Computer Physics Communications 253 (2020) 107178

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

# Generalized quadrature for finite temperature Green's function methods☆

Jie Gu [a], Jia Chen [a], Yang Wang [b], X.-G. Zhang [a,*]

[a] Department of Physics, the Quantum Theory Project and the Center for Molecular Magnetic Quantum Materials, University of Florida, Gainesville, USA
[b] Pittsburgh Supercomputing Center, Pittsburgh, USA

ARTICLE INFO

*Article history:*
Received 15 July 2019
Received in revised form 22 November 2019
Accepted 17 January 2020
Available online 24 January 2020

*Keywords:*
Gaussian quadrature
Green's function
Fermi–Dirac distribution function
Finite temperature

ABSTRACT

In electronic structure and quantum transport calculations, many physical quantities are integrations over the electron energy, weighted by the Fermi–Dirac distribution function. Green's function based approaches commonly circumvent the numerically difficult real energy integration by extending the integrand analytically into the complex energy plane, and using a Gaussian quadrature integration over a complex energy contour for zero temperature. For finite temperatures, a much slower convergent sum over the Matsubara frequencies is necessary. We present a generalized quadrature method that uses orthogonal polynomials on the manifold of Matsubara frequencies to enable rapid convergence. Both Gaussian quadrature integration and Matsubara frequency summation methods are shown to be limiting cases of the generalized method. Tests on an all-electron *ab initio* code and total energy calculation of interacting Anderson impurity model show convergence with a small number of energy mesh points.



## 1. Introduction

Green's function is essential to many theoretical methods in condensed matter physics. In such methods, evaluations of the total energy and other physical quantities usually require integrals over the electron energy of the Green's function weighted by the Fermi–Dirac distribution function. For example, in the Green's function approach to the *ab initio* calculation of electron density $\rho$, the following integral over the energy parameter $\epsilon$ is evaluated,

$$\rho(\mathbf{r}) = -\frac{2}{\pi}\mathrm{Im}\int_{-\infty}^{\infty} f(\epsilon, T)G(\mathbf{r}, \mathbf{r}; \epsilon)d\epsilon \tag{1}$$

where $f(\epsilon, T) = 1/(1 + e^{(\epsilon-\mu)/k_BT})$ is the Fermi–Dirac distribution function, $\mu$ the chemical potential, $T$ the temperature, and $G$ the Green's function of the Kohn–Sham equation. Such a calculation is needed in Green's function based *ab initio* electronic structure and transport calculations, such as the Korringa–Kohn–Rostoker (KKR) electronic structure method [1–3], the GW method [4], and the nonequilibrium Green's Function (NEGF) method for electron transport [5].

In finite-temperature many-body calculations, observables are usually calculated from the Matsubara Green's function,

$$k_BT\sum_n G(i\Omega_n) = \frac{1}{2\pi i}\int_C dzG(z)f(z). \tag{2}$$

The Matsubara frequencies arise from the poles $z = i\Omega_n$ of the Fermi–Dirac function $f(z)$. This allows sum over the discrete Matsubara frequencies $\Omega_n = (2n + 1)\pi k_BT$ for fermions to be evaluated as a contour integral [6]. In Migdal–Eliashberg theory [7,8] for phonon driven superconductivity and dynamical mean-field theory [9] (DMFT) for strongly correlated electrons, equations are constructed to calculate Matsubara Green's function. Those techniques have been extended to *ab-initio* electronic structure theories. A case in point is dynamical mean-field theory combined with density functional theory (DFT+DMFT) [10], KKR (KKR+DMFT) [11], or exact muffin-tin orbitals (EMTO+ DMFT) [12], which have been indispensable tools for strongly correlated materials. Recently, Matsubara Green's function based methods have also been applied to finite systems [13,14], which makes it useful for quantum chemistry.

To speed up calculations involving Eqs. (1) and (2), several approximate Fermi–Dirac functions have been proposed [15–24]. Among these approximations, Nicholson and Zhang [16] provided a family of approximate distribution functions that accurately represent the Fermi–Dirac function over the valence band region.

☆ The review of this paper was arranged by Prof. D.P. Landau.

* Corresponding author.
*E-mail address:* xgz@ufl.edu (X.-G. Zhang).

https://doi.org/10.1016/j.cpc.2020.107178
0010-4655/

By approximating in an appropriate manner, e.g., $f(E) \approx f_N(E)$, with $N$ indicating an $N$th order approximation of $f(E)$, one can approximate the infinite Matsubara frequency sum with a finite sum,

$$\int_C dz G(z) f(z) \approx \int_C dz G(z) f_N(z) = 2\pi i \sum_{n=1}^{N} G(i\epsilon_n) w_n \tag{3}$$

where $i\epsilon_n$ are the poles of $f_N(E)$ and $w_n$ the corresponding residual. It was demonstrated on a model system for which the calculated free energy at finite temperature is given accurately. For low temperatures, when a large number of Matsubara frequencies needs to be summed to yield accurate results, approximating with $f_N$ may only partly alleviate the difficulty because $N$ scales inversely with $T$ and may also be large.

Inspired by the commonly used Gaussian quadrature method for continuous integration for zero temperature [25], here we devise a generalized quadrature using orthogonal polynomials over a discrete sum rather than a continuous integration for finite temperature calculations. The quadrature can be applied to the discrete Matsubara sum and reduce the number of terms by orders of magnitude without sacrificing accuracy. A similar quadrature can also be devised for the Nicholson–Zhang method [16] to remove the dependence of the number of mesh points on the temperature.[1] We demonstrate the power and accuracy of the method with two numerical examples, KKR calculations of the total energy in fcc Cu and bcc Fe, and a numerical solution of the Anderson impurity model.

## 2. Quadrature rule

### 2.1. Generalized quadrature rule for a discrete sum

We look for an approximation to the discrete sum $\sum_n^N G(\phi_n) W(\phi_n)$ where $W$ is a weight function, analogous to the Gaussian quadrature rule for a one-dimensional integral $\int_0^1 G(\phi) W(\phi) d\phi$. Similar to the Gaussian quadrature rule, we will generate the quadrature from a set of orthogonal polynomials. To do this, we define an inner product,

$$\langle F(\phi), H(\phi) \rangle \equiv \sum_n^N F(\phi_n) H(\phi_n) W(\phi_n). \tag{4}$$

The construction of the orthogonal polynomials is similar to the Gaussian quadrature method. The zeroth and first order polynomials are

$$P_0(\phi) = 1, \tag{5}$$

and

$$P_1(\phi) = \phi - \frac{\langle \phi, P_0 \rangle}{\langle P_0, P_0 \rangle}, \tag{6}$$

Higher order polynomials are obtained using the recursion relation,

$$P_{n+1}(\phi) = (\phi - a_n) P_n - b_n P_{n-1}, \tag{7}$$

where

$$a_n = \frac{\langle \phi P_n, P_n \rangle}{\langle P_n, P_n \rangle}, \quad b_n = \frac{\langle P_n, P_n \rangle}{\langle P_{n-1}, P_{n-1} \rangle}. \tag{8}$$

[1] Subroutines are available at https://github.com/xgzhanggroup/QuadratureFermi.

Once the coefficients are determined, the Golub–Welsch algorithm can be used to find the quadrature rule [26]. A symmetric tridiagonal matrix $J$ is constructed using $a_n$ and $b_n$,

$$J = \begin{pmatrix} a_0 & \sqrt{b_1} & 0 & \cdots & \cdots & \cdots \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & 0 & \cdots & \cdots \\ 0 & \sqrt{b_2} & a_2 & \sqrt{b_3} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & \sqrt{b_{n-2}} & a_{n-2} & \sqrt{b_{n-1}} \\ \cdots & \cdots & \cdots & 0 & \sqrt{b_{n-1}} & a_{n-1} \end{pmatrix} \tag{9}$$

The matrix is diagonlized and the quadrature points $\phi_j'$'s are the eigenvalues. If $\eta^{(j)}$ is the normalized eigenvector corresponding to $\phi_j'$, a weight $w_j$ is defined in terms of the first component of $\eta^{(j)}$:

$$w_j = \left(\eta_1^{(j)}\right)^2 \sum_n^N W(\phi_n). \tag{10}$$

An $N_Q$th order orthogonal polynomial yields the quadrature rule,

$$\sum_n^N G(\phi_n) W(\phi_n) \approx \sum_j^{N_Q} G\left(\phi_j\right) w_j, \quad N_Q < N, \tag{11}$$

where $\phi_j$ and $w_j$ are the quadrature points and weights.

### 2.2. Imaginary frequency sum

We first apply the generalized quadrature rule to reduce the infinite imaginary frequency sum,

$$\sum_{n=0}^{\infty} G(i\Omega_n), \tag{12}$$

with a small number of terms, while preserving sufficient numerical accuracy. Because this is an infinite sum, the generalized quadrature rule cannot be applied directly with $W = 1$.

