2^\dagger \\ \hat{a}_3^\dagger\end{bmatrix}^T \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix}\hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3\end{bmatrix}, \tag{35}$$

$$\hat{X}_8^L = \hat{\sigma}_x^{(1)} \otimes \hat{\sigma}_y^{(2)} \to \begin{bmatrix}\hat{a}_0^\dagger \\ \hat{a}_1^\dagger \\ \hat{a}_2^\dagger \\ \hat{a}_3^\dagger\end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix}\hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3\end{bmatrix}.$$
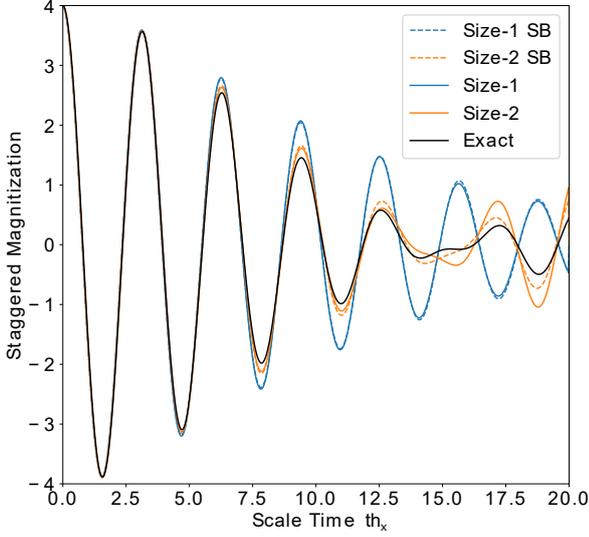
FIG. 4. **Dynamics of the staggered magnetization in the 4-spin Transverse Ising example** (see text for details). Solid black represents the exact result, while the solid (dashed) colored lines are the results of the Operator (Schwinger Boson) CTWA. The staggered magnetization is represented by $\overline{x_3^L + x_3^R - x_6^L - x_6^R}$. CTWA dynamics approach exact results with increasing cluster size.

The classical Hamiltonian H_W becomes:

$$H_W(a, b) = (a_i^* a_j + b_i^* b_j)(JT_{15}^{ij} + h_x T_3^{ij} + h_x T_4^{ij})$$
$$+ J a_i^* a_j b_k^* b_l T_6^{ij} T_3^{kl}. \quad (36)$$

2. Draw initial conditions for $(\vec{a}, \vec{b})$. For the state $|\uparrow\downarrow\uparrow\downarrow\rangle$ the initial condition is $|\uparrow\downarrow\rangle$ per cluster corresponds to $a_1, b_1 = \sqrt{1 + \delta_0}$ and a_q, b_q with $q = 0, 2, 3$ random numbers drawn from the Gaussian probability distribution (18) with zero mean and the variance $\delta_0 = \frac{\sqrt{1 + 2^2} - 1}{2^2} \approx 0.30902$. Note that because the initial state here is $\uparrow\downarrow$ and not $|\uparrow\uparrow\rangle$, it is the Schwinger bosons a_1, b_1 (not a_0, b_0), which are special. Similarly for $|\downarrow\uparrow\rangle$ ($|\downarrow\downarrow\rangle$) states the bosons a_2, b_2 (a_3, b_3) will be special.

3. Evolve each point $(\vec{a}(t), \vec{b}(t))$ via Gross-Pitaevskii equations (19), which explicitly read:

$$i \frac{\partial}{\partial t} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} J & h & h & 0 \\ h & -J & 0 & h \\ h & 0 & -J & h \\ 0 & h & h & J \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$
$$+ J(|b_0|^2 + |b_1|^2 - |b_2|^2 - |b_3|^2) \begin{bmatrix} a_0 \\ -a_1 \\ a_2 \\ -a_3 \end{bmatrix} \quad (37)$$

and similarly for the b-bosons.

4. We can compute expectation values of observables by averaging the corresponding Weyl symbols computed on time-dependent phase space points over the initial conditions. For example:

$$\langle \hat{\sigma}_z^{(1)}(t) \hat{\sigma}_z^{(2)}(t) \rangle \approx \overline{a_a^* a_b} T_{15}^{ab}, \quad (38)$$
$$\langle \hat{\sigma}_z^{(2)}(t) \hat{\sigma}_z^{(3)}(t) \rangle \approx \overline{a_a^* a_b b_x^* b_y} T_6^{ab} T_3^{xy},$$
$$\langle \hat{\sigma}_z^{(1)}(t) \rangle \approx \overline{a_a^* a_b} T_3^{ab}.$$

The results of the simulations of the Schwinger boson CTWA are shown in Fig. (4) using dashed lines. They are almost indistinguishable from the operator CTWA results.

D. CTWA and the Dirac Mean Field Approximation.

The basic idea of the CTWA (as of any other TWA) is best summarized in Fig. 1. One first generates an initial condition drawn from a Gaussian distribution (8) or (18). Then the phase space variables evolve in time according to the semiclassical (or equivalently mean field) equations of motion (10) or (19). After that, one evaluates observables of interest on these time evolved trajectories and repeats the procedure; effectively averaging over initial conditions. Note that even though classical and mean field equations are identical, TWA is not equivalent to the mean field approximation and generally outperforms mean field in non-linear (interacting) systems. TWA reduces to the mean-field approximation only if the initial Winger function contracts to a point in phase space and thereby suppress the initial quantum noise to zero.

The reason for such nonequivalence between TWA and mean-field approaches in non-linear systems is that averaging over initial conditions and time evolution are non-commuting operations. Therefore time evolving an average trajectory is generally very different from first solving nonlinear equations of motion for randomly chosen initial conditions and then averaging the observable of over the initial noise. In fact using short time perturbation theory one can show that CTWA correctly reproduces exact dynamics of observables up to the order t^2, while the mean field dynamics is accurate only up to the order t (For more details, see section V A). This observation also justifies sufficiency of the Gaussian initial conditions because even if we improve further the Wigner function matching the higher initial moments of phase point operators, the accuracy of CTWA will remain the same because there will be corrections of the order of t^3 coming from quantum jumps, which are beyond TWA[4].

In situations where one has large statistical fluctuations, e.g. the initial density matrix represents an infinite temperature state like it was done in Ref. 11 the statistical noise dominates over the quantum noise. This is reflected in the Wigner function of the infinite temperature density matrix being essentially completely random. Even in those cases, we believe, referring to the method proposed by the authors as variational mean field approximation is somewhat misleading, because each initial condition is evolved independently

and the averaging is done in the end. Mean field dynamics, by construction implies that a single initial condition represents the dynamics of the whole system.

CTWA is also similar in spirit to the variational approach proposed in Ref. 12 if we assume that the Wigner function remains Gaussian at all time. Then instead of averaging over initial conditions one simply has to propagate both average and variance of phase space variables. We comment, however, that this would introduce another approximation into the CTWA method, because for a nonlinear evolution any initial Gaussian distribution function becomes non-Gaussian. In some situations, for example where classical dynamics becomes unstable, the Gaussian approximation can become very inaccurate. A simple example where their approximation will fail would be dynamics of a particle initialized on the top of a Mexican hat potential (see Refs. 4 and 17). In other situations the Gaussian approximation for the Wigner function might remain accurate even after long times.

III. APPLICATION OF CTWA TO THE NON-INTEGRABLE ISING MODEL

As a first nontrivial example, we present the method as applied to the Ising model in a transverse and longitudinal field, which is known to have a quantum chaotic regime. Here, we consider the setup similar to that of Ref. 11 where we initially polarize a single spin in the up state and follow time decay of the magnetization. However, unlike Ref. 11 we consider two setups of the "remaining" bath spins where i) they are prepared in a pure polarized state and ii) they are prepared in a mixed state. We consider the Hamiltonian:

$$\hat{H} = \sum_{i=0}^{19} \hat{\sigma}_x^{(i)}\hat{\sigma}_x^{(i+1)} + 0.8090\hat{\sigma}_x^{(i)} - 0.9045\hat{\sigma}_z^{(i)}, \quad (39)$$

with periodic boundary conditions. This set of couplings is known to lead to the chaotic Hamiltonian satisfying the eigenstate thermalization hypothesis[18]. To benchmark our CTWA results we compare them to exact quantum dynamics for a system of 20 spins. For the pure state initial condition we choose $|\psi\rangle = |\uparrow\downarrow\downarrow\ldots\downarrow\rangle$, i.e. all spins except the first one are polarized along the $-z$ direction. For the mixed state initial condition we uniformly sample all the bath spins from the Bloch sphere keeping always the first spin polarized along the positive z-axis. The latter setup is identical to the one studied in Ref. 11.

