Numerical Algorithm for Exact Finite Temperature Spectra and Its Application to Frustrated Quantum Spin Systems

Youhei Yamaji, 1, 2 Takafumi Suzuki, 3 and Mitsuki Kawamura 4

1 Department of Applied Physics, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
2 JST PRESTO, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
3 Graduate School of Engineering, University of Hyogo, Hyogo, 670-2280, Japan
4 The Institute for Solid State Physics, The University of Tokyo, Kashiwa-shi, Chiba, 277-8581, Japan

(Dated: February 9, 2018)

A numerical algorithm to calculate exact finite-temperature spectra of many-body lattice Hamiltonians is formulated by combining the typicality approach and the shifted Krylov subspace method. The combined algorithm, which we name finite-temperature shifted Krylov subspace method for simulating spectra (FTKω), efficiently constructs typical pure states in microcanonical shells and reproduces the canonical-ensemble probability distribution at finite temperatures with the computational cost proportional to the Fock space dimension. The present FTKω enables us to exactly calculate finite-temperature spectra of many-body systems whose system sizes are twice larger than those handled by the canonical ensemble average and allows us to access the frequency domain directly without sequential real-time evolution often used in previous studies. By employing the reweighting method with the present algorithm, we obtain significant reduction of the numerical costs for temperature sweeps. Application to a representative frustrated quantum spin system, namely the Kitaev-Heisenberg model on a honeycomb lattice, demonstrates the capability of the FTKω. The Kitaev-Heisenberg model shows quantum phase transitions from the quantum spin liquid phase exactly obtained for the Kitaev model to magnetically ordered phases when the finite amplitude of the Heisenberg exchange coupling is introduced. We examine temperature dependence of dynamical spin structure factors of the Kitaev-Heisenberg model in proximity to the quantum spin liquid. It is clarified that the crossover from a spin-excitation continuum, which is a characteristic of the quantum spin liquid, to a damped high-energy magnon mode occurs at temperatures higher than the energy scale of the Heisenberg exchange couplings or the spin gap that is a signature of the quantum spin liquid at zero temperature. The crossover and the closeness to the Kitaev’s quantum spin liquid are quantitatively measured by the width of the excitation continuum or the magnon spectrum. The present results shed new light on analysis of neutron scattering and other spectroscopy measurements on spin-liquid candidates.

I. INTRODUCTION

An expectation value of an observable $\hat{O}$ in equilibrium at inverse temperature $\beta$ is given by the canonical ensemble average,

$$
\langle \hat{O} \rangle^{\text{ens}}_{\beta} = \sum_{\nu} \frac{e^{-\beta E_{\nu}}}{Z(\beta)} \langle \nu | \hat{O} | \nu \rangle
$$

for the many-body quantum system described by the Hamiltonian $\hat{H}$, where $E_{\nu}$ and $| \nu \rangle$ are eigenvalues and orthonormalized eigenvectors of $\hat{H}$, respectively. Here, $Z(\beta)$ is the partition function given by $Z(\beta) = \sum_{\nu} e^{-\beta E_{\nu}}$. Although the formula is simple, the evaluation of it is not straightforward. Even if modern supercomputers are employed, it remains difficult due to exponential wall. 6

The Fock space dimension $N_{F}$ of the many-body quantum system increases exponentially with the total number of particles in or the size of the target system. When the system is composed of mutually interacting $S=1/2$ quantum spins (or interacting qubits), the Fock space dimension is given as $N_{F} = 2^{N}$, where $N$ is the total number of the spins (or qubits). Since the straightforward evaluation of the canonical ensemble average requires every eigenvalue and eigenvector of $\hat{H}$, the memory cost for it is scaled by $N^{2}$ and the computational cost of it is scaled by $N^{3}$. For example, when the target system consists of $N = 24$ $S=1/2$ quantum spins, it requires storing $2^{48} \sim 3 \times 10^{14}$ complex numbers ($\sim$ 4PB) on memory and, at least, $2^{72} \sim 4 \times 10^{21}$ floating-point operations. Thus, even for the largest-ever supercomputer, it has remained a formidable task.

Although the ensemble average is the canonical prescription of statistical mechanics, it has been found and elucidated that the canonical or microcanonical ensemble is neither the only way to calculate the equilibrium expectation value of the observable at finite temperatures nor to construct statistical mechanics. 7

Indeed, in the thermodynamic limit, it has been proven that a single pure state, which is called a typical pure state, replaces the canonical ensemble. 8

Even for finite-size and finite $N_{F}$ systems, the computational cost of calculating $\langle \hat{O} \rangle^{\text{ens}}_{\beta}$ is reduced from $O(N_{F}^{3})$ to $O(N_{F}^{2})$ A typical pure state is constructed by utilizing imaginary time evolution of a random vector. 9

The pure state approach may recall the thermo field dynamics that replaces the canonical ensemble with a statistical pure state. 10 However, we note that the pure state approach does not assume equal a priori probability while the thermo field dynamics is constructed by keeping the equal a priori probability.
The $O(N_F)$ method enables us to simulate the systems with twice larger number of particles or spins than the conventional ensemble average. If every eigenstate is stored on memory, the memory cost is scaled by $N_F^2$. However, if a single pure-state wave function reproduces finite-temperature expectation values of observables, the memory cost is scaled by $N_F$. Therefore, even though several pure states are required in practical simulation, due to the exponential dependence of $N_F$ on $N$, the typical pure state approach can handle twice larger system size.

In addition to thermodynamic quantities in equilibrium, excitation spectra at finite temperatures are also accessible with computational costs of $O(N_F)$. The excitation spectra at finite temperatures are calculated by constructing a set of excited states by the Lanczos method\ref{11,12} by simulating real-time evolution of a typical pure state\textsuperscript{16–23} or by constructing a microcanonical ensemble\textsuperscript{24,25,26}, whose computational costs are of $O(N_F)$. The first method has been considered that hundreds of initial random vectors are necessary to obtain accurate results. In the second method, the excitation spectra are achieved by the Fourier transformation of the real-time evolution with an appropriate perturbation. The accuracy of the approach is guaranteed by typicality in the real-time evolution.\textsuperscript{17,18,20,21} However, longer real-time simulation is required to obtain lower energy spectra in this approach. The third method requires determining the temperature that corresponds to the obtained microcanonical shell independently. A more efficient and self-contained method has been desirable.

A growing demand for finite-temperature simulation of excitation spectra has originated from experimental researches on many-body quantum systems. As a typical example, the electron spin resonance in strongly correlated electrons has stimulated not only theoretical studies\textsuperscript{26–29}, but also, numerical studies\textsuperscript{16,19,31} on the finite-temperature excitation spectra.

Raman scattering and inelastic neutron scattering measurements on a class of quantum magnets, so-called Kitaev materials\textsuperscript{32}, have brought renewed attention to the temperature dependence of the excitation spectra\textsuperscript{31,33}. Pioneered by Kitaev, a family of exactly solvable two-dimensional quantum spin Hamiltonians has been found\textsuperscript{32,33}, which does not exhibit any spontaneous symmetry breaking down to zero temperature and, thus, shows spin liquid ground state\textsuperscript{32}. The models in the family are generically called the Kitaev models. Shortly after the findings, it has been proposed\textsuperscript{33} that the Kitaev model on the two dimensional honeycomb lattice captures low-energy spin degrees of freedom in heavy transition metal oxides with honeycomb networks of transition metal ions, which are typified by an iridium oxide $\alpha$-Na$_2$IrO$_3$\textsuperscript{34,35}. So far, the Kitaev materials including $\alpha$-A$_2$IrO$_3$ ($A=$ Na, Li), $\alpha$-RuCl$_3$, hyperhoneycomb iridate $\beta$-Li$_2$IrO$_3$\textsuperscript{36}, and stripy-honeycomb iridate $\gamma$-Li$_2$IrO$_3$\textsuperscript{37} exhibit spontaneous time-reversal symmetry breakings. However, these materials expected in proximity to the Kitaev’s spin liquid stimulate the experimental research on the excitation spectra at finite temperatures, which requires theoretical counterparts. In addition to the theoretical studies on Raman spectra and dynamical spin structure factors of the simple Kitaev model at zero temperature\textsuperscript{34,35} and finite temperatures\textsuperscript{37–39}, theoretical and numerical studies on the Kitaev-like Hamiltonian on variety of tricoordinate networks\textsuperscript{40} and more realistic effective Hamiltonians\textsuperscript{40,41,42} are highly desirable.

In this paper, we propose an $O(N_F)$ algorithm for simulating exact finite-temperature excitation spectra in frequency domain by combining the typical pure state approach and the shifted Krylov subspace method\textsuperscript{43}. The present algorithm rewrites the Lehmann representation of the Green’s function by utilizing randomly-taken linear combination of eigenstates in an equi-energy shell (a microcanonical shell), instead of eigenstates. The combination of the typical pure states and the shifted Krylov subspace method leads to $O(N_F)$ construction of the linear combination of eigenstates in a single equi-energy shell. The linear combination is a typical pure state that corresponds to a microcanonical shell.\textsuperscript{43} We name the present algorithm finite-temperature shifted Krylov subspace method for simulating spectra (FTK$\omega$). The construction of the microcanonical shells is compatible with the reweighting method\textsuperscript{44} that reduces computational costs for tuning temperature. It also ensures parallelizability of the FTK$\omega$ algorithm and makes it suitable for massively parallel environments.

We show an application of the FTK$\omega$ to the simplest effective Hamiltonian of the two-dimensional Kitaev-like systems on honeycomb lattices, namely, the Kitaev-Heisenberg model\textsuperscript{45}. When the Heisenberg exchange coupling, which breaks the integrability of the Kitaev model, is introduced, quantum phase transitions between the Kitaev’s spin liquid and magnetically ordered states are realized\textsuperscript{46,47,48}. We focus on the proximity of the phase boundary between the Kitaev’s spin liquid phase and a magnetically ordered phase, where characteristics of the Kitaev’s spin liquid, such as the thermal fractionalization\textsuperscript{49} are observed by heating the magnetically ordered ground states.\textsuperscript{50} We examine temperature dependence of dynamical spin structure factors of the Kitaev-Heisenberg model and clarify that the crossover from a spin-excitation continuum, which is a characteristics of the quantum spin liquid, to a damped high-energy magnon mode occurs at temperatures higher than the energy scale of the Heisenberg exchange couplings or the spin gap that is a signature of the quantum spin liquid at zero temperature. The crossover and the closeness to the quantum spin liquid are quantitatively measured by a dimensionless ratio of the width of the excitation continuum or the damped magnon spectrum and the energy at which the spectral weight becomes maximum. The present results shed new light on analysis of neutron scattering and other spectroscopy measurements on the spin-liquid candidates.

The rest of the paper is organized as follows. We review
the typical pure state approach and the shifted Krylov subspace method in section II. In section III the \(O(N_F)\) algorithm is detailed. The computational costs and parallelizability are examined in section IV. The application of the present algorithm to the Kitaev and Kitaev-Heisenberg models is shown in section V. Section VI is devoted to the summary and discussion.

II. PRELIMINARIES

Before going to the formulation of the \(O(N_F)\) FTK\(\omega\) algorithm, we briefly review the two building blocks of the algorithm to make this paper self-contained: The typical pure state approach and the shifted Krylov subspace method are explained in the following.

A. Typical pure state approach

First, we explain that a typical pure state indeed replaces the canonical ensemble at infinite temperature. At \(\beta = 0\), a typical pure state is nothing but a random vector \(\{c_x\}\) as shown below. A random vector is easily constructed by employing real space configurations \(\{|x\}\) as,

\[
|\phi_0\rangle = \sum_x c_x |x\rangle,
\]

where \(\{c_x\}\) is a set of random complex numbers that satisfies the normalization condition \(\sum_x |c_x|^2 = 1\). If we focus on a quantum lattice model consisting of \(N_S=1/2\) spins, the real space configurations are simply given as sets of binary bits, such as \(|x\rangle = |\sigma_0\sigma_\uparrow \cdots \sigma_{N-1}\rangle\) \((\sigma_j = 0, 1, \sigma_j = \uparrow, \downarrow)\), which are easy to implement. Uniform distribution on the unit sphere in \(\mathbb{R}^{2N_F}\) is often used as probability distribution for the set of the random numbers \(\{c_x\}\). Then, the average of the expectation value \(\langle \phi_0 | \hat{O} | \phi_0 \rangle\) over the uniform probability is trivially equal to the ensemble average \(\langle \hat{O} \rangle^\text{ens}_{\beta=0}\) as

\[
\mathbb{E}[\langle \phi_0 | \hat{O} | \phi_0 \rangle] = N_F^{-1} \sum_{\nu} \langle \nu | \hat{O} | \nu \rangle = \langle \hat{O} \rangle^\text{ens}_{\beta=0},
\]

where \(\mathbb{E}[\{c_x\}]\) denotes the average of a function \(f(\{c_x\})\) over the uniform probability, \(\mathbb{E}[|c_x|^2] = N_F^{-1}\), and the unitary transformation \(|\nu\rangle = \sum_x U_{x\nu} |x\rangle\) are used.

A non-trivial fact is exponentially small variance of the difference \(\delta\hat{O} = \langle \phi_0 | \hat{O} | \phi_0 \rangle - \langle \hat{O} \rangle^\text{ens}_{\beta=0}\), which is given by

\[
\mathbb{E}[\delta\hat{O}] = \frac{N_F^{-1} \text{Tr} [\hat{O}^2] - N_F^{-2} \text{Tr} [\hat{O}]^2}{N_F + 1},
\]

where \(\text{Tr} \hat{O} = \sum_{\nu} \langle \nu | \hat{O} | \nu \rangle\). The above formula Eq.[4] is found by Hams and De Raedt. Later, Sugita and Reimand obtained essentially the same results independently. If we set \(\hat{O} = H\) and recall that width of energy distribution is scaled by \(N\) as

\[
N_F^{-1} \text{Tr} [\hat{H}^2] - N_F^{-2} \text{Tr} [\hat{H}]^2 \propto N,
\]

we obtain exponentially small variance of energy estimated by the typical state \(|\phi_0\rangle\): The variance \(\mathbb{E}[(\delta H^2)]\) turns out to be proportional to \(N/(N_F + 1) < N2^{-N}\) for \(N\) quantum spins by utilizing Eq.[4].

