



Symbolic determinant construction of perturbative expansionsIbsal Assi* and J. P. F. LeBlanc *Department of Physics and Physical Oceanography, Memorial University of Newfoundland,
St. John's, Newfoundland & Labrador, Canada A1B 3X7* (Received 16 May 2023; revised 11 December 2023; accepted 14 February 2024; published 25 March 2024)

We present a symbolic algorithm for the fully analytic treatment of perturbative expansions of Hamiltonians with general two-body interactions. The method merges well-known analytics with the recently developed symbolic integration tool, algorithmic Matsubara integration, that allows for the evaluation of the imaginary frequency/time integrals. By symbolically constructing Wick contractions at each order of the perturbative expansion we order by order construct the fully analytic solution of the Green's function and self-energy expansions. A key component of this process is the assignment of momentum/frequency conserving labels for each contraction that motivates us to present a fully symbolic Fourier transform procedure which accomplishes this feat. These solutions can be applied to a broad class of quantum chemistry problems and are valid at arbitrary temperatures and on both the real- and Matsubara-frequency axis. To demonstrate the utility of this approach, we present results for simple molecular systems as well as model lattice Hamiltonians. We highlight the case of molecular problems where our results at each order are numerically exact with no stochastic uncertainty.

DOI: [10.1103/PhysRevB.109.125143](https://doi.org/10.1103/PhysRevB.109.125143)**I. INTRODUCTION**

Quantum many-body systems (QMBS) are generally difficult to treat due to the large number of degrees of freedom associated with them. The number of electrons in condensed matter systems puts limitations on what information can be accessed from the full Hilbert space (HS). In practice, only a small portion of the HS is accessible via the numerical tools. One of the famous examples of those numerical methods is the density functional theory (DFT), which is generally used to study the structure and interactions of different physical systems [1]; on the other hand, DFT fails in studying strong correlation effects in various materials [2].

Perturbation theories are a fundamental tool in a physicist's arsenal for tackling interacting electron systems. In many body perturbation theory (MBPT), physical observables are expressed as an infinite series where each subsequent order is represented by an exponentially large number of contractions generated from Wick's theorem. Each contraction requires the evaluation of integrals over the set of all internal variables. There are several ways to treat MBPT numerically, with the most popular perhaps being diagrammatic Monte Carlo (DiagMC) algorithms [3–6]. Standard DiagMC methods suffer from the fermionic sign problem that results from the large number of contractions (diagrams) with alternating sign [3,4]. In recent years, determinant methods have been introduced that can somewhat mitigate this issue [7–9]. The connected determinant diagrammatic Monte Carlo (CDet) method was introduced to treat perturbative expansions and avoids the factorial scaling of diagrams at exponential cost [10–12].

Those methods always have stochastic uncertainties and since they are based on the Matsubara formalism for finite temperatures they require numerical forms of analytic continuation in order to produce dynamical properties in real frequency or real time. More recently the advent of the algorithmic Matsubara integration (AMI) [13,14] method allows us to symbolically evaluate summations over Matsubara frequencies and has been successfully applied to a number of physical problems such as the 2D Hubbard model [15–18] as well as the uniform electron gas [19,20]. AMI provides access to real frequency calculations via textbook analytic continuation, the replacement $i\omega_n \rightarrow \omega + i0^+$, which avoids ill-posed numerical analytic continuation schemes [21]. It reduces the sampling space of internal variables minimizing the effect of the curse of dimensionality and reducing overall numerical uncertainty.

In this work, we build on determinantal methods by introducing a fully algorithmic approach which we call the symbolic determinant method (symDET) which applies MBPT to extremely general Hamiltonians relevant to quantum chemistry and condensed matter physics. We start by generating all Wick contractions symbolically from a determinant and then proceed to Fourier transform those contractions also symbolically. We then perform the integrals over the internal variables with the use of AMI for evaluating the Matsubara summations. In the next section, we introduce the elements of this algorithm in detail. We then provide several applications in the following section and provide a summary.

II. MODEL AND METHODS**A. Two-body Hamiltonian**

We discuss the evaluation of a very general two-body Hamiltonian with two terms: a single-particle term, H_0 , and

*iassi@mun.ca

a generalized four-operator interaction term, H_V . These are given by

$$H = \underbrace{\sum_{ab} h_{ab} c_a^\dagger c_b}_{H_0} + \frac{1}{2} \underbrace{\sum_{abcd} U_{abcd} c_a^\dagger c_c^\dagger c_d c_b}_{H_V}. \quad (1)$$

Here a and b are arbitrary band indices—that might also include momenta or spin degrees of freedom—and the c_i^\dagger and c_i represent standard creation and annihilation operators in the state i , respectively, and the values of h_{ab} represent one-electron integrals while U_{abcd} is the two-electron interaction matrix. The presumption for finding solutions to the model are that the single-particle term, H_0 , is known and diagonal allowing us to perform an expansion in powers of the interaction term.

B. Perturbative expansion of Green's function

We define the noninteracting Green's function

$$g_{ba}(\tau) = -\langle c_b(\tau) c_a^\dagger(0) \rangle = [(-\partial_\tau + \mu)\mathbb{1} - \mathbf{h}]_{ba}^{-1}, \quad (2)$$

here written in imaginary time, τ . Later we will perform the Fourier transform to represent the Green's function for Matsubara frequency, $i\nu_n$. In general, h_{ab} may not be diagonal which results in a nondiagonal Green's function. Without loss of generality we simplify the problem by presuming that H_0 can be represented in a diagonal basis and that the interaction U_{abcd} is known in that diagonal basis. Thus we can rewrite the diagonal Green's function on the Matsubara axis as

$$g_{ab}(i\nu_n) = \frac{\delta_{ab}}{i\nu_n - h_{ab}}, \quad (3)$$

where δ_{ab} is the Kronecker delta. In this representation the poles of the Green's function can be symbolically determined and this is necessary when implementing the AMI method [13]. If h_{ab} is not diagonal, then \mathbf{g} is not diagonal and the pole structure of the Green's function becomes obfuscated by the matrix inversion process.