In Fig. 5 we show time dependence of the z-magnetization of the initially z-polarized spin for the pure state of the bath (top) and the mixed state of the bath (bottom). For the pure initial state the CTWA reproduces short-time coherent oscillations, as well as long-time diffusive decay and thermodynamic equilibrium. The only thing CTWA misses are the intermediate time oscillations of the magnetization; this mistake decreases for larger clusters. The agreement of CTWA with exact dynamics is even better for the mixed initial state of the bath.

While the accuracy of CTWA at short times can be generally expected as in standard TWA approximations, correctly capturing the long time dynamics is highly nontrivial as TWA often leads to an uncontrolled error[4]. Intuitively this feature of CTWA is not that surprising as cluster variables correctly capture short-distance quantum correlations, while the long time/long distance dynamics is generally expected to be classical. At least for thermalizing, nonintegrable, systems CTWA is expected to predict accurately both short time and long time results with the increasing cluster size as seen in the figure (see also Ref. 6, where similar long time accuracy of the fermionic TWA was observed).

Short time dynamics. Like the traditional TWA (see Ref. 4 for details) CTWA is asymptotically exact at short times, up to the order $\mathcal{O}(t^2)$. This is guaranteed by the short time perturbation theory for all observables. Specifically any short time response of an observable $\hat{O}$ can be Taylor expanded in time, as is detailed in section V A. Similarly, one may heuristically argue that dynamics are exact until entanglement between clusters, which is treated approximately, become relevant.

Long time dynamics. In CTWA we generically deal with classical nonlinear and hence chaotic systems. For linear systems CTWA is guaranteed to be exact at all times and the case of nonlinear integrable systems goes beyond the scope of this paper and requires a special attention. Then the long time steady state is expected to be described by the thermal equilibrium:

$$\langle\hat{O}(t)\rangle_{t\to\infty} \to= Z^{-1}\int d\vec{x}\,e^{-\beta H_W(\{x\})}O_W(\{x\}), \quad (40)$$

where β is the temperature set by the initial energy of the system. At least for high temperatures this Gibbs distribution is equivalent to the quantum thermal distribution $\langle\hat{O}\rangle = \text{Tr}[\hat{O}e^{-\beta(E)\hat{H}}]$. Moreover the long time approach to the thermal equilibrium, is usually described within the classical hydrodynamic framework and thus is compatible with CTWA. While we do not have more mathematically rigorous arguments of asymptotic equivalence of thermal classical (in terms of cluster variables) and quantum distributions, we observed that in all ergodic regimes as we increase the cluster size CTWA correctly predicts long time behavior of chaotic quantum systems.

Intermediate time dynamics. CTWA may miss intermediate time dynamics of the system, as illustrated in Fig. 5. The mistake is usually can be controlled by increasing the cluster size. For sufficiently large clusters the system might enter the classical hydrodynamic regime before the mistake due to the truncation of quantum dynamics kicks in. Then the CTWA becomes essentially exact at all times like it e.g. happens for the 8-site CTWA in the bottom plot of Fig. 5. This usually happens at high temperatures and away from integrability. Close to the ground state the coherent quantum dynamics is expected to persist for very long times and as a result one needs to use very large clusters to capture dynamics correctly at long times. This can be potentially achieved by combining CTWA with DMRG techniques and is a subject of future work.
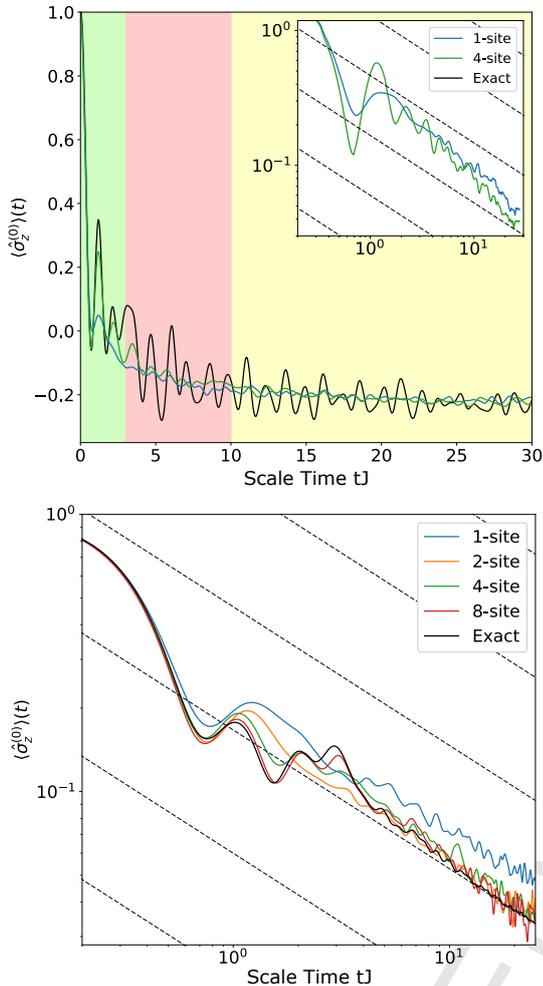
FIG. 5. **Magnetization decay in a spin chain.** Time dependence of $\langle \hat{\sigma}_z^{(0)}(t) \rangle$ of the initially polarized spin coupled to other spins forming a "bath". The top plot corresponds to the initial pure state $|\uparrow\downarrow\ldots\downarrow\rangle$. The bottom plot corresponds to the mixed, infinite temperature, state of the bath, where all spins except one are prepared in a random initial state. The system size is $N = 20$ for the pure state and $N = 16$ for the mixed state; the Hamiltonian of the system is given by Eq. (39). The inset in the top plot shows $\langle \hat{\sigma}_z^{(0)}(t) \rangle - 1/N \sum_j \langle \sigma_z^{(j)} \rangle$. Both setups show initial coherent oscillations of the magnetization followed by long-time diffusive relaxation to the thermal state. The accuracy of the CTWA clearly improves with the cluster size. Dashed lines are $1/\sqrt{t}$ diffusive asymptotes.

IV. APPLICATION OF CTWA TO THE DISORDERED 1D HEISENBERG CHAIN

Another system to test CTWA is the 1d Heisenberg model with evenly distributed disorder in the z-component of the magnetic field and periodic boundary conditions:

$$\hat{H} = J \sum_{\langle ij \rangle} \left(\hat{\sigma}_x^{(i)} \hat{\sigma}_x^{(j)} + \hat{\sigma}_y^{(i)} \hat{\sigma}_y^{(j)} + \hat{\sigma}_z^{(i)} \hat{\sigma}_z^{(j)} \right) + h_z \sum_{i=1}^{N} \delta^i \hat{\sigma}_z^{(i)},$$

$$(41)$$

where $\delta^i \in [-1, 1]$. It is well known phenomenologically that this model undergoes a transition to the Many Body Localized (MBL) phase[19] above some critical disorder strength $h_z \approx 3.5$. This phase is characterized, amongst other things, by a "memory" of initial conditions which persists indefinitely, as well as a logarithmic growth of entanglement in time after a quench. A natural order parameter to measure the transition to MBL is the staggered z-magnetization $\hat{M}_z$[11,19], which is maximal at $t = 0$ for a Neel initial state $|\psi_0\rangle$:

$$|\psi(0)\rangle = |\ldots \uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \ldots\rangle \quad ; \quad \hat{M}_z = \sum_{i=1}^{N} (-1)^i \hat{\sigma}_z^{(i)}. \quad (42)$$

If the system thermalizes then this order parameter decays to zero; while in MBL non-thermal regime it decays to some non-zero value.