We assume that the sum is polynomial convergent, i.e., for large $n$,

$$G(i\Omega_n) \propto \frac{1}{\Omega_n^{1+\varepsilon}}, \tag{13}$$

with $\varepsilon > 0$ to ensure convergence of the sum. The weight function is defined as $W(i\Omega_n) = 1/\Omega_n^{1+\varepsilon}$, and

$$\sum_{n=0}^{\infty} G(i\Omega_n) = \sum_{n=0}^{n_0} G(i\Omega_n) + \sum_{n=n_0+1}^{\infty} \bar{G}(\phi_n) \frac{1}{\Omega_n^{1+\varepsilon}}, \tag{14}$$

where $\bar{G}(\phi_n) = G(i\Omega_n)\Omega_n^{1+\varepsilon}$ is defined as a function of $\phi_n = 1/\Omega_n^{\varepsilon}$ and the weight function $\bar{W}(\phi_n) = W(i\Omega_n) = \phi_n^{(1+\varepsilon)/\varepsilon}$. The first $n_0$ terms are summed as is, in order to avoid the numerical instability due to the dominance of the first few terms, and the generalized quadrature is now applied with $\phi_n$, starting from $n = n_0 + 1$. The quadrature points $\Omega_j'$ and weights $w_j$ are

$$\Omega_j' = (\phi_j')^{-1/\epsilon} \tag{15}$$

$$w_j = \left(\eta_1^{(j)}\right)^2 (\Omega_j')^{1+\varepsilon} \sum_{n=n_0+1}^{\infty} \frac{1}{\Omega_n^{1+\varepsilon}} \tag{16}$$

where $\phi_j'$ and $\eta_1^{(j)}$ are obtained as described in Section 2.1.

The summation to find the inner product defined in Eq. (4) can be truncated at a cutoff $N$, whose error can be estimated from

$$\int_N^{\infty} \frac{1}{x^{1+\varepsilon}} dx = \frac{1}{\varepsilon N^{\varepsilon}} = \eta, \tag{17}$$

where $\eta$ is the tolerance. For $\varepsilon = 1$, $N = 1/\eta$. If the desired $\eta$ is small, it requires a large $N$ and the summations in the inner products may become computationally slow. For such cases we need an approximate method to evaluate the truncated terms to allow a smaller $N$. We note that the discrete sum beyond truncation can be converted into an integral using the trapezoidal rule,

$$\begin{aligned}&\sum_{n=N}^{\infty} F(\phi_n)W(\phi_n),\\ &= \frac{1}{2}F(\phi_N)W(\phi_N) + \frac{1}{2\varepsilon\Omega_0}\int_0^{\phi_N} F(\phi)d\phi\\ &= \frac{1}{2}F(\phi_N)W(\phi_N) + \frac{1}{2\varepsilon\Omega_0}\left[F(\phi_N)\phi_N + \frac{1}{2}\frac{dF(\phi_N)}{d\phi}\phi_N^2\right].\end{aligned} \tag{18}$$

Likewise,

$$\begin{aligned}&\sum_{n=N}^{\infty} \phi F(\phi_n)W(\phi_n)\\ &= \frac{1}{2}\phi_N F(\phi_N)W(\phi_N)\\ &+ \frac{1}{2\varepsilon\Omega_0}\left[\frac{1}{2}F(\phi_N)\phi_N^2 + \frac{1}{3}\frac{dF(\phi_N)}{d\phi}\phi_N^3\right].\end{aligned} \tag{19}$$

### 2.3. Generalized complex quadrature

More generally, one can first reduce the infinite Matsubara frequency sum to a finite sum, then apply the generalized quadrature rule to further reduce the number of terms. Recall the approximate Fermi–Dirac distribution function used in Ref. [16],

$$f_N^{-1} - 1 = \frac{\left[1 + x\left(\frac{1+\gamma}{4N}\right)\right]^{2N}}{\left[1 - x\left(\frac{1-\gamma}{2N}\right)\right]^{N}}, \tag{20}$$

where $\gamma = 3 - \sqrt{8}$ and $x = (\epsilon - \mu)/k_BT$. $f_N$ approaches the exact Fermi–Dirac function as $N \to \infty$. When $N$ is finite, $f_N$ significantly deviates from the exact function as $x(1-\gamma)/(2N) \to 1$. Therefore, the necessary condition of a good approximation is,

$$N \gtrsim \frac{W(1-\gamma)}{2k_BT} \approx \frac{0.4W}{k_BT}, \tag{21}$$

where $W$ is the bandwidth.

When the integral along the semicircle $z = Re^{i\theta}(R \to \infty)$ vanishes, the integral containing the approximate Fermi–Dirac function is equal to a sum over the residues at corresponding poles, given by

$$\int_{-\infty}^{\infty} f_N(x)\, g(x)\, dx = 2\pi i \sum_n^N g(z_n)w(z_n), \tag{22}$$

where $g(x)$ is an integrand which has no poles in the upper half plane, and $z_n$ and $w(z_n)$ are the poles and residues of $f_N$ respectively.