With the target of generating the order-by-order expansion of H_V we start by following the standard construction of the m th order correction to the imaginary time Green's function as

$$G_{ba}^{(m)}(\tau) = \frac{(-1)^m}{m!} \left\langle \mathcal{T} \left[\prod_{\ell=1}^m \int_0^\beta d\tau_\ell H_V(\tau_\ell) \right] c_b(\tau) c_a^\dagger(0) \right\rangle_{0c}, \quad (4)$$

where \mathcal{T} is the time ordering operator and $\beta = T^{-1}$ is the inverse temperature in units of the Boltzmann constant k_B . In Eq. (4) the expectation value is with respect to the unperturbed Hamiltonian and should include only connected diagram topologies. We see that at order m we must compute the expectation value of a sequence of $4m$ creation and annihilation operators attached to times τ_m , in addition to the external operators $c_b(\tau)$ and $c_a^\dagger(0)$.

This expectation value can be evaluated using Wick's theorem, replacing the expectation value with a sum of all possible contractions of creation and annihilation operators. This is typically accomplished in matrix form with rows and columns

represented by annihilation and creation operators, respectively. One can then generate all possible contractions—while also keeping correct track of the fermionic sign arising from commuting fermionic operators—by just taking the determinant of said matrix [10,12].

For this we define $\mathbf{G}^{(m)}$ to be a $(2m+1) \times (2m+1)$ matrix in which the rows (columns) correspond to the $2m$ annihilation (creation) operators plus an additional entry in each for the external vertices. We introduce column and row indices α, β such that

$$\begin{aligned} \{a_\alpha\} &:= \{a_1, c_1, a_2, c_2, \dots, a_m, c_m, a_{out}\}, \\ \{b_\beta\} &:= \{b_1, d_1, b_2, d_2, \dots, b_m, d_m, b_{in}\}, \end{aligned} \quad (5)$$

and define the matrix elements [11]

$$\mathbf{G}_{\beta\alpha}^{(m)} := -\langle c_{b_\beta}(\tau_\beta) c_{a_\alpha}^\dagger(\tau_\alpha) \rangle_0 = g_{b_\beta a_\alpha}(\tau_\beta - \tau_\alpha + 0^-) = g_{\alpha\beta}. \quad (6)$$

The full matrix can then be written

$$\mathbf{G}^{(m)} := \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} & g_{1a} \\ g_{21} & g_{22} & \cdots & g_{2n} & g_{2a} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} & g_{na} \\ g_{b1} & g_{b2} & \cdots & g_{bn} & g_{ba} \end{bmatrix}, \quad (7)$$

where $n = 2m$.

This construction has been presented numerous times and forms the basis for determinant Monte Carlo methods applied to many-body systems [10–12,22]. In the standard prescription, the \mathbf{G} matrix is populated in the real space and imaginary time. The determinant procedure is typically evaluated numerically by inserting numerical values for the imaginary time Green's function and sampling over all continuous times τ_m . The one caveat to doing this is that the terms generated represent both connected and disconnected Feynman graphs. Removing the disconnected components can be accomplished with the recent method described by Rossi *et al.* [10]. However, in our work, contractions corresponding to disconnected diagrams will be removed via our symbolic Fourier transformation, a process that does the two tasks at once as detailed later.

Wick's theorem implies that the expectation value in Eq. (4) can be written in the form of a determinant of $\mathbf{G}^{(m)}$. However, doing so is difficult in general because the expectation value in Eq. (4) should allow only connected topologies in the Wick contractions. For now we rewrite Eq. (4) as

$$\begin{aligned} G_{ba}^{(m)}(\tau) &= \frac{(-1)^m}{2^m m!} \left[\prod_{p=1}^{p=m} \sum_{a_p b_p c_p d_p} U_{a_p b_p c_p d_p} \right] \\ &\times \left[\prod_{\ell=1}^m \int_0^\beta d\tau_\ell \right] \det(\mathbf{G}^{(m)})_c, \end{aligned} \quad (8)$$

where the c subscript on the determinant implies that we include only the connected terms in the determinant. Such a “connected-determinant” function does not in general exist for an arbitrary matrix and this issue has driven the development of numerical schemes [10]. In what follows we devise a symbolic approach to evaluating Eq. (8) which allows us to quickly identify and remove disconnected Wick contractions.

C. Algorithmic Matsubara integration

The method of algorithmic Matsubara integration, introduced in Ref. [13], was presented as a general procedure for the analytic evaluation of the temporal integrals of arbitrary Feynman diagram expansions. In essence, AMI is a straightforward application of residue theorem that stores the minimal information required to construct the analytic solution for an arbitrarily complex integrand comprised of a product of bare Green's functions. The Matsubara integrals are not conceptually challenging to perform and are the topic of numerous textbook exercises. The difficulty in performing those contour integrals lies only in that the number of poles and number of resulting analytic terms grows exponentially with diagram order.

Using the existing AMI library [14] the result of AMI is stored in three nested arrays: signs/prefactors S , complex poles P , and Green's functions R . From these three objects, whose storage is quite minimal, one can then construct the analytic expression symbolically through elementary algebraic operations [13].

The beauty of such an approach is that the result is analytic in external variables, allowing for true analytic continuation of $iv_n \rightarrow v + i0^+$, and is also an explicit function of temperature, T . Further, for a given graph topology the AMI procedure need only be performed once and is valid for any choice of dispersion in any dimensionality and can be applied to model systems for a wide variety of Feynman diagrammatic expansions [17,18,20,23,24]. For the present work, we use AMI as a method for evaluating Matsubara integrands and the determinant construction replaces the usual Feynman diagram representation.

D. Symbolic determinant method—symDET

In this work we deviate from the standard determinantal scheme mentioned in Sec. II B. Here we will outline a procedure to generate the perturbative expansion in terms of bare propagators such that the integrands of Eq. (4) are in a form suitable for AMI [13]. AMI operates in the energy(momentum)/frequency basis and cannot be applied to imaginary or real-time Green's functions—though there exist nonalgorithmic variants designed in the same spirit that may perhaps overcome this barrier [25]. Each term in Eq. (4) is represented as a function of a set of imaginary times. To translate these to a form amenable to AMI we require tools to (1) perform the symbolic Wick's contractions for each term in Eq. (4), (2) identify and remove disconnected topologies, and (3) perform the nested sequence of Fourier transform from $\tau \rightarrow iv_n$ symbolically.