The partition function at a finite temperature is also obtained by the typical state \(|\phi_0\rangle\). If we set \(\hat{O} = e^{-\beta \hat{H}}\), we obtain \(Z(\beta) = N_F \mathbb{E}[(\phi_0 | e^{-\beta \hat{H}} | \phi_0 )]\). The variance of the typical-state estimate \(N_F \mathbb{E}[(\phi_0 | e^{-\beta \hat{H}} | \phi_0 )]^2\) is also given by Eq.[4]: The upper bound of the variance is estimated as

\[
\mathbb{E} \left[ \left( N_F |\phi_0 | e^{-\beta \hat{H}} | \phi_0 \rangle - Z(\beta) \right)^2 \right] < e^{-S(\beta^*)},
\]

where \(S(\beta)\) is entropy at the inverse temperature \(\beta\) and \(\beta^*\) is a constant that satisfies \(\beta < \beta^* < 2\beta\).

Ensemble average of an observable \(\hat{O}\) other than \(e^{-\beta \hat{H}}\) is replaced by the expectation value with the following typical pure state \(|\phi_\beta\rangle\),

\[
|\phi_\beta\rangle = e^{-\beta \hat{H}/2} |\phi_0\rangle,
\]

which is obtained through imaginary-time evolution initialized with \(|\phi_0\rangle\). The details of the imaginary-time evolution are given in Appendix A. The ensemble average of \(\hat{O}\) is given by

\[
\langle \hat{O} \rangle^\text{ens}_{\beta} = \frac{\mathbb{E} [\langle \phi_\beta | \hat{O} | \phi_\beta \rangle]}{\mathbb{E} [\langle \phi_\beta | \phi_\beta \rangle]}.
\]

Variance of the estimate by the typical pure state \(|\phi_\beta\rangle\),

\[
\sigma^2_{\hat{O}} = \mathbb{E} \left[ \left( \frac{\langle \phi_\beta | \hat{O} | \phi_\beta \rangle}{\langle \phi_\beta | \phi_\beta \rangle} - \langle \hat{O} \rangle^\text{ens}_{\beta} \right)^2 \right],
\]

is also bounded as

\[
\sigma^2_{\hat{O}} \leq \left( \frac{\langle \hat{O} \rangle^\text{ens}_{\beta} - \langle \hat{O} \rangle^\text{ens}_{\beta}}{\langle \hat{O} \rangle^\text{ens}_{\beta}} \right)^2 \text{exp}[2\beta \{F(2\beta) - F(\beta)\}],
\]

which is derived by Sugiura and Shimizu. Here, \(F(\beta)\) is free energy given by \(F(\beta) = -k_B T \ln Z(\beta)\).

For later use, we briefly explain the evaluations of entropy and heat capacity by the typical pure states. The strict definition for entropy and heat capacity is as follows: Entropy \(S(\beta)\) is given by using the identity \(F(\beta) = \langle \hat{H} \rangle^\text{ens}_{\beta} - TS(\beta)\) as

\[
S(\beta) = \frac{\mathbb{E} [\langle \phi_\beta | \hat{H} | \phi_\beta \rangle]}{\beta \mathbb{E} [\langle \phi_\beta | \phi_\beta \rangle]} + k_B \ln (N_F \mathbb{E} [\langle \phi_\beta | \phi_\beta \rangle]),
\]

and heat capacity \(C(\beta)\) is simply given by

\[
C(\beta) = \frac{\mathbb{E} [\langle \phi_\beta | \hat{H}^2 | \phi_\beta \rangle] - \mathbb{E} [\langle \phi_\beta | \hat{H} | \phi_\beta \rangle]^2}{k_B T^2}. \]
However, for simplicity, we use simplified estimates of entropy and heat capacity as
\[
S(\beta) = E \left[ \frac{\langle \phi_\beta | \hat{H} | \phi_\beta \rangle}{T \langle \phi_\beta | \phi_\beta \rangle} + k_B \ln (N_F \langle \phi_\beta | \phi_\beta \rangle) \right], \quad (13)
\]
and
\[
C(\beta) = \frac{E \left[ \frac{\langle \phi_\beta | \hat{H} | \phi_\beta \rangle^2}{\langle \phi_\beta | \phi_\beta \rangle} \right] - \langle \phi_\beta | \hat{H} | \phi_\beta \rangle^2}{k_BT^2}. \quad (14)
\]
These simplified formulae do not cause deviation from the exact ones given by Eqs. \[11\] and \[12\] beyond standard deviation and standard errors at least in the application to frustrated magnets.

B. Shifted Krylov subspace method

Excitation spectra are given by taking imaginary parts of Green’s functions in the linear response theory. We start with the following Green’s function at zero temperature,
\[
G^{AB}(\zeta) = \langle 0 | \hat{A}^\dagger (\zeta - \hat{H})^{-1} \hat{B} | 0 \rangle, \quad (15)
\]
where $|0\rangle$ is the ground state. Although the standard $O(N_F)$ approach to calculate the excitation spectra of correlated electrons systems is the Lanczos method\[^{53}\] here, we review an alternative approach below.

To evaluate the above formula, we solve a linear equation by employing a conjugate gradient (CG) method, instead of calculating the expectation value of the resolvent of $\hat{H}$ by the Lanczos method. The CG methods find a solution in a subspace of the Fock space as follows. First, by introducing the following three vectors,
\[
\begin{align*}
|\lambda\rangle &= \hat{A} |0\rangle, \\
|\rho\rangle &= \hat{B} |0\rangle, \\
|\chi(\zeta)\rangle &= (\zeta - \hat{H})^{-1} |\rho\rangle,
\end{align*}
\]
we rewrite $G^{AB}(\zeta)$ as
\[
G^{AB}(\zeta) = \langle \lambda | \chi(\zeta) \rangle, \quad (19)
\]
where $|\chi(\zeta)\rangle$ is unknown. To obtain the unknown vector $|\chi(\zeta)\rangle$, we solve the following linear equation,
\[
(\zeta - \hat{H}) |\chi(\zeta)\rangle = |\rho\rangle. \quad (20)
\]

When the matrix $\hat{H}$ is not able to be stored in the memory but a few wave functions can be stored, the linear equation is solved iteratively, for example, by using the CG methods. At $n$th iteration, the CG algorithm initialized with $|\chi_0(\zeta)\rangle = |\rho\rangle$ finds an approximate solution $|\chi_n(\zeta)\rangle$ within a $n$-dimensional Krylov subspace $K_n(\zeta - \hat{H}, |\rho\rangle) = \text{span}\{|\rho\rangle, (\zeta - \hat{H}) |\rho\rangle, \ldots, (\zeta - \hat{H})^{n-1} |\rho\rangle\}$. For complex matrices and vectors, a variation of the CG algorithm, the biconjugate gradient (BiCG), is employed. At each steps, the BiCG algorithm searches an approximate solution $|\chi_n(\zeta)\rangle$ by utilizing a biorthogonal basis set. The biorthogonal basis set consists of the residual vectors, $\{|\rho_0(\zeta)\rangle, |\rho_1(\zeta)\rangle, \ldots, |\rho_{n-1}(\zeta)\rangle\}$ and $\{|\tilde{\rho}_0(\zeta)\rangle, |\tilde{\rho}_1(\zeta)\rangle, \ldots, |\tilde{\rho}_{n-1}(\zeta)\rangle\}$ that satisfy $\langle \rho_k(\zeta) | \rho_{k'}(\zeta) \rangle \propto \delta_{kk'}$, is iteratively generated by
\[
|\rho_k(\zeta)\rangle = (\zeta - \hat{H}) |\chi_k(\zeta)\rangle - |\rho\rangle, \quad (21)
\]
and
\[
|\tilde{\rho}_k(\zeta)\rangle = (\tilde{\chi}_k(\zeta) | (\zeta^* - \hat{H}) - |\tilde{\rho}\rangle, \quad (22)
\]
where $|\tilde{\rho}\rangle$ is an arbitrary vector with a non-zero 2-norm and a finite internal product $\langle \tilde{\rho} | \rho \rangle$. The approximate solution $|\chi_n(\zeta)\rangle$ is found in the basis set $\{|\rho_0(\zeta)\rangle, |\rho_1(\zeta)\rangle, \ldots, |\rho_{n-1}(\zeta)\rangle\}$ \[\{(\tilde{\rho}_0(\zeta)), (\tilde{\rho}_1(\zeta)), \ldots, (\tilde{\rho}_{n-1}(\zeta))\}\] as
\[
2 \text{norm and a finite internal product } \langle \tilde{\rho} | \rho \rangle. \quad (23)
\]
We note that one needs to solve Eq.\[20\] essentially once at a fixed complex number $\zeta = \hbar \omega + i\eta$ to obtain whole spectrum $-\text{Im}G^{AB}(\hbar \omega + i\eta)$. Due to the shift invariance of the Krylov subspace, namely, $K_n(\zeta - \hat{H}, |\phi\rangle) = K_n(\zeta' - \hat{H}, |\phi\rangle)$ for any complex number $\zeta' \neq \zeta$, the biorthogonal bases $|\rho_k(\zeta)\rangle$ and $|\tilde{\rho}_k(\zeta)\rangle$ are proportional to the other biorthogonal bases $|\rho_k(\zeta')\rangle$ and $|\tilde{\rho}_k(\zeta')\rangle$, respectively\[^{54}\]. Then, one can obtain $|\chi_n(\zeta')\rangle$ from $|\chi_n(\zeta)\rangle$ without the matrix-vector multiplication\[^{54}\]. The Krylov subspace methods utilizing the shift invariance are called the shifted Krylov subspace methods.

In this study, we employ the shifted BiCG method\[^{54}\] implemented in a numerical library $K_\omega$ for the shifted Krylov subspace method\[^{54}\]. The condition for truncating the shifted BiCG iteration and the dimension of the Krylov subspace required for the convergence are examined later. Typical dimension of the Krylov subspace is of the order of ten thousand at most in the present application.

III. FINITE-TEMPERATURE DYNAMICAL SPIN STRUCTURE FACTOR

If every eigenvalue $\{E_\nu\}$ and eigenvector $\{|\nu\rangle\}$ of the Hamiltonian $\hat{H}$ are known, the Green’s function at a finite temperature $\beta^{-1}$ is given as
\[
G^{AB}_\beta(\hbar \omega + i\eta) = \sum_{\nu,\mu} \frac{e^{-\beta E_\nu}}{Z(\beta) \hbar \omega + i\eta + E_\nu - E_\mu} |\nu\rangle \hat{A}^\dagger |\mu\rangle |\mu\rangle \hat{B} |\nu\rangle, \quad (23)
\]
where $Z$ is the partition function of the system defined as $Z(\beta) = \sum_\nu e^{-\beta E_\nu}$. In the following sections, we will formulate an algorithm to estimate $G^{AB}_\beta$ with computational costs of $O(N_F)$, and give upper bounds of errors in the estimate. For later use, we rewrite the above expression of $G^{AB}_\beta$ as
\[
G^{AB}_\beta(\hbar \omega + i\eta) = \sum_\nu \frac{e^{-\beta E_\nu}}{Z(\beta) \hbar \omega + i\eta + E_\nu - \hat{H}} |\nu\rangle \hat{A} \frac{1}{\hbar \omega + i\eta + E_\nu - \hat{H}} \hat{B} |\nu\rangle. \quad (24)
\]
A. Intuitive overview

Here, we reformulate Eq. (24) with a typical pure state $|\phi_\beta\rangle$ to avoid using the whole set of $E_\nu$ and $|\nu\rangle$. First, we note that the normalized typical state is naively expected to behave as

$$
|\psi_\beta\rangle \equiv \frac{|\phi_\beta\rangle}{\sqrt{\langle \phi_\beta | \phi_\beta \rangle}} \sim \sum_\nu e^{i \varphi_\nu} e^{-\frac{\beta}{2} E_\nu} \sqrt{Z(\beta)} |\nu\rangle, \quad (25)
$$

where each $\varphi_\nu$ is a random variable distributed over the interval $[0, 2\pi)$. By introducing a projection operator,

$$
\hat{P}_\nu = |\nu\rangle \langle \nu|, \quad (26)
$$

we rewrite the formula based on canonical ensemble, Eq. (24), as

$$
G_{AB}^\beta(\zeta) \sim \sum_\nu \langle \psi_\beta | \hat{P}_\nu \hat{A}^\dagger \frac{1}{\zeta + E_\nu - H} \hat{B} \hat{P}_\nu | \phi_\beta \rangle. \quad (27)
$$

Thus far, there is no reduction of computational costs from the original formula Eq. (24), since the exact projection operator $\hat{P}_\nu$ requires the whole set of $|\nu\rangle$.

The important step is to find an economical and practical implementation of the projection operator $\hat{P}_\nu$. Although there is no $O(N_F)$ implementation of the exact $\hat{P}_\nu$ in the literature as far as we know, there is a filter operator $\{\hat{P}_\nu\}$ that constructs equi-energy shells and is realizable with the computational cost of $O(N_F)$ by employing the shifted Krylov method, as follows.

B. Filter operator and equi-energy shells

The filter operator $\{\hat{P}_\nu\}$ is defined by integrating the resolvent of $H$ along a contour $C_{\gamma, \rho}$ defined by $z = \rho e^{i\theta} + \gamma$ with $0 \leq \theta < 2\pi$ as

$$
\hat{P}_{\gamma, \rho} = \frac{1}{2\pi i} \oint_{C_{\gamma, \rho}} \frac{dz}{z-H}. \quad (28)
$$

If the filter operator is applied to an arbitrary wave function $|\phi\rangle = \sum_\nu d_\nu |\nu\rangle$, the operator filters the eigenvectors with the eigenvalues $E_\nu \not\in (\gamma - \rho, \gamma + \rho)$ as

$$
\hat{P}_{\gamma, \rho} |\phi\rangle = \sum_{E_\nu \in (\gamma - \rho, \gamma + \rho)} d_\nu |\nu\rangle. \quad (29)
$$

When a small $\rho$ limit is taken, the filter operator realizes a microcanonical ensemble. The filter operator is practically implemented as a Reimann sum $\sum_{\nu}$. The discretized filter operator is defined as

$$
\hat{P}_{\gamma, \rho, M} = \frac{1}{M} \sum_{j=1}^{M} \rho e^{i\theta_j} + \gamma - H, \quad (30)
$$

where $\theta_j = 2\pi (j - 1/2)/M$. The discretized contour is illustrated in Fig. 1. Multiplication of $\hat{P}_{\gamma, \rho, M}$ to a wave function is simply realized by the shifted Krylov subspace method while it is hardly achievable by the standard Lanczos algorithm.