Because averaging over disorder and quantum fluctuations are commuting operations we can simultaneously update initial conditions according to the Wigner function and the disorder realization of the Hamiltonian. This allows one to improve numerical convergence of the results and parallelize computations if necessary.

In Fig. 6 we plot the normalized staggered magnetization

$$m_z = \frac{1}{N} \langle \psi(t) | \hat{M}_z | \psi(t) \rangle,$$

for the initial Neel state and different strengths of disorder. The left two plots show the results of the operator CTWA and the right plots show wave function (Schwinger Boson) CTWA. Both methods clearly lead to nearly identical curves. On the top two plots we show results for a relatively small system size $N = 16$, which allows for comparison with exact diagonalization. The bottom plots show CTWA results for a larger system size $N = 64$.

At short times, CTWA reproduces very well exact quantum dynamics. Accuracy of CTWA clearly improves with the cluster size. CTWA works well both at low disorder and at high disorder, but it clearly fails to reproduce localization at intermediate values of disorder. This is consistent with earlier work[20,21] observing that classical disordered spin chains in thermodynamic limit eventually show ergodic diffusive behavior. Our numerical results (bottom plots) indeed indicate that irrespective of the cluster size long time dynamics with CTWA approximation is always diffusive. Nevertheless CTWA indicates presence of MBL regime showing longer and longer localized prethermalization plateaus with the increasing cluster size.

This example illustrates limitations of CTWA to correctly capture long time non-ergodic behavior. We believe that this is not a fundamental limitation though and even in this case CTWA can be improved by choosing a better operator basis, for example using not cluster operators constructed from Pauli matrices but from so called l-bits[22], which also form a complete operator basis or using the basis of the matrix product operators.[23]

A. Entanglement entropy within CTWA

Interestingly CTWA allows one not only to extract expectation values of local observables but also get information about
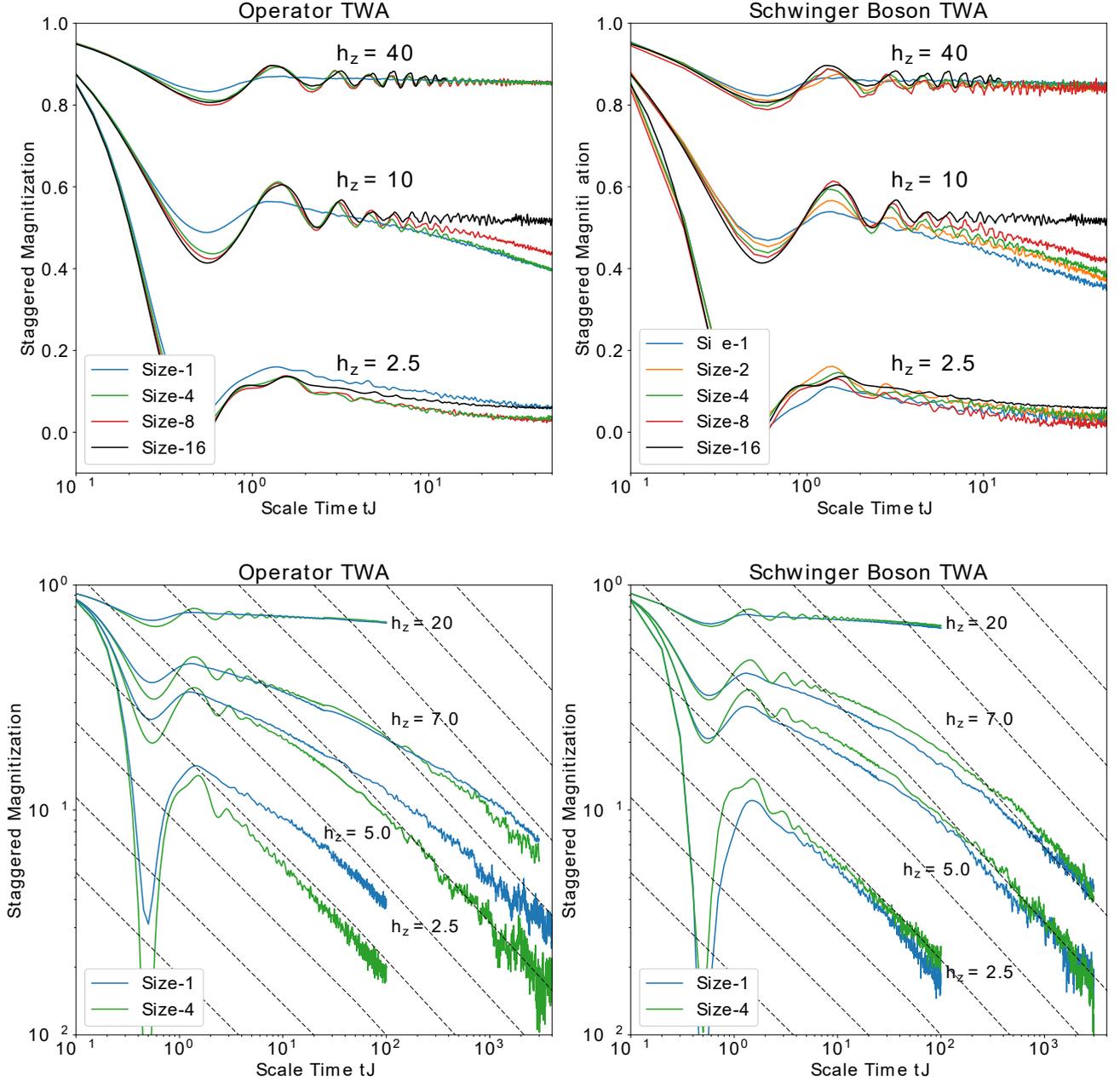
FIG. 6. **Dynamics of the staggered magnetization in the 1d Disordered Heisenberg Chain**. TOP: Comparison of exact dynamics vs operator CTWA (left) and Schwinger Boson CTWA (Right) for varying cluster sizes and a 16-site Heisenberg chain (see text for details). BOTTOM: long time decay of the staggered magnetization within the cluster CTWA (Left) and Schwinger Boson CTWA (Right) for a 64-site 1d Heisenberg chain. Dashed lines are $1/\sqrt{t}$ diffusive asymptotes.

entanglement of subsystems. In order to compute the entanglement entropy we need to treat the full density matrix of a subsystem as an observable. For simplicity we will only consider the density matrices which are either confined to a single cluster or span two clusters. Any density matrix confined to a cluster can be represented as

$$\hat{\rho}^{i'}(t) = \sum_{\alpha} c_{\alpha}^{i'}(t) \hat{X}_{\alpha}^{i'},$$

with some time dependent coefficient. From the orthonormality of the basis we obtain

$$c_{\alpha}^{i'} = \frac{1}{D}\text{Tr}[\hat{\rho}^{i'}(t)\hat{X}_{\alpha}^{i'}] = \frac{1}{D}\text{Tr}[\hat{\rho}(t)\hat{X}_{\alpha}^{i'}]$$

$$= \frac{1}{D}\text{Tr}[\hat{\rho}\hat{X}_{\alpha}^{i'}(t)] = \frac{1}{D}\langle \hat{X}_{\alpha}^{i'}(t)\rangle \approx \frac{1}{D}\overline{x_{\alpha}^{i'}(t)},$$

where the last equality holds only within the accuracy of CTWA approximation. Therefore we can rewrite the reduced
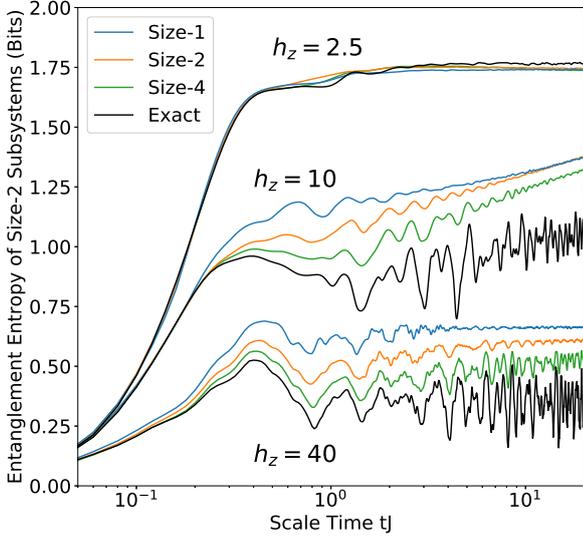
FIG. 7. **Entropy Growth in the Disordered Heisenberg Model**
Showing the Von Neumann entropy (equation (45)) for size-2 subsystems of a size-16 Heisenberg chain and fixed disorder realization. Entropy is averaged over all adjacent size-2 subsystems. Black is exact dynamics, colored are CTWA dynamics for various cluster sizes. All cluster offsets were used, (eg, 4 offsets for size-4 clusters). Both intra (equation (43)) and inter (equation (44)) cluster subsystems were averaged. Clearly the entanglement entropy approaches the exact result for larger cluster sizes.