The poles of $f_N$ (called Nicholson–Zhang poles hereafter) can be easily calculated by solving the quadratic equation

$$\frac{\left[1 + z\left(\frac{1+\gamma}{4N}\right)\right]^2}{1 - z\left(\frac{1-\gamma}{2N}\right)} = e^{2i\phi_n}, \tag{23}$$

where $\phi_n = (2n-1)\pi/(2N), n = 1, 2, \ldots, N$. $N$ is the number of poles in the upper half plane, and also the number of terms in the sum. Explicitly,

$$z_n = \frac{2(c-1)}{\sqrt{(2a+bc)^2 + 4a^2(c-1)} \pm (2a+bc)}. \tag{24}$$

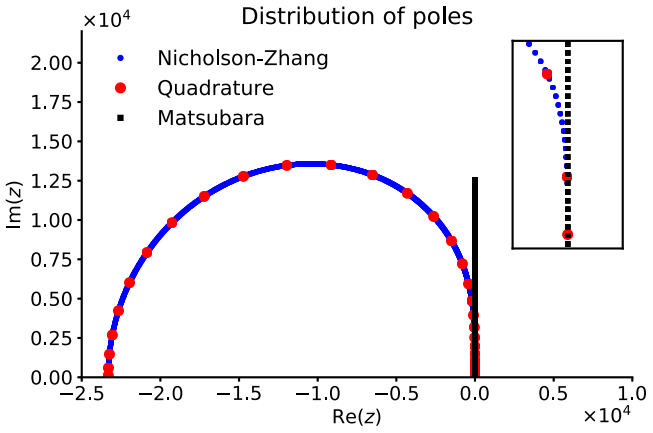


**Fig. 1.** Distribution of poles for Matsubara sum, Nicholson–Zhang's method and the generalized complex quadrature. The inset enlarges the region near the origin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**
Asymptotics of poles and residuals in the large $N$ limit.

| Sign | Pole | Residual |
|---|---|---|
| + | $i(2j-1)\pi + O(1/N)$ | $1 + O(1/N)$ |
| − | $O(N)$ | $O(1)$ |

where $a = (1+\gamma)/(4N)$, $b = (1-\gamma)/(2N)$, $c = \exp(2i\phi_n)$.

The Nicholson–Zhang poles for $N = 2000$ are shown in Fig. 1 (blue dots). They are so dense that a line is formed visually. A comparison with the first 2000 Matsubara poles (black squares) is also plotted.

A fast approximate approach to calculate the sum in Eq. (22) is desired in the form below,

$$\sum_n^N g(z_n)\, w(z_n) \approx \sum_j^{N_Q} g\left(z_j'\right) w_j', \quad N_Q \ll N, \tag{25}$$

where $z_n$ and $w(z_n)$ are respectively the poles and residues of the approximate Fermi distribution function, and the variables with prime define the quadrature, which should not depend on the form of the integrand $g$. $N_Q$ is the order of the generalized quadrature.

To find the quadrature, we change the variable $z_n$ to $\phi_n$ using $z_n = z(\phi_n)$. The left hand side of Eq. (25) becomes a function of $\phi_n$ only. This allows us to evaluate $\phi_j$ and $w_j$ using the method described in Section 2.1,

$$\begin{aligned}\sum_n^N g(z_n)\, w(z_n) &= \sum_j^{N_Q} g\left(z\left(\phi_j\right)\right) w(z(\phi_j))w_j\\ &= \sum_j^{N_Q} g\left(z_j'\right) w_j',\end{aligned} \tag{26}$$

where $z_j' \equiv z\left(\phi_j\right)$ and $w_j' = w(z_j)w_j$ are the quadrature points and weights needed in Eq. (25). These quadrature points lie on the same arc with the Nicholson–Zhang poles, as shown in Fig. 1 (red dots).

#### 2.3.1. Large N limit at finite temperature

For terms with $j \ll N$, the asymptotics of the poles and residuals are given in Table 1 according to the sign in the denominator of the poles (see Eq. (24)). Since we require the integral along the semicircle $z = Re^{i\theta}(R \to \infty)$ to vanish, the poles corresponding to the negative sign (bottom row in Table 1) do not contribute

to the sum. Therefore, in the large $N$ limit effectively all of the Nicholson–Zhang poles converge to the Matsubara poles.

*2.3.2. Zero-temperature limit*

At zero temperature, the continuous integral over energy $\int_{-\infty}^{\epsilon_F} g(\epsilon)d\epsilon$ is evaluated, where $\epsilon_F$ is the Fermi energy. The standard method to calculate this integral is using the residual theorem to transform the integral on the real axis to that on a semicircular contour connecting the valence band bottom and the Fermi energy, then Gauss–Legendre quadrature can be used (see, e.g., Ref. [25]), i.e.,

$$\begin{aligned}\int_{-\infty}^{\epsilon_F} g(\epsilon)d\epsilon &= -\int_0^{\pi} g(z(\phi))z'(\phi)d\phi \\ &= -\sum_j^{N_Q} g(z(\phi_j))z'(\phi_j)w_j \end{aligned} \tag{27}$$

where $z(\phi) = R_0 + Re^{i\phi}$, $\phi_j$ and $w_j$ are the Gaussian–Legendre quadrature rule.