By introducing an appropriate grid measured from the low-energy onset $E_b$ in energy axis,

$$
\epsilon_m = E_b + (2m + 1)\epsilon, \quad (31)
$$

the set of the filter operators $\{\hat{P}_{\epsilon_m, \epsilon, M}\}$ with the discretization parameters,

$$
\delta = (E_b, \epsilon, M), \quad (32)
$$

indeed replace the projection operators $\{\hat{P}_\nu\}$. The filtered typical state given by

$$
|\phi_{\beta, \delta}^m\rangle = \hat{P}_{\epsilon_m, \epsilon, M} |\phi_\beta\rangle \quad (33)
$$

is a random vector residing in an equi-energy shell $(\epsilon_m - \epsilon, \epsilon_m + \epsilon)$, which corresponds to a microcanonical ensemble. From the filtered typical pure states, we obtain a discretized formula for probability distribution as

$$
\tilde{P}_{\beta, \delta}^m (\epsilon_m) = \frac{1}{2\epsilon} \frac{\langle \phi_{\beta, \delta}^m | \hat{A}^\dagger \frac{1}{\zeta + \epsilon_m - H} \hat{B} | \phi_{\beta, \delta}^m \rangle}{\langle \phi_{\beta} | \phi_{\beta} \rangle}. \quad (34)
$$

In the following application, we prepare the $L$ filter operators to cover an energy range $[E_b, E_b + 2L\epsilon]$. Here, $E_b$ and $L$ are chosen to keep the probability distribution $\tilde{P}_{\beta, \delta}^m (E)$ smaller than $10^{-14}$ outside the energy range $[E_b, E_b + 2L\epsilon]$.

C. Green’s function

A representation of the Green’s function is thus achieved by employing the filtered typical pure states $\{|\phi_{\beta, \delta}^m\rangle\}$ as

$$
\tilde{G}_{\beta, \delta}^{AB}(\zeta) = \sum_{m=0}^{L-1} \frac{\langle \phi_{\beta, \delta}^m | \hat{A}^\dagger \frac{1}{\zeta + \epsilon_m - H} \hat{B} | \phi_{\beta, \delta}^m \rangle}{\langle \phi_{\beta} | \phi_{\beta} \rangle}. \quad (35)
$$

After taking appropriate limits and average over the distribution of initial random vectors $|\phi_0\rangle$, we indeed replace the canonical ensemble prescription by combining
the typical pure states and the shifted Krylov subspace method as
\[
G^{AB}_{\beta}(\zeta) = \lim_{\epsilon \to +0} \lim_{M \to +\infty} \frac{E[\langle \phi_\beta | \phi_\beta \rangle G^{AB}_{\beta}(\zeta)]}{E[\langle \phi_\beta | \phi_\beta \rangle]}.
\]
(36)

For simplicity, we use the normalized filtered typical pure state,
\[
|\psi^m_{\beta, \delta}\rangle = \frac{|\phi^m_{\beta, \delta}\rangle}{\sqrt{\langle \phi_\beta | \phi_\beta \rangle}},
\]
(37)

instead of \(|\phi^m_{\beta, \delta}\rangle\) in Eq. (35) and replace the denominator of the righthand side with unity.

The deviation between the pure-state representation \(G^{AB}_{\beta}(\zeta)\) and the canonical-ensemble representation \(\tilde{G}^{AB}_{\beta}(\zeta)\) is bounded as follows. The source of the deviation is twofold: The discretization parameters \(\delta = (e_B, \epsilon, M)\) and the variance of the stochastic variables \(\{e_m\}\). The former source can be examined by changing the set of the discretization parameters \(\delta\).

Ref. 6 and 14, the upper bound of the variance between \(G^{AB}_{\beta}(\zeta)\) and \(\tilde{G}^{AB}_{\beta}(\zeta)\) due to the variance of \(\{e_m\}\) is estimated as
\[
E[|G^{AB}_{\beta}(h \omega + i \eta) - \tilde{G}^{AB}_{\beta}(h \omega + i \eta)|^2] 
\leq e^{-2\beta \{F(2\beta) - F(\beta)\} - \pi \frac{1}{2}(B^\dagger A A^\dagger B)_{\text{ens}}^{2\delta}} \left[ \frac{\tau}{\eta} - \frac{\eta + h/\tau}{\tau} \right] + |G^{AB}_{\beta}(h \omega + i \eta) - \tilde{G}^{AB}_{\beta}(h \omega + i \eta)|^2,
\]
(38)

where constants \(\eta\) and \(\tau\) satisfy \(\min \{\eta, \epsilon\} < \eta < O(N, J_0)\) and \(\tau > 0\), respectively. The details of the derivation is given in Appendix A. As pointed out in Ref. 14, the prefactor \(e^{-2\beta \{F(2\beta) - F(\beta)\}}\) in the upper bound of the standard deviation is estimated by entropy: Due to the convex nature of free energy, there is an inverse temperature \(\beta^* \in (\beta, 2\beta)\) that satisfies \(S(\beta^*) = 2\beta \{F(2\beta) - F(\beta)\}\), where \(S\) is entropy. Therefore, the prefactor exponentially decreases with increasing \(N\), since entropy is extensive quantity. The variance Eq. (38) is a counterpart in frequency domain of the variance in time domain.

D. Reweighting

The present FTK_\omega algorithm based on the filtered typical states resembles histogram techniques, \(\Phi[10]\), which were introduced to exploit Monte Carlo simulation data. Indeed, by reweighting the Boltzmann factors in the filtered typical states \(\{|\phi^m_{\beta, \delta}\rangle\}\), we can calculate a finite-temperature expectation value of any operator \(\hat{O}\) at an inverse temperature \(\beta^*\) different from \(\beta_0\). The filtered typical states for \(\beta^*\) are given by
\[
|\phi^m_{\beta^*, \delta}\rangle = e^{\frac{\beta^* - \beta}{\beta_0} \epsilon_m} |\phi^m_{\beta, \delta}\rangle,
\]
(39)

where the factor \(\exp[(\beta - \beta^*)\epsilon_m/2]\) is a c-number. Since the c-number and any operators commute, the expectation value of any operator \(\hat{O}\) taken by the filtered state \(|\phi^m_{\beta^*, \delta}\rangle\) is given by
\[
\langle \phi^m_{\beta^*, \delta} | \hat{O} | \phi^m_{\beta^*, \delta} \rangle = e^{(\beta - \beta^*)\epsilon_m} \langle \phi^m_{\beta, \delta} | \hat{O} | \phi^m_{\beta, \delta} \rangle.
\]
(40)

Therefore, if once we calculate \(\langle \phi^m_{\beta, \delta} | \hat{O} | \phi^m_{\beta, \delta} \rangle\) and \(\langle \phi^m_{\beta^*, \delta} | \hat{O} | \phi^m_{\beta^*, \delta} \rangle\) for every \(m \in [0, L]\), we can estimate the expectation value of \(\hat{O}\) at \(\beta^*\) \((\neq \beta)\) by the following simple expression,
\[
\langle \hat{O} \rangle_{\beta^*} = \frac{\sum_{m=0}^{L-1} e^{(\beta - \beta^*) \epsilon_m} \langle \phi^m_{\beta^*, \delta} | \hat{O} | \phi^m_{\beta^*, \delta} \rangle}{\sum_{\ell=0}^{L-1} e^{(\beta - \beta^*) \epsilon_\ell} \langle \phi^\ell_{\beta^*, \delta} | \hat{O} | \phi^\ell_{\beta^*, \delta} \rangle}.
\]
(41)

The probability distribution for \(\beta^*\) is also obtained as
\[
\tilde{P}_{\beta^*, \delta}(\epsilon_m) = \frac{1}{2K} \sum_{\ell=0}^{L-1} e^{(\beta - \beta^*) \epsilon_\ell} \langle \phi^\ell_{\beta^*, \delta} | \hat{O} | \phi^\ell_{\beta^*, \delta} \rangle.
\]
(42)

The reweighting method works and significantly reduces the computational cost for tuning temperature if the probability distributions at the original temperature \(1/\beta\) and the target temperature \(1/\beta^*\) are overlapped each other. The ratio of the difference in internal energy and width of these distributions determines whether these distributions are overlapped or not: If the ratio,
\[
r = \frac{|\langle \hat{H} \rangle_{\text{ens}}^{\beta} - \langle \hat{H} \rangle_{\text{ens}}^{\beta^*}|}{\sqrt{C(\beta)/\langle L \rangle_k B^2} + \sqrt{C(\beta^*)/\langle L \rangle_k B^2}},
\]
(43)
is small, the probability distributions at \(\beta\) and \(\beta^*\) are overlapped each other. Here, we approximate the probability distribution at \(\beta\) by a Gaussian distribution \(N(\beta) = e^{-(\beta - \beta_0)^2/2\sigma^2}/\sqrt{2\pi\sigma^2}\) with the variance \(\sigma^2 = C(\beta)/\langle L \rangle_k B^2\). When the system size \(N\) increases with fixed \(\beta\), the ratio \(r\) increases at the square root of \(N\). Therefore, if we enlarge the system size with the fixed difference \(1/\beta - 1/\beta^*\), the overlap decreases. As shown in Sec. V, the reweighting method works for the finite-size spin clusters up to \(N = 24\) if the difference \(1/\beta - 1/\beta^*\) is appropriate. In the present application, we choose the difference \(1/\beta - 1/\beta^*\) that corresponds to \(r \approx 0.6\) and confirm the accuracy of the reweighting by comparing the reweighted spectrum at \(\beta^*\) starting from \(\beta\) with the spectrum directly calculated at \(\beta^*\).
IV. COSTS AND PARALLELIZABILITY

A. Numerical cost of the present algorithm

We examine numerical costs and parallelizability of the present method summarized in Eqs. (33) and (35). The most time-consuming part is shared by the present method and other related ones,\textsuperscript{16,19,21–25} It is multiplication between the Hamiltonian matrix $\hat{H}$ and a wave function, which is the most time-consuming operation of a single Lanczos step in the Lanczos method. Thus, the numerical costs of these methods are measured by the number of the Lanczos steps (or matrix-vector multiplications).

The numerical cost of the present method is determined by the number of the Lanczos steps for the imaginary-time evolution in Eq. (7), the filter operation in Eqs. (33), the calculation of the equi-energy Green’s function $\langle \psi^m_{\delta \alpha} | \hat{A}^\dagger (\zeta + \mathcal{E}_m - \hat{H})^{-1} \hat{B} | \psi^m_{\delta \alpha} \rangle$ in Eq. (35), and the number of the equi-energy shells, where these numbers are denoted by $N^{(r)}_L$, $N^{(P)}_L$, $N^{(eG)}_L$, and $N_E (= L)$, respectively. Then the numerical cost is scaled by

$$ N^{(r)}_L + (\alpha_K N^{(P)}_L + N^{(eG)}_L) \times N_E, \tag{44} $$

where $\alpha_K$ is a factor larger than 1 due to the additional cost of the shifted Krylov subspace method. Here, we note that the naive implementation of the imaginary time evolution (see Appendix A) is accurate but less efficient than the implementation by polynomial expansion of the imaginary-time-evolution operator,\textsuperscript{13} and the microcanonical TPQ algorithm.\textsuperscript{12} However, $N^{(r)}_L$ is negligible in the practical simulations at high and moderate temperatures.

B. Numerical cost of related methods

In comparison, we also estimate the numerical costs of the closely related methods, namely, the Boltzmann-weighted time-dependent method (BWTDM)\textsuperscript{12} and the microcanonical Lanczos method (MCLM).\textsuperscript{22} The BWTDM consists of the imaginary-time and real-time evolution. Therefore, the numerical cost of the BWTDM is scaled by $N^{(r)}_L + N^{(T)}_L$, where $N^{(T)}_L$ is the number of the Lanczos steps for the real-time evolution and is proportional to the number of the time steps. The computational cost of the other related methods may be the same.\textsuperscript{16,21–22} The MCLM, instead, consists of projection to obtain a pure state in a microcanonical shell and the calculation of Green’s function by employing the standard Lanczos method. The projection is realized through obtaining an eigenstate with the lowest eigenvalue of $(\hat{H} - \lambda)^2$ by the Lanczos method, where $\lambda$ is the target energy. In the practical applications of the MCLM, a single microcanonical shell is used for finite-temperature simulation,\textsuperscript{22} although the MCLM can be used to take a canonical ensemble average by constructing multiple microcanonical shells.\textsuperscript{11} Thus, the numerical cost of MCLM is scaled by $N^{(MC)}_L + N^{(eG)}_L$, where $N^{(MC)}_L$ is the number of the Lanczos step for the projection by $(\hat{H} - \lambda)^2$. If we set $N^{(r)}_L = 0$ and $N_E = 1$, the present method essentially reproduces the same results obtained by the MCLM although the projection and the present filter operation are quantitatively different.

C. Advantage of the present algorithm

The present algorithm seems to be more computationally demanding than the previous related methods.\textsuperscript{16,19,21–25} However, the present method has several advantages over the previous ones. The most striking difference between the present algorithm and the related previous approaches is the computational cost for tuning temperature. The previous approaches require to repeat the entire simulation to obtain the linear responses at different temperature. None of the previous approaches is compatible with the reweighting method. In contrast, the present algorithm interpolates potentially exact spectra at temperatures between two adjacent temperature points with negligible numerical costs by employing the reweighting method. The reweighting method works if the overlap of the probability distributions at these two temperatures is significant.

D. Parallelizability

The present algorithm is more parallelizable than the real-time evolution of the typical pure state. The difference in the parallelizability of these two approaches becomes evident, when higher resolution (or smaller broadening factor $\eta$) is required. To obtain higher resolution in frequency, the present FTK needs more filter operators, or larger $N_E (= L)$ and smaller $\epsilon$. On the other hand, the methods based on real-time evolution need longer time steps, or larger $N^{(T)}_L$. While the construction of each filter operator is parallelizable, the real-time evolution is sequential and not parallelizable.

By taking a simple hybrid parallelization scheme, we examine the parallelizability. Aside from parallel efficiency, the numerical costs with the $N_{th}$ threads and $N_{pr}$ processes are scaled as follows if certain schemes of parallelization are chosen. Here, we choose a simple scheme that parallelizes a single Lanczos step by shared memory parallelization with $N_{th}$ threads. The numerical cost of the present method may be scaled as

$$ \frac{N^{(r)}_L}{N_{th}} + \left( \frac{\alpha_K N^{(P)}_L + N^{(eG)}_L}{N_{th}} \right) \times \frac{N_E}{N_{pr}}. \tag{45} $$

In addition to parallelization of an every single Lanczos step, the summation over $m$ in Eq. (35) can be parallelized.
efficiently. The only way to parallelize the BWTDM, in contrast, is the parallelization of the Lanczos step. Therefore, the long-time sequential simulation of the BWTDM required to obtain the accurate low-energy spectra, is bottlenecked by parallelization efficiency of the single Lanczos step.