density matrix of a cluster (or similarly any sub-cluster of a larger cluster) as

$$\hat{\rho}^{i'}(t) \approx \frac{1}{D} \sum_{\alpha} \overline{x_{\alpha}^{i'}(t)} \hat{X}_{\alpha}^{i'}. \tag{43}$$

Likewise for a reduced density spanning two clusters (or similarly two sub-clusters) we find

$$\hat{\rho}^{i'j'} \approx \frac{1}{D^2} \sum_{\alpha\beta} \overline{x_{\alpha}^{i'}(t) x_{\beta}^{j'}(t)} \hat{X}_{\alpha}^{i'} \otimes \hat{X}_{\beta}^{j'}. \tag{44}$$

Note that generically for any interacting (nonlinear) Hamiltonian the averages $\overline{x_{\alpha}^{i'}(t) x_{\beta}^{j'}(t)}$ do not factorize so one can get entanglement between clusters even within CTWA like it was recently analyzed in Ref. 21. We note that to get the inter-cluster entanglement the presence of the noise is crucial because in any noiseless mean-filed type approximations the equality $\overline{x_{\alpha}^{i'}(t) x_{\beta}^{j'}(t)} = \overline{x_{\alpha}^{i'}(t)}\, \overline{x_{\beta}^{j'}(t)}$ always holds. The entanglement entropy of the particular subsystem A, which can include one or more clusters, is evaluated according to the standard rule:

$$S^{i'}(t) = \text{Tr}[\hat{\rho}_A(t) \log_2(\hat{\rho}_A(t))]. \tag{45}$$

In Fig. 7 we show the calculated entropy density for the disordered Heisenberg model of size $N = 16$. At low disorder, the reduced entropy density is known to grow ballistically ($\sim t$) before saturating at a thermal value. At high
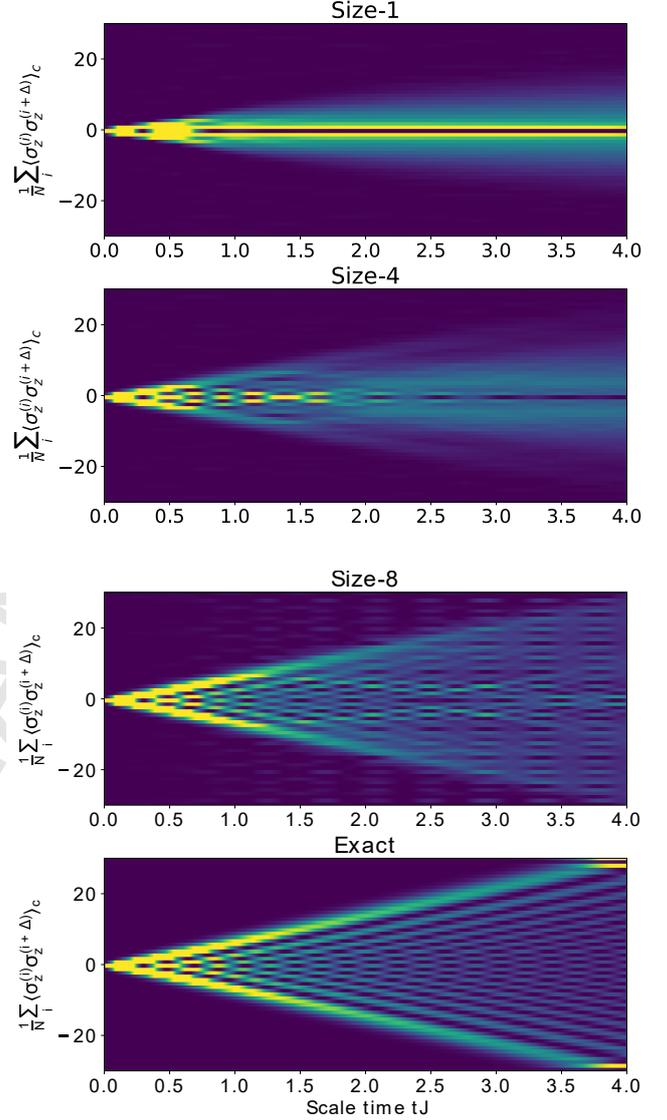

FIG. 8. **The connected correlator for the 1d XY spin chain** as outlined in Eqs. (46), (49). Ballistic growth of entanglement is well encoded in the correlations of the semiclassical phase space, improving with cluster size. Color encodes strength of the expectation value of the connected correlator $\langle \hat{\sigma}_z^{(i)} \sigma_z^{(i+\Delta)} \rangle_c$.

disorder, i.e. in the MBL regime, the entanglement entropy shows initial growth followed by intermediate time saturation at a prethermal plateau and then eventual logarithmic growth, consistent with the results of the exact diagonalization (see the review. 24 and refs. therein for further details). As with other observables accuracy of CTWA improves with increasing cluster size. In contrast to the exact dynamics, the slow logarithmic growth of the entanglement entropy in the CTWA goes hand in hand with decay of the order parameter. For this reason the CTWA likely overestimates the slope of asymptotic growth of the entropy with $\log(t)$.

Using Eq. (11) CTWA also allows one to compute various local and non-local correlation functions. Suppose we want

compute the equal time connected correlation function of z-magnetization between sites i and j. If these sites belong to different clusters then

$$\langle \hat{\sigma}_z^{(i)}(t) \hat{\sigma}_z^{(j)}(t) \rangle - \langle \hat{\sigma}_z^{(i)}(t) \rangle \langle \hat{\sigma}_z^{(i+\Delta)}(t) \rangle \quad (46)$$

$$= \langle \hat{X}_\alpha^{(i')}(t) \hat{X}_\beta^{(j')}(t) \rangle - \langle \hat{X}_\alpha^{(i')}(t) \rangle \langle \hat{X}_\beta^{(j')}(t) \rangle \quad (47)$$

$$= \overline{x_\alpha^{i'}(t) x_\beta^{j'}(t)} - \overline{x_\alpha^{i'}(t)} \; \overline{x_\beta^{j'}(t)}. \quad (48)$$

where i' (j') is the cluster containing site i (j) and α and β are the indices corresponding to representation of Pauli matrices through the cluster basic operators (in general the observables of interest can be represented through linear combinations of the basis operators). If the sites i and j belong to the same cluster i' then computing the correlation function is even simpler as the whole product $\langle \hat{\sigma}_z^{(i)}(t) \hat{\sigma}_z^{(j)}(t) \rangle$ is represented through a linear combination of the basis operators $\hat{X}_\alpha^{(i')}$.

These correlation functions are sensitive to entanglement/mutual information between different sites and like the entanglement entropy can not be captured within mean-field approximations as they arise, within CTWA, solely due to the initial fluctuations which propagate differently in space-time. In Fig. 8 we show this correlation function for the periodic 64-site XY model at zero disorder:

$$\hat{H} = \sum_{i=0}^{64} \hat{\sigma}_x^{(i)} \hat{\sigma}_x^{(i+1)} + \hat{\sigma}_y^{(i)} \hat{\sigma}_y^{(i+1)}. \quad (49)$$

This Hamiltonian is well-known to be integrable and mappable to free fermions. In fact under an alternate operator basis introduced in Ref. 6 the TWA dynamics for this model become exact. This emphasizes that basis choice $\{\hat{X}_\alpha^{i'}\}$ can be important for the accuracy of CTWA. The connected correlator clearly approaches the exact result for larger cluster sizes, and is nonzero even far outside the span of a cluster, indicating that entanglement information can be well encoded in the phase space.

