Nested Iterative Shift-invert Diagonalization for Many-body Localization in the Random-field Heisenberg Chain

Taito Kutsuzawa\textsuperscript{1} and Synge Todo\textsuperscript{1,2,3}

\textsuperscript{1}Department of Physics, The University of Tokyo, Tokyo 113-0033, Japan
\textsuperscript{2}Institute for Physics of Intelligence, The University of Tokyo, Tokyo 113-0033, Japan
\textsuperscript{3}Institute for Solid State Physics, The University of Tokyo, Kashiwa 277-8581, Japan

(Dated: April 1, 2022)

We study the many-body localization of the random-field Heisenberg chain using the nested shift-invert Lanczos method with an iterative linear solver. We use the minimum residual method (MINRES) inside each Lanczos iteration. The memory consumption of the proposed method is only proportional to the dimension of the Hilbert space. We also introduce a preconditioner that takes into account the effects of disorder and interaction in the random-field Heisenberg chain. As a probe of many-body localization transition, we propose a unitary operator called the twist operator, which has a clear interpretation in the real space. We discuss its behavior for thermal and localized eigenstates. We demonstrate the efficiency of the nested iterative shift-invert diagonalization method with the proposed preconditioner for the many-body localization problem and estimate the transition point of the random-field Heisenberg chain more precisely based on the finite-size analysis of the expectation value of the twist operator.

I. INTRODUCTION

Understanding the mechanism of thermalization is one of the central problems in condensed matter physics \cite{1}. A natural question is whether a closed quantum system thermalizes or not by its own unitary dynamics. The eigenstate thermalization hypothesis (ETH) \cite{2,3} is a sufficient condition that explains thermalization from the microscopic level. However, it is known that there are several counterexamples, such as integrable systems \cite{5–9} and quantum many-body scar \cite{10–12}, where the system does not thermalize by itself. Many-body localization \cite{13–16} has been investigated extensively as another possible mechanism that breaks ETH. A many-body localized system fails to thermalize under strong disorder but can also delocalize and recover thermalization due to interaction. The phase transition between thermal and localized phases is called many-body localization transition \cite{17,18}.

The Hamiltonian of the spin-1/2 random-field Heisenberg chain is given by

\[ \hat{H} = \sum_{j=1}^{L} (\hat{S}_j \cdot \hat{S}_{j+1} - h_j \hat{S}_j^z), \]

(1)

where \( \hat{S}_j = (\hat{S}_j^x, \hat{S}_j^y, \hat{S}_j^z) \) is a spin-1/2 operator at site \( j = 1, 2, \ldots, L \). We impose the periodic boundary conditions and identify \( \hat{S}_{L+1} \) with \( \hat{S}_1 \). The Hamiltonian \( \hat{H} \) commutes with \( \hat{S}_z := \sum_{j=1}^{L} \hat{S}_j^z \). We consider a chain of even length \( L \) and eigenstates in the \( \hat{S}_z = 0 \) sector. The magnetic fields \( h_1, h_2, \ldots, h_L \) are independent and identically distributed random variables with the uniform distribution over the interval \([ -h, h ] \), where \( h \geq 0 \) represents the disorder strength.

Let us consider highly excited states in the middle of the energy spectrum of the model \cite{1}. It has been argued that they behave qualitatively differently in the weak disorder regime \( h \ll J \) and the strong disorder regime \( h \gg J \) \cite{17,18}. Under the weak disorder, the energy eigenstates are extended and have volume-law entanglement (thermal phase). On the other hand, under the strong disorder, the energy eigenstates localize and obey the area law of entanglement (many-body localized phase), where local integrals of motion emerge \cite{19} but the interaction causes dephasing dynamics that yields the logarithmic growth of entanglement \cite{20,21}.

The many-body localization has been investigated by using various numerical methods. As long as the disorder is strong enough \( h \gg J \), the tensor network method \cite{22} can simulate many-body localized states efficiently as the entanglement entropy is small. However, as this method becomes rapidly ineffective near the transition between volume-law and area-law entangled eigenstates, the exact diagonalization has been used mainly so far \cite{17,18,23,24}. Unfortunately, the system size that the exact diagonalization can handle is limited, and the numerical results suffer from a severe finite-size effect. Thus, the nature of the many-body localization transition in the thermodynamic limit is still under active study \cite{25,26}. To assess the existence of the many-body localized phase in the infinite system, we must study larger systems with the help of more sophisticated numerical methods.

In this paper, we introduce the shift-invert Lanczos method \cite{30,31} with an iterative linear solver, as with the Anderson localization problem \cite{32}. The most computationally demanding part of the algorithm is to solve a linear equation. The coefficient matrix is the energy-shifted Hamiltonian, which is Hermitian indefinite. For this purpose, we use the minimum residual method (MINRES) \cite{33}, which is one of the Krylov subspace methods \cite{34}. The memory usage becomes lower than direct methods \cite{35,36} by a factor of the exponential of system size.

The computation time becomes longer than the direct method in exchange for lower memory usage. For Krylov
subspace methods to converge successfully, preconditioning is essential. We propose a simple preconditioner with both disorder and interaction taken into account, which we find effective for the present random-field Heisenberg chain.

By using the proposed nested iterative method, we successfully reproduce the results of the bipartite entanglement entropy and the level spacing ratio [17, 18, 45] across the many-body localization transition of the random-field Heisenberg chain [19]. In addition, we propose to use the twist operator [10, 17] as a probe of many-body localization transition. The twist operator has a clear meaning in the real space that it rotates spins over the chain with gradually increasing angles. We discuss the finite-size effect of the expectation value of the twist operator for thermal and localized eigenstates. Furthermore, we estimate the finite-size effect of the expectation value of the twist operator. In the present algorithm and give a finite-size analysis of its expectation value in thermal many-body localization transition point of the random-field Heisenberg chain with gradually increasing angles. We discuss the finite-size effect of the expectation value of the twist operator for thermal and localized eigenstates. We calculate the expectation value of the twist operator for each energy eigenstate to observe how thermal and localized eigenstates respond to the global perturbation induced by the twist operator. Furthermore, we estimate the many-body localization transition point of the random-field Heisenberg chain by the finite-size analysis of the many-body localization transition by using the twist operator. Finally, we give a conclusion in Sec.V.