![Diagram of honeycomb clusters](image)

**FIG. 2.** (color online): Finite-size honeycomb clusters with periodic boundary conditions. Bonds along the three different directions are labeled as the $x$, $y$, and $z$ bond, which are along $-60^\circ$, $60^\circ$, and $0^\circ$ (horizontal) directions, respectively. The 12 site and 18 site clusters shown in (a) and (b), respectively, are used for comparison with results by canonical ensemble. Comparison with the thermodynamic limit of the Kitaev model and application to the Kitaev-Heisenberg model are done for the 24 site cluster depicted in (c).

**V. NUMERICAL RESULTS**

Here, we examine the accuracy of the present $O(N_F)$ algorithm with practical choices of the parameter $\delta$ and show an application to the Kitaev-Heisenberg model. We will calculate dynamical spin structure factors (DSFs) at finite temperature on finite-size honeycomb clusters with periodic boundary conditions, which are illustrated in Fig. 2. By setting $\zeta = \hbar \omega + i \eta$ and

$$\hat{A} = \hat{B} = \hat{S}_{i}^{\alpha}Q = N^{-1/2} \sum_{\ell} e^{+iQ \cdot \mathbf{R}_\ell} \hat{S}_{\ell}^{\alpha},$$

in Eq. (36), we obtain the DSFs at a momentum $Q$ and a frequency $\omega$ as

$$\tilde{S}_{\beta,\delta}(Q, \omega) = -\frac{1}{\pi} \text{Im} \sum_{\alpha=x,y,z} \sum_{m=0}^{L-1} \langle \psi_m^{\beta,\delta} | \hat{S}_{\alpha}^{\delta} - Q \times \hat{H} | \psi_m^{\beta,\delta} \rangle,$$

where $\hat{S}_{\alpha}^{\delta}$ ($\alpha = x, y, z$) is an $S=1/2$ spin operator. We note that $\mathbf{R}_\ell$ is the real space coordinate of the $\ell$ th spin, instead of the position of the unit cell that contains the $\ell$ th spin.

**A. Target Hamiltonian**

The Kitaev-Heisenberg model on a honeycomb lattice\cite{24} consists of $S=1/2$ spins that mutually interact with two types of the nearest-neighbor exchange couplings: The Kitaev coupling\cite{36} and the Heisenberg exchange coupling. The nearest-neighbor bonds on the honeycomb lattice have three different directions. When the three bonds are labeled as $x$, $y$, and $z$, the Kitaev-Heisenberg Hamiltonian,

$$\hat{H} = \sum_{\gamma=x,y,z} \sum_{(i,j) \in \gamma} \hat{H}_{ij}^{(\gamma)},$$

is defined by the exchange coupling for the $\gamma$ ($= x, y, z$) bond,

$$\hat{H}_{ij}^{(\gamma)} = J \hat{S}_{i}^{x} \cdot \hat{S}_{j}^{x} + K \hat{S}_{i}^{z} \hat{S}_{j}^{z},$$

where $K = J_0 \sin \varphi$ is the Kitaev coupling constant and $J = (J_0/2) \cos \varphi$ is the Heisenberg exchange coupling constant. Below, we set the energy unit as $J_0 = 1$. 

![Temperature dependence of $-S(\beta^*)/2k_B$](image)

**FIG. 3.** (color online): Temperature dependence of $-S(\beta^*)/2k_B$ for $\varphi = 90^\circ$. The red (blue) solid curve represents $\exp[-S(\beta^*)/2k_B]$ of the 18 site (24 site) cluster with the periodic boundary condition illustrated in Fig. 2. The shaded gray belts illustrate the standard deviation estimated by four initial random vectors. The vertical red (blue) dashed line indicates $\exp[-(N \ln 2)/2]^{1/2}$ for $N = 18$ ($N = 24$).
The phase diagram of the Kitaev-Heisenberg model has been numerically clarified. The exact diagonalization for a 24 site cluster shows that the striped, Néel, zigzag, and ferromagnetic ordered phases are the ground states for $-76.1^\circ \leq \phi \leq -33.8^\circ$, $-33.8^\circ \leq \phi \leq 87.5^\circ$, $92.2^\circ \leq \phi \leq 161.8^\circ$, and $161.8^\circ \leq \phi \leq 251.8^\circ$, respectively. The spin liquid phase is stabilized for $87.5^\circ \leq \phi \leq 92.2^\circ$ and $251.8^\circ \leq \phi \leq 283.9^\circ = -76.1^\circ$. These phase boundaries are consistent with the previous tensor-network study.

In the following section, we demonstrate capability and efficiency of the present algorithm by calculating DSFs around the phase boundary between the spin liquid phase and the zigzag ordered phase at $\phi \sim 92.2^\circ$.

B. Variance and statistical errors

As clarified in the literature and the present paper, there is an upper bound on variance of finite-temperature physical quantities calculated by the typical pure state approach. The upper bound has been found to be proportional to $\exp[-S(\beta^*)/k_B] = \exp[-2\beta F(\beta) - F(\beta)]$, regardless of whether physical quantities are static or dynamical.

Here, we note that the square root of the variance (standard deviation) is not an estimate of statistical errors in the physical quantities due to the distribution of the random initial vectors. The standard error proportional to $\exp[-S(\beta^*)/k_B]/N_s$ is the estimate of the deviation from physical quantities by canonical ensemble, when the average is taken over $N_s$ initial random vectors.

In frustrated quantum spin systems, sizable entropy often remains even at low temperature. The Kitaev model is an example of such frustrated magnets. The Kitaev model has been shown to exhibit a half plateau in temperature dependence of entropy. The 24 site cluster of the Kitaev model employed in the following analysis shows the half plateau in the temperature range $0.02 \leq k_B T \leq 0.4$. As inferred from the plateau, the factor $(\exp[-S(\beta^*)/k_B])^{1/2}$ remains as small as 0.01 even below $k_B T \sim O(J_0/10)$ for $N = 24$. In Fig. 4 temperature dependence of $(\exp[-S(\beta^*)/k_B])^{1/2}$ is shown for the antiferromagnetic Kitaev model ($\phi = 90^\circ$). The extensive properties of entropy are reflected in the size dependence of $(\exp[-S(\beta^*)/k_B])^{1/2}$, which is evident in difference between the results for the 24 site and 18 site cluster.

At lower temperature than the plateau region, the standard deviation has been known to be much smaller than $(\exp[-S(\beta^*)/k_B])/N_s$ that is a monotonically increasing function of $T$. The reduction of the standard deviation originates from prefactors in the variance.

C. Probability distribution

By applying the filter operators to the typical pure states, we obtain an accurate estimate of probability distribution for the antiferromagnetic Kitaev model ($\phi = 90^\circ$). The results are shown in Fig. 4 for $k_B T = 0.1$, 0.2, 0.5, 1, and $k_B T \rightarrow +\infty$. At high temperature, $P_{\beta,\delta}$ resembles the Gaussian distribution with a width proportional to $\sqrt{\Gamma(\beta)}/k_B \beta^2$. At the high temperature limit, $P_{\beta,\delta}$ becomes nothing but density of states. The discretization parameter $\delta = (E_b, \epsilon, M)$ for the filter operator is chosen as summarized in Appendix C.

As seen in Fig. 4 the probability distributions for these temperatures overlap each other. The overlap among them guarantees that the reweighting method interpolates temperature dependence of physical quantities between the two adjacent temperatures.

D. Comparison with canonical ensemble

To examine accuracy of the present algorithm with a practical choice of discretization parameters $\delta$, we compare the DSFs obtained by the present FTK algorithm denoted by $S_{\beta,\delta}(\Gamma, \omega)$ with those obtained by canonical ensemble denoted by $S_{c}(\Gamma, \omega)$. We employ a 12 site and 18 site clusters, which is illustrated in Fig. 2(a) and (b). The 18 site cluster is practically one of the largest system we can directly take a canonical ensemble average for the Kitaev-Heisenberg model so far.

In Fig. 5 we show the DSFs at $Q = \Gamma$ obtained by the present algorithm and canonical ensemble average for the antiferromagnetic Kitaev model ($\phi = 90^\circ$). Here, the broadening factor $\eta$ is fixed at 0.02. Convergence of the results is examined by changing the discretization parameter $\delta = (E_b, \epsilon, M)$. From $k_B T = 1$ to $k_B T = 0.1$.
0.01, the present FTKω indeed reproduces the canonical ensemble average.

The DSFs of the 12 site cluster at $Q = \Gamma$ are shown in the left panel of Fig. 5. By the binning analysis, we estimate standard error $\sigma_E$ at each frequency: First, we prepare 8 sets of 8 samples (8 sets of 16 samples) at $k_B T = 1$ ($k_B T = 0.1$ and 0.01). Then, we take averages over the samples within each set and estimate the standard error $\sigma_E$ by the standard deviation of the sets of these averages. As the result, we find that $S_c(\Gamma, \omega)$, and $\tilde{S}_{\beta, \delta}(\Gamma, \omega)$ agree within 1 standard error or, at least, 2 standard error.

Since the simulation for the 18 site cluster is more computationally demanding, standard deviation $\sigma_S$ of 8 samples, instead of the standard error, is estimated. The number of the available samples is too small to estimate the standard error. In the right panel of Fig. 5, we compare $S_c(\Gamma, \omega)$ and $\tilde{S}_{\beta, \delta}(\Gamma, \omega)$ of the 18 site cluster. Within 1 standard deviation, $S_c(\Gamma, \omega)$ and $\tilde{S}_{\beta, \delta}(\Gamma, \omega)$ agree.

E. Temperature evolution of spectra

To examine temperature evolution of the DSF for finite size clusters, we compare with $S(Q, \omega)$ of the antiferromagnetic Kitaev model ($\varphi = 90^\circ$) obtained by available numerical results at the thermodynamic limit. In Fig. 6, the DSFs at two typical momenta ($Q = \Gamma, M$) by FTKω for the 24 site cluster are compared with those at the thermodynamic limit obtained by the Majorana-fermion cluster dynamical mean field theory (cDMFT) with the maximum entropy method (MaxEnt) reported in Ref. [48] and Ref. [73]. Here, the broadening factor $\eta$ is set to 0.02 for the FTKω results.

Even though there is quantitative difference between the spectra of the finite size cluster and the cDMFT, the DSF for the 24 site cluster captures shifts in spectral weight upon cooling down to $k_B T = 0.1$, qualitatively. At the $\Gamma$ point ($Q = 0$), the FTKω shows the shift of the spectral weight from low frequency to high frequency ($\hbar \omega \gtrsim 1$), which is not hampered by the finite-size effect.
and consistent with the cDMFT. At the $M$ point, the formation of the low-energy peak and high-energy shoulder upon cooling is obtained by the 24-site simulation, which is again consistent with the cDMFT. Although there are detailed peak structures due to the finite-size effect for $k_B T \lesssim 0.2$, the present finite-size simulation presumably captures the temperature evolution of the DSFs.

F. Proximity of Kitaev’s spin liquid

The advantages of the present FTKω are its applicability to frustrated systems, which are hardly tractable by the quantum Monte Carlo methods except special limits, and its compatibility with the reweighting techniques that enables us to sweep a range of temperatures with reasonable numerical costs. The antiferromagnetic and ferromagnetic Kitaev models ($\varphi = 90^\circ$ and $270^\circ$, respectively) are the special limits that are tractable by Majorana-fermion quantum Monte Carlo methods\cite{33,34,75}. In the section, we demonstrate the advantages and capability of the FTKω by simulating the DSFs of the Kitaev-Heisenberg model and sweeping a range of temperatures from $k_B T = 10$ to 0.1.

Here, we focus on a proximity of the Kitaev’s spin liquid. As clarified in Ref\cite{60} vicinity to the spin liquid is observable as two peak structures in temperature dependence of heat capacity, even in the magnetically ordered phases. A concrete example is the Kitaev-Heisenberg model at $\varphi = 100^\circ$. For this choice of $\varphi$, temperature dependence of heat capacity has the two peak structure, which is in a close resemblance to that of the Kitaev model, although the ground state has been known to show the zigzag order.

Temperature dependence of heat capacity is informative\cite{24}. However, it is not straightforward to extract an electronic contribution from total heat capacity that may be dominated by the lattice contribution. Alternative approaches to verifying the emergence or proximity of quantum spin liquids are highly desirable. Spectroscopic measurement is one of the promising approaches. Especially, the DSFs have attracted much attention due to recent inelastic neutron scattering measurements on a Kitaev material $\alpha$-RuCl$_3$\cite{34,35,75}.

Before going to the details of the present simulation, we note that finite size effects are plausibly weak at temperatures above $k_B T \sim 0.1$. As shown in the literature\cite{59,60} on the Kitaev-Heisenberg model, there are two temperature scales $T_h$ and $T_l$ at which the temperature dependence of the heat capacity shows local maxima as the function of temperature, in the proximity of the Kitaev’s quantum spin liquid. In Ref\cite{60} it is clarified that the system size dependence of the heat capacity becomes negligible for the 24 site or larger clusters at temperatures well above $T_l$, where $k_B T_l$ is smaller than 0.1 for $\varphi \lesssim 100^\circ$, while, at the Kitaev limit, $\varphi = 90^\circ$, the system size dependence has been shown to be significant around and below the temperature scale $T_l$. Therefore, in the following, we
show the DSFs of the 24 site cluster for \( k_B T \geq 0.1 \) and expect that these results are robust against the finite-size effects.

1. Spectral weight evolution at typical momenta

To capture the proximity of the Kitaev’s spin liquid in the Kitaev-Heisenberg model, we simulate the temperature dependence of the DSFs of the 24 site cluster. First, the temperature evolution of the spectra is examined for \( \varphi = 100^\circ \) at typical momenta \( Q = \Gamma \) and \( M \). Then, the finite-temperature spectra at \( \varphi = 100^\circ \) are compared with the spectra at the Kiteav limit (\( \varphi = 90^\circ \)).

In Fig. 8, the temperature evolution of \( \tilde{S}_{\beta,\delta}(\Gamma, \omega) \) and \( \tilde{S}_{\beta,\delta}(M, \omega) \) for \( \varphi = 100^\circ \) is shown by using the reweighting method. At the \( \Gamma \) point, the spectral weight shifts from \( \hbar \omega \sim 0 \) to \( \hbar \omega \gtrsim 1 \) upon cooling while the low-energy
peak below $\hbar \omega \sim 0.5$ and the high-energy shoulder above $\hbar \omega \sim 1$ develop at the $M$ point at low temperatures below $k_B T \sim 0.2$. These temperature dependences seemingly resemble those of the Kitaev model obtained by the cluster dynamical mean-field theory\cite{22}. However, as detailed below, there is substantial difference between the spectra at $\varphi = 100^\circ$ and that at $\varphi = 90^\circ$. Here, the continuous temperature dependence is obtained by the reweighting method. The filter operators are constructed at $k_B T = +\infty$, 1, 0.5, and 0.2. The reweighting method accurately reproduces the spectra for the temperature ranges $1 < k_B T \leq 10$, $0.5 < k_B T \leq 1$, $0.2 < k_B T \leq 0.5$, and $0.1 < k_B T \leq 0.2$ by starting with the filtered typical pure states at $k_B T = +\infty$, 1, 0.5, and 0.2, respectively.

### 2. Comparison with the Kitaev limit

To contrast the DSF for $\varphi = 100^\circ$ obtained by the FTK$\omega$, we compare that with $S_{\delta}(Q, \omega)$ for $\varphi = 90^\circ$ in Fig. 8. The momentum dependence of the equi-energy slices are shown by changing temperature and frequency. The equi-energy slices are prepared by averaging the spectra within an energy window whose width is 0.1. The momentum dependence is numerically interpolated for visibility without changing the simulation results at the discrete momenta $Q$ compatible with the finite size cluster.

At $k_B T = 0.5$, the DSFs for $\varphi = 90^\circ$ and $\varphi = 100^\circ$ are almost the same, as shown in Figs. 8(a) and (b), respectively. However, below $k_B T = 0.2$, not only the low-energy spectrum at $Q = M$ but also the high-energy spectrum at $Q = \Gamma$ for these two parameters show a stark contrast. The spectral weight at $Q = M$ below $\hbar \omega \sim 0.2$ grows significantly for $\varphi = 100^\circ$. The growth signals development of the zigzag correlations, which is consistent with the temperature dependence of the static spin structure factor\cite{21} at $Q = M$\cite{212}. In addition, the spectral weight at $Q = \Gamma$ for $1 \lesssim \hbar \omega \lesssim 1.5$ grows.

If we recall how small $|J|$ is for $\varphi = 100^\circ$, one may naively wonder why the intensity growth at such high energy region occurs. Indeed, the onset temperature of the intensity growth is more than twice of $|J| = (1/2)|\cos 100^\circ| \sim 0.087$, and the energy scale $1 \lesssim \hbar \omega \lesssim 1.5$ is far beyond the scale of the perturbation $J$.

### 3. Crossover from continuum to damped magnon mode

The intensity growth at $Q = \Gamma$ for $1 \lesssim \hbar \omega \lesssim 1.5$ is quantitatively captured by analyzing width of the broad peak in the spectra. To extract the peak width, we fit the high energy peak by an asymmetric Lorentzian function,

$$s(\omega) = \frac{\gamma \ell}{\gamma \ell + (\hbar \omega - \epsilon \ell)^2} \left\{ a + \frac{b}{1 + e^{(\hbar \omega - \epsilon \ell)}} \right\}, \quad (50)$$

where $\gamma$, $\ell$, $a$, $b$, and $\epsilon \ell$ are fitting parameters. For the fitting, we choose a energy window $1 \lesssim \hbar \omega \lesssim 2$ to exclude a contribution of low energy tails. As shown in Fig. 8(a), the asymmetric Lorentzian function well fit the high energy broad peak for $90^\circ \leq \varphi \leq 100^\circ$. Then, the
dimensionless ratio $\gamma/\epsilon_1$ is a measure of the peak width. In the standard analysis of the magnon spectrum, the dimensionless measure of the peak width is given by the ratio of the full width at half maximum $\Gamma_p$ and the peak energy $E_p$, $\Gamma_p/E_p$. Here, the full width at half maximum $\Gamma_p$ is approximately twice of the imaginary part of the magnon self-energy at $\hbar \omega = E_p$. The present measurement $\gamma/\epsilon_1$ is qualitatively similar to $\Gamma_p/E_p$ although $\epsilon_1 > E_p$ and $\gamma < \Gamma_p$ hold for the fitting function $s(\omega)$. Thus, the ratio $\gamma/\epsilon_1$ gives a lower bound for $\Gamma_p/E_p$.

The dimensionless peak width $\gamma/\epsilon_1$ at the Kitaev limit ($\varphi = 90^\circ$) shows the temperature dependence distinct from those for $\varphi \geq 92.2^\circ$, as shown in Fig. 9(b). In the quantum spin liquid phase, $\gamma/\epsilon_1$ seems to be always larger than 0.3. In contrast, for $\varphi = 95^\circ$ and $100^\circ$, $\gamma/\epsilon_1$ becomes smaller than 0.3 around $k_B T/\epsilon_1 \sim 0.1$, where $\epsilon_1 \sim 1.4$ for $90^\circ \leq \varphi \leq 100^\circ$. The distinct temperature dependence of $\gamma/\epsilon_1$ is coincide with the quantum phase transition at $\varphi \sim 92.2^\circ$ between the Kitaev’s quantum spin liquid phase and the magnetically ordered phase. However, we note that $\gamma/\epsilon_1$ for $\varphi = 100^\circ$ at $k_B T/\epsilon_1 \sim 0.1$ is at least twice larger than the observed upper limit of $\Gamma_p/E_p$ in non-frustrated magnets: The experimental and theoretical studies on the magnon peak width of non-frustrated square-lattice antiferromagnets show $\Gamma_p/E_p \lesssim 0.1$ at the top of the magnon dispersion for $k_B T/E_p \lesssim 0.14$.[66] Thus, the peak width $\gamma/\epsilon_1$ at temperatures around $k_B T/\epsilon_1 \sim 0.1$ is a good measure of frustration.

The peak narrowing at finite temperatures is associated with the quantum phase transition from the spin liquid to the zigzag order. To clarify the relation between the peak narrowing and the quantum phase transition, we summarize the temperature and $\varphi$ dependences of the peak width $\gamma/\epsilon_1$ in a $\varphi$-$T$ phase diagram of the Kitaev-Heisenberg model for $90^\circ \leq \varphi \leq 100^\circ$, which is shown in Fig. 10. The peak width $\gamma/\epsilon_1$ for $k_B T/\epsilon_1 \gtrsim 0.1$ reflects the quantum phase transition at zero temperature. Thus, we attribute the peak narrowing to a finite-temperature crossover from the spin-excitation continuum at the Kitaev’s spin liquid phase ($\varphi = 90^\circ$) to the high-energy damped magnon mode that signal the magnetically ordered ground state.

It has already been revealed in the literature that there are two characteristic temperature scales in the proximity of the Kitaev’s spin liquid phase,[59] as illustrated in Fig. 10. As found in Ref.[59] for the Kitaev model and later for the Kitaev-Heisenberg model,[60] there are two temperature scales $T_h$ and $T_c$ at which temperature dependence of heat capacity shows local maxima as the function of temperature. As clarified for the Kitaev model,[59] nearest-neighbor spin-spin correlations develop upon cooling around the high-temperature scale $T_h$, while the spin gap starts to develop below the low-temperature scale $T_c$. In contrast to the Kitaev limit, in the Kitaev-Heisenberg model, spin-spin correlations start to develop or long-range magnetic orders appear via an order-by-disorder mechanism.[59] at temperatures below the low-temperature scale $T_c$. While the ratio $T_h/T_c$ has been proposed as a measure of distance from the Kitaev’s spin liquid phase in Ref.[60] the peak width $\gamma/\epsilon_1$ at moderately high temperatures far above the low-temperature scale $T_c$ offers another measure of the closeness to the Kitaev’s spin liquid.

We note that the classical Kitaev model shows qualitatively similar dynamics to the quantum counterpart.[61] The semiclassical dynamics of the classical antiferromagnetic Kitaev model reproduces the high-energy continuum of the quantum Kitaev model at $\varphi = 90^\circ$ except the difference in the energy scale due to the different spin amplitude, although development of the spin gap in the quantum Kitaev model signals the breakdown of the similarity between semiclassical and finite-temperature quantum dynamics.[61] When the finite Heisenberg exchange coupling $J/|K| = -0.1$ is introduced in the classical model, a crossover from the a high-energy continuum to a high-energy damped magnon mode is found at $Q = \Gamma$ upon decreasing temperature.[21] across the transition temperature of the order by disorder, which seems to be consistent with the present results for the quantum counterpart. However, here, we note that there is a significant difference between the semiclassical dynamics and quantum dynamics of the Kitaev-Heisenberg model if we associate the transition temperature of the order by disorder in the classical model with the temperature scale $T_c$ in the quantum counterpart: Although the semiclassical dynamics of the Kitaev-Heisenberg model seems to show the continuum as broad as that in the zero-temperature Kitaev limit above the transition temperature (and below $k_B T \sim 1$), the quantum dynamics shows the high-energy excitation peak at $Q = \Gamma$ narrower than that in the Kitaev limit far above the low-temperature scale $T_c$, as shown in Fig. 10. Detailed comparison between the temperature dependences of the classical and quantum dynamics are left for future studies.

The crossover from the spin-excitation continuum to damped magnon mode at high energy is plausibly ubiquitous in the proximity of the Kitaev’s quantum spin liquid phase. The intensity growth and line shape narrowing in high-energy spin excitation spectra are expected to be independent of specific choice of perturbation that drives the Kitaev’s quantum spin liquid to magnetically ordered states. Although the threshold value of $\gamma/\epsilon_1$, which separates the Kitaev’s quantum spin liquid and a frustrated magnet with the magnetically ordered ground state, depends on momenta and the Hamiltonian, the temperature dependence of $\gamma/\epsilon_1$ may offer a common measure of the distance to the Kitaev’s quantum spin liquid.

VI. SUMMARY AND DISCUSSION

In the present paper, we have proposed an $O(N_F)$ algorithm for simulating finite-temperature spectra, called FTdNG, by combining the typical pure state approach and the shifted Krylov subspace method. The
The present algorithm enables us to obtain spectra directly in the frequency domain without the aid of real-time evolution of typical pure states employed in the previous studies.\cite{19, 21, 25} Probability distribution obtained by utilizing the shifted Krylov subspace method, which is essential to the present algorithm, makes possible a use of the reweighting method to finely tune temperature. The reweighting method significantly reduces computational costs to study temperature dependence of the spectra. From the probability distribution at the set of the discrete temperatures, the FTK$\omega$ interpolates potentially exact spectra at temperatures between the two adjacent discrete temperatures with negligible costs.

The present FTK$\omega$ is implemented by using a complete orthonormal basis set of the Fock space in this paper. As a next step, implementation by a compressed basis set is highly desirable to simulate much larger systems. The typical pure state approaches for static observables have already been implemented by variational wave functions.\cite{19, 21, 25} There have also been several studies on the Krylov subspace method by using various variational basis sets such as tensor-network states.\cite{53} The variational bases that are compatible with volume law entanglement will realize the compressed-basis FTK$\omega$.

The capability of the FTK$\omega$ is demonstrated by simulating finite-temperature dynamical spin structure factors of the Kitaev-Heisenberg model. We have found that, even though the absolute value of the ratio of the Heisenberg exchange coupling $J$ and the Kitaev couplings, $|J/K|$, is small for $\varphi = 100^\circ$, temperature dependence of the dynamical spin structure factor shows substantial deviation from that of the Kitaev model not only in the low-energy spectrum at the $M$ point, which signals the onset of the zigzag correlation, but also in the high-energy spectrum at the $\Gamma$ point, even at the temperatures twice larger than $|J|$. The perturbative approaches fail in explaining the deviation. The present exact temperature dependence of dynamical spin structure factors set constraint on approximations, even though the present results are limited for the finite size clusters. At least for $k_B T \gtrsim 0.1$, the present finite-size simulation essentially captures the temperature evolution of the spectral weight that is consistent with that at the thermodynamic limit obtained by the cluster dynamical mean-field theory.\cite{55}

The finite-temperature dynamical spin structure factors shed new light on emergent temperature scales in the proximity of the Kitaev’s quantum spin liquid. As found in Ref.\cite{60} for the Kitaev model and later for the Kitaev-Heisenberg model\cite{60} in the proximity of the Kitaev quantum spin liquid, there are two temperature scales $T_h$ and $T_t$ at which temperature dependence of heat capacity shows peak structures. In the parameter range $90^\circ \leq \varphi \leq 100^\circ$, there are two peak structures in heat capacity of the 24 site cluster as illustrated in Fig. 10. As clarified for the Kitaev model,\cite{58} around $T = T_h$, the nearest-neighbor spin-spin correlations develop while the spin gap starts to develop below $T = T_t$. In contrast to the Kitaev limit, static structure factors for $\varphi = 100^\circ$ grow at the $M$ point around $T = T_t \sim 0.05$. If there is small but finite magnetic anisotropy or a three-dimensional coupling, or if the order-by-disorder mechanism found in the classical Kitaev-Heisenberg model,\cite{70, 80} is relevant to the quantum model, spontaneous time-reversal symmetry breakings will occur below the low-temperature scale. In Fig. 10, the parameter region where we expect the spontaneous symmetry breaking is illustrated as shaded region below $T \sim T_t$ for $\varphi \gtrsim 92.2^\circ$ in the $\varphi-T$ phase diagram. From the measurements on heat capacity and static magnetic orders, these two temperature scales seem to characterize the magnetism in the proximity of the Kitaev’s quantum spin liquid. For example, the ratio $T_t/T_h$ has been proposed as a measure of distance from the Kitaev’s spin liquid phase.\cite{60}

The present results on the dynamical spin structure factors reveal that there is the crossover from the spin-excitation continuum at the Kitaev limit to the damped magnon modes at high energy. The dynamical spin structure factors $S(Q, \omega)$ show significant deviations from those at the Kitaev limit, even at high energy $(1 \lesssim \omega \lesssim 1.5)$. For $\varphi \gtrsim 92.2^\circ$, $S(Q = \Gamma, \omega)$ deviates from that for $\varphi = 90^\circ$ below $T \sim T_h$, which is another precursor of the magnetically ordered ground state, in addition to development of the low-energy spectral weight at the $M$ point due to short-range magnetic correlations. Although a finite-size cluster is employed in the present simulation, the high-energy spectra well above the low-temperature scale $T_t$ are reliable since the finite-size effects become negligible for $N \geq 24$ above the temperature scale $T_h$.\cite{60}

In addition to thermodynamic measurements such as heat capacity,\cite{60} the spectroscopic measurements are found to be useful to measure the distance between a given Kitaev material and the Kitaev limit. The present flexible algorithm is also applicable to other linear responses such as thermal conductivity,\cite{20, 24} which will contribute to an understanding of the proximity of not only the Kitaev’s spin liquid but also other spin liquid candidates.\cite{53, 61, 64}

ACKNOWLEDGMENTS

Y. Y. gratefully thanks Takeo Hoshi and Tomohiro Sugabe for continued collaboration on the shifted Krylov subspace method, which leads him to the present study. Y. Y. also thanks Naoki Kawashima, Seiji Miyashita, and Hans De Raedt for their enlightening discussions and letting him know important references of the typical state approaches. Y. Y. further thanks Masatoshi Imada for carefully reading the manuscript and helpful comments on it. Karen Hallberg for the enlightening discussion on numerical approaches for simulating excitation spectra of correlated electron systems, Synge Todo for his useful comments on statistical treatment in the typical state approaches, Takahiro Misawa for discussions about Ref.\cite{15}
and thank Tsuyoshi Okubo for bringing his attention to the reweighting method. We thank Junki Yoshitake and Yukitoshi Motome for providing us their numerical data on finite-temperature dynamical structure factors of the Kitaev model. In addition, Y. Y. thanks Junki Yoshitake for discussions on the maximum entropy method. Yukitoshi Motome for his comment on the temperature scales of the Kitaev-Heisenberg model. Y. Y. was supported by JSPS KAKENHI (Grant Nos. 15K17702 and 16H06345) and was supported by PRESTO, JST (JPMJPR15NF). This research was supported by MEXT as “Priority Issue on Post-K computer” (Creation of New Functional Devices and High-Performance Materials to Support Next-Generation Industries) and “Exploratory Challenge on Post-K computer” (Frontiers of Basic Science: Challenger (The Limits)). T. S. was supported by JSPS KAKENHI (Grant No. 16K17751). M. K. acknowledges support by Building of Consortia for the Development of Human Resources in Science and Technology from the MEXT of Japan. Our numerical calculation was partly carried out at the Supercomputer Center, Institute for Solid State Physics, University of Tokyo. The exact diagonalization (ED) calculations are partly double-checked by using an open-source ED program package $\text{H$q^2$D}$[88,89]. The excitation spectra are calculated by employing a numerical library $K\omega$ for shifted Krylov subspace methods[13].

Appendix A: Imaginary-time evolution

In the present paper, we calculate the typical pure state (or canonical thermal pure quantum state[13]) at inverse temperature $\beta$ by following Ref[14] as

$$e^{N\beta \hat{H}/2}|\psi_\beta\rangle = e^{N\beta (\hat{\mathcal{H}} - \hat{\mathcal{H}}^2)/2}|\psi_0\rangle = \sum_{k=0}^{\infty} \frac{(N\beta/2)^k}{k!} (\hat{\mathcal{H}} - \hat{\mathcal{H}})^k |\psi_0\rangle,$$  \hfill (A1)

where $\hat{\mathcal{H}} = \hat{\mathcal{H}}/N$ is used. The above formula is not suitable for the numerical simulation, since terms in the rightmost hand side of Eq. (A1) become too large for $\beta \gg 1$ and will introduce cancellation of significant digits.

To avoid the cancellation of significant digits, we split the imaginary time evolution and divide it into $M_d$ steps. The step size $\beta/M_d$ is determined by the following estimation. First, we estimate the amplitude of the largest order term is bounded by

$$\left\| \frac{(N\beta/2)^k}{k!} (\hat{\mathcal{H}} - \hat{\mathcal{H}})^k |\psi_0\rangle \right\| \leq \frac{(N\beta/2)^k (\ell + |\epsilon_0|)^k}{k!} \sim \frac{(N\beta/2)^k (\ell + |\epsilon_0|)^k}{\sqrt{2\pi k}(k/e)^k}. \hfill (A2)$$

Then, by differentiating the term in the rightmost side of Eq. (A2) with respect to $k$, we find that, when $k = X - 1/2 + \mathcal{O}(1/X)$, where $X = (N\beta/2M_d)(\ell + |\epsilon_0|)$, the $k$ th term becomes maximum among the series expansion. For $k = X - 1/2 + \mathcal{O}(1/X)$, we obtain the asymptotic formula for the extremum as

$$\max_k \left\{ \frac{(N\beta/2M_d)^k (\ell + |\epsilon_0|)^k}{\sqrt{2\pi k}(k/e)^k} \right\} \sim \exp(X). \hfill (A3)$$

When we set an upper limit $\Lambda$ for $e^X/(2\pi X)^{-1/2}$, we can determine an appropriate $M_d$ through iteratively solving $e^X/(2\pi X)^{-1/2} = \Lambda$. \hfill (A4)

By using the following recurrence relation for $k \geq 0$ initialized with $X_0 = \Lambda$,

$$X_{k+1} = \ln \Lambda + \frac{1}{2} \ln(2\pi X_k), \hfill (A5)$$

we obtain $M_d$ as

$$M_d = \frac{N\beta/2}{\lim_{k\to+\infty} X_k} (\ell + |\epsilon_0|). \hfill (A6)$$

Appendix B: Upper bounds of variance

The source of the deviation between $\mathcal{G}_{\beta}^{AB}(\zeta)$ and $\tilde{\mathcal{G}}_{\beta,\delta}^{AB}(\zeta)$ is twofold: The discretization parameters $\delta = (E_b, \epsilon, M)$ and variance of $\{c_n\}$. The former source can be examined by changing the set of the discretization parameters $\delta$. Therefore, we here focus on the deviation originating from the variance of stochastic variables $\{c_n\}$ and give the upper bounds of the variance of the present $\mathcal{O}(N_F)$ algorithm by following Refs[6] and [14].

We start with rewriting Eq. (35) as

$$\tilde{\mathcal{G}}_{\beta,\delta}^{AB}(\zeta) = \sum_{n} c_n^2 e^{\beta E_n} e^{\epsilon(E_n - E_m)} \mathcal{E}_m^{E_n}(E|\hat{O}_m(\zeta)|E')$$

$$\leq \sum_{n} c_n^2 e^{\beta E_n} e^{\epsilon(E_n - E_m)} \mathcal{E}_m^{E_n}(E|\hat{O}_m(\zeta)|E'), \hfill (B1)$$

where we define a equi-energy shell as $\Delta_m = [\mathcal{E}_m - \epsilon, \mathcal{E}_m + \epsilon]$ and an operator as

$$\hat{O}_m(\zeta) = \hat{\mathcal{F}}(\zeta - \hat{H} + \mathcal{E}_m)^{-1} \hat{\mathcal{B}}. \hfill (B2)$$

For later usage, we introduce the following shorthand for expectation values as

$$\bar{f}_m(\zeta) = e^{\beta E_m} \sum_{E \in \Delta_m} \langle E|\hat{O}_m(\zeta)|E \rangle, \hfill (B3)$$

and

$$\bar{g} = Z(\beta), \hfill (B4)$$

and for stochastic variables as

$$\bar{f}_m(\zeta) + \delta f_m(\zeta) = e^{\beta E_m} N_F \sum_{E, E' \in \Delta_m} c_E c_{E'} \langle E|\hat{O}_m(\zeta)|E' \rangle, \hfill (B5)$$
and
\[ g + \delta g = N_F \sum_n |c_n|^2 e^{-\beta E_n}, \]
(B6)

where \( \tilde{G}_{AB}^{\beta,\delta}(\zeta) = \sum_m (\mathcal{J}_m(\zeta) + \delta f_m(\zeta))/[g + \delta g] \).

Then, the variance of \( \tilde{G}_{AB}^{\beta,\delta}(\zeta) \) is given by
\[ \sigma^2(\zeta) = E \left[ \left( \frac{\mathcal{J}_m(\zeta) + \delta f_m(\zeta)}{g + \delta g} - \frac{\mathcal{J}_m(\zeta)}{g} \right)^2 \right] \]
\[ \simeq \sum_{m,m'} \frac{1}{g^2} E[\delta f_m(\zeta)\delta f_{m'}(\zeta)] \]
\[ + \sum_{m,m'} \frac{\mathcal{J}_m(\zeta)\mathcal{J}_{m'}(\zeta)}{g^3} E[\delta g^2] \]
\[ - 2\text{Re} \sum_{m,m'} \frac{\mathcal{J}_m(\zeta)^*}{g^2} E[\delta f_m(\zeta)\delta g]. \]
(B7)

Below, we evaluate \( \sigma^2(\zeta) \) term by term: By using the following formulae:
\[ E[|c_n|^2] = 1/N_F, \]
(88)
\[ E[|c_k|^2|c_l|^2] = 1/N_F/(N_F + 1), \]
(89)
\[ E[|c_n|^4] = 2/N_F/(N_F + 1), \]
(B10)

we obtain the following expectation values,
\[ E[\delta f_m(\zeta)\delta f_{m'}(\zeta)] = e^{-\beta(\varepsilon_m+\varepsilon_{m'})}/N_F + 1 \]
\[ \times \delta_{m,m'} \sum_{E,E'} \langle E|\hat{O}_m(\zeta)^\dagger|E'\rangle\langle E'|\hat{O}_m(\zeta)|E \rangle \]
\[ - \left( \sum_{E\in\Delta_m} \langle E|\hat{O}_m(\zeta)^\dagger|E \rangle \right) \]
\[ \times \left( \sum_{E'\in\Delta_{m'}} \langle E'|\hat{O}_{m'}(\zeta)|E' \rangle \right), \]
(B11)

\[ E[\delta g^2] = \frac{N_F}{N_F + 1} \left[ Z(\beta)^2 - \frac{Z(\beta)^2}{N_F} \right], \]
(B12)

and
\[ E[\delta f_m(\zeta)\delta g] = e^{-\beta\varepsilon_m} N_F e^{-\beta\varepsilon_m} - Z(\beta)/N_F + 1 \]
\[ \sum_{E\in\Delta_m} \langle E|\hat{O}_m(\zeta)^\dagger|E \rangle. \]
(B13)

After straightforward calculations, we reach the following expression,
\[ \sigma^2(\zeta) \simeq \frac{N_F}{N_F + 1} \frac{Z(\beta)^2}{Z(\beta)^2} [A_{2a}(\zeta) + |G_{AB}^{\beta}(\zeta)|^2 - 2\text{Re}\{G_{AB}^{\beta}(\zeta)G_{AB}^{\beta}(\zeta)^*\}]. \]
(B14)

where we introduce the following shorthand,
\[ A_{\beta}(\zeta) = \sum_m e^{-\beta\varepsilon_m} \sum_{E,E'} \langle E|\hat{O}_m(\zeta)^\dagger|E'\rangle\langle E'|\hat{O}_m(\zeta)|E \rangle. \]
(B15)

Generally, estimate of \( A_{\beta}(\zeta) \) in Eq. (B14) is not tractable. Below, we give an upper bound of \( A_{\beta}(\zeta) \). First, we use the following inequality: There is a positive constant \( \min\{\eta, \epsilon\} < \bar{\eta} < \mathcal{O}(NJ_0) \) that satisfies the inequality
\[ A_{\beta}(\hbar\omega + i\eta) < \frac{1}{\bar{\eta}} \int_{-\infty}^{+\infty} d\omega' A_{\beta}(\hbar\omega' + i\eta). \]
(B16)

Then, we estimate upper bounds of the following integral as
\[ \int_{-\infty}^{+\infty} \hbar\omega \sum_{E,E'} \langle E|\hat{O}_m(\hbar\omega + i\eta)^\dagger|E'\rangle\langle E'|\hat{O}_m(\hbar\omega + i\eta)|E \rangle \]
\[ = \frac{4\pi\hbar\eta}{E - E'} \sum_{E\in\Delta_m} \sum_{k,t} \langle E|\hat{B}_t^+|k\rangle\langle k|\hat{A}_t|E'\rangle\langle E'|\hat{A}_t^\dagger|k\rangle\langle k|\hat{B}_t|E \rangle \]
\[ < \frac{\pi\hbar}{\eta} \sum_{E\in\Delta_m} \int_{-\infty}^{+\infty} dt e^{-2\eta|t|/\hbar} \langle E|\hat{B}_t^+\hat{A}(t)^\dagger|E'\rangle\langle E'|\hat{A}(t)^\dagger\hat{B}_t|E \rangle, \]
(B17)

where we use the following transformation,
\[ \langle E|\hat{B}_t^+\hat{A}(t)^\dagger|E'\rangle \]
\[ = \langle E|\hat{B}_t^+ e^{i\hat{H}_t/\hbar} \hat{A} e^{-i\hat{H}_t/\hbar}|E'\rangle \]
\[ = \sum_k e^{-i\varepsilon_m t/\hbar + i\varepsilon_k t/\hbar} \langle E|\hat{B}_t^+|k\rangle\langle k|\hat{A}|E'\rangle. \]
(B18)

The integral in the last line of Eq. (B17) is bounded: There is a positive constant \( \tau \) that satisfies
\[ \int_{-\infty}^{+\infty} dt e^{-2\eta|t|/\hbar} \langle E|\hat{B}_t^+\hat{A}(t)^\dagger\hat{B}_t|E \rangle \]
\[ < \langle E|\hat{B}_t^+ \hat{A}\hat{A}^\dagger \hat{B}_t|E \rangle \int_{-\infty}^{+\infty} dt e^{-2(\eta/\hbar + 2/\tau)|t|} \]
\[ = \langle E|\hat{B}_t^+ \hat{A}\hat{A}^\dagger \hat{B}_t|E \rangle. \]
(B19)

The positive constant \( \tau \) simply corresponds to a correlation time that characterizes the correlation function \( \langle E|\hat{B}_t^+\hat{A}(t)^\dagger\hat{B}_t|E \rangle \).
Finally, we obtain an upper bounds for $\sigma^2(\omega + i\eta)$ as

$$
\sigma^2(h\omega + i\eta) 
\leq \frac{Z(2\beta)}{Z(\beta)^2} \left[ \pi \sum_{n} e^{-\beta e_m} \frac{\langle e|B^1\hat{A}\hat{A}^1|B\rangle E}{\eta + h/\tau} \right]
\leq +|\hat{g}^{AB}_\beta(h\omega + i\eta)|^2 - 2Re\{\hat{g}^{AB}_\beta(h\omega + i\eta)\hat{g}^{AB}_\beta(h\omega + i\eta)\}
\leq \frac{Z(2\beta)}{Z(\beta)^2} \left[ \pi \frac{\langle B^1\hat{A}\hat{A}^1|B\rangle^{\text{ens}}}{\eta + h/\tau} \right]
\leq +|\hat{g}^{AB}_\beta(h\omega + i\eta) - \hat{g}^{AB}_\beta(h\omega + i\eta)|^2.
\tag{B20}
$$

The factor $Z(2\beta)/Z(\beta)^2$ is known to be exponentially small, when the system size $N$ grows [13]. The temperature dependence of $Z(2\beta)/Z(\beta)^2$ is trivially grasped as follows. First, we take the simple limits $\beta \to +0$ and $\beta \to +\infty$ as

$$
\lim_{\beta \to +0} \frac{Z(2\beta)}{Z(\beta)^2} = \frac{1}{N_F},
\tag{B21}
$$

and

$$
\lim_{\beta \to +\infty} \frac{Z(2\beta)}{Z(\beta)^2} = \frac{1}{D},
\tag{B22}
$$

where $D$ is the degeneracy of the ground state. By taking temperature derivative of $Z(2\beta)/Z(\beta)^2$ as,

$$
\frac{\partial}{\partial T} \left( \frac{Z(2\beta)}{Z(\beta)^2} \right) = 2\beta^2 \frac{Z(2\beta)}{Z(\beta)^2} \left[ \langle E^{\text{ens}} \rangle - \langle E \rangle^2 \right],
\tag{B23}
$$

and remembering that $\langle E \rangle^\beta$ is a monotonically decreasing function of temperature, we can prove the following relation,

$$
\frac{1}{N_F} \leq Z(2\beta)/Z(\beta)^2 \leq \frac{1}{D}.
\tag{B24}
$$

If we use standard relations for free energy $F(\beta)$ and entropy $S(\beta)$, $F(\beta) = -\beta^{-1} \ln Z(\beta)$ and $dF(\beta)/dT = -S(\beta)$, we obtain the following relation: There is an inverse temperature that satisfies $\beta \leq \beta^* \leq 2\beta$ and

$$
Z(2\beta)/Z(\beta)^2 = e^{-2\beta[F(2\beta) - F(\beta)]} = e^{-S(\beta^*)}.
\tag{B25}
$$

At finite temperature, $S(\beta)$ is finite and extensive. Since the entropy is proportional to $N$, the factor $Z(2\beta)/Z(\beta)^2$ is exponentially small.

**Appendix C: Convergence of shifted BiCG method and dependence on discretization**

The shifted BiCG method is employed in the present FTK$\omega$ algorithm to implement the multiplication of $(\zeta - \hat{H})^{-1}$ and the filter operator. The convergence of the CG methods is verified by the 2-norm of the residual vectors. By setting the upper bound on the 2-norm

$$
\|\rho_n(\zeta)\|_2 = \sqrt{\langle \rho_n(\zeta)|\rho_n(\zeta) \rangle},
$$

can we truncate the CG steps in a controlled fashion. In the shifted Krylov subspace method that handles a set of shifts or complex numbers $\{\zeta\}$, we need to choose a residual vector for the truncation. A choice that guarantees the quality of the convergence is the residual vector with the largest 2-norm, which is realized by the seed switching method [400].

In this section, we examine the CG-step dependence of the 2-norm in the present application. We choose examples from the calculations for the 18 site and 24 site clusters of the antiferromagnetic Kitaev model ($\varphi = 90^\circ$) at $k_B T = 0.1$. The examples are chosen from the construction of the filtered typical states defined in Eq.(33) since the construction is the most time-consuming part of the present FTK$\omega$ algorithm.

The shifted linear equation $|\chi(\zeta)\rangle = (\zeta - \hat{H})|\phi_\beta\rangle$ is solved for the set of shifts $\{\zeta\}$ to construct the $L$ filtered typical states. The shifted BiCG method is applied to each filter operator separately in the present implementation and the contour integral in every single filter operator is approximated by the Riemann sum with the $M$ discrete points.

In Fig.11 we show typical examples of the CG-step dependence of the maximum 2-norm in the constructions of the filter operators. For each filter operator (each $m \in [0, L]$), at each CG step, the maximum 2-norm $\max_{\zeta}\{||\rho_n(\zeta)||_2\}$ is chosen from the set of the $M$ discrete points $\zeta \in \{\epsilon \omega^m + \epsilon m \}_{1\leq m < L}$ along the contour illustrated in Fig.1. The upper bound of the 2-norm is set to $10^{-4}$ or smaller for the 18 site and 24 site clusters, which
guarantees convergence of expectation values taken by the filtered typical pure states. For the 18 site cluster, we also show how the CG-step dependence of the 2-norm depends on the discretization δ = (E_L, ϵ, M).

The CG-step dependence is shown for E_m at which the probability distribution P_δ,δ becomes maximum in Fig. 1(a). When we choose E_m that requires the largest number of the CG steps, the speed of the convergence also depends on \min_{|\zeta|} \{ |\text{Im} \zeta| \} for \zeta \in \{ ϵ_\text{e}^{iθ} + E_m \}j. As shown in Fig. 1(b), larger L and M, which decrease \min_{|\zeta|} \{ |\text{Im} \zeta| \}, may require more CG steps. When both L and M are doubled, which give four times smaller \min_{|\zeta|} \{ |\text{Im} \zeta| \}, the number of the CG steps required for the convergence increases by around 10 percent.

Here, we note that the number of the required CG steps depends on the density of states at E_m while we only show the typical examples in Fig. 1(a). This dependence is inferred from the convergence theorem of the Lanczos method. The number of the Lanczos steps required to obtain an eigenstate and eigenvalue becomes larger as the density of states at the target eigenvalue becomes larger. At k_BT = 0.1, as shown in Fig. 2, the peak of the probability distribution is located nearby the lower edge of the eigenvalue distribution. Both the 18 site and 24 site clusters show faster convergences for E_m close to the edge of the probability distribution compared with the other choice of E_m. The sparse density of states nearby the lowest eigenvalue naturally explains the faster convergence.

The system size dependence of the convergence is of practical importance. As evident in Fig. 1(b), the system size affects the CG-step dependence of the 2-norm. However, when we lower the upper bound of the 2-norm to ensure the exponential decay of the CG-step dependence of the 2-norm, we observe that ten thousand CG steps are practically enough to obtain the convergence for any E_m, irrespective of the system size. Only nearby the edge of the probability distribution, the exponential decay is sensitive to the discretization and the system size.

Then, we examine how the discretization δ affects the spectra for the fixed broadening factor ϵ = 0.2. As formulated in Eq. (26), the FTKω exactly reproduces the finite-temperature spectra by the canonical ensemble average after taking the average over the initial random vectors and the two limits, M → +∞ and ϵ → +0. The large M limit should be taken before the small ϵ or large L limit. Here, we note that the interval of the discrete energy grid, ϵ, is set to be comparable to or smaller than the broadening factor δ. To find a reasonable choice of δ, we examine the δ dependence of the dynamical spin structure factor for the 12 site and 18 site clusters of the antiferromagnetic Kiteav model (ϕ = 90°) at the Γ point. First, we examine the convergence when M is increased. For L = 200 and 400, the M dependence of the spectrum is examined in Figs. 12(a) and (b) when ϵL is fixed. The spectrum calculated with larger L requires larger M to converge. Second, we examine the L dependence of the spectrum with an appropriate M and keeping ϵL constant. As shown in Fig. 12(c), we obtain a converged result for the 12 site cluster for ϕ = 90° by increasing L and choosing an appropriate M. The system size also affects the convergence. When the system size is increased from N = 12 to 18, the L dependence becomes smaller as shown in Fig. 12(d).

The discretization parameters used in Sec. IV are summarized in Table 1. Here, we choose the discretization parameters to obtain converged results for the 12 site and 18 site clusters. For the 24 site cluster, to take a balance of accuracy and a numerical cost, a practical parameter set is chosen based on the L and M dependence of the dynamical spin structure factors for N = 12 and 18. As shown in Fig. 12(d), the (L, M) dependence is already small for N = 18, at least, for L ≥ 100. Thus, we choose L = 128 and confirm that M = 16 is enough to obtain reasonable results.

---

1 W. Kohn, “Nobel lecture: Electronic structure of matter–Wave functions and density functionals,” Rev. Mod. Phys. 71, 1253–1266 (1999).
2 M. Imada and M. Takahashi, “Quantum transfer Monte Carlo method for finite temperature properties and quantum molecular dynamics method for dynamical correlation functions,” J. Phys. Soc. Jpn. 55, 3354 (1986).
3 John Skilling, “Maximum entropy and bayesian methods: Cambridge, england, 1988,” (Springer Science & Business Media, 2013) p. 455.
4 Pedro de Vries and Hans De Raedt, “Solution of the time-dependent Schrödinger equation for two-dimensional spin-1/2 Heisenberg systems,” Phys. Rev. B 47, 7929–7937 (1993).
5 J. Jaklić and P. Prelovšek, “Lanczos method for the calculation of finite-temperature quantities in correlated systems,” Phys. Rev. B 49, 5065–5068 (1994).
6 Anthony Hams and Hans De Raedt, “Fast algorithm for finding the eigenvalue distribution of very large matrices,” Phys. Rev. E 62, 4365–4377 (2000).
7 The studies may recall the maximum-entropy approach and the forced oscillator method as predecessors of them.
8 Hal Tasaki, “From quantum dynamics to the canonical distribution: General picture and a rigorous example,” Phys. Rev. Lett. 80, 1373–1376 (1998).
9 Sandu Popescu, Anthony J Short, and Andreas Winter, “Entanglement and the foundations of statistical mechanics,” Nature Physics 2, 754–758 (2006).
10 Sheldon Goldstein, Joel L. Lebowitz, Roderich Tumulka, and Nino Zanghí, “Canonical typicality,” Phys. Rev. Lett. 96, 050403 (2006).
11 A Sugita, “On the basis of quantum statistical mechanics.”
FIG. 11. (color online): CG-step dependence of the maximum 2-norm of the residual vectors for the 18 site and 24 site clusters. (a) The CG-step dependence is shown for $E_m$ at which the probability distribution $\tilde{\mathcal{P}}_{\beta, \delta}$ becomes maximum. (b) The CG-step dependence for $E_m$ that requires the largest number of the CG steps.

Nonl. Phen. Compl. Sys. 10, 192–195 (2007).

12 Peter Reimann, “Typicality for generalized microcanonical ensembles,” Phys. Rev. Lett. 99, 160404 (2007).

13 Sho Sugiura and Akira Shimizu, “Thermal pure quantum states at finite temperature,” Phys. Rev. Lett. 108, 240401 (2012).

14 Sho Sugiura and Akira Shimizu, “Canonical thermal pure quantum state,” Phys. Rev. Lett. 111, 010401 (2013).

15 Y. Takahashi and H. Umezawa, “Thermo field dynamics,” Collect. Phenom. 2, 55 (1975).

16 Toshiaki Iita and Toshihazu Ebisuzaki, “Algorithm for linear response functions at finite temperatures: Application to ESR spectrum of $S = \frac{1}{2}$ antiferromagnet Cu benzoate,” Phys. Rev. Lett. 90, 047203 (2003).

17 Christian Bartsch and Jochen Gemmer, “Dynamical typicality of quantum expectation values,” Phys. Rev. Lett. 102, 110403 (2009).

18 Tarek A. Elsayed and Boris V. Fine, “Regression relation for pure quantum states and its implications for efficient computing,” Phys. Rev. Lett. 110, 070404 (2013).

19 Robin Steinigeweg, Jochen Gemmer, and Wolfram Brenig, “Spin-current autocorrelations from single pure-state propagation,” Phys. Rev. Lett. 112, 120601 (2014).

20 R. Steinigeweg, A. Khodja, H. Niemeyer, C. Gogolin, and J. Gemmer, “Pushing the limits of the eigenstate thermalization hypothesis towards mesoscopic quantum systems,” Phys. Rev. Lett. 112, 130403 (2014).

21 Takaaki Monnai and Ayumu Sugita, “Typical pure states and nonequilibrium processes in quantum many-body systems,” Phys. Rev. Lett. 112, 130403 (2014).
tems,” J. Phys. Soc. Jpn. 83, 094001 (2014).
22 C. Karrasch, D. M. Kennes, and J. E. Moore, “Transport properties of the one-dimensional Hubbard model at finite temperature,” Phys. Rev. B 90, 155104 (2014).
23 F. Jin, R. Steinigeweg, F. Heidrich-Meisner, K. Michielsen, and H. De Raedt, “Finite-temperature charge transport in the one-dimensional Hubbard model,” Phys. Rev. B 92, 205103 (2015).
24 M. W. Long, P. Prelovšek, S. El Shawish, J. Karadamoglou, and X. Zotos, “Finite-temperature dynamical correlations using the microcanonical ensemble and the Lanczos algorithm,” Phys. Rev. B 68, 235106 (2003).
25 X. Zotos, “High temperature thermal conductivity of two-leg spin-1/2 ladders,” Phys. Rev. Lett. 92, 067202 (2004).
26 Ryogo Kubo and Kazuhsisa Tomita, “A general theory of magnetic resonance absorption,” J. Phys. Soc. Jpn. 9, 888–919 (1954).
27 Hazime Mori and Kyozi Kawasaki, “Theory of dynamical behaviors of ferromagnetic spins,” Progress of Theoretical Physics 27, 529 (1962).
28 R. E. Dietz, F. R. Merritt, R. Dingle, Daniel Hone, B. G. Silbernagel, and Peter M. Richards, “Exchange narrowing in one-dimensional systems,” Phys. Rev. Lett. 26, 1186–1188 (1971).
29 Masaki Oshikawa and Ian Affleck, “Electron spin resonance in $S = \frac{1}{2}$ antiferromagnetic chains,” Phys. Rev. B 65, 134410 (2002).
30 Manabu Machida, Toshihiko Itaka, and Seiji Miyashita, “ESR intensity and the Dzyaloshinsky-Moriya interaction of the nanoscale molecular magnet $V_{15}$,” Phys. Rev. B 86, 224412 (2012).
31 S. El Shawish, O. Cépas, and S. Miyashita, “Electron spin resonance in $S = \frac{1}{2}$ antiferromagnets at high temperature,” Phys. Rev. B 81, 224421 (2010).
32 As reviews, see Refs. 35 and 36.
33 Luke J. Sandilands, Yao Tian, Kemp W. Plumb, Young-June Kim, and Kenneth S. Burch, “Scattering continuum and possible fractionalized excitations in $\alpha$-RuCl$_3$,” Phys. Rev. Lett. 114, 147201 (2015).
34 A Banerjee, CA Bridges, J-Q Yan, AA Aczel, L Li, MB Stone, GE Granroth, MD Lumsden, Y Yiu, J Knolle, et al., “Proximate Kitaev quantum spin liquid behaviour in a honeycomb magnet,” Nature materials 15, 733–740 (2016).
35 Arnab Banerjee, Jiaqiang Yan, Johannes Knolle, Craig A. Bridges, Matthew B. Stone, Mark D. Lumsden, David G. Mandrus, David A. Tennant, Roderich Moessner, and Stephen E. Nagler, “Neutron scattering in the proximate quantum spin liquid $\alpha$-RuCl$_3$,” Science 356, 1055–1059 (2017).
36 A. Kitaev, “Anyons in an exactly solved model and beyond,” Annals Phys. 321, 2 (2006).
37 Saptarshi Mandal and Naveen Surendran, “Exactly solvable Kitaev model in three dimensions,” Phys. Rev. B 79, 024426 (2009).
38 As a review on quantum spin liquids, see Ref. 37.
39 G. Jackeli and G. Khaliullin, “Mott insulators in the strong spin-orbit coupling limit: From Heisenberg to a quantum compass and Kitaev models,” Phys. Rev. Lett. 102, 017205 (2009).
40 Jiri Chaloupka, George Jackeli, and Giniyat Khaliullin, “Kitaev-Heisenberg model on a honeycomb lattice: Possible exotic phases in iridium oxides $A_2$IrO$_3$,” Phys. Rev. Lett. 105, 027204 (2010).
41 Atsuo Shitade, Hosho Katsura, Jan Kuneš, Xiao-Liang Qi, Shou-Cheng Zhang, and Naoto Nagaosa, “Quantum spin Hall effect in a transition metal oxide Na$_2$IrO$_3$,” Phys. Rev. Lett. 102, 256403 (2009).
42 Yugesh Singh and P. Gegenwart, “Antiferromagnetic mott insulating state in single crystals of the honeycomb lattice material Na$_2$IrO$_3$,” Phys. Rev. B 82, 064412 (2010).
43 T. Takayama, A. Kato, R. Dinnebier, J. Nuss, H. Kono, L. S. I. Veiga, G. Fabbri, D. Haskel, and H. Takagi, “Hyperf honeycomb iridate $\beta$-Li$_2$IrO$_3$ as a platform for Kitaev magnetism,” Phys. Rev. Lett. 114, 077202 (2015).
44 Kimberly A. Modic, Tess E. Smidt, Itamar Kimchi, Nicholas P. Breznay, Alun Biffin, Sunghyun Choi, Roger D. Johnson, Radu Coldea, Pilanda Watkins-Curry, Gregory T. McCandless, Julia Y. Chan, Felipe Gandara, Z. Islam, Ashvin Vishwanath, Arkady Shekhter, Ross D. McDonald, and James G. Analytis, “Realization of a three-dimensional spin-anisotropic harmonic honeycomb iridate,” Nature commun. 5, 4203 (2014).
45 J. Knolle, D. L. Kovrizhin, J. T. Chalker, and R. Moessner, “Dynamics of a two-dimensional quantum spin liquid: Signatures of emergent Majorana fermions and fluxes,” Phys. Rev. Lett. 112, 207203 (2014).
46 J. Knolle, Gia-Wei Chen, D. L. Kovrizhin, R. Moessner, and N. B. Perkins, “Raman scattering signatures of Kitaev spin liquids in $A_2$IrO$_3$ iridates with $A$=Na or Li,” Phys. Rev. Lett. 113, 187201 (2014).
47 J Nasu, J Knolle, DL Kovrizhin, Y Motome, and Roderich Moessner, “Fermionic response from fractionalization in an insulating two-dimensional magnet,” Nature Physics 12, 912–915 (2016).
48 Junki Yoshitake, Joji Nasu, and Yukitoshi Motome, “Fractional spin fluctuations as a precursor of quantum spin liquids: Majorana dynamical mean-field study for the Kitaev model,” Phys. Rev. Lett. 117, 157203 (2016).
49 M. Herrmanns and S. Trebst, “Quantum spin liquid with a Majorana fermi surface on the three-dimensional hypercognaton lattice,” Phys. Rev. B 89, 235102 (2014).
50 Jiri Chaloupka, George Jackeli, and Giniyat Khaliullin, “Zigzag magnetic order in the iridium oxide Na$_2$IrO$_3$,”Phys. Rev. Lett. 110, 097204 (2013).
51 V. K. Katukuri, S. Nishimoto, V. Yushkhankhi, A. Stoyanova, H. Kandpal, C. Sungkyun, R. Coldea, I. Rouschatzakis, L. Hozoi, and J. van den Brink, “Kitaev interactions between $j=1/2$ moments in honeycomb Na$_2$IrO$_3$ are large and ferromagnetic: insights from ab initio quantum chemistry calculations,” New J. Phys. 16, 013056 (2014).
52 Jeffrey G. Rau, Eric Kin-Ho Lee, and Hae-Young Kee, “Generic spin model for the honeycomb iridates beyond the Kitaev limit,” Phys. Rev. Lett. 112, 077204 (2014).
53 Youhei Yamaji, Yusuke Nomura, Moyuru Kurita, Ryotaro Arita, and Masatoshi Imada, “First-principles study of the honeycomb-lattice iridates Na$_2$IrO$_3$ in the presence of strong spin-orbit interaction and electron correlations,” Phys. Rev. Lett. 113, 107201 (2014).
54 Andreas Frommer, “BiCGstab ($\ell$) for families of shifted linear systems,” Computing 70, 87–109 (2003).
55 The construction of the microcanonical shell may remind the readers of the microcanonical thermal pure quantum (TPQ) state proposed in Ref. 13. To avoid possible confusion, we note that the microcanonical TPQ state does not correspond to a microcanonical shell. As proven in Ref. 13.
the microcanonical thermal pure quantum state reproduces probability distribution of the canonical ensemble.

Alan M. Ferrenberg and D. P. Landau, “Critical behavior of the three-dimensional Ising model: A high-resolution Monte Carlo study,” Phys. Rev. B 44, 5081–5091 (1991).

Hong-Chen Jiang, Zheng-Cheng Gu, Xiao-Liang Qi, and Simon Trebst, “Possible proximity of the mott insulating iridate Na$_2$IrO$_3$ to a topological phase: Phase diagram of the heisenberg-kitaev model in a magnetic field,” Phys. Rev. B 83, 245104 (2011).

Juan Osorio Iregui, Philippe Corboz, and Matthias Troyer, “Probing the stability of the spin-liquid phases in the Kitaev-Heisenberg model using tensor network algorithms,” Phys. Rev. B 90, 195102 (2014).

Joji Nasu, Masafumi Udagawa, and Yukitoshi Motome, “Thermal fractionalization of quantum spins in a kitaev model: Temperature-linear specific heat and coherent transport of majorana fermions,” Phys. Rev. B 92, 115122 (2015).

Youhei Yamaji, Takafumi Suzuki, Takuto Yamada, Seiichiro Suga, Naoki Kawashima, and Masatoshi Imada, “Chaos and criteria for designing a kitaev spin liquid revealed by thermal and spin excitations of the honeycomb iridate Na$_2$IrO$_3$,” Phys. Rev. B 93, 174425 (2016).

Nazakat Ullah, “Invariance hypothesis and higher correlations of hamiltonian matrix elements,” Nuclear Physics B 85, 65 – 71 (1964).

S. Miyashita and H. De Raedt, private communication.

E. R. Gagliano and C. A. Balseiro, “Dynamical properties of quantum many-body systems at zero temperature,” Phys. Rev. Lett. 59, 2999–3002 (1987).

A numerical library for the shifted Krylov methods, $K_a$, developed by M. Kawamura based on Refs.95 and 96 is available through https://github.com/issp-center-dev/Komega.

Tosio Kato, “On the convergence of the perturbation method, i,” Progress of Theoretical Physics 4, 514 (1949).

Tetsuya Sakurai and Hiroshi Sugiiura, “A projection method for generalized eigenvalue problems using numerical integration,” Journal of Computational and Applied Mathematics 159, 119–128 (2003).

Tsunomu Iegami, Tetsuya Sakurai, and Umpei Nagashima, “A filter diagonalization for generalized eigenvalue problems based on the Sakurai–Sugiura projection method,” Journal of Computational and Applied Mathematics 233, 1927–1936 (2010).

Noritaka Shimizu, Yutaka Utsuno, Yasunori Futamura, Tetsuya Sakurai, Takahiro Mizusaki, and Takaharu Otsuka, “Stochastic estimation of nuclear level density in the nuclear shell model: An application to parity-dependent level density in 58Ni,” Physics Letters B 753, 13 – 17 (2016).

M. Falcioni, E. Marinari, M.L. Paciello, G. Parisi, and B. Taglienti, “Complex zeros in the partition function of the four-dimensional SU(2) lattice gauge model,” Physics Letters B 108, 331 – 332 (1982).

Enzo Marinari, “Complex zeroes of the d = 3 Ising model: Finite-size scaling and critical amplitudes,” Nuclear Physics B 235, 123 – 134 (1984).

P Prelovšek and J Bonča, “Ground state and finite temperature Lanczos methods,” in Strongly Correlated Systems (Springer, 2013) pp. 1–30.

Robin Steinigeweg, Jacek Herbrych, Xenophon Zotos, and Wolfram Breig, “Heat conductivity of the heisenberg spin-1/2 ladder: From weak to strong breaking of integrability,” Phys. Rev. Lett. 116, 017202 (2016).

Junki Yoshitake, Joji Nasu, Yasuyuki Kato, and Yukitoshi Motome, “Majorana dynamical mean-field study of spin dynamics at finite temperatures in the honeycomb kitaev model,” Phys. Rev. B 96, 024438 (2017).

Kavita Mehlawat, A. Thamizhavel, and Yogesh Singh, “Heat capacity evidence for proximity to the Kitaev quantum spin liquid in A$_2$IrO$_3$ (A=Na, Li),” Phys. Rev. B 95, 144406 (2017).

Keijing Ran, Jinghui Wang, Wei Wang, Zhao-Yang Dong, Xiao Ren, Song Bao, Shichao Li, Zhen Ma, Yuan Gan, Youtian Zhang, J. T. Park, Guochu Deng, S. Danilkin, Shun-Li Yu, Jian-Xin Li, and Jinsheng Wen, “Spin-wave excitations evidencing the kitaev interaction in single crystalline α–RuCl$_3$,” Phys. Rev. Lett. 118, 107203 (2017).

We note that the energy unit $\alpha$ in Ref. 40 is half of the present energy unit. The label of the typical momenta is also different: The $Y$ point in Ref. 40 is denoted by $M$ in the present paper.

R. C. Coleka, S. M. Hayden, G. Aeppli, T. G. Perring, C. D. Frost, T. E. Mason, S.-W. Cheong, and Z. Fisk, “Spin waves and electronic interactions in La$_x$CuO$_y$,” Phys. Rev. Lett. 86, 5377–5380 (2001).

H. Shao, Y. Q. Qin, S. Capponi, S. Chesi, Z. Y. Meng, and A. W. Sandvik, “Nearly deconfined spinon excitations in the square-lattice spin-1/2 heisenberg antiferromagnet,” arXiv:1708.03232.

Craig C. Price and Natalia B. Perkins, “Critical properties of the kitaev-heisenberg model,” Phys. Rev. Lett. 109, 187201 (2012).

Craig Price and Natalia B. Perkins, “Finite-temperature phase diagram of the classical kitaev-heisenberg model,” Phys. Rev. B 88, 024410 (2013).

A. M. Samarakoon, A. Banerjee, S.-S. Zhang, Y. Kambara, S. E. Nagler, D. A. Tennant, S.-H. Lee, and C. D. Batista, “Comprehensive study of the dynamics of a classical kitaev spin liquid,” Phys. Rev. B 96, 134408 (2017).

Kensaku Takai, Kota Ido, Takahiro Misawa, Youhei Yamaji, Masatoshi Imada, “Finite-temperature variational monte carlo method for strongly correlated electron systems,” Journal of the Physical Society of Japan 85, 034601 (2016).

Rui-Zhen Huang, Hai-Jun Liao, Zhi-Yuan Liao, Hai-Dong Xie, Zhi-Yuan Xie, Hui-Hai Zhao, Jing Chen, and Tao Xiang, “A generalized lanczos method for systematic optimization of tensor network states,” arXiv preprint arXiv:1611.09574 (2016).

Joji Nasu, Junki Yoshitake, and Yukitoshi Motome, “Thermal transport in the kitaev model,” Physical Review Letters 119, 127204 (2017).

Ioannis Rousochatzakis and Natalia B. Perkins, “Classical spin liquid instability driven by off-diagonal exchange in strong spin-orbit magnets,” Phys. Rev. Lett. 118, 147204 (2017).

Andrei Catuneanu, Youhei Yamaji, Gideon Wachtel, Hae-Young Kee, and Yong Baek Kim, “Realizing quantum spin liquid phases in spin-orbit driven correlated materials,” arXiv:1701.07837.

Matthias Göhikel, Gideon Wachtel, Youhei Yamaji, Frank Pollmann, and Yong Baek Kim, “Signatures of quantum spin liquid in kitaev-like frustrated magnets,” arXiv:1706.09908.

An ED program package $\Phi$ is available through arXiv:1708.03232.
https://github.com/QLMS/HPhi.

89 Mitsuaki Kawamura, Kazuyoshi Yoshimi, Takahiro Misawa, Youhei Yamaji, Synge Todo, and Naoki Kawashima, “Quantum lattice model solver $\hat{H}\Phi$,” Computer Physics Communications 217, 180 – 192 (2017).

90 Susumu Yamamoto, Tomohiro Sogabe, Takeo Hoshi, Shao-Liang Zhang, and Takeo Fujii, “Shifted conjugate-orthogonal conjugate-gradient method and its application to double orbital extended Hubbard model,” J. Phys. Soc. Jpn. 77, 114713 (2008).

91 Yousef Saad, Numerical Methods for Large Eigenvalue Problems (SIAM, 2011).

92 Lawrence R Mead and Nikos Papanicolaou, “Maximum entropy in the problem of moments,” Journal of Mathematical Physics 25, 2404–2417 (1984).

93 Michael L. Williams and Humphrey J. Maris, “Numerical study of phonon localization in disordered systems,” Phys. Rev. B 31, 4508–4515 (1985).

94 K. Yakubo and T. Nakayama, “Absence of the hump in the density of states of percolating clusters,” Phys. Rev. B 36, 8933–8936 (1987).

95 Philipp Gegenwart and Simon Trebst, “Spin-orbit physics: Kitaev matter,” Nature Physics 11, 444–445 (2015).

96 Simon Trebst, “Kitaev materials,” [arXiv:1701.07056](https://arxiv.org/abs/1701.07056)

97 Leon Balents, “Spin liquids in frustrated magnets,” Nature 464, 1288–1295 (2010).

98 Tomohiro Sogabe, Takeo Hoshi, Shao-Liang Zhang, and Takeo Fujiiwara, “A numerical method for calculating the Green’s function arising from electronic structure theory,” in Frontiers of Computational Science (Springer, 2007) pp. 189–195.