V. CTWA DETAILS

In this section we elaborate some details of CTWA, starting from how one can define and sample the Gaussian Wigner function as well as detail implementing efficient CTWA dynamics. We will also provide some additional justifications for the choice of initial conditions and comment on connections between operator and Schwinger boson CTWA schemes.

A. The Gaussian Wigner function for operator CTWA.

We formally introduced the Gaussian Wigner function in Eqs. (8) and (9). For simplicity we will consider here, as before, only product state initial density matrices such that the Wigner function also factorizes. For this reason we can drop the cluster index in the notation Wigner function in this section. For completeness we repeat its definition here:

$$W(\{x\}) = \frac{1}{Z} \exp \left[(x_\alpha - \rho_\alpha) \Sigma_{\alpha\beta} (x_\beta - \rho_\beta) \right], \quad (50)$$

$$\text{Tr}[\hat{\rho} \hat{X}_\beta] = \int \prod_\alpha dx_\alpha \, x_\beta \, W(\{x\}),$$

$$\text{Tr}[\hat{\rho}(\hat{X}_\beta \hat{X}_\gamma + \hat{X}_\gamma \hat{X}_\beta)] = 2 \int \prod_\alpha dx_\alpha \, x_\beta x_\gamma W(\{x\}).$$

While our Gaussian choice for the Wigner function looks a little ad-hoc, there are several reasons justifying such choice (beyond the fact that it is easy to sample):

- Standard TWA, which is formally controlled by $\hbar$ or $1/S$ in the spin models with large S corresponding to the classical limit is only accurate to the order $\hbar^2$ or $1/S^2$ (see Ref. 4). For this reason one can approximate the Wigner function to the same order without the loss of accuracy. For pure spin states, e.g. states polarized along the z-axis, higher order cumulants are suppressed[25] by powers of $1/S$ and so it is only necessary to correctly describe first two cumulants: the mean and the variance. Any mixed state can be represented as a statistical mixture of pure states, so the associated Wigner function can be approximated as a sum of Gaussian distributions with non-negative weights. Interestingly in Ref. 5 it was observed numerically that in the $SU(3)$ case the Gaussian Wigner function results in a slightly better approximation to dynamics than the exact Wigner function. On passing we note the Gaussian approximations to the density matrices representing quantum statistical ensembles close to the classical limit also correctly capture leading quantum corrections to the corresponding classical probability distributions[26].

- Fixing second moments of the Wigner Function ensures that CTWA is asymptotically exact at short times up to the order $\mathcal{O}(t^2)$, which is the same accuracy as normal TWA. Similarly, the mean-field approximation, which sets all fluctuations to zero, is only accurate to the linear order in time. This can be seen by Taylor expanding the evolution of some operator $\hat{O}$. In linear order the response will be determined by the initial expectation value of the commutator $\langle \psi_0 | [\hat{O}, \hat{H}] | \psi_0 \rangle$, which is by construction is linear in cluster variables for any product state $|\psi_0\rangle$. In the next t^2 order, the response will involve expectations values of the higher order commutators like $\langle \psi_0 | [[\hat{O}, \hat{H}], \hat{H}] | \psi_0 \rangle$. It is straightforward to check that such double commutators will involve only linear or quadratic terms in the cluster operators, which are again guaranteed to be exactly reproduced by the Gaussian Wigner function. We point in this respect that absence of fluctuations in initial conditions in dynamic mean-field approximations generally leads to the mistake in the expectation value of the double commutator and hence leads only to accuracy up to the linear order in t. Therefore initial quantum fluctuations encoded in the width of the Wigner function guarantee the correct short time dynamics. Beyond
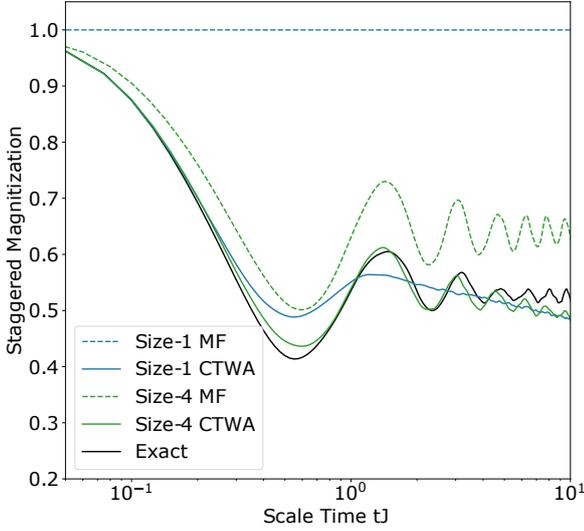
FIG. 9. **Mean Field dynamics for the Disordered Heisenberg model** (41), Neel initial condition, and staggered magnetization observable (42), averaged over disorder $h_z = 10$. Clearly for increased cluster size dynamics approach exact results, and fluctuations are required to reproduce dynamics at short times.

the order t^2 there are generally quantum jump contributions to the dynamics, which go beyond TWA[4], and so going beyond the Gaussian approximation of the Wigner function will generally not improve the accuracy of CTWA, at least at short times.

The importance of fluctuations is shown in figure 9, where dynamics of the disordered Heisenberg model are compared to that of mean-field CTWA. Clearly, without fluctuations, the second derivative of the observable (eg, $\mathcal{O}(t^2)$) is not correctly approximated.

- The exact Wigner function, which can be e.g. obtained from a Schwinger boson representation of the basis operators $\hat{X}_\alpha$ has a huge (exponential in the cluster size) amount of noise in the system. It follows from the fact the each unoccupied Schwinger boson has a $1/2$ quantum noise, and the number of Schwinger bosons is equal to the Hilbert space dimension of the cluster. This huge noise is both unphysical and would make long time dynamics uncontrollable leading to the same issues which traditional TWA faces due to diverging ultraviolet noise coming from zero point fluctuations of vacuum modes[3]. The Gaussian Wigner function we are using here does not have this problem and the quantum noise does not diverge with the increasing cluster size. Empirically, we find that this not only allows us to do controlled accurate simulations of quantum dynamics, but in many cases CTWA with the Gaussian Wigner function accurately predicts both the short time dynamics and the long time thermalization.

B. Alternate Operator Basis for Operator CTWA

In the previous sections we use an operator basis of products of Pauli matrices (see Sec. II C 1) because it has transparent physical interpretation. However for sampling purposes it is more convenient to use the basis of rank 1 matrices $\hat{Y}_{nm}$, $n = 0, 1 \dots D - 1$, $m = 1, 2 \dots n$, defined as:

$$m = n \quad \hat{Y}_{nn} = |n\rangle\langle n|,$$
$$m < n \quad \hat{Y}_{nm}^{(+)} = \frac{1}{\sqrt{2}}(|n\rangle\langle m| + |m\rangle\langle n|),$$
$$\hat{Y}_{nm}^{(-)} = \frac{i}{\sqrt{2}}(|n\rangle\langle m| - |m\rangle\langle n|),$$

where $|n\rangle$ is the n-th state in the standard product basis of spins polarized in the z-direction. As in the main text we use the convention that $|0\rangle = |\uparrow\uparrow \dots \uparrow\rangle$, $|1\rangle = |\downarrow\uparrow \dots \uparrow\rangle$, ..., These operators are explicitly Hermitian clearly forming a complete operator basis. The structure constants for these operators are easy to find from the commutation relations, e.g.

$$[\hat{Y}_{nm}^{(+)}, \hat{Y}_{pq}^{(-)}] =$$
$$\frac{i}{\sqrt{2}} \left(\delta_{mp}\hat{Y}_{nq}^{(+)} + \delta_{np}\hat{Y}_{mq}^{(+)} - \delta_{nq}\hat{Y}_{mp}^{(+)} - \delta_{mq}\hat{Y}_{np}^{(+)} \right). \quad (51)$$