II. NESTED ITERATIVE SHIFT-INVERT DIAGONALIZATION

Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian sparse matrix. Our goal is to calculate a few eigenvectors of $A$ with eigenvalues closest to an arbitrary target value $\sigma$, which is typically in the middle of the spectrum of $A$.

We use the shift-invert Lanczos method [39, 40]. In other words, we apply the Lanczos method [48] on the matrix $(A - \sigma I)^{-1}$, where $I \in \mathbb{C}^{n \times n}$ is the identity matrix. The most time-consuming part of the algorithm is the matrix inversion:

$$v \mapsto (A - \sigma I)^{-1}v \quad (2)$$

for a given vector $v \in \mathbb{C}^n$. To execute the matrix inversion, a previous study [18] used direct solvers, which compute the LU decomposition of the coefficient matrix $A - \sigma I$.

Instead, we make use of the minimum residual method (MINRES) [42], which is a Krylov subspace method that solves a linear equation $Ax = b$ with a Hermitian coefficient matrix $A \in \mathbb{C}^{n \times n}$ by minimizing the norm of the residual vector $b - Ax$ over the Krylov subspace (see appendix [4]). Algorithm [1] shows the pseudocode of the nested iterative shift-invert diagonalization, where $v \in \mathbb{C}^n$ is an initial vector and $m > 0$ the maximum number of Lanczos iterations. The superscript $\dagger$ means Hermitian conjugate. The iterative algorithm MINRES is nested inside each iteration of the Lanczos method.

We substitute the iterative solver for direct solvers because the former has a great advantage in memory usage against the latter. In direct algorithms, as the coefficient matrix is factorized into lower and upper triangular matrices, some zero entries in the original matrix become nonzero. These new nonzero elements are called fill-in. Let us define the fill-in ratio as the number of nonzero entries of the factorized matrices divided by that of the original matrix. Direct solvers attempt to keep the fill-in ratio as small as possible. However, it is reported [44] that the fill-in ratio grows exponentially with the system size for the random-field Heisenberg model’s Hamiltonian [1] represented in the $\hat{S}_z$ basis, where all $\hat{S}_z$ are diagonal. In contrast to direct solvers, iterative algorithms do not modify the coefficient matrix $A$ but use it only as matrix-vector products $v \mapsto Av$. We do not even need to retain the Hamiltonian on memory if the matrix elements can be generated on the fly during the matrix-

| Algorithm 1 Nested iterative shift-invert diagonalization |
|----------------------------------------------------------|
| 1: function SIMINRES($A$, $\sigma$, $v$, $m$) |
| 2: $v_0 \leftarrow 0$ |
| 3: $\beta_1 \leftarrow ||v||$ |
| 4: $v_1 \leftarrow v/\beta_1$ |
| 5: $k \leftarrow 1$ |
| 6: loop |
| 7: /* Solve $(A - \sigma I)x = v_k$ by MINRES */ |
| 8: $x \leftarrow (A - \sigma I)^{-1}v_k$ |
| 9: |
| 10: /* Orthogonalization */ |
| 11: $w \leftarrow x - \beta_k v_{k-1}$ |
| 12: $\alpha_k \leftarrow v_k^T w$ |
| 13: $w \leftarrow w - \alpha_k v_k$ |
| 14: |
| 15: /* Normalization */ |
| 16: $\beta_{k+1} \leftarrow \sqrt{w^T w}$ |
| 17: if $\beta_{k+1} = 0$ or $k = m$ then |
| 18: break |
| 19: end if |
| 20: $v_{k+1} \leftarrow w/\beta_{k+1}$ |
| 21: |
| 22: $k \leftarrow k + 1$ |
| 23: end loop |
| 24: Let $T_k$ be the $k \times k$ real symmetric tridiagonal matrix with diagonals $\alpha_1, \ldots, \alpha_k$ and subdiagonals $\beta_2, \ldots, \beta_k$. |
| 25: Diagonalize $T_k$ ($T_k y_i^{(k)} = \theta_i^{(k)} y_i^{(k)}$ ($i = 1, 2, \ldots, k$)). |
| 26: return $\sigma + \frac{1}{\theta_i^{(k)}} V_k y_i^{(k)}$ ($i = 1, 2, \ldots, k$) |
| 27: end function |
vector product. Moreover, Krylov subspace methods for Hermitian coefficient matrices, including MINRES, usually retain only a fixed number of vectors as a result of the Lanczos tridiagonalization. Thus, the memory usage can be kept down to $O(n)$.

On the other hand, there is a possibility that iterative algorithms do not give solutions in a reasonable time or even do not converge at all. The convergence of the Krylov subspace methods is governed by the condition numbers of coefficient matrices. If the coefficient matrix is ill-conditioned, that is, its condition number is large, it takes many iterations for iterative algorithms to converge. In fact, our problem becomes ill-conditioned exponentially with the system size because the target energy $\sigma$ is in the middle of the energy spectrum where the density of states grows exponentially. However, the convergence of iterative algorithms is improved significantly if we can find an effective preconditioner. In order to make the algorithm practical, it is essential to develop a preconditioner, tailored to our coefficient matrix, which makes the algorithm practical, it is essential to develop a preconditioner, tailored to our coefficient matrix, which transforms the matrix into one as close to the identity matrix as possible.