Here we show that in the zero-temperature limit our method is equivalent to the conventional method except for a slightly different complex energy contour. Fixing $Nk_BT = W(1-\gamma)/2$ (see Eq. (21)), where $W$ is the bandwidth, and taking the limit $T \to 0$ and $N \to \infty$, the inner product defined in Eq. (4) has the limit,

$$\begin{aligned}\langle F(\phi), H(\phi)\rangle &= \frac{N}{\pi}\sum_i^N \frac{\pi}{N} F(\phi_i) H(\phi_i) \\ &\to \frac{N}{\pi}\int_0^{\pi} FHd\phi. \end{aligned} \tag{28}$$

Since all coefficients in polynomials are defined in terms of the ratios of inner products, the constant $N/\pi$ is irrelevant here. $\int_0^{\pi} FHd\phi$ is exactly the inner product defined for Gauss–Legendre quadrature over the interval $[0, \pi]$, so are all of the coefficients in these polynomials. Thus, for $N \to \infty$, the quadrature points $\phi_j$ are exactly their counterparts in Gauss–Legendre quadrature over the interval $[0, \pi]$, with the weights $w_j$ having a prefactor of $N/\pi$. It is easy to verify that

$$w(z(\phi_j)) = -\frac{z'(\phi_j)}{2Ni}. \tag{29}$$

Taking into account the prefactor of $N/\pi$ and $2\pi i$ for residual theorem and comparing Eqs. (26) and (27), the generalized complex quadrature in the zero temperature limit is equivalent to the conventional method except for a different function $z(\phi)$ and the corresponding complex energy contour.

## 3. Numerical examples

### 3.1. Anderson impurity model

To showcase this method in interacting systems, we apply the method described in Section 2.2, imaginary frequency sum, to the Anderson impurity model (AIM) [27]. For our purpose, the AIM is simple enough to yield exact ground state energy and Matsubara Green's function by exact diagonalization method [28], allowing us to benchmark the generalized quadrature method. In addition, it offers a non-trivial test with similar numerical procedures found in other many-body calculations based on Matsubara Green's function.

Here, we consider the AIM with one impurity orbital and four bath sites,

$$\begin{aligned}\mathcal{H}_{\text{imp}} = &\sum_{\sigma} \epsilon_c c_\sigma^\dagger c_\sigma + U n_\uparrow n_\downarrow \\ &+ \sum_{l,\sigma}\left(V_{l,\sigma} a_{l,\sigma}^\dagger c_\sigma + V_{l,\sigma}^* c_\sigma^\dagger a_{l,\sigma}\right) + \sum_{l,\sigma} \epsilon_l a_{l,\sigma}^\dagger a_{l,\sigma} \end{aligned}. \tag{30}$$

Matsubara Green's function of the impurity orbital can be constructed as,

$$\begin{aligned}\mathcal{G}_\sigma(i\Omega_n) = &\langle 0|c_\sigma \frac{1}{i\Omega_n + E_0 - \mathcal{H}} c_\sigma^\dagger|0\rangle \\ &+ \langle 0|c_\sigma^\dagger \frac{1}{i\Omega_n - E_0 + \mathcal{H}} c_\sigma|0\rangle, \end{aligned} \tag{31}$$

where $|0\rangle$ is ground state vector in Fock space and $E_0$ is ground state energy.

Non-interacting Green's function can be calculated by,

$$\mathcal{G}^0_{\mu,\nu}(i\Omega_n) = (i\Omega_n - \mathcal{H}_0)^{-1}, \tag{32}$$

where $\mathcal{H}_0$ is the impurity Hamiltonian without $Un_\uparrow n_\downarrow$ term.

Self-energy of the impurity orbital can be calculated as,

$$\Sigma(i\Omega_n) = \frac{1}{\mathcal{G}^0(i\Omega_n)} - \frac{1}{\mathcal{G}(i\Omega_n)}. \tag{33}$$

It is important in calculations to separate the self-energy into two parts, the static part which is the asymptotic limit of self-energy, and the frequency dependent part called dynamical self-energy [29],

$$\Sigma(i\Omega_n) = \Sigma^\infty + \Sigma^{\text{dyn}}(i\Omega_n). \tag{34}$$

For our model, the static part of self-energy is the Hartree potential of electron–electron interaction,

$$\Sigma^\infty = \frac{n_\uparrow + n_\downarrow}{2} \times U. \tag{35}$$

Total energy of the system can be calculated based on the Matsubara Green's function and the self-energy,

$$\begin{aligned}\langle \mathcal{H}\rangle = &\text{Tr}[\mathcal{H}_0\mathcal{P}] + \Sigma^\infty \frac{n_\uparrow + n_\downarrow}{2} \\ &+ \text{Re}\frac{1}{\beta}\sum_n \mathcal{G}(i\Omega_n)\Sigma^{\text{dyn}}(i\Omega_n), \end{aligned} \tag{36}$$

where $\mathcal{P}$ is the single particle density matrix, which is usually calculated as a frequency sum of the Green's function $\mathcal{P} = \sum_n \mathcal{G}(i\Omega_n)/\beta$. In this work, to avoid error due to sum of Matsubara frequency, we used the exact ground state to calculate the single particle density matrix. Usually, the last term has to be calculated as sum of Matsubara frequency. Here, we applied the generalized quadrature method to evaluate this term,

$$E^{\text{dyn}} = \text{Re}\frac{1}{\beta}\sum_j^{N_Q} \mathcal{G}(z_j')\Sigma^{\text{dyn}}(z_j')w_j', \tag{37}$$

where quadrature points $z_j'$ and weights $w_j'$ were obtained using the imaginary frequency sum with $\varepsilon = 1$.

To make sure that the excited states do not make contribution to $\langle\mathcal{H}\rangle$, we use a very low temperature $\beta = 512$. All hoppings between the impurity and the bath are set to $-1.0$; on-site energies of the four bath orbitals are $\pm 1.0$ and $\pm 2.0$. When $U = 3.0$ and impurity on-site energy is $\epsilon_c = -1.5$, the impurity is half-filled. To stay away from half-filling, parameters of $U = 2.0$, $\epsilon_c = -0.5$ are chosen so we have four electrons in the model and the occupation number of the impurity is $n = 0.67$.

The total energy is calculated by exact diagonalization with summations over the Matsubara frequencies and the quadrature points via Eqs. (36) and (37). Absolute error of the total energy is plotted in Fig. 2. As one can see, in order to achieve accuracy around $1.0 \times 10^{-5}$, only 20 quadrature points are necessary in addition to the first $n_0 = 10$ Matsubara frequencies which are summed directly. In contrast, to reach the same level of accuracy using direct summation over Matsubara frequencies, about $10^7$ points are needed. Thus we conclude that the generalized quadrature method can greatly speed up finite-temperature many-body calculations.

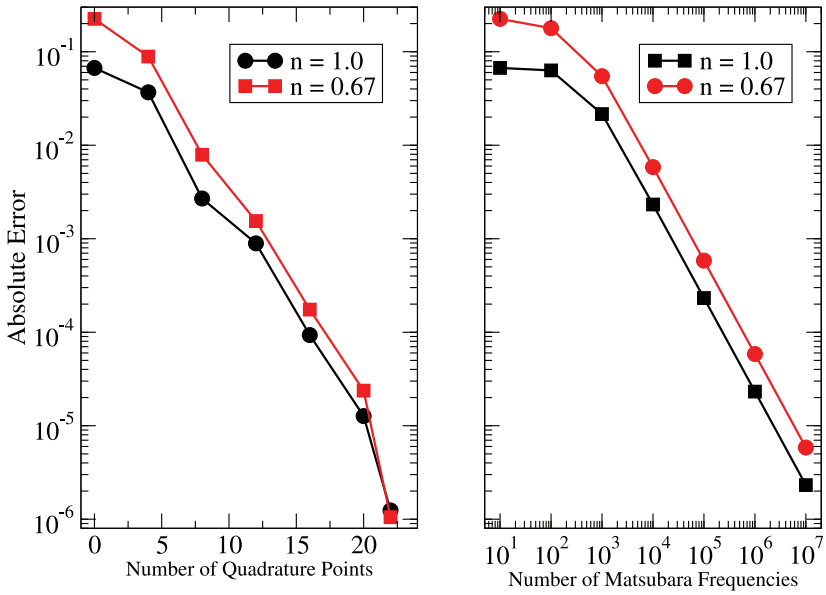


**Fig. 2.** (left) Total energy absolute error of impurity model as a function of the number of quadrature points; (right) total energy absolute error of impurity model as a function of the number of Matsubara frequencies.

**Table 2**
The total wall clock time ($t$) of self-consistent field (SCF) *ab initio* electronic structure calculations for Cu and Fe at finite temperatures. In the table, $N$ is the number of poles used in the Nicholson–Zhang approximate distribution function, and $N_Q$ is the number of quadrature points used in Eq. (25). The wall clock time is determined by the number of poles or the number of quadrature points used in each SCF iteration, as well as the total number of iterations to achieve convergence.

| $T$(K) | Cu (LSMS) | | | | Fe (spin-polarized, KKR) | | | |
|---|---|---|---|---|---|---|---|---|
| | Nicholson–Zhang | | Quadrature | | Nicholson–Zhang | | Quadrature | |
| | $N$ | $t$ (s) | $N_Q$ | $t$ (s) | $N$ | $t$ (s) | $N_Q$ | $t$ (s) |
| 300 | 199 | 539.993 | 15 | 153.313 | 199 | 1468.811 | 15 | 110.520 |
| 1000 | 61 | 200.571 | 10 | 100.038 | 61 | 521.090 | 10 | 100.673 |
| 1500 | 41 | 147.253 | 10 | 100.223 | 41 | 335.802 | 10 | 125.965 |

### 3.2. Ab-initio code

The generalized complex quadrature method described in Section 2.3 has been implemented in the MST2 package [30], an all-electron *ab initio* code based on multiple scattering theory capable of performing full-potential KKR or linear scaling LSMS electronic structure calculations [31]. This will enable the *ab initio* code to calculate the electronic free energy and other physical quantities at finite temperatures.

To compare the code performance for using the Nicholson–Zhang method and the quadrature method for finite temperature calculation, we show the timing of the self-consistent calculation for fcc Cu and bcc Fe at temperatures 300 K, 1000 K, and 1500 K. Specifically, we performed LSMS calculation for fcc Cu, and spin-polarized KKR calculation for bcc Fe. In these finite temperature calculations, the integral in Eq. (1) is reduced to a summation of the Green functions either evaluated at the Nicholson–Zhang poles or the quadrature points. The results are presented in Table 2, in which the timing is the wall clock time for each job running on 8 CPU cores. For the LSMS calculation for Cu, the local interaction zone size is chosen to be 88 atoms, which is sufficient to converge the calculations with total energy difference $<5 \times 10^{-8}$ Ry and the rms of the LDA potential is less than $1 \times 10^{-7}$ Ry. For the spin-polarized KKR calculation for Fe, the number of special k-points in the irreducible Brillouin zone is 140, which is sufficient to converge the calculations with total energy difference $<5 \times 10^{-9}$ Ry and the rms of the LDA potential $<1 \times 10^{-8}$ Ry. The computational results show that the quadrature method improves the code performance at low temperatures, due to the fact that a significant number of Nicholson–Zhang poles ($N$) are required at low temperatures, compared to much smaller number of quadrature points ($N_Q$) need to be used in order to achieve a similar accuracy.

## 4. Conclusion

In summary, two new integration methods containing Fermi–Dirac distribution function, which extend the Gaussian quadrature method to discrete points in the complex energy plane, have been presented. Their efficiency and accuracy have been demonstrated in total energy calculation of interacting Anderson impurity model and in the all-electron *ab initio* code MST2 package respectively. These methods will allow for *ab initio* calculation of the electronic free energy at finite temperatures, especially for codes to have better performance and numerical stability when the material under investigation exhibits large density of states fluctuation near the Fermi energy. The quadrature is general and can be applied to other approximate Fermi–Dirac functions, and even integrals involving functions other than the Fermi–Dirac function.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Jie Gu:** Methodology, Validation, Writing - original draft. **Jia Chen:** Validation, Writing - original draft. **Yang Wang:** Validation, Writing - original draft. **X.-G. Zhang:** Conceptualization, Methodology, Writing - review & editing.

## Acknowledgments

This work was supported as part of the Center for Molecular Magnetic Quantum Materials ($M^2QM$), an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0019330. Part of this work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation, USA grant number ACI-1548562. Specifically, it used the Bridges system, which is supported by National Science Foundation, USA award number ACI-1445606, at the Pittsburgh Supercomputing Center (PSC). The authors acknowledge University of Florida Research Computing for providing computational resources and support that have contributed to the research results reported in this publication.

## References

[1] J. Korringa, Physica 13 (6–7) (1947) 392–400.
[2] W. Kohn, N. Rostoker, Phys. Rev. 94 (5) (1954) 1111.
[3] J. MacLaren, X.-G. Zhang, W. Butler, X. Wang, Phys. Rev. B 59 (8) (1999) 5470.
[4] L. Hedin, Phys. Rev. 139 (3A) (1965) A796.
[5] L.V. Keldysh, et al., Sov. Phys. J. Exp. Theor. Phys. 20 (4) (1965) 1018–1026.
[6] P. Coleman, Introduction to Many-Body Physics, Cambridge University Press, 2017.
[7] A.B. Migdal, J. Exp. Theor. Phys. 7 (6) (1958) 996–1000.
[8] G. Eliashberg, J. Exp. Theor. Phys. 11 (3) (1960) 696–702.
[9] A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Modern Phys. 68 (1996) 13–125, http://dx.doi.org/10.1103/RevModPhys.68.13, URL https://link.aps.org/doi/10.1103/RevModPhys.68.13.
[10] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Modern Phys. 78 (2006) 865–951, http://dx.doi.org/10.1103/RevModPhys.78.865, URL https://link.aps.org/doi/10.1103/RevModPhys.78.865.
[11] J. Minar, L. Chioncel, A. Perlov, H. Ebert, M. Katsnelson, A. Lichtenstein, Phys. Rev. B 72 (4) (2005) 045125.
[12] L. Chioncel, L. Vitos, I. Abrikosov, J. Kollar, M. Katsnelson, A. Lichtenstein, Phys. Rev. B 67 (23) (2003) 235106.
[13] N. Lin, C.A. Marianetti, A.J. Millis, D.R. Reichman, Phys. Rev. Lett. 106 (2011) 096402.
[14] D. Neuhauser, R. Baer, D. Zgid, J. Chem. Theory Comput. 13 (11) (2017) 5396–5403, http://dx.doi.org/10.1021/acs.jctc.7b00792, pMID: 28961398.
[15] S. Goedecker, Phys. Rev. B 48 (23) (1993) 17573–17575, http://dx.doi.org/10.1103/PhysRevB.48.17573, URL https://link.aps.org/doi/10.1103/PhysRevB.48.17573.
[16] D. Nicholson, X.-G. Zhang, Phys. Rev. B 56 (20) (1997) 12805–12810, http://dx.doi.org/10.1103/PhysRevB.56.12805.
[17] F. Gagel, J. Comput. Phys. 405 (1997) 6, http://dx.doi.org/10.1006/jcph.1997.5871.
[18] T. Ozaki, Phys. Rev. B 75 (3) (2007) 035123, http://dx.doi.org/10.1103/PhysRevB.75.035123, URL https://link.aps.org/doi/10.1103/PhysRevB.75.035123.
[19] A. Croy, U. Saalmann, Phys. Rev. B 80 (7) (2009) 073102, http://dx.doi.org/10.1103/PhysRevB.80.073102, URL https://link.aps.org/doi/10.1103/PhysRevB.80.073102.
[20] L. Lin, J. Lu, L. Ying, E. Weinan, Chinese Ann. Math. Ser. B 30 (6) (2009) 729–742, http://dx.doi.org/10.1007/s11401-009-0201-7, URL http://link.springer.com/10.1007/s11401-009-0201-7.
[21] C. Karrasch, V. Meden, K. Schönhammer, Phys. Rev. B 82 (12) (2010) 125114, http://dx.doi.org/10.1103/PhysRevB.82.125114, URL https://link.aps.org/doi/10.1103/PhysRevB.82.125114.
[22] J. Hu, R.-X. Xu, Y. Yan, J. Chem. Phys. 133 (10) (2010) 101106, http://dx.doi.org/10.1063/1.3484491, URL http://aip.scitation.org/doi/10.1063/1.3484491.
[23] J. Hu, M. Luo, F. Jiang, R.-X. Xu, Y. Yan, J. Chem. Phys. 134 (24) (2011) 244106, http://dx.doi.org/10.1063/1.3602466, URL http://aip.scitation.org/doi/10.1063/1.3602466.
[24] R.B. Sidje, Y. Saad, Numer. Algorithms 56 (3) (2011) 455–479, http://dx.doi.org/10.1007/s11075-010-9397-6, URL http://link.springer.com/10.1007/s11075-010-9397-6.
[25] K. Wildberger, P. Lang, R. Zeller, P.H. Dederichs, Phys. Rev. B 52 (15) (1995) 11502–11508, http://dx.doi.org/10.1103/PhysRevB.52.11502, URL https://link.aps.org/doi/10.1103/PhysRevB.52.11502.
[26] G.H. Golub, J.H. Welsch, Math. Comp. 23 (106) (1969) 221–230, http://dx.doi.org/10.1090/S0025-5718-69-99647-1.
[27] P.W. Anderson, Phys. Rev. 124 (1961) 41–53, http://dx.doi.org/10.1103/PhysRev.124.41, URL https://link.aps.org/doi/10.1103/PhysRev.124.41.
[28] M. Caffarel, W. Krauth, Phys. Rev. Lett. 72 (1994) 1545–1548, http://dx.doi.org/10.1103/PhysRevLett.72.1545, URL https://link.aps.org/doi/10.1103/PhysRevLett.72.1545.
[29] H. Park, A.J. Millis, C.A. Marianetti, Phys. Rev. B 90 (2014) 235103, http://dx.doi.org/10.1103/PhysRevB.90.235103, URL https://link.aps.org/doi/10.1103/PhysRevB.90.235103.
[30] The MST2 package is available at https://github.com/mstsuite/MST2.
[31] Y. Wang, G.M. Stocks, W.A. Shelton, D.M.C. Nicholson, Z. Szotek, W.M. Temmerman, Phys. Rev. Lett. 75 (15) (1995) 2867–2870, http://dx.doi.org/10.1103/PhysRevLett.75.2867, URL https://link.aps.org/doi/10.1103/PhysRevLett.75.2867.