As is shown in section V C, any product state wave function initial condition can be mapped from an initially Z-polarized state $|\psi\rangle = |0\rangle$. Thus, defining the Wigner function for this state is sufficient to define any Wigner function. For this state, the correct sampling of the phase space coordinates y_{nm} is given by:

$$y_{00} = 1, \ y_{n0}^{(+)} = \frac{\delta_n}{\sqrt{2}}, \ y_{n0}^{(-)} = \frac{\sigma_n}{\sqrt{2}}, \ y_{nm} = 0, \quad m \neq 0. \quad (52)$$

Here δ_n and σ_n are independent Gaussian random variables with zero mean and unit variance. The proof of Eq. (52) is rather straightforward. First of all let us observe that in the state $|\psi\rangle = |0\rangle$ only the operator $\hat{Y}_{00}$ has a non-vanishing expectation value

$$\langle 0|\hat{Y}_{00}|0\rangle = \langle 0|0\rangle\langle 0|0\rangle = 1 = \overline{y_{00}}.$$

This expectation value is obviously reproduced by the chosen distribution function

$$\langle 0|\hat{Y}_{00}^2|0\rangle = 1 \qquad \overline{y_{00}^2} = 1$$
$$\langle 0|\{\hat{Y}_{0p}^{(+)}, Y_{0q}^{(+)}\}_+|0\rangle = \delta_{pq} \qquad 2\overline{y_{0p}^{(+)} y_{0q}^{(+)}} = \delta_{pq} \quad (53)$$
$$\langle 0|\{\hat{Y}_{0p}^{(-)}, Y_{0q}^{(-)}\}_+|0\rangle = \delta_{pq} \qquad 2\overline{y_{0p}^{(-)} y_{0q}^{(-)}} = \delta_{pq}.$$

It is easy to verify that the expectation value of all other anticommutators vanish; likewise all other correlations of the associated phase space variables vanish too.

C. General initial conditions for Operator CTWA

As mentioned above, any product state within a cluster can be mapped from the Z-polarized state via some rotation.

This may generally be done via a unitary coordinate rotation $x_\alpha \to U_{\alpha\beta} x_\beta$. Sometimes it is easier to to do instead a local (intra-cluster) Hamiltonian evolution, where CTWA is exact. For example, for the state $|\downarrow\uparrow\uparrow\rangle$ we can sample a point from the Wigner function corresponding to the state $|0\rangle = |\uparrow\uparrow\uparrow\rangle$, then evolve the Hamiltonian $\hat{H} = \sigma_x^{(0)}$ for time $t = \pi$. It is easy to prove that this way of sampling the initial conditions is equivalent to sampling from the Gaussian Wigner function corresponding to the state $|\downarrow\uparrow\uparrow\rangle$. In this way, the initial states mentioned in the paper can be readily generated. E.g. for the Neel state one can first sample the state $|0\rangle$ and then apply the Hamiltonian $\hat{H} = \sum_{i \, \text{even}} \hat{\sigma}_x^{(i)}$ for time $t = \pi$. The Wigner function for a random state of the Ising example was produced in a similar fashion by applying the Hamiltonian $\hat{H} = \sum_i h_x^i \hat{\sigma}_x^{(i)} + h_y^i \hat{\sigma}_y^{(i)} + h_z^i \hat{\sigma}_z^{(i)}$ for $t = \pi$ with the randomly chosen magnetic fields $\vec{h}$ and so on.

D. Dimensional Reduction for Operator CTWA

Intriguingly, for a Gaussian Wigner and pure initial state, the complexity of the system can be reduced from D^2 to $2D$. To see this let us define a Hermitian matrix y consisting of elements y_{nn} on diagonal , $y_{mn} = 1/\sqrt{2}(y_{mn}^{(+)} - i y_{mn}^{(-)})$ on the upper triangular and $y_{nm} = 1/\sqrt{2}(y_{mn}^{(+)} + i y_{mn}^{(-)})$ on the lower triangular. From Eq. (53) it follows that the only nonzero matrix elements of y are $y_{00} = 1$, y_{0n} and y_{n0}:

$$ y = \begin{pmatrix} 1 & \frac{\delta_1 - i\sigma_1}{2} & \frac{\delta_2 - i\sigma_2}{2} & \cdots & \frac{\delta_D - i\sigma_D}{2} \\ \frac{\delta_1 + i\sigma_1}{2} & 0 & 0 & \cdots & 0 \\ \frac{\delta_2 + i\sigma_2}{2} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\delta_D + i\sigma_D}{2} & 0 & 0 & \cdots & 0 \end{pmatrix}. \quad (54) $$

This matrix has only 2 non-zero eigenvalues $\lambda_\pm$, allowing us to express it as

$$ y = \lambda_+ |\psi_+\rangle\langle\psi_+| + \lambda_- |\psi_-\rangle\langle\psi_-|, \quad (55) $$

where $|\psi_\pm\rangle$ are the eigenvectors of y corresponding to the two non-zero eigenvalues $\lambda_\pm$, which can be obtained from the characteristic polynomial:

$$ f(\lambda) = (-\lambda)^{D-2}\left((1-\lambda)(-\lambda) - \frac{1}{4}\sum_{i=1}^{D-1}(\delta_i^2 + \sigma_i^2)\right). \quad (56) $$

Then the only two non-zero eigenvalues are

$$ \lambda_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 + \sum_{i=1}^{D-1}(\delta_i^2 + \sigma_i^2)}. \quad (57) $$

The eigenvalues can be interpreted as quasi-probabilities, since $\lambda_+ + \lambda_- = 1$. Note that, one the eigenvalues is always positive and the other is always negative. Alternatively one can think of them as the components of an auxiliary spin 1/2 degree of freedom per cluster. Note that the eigenvalue only

depends on the sum of all the δ_i and σ_i. Given that they are Gaussian i.i.d. variables, the eigenvalues converge to

$$ \lambda_\pm \approx \frac{1}{2} \pm \frac{1}{2}\sqrt{2D-1} \approx \frac{1}{2} \pm \sqrt{\frac{D}{2}}, \quad (58) $$

in the large D limit. Fluctuations in the eigenvalues from realization to realization are subleading and of order $O(D^{1/4})$. The eigenvectors on the other hand become,

$$ \langle 0|\psi_\pm\rangle = \sqrt{\frac{\lambda_\pm^2}{\lambda_\pm^2 + 1/2}}, $$
$$ \langle i|\psi_\pm\rangle = \sqrt{\frac{\lambda_\pm^2}{\lambda_\pm^2 + 1/2}} \frac{\delta_i + i\sigma_i}{2\lambda_\pm} \quad \forall i > 0. \quad (59) $$

For sufficiently large clusters this behaves as

$$ \langle 0|\psi_\pm\rangle \approx 1, $$
$$ \langle i|\psi_\pm\rangle \approx \pm \frac{\delta_i + i\sigma_i}{\sqrt{2D}} \quad \forall i > 0. \quad (60) $$

Both eigenvectors of y are therefore the initial state $|0\rangle$ supplemented by exponentially small noise in cluster size. This is the primary reason why CTWA doesn't suffer from any problematic noise accumulation. Indeed, from (54) it might have appeared that CTWA is adding an exponential amount of noise to the initial conditions. However, the above analysis shows that it actually results in exponentially large eigenvalues with exponential suppression of the fluctuations in each of the eigenvectors. It can be shown using the equations of motion (10) that these eigenvalues are conserved in time, and thus the noise remains confined at all times.

We can use conservation of $\lambda_\pm$ to further reduce the number of equations of motion for the phase space variables from naive $\mathcal{O}(ND^2)$, where N is the number of clusters to $\mathcal{O}(ND)$. For example, if we have two clusters denoted by L and R and a Hamiltonian:

$$ \hat{H} = \hat{H}_L + \hat{H}_R + J\hat{V}_L\hat{V}_R $$

then equations of motion for the variables y_{nm} are equivalent to equations of motion for the states $|\psi_\pm\rangle$ with each cluster. Thus:

$$ i\partial_t|\psi_\pm^L(t)\rangle = \hat{H}_L|\psi_\pm^L(t)\rangle + J\hat{V}_L|\psi_\pm^L(t)\rangle \times $$
$$ \times \left(\lambda_+^{(R)}\langle\psi_+^R(t)|\hat{V}_R|\psi_+^R(t)\rangle + \lambda_-^{(R)}\langle\psi_-^R(t)|\hat{V}_R|\psi_-^R(t)\rangle\right), \quad (61) $$

and similarly for the right two eigenvectors. At the end of the evolution the phase space variables $y_{nm}(t)$ for each for the clusters are obtained from the state vectors $|\psi_\pm^{(L,R)}\rangle$ according to Eq. (55). Let us comment that clearly this way of simulating dynamics of cluster variables bears many parallels with Schwinger boson CTWA simulations, which we will discuss in some detail below.