Let us consider solving a preconditioned system

$$(C^{-1}(A - \sigma I)(C^{-1})^*)C^t x = C^{-1}v$$

by MINRES. The preconditioning matrix is defined to be $M := CC^t$, which must be easily invertible and positive definite. We consider diagonal matrices with positive diagonal elements to satisfy these requirements. Let $A$ be the Hamiltonian represented in the $\hat{S}^z$ basis. If the matrix $A - \sigma I = (a_{ij} - \sigma \delta_{ij})_{1 \leq i, j \leq n}$ is diagonally dominant, then scaling the matrix by its diagonal elements should be effective. This idea gives rise to the Jacobi preconditioner $M_{\text{Jacobi}}$, namely

$$M_{\text{Jacobi}} := \text{diag}(\{a_{11} - \sigma, a_{22} - \sigma, \ldots, a_{nn} - \sigma\}),$$

where we assume $a_{ii} \neq \sigma$ ($i = 1, 2, \ldots, n$) and take the absolute values to ensure the positive definiteness.

The Hamiltonian does become diagonally dominant when the disorder is strong. However, the off-diagonal elements play an essential role in the delocalization transition to the thermal phase as the disorder is weakened. For this reason, we attempt to incorporate off-diagonal elements into the Jacobi preconditioner by considering the following preconditioner:

$$M_{\text{norm}} := \text{diag}(r_1, r_2, \ldots, r_n),$$

where $r_i$ ($i = 1, 2, \ldots, n$) is the 2-norm of the $i$-th row of the matrix $A - \sigma I$:

$$r_i := \sqrt{|a_{ii} - \sigma|^2 + \sum_{k \neq i} |a_{ik}|^2}.$$  

We demonstrate the effectiveness of the proposed preconditioner in Sec. IV.

### III. TWIST OPERATOR

The twist operator

$$\hat{U}_{\text{twist}} := \exp \left[ \frac{2\pi}{L} \sum_{j=1}^{L} j \hat{S}_j^z \right]$$

is a unitary operator that generates a spin-wave-like excitation by rotating spins around the $z$ axis with angles $\theta_j := \frac{2\pi}{L} j$ gradually increasing over sites $j = 1, 2, \ldots, L$. The twist operator appeared in the proof of the Lieb-Schultz-Mattis theorem to create a trial state orthogonal to the ground state and has excitation energy of $O(L^{-1})$. Nakamura and Todo utilized the twist operator as an order parameter to detect quantum phase transitions.

We propose the following quantity for detecting the many-body localization transition between thermal and localized eigenstates:

$$z := \langle \psi | \hat{U}_{\text{twist}} | \psi \rangle,$$  

where $| \psi \rangle$ is an eigenstate of the Hamiltonian. This quantity, twist overlap $z$, measures how much the twisted state $\hat{U}_{\text{twist}} | \psi \rangle$ overlaps with the original eigenstate $| \psi \rangle$. For thermal eigenstates, we expect that the twist operator generates a spin-wave-like excitation and creates a new state which is orthogonal to the original one, as in the case without randomness. On the other hand, we expect that localized eigenstates are not affected by the long-wave-length perturbation by the twist operator but are responsive to modifications of the local degrees of freedom. In the rest of this section, we discuss the excitation energy induced by the twist operator and the finite-size behavior of the twist overlap for thermal and localized eigenstates.

First, let us confirm that the energy difference between the twisted state $\hat{U}_{\text{twist}} | \psi \rangle$ and the original eigenstate $| \psi \rangle$ vanishes in the thermodynamic limit for the Hamiltonian. We can calculate the energy difference as in the case without randomness since the random magnetic-field terms in the Hamiltonian are invariant under the transformation by the twist operator. By using the following identities,

$$\hat{U}_{\text{twist}}^\dagger \hat{S}_j^z \hat{U}_{\text{twist}} = \hat{S}_j^z,$$

$$\hat{U}_{\text{twist}}^\dagger \hat{S}_j^+ \hat{U}_{\text{twist}} = e^{-i\theta_j} \hat{S}_j^+,$$

$$\hat{U}_{\text{twist}}^\dagger \hat{S}_j^- \hat{U}_{\text{twist}} = e^{i\theta_j} \hat{S}_j^-,$$

where $\theta_j := \frac{2\pi}{L} j$, we can write

$$\hat{U}_{\text{twist}}^\dagger \hat{H} \hat{U}_{\text{twist}} - \hat{H} = \frac{J}{2} \left( \cos \left( \frac{2\pi}{L} j \right) - 1 \right) \sum_{j=1}^{L} (\hat{S}_j^+ \hat{S}_{j+1}^- + \hat{S}_j^- \hat{S}_{j+1}^+) + i \frac{J}{2} \sin \left( \frac{2\pi}{L} j \right) \sum_{j=1}^{L} (\hat{S}_j^+ \hat{S}_{j+1}^- - \hat{S}_j^- \hat{S}_{j+1}^+).$$
Let $|\psi\rangle$ be an eigenstate of the Hamiltonian (1). The Hamiltonian is real symmetric in the $\hat{S}^z$ basis. Hence, the expansion coefficients of the eigenstate $|\psi\rangle$ can be chosen to be real and so are those of $\hat{S}^+_j \hat{S}^-_{j+1} |\psi\rangle$ and $\hat{S}^-_j \hat{S}^+_j |\psi\rangle$. As a result, we obtain the energy difference as follows:

$$\langle \psi| (\hat{U}_{\text{twist}}^\dagger \hat{H} \hat{U}_{\text{twist}} - \hat{H}) |\psi\rangle = \frac{J}{2} \left( \cos \left( \frac{2\pi}{L} \right) - 1 \right) \times \sum_{j=1}^L \langle \psi| (\hat{S}^+_j \hat{S}^-_{j+1} + \hat{S}^-_j \hat{S}^+_j) |\psi\rangle. \quad (13)$$

An upper bound can be given as follows assuming the eigenstate $|\psi\rangle$ is normalized:

$$| \langle \psi| (\hat{U}_{\text{twist}}^\dagger \hat{H} \hat{U}_{\text{twist}} - \hat{H}) |\psi\rangle | \leq \frac{|J|}{4} \left( \frac{2\pi}{L} \right)^2 L = \frac{|J|\pi^2}{L}, \quad (14)$$

which vanishes in the limit $L \to \infty$.