We provide the solution to each issue in the following subsections.

1. Symbolic Wick's contractions

When creating a symbolic representation of the matrix form of Eq. (7) each element with row and column indices α and β is just a function of those indices. We can therefore generate a symbolic representation by replacing the entries with their row and column indices, $\mathbf{G}_{\alpha\beta} \rightarrow (\alpha, \beta)$.

If we can take a determinant of this matrix and store each term separately, we will have generated the expressions that represent the $n!$ connected and disconnected diagrams. While the evaluation of numerical determinants can be accomplished in $O(n^3)$ time, an advantage of modern determinantal methods [10,26,27] here is that we want to proceed symbolically; there is no obvious route to such fast evaluations. Instead, we take the most pedantic approach and simply store the explicit parameters of each term in the determinant. While this factorial scaling sounds problematic the tradeoff is an analytic expression that is exact to machine precision. This is in lieu of stochastic methods that, while they can evaluate determinants quickly, must perform temporal integrals via Monte Carlo sampling, a process that for high accuracy requires typically $10^6 \rightarrow 10^8$ samples. We expect that for low orders we will arrive at a precise numerical result with fewer operations despite this factorial scaling.

To proceed we use the Leibniz formula for an $n \times n$ matrix, A , with elements $a_{i,j}$:

$$\det(A) = \sum_{p \in \mathcal{P}_n} \left(\text{sgn}(p) \prod_{i=1}^n a_{i,p_i} \right). \quad (9)$$

In this expression, $p = (p_1, p_2, \dots, p_n)$ is a permutation of the set $\{1, 2, \dots, n\}$ and \mathcal{P}_n is the set of all such permutations. $\text{sgn}(p)$ is the signature of p defined as $+1$ whenever the reordering requires an even number of interchanges and -1 when an odd number is required. Finding the permutations of p and the associated signs is a straightforward computational problem. To do this symbolically we generate a permutation p and then store the indices of $a_{i,p_i} = (i, p_i)$ for each i . Each term in Eq. (9) is then completely defined by a vector of such pairs and a single $+1/-1$ sign prefactor.

This represents a major departure from typical determinantal QMC methods [28,29] where such a matrix is filled with numerical values. In our case we have yet to assign values to the entries and instead we want to store the information required to later symbolically construct the expression.

2. Two in one: Symbolic Fourier transform (SFT)

Motivated by the AMI algorithm [13], we proceed to work in the frequency domain; hence the need to do the Fourier transform of Eq. (4). We should recall that for Feynman diagrams the set of possible diagram topologies is independent of coordinate and temporal labeling of each vertex. However, in the contractions of Eq. (7) each topology may appear multiple times—as is famously the case for a single-band problem where the $m!$ denominator is precisely canceled by $m!$ duplicates of each topology. Since we have each contraction, we are free to represent each as a graph in momentum and Matsubara frequency space. However, in doing so one would need to develop an internally consistent labeling of each graph—a process that is fundamentally nonlocal in diagram topology and also is not unique.

Instead we choose to mimic the analytic process and have devised an analytic representation of the temporal Fourier transform. The procedure, detailed in the Appendix, performs the Fourier transform by first sorting the contraction pairs (i, j) that represent imaginary time Green's functions

spanning between times $\tau_{\lfloor \frac{i}{2} \rfloor}$ and $\tau_{\lfloor \frac{j}{2} \rfloor}$. The pairs are then separated into three sublists A , B , and C based on criterion discussed in the Appendix. Since the contraction pairs are effectively source/target sets, the connectivity of the contraction can be determined directly as is done in graph theory, identical to a depth first search, at minimal expense, which scales linear with the perturbation order n . If at the end of the process the number of pairs in A is $n - 1$, in B is 2, and in C we have n pairs, then the contraction is a connected diagram and if not it can be excluded. Hence a by-product of performing the Fourier transform is a mechanism to exclude disconnected diagrams and this can then be used to evaluate Eq. (8). The symbolic Fourier transform of the time integrals is done by simply converting those three lists to matrices as described in the Appendix. The advantage of this is that one obtains a unique set of internal labels that obey frequency and momentum conservation at all vertices. At the end, we obtain the frequency labeling matrix that serves as an input for the AMI algorithm [13]

$$\omega_{\text{AMI}} = \left(\frac{\beta}{I_n} \mid \frac{\alpha}{0} \right), \quad (10)$$

where α and β are matrices with entries $\{0, \pm 1\}$ [see Eq. (A6)], I_n is an $n \times n$ identity matrix, and 0 here represents an n -dimensional zero vector.

E. Evaluation

The frequency domain version of Eq. (8) is of the form

$$G_{ba}^{(n)}(i\omega_{\text{ex}}) = \frac{(-1)^n}{2^n n!} g_b(i\omega_{\text{ex}}) g_a(i\omega_{\text{ex}}) \times \sum_{c \in \mathcal{C}} \prod_{p=1}^p \sum_{a_p b_p c_p d_p} U_{a_p b_p c_p d_p} \sum_{\{i\Omega_n\}} \prod_j^{2n-1} g_{e_j}^j(\alpha_j \cdot \omega), \quad (11)$$

where the first summation is over the internal (nonfrequency) variables $\{a_i, b_i, c_i, d_i\}$, the second summation is over the set of internal Matsubara frequencies $\{i\Omega_n\}$, the last is over all contractions belonging to the set \mathcal{C} (those which are equivalent to connected diagrams), and e_j is the nonfrequency label attached to the diagonal noninteracting propagator $g_{e_j}^j(\alpha_j \cdot \omega)$. Here α_j is the j th row in Eq. (10), $\omega = (i\Omega_1, \dots, i\Omega_n, i\omega_{\text{ex}})^T$, and

$$g_{e_j}^j(\alpha_j \cdot \omega) = \frac{1}{\alpha_j \cdot \omega - \varepsilon_{e_j}} \quad (12)$$

is the Fourier transformed free propagator. In the case of molecular problems, or generically *discrete* systems, one performs the e_j summations directly such that our algorithm gives the exact value of the perturbative expansion. However, in the case of lattice problems, we use stochastic sampling over momenta and we obtain results with stochastic error bars. In both cases, the Matsubara summations are evaluated exactly.

Evaluation of Eq. (11) involves a number of steps as follows.

- (1) A contraction is selected from the determinant.
- (2) A set of band indices, $\{a_i, b_i, c_i, d_i\}$, are chosen.

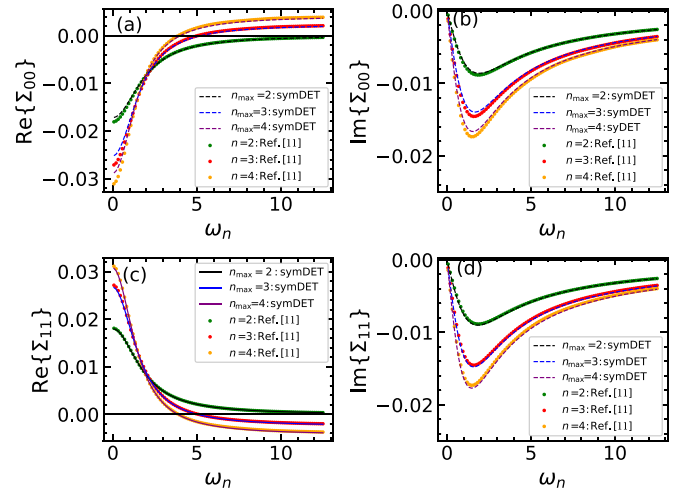


FIG. 1. (a), (b) Real and imaginary parts of the self-energy for H_2 in the STO-6g basis with external band indices $a_{\text{ex}} = b_{\text{ex}} = 0$. (c), (d) Plots of the real and imaginary parts of Σ^{H_2} for $a_{\text{ex}} = b_{\text{ex}} = 1$. Here we took $\beta = 50.0 E_h^{-1}$.

- (3) Energy conserving labels are generated for this contraction/ band index via symbolic Fourier transform.
- (4) Symbolic expression for the sum over $\{i\Omega_n\}$ is generated via AMI.
- (5) Analytic expression is evaluated and its result is stored.

III. APPLICATIONS

A. Application to molecular chemistry— H_2

Molecular hydrogen is the simplest system to consider as a testbed for method development and here we start with the simplest representation in the STO-6g basis which describes the interaction between the two hydrogen atoms having only $1s$ orbitals. In particular, we will see later in Sec. III C that the two state problem is the basic component of a single band with spin \uparrow/\downarrow and therefore correct results for the STO-6g basis are paramount in developing the method beyond simple problems. We use the PYSCF package [30] to obtain the Hartree-Fock solutions for the STO-6g basis in units of E_h from which we compute the self-energy on the Matsubara axis illustrated in Fig. 1. We have compared our results in detail to those in Ref. [11] and find that our exact result is within stochastic error bars of that work. Different from their result, our starting eigenstates are asymmetric resulting in distinct values of Σ_{00} and Σ_{11} , while the off-diagonal self-energy terms are zero in this case. While we stop at fourth order, there is no conceptual hurdle to evaluating higher orders or larger orbital basis sets. However, the computational expense is factorial in order and exponential in basis. Nevertheless, the procedure is easily parallelizable.

The real advantage to our approach is the direct evaluation of real frequency properties. By symbolically replacing $i\omega_n \rightarrow \omega + i\Gamma$ we can plot the self-energy in real frequencies shown in Fig. 2 for a particular choice of Γ that can be made arbitrarily small. Here we focus on a relevant frequency range where there is an expected new peak that is created by a sharp feature in $\text{Re}\Sigma(\omega)$ such that the interacting Green's

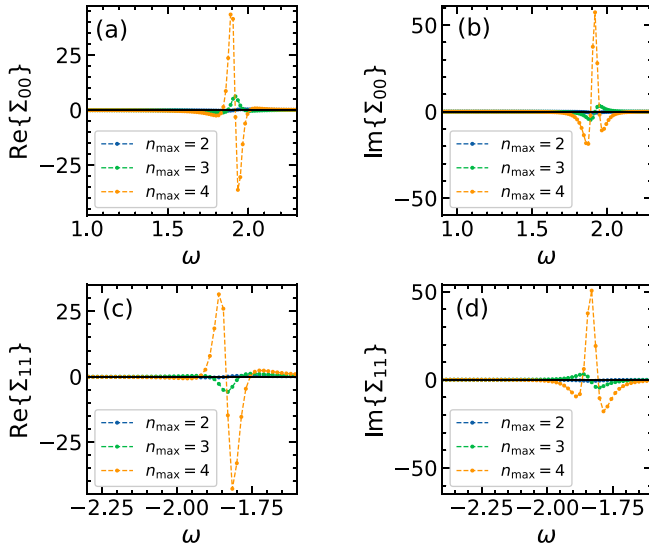


FIG. 2. (a), (b) Real and imaginary parts of Σ_{00} , while (c) and (d) are the components of the self-energy for the second band for H_2 (in the STO-6g basis) on the real frequency axes in units of E_h . Here we took the regulator $\Gamma = 0.05$.

function gains one or more additional poles. This is seen in the spectral function as shown in Fig. 3. The dominant peaks remain those of the noninteracting dispersion while additional peaks—shown in the insets—appear at energies offset by the peak difference $\Delta E = h_{11} - h_{00}$, which is expected based on the second order expansion. At fourth order shown, there are two additional poles instead of a single peak near $\omega = \pm 2$.

As an example for a larger orbital basis set, we compute the self-energy for H_2 in the 10 orbital cc-pVDZ basis representation as shown Fig. 4. This basis is five times larger than its STO-6g counterpart, stressing our ability to study larger molecules with symDET. We should stress that the imaginary frequency and real frequency calculations for such small systems can be done via exact diagonalization (ED) as done in other works [11]; however, ED is restricted by systems

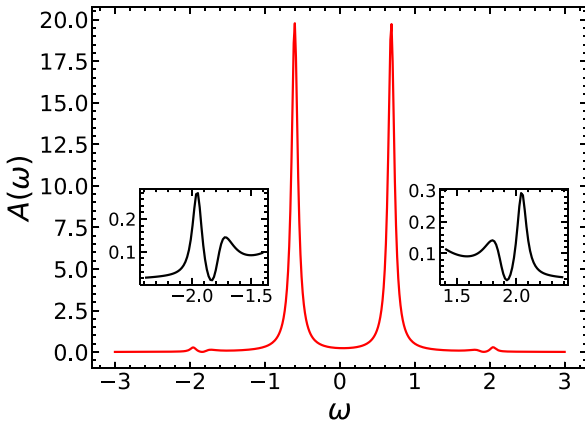


FIG. 3. Spectral function for H_2 in the STO-6g basis truncated at fourth order. Inset data is a zoom out of the extra peaks with lower intensity. Here we took the regulator 0.05.

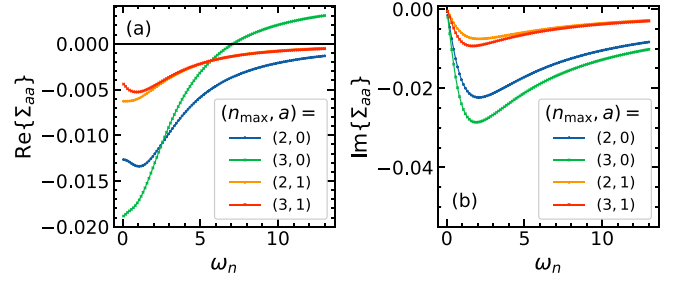


FIG. 4. Matsubara self-energy for hydrogen in the cc-pVDZ basis versus the Matsubara frequency with $\beta = 50.0$. (a) The real part of the self-energy components (0,0) and (1,1) truncated at second ($n_{\max} = 2$) and third ($n_{\max} = 3$) orders and (b) are the imaginary counterparts.

size [31,32], something that is not a hurdle for many-body perturbative methods.

In principle our method could also be used to obtain other observables, for example, the binding energy for molecules can be obtained by generating the poles of the full propagator, which is straightforward via the AMI part of our code. An interesting implication of these calculations is the possibility of performing self-consistent perturbation theory beyond the well-known second order Green's function perturbation theory (GF₂) [28,33,34]. Since we can obtain results on the real frequency axis one could in principle merge our scheme with a spectral representation as done in [19] or with a Prony pole approximation as in [35].

B. Hubbard-Dimer model

To demonstrate the versatility of our approach we study the Hubbard dimer. The model consists of two sites each having a spin 1/2 particle. The model we use is described below [36]:

$$H = H_0 + H_U + H_H + H_{SB} - \mu \sum_{i,\sigma} c_{i\sigma}^\dagger c_{i\sigma}, \quad (13)$$

where $H_0 = -t \sum_{\sigma=\uparrow,\downarrow} (c_{0\sigma}^\dagger c_{1\sigma} + c_{0\sigma} c_{1\sigma}^\dagger)$ is the hopping term for electrons between the two sites, $H_U = U \sum_i n_{i\uparrow} n_{i\downarrow} - \frac{U}{2} \sum_{i\sigma} n_{i\sigma}$ describes the on-site interaction, $H_H = H \sum_i (n_{i\uparrow} - n_{i\downarrow})$ is the interaction due to an applied magnetic field, and $H_{SB} = U_a (n_{0\uparrow} n_{0\downarrow} - n_{1\uparrow} n_{1\downarrow}) + \mu_a (n_{0\uparrow} + n_{0\downarrow} - n_{1\uparrow} - n_{1\downarrow}) + H_a (n_{0\uparrow} - n_{0\downarrow} - n_{1\uparrow} + n_{1\downarrow})$ is a symmetry-breaking term. By diagonalizing the quadratic part of the full Hamiltonian, we can rewrite the above Hamiltonian in the usual form

$$H = \sum_{a=1}^4 \varepsilon_a f_a^\dagger f_a + \frac{1}{2} \sum_{abcd} V_{abcd} f_a^\dagger f_c^\dagger f_d f_b, \quad (14)$$

where ε_a is the effective dispersion, f_a^\dagger (f_a) are the creation (annihilation) fermionic operators, and V_{abcd} is the effective interaction, where both ε_a and V_{abcd} can be obtained analytically for this four-band system. In this example, the self-energy in this basis is not diagonal (rather a block diagonal). As an illustration, we plot the imaginary and real parts of Σ_{00} and Σ_{01} up to fourth order for $U = 2.5t$, $\mu = 0.7$,

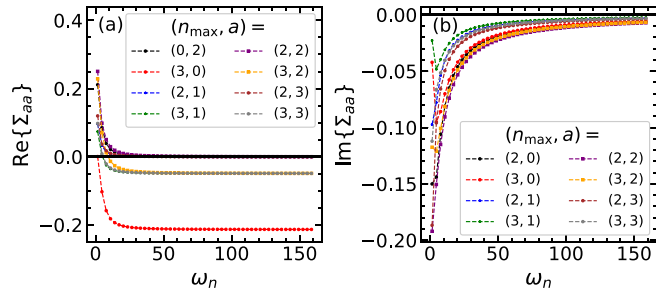


FIG. 5. (a), (b) Real and imaginary parts of the diagonal elements of the self-energy matrix for $n_{\max} = 2, 3$ for the Hubbard dimer model. Here we took $t = 1.0, U = 2.5, \mu = 0.70, H = 0.30, U_a = 0.50, \mu_a = 0.20, H_a = 0.030$, and $\beta = 2.0$.

$H = 0.30, U_a = 0.5, \mu_a = 0.20, H_a = 0.030$, and $\beta = 2.0$. (See Fig. 5.)

C. Single-band Hubbard model

The simplest starting point for considering a lattice Hamiltonian is the single-band Hubbard model of spin-1/2 fermions on a square lattice. The model is typically written in real-space notation as

$$H = \sum_{\langle ij \rangle \sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (15)$$

where t_{ij} is the hopping amplitude, $c_{i\sigma}^{(\dagger)}$ is the annihilation (creation) operator at site i , $\sigma \in \{\uparrow, \downarrow\}$ is the spin, U is the on-site Hubbard interaction, $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is the number operator, μ is the chemical potential, and $\langle ij \rangle$ restricts the sum to nearest neighbors. For a 2D square lattice we take $t_{ij} = -t$, resulting in the free particle energy

$$\epsilon(\mathbf{k}) = -2t[\cos(k_x) + \cos(k_y)] - \mu. \quad (16)$$

Mapping this problem to Eq. (1) leads to an effective problem of two degenerate bands with states $\uparrow = (k, \sigma = \uparrow)$ and $\downarrow = (k, \sigma = \downarrow)$ and the band indices are then summed over up and down bases. This leads to a diagonal and spin independent $h_{ab} = \epsilon_k \delta_{ab}$ and an interaction term independent of momentum with entries $U_{\uparrow\uparrow\downarrow} = U_{\downarrow\downarrow\uparrow} = U$ and all other U elements zero.

Due to the additional k indices, after processing with AMI each m th order Wick contraction contains an m -dimensional integral over internal momentum vectors which requires approximate numerical integration methods to evaluate. We proceed stochastically and therefore the results are no longer fully analytic. Otherwise, the procedure is unchanged from the two-band case of H_2 in the STO-6g basis which highlights the importance of that problem as a benchmark. As an illustration, we have calculated the self-energy for the 2D square lattice on the Matsubara axis shown in Fig. 6 for doped cases $\mu \neq 0$. Moreover, the exact same expressions can be used to generate the matching real-frequency results which we show in Fig. 7.

IV. CONCLUSIONS

In this work we have developed an algorithm that can handle single and multiband problems for general two-body

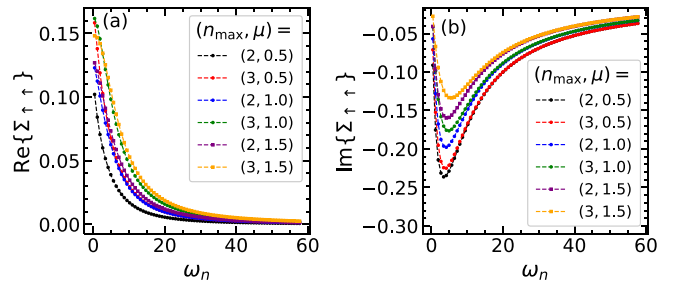


FIG. 6. (a) Real part of the (spin up) self-energy of the two dimensional Hubbard model for $t = 1.0, U = 3.0, \beta = 8.33, \vec{k} = (0, \pi)$, and at different values of μ as indicated, and (b) are the imaginary counterparts.

interaction models at equilibrium. The steps to our determinant method are (1) generating contractions by evaluating the proper determinant, (2) performing the symbolic Fourier transform, (3) using the AMI to evaluate the Matsubara summations exactly, and (4) sum or sample any remaining internal degrees of freedom.

We have applied our algorithm to a variety of problems from molecular chemistry to lattice models up to fourth order perturbation theory. The method is therefore flexible and can solve different models in both real and imaginary frequency domains allowing it to be of great importance for both quantum chemistry and lattice system applications. The bottleneck in computation of lattice systems remains the numerical integration over remaining spatial degrees of freedom. When the numerical regulator Γ is small this becomes difficult due to the sharp nature of the integrands. The use of renormalized perturbation theory might help alleviate these difficulties [18]. Finally, our algorithm, equivalent to a single shot GFn [28], exceeds what is currently available. Although we limited ourselves to fourth order calculations, higher order corrections can be achieved, since the algorithm is valid at any arbitrary perturbation order and system size. Of particular interest is molecular problems where we are able to evaluate each perturbative order exactly to machine precision. In these cases, regardless of the computational expense of higher orders, since the result is exact, it need only *ever* be computed once.

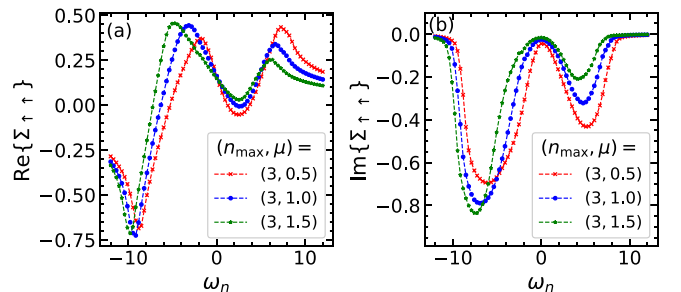


FIG. 7. (a), (b) Real and imaginary parts of the self-energy (truncated at third order) versus the real frequency for the 2D Hubbard model evaluated for the parameter's choice: $U = 3t = 3.0, \beta = 8.33$, and $\vec{k} = (0, \pi)$ with different values of μ as indicated. We took a Monte Carlo sample of size 1×10^8 and the regulator $\Gamma = 0.2$.

ACKNOWLEDGMENTS

J.P.F.L. would like to thank G. Booth and O. Backhouse for helpful discussion. We would like to thank as well J. Li and E. Gull who were instrumental at the onset of this work. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-2022-03882 and support from the Simons Collaboration on the Many Electron Problem. Computational resources were provided by the Digital Research Alliance of Canada. Our Monte Carlo codes make use of the open source ALP-SCore framework [37,38] and the LIBAMI package [14].

APPENDIX: STEPS OF SYMBOLIC FOURIER TRANSFORMATION

1. Sorting Wick's contractions

An important step to perform the symbolic Fourier transformation is to sort the given Wick contraction that corresponds to a connected diagram as follows. First, let us represent the given contraction as $\mathcal{C} = [\vec{P}, s]$, where $\vec{P} = (p_1, p_2, \dots, p_{2n+1})$ is a vector of pairs representing each fermion line with $p_j = (\tau_0^j, \tau_1^j)$ and s is the sign of the contraction. In the language of graph theory, \vec{P} contains the edges of the graph. To check if the diagram is connected or not, one can use the depth first search (DFS), which requires \vec{P} as an input. If \mathcal{C} is connected, then we introduce three vectors of pairs \vec{A} , \vec{B} , and \vec{C} , where we store the pairs from \vec{P} into these three vectors based on the following convention. The pairs representing connection with external vertices are stored in \vec{B} and the pairs which represent loops, i.e., tadpole/clamshell structures, are stored in \vec{C} .

The next step is to reduce the number of pairs in \vec{A} to $n - 1$, which is adopted from the basic graph theory fact that a given connected graph with n vertices has $n - 1$ edges connecting all the vertices together (plus the extra edges). This can be done recursively using the DFS by removing one pair at a time from \vec{A} and applying the DFS to check if the remaining pairs keep all the vertices connected or not. If the removal of a given pair does not affect the connectedness, then the pair should be added to \vec{C} ; otherwise, it should be put back into \vec{A} and then move to the next pair in \vec{A} and do the same steps until the number of pairs is $n - 1$. At this moment, the numbers of pairs in \vec{C} is n , with the total number of pairs in all the three vectors $2n + 1$ as expected. The contraction \mathcal{C} will have the form

$$\mathcal{C} = [\vec{A}, \vec{B}, \vec{C}, s]. \quad (\text{A1})$$

2. Array representation of the noninteracting Green's function

Let us assume that the fermionic line connecting two vertices τ_i and τ_j in an n th order Feynman diagram is represented by a Green's function of the form $g(\eta; \tau_i - \tau_j)$, where η is a set of quantum labels attached to the corresponding Green's function. We introduce the following useful array representation of $g(\eta; \tau_i - \tau_j)$:

$$g(\eta; \tau_i - \tau_j) := [V_j(1 - \delta_{ij}), \eta], \quad (\text{A2})$$

where $V_j \in \mathbb{R}^n$ is an n -dimensional vector defined in the following way.

(i) If the fermionic line is connecting two different internal vertices, then V_j has $+1$ at the i th row, -1 at the j th row, and zeros elsewhere.

(ii) V_j is the n -dimensional zero vector if $\tau_i = \tau_j$. This is guaranteed by δ_{ij} in the equation above.

(iii) The two external fermionic lines are represented with an n -dimensional vector with only one nonzero entry ± 1 . Basically, when τ_j is external time and τ_i is internal time, then V_j has entry of $+1$ at the i th row and zeros elsewhere. On the other hand, if τ_i is the external time and τ_j is an internal time, then V_j is an entry -1 at the i th row and zeros elsewhere.

Following this notation, we can represent a Wick contraction (A1) as

$$\mathcal{C} = [M, s], \quad (\text{A3})$$

where $M = (A|B|C)$ is an $n \times 2n + 1$ matrix obtained by mapping the pairs in $\{\vec{A}, \vec{B}, \vec{C}\}$ into column vectors using the convention explained above. Basically, the $n - 1$ pairs in \vec{A} form an $n \times n - 1$ matrix A , the 2 pairs in \vec{B} form an $n \times 2$ matrix B , and the n pairs in \vec{C} form an $n \times n$ matrix C . In the next section, we will use this result to obtain the Fourier transformation of the contraction \mathcal{C} .

3. Symbolic Fourier transform

Let us assume that the fermionic lines whose vectors stored in A have the dependent Matsubara frequencies $\{\omega_1, \omega_2, \dots, \omega_{n-1}\}$, the ones stored in B have the external frequency ω_{ex} , and the vectors stored in C have the independent Matsubara frequencies $\{\Omega_1, \Omega_2, \dots, \Omega_n\}$. Defining $\vec{\Omega} = (\omega_1, \omega_2, \dots, \omega_{n-1}, \omega_{\text{ex}}, \omega_{\text{ex}}, \Omega_1, \Omega_2, \dots, \Omega_n)^t$, then one can show that the equation that connects all the frequencies together is

$$M\vec{\Omega} = \vec{0}. \quad (\text{A4})$$

The above equation is thought of as the set of delta functions which act to enforce conservation laws at each vertex so long as Eq. (A4) is satisfied. Our task is to represent the dependent frequencies in terms of the other frequencies, which is obtained using the above equation, giving

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \vdots \\ \omega_{n-1} \end{pmatrix} = \alpha \omega_{\text{ex}} + \beta \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_n \end{pmatrix}, \quad (\text{A5})$$

where

$$\alpha = -J^{-1}A^T B \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \beta = -J^{-1}A^T C, \quad (\text{A6})$$

with $J = A^T A$ being an $(n - 1) \times (n - 1)$ matrix. The above Eq. (A5) gives a unique representation of the frequency labels which satisfy the conservation laws at all internal vertices. Using this notation, a Green's function with a dependent frequency ω_j , i.e., $g(\eta; \omega_j)$, will be represented as

$$g_k(\eta_k; \omega_j) = \frac{1}{i\beta_j \cdot \vec{\Omega}_{\text{ind}} + i\alpha_j \omega_{\text{ex}} - \varepsilon_{\eta_k}}, \quad (\text{A7})$$

where β_j is the j th row in β , α_j is the j th entry in α , and $\bar{\Omega}_{\text{ind}} = (\Omega_1, \Omega_2, \dots, \Omega_n)^T$. Consequently, we introduce the Fourier transformation of the Wick contraction (A1) as

$$\mathcal{F}[C] := [g_A, g_B, g_C, s], \quad (\text{A8})$$

where

$$g_A = [g_1(\eta_1; \omega_1), g_2(\eta_2; \omega_2), \dots, g_{n-1}(\eta_{n-1}; \omega_{n-1})], \quad (\text{A9})$$

$$g_B = [g_n(\eta_n; \omega_{\text{ex}}), g_2(\eta_{n+1}; \omega_{\text{ex}})], \quad (\text{A10})$$

and

$$g_C = [g_{n+2}(\eta_{n+2}; \Omega_1), \dots, g_{2n+1}(\eta_{2n+1}; \Omega_n)], \quad (\text{A11})$$

where the Fourier transformed Green's functions in $\{g_B, g_C\}$ take the following simple form:

$$g_\ell(\eta_\ell; \omega) = \frac{1}{i\omega - \varepsilon_{\eta_\ell}}. \quad (\text{A12})$$

Theorem 1. Let M be an $n \times 2n + 1$ matrix representing one particular contraction belonging to specific topology \mathcal{T} with M satisfying (A4); then the frequency matrix ω_{AMI} (10) is unique for all contractions belonging to the same \mathcal{T} .

Proof. We know that there are $2^n n!$ contractions per topology \mathcal{T} at n th order. The factor 2^n coming from inverting the interaction line at each vertex essentially keeps M invariant. The factorial part coming from relabelling the vertices is equivalent to re-arranging the rows in M . Let P be an $n \times n$ orthogonal matrix that permutes the rows in M bringing it to a new matrix $\tilde{M} := (\tilde{A}|\tilde{B}|\tilde{C}) = PM$. This is equivalent to setting $\tilde{A} = PA$, $\tilde{B} = PB$, and $\tilde{C} = PC$. Clearly, $\tilde{J} = \tilde{A}^T \tilde{A} = J$, $\tilde{A}^T \tilde{B} = A^T B$, and $\tilde{A}^T \tilde{C} = A^T C$. Thus $\tilde{\alpha} = \alpha$ and $\tilde{\beta} = \beta$.

The frequency labels can be not unique for a given diagram due to the several possible options of our choice of A and equivalently C . In graph theory language, this has to do with the existence of several *directed trees* that are consisting of $n - 1$ edges connecting the n vertices. Regardless of this starting choice, the above theorem implies that all of the sibling diagrams in the same topology will always have the same frequency labels once the labels are fixed for one diagram [the AMI input matrix Eq. (10)].

-
- [1] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).
- [2] A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Chem. Rev.* **112**, 289 (2012).
- [3] N. V. Prokof'ev and B. V. Svistunov, *Phys. Rev. B* **77**, 125101 (2008).
- [4] K. Van Houcke, F. Werner, E. Kozik, N. Prokof'ev, B. Svistunov, M. J. H. Ku, A. T. Sommer, L. W. Cheuk, A. Schirotzek, and M. W. Zwierlein, *Nat. Phys.* **8**, 366 (2012).
- [5] K. Chen and K. Haule, *Nat. Commun.* **10**, 3725 (2019).
- [6] E. Kozik, K. V. Houcke, E. Gull, L. Pollet, N. Prokof'ev, B. Svistunov, and M. Troyer, *Europhys. Lett.* **90**, 10004 (2010).
- [7] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, *Phys. Rev. B* **72**, 035122 (2005).
- [8] E. Burovski, N. Prokof'ev, B. Svistunov, and M. Troyer, *Phys. Rev. Lett.* **96**, 160402 (2006).
- [9] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).
- [10] R. Rossi, *Phys. Rev. Lett.* **119**, 045701 (2017).
- [11] J. Li, M. Wallerberger, and E. Gull, *Phys. Rev. Res.* **2**, 033211 (2020).
- [12] F. Šimkovic IV and M. Ferrero, *Phys. Rev. B* **105**, 125104 (2022).
- [13] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, *Phys. Rev. B* **99**, 035120 (2019).
- [14] H. Elazab, B. McNiven, and J. LeBlanc, *Comput. Phys. Commun.* **280**, 108469 (2022).
- [15] B. D. E. McNiven, G. T. Andrews, and J. P. F. LeBlanc, *Phys. Rev. B* **104**, 125114 (2021).
- [16] B. D. E. McNiven, H. Terletska, G. T. Andrews, and J. P. F. LeBlanc, *Phys. Rev. B* **106**, 035145 (2022).
- [17] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, *Phys. Rev. B* **102**, 045115 (2020).
- [18] M. D. Burke, M. Grandadam, and J. P. F. LeBlanc, *Phys. Rev. B* **107**, 115151 (2023).
- [19] I. S. Tupitsyn, A. M. Tsvelik, R. M. Konik, and N. V. Prokof'ev, *Phys. Rev. Lett.* **127**, 026403 (2021).
- [20] J. P. F. LeBlanc, K. Chen, K. Haule, N. V. Prokof'ev, and I. S. Tupitsyn, *Phys. Rev. Lett.* **129**, 246401 (2022).
- [21] R. Levy, J. P. F. LeBlanc, and E. Gull, *Comput. Phys. Commun.* **215**, 149 (2017).
- [22] F. Šimkovic IV, J. P. F. LeBlanc, A. J. Kim, Y. Deng, N. V. Prokof'ev, B. V. Svistunov, and E. Kozik, *Phys. Rev. Lett.* **124**, 017003 (2020).
- [23] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, *Phys. Rev. B* **101**, 125109 (2020).
- [24] R. Farid, M. Grandadam, and J. P. F. LeBlanc, *Phys. Rev. B* **107**, 195138 (2023).
- [25] J. Vučićević, P. Stipsić, and M. Ferrero, *Phys. Rev. Res.* **3**, 023082 (2021).
- [26] R. Rossi, F. Werner, N. Prokof'ev, and B. Svistunov, *Phys. Rev. B* **93**, 161102(R) (2016).
- [27] R. Rossi, T. Ohgoe, K. Van Houcke, and F. Werner, *Phys. Rev. Lett.* **121**, 130405 (2018).
- [28] S. Hirata, M. R. Hermes, J. Simons, and J. V. Ortiz, *J. Chem. Theory Comput.* **11**, 1595 (2015).
- [29] A. Moutenet, W. Wu, and M. Ferrero, *Phys. Rev. B* **97**, 085117 (2018).
- [30] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).
- [31] J. V. Roth, D. Schweitzer, L. J. Sieke, and L. von Smekal, *Phys. Rev. D* **105**, 116017 (2022).
- [32] M. Caffarel and W. Krauth, *Phys. Rev. Lett.* **72**, 1545 (1994).

- [33] N. E. Dahlen and R. van Leeuwen, *J. Chem. Phys.* **122**, 164102 (2005).
- [34] J. J. Phillips and D. Zgid, *J. Chem. Phys.* **140**, 241101 (2014).
- [35] L. Zhang and E. Gull, [arXiv:2312.10576](https://arxiv.org/abs/2312.10576).
- [36] J. Fei, C.-N. Yeh, D. Zgid, and E. Gull, *Phys. Rev. B* **104**, 165111 (2021).
- [37] A. Gaenko, A. E. Antipov, G. Carcassi, T. Chen, X. Chen, Q. Dong, L. Gamper, J. Gukelberger, R. Igarashi, S. Isakov *et al.*, *Comput. Phys. Commun.* **213**, 235 (2017).
- [38] M. Wallerberger, S. Isakov, A. Gaenko, J. Kleinhenz, I. Krivenko, R. Levy, J. Li, H. Shinaoka, S. Todo, T. Chen, X. Chen, J. P. F. LeBlanc, J. E. Paki, H. Terletska, M. Troyer, and E. Gull, [arXiv:1811.08331](https://arxiv.org/abs/1811.08331).