E. Gaussian Wigner function for Schwinger bosons

In the previous section we discussed that the 4^N complexity of the cluster variables can be reduced to 2^N complexity

by the appropriate sampling of the matrices by vectors and evolving these vectors like wave functions (c.f. Eq. (55) and the following discussion). It is interesting that the two eigenvalues $\lambda_\pm$ appearing in this sampling (c.f. Eqs. (55) and (61)) can be interpreted as discrete quasi-probabilities as they do not evolve in time. As we discussed $\lambda_- < 0$ so we are effectively dealing with negative quasi-probabilities (alternatively the two numbers λ_+ and λ_- represent couplings to z-components of one auxiliary spin per cluster, which does not evolve in time). This of course does not create a severe sign problem because we only have a single negative number per cluster containing exponentially many variables.

Instead of this sampling of the operator variables, one can directly work with Schwinger bosons (see Sec. IIB). There we are dealing with $\mathscr{O}(2^N)$ independent complex numbers from the beginning. It is interesting that reproducing all fluctuations of the basis operators in the language of Schwinger bosons also requires a single negative quasi-probability per cluster. This is mathematically related to the fact that we are dealing with the Fock state containing one boson.

In the Schwinger boson representation the basis operators are expressed through the bilinears of creation and annihilation operators $\hat{b}_a^\dagger$, $\hat{b}_a$ and the corresponding phase space variables are represented similarly through the complex bosonic fields b_a^*, b_a satisfying canonical Poisson bracket relations $\{b_a^*, b_b\} = -i\delta_{ab}$ (c.f. Refs. 4 and 5):

$$\hat{X}_\alpha = \hat{b}_a^\dagger T_\alpha^{ab} \hat{b}_b, \quad x_\alpha = b_a^* T_\alpha^{ab} b_b. \tag{62}$$

Here T_α are the D-dimensional matrices realizing the $SU(D)$ algebra with the structure constants $f_{\alpha\beta\gamma}$:

$$[T_\alpha, T_\beta] = if_{\alpha\beta\gamma}T_\gamma.$$

In the fundamental (spin one/half) representation of the group we work with the total number of Schwinger bosons should be one:

$$\hat{n} = \sum_a \hat{b}_a^\dagger \hat{b}_a = \hat{\mathbf{I}}.$$

Because $\hat{n}$ is conserved in time for any Hamiltonian expressed through the spin operators $\hat{X}_\alpha$ one only needs to impose this constraint at initial conditions. Physically the Schwinger boson simply labels the expansion of the wave function of the system in some basis, e.g. given by Eq. (14) and the matrices T_α are simply given by the matrix elements of the corresponding basis operators $\hat{X}_\alpha$ in this basis: $T_\alpha^{ab} = \langle a|\hat{X}_\alpha|b\rangle$.

As in the case with clusters in order to reproduce correctly initial conditions for any pure state it is sufficient to reproduce them for the state $|0\rangle = |\uparrow\uparrow\ldots\uparrow\rangle$, which is in the language of Schwinger bosons is represented as $|100\ldots0\rangle$. In principle one can use the exact Wigner function for this state, which is given by the product of the Wigner function of the Fock state with one boson for the boson b_0 and of the Wigner functions corresponding to the vacuum state for all other bosons:

$$W_{sb}(\vec{b}, \vec{b}^*) = L_1(4|b_0|^2) \prod_{a=0}^{D-1} e^{-2|b_a|^2} \tag{63}$$

This choice of the Wigner function is very inconvenient for large clusters as it leads to exponentially hard sampling and very large errors in the long time limit. The reason behind is that each vacuum mode contributes noise of the order of unity to any basis operator and hence the total number of noise in the system scales exponentially with the cluster size N. Moreover, we checked numerically that even if we can circumvent the sampling problem for moderate size clusters, this Wigner function leads to large, essentially uncontrolled, mistakes at long times. The reason is that under nonlinear TWA equations this noise describing virtual vacuum excitation can deplete leading to negative occupation numbers of empty modes. In fact, this problem is very well known in standard bosonic TWA through spurious dependence of long time TWA dynamics on ultraviolet cutoff[3].

We find that it is much more convenient and accurate to use a phenomenological Wigner function, which fixes first and second moments of the basis operators in the initial state:

$$\langle \hat{X}_\alpha \rangle = \sum_{ab} T_\alpha^{ab} \int d\vec{b}d\vec{b}^* W(\vec{b}, \vec{b}^*) b_a^* b_b, \tag{64}$$

$$\langle \{\hat{X}_\alpha, \hat{X}_\beta\}_+ \rangle = 2 \sum_{ab,cd} T_\alpha^{ab} T_\beta^{cd} \int d\vec{b}d\vec{b}^* W(\vec{b}, \vec{b}^*) b_a^* b_b b_c^* b_d.$$

It is easy to check that for the state $|100\ldots0\rangle$ the Wigner function like in the exact case factorizes into the product of $w_0(b_0, b_0^*)$ corresponding to the occupied boson and the product of identical Wigner functions describing empty modes:

$$W(\vec{b}, \vec{b}^*) = w_0(b_0, b_0^*) \prod_{a=1}^{D-1} w_1(b_a^*, b_a). \tag{65}$$

In order to fix the correct expectation values and fluctuations of the basis operators in the $|\uparrow\ldots\uparrow\rangle$ state it is sufficient to consider only one diagonal and one off-diagonal operator, e.g. $\hat{\sigma}_z^{(1)}$ and $\hat{\sigma}_x^{(1)}$. It is easy to see that all other diagonal and off-diagonal operators are obtained from these two by permutations of Schwinger bosons $b_1, \ldots b_{D-1}$. Because those have identical distributions such permutations do not affect any expectation values. Then the four non-equivalent expectation values we need to reproduce are

$$1 = \langle \sigma_z^{(1)} \rangle = \overline{|b_0|^2 - |b_1|^2 + |b_2|^2 - |b_3|^2 + \ldots} = \overline{|b_0|^2} - \overline{|b_1|^2},$$
$$0 = \langle \sigma_x^{(1)} \rangle = \overline{b_0^* b_1 + b_1^* b_0 + b_2^* b_3 + b_3^* b_2 + \ldots},$$
$$1 = \langle \left(\sigma_z^{(1)}\right)^2 \rangle = \overline{|b_0|^4} + (D-1)\overline{|b_1|^4} - 2\overline{|b_0|^2}\,\overline{|b_1|^2} - D(\overline{|b_1|^2})^2,$$
$$1 = \langle \left(\sigma_x^{(1)}\right)^2 \rangle = 2\overline{|b_0|^2}\,\overline{|b_1|^2} + (D-2)(\overline{|b_1|^2})^2.$$

The second equation is trivially satisfied because for the Fock state the Wigner function cannot depend on the phases of the bosons. Combining the first and the last equation and assuming that for the empty modes (b_1) we can use the Gaussian ansatz such that $\overline{|b_1|^4} = 2(\overline{|b_1|^2})^2 \equiv 2\sigma_1^4$ we find

$$D\sigma_1^4 + 2\sigma_1^2 - 1 = 0 \Rightarrow \sigma_1^2 = \frac{\sqrt{1+D}-1}{D}. \tag{66}$$

exactly as in Eq. (18). Plugging this into the remaining two equations we find

$$\overline{|b_0|^2} = 1 + \sigma_1^2,$$
$$\overline{|b_0|^4} = 2 - 2(D-2)\sigma_1^4. \tag{67}$$

It is easy to see that this condition can not be satisfied with a positive probability distribution for any $D \geq 2$. We find empirically in all examples we simulated that one can simply suppress fluctuations in $\overline{|b_0|^2}$ all the way to zero like we did in Eq. (18) without affecting accuracy of TWA both at short and at long times. This means that Eq. (64) is not fully obeyed. In particular, the expectation value of $\hat{\sigma}_z^2$ is not reproduced correctly by the simplified Wigner function. However, this expectation value does not enter the linear response and thus does not affect the accuracy of the short time dynamics. It is equally unimportant in the long time limit. One can check that in order to satisfy Eq. (67) one can choose an appropriate discrete probability distribution for $|b_0|^2$ characterized by one positive λ_+ and one negative λ_- probabilities in direct analogy with eigenvalues $\lambda_\pm$ discussed in the previous section. Interestingly the expectation value of the identity operator $\langle 0|\hat{I}|0\rangle = \sum_j \overline{|b_j|^2}$ scales as $\sqrt{D}$ in the large D limit, which implies that one has to be cautious in literary identifying Schwinger bosons with the components of the wave function. However, the identity operator (equal to the total Schwinger boson number operator) is conserved in time and does not affect expectation values of any traceless operators we are interested in.

As a final comment, remarkably the Schwinger boson Wigner function with the noise set by Eq. (66) is much more physical than the exact Wigner function. At larger D we occupy each empty mode with $1/\sqrt{D}$ noise such that the total amount of noise in the system is always of the order of one, independent of the cluster size, as in the operator CTWA method.

VI. CONCLUSIONS

In this work, we developed a general method for simulating dynamics of interacting quantum systems using cluster semiclassical (truncated Wigner) approximation, which we termed CTWA. The method is based on first including various local correlations within a cluster as additional degrees of freedom effectively increasing the phase space dimensionality and then utilizing the TWA in this higher dimensional space. As particular illustrations of CTWA we focus on quantum spin one-half chains, which we split into clusters of size N. Each cluster is spanned by $D^2 = 4^N$ basis operators forming a closed $SU(D)$ algebra. These operators are then treated as high-dimensional spin components. The number of degrees of freedom can be further reduced to $D = 2^N$ either using the Schwinger boson representation of the operators or using other decompositions of operators into products of D-dimensional vectors, which can be alternatively sampled and propagated. Formally CTWA equations are identical to the mean field equations for the density matrix if we use operators as classical variables and to the mean field equations for the wave function if we use the Schwinger Boson representation.

A crucial difference between CTWA and mean field is the presence of fluctuations in initial conditions encoded in the initial Wigner function. These fluctuations allow CTWA to correctly reproduce short time dynamics to a higher order

in time than the mean field. They also allow capturing entanglement between subsystems and accurately reproducing local and non-local connected correlation functions of various observables via the classical mutual information in phase space, which are beyond mean field approximations. To avoid the unphysical exponential in the cluster size noise in initial conditions we use the Gaussian approximation to the Wigner function. The parameters of the Gaussian are chosen such that the Gaussian Wigner function reproduces the expectation values and fluctuations of the basis operators in the initial state.

We demonstrate the method simulating dynamics in various strongly coupled spin-chains with and without disorder. As computational difficulty scales linearly in system size, we simulate both small 1D systems (to allow for comparison with exact diagonalization) as well as in larger systems. In all the cases CTWA predictions approach exact results with an increasing cluster size. In thermalizing regimes we find numerically that CTWA accurately captures the long time thermal phase and the hydrodynamic (diffusive) approach to the equilibrium. In the MBL regime CTWA reproduces the intermediate time localization plateau but eventually predicts thermalization, with thermalization time increasing with the cluster size. We also demonstrate that CTWA predictions for entanglement closely resemble exact results.

While in this work we focused on a particular cluster choice of phase space variables we note that the method can be applied to any operator basis $\{\hat{X}_\alpha\}$ which forms a closed algebra, as long as the Hamiltonian and observables of interest can be expressed through sums and products of these basis operators. The basis operators then map to phase space variables, which evolve in time according to the corresponding classical Hamiltonian equations of motion. Note that as the operator algebra is invariant under any unitary transformation, one can take some fixed operator basis unitarily evolve $\{\hat{X}_\alpha\} \rightarrow \{\hat{U}^\dagger \hat{X}_\alpha \hat{U}\}$ to form a new (nonlocal) operator basis and then apply CTWA. We expect that in this way one can further improve accuracy of the method. One can also generalize CTWA to the situations where we deal with the approximate operator basis, which does not form a closed algebra, e.g. a Matrix Product State basis like the one used in Refs. 10–12 in the context of time-dependent variational method. Technically this can be done by using the second of Eqs. (4) defining the structure constants. In this way one effectively projects the commutator between any two basis operators to the subspace spanned by these operators. Finally one can combine the cluster CTWA with the fermionic TWA developed in Ref. 6 to describe dynamics of strongly interacting fermionic systems.

[1] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, Physics Reports **106**, 121 (1984).

[2] M. J. Steel, M. K. Olsen, L. I. Plimak, P. D. Drummond, S. M. Tan, M. J. Collett, D. F. Walls, and R. Graham, Phys. Rev. A **58**, 4824 (1998).

[3] P. Blakie, A. Bradley, M. Davis, R. Ballagh, and C. Gardiner, Advances in Physics **57**, 363 (2008), https://doi.org/10.1080/00018730802564254.

[4] A. Polkovnikov, Annals of Physics **325**, 1790 (2010), arXiv:0905.3384.

[5] S. M. Davidson and A. Polkovnikov, Physical Review Letters **114** (2015), 10.1103/PhysRevLett.114.045701, arXiv:1406.1192.

[6] S. M. Davidson, D. Sels, and A. Polkovnikov, Annals of Physics **384**, 128 (2017).

[7] A. P. Orioli, A. Safavi-Naini, M. L. Wall, and A. M. Rey, , 1 (2017), arXiv:1705.06203.

[8] M. Schmitt, D. Sels, S. Kehrein, and A. Polkovnikov, arXiv:1802.06796 (2018).

[9] L. G. Yaffe, Rev. Mod. Phys. **54**, 407 (1982).

[10] P. Kramer, Journal of Physics: Conference Series **99**, 012009 (2008).

[11] E. Leviatan, F. Pollmann, J. H. Bardarson, and E. Altman, arXiv , 1 (2017), arXiv:1702.08894.

[12] T. Shi, E. Demler, and J. I. Cirac, , 1 (2017), arXiv:1707.05902.

[13] F. Verstraete, V. Murg, and J. Cirac, Advances in Physics **57**, 143 (2008), https://doi.org/10.1080/14789940801912366.

[14] R. Orús, Annals of Physics **349**, 117 (2014).

[15] D. Pekker and B. K. Clark, Physical Review B **95**, 1 (2017), arXiv:1410.2224.

[16] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Physics Reports **697**, 1 (2017), geometry and non-adiabatic response in quantum and classical systems.

[17] D. Sels, F. Brosens, and W. Magnus, Physics Letters A **376**, 809 (2012).

[18] H. Kim, T. N. Ikeda, and D. A. Huse, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **90**, 1 (2014), arXiv:1408.0535.

[19] D. J. Luitz, N. Laflorencie, and F. Alet, Physical Review B - Condensed Matter and Materials Physics **91**, 1 (2015), arXiv:1411.0660.

[20] V. Oganesyan, A. Pal, and D. Huse, Physical Review B **80**, 1 (2009), arXiv:0905.4112.

[21] O. L. Acevedo, A. Safavi-Naini, J. Schachenmayer, M. L. Wall, R. Nandkishore, and A. M. Rey, Phys. Rev. A **96**, 033604 (2017).

[22] A. K. Kulshreshtha, A. Pal, T. B. Wahl, and S. H. Simon, arXiv:1707.05362 (2017).

[23] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, New Journal of Physics **12**, 025012 (2010).

[24] E. Altman and R. Vosk, Annual Review of Condensed Matter Physics **6**, 383 (2015).

[25] For standard $SU(2)$ spins this fact trivially follows that in the classical limit the polarized spin is represented by the non-fluctuating δ-function distribution and fluctuations are quantum coming from non-commutativity of spin components. Same arguments can be applied to higher-dimensional spins.

[26] H. Kim, A. Polkovnikov, and E. A. Yuzbashyan, Annals of Physics **381**, 107 (2017).