Next, we discuss the behavior of the twist overlap $z$ in the thermal phase. Let us expand a thermal eigenstate $|\psi\rangle$ as

$$|\psi\rangle = \sum_{\alpha=1}^n c_\alpha |\alpha\rangle \quad (c_\alpha \in \mathbb{R}), \quad (15)$$

where $\{|\alpha\rangle\}$ is the set of $\hat{S}^z$ basis vectors in the $\hat{S}^z = 0$ sector of dimension $n = (L/2)^2$. As an approximation of extended eigenstates, we assume that $c_\alpha$ ($\alpha = 1, 2, \ldots, n$) are drawn independently from the normal distribution with zero mean and variance $1/n$. The variance comes from the normalization of $|\psi\rangle$. Then, the average of the squared norm of the expectation value of the twist operator becomes (see appendix [3])

$$\overline{| \langle \psi| \hat{U}_{\text{twist}} |\psi\rangle |^2} = \frac{2}{n}, \quad (16)$$

which decreases exponentially as the system size $L$ grows.

Lastly, we discuss the behavior of the twist overlap $z$ in the localized phase by perturbation theory. We assume $\hbar \gg J$ and divide the Hamiltonian (1) into an unperturbed part $\hat{H}_0$ and a perturbation part $\hat{V}$ as follows:

$$\hat{H} = \hat{H}_0 + \hat{V}, \quad (17)$$

$$\hat{H}_0 = -\sum_{j=1}^L h_j \hat{S}^z_j, \quad (18)$$

$$\hat{V} = J \sum_{j=1}^L \hat{S}_j \cdot \hat{S}_{j+1}. \quad (19)$$

The unperturbed part $\hat{H}_0$ is diagonal in the $\hat{S}^z$ basis $\{|\alpha\rangle\}$:

$$\hat{H}_0 |\alpha\rangle = E_{\alpha}^{(0)} |\alpha\rangle. \quad (20)$$

Let us consider the following eigenstate $|\psi\rangle$ of the perturbed Hamiltonian $\hat{H}$:

$$|\psi\rangle = |\alpha\rangle + \sum_{\beta \neq \alpha} c_{\beta}^{(1)} |\beta\rangle + \mathcal{O}(J^2), \quad (21)$$

which is expressed as a linear combination of the unperturbed eigenstates $\{|\alpha\rangle\}$. We fix the coefficient of $|\alpha\rangle$ to be unity. $c_{\beta}^{(1)}$ is the $\mathcal{O}(J)$ coefficient of $|\beta\rangle$ in $|\psi\rangle$:

$$c_{\beta}^{(1)} = \frac{\langle \beta| \hat{V} |\alpha\rangle}{E_{\alpha}^{(0)} - E_{\beta}^{(0)}}. \quad (22)$$

Then, we can evaluate the expectation value of the twist operator as follows (see appendix [C]):

$$\left| \langle \psi| \hat{U}_{\text{twist}} |\psi\rangle \right|^2 = \frac{\langle \psi|\hat{U}_{\text{twist}}|\psi\rangle^2}{1 + \langle \psi|\psi\rangle^2} - \langle \alpha| \hat{U}_{\text{twist}} |\alpha\rangle \right| \leq 2 \sin \left( \frac{\pi}{L} \right) \frac{\langle \psi|\psi\rangle^2}{1 + \langle \psi|\psi\rangle^2} + \mathcal{O}(J^3), \quad (23)$$

where $\langle \psi|\psi\rangle^2$ is the $\mathcal{O}(J^2)$ term in the normalization constant $\langle \psi|\psi\rangle$:

$$\langle \psi|\psi\rangle^2 = \sum_{\beta \neq \alpha} |c_{\beta}^{(1)}|^2. \quad (24)$$

The $\mathcal{O}(J^2)$ term in the right hand side of Eq. (23) vanishes as $\mathcal{O}(L^{-1})$ in the limit $L \to \infty$.

IV. RESULTS

A. Benchmark

We implement the outer Lanczos loop on a central processing unit (CPU). The most computationally intensive part within the Lanczos loop is the matrix inversion using MINRES. The computation of MINRES consists of matrix-vector products, vector inner products, vector additions, and scalar-vector multiplications, which are memory-bound and can be performed embarrassingly parallel. To accelerate the matrix inversion, we offload the execution of the MINRES algorithm to a graphical processing unit (GPU). A GPU is a parallel computer equipped with a massive number of computation cores and high memory bandwidth. We use the NVIDIA GPU (V100 and A100) and its programming model CUDA [50]. Figure 1 shows the concept of our implementation.

For the Heisenberg model's Hamiltonian (1), we measure how much memory and time our program spends to obtain 50 eigenstates with eigenvalues closest to the target energy $\sigma = (E_{\text{min}} + E_{\text{max}})/2$, where $E_{\text{min}}$ and $E_{\text{max}}$ are the smallest and largest eigenvalues, respectively. First, we measure the CPU and GPU memory for
a single run of the program without any preconditioner. We fix the disorder strength to be $h = 1$. Figure 2 shows the result. The memory usage is expected to be modeled as

$$f_{\text{CPU}}(n) \approx 8 \text{ bytes} \times m \times n + C_{\text{CPU}},$$

$$f_{\text{GPU}}(n) \approx 8 \text{ bytes} \times 9 \times n + C_{\text{GPU}},$$

for CPU and GPU, respectively. Here, $n$ is the dimension of the Hilbert space, i.e., the size of the matrix, and $m$ is the number of the iterations of the Lanczos algorithm (see Alg. 1). Factors 8 and 9 in Eqs. (25) and (26) denote the size of double-precision floating number and the number of vectors required for MINRES, respectively. We observe that the number of Lanczos iterations $m$ is almost constant regardless of the system size $L$ and the disorder strength $h$. Thus, we fix it as a constant, $m = 130$, for modeling the CPU memory usage by Eq. (25). The constant terms $C_{\text{CPU}}$ and $C_{\text{GPU}}$ are determined from the memory usage measured for $L = 12$ ($n = 924$). As shown in Fig. 2, our linear memory-usage models describe the measured values very well, and we expect that the models are valid for larger $n$.

Our implementation of Alg. 1 uses several orders of magnitude less memory than the shift-invert Lanczos method with direct linear solvers. In the latter, it was reported to be 244 GB for $L = 22$ [44], while in the former is only 1.85 GB (CPU) plus 0.89 GB (GPU). Even for system size $L = 30$, we estimate that the memory usage of Alg. 1 is 162 GB (CPU) plus 12.6 GB (GPU), which fits well in modern workstations. This significant reduction of memory usage is the essential advantage of the present algorithm.

Next, we measure the elapsed time (wall time) spent in a single program run. We measure the execution time for the cases without preconditioner and with the row-norm preconditioner [5]. Figure 3 shows the results for $h = 3$ and 4. The time complexity of the algorithm is expected to be $O(n^2 L m)$. As $m$ is almost constant regardless of $L$, the execution time of order $n^2$ for large matrix sizes $n = 184,756, 705,432, 2,704,156$ ($L = 20, 22, 24$, respectively). The factor $L$ in the expected time complexity $O(n^2 L m)$ can also be treated as a constant since $n$ increases exponentially with respect to $L$. We also confirm that the row-norm preconditioner [4] significantly reduces the execution time even for $h = 3$ and 4, which is in the thermal phase as we see below.
B. Entanglement and level spacing

By using Alg. 1, we calculate the bipartite entanglement entropy

$$S = - \text{Tr} \rho_A \log \rho_A$$

and the level spacing ratio

$$r_k = \frac{\min\{\delta_k, \delta_{k-1}\}}{\max\{\delta_k, \delta_{k-1}\}}$$

for the random-field Heisenberg chain (1). Here, $\rho_A$ is the reduced density matrix for the subsystem of sites $j = 1, 2, \ldots, L/2$. The level spacing is defined as the difference between adjacent eigenvalues, $\delta_k = E_{k+1} - E_k$. For each pair of system size $L$ and disorder strength $h$, the number of disorder realizations is 10 000 for $L \leq 16$, 5 000 for $L = 18$, 1 000 for $L = 20$, and 100 for $L = 22$. For each realization of disorder, 50 eigenstates with eigenvalues closest to the target energy $\sigma = (E_{\text{min}} + E_{\text{max}})/2$ are computed, where $E_{\text{min}}$ and $E_{\text{max}}$ are the smallest and largest eigenvalues, respectively. The standard deviation is evaluated only over disorder realizations because eigenstates of the same disorder realization can be correlated.

Figure 4 shows the system-size dependence of the average bipartite entanglement entropy $\bar{S}$. The entanglement $\bar{S}$ is confirmed to be proportional to the system volume $L$ for small $h$, which is expected behavior for thermal states. We can see that the random-matrix-theory (RMT) value

$$S_{\text{RMT}}(L) = \ln(2) \frac{L}{2} + \frac{1}{2} \ln \left( \frac{1}{2} \right) - \frac{1}{4}$$

(29)

gives a decent estimate for the bipartite entanglement entropy for small enough $h$ even at the small system sizes $L \leq 22$. On the other hand, we can see that the entanglement $\bar{S}$ scales in a sub-volume manner for large enough $h$.

Figure 5 shows the disorder-strength dependence of the average level spacing ratio $\bar{r}$. We can confirm that the level spacing ratio $\bar{r}$ takes the Gaussian orthogonal ensemble (GOE) value $r_{\text{GOE}} = 0.5307(1)$ for small $h$. We can also observe that the level spacing ratio $\bar{r}$ decreases down to the Poisson value $r_{\text{Poisson}} = 2\ln(2) - 1 \approx 0.38629$ as the disorder strength $h$ increases.
We compute the eigenstate expectation value of the twist operator by using the same set of eigenvectors as Figs. 4 and 5. First, the squared norm of the expectation value of the twist operator is calculated for each eigenstate. Second, the computed values are averaged over eigenstates of the same disorder realization. Lastly, from these intra-spectrum-averaged values, we calculate the average and standard deviation.

Figure 6 shows the mean square $|z|^2$ of the magnitudes of the twist overlap. As the system size $L$ increases, we can observe that $|z|^2$ decreases in the weak disorder $h$, but on the other hand, it increases in the strong disorder $h$. The intersections of the curves of $|z|^2$ for different system sizes $L$ drift slowly towards larger disorder strength $h$. Compared with the level spacing ratio (Fig. 5), the disorder-strength dependence of the eigenstate expectation value of the twist operator is smoother and we can see the system-size dependence more clearly.

Let us examine the system-size scaling of $|z|^2$ in the weak and strong disorder regimes. Figure 7 shows $|z|^2$ versus the dimension of the Hilbert space $n$ under the weak disorder. It can be seen that $|z|^2$ decreases with the power of $n$, i.e., exponentially with the system size $L$, for small $h$. For $h = 2$, the exponent of decay is evaluated as $-0.9$, which is close to the anticipated $n^{-1}$ scaling. Figure 8 shows the system-size dependence of $1 - |z|^2$ in the strong disorder regime. It can be seen that $1 - |z|^2$ decreases with the power of $L$ for large $h$. For $h = 10$, the exponent of the power law is evaluated as $-0.9$, which is consistent with the predicted $L^{-1}$ scaling.

Next, we estimate the many-body localization transition point of the model (1). In the same approach with a previous research [26], we interpolate $|z|^2$ with a polynomial of degree three as a function of the disorder strength $h$ for each system size $L$. Let $f_L(h)$ be the interpolation of $|z|^2$ for system size $L$. We calculate $h_c^L(L)$ as the intersection of interpolations $f_{L-1}(h)$ and $f_{L+1}(h)$. In the same way, we define $h_c^L(L)$ as the intersection of interpolations $f_{L-2}(h)$ and $f_{L+2}(h)$. As shown in Fig. 9 we fit $h_c^L(L)$ and $h_c^L(L)$ as linear functions of $L^{-1}$ and extrapolate them to the thermodynamic limit $L \to \infty$. We thus conclude that the transition point is $h_c^{\text{twist}} \approx 5$. This value is close to the value $h_c \approx 5.4$ [26] estimated by using the bipartite entanglement entropy and the level spacing ratio. On the other hand, the present estimate is significantly greater than $h_c \approx 3.7$ [18] concluded by the finite-size data collapse. According to Fig. 9 we can say that the disorder strength $h \approx 3.7$ is likely to be in the thermal region for larger systems.

C. Twist operator

In the same approach with a previous research [26], we interpolate $|z|^2$ with a polynomial of degree three as a function of the disorder strength $h$ for each system size $L$. Let $f_L(h)$ be the interpolation of $|z|^2$ for system size $L$. We calculate $h_c^L(L)$ as the intersection of interpolations $f_{L-1}(h)$ and $f_{L+1}(h)$. In the same way, we define $h_c^L(L)$ as the intersection of interpolations $f_{L-2}(h)$ and $f_{L+2}(h)$. As shown in Fig. 9 we fit $h_c^L(L)$ and $h_c^L(L)$ as linear functions of $L^{-1}$ and extrapolate them to the thermodynamic limit $L \to \infty$. We thus conclude that the transition point is $h_c^{\text{twist}} \approx 5$. This value is close to the value $h_c \approx 5.4$ [26] estimated by using the bipartite entanglement entropy and the level spacing ratio. On the other hand, the present estimate is significantly greater than $h_c \approx 3.7$ [18] concluded by the finite-size data collapse. According to Fig. 9 we can say that the disorder strength $h \approx 3.7$ is likely to be in the thermal region for larger systems.
V. CONCLUSION

In this paper, we proposed the nested iterative shift-invert Lanczos method with the minimum residual method (MINRES) for studying many-body localization. As a result of substituting the iterative linear solver MINRES for direct linear solvers, the memory usage is reduced by several orders of magnitude, beginning proportional to the dimension of a Hilbert space. In addition, we introduced a simple preconditioner for the Hamiltonian of the random-field Heisenberg chain. The preconditioner helps us to handle a large number of disorder realizations by decreasing the execution time of the algorithm. We performed a large-scale exact diagonalization of the random-field Heisenberg chain with system size up to 22 and $10^2 - 10^4$ disorder realizations. We consider that the algorithm can reach even larger systems with more problem-specific elaborate preconditioners and thus contribute to understanding the fate of the many-body localized phase in the thermodynamic limit.

Furthermore, we proposed the twist operator as a probe of many-body localization transition with a clear interpretation in the real space. We discussed that the magnitude of the expectation value of the twist operator decreases for thermal eigenstates but increases for localized eigenstates as the system size grows. We evaluated the twist overlap numerically for the random-field Heisenberg chain using the nested iterative shift-invert diagonalization method. Also, we estimated the many-body localization transition point of the random-field Heisenberg chain from the system-size dependence of the mean twist overlap and obtained $h_c^{\text{twist}} \approx 5$. It is left as future work to examine further the twist operator’s capability to detect many-body localization transitions in other strongly-correlated quantum systems.

ACKNOWLEDGMENTS

The computation was performed on the workstations at the Institute for Physics of Intelligence (iπ), the University of Tokyo, and the supercomputing system MASAMUNE-IMR at the Center for Computational Materials Science, the Tohoku University.

Appendix A: Minimum residual method

The minimum residual method (MINRES) \cite{Saunders} is an iterative algorithm to solve a system of linear equations $Ax = b$ where $A \in \mathbb{C}^{n \times n}$ is an $n \times n$ Hermitian indefinite matrix and $b \in \mathbb{C}^n$ a vector. Let $x_0 \in \mathbb{C}^n$ be the initial guess of the solution, $r_0 := b - Ax_0$ the initial residual vector, and

$$
K_k(A, r_0) := \text{span}(r_0, Ar_0, A^2r_0, \ldots, A^{k-1}r_0) \subset \mathbb{C}^n
$$

the $k$-dimensional Krylov subspace. The solution $x_k$ that MINRES yields at the $k$-th step of the algorithm is the least-square solution within the $k$-dimensional Krylov subspace translated by $x_0$:

$$
x_k := \arg \min_{x \in x_0 + K_k(A, r_0)} \|b - Ax\|. \quad (A2)
$$

This least-square problem can be solved by reducing the coefficient matrix $A$ into a tridiagonal matrix with the Lanczos three-term recurrence relation and eliminating the subdiagonals by a simple Givens rotation \cite{Kato:1995}. Algorithm 2 shows the pseudocode of MINRES that solves a preconditioned system

$$
(C^{-1}A(C^{-1})^\dagger)(C^\dagger x) = C^{-1}v. \quad (A3)
$$

The preconditioning matrix is defined as $M := CC^\dagger$. 

FIG. 8. System-size dependence of the expectation value of the twist operator in the strong disorder regime. The horizontal axis represents the system size $L$. The dotted line indicates the $L^{-1}$ scaling for comparison.

FIG. 9. Estimation of many-body localization transition point of the random-field Heisenberg chain. The horizontal axis represents the system size $L$. The system sizes corresponding to the points in the figure are $L = 13, 14, 15, 16, 17, 18, 19$. We extrapolate $h_c^I(L)$ and $h_c^J(L)$ to the infinite system size $L$ as linear functions of $L^{-1}$. 

FIG. 10. Results for the twist overlap $|z|_2$ for $h = 4.0, 5.0, 6.0, 10.0$. The solid line indicates the $L^{-1}$ scaling for comparison.
Algorithm 2 Preconditioned minimum residual method

1: function MINRES(A, x₀, b, M, k_max, tolerance)
2:   z₀ ← 0, z₁ ← b − Ax₀
3:   q₁ ← M⁻¹ z₁
4:   β₀ ← 1, β₁ ← \sqrt{q₁ᵀ z₁}
5:   φ₀ ← β₁, δ₁ ← 0
6:   c₀ ← 1, s₀ ← 0
7:   d₋₁ ← 0, d₀ ← 0
8:   k ← 0
9:   while |φ_k| ≥ tolerance and k < k_max do
10:     k ← k + 1
11:     p_k ← Aq_k
12:     α_k ← q_kᵀp_k/β_k²
13:     z_k+₁ ← \frac{1}{α_k} p_k - α_k z_k - \frac{β_k}{δ_k} z_{k-1}
14:     q_k+₁ ← M⁻¹ z_{k+1}
15:     β_k+₁ ← \sqrt{q_k+₁ᵀ z_{k+1}}
16:     δ_k ← c_{k-1} δ_k + s_{k-1} α_k
17:     γ_k ← -s_{k-1} δ_k + c_{k-1} α_k
18:     θ_k ← s_{k-1} θ_{k+1}
19:     δ_k+₁ ← c_{k} δ_k + s_{k} θ_k
20:     γ_k ← \frac{γ_k^2}{δ_k} + \frac{β_k^2}{δ_{k+1}}
21:     c_k ← γ_k/γ_k
22:     s_k ← δ_k+₁/γ_k
23:     τ_k ← c_k φ_{k-1}
24:     φ_k ← -s_k φ_k
25:     d_k ← (\frac{1}{δ_k} q_k - δ_k d_{k-1} - ε_k d_{k-2})/γ_k
26:     x_k ← x_{k-1} + τ_k d_k
27: end while
28: return x_k
29: end function

Appendix B: Derivation of Eq. \[16\]

The expectation value of the twist operator can be written as follows:

\[ \langle \psi | \hat{U}_{\text{twist}} | \psi \rangle = \sum_{\alpha=1}^{n} c_\alpha^2 u_\alpha \quad (u_\alpha := \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle), \]

(B1)

where \(|u_\alpha| = 1\). The average of the squared norm can be calculated as follows:

\[ |\langle \psi | \hat{U}_{\text{twist}} | \psi \rangle|^2 = \sum_{\alpha=1}^{n} c_\alpha^4 + \sum_{\alpha \neq \beta} c_\alpha^2 c_\beta^2, \]

(B2)

where we used \(|u_\alpha| = 1\). We can now use the following identity:

\[ n + \sum_{\alpha \neq \beta} u_\alpha^* u_\beta = 0, \]

(B4)

which can be obtained by taking the squared norm of the both sides of Eq. \[15\] in the following lemma.

Lemma 1.

\[ \sum_{\alpha=1}^{n} u_\alpha = 0. \]

(B5)

Proof. We consider the following state:

\[ |\phi\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{n} |\alpha\rangle, \]

(B6)

for which the expectation value of the twist operator is

\[ \langle \phi | \hat{U}_{\text{twist}} | \phi \rangle = \frac{1}{n} \sum_{\alpha=1}^{n} u_\alpha. \]

(B7)

Let \(\hat{T}\) be the translation operator which satisfies

\[ \hat{T} \hat{S}_j \hat{T} = \hat{S}_{j-1}, \]

(B8)

where we identify \(\hat{S}_0\) with \(\hat{S}_L\). The state \(|\psi\rangle\) is translation invariant, hence we have

\[ \langle \phi | \hat{T} \hat{U}_{\text{twist}} | \phi \rangle = \langle \phi | \hat{U}_{\text{twist}} | \phi \rangle. \]

(B9)

On the other hand, if we act the translation operator on the twist operator, we obtain

\[ \hat{T} \hat{U}_{\text{twist}} \hat{T} = \exp \left[ \frac{2\pi i}{L} \sum_{j=1}^{L} j \hat{S}_j^z \right] \]

\[ = \exp \left[ \frac{2\pi i}{L} \sum_{j=1}^{L} \{ (j - 1) + 1 \} \hat{S}_j^z \right] \]

\[ = \exp \left[ \frac{2\pi i}{L} \sum_{j=1}^{L} j \hat{S}_j^z \right] \exp \left[ \frac{2\pi i}{L} \sum_{j=1}^{L} \hat{S}_j^z \right] \exp \left[ \frac{2\pi i}{L} \sum_{j=1}^{L} \hat{S}_j^z \right] \]

\[ = -\hat{U}_{\text{twist}}, \]

(B10)

where we used \(e^{-2\pi i \hat{S}_L^z} = -1\) and \(\sum_{j=1}^{L} \hat{S}_j^z = 0\). Therefore, we also have

\[ \langle \phi | \hat{T} \hat{U}_{\text{twist}} | \phi \rangle = -\langle \phi | \hat{U}_{\text{twist}} | \phi \rangle. \]

(B11)

Eqs. \[B7\], \[B9\], and \[B11\] give Eq. \[B5\].
Appendix C: Derivation of Eq. (23)

Since the twist operator $\hat{U}_{\text{twist}}$ is diagonal with respect to the unperturbed eigenstates, we can write the expectation value of the twist operator as follows:

$$\langle \psi | \hat{U}_{\text{twist}} | \psi \rangle = \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle + \sum_{\beta \neq \alpha} \left| c^{(1)}_{\beta} \right|^2 \langle \beta | \hat{U}_{\text{twist}} | \beta \rangle + \mathcal{O}(J^3). \quad (C1)$$

Let us evaluate the leading $\mathcal{O}(J^3)$ perturbation. The off-diagonal element $\langle \beta | \hat{V} | \alpha \rangle$ is nonzero if and only if there is a site $j$ such that the $j$-th and $(j+1)$-th spins of $|\alpha\rangle$ are opposite and $|\beta\rangle$ is the same with $|\alpha\rangle$ except the $j$-th and $(j+1)$-th spins are flipped. With this flip operation, the expectation value of the twist operator gains a phase factor $e^{i \frac{\pi}{2} \Delta_{\alpha\beta}}$. For example, if the $j$-th and $(j+1)$-th spins of $|\alpha\rangle$ are up and down, respectively, and $|\beta\rangle$ is $|\alpha\rangle$ with the two spins flipped, then we have

$$\langle \beta | \hat{U}_{\text{twist}} | \beta \rangle = \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle e^{i \frac{\pi}{2} \Delta_{\alpha\beta}}. \quad (C2)$$

Let us define an integer $d_{\alpha\beta}$ ($\beta \neq \alpha$) as follows:

$$\langle \beta | \hat{U}_{\text{twist}} | \beta \rangle = \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle e^{i \frac{\pi}{2} \Delta_{\alpha\beta}}, \quad (C3)$$

then the integer $d_{\alpha\beta}$ satisfies $d_{\alpha\beta} = 1$ if the expansion coefficient $c^{(1)}_{\beta}$ is nonzero. With this integer $d_{\alpha\beta}$, Eq. (C1) is rewritten as follows:

$$\begin{align*}
\langle \psi | \hat{U}_{\text{twist}} | \psi \rangle &= \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle \left( 1 + \sum_{\beta \neq \alpha} \left| c^{(1)}_{\beta} \right|^2 e^{i \frac{\pi}{2} \Delta_{\alpha\beta}} \right) + \mathcal{O}(J^3) \\
&= \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle \\
&\times \left( 1 + \langle \psi | \psi \rangle \langle \psi \rangle + \sum_{\beta \neq \alpha} \left| c^{(1)}_{\beta} \right|^2 (e^{i \frac{\pi}{2} \Delta_{\alpha\beta}} - 1) \right) + \mathcal{O}(J^3). \quad (C4)
\end{align*}$$

We can write

$$\begin{align*}
\left| \langle \psi | \hat{U}_{\text{twist}} | \psi \rangle - \langle \alpha | \hat{U}_{\text{twist}} | \alpha \rangle \left( 1 + \langle \psi | \psi \rangle \langle \psi \rangle \right) \right| &= \left| \sum_{\beta \neq \alpha} \left| c^{(1)}_{\beta} \right|^2 (e^{i \frac{\pi}{2} \Delta_{\alpha\beta}} - 1) \right| + \mathcal{O}(J^3) \\
&\leq \sum_{\beta \neq \alpha} \left| c^{(1)}_{\beta} \right|^2 2 \sin \left( \frac{\pi}{L} d_{\alpha\beta} \right) + \mathcal{O}(J^3) \\
&= 2 \sin \left( \frac{\pi}{L} \langle \psi | \psi \rangle \right) + \mathcal{O}(J^3),
\end{align*}$$

from which we obtain Eq. (23).
[30] R. Vosk, D. A. Huse, and E. Altman, Phys. Rev. X 5, 031032 (2015).
[31] A. C. Potter, R. Vasseur, and S. A. Parameswaran, Phys. Rev. X 5, 031033 (2015).
[32] P. T. Dumitrescu, R. Vasseur, and A. C. Potter, Phys. Rev. Lett. 119, 110604 (2017).
[33] T. Thiery, F. Huveneers, M. Müller, and W. De Roeck, Phys. Rev. Lett. 121, 140601 (2018).
[34] L. Zhang, B. Zhao, T. Devakul, and D. A. Huse, Phys. Rev. B 93, 224201 (2016).
[35] A. Goremykina, R. Vasseur, and M. Serbyn, Phys. Rev. B 99, 224205 (2019).
[36] A. Morningstar and D. A. Huse, Phys. Rev. B 99, 224205 (2019).
[37] A. Morningstar, D. A. Huse, and J. Z. Imbrie, Phys. Rev. B 102, 125134 (2020).
[38] P. T. Dumitrescu, A. Goremykina, S. A. Parameswaran, M. Serbyn, and R. Vasseur, Phys. Rev. B 99, 094205 (2019).
[39] T. Ericsson and A. Ruhe, Math. Comp. 35, 1251 (1980).
[40] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, eds., Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide (SIAM, Philadelphia, 2000).
[41] O. Schenk, M. Bollhöfer, and R. A. Römer, SIAM Review 50, 91 (2008).
[42] C. C. Paige and M. A. Saunders, SIAM J. Numer. Anal. 12, 617 (1975).
[43] C. F. Van Loan and G. H. Golub, Matrix Computations, 4th ed. (Johns Hopkins University Press, Baltimore, 2013).
[44] F. Pietracaprina, N. Macé, D. J. Luitz, and F. Alet, SciPost Phys. 5, 45 (2018).
[45] V. Oganesyan and D. A. Huse, Phys. Rev. B 75, 155111 (2007).
[46] E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. 16, 407 (1961).
[47] M. Nakamura and S. Todo, Phys. Rev. Lett. 89, 077204 (2002).
[48] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, and N. Kawashima, Comput. Phys. Commun. 217, 180 (2017).
[49] I. Affleck and E. H. Lieb, Lett. Math. Phys. 12, 57 (1986).
[50] Programming Guide :: CUDA Toolkit Documentation, https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html (2021), (Accessed on 12/05/2021).
[51] L. Vidmar and M. Rigol, Phys. Rev. Lett. 119, 220603 (2017).
[52] Y. Huang, Nucl. Phys. B 938, 594 (2019).
[53] Y. Huang, Nucl. Phys. B 966, 115373 (2021).
[54] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Phys. Rev. Lett. 110, 084101 (2013).
[55] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. V. van der Vorst, Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition (SIAM, Philadelphia, PA, 1994).