Inverse Volume Scaling of Finite-Size Error in Periodic Coupled Cluster Theory

Xin Xing[1] and Lin Lin[1,2]

[1]*Department of Mathematics, University of California, Berkeley, California 94720, USA*
[2]*Applied Mathematics and Computational Research Division,*
Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Coupled cluster theory is one of the most popular post-Hartree-Fock methods for *ab initio* molecular quantum chemistry. The finite-size error of the correlation energy in periodic coupled cluster calculations for three-dimensional insulating systems has been observed to satisfy the inverse volume scaling, even in the absence of any correction schemes. This is surprising, as simpler theories that utilize only a subset of the coupled cluster diagrams exhibit much slower decay of the finite-size error, which scales inversely with the length of the system. In this study, we review the current understanding of finite-size error in quantum chemistry methods for periodic systems. We introduce new tools that elucidate the mechanisms behind this phenomenon in the context of coupled cluster doubles calculations. This reconciles some seemingly paradoxical statements related to finite-size scaling. Our findings also show that singularity subtraction can be a powerful method to effectively reduce finite-size errors in practical quantum chemistry calculations for periodic systems.

Subject Areas: Chemical Physics,
Computational Physics,
Quantum Physics

I. INTRODUCTION

In the past few decades, *ab initio* methods for quantum many-body systems, such as density functional theory (DFT), quantum Monte Carlo methods, and quantum chemistry wave function methods, are becoming increasingly accurate and applied to ever larger range of systems [1,2]. Unlike molecular systems, periodic systems, including solids and surfaces, require calculating properties in the thermodynamic limit (TDL), a theoretical state in which the system size approaches infinity. However, the TDL cannot be directly accessed in practical applications. Finite-sized computational supercells are employed to approximate this limit, which introduces finite-size errors into the calculations. Finite-size errors can significantly affect the accuracy of calculations, even for systems with thousands of atoms. An extreme case is a moiré system such as magic angle twisted bilayer graphene (MATBG), where each computational unit cell consists of approximately 10 000 atoms, and the supercell needs to have more than 100 000 atoms to capture subtle correlation effects [3,4]. Directly tackling finite-size effects by enlarging the supercell size is very demanding, even for relatively inexpensive DFT calculations with modern-day supercomputers. For more accurate theories, this task is often computationally intractable. Understanding the finite-size scaling, i.e., the scaling of the finite-size error with respect to the system size, and employing finite-size error correction schemes are, therefore, crucial for obtaining accurate results using moderate-sized calculations.

The sources of finite-size errors in *ab initio* calculations are multifaceted and complex [5–7]. These errors are influenced by numerous factors, including system characteristics such as whether it is insulating or metallic or whether it is a three-dimensional bulk system versus a low-dimensional system. Calculations of electron kinetic energy, electron-ion interaction energy, Hartree energy, Fock exchange energy, and electron correlation energy can all contribute to finite-size errors. The first four types are predominantly single particle in nature, while the electron correlation energy is significantly more complex. To a large extent, electronic correlation is short-ranged, and this characteristic has spurred the development of local correlation methods, whose computational cost may scale linearly with system size. However, for *ab initio* methods to be accurate, they must also effectively account for van der Waals (vdW) interactions [8]. In solids, the cumulative effect of weak van der Waals interactions can become a significant contributor to the energy. The convergence of vdW energy follows an inverse volume scaling, implying that the finite-size error is inversely proportional to the volume of the supercell. The origin of finite-size error is not

solely confined to the characteristics of a physical system. The finite-size scaling can vary significantly across different theoretical frameworks even when applied to the same physical system. Sometimes, the finite-size scaling can manifest differently even when using the same theoretical framework for the same system, simply by altering the shape of the computational unit cell. All these complexities require careful analysis under varying scenarios, and this is important for accurately extrapolating toward the TDL or, preferably, for devising improved correction schemes aimed at reducing finite-size errors.

In recent years, there has been a growing interest in employing quantum chemistry wave function methods, such as Møller-Plesset perturbation theory and coupled cluster (CC) theory [9], to compute ground-state and excited-state properties for periodic systems [7,10–25]. Originally developed to study nuclear physics [26,27], CC theory has become one of the most popular methods in quantum chemistry that involves electronic correlation [28]. The coupled cluster singles, doubles, and perturbative triples [CCSD(T)] theory is often referred to as the "gold standard" in molecular quantum chemistry. In these methods, there are two primary strategies to approximate the TDL. The first involves expanding the computational supercells within real space. The second strategy involves performing calculations using a fixed unit cell and refining discretization of the Brillouin zone within reciprocal space using a $\mathbf{k}$-point mesh, such as the Monkhorst-Pack mesh [29]. This paper focuses on the latter approach, where the number of $\mathbf{k}$ points is denoted by $N_{\mathbf{k}}$. If the Monkhorst-Pack mesh includes the Gamma point of the Brillouin zone, this approach is equivalent to using a supercell comprised of $N_{\mathbf{k}}$ unit cells. The convergence toward the TDL can be studied by increasing this single parameter $N_{\mathbf{k}}$ toward infinity. The computational expense of quantum chemistry methods can rise sharply with respect to the system size. For example, in coupled cluster singles and doubles (CCSD) calculations, the computational cost scaling is $\mathcal{O}(N_{\mathbf{k}}^6)$ in real-space implementations [9] and $\mathcal{O}(N_{\mathbf{k}}^4)$ in reciprocal-space implementations [30]. Therefore, even moderate-sized calculations with $N_{\mathbf{k}} = 2 \times 2 \times 2$ or $3 \times 3 \times 3$ can already be computationally challenging. In CCSD(T) calculations, the computational cost of the real-space implementation can escalate to $\mathcal{O}(N_{\mathbf{k}}^7)$ in the worst-case scenario. This implies that even a moderate refinement of the Monkhorst-Pack grid by a factor of 2 along each dimension can lead to an increase in computational cost by $(2^3)^7 \approx 2$ million folds. Given this computational challenge imposed by system size, it is mostly impractical to estimate the finite-size scaling using power-law fitting over calculations on increasingly large systems and then determine the TDL value by extrapolation. Instead, a more feasible approach is to acquire the exact finite-size scaling through rigorous mathematical analysis and subsequently utilize power-law extrapolation to estimate the TDL value.

Although numerous empirical studies have examined the finite-size scaling of some quantum chemistry methods, there has been a notable lack of rigorous analysis. To the best of our knowledge, the first rigorous analysis of finite-size error in Hartree-Fock (HF) theory and second-order Møller-Plesset perturbation theory (MP2) for insulating systems has been conducted only recently [31]. The principal findings in Ref. [31] include the following.

(1) In the absence of any corrections, the finite-size error in HF scales as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$. As $N_{\mathbf{k}}$ is proportional to the supercell volume, $N_{\mathbf{k}}^{1/3}$ is proportional to the length of the supercell (we always assume uniform refinement of the $\mathbf{k}$-point mesh along all three dimensions for any shaped unit cell). This scaling is, thus, also referred to as inverse length scaling.

(2) By applying the Madelung constant correction, the finite-size error in HF improves to $\mathcal{O}(N_{\mathbf{k}}^{-1})$, i.e., the inverse volume scaling.

(3) The finite-size error of the MP2 correlation energy also satisfies the inverse volume scaling $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

As can be seen from HF and MP2, the specific finite-size scaling depends on the level of computational theory and numerical treatment employed in the calculations. In addition to confirming the empirically observed scalings, this rigorous analysis offers significant additional insights. It elucidates the nature of the singularity of the electron repulsion integral (ERI) due to the Coulomb kernel and its impact on finite-size scaling; it explains why the finite-size error can depend on the shapes and symmetries of the unit cell; and it has led to the development of new correction schemes, such as the staggered mesh method, which can expedite the convergence of HF, MP2, and random phase approximation (RPA) calculations toward TDL [32,33].

An innovative theoretical observation provided in Ref. [31] is that the finite-size error can be largely comprehended if assuming HF single-particle orbitals can be acquired exactly at any given $\mathbf{k}$ point in the Monkhorst-Pack mesh. This perspective helps to disentangle the contribution due to the relatively manageable single-particle effects from the collective and more complex electron correlation effects. Based on this assumption, the finite-size error can be rigorously examined from a numerical quadrature perspective. Specifically, the value of a physical observable at the TDL can often be written as a multidimensional integral over the Brillouin zone. As the Monkhorst-Pack mesh forms a uniform grid discretizing the Brillouin zone, the analysis simplifies to investigating the quadrature error of certain trapezoidal rule in a periodic region, a topic widely discussed in numerical analysis literature. The novelty here lies in the recognition that the associated integrands possess a unique singularity structure that is asymptotically of a specific fractional form. Reference [31], therefore, develops a new Euler-MacLaurin type of analysis that facilitates the study of

the finite-size error associated with HF and MP2 methods, taking into account this fractional form singularity.

Compared to HF and MP2, the finite-size error analysis and finite-size error correction methods in CC methods remain nascent. Early works in Refs. [12,34] focus on an important intermediate quantity called the structure factor and develop finite-size correction methods based on different types of polynomial interpolation of the structure factor near the Coulomb singularity. Subsequent studies on the uniform electron gas system [35,36] numerically examine the power-law scaling of the finite-size error in CC theory, offering new analytical perspective on the structure factor. These works also provide new error correction methods, such as twist-angle techniques [37,38] and structure factor interpolation [39] for these metallic systems. Very recently, Ref. [40] replaces the Coulomb kernel in CC calculations with an averaged Coulomb kernel in each quadrature element to reduce the finite-size error for anisotropic systems. Despite these developments, theoretical understanding of the finite-size error scaling in CC methods (with and without error correction methods) has been lacking.

Recently, by applying a similar approach, we have expanded the finite-size error analysis to include the correlation energy of the third-order Møller-Plesset perturbation theory (MP3) and CC theory [41]. The simplest CC theory is the coupled cluster singles (CCS) theory. However, due to Thouless's theorem [42], CCS only rotates the Slater determinant to another Slater determinant. For a nondegenerate closed-shell Hartree-Fock reference, the correlation energy from the CCS theory vanishes. Therefore, we focus on the coupled cluster doubles (CCD) theory, which is mathematically the simplest and representative form of CC theory. From a diagrammatic perspective, CC diagrams encompass all Møller-Plesset perturbation diagrams. When the CCD amplitude equation is solved iteratively with n fixed point iterations [referred to as the CCD(n) scheme], the MP2 and MP3 diagrams can be generated from CCD(1) and CCD(2), respectively. More generally, CCD(n) consists of a finite subset of Møller-Plesset perturbation diagrams. It is worth pointing out that there is a method called CCSD(2) which uses second-order perturbation to rectify CCSD energies for multireference and open-shell systems [43], and its meaning should, therefore, be distinguished from our CCD(n). Our analysis uncovers that the finite-size error in MP3 and CC is fundamentally different from that in MP2. Specifically, theories such as MP2 and RPA incorporate only "particle-hole" types of diagrams. The integrands corresponding to these diagrams are singular, but the singularity is relatively mild. However, starting from MP3, additional perturbative terms, namely, the "particle-particle" and "hole-hole" diagrams, must be considered. These terms introduce much stronger singularities, necessitating the development of new analytical tools. Our quadrature error analysis adapts the Poisson summation formula in a new setting and aligns

with a recently developed trapezoidal quadrature analysis for certain singular integrals [44]. This approach provides more accurate estimates and can be more applicable than our previous quadrature analysis based on the Euler-MacLaurin formula for HF and MP2 in Ref. [31]. In the absence of finite-size correction schemes and assuming exact orbital energies at any **k** point, the study in Ref. [41] concludes that the finite-size errors of MP3 and CCD both satisfy the inverse length scaling $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$.

Interestingly, for CCD, earlier numerical calculations did not provide conclusive evidence regarding its finite-size scaling, with different studies suggesting either an inverse volume scaling [15,34] or inverse length scaling [17]. More recent calculations demonstrate that the electron correlation energy in periodic coupled cluster calculations should follow an inverse volume scaling, even in the absence of finite-size correction schemes. This observation points to a significant gap in the theoretical understanding of the finite-size error and prompts the *central question* we address in this paper: *How can we reconcile the following seemingly paradoxical facts?*

(1) Without finite-size corrections, the finite-size error in CCD exhibits inverse volume scaling.

(2) Without finite-size corrections, the finite-size error in MP3 exhibits inverse length scaling. This rate is sharp and cannot be improved.

(3) All MP3 diagrams are encompassed within the CCD formulation.

There are several often-cited physical justifications for expecting that the CCD method, and CC theory more generally, may exhibit superior behavior when applied to periodic systems. One such reason stems from the size extensivity of CC theory. A theory is *size extensive* when the total energy of two noninteracting identical systems, calculated as a combined system, equates to twice the energy of one system computed independently. Unlike theories such as truncated configuration interaction methods, which are not size extensive, truncated CC theory (such as CCD) possesses this advantageous characteristic. Another reason is that CC theory can be formulated in such a way that it does not explicitly depend on orbital energies. One practical consequence of this is that, upon the convergence of the coupled cluster iterations, the Madelung constant correction, often used to reduce finite-size effects in many-body simulations, cancels out naturally. Therefore, in this scenario, CC without the Madelung constant correction is equivalent to that with the Madelung constant correction.

We would like to clarify that neither size extensivity nor the cancellation of the Madelung constant alone is sufficient to address the aforementioned question. While size extensivity is indeed a desirable property, many methods such as HF, MP2, and MP3, among others, are all size extensive. In fact, given that periodic systems are infinitely large, size extensivity should be viewed as a *necessary*

condition for the applicability of any numerical method to such systems. However, this property does not provide insight into the convergence rate of the finite-size error. The cancellation of the Madelung constant plays a pivotal role here. Nevertheless, demonstrating that CCD calculation (or even MP3) with the Madelung constant correction exhibits inverse volume scaling in finite-size error is itself a significant challenge. This requires the development of new technical tools not currently available in existing literature. Indeed, the development and application of these tools are the main technical contribution of this paper.

In this paper, we elucidate the origin of the inverse volume scaling behavior. Our analysis consists of two steps. First, we investigate the structure of the CCD amplitude equation. We show that the Madelung constant correction, commonly used to reduce finite-size errors in Fock exchange energy and orbital energies, can also be applied to reduce the finite-size error in ERI contractions within the CCD amplitude equation. We establish a connection between the Madelung constant correction and a quadrature error reduction technique known as the singularity subtraction method [45]. By subtracting the leading singular terms from the integrands in the numerical quadratures, the Madelung constant correction reduces the finite-size errors in both the ERI contractions and the orbital energies from $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ to $\mathcal{O}(N_{\mathbf{k}}^{-1})$. Furthermore, we demonstrate that, with the Madelung constant correction, the finite-size errors in $\mathrm{CCD}(n)$ and converged CCD calculations satisfy the desired inverse volume scaling.

In the second step of our analysis, we observe that, upon convergence of the CCD amplitude equations, the Madelung constant corrections to both orbital energies and ERI contractions perfectly cancel each other out for any finite-sized system. This cancellation ensures that the CCD correlation energy remains the same, regardless of whether the Madelung constant correction is applied. Combining this result with the first step, we conclude that the finite-size error of the correlation energy in converged CCD calculations satisfies the desired inverse volume scaling without the need for any additional correction schemes. However, prior to the convergence of the amplitude equations, this perfect cancellation does not occur, and the finite-size error of $\mathrm{CCD}(n)$ calculations remains $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$. A similar lack of cancellation occurs when the orbital energies take their exact value at the TDL but the ERI contractions are not corrected, resulting in an $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ finite-size error for both converged CCD and $\mathrm{CCD}(n)$ calculations studied in Ref. [41].

To validate our theoretical analysis, we perform CCD calculations on a 3D periodic hydrogen dimer system using the PYSCF software package [46]. Our numerical results support the conclusions drawn from the theoretical analysis and provide further evidence for the finite-size scaling behavior summarized in Table I.

The paper is organized as follows. Section II introduces background knowledge and basic notations. Section III first decomposes the finite-size error in $\mathrm{CCD}(n)$ calculations into the errors in four basic components and then describes how the four components contribute to the overall finite-size error with the possible Madelung constant correction. Section IV explains the key ideas of how the Madelung constant correction can reduce the finite-size error in orbital energies and ERI contractions from a numerical quadrature perspective. Section V illustrates the numerical results that corroborate our error estimate. Lastly, Sec. VI discusses the implication of our theoretical study and future directions.

TABLE I. Finite-size scaling of different computational theories with and without corrections to orbital energies and ERI contractions. The first two rows refer to the finite-size scaling of the Hartree-Fock (HF) exchange energy, and the remaining rows refer to the finite-size scaling of the correlation energy (excluding the HF exchange). While we focus on the Madelung constant correction in Ref. [31] and this work, any correction schemes that reduces the finite-size errors in orbital energies and ERI contractions to $\mathcal{O}(N_{\mathbf{k}}^{-1})$ can also be applied, and the conclusions drawn here remain valid. In particular, exact values of single particle orbital energies at the TDL satisfy the condition above. N/A means that Madelung constant correction does not apply within the theory.

Theory	Correction to orbital energies	Correction to ERI contractions	Finite-size scaling	References
HF	N/A	✗	$N_{\mathbf{k}}^{-1/3}$	[31, Theorem 3.1] and [6,17,47]
HF	N/A	✓	$N_{\mathbf{k}}^{-1}$	[31, Theorem 5.1] and [47]
MP2	✓	N/A	$N_{\mathbf{k}}^{-1}$	[31, Theorem 4.1] and [15,34]
RPA, SOSEX, drCCD	✓	N/A	$N_{\mathbf{k}}^{-1}$	[33,34]
MP3	✓	✗	$N_{\mathbf{k}}^{-1/3}$	[41]
MP3	✓	✓	$N_{\mathbf{k}}^{-1}$	Theorem 1
CCD(n)/CCD	✓	✗	$N_{\mathbf{k}}^{-1/3}/N_{\mathbf{k}}^{-1/3}$	[41]
CCD(n)/CCD	✓	✓	$N_{\mathbf{k}}^{-1}/N_{\mathbf{k}}^{-1}$	Theorem 1/Theorem 2
CCD(n)/CCD	✗	✓	$N_{\mathbf{k}}^{-1/3}/N_{\mathbf{k}}^{-1/3}$	Theorem 1/Theorem 2
CCD(n)/CCD	✗	✗	$N_{\mathbf{k}}^{-1/3}/N_{\mathbf{k}}^{-1}$	Theorem 1/Corollary 3

II. BACKGROUND

Consider a unit cell and its Brillouin zone denoted as Ω and Ω^*, respectively. Denote the associated real- and reciprocal-space lattices as $\mathbb{L}$ and $\mathbb{L}^*$, respectively. To model such a periodic system, the Brillouin zone Ω^* is commonly discretized using a uniform mesh $\mathcal{K}$ of size $N_{\mathbf{k}}$ (known as the Monkhorst-Pack grid [29]). The orbitals and orbital energies (also called bands and band energies) $\{\psi_{n\mathbf{k}}, \varepsilon_{n\mathbf{k}}\}$ indexed by orbital indices n and momentum vectors $\mathbf{k} \in \mathcal{K}$ can be solved by the HF method. As a common practice, $n \in \{i, j\}$ refers to an occupied orbital and $n \in \{a, b\}$ refers to an virtual orbital. Throughout this paper, we use the *normalized* ERI:

$$
\langle n_1\mathbf{k}_1, n_2\mathbf{k}_2 | n_3\mathbf{k}_3, n_4\mathbf{k}_4 \rangle
$$
$$
= \frac{4\pi}{|\Omega|} \sideset{}{'}\sum_{\mathbf{G} \in \mathbb{L}^*} \frac{1}{|\mathbf{q} + \mathbf{G}|^2} \hat{\varrho}_{n_1\mathbf{k}_1, n_3(\mathbf{k}_1+\mathbf{q})}(\mathbf{G}) \hat{\varrho}_{n_2\mathbf{k}_2, n_4(\mathbf{k}_2-\mathbf{q})}(-\mathbf{G}),
$$

(1)

where $\mathbf{q} = \mathbf{k}_3 - \mathbf{k}_1$ is the momentum transfer vector, the crystal momentum conservation $\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4 \in \mathbb{L}^*$ is assumed implicitly, and $\hat{\varrho}_{n'\mathbf{k}',n\mathbf{k}}(\mathbf{G}) = \langle \psi_{n'\mathbf{k}'} | e^{i(\mathbf{k}'-\mathbf{k}-\mathbf{G})\cdot\mathbf{r}} | \psi_{n\mathbf{k}} \rangle$ is Fourier representation of the pair product. The primed summation over $\mathbf{G}$ means that the possible term with $\mathbf{q} + \mathbf{G} = \mathbf{0}$ is excluded in the numerical calculation. Using a finite mesh $\mathcal{K}$, the HF orbital energy without any correction is computed as

$$
\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} = \langle n\mathbf{k}|\hat{H}_0|n\mathbf{k}\rangle
$$
$$
+ \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_i \in \mathcal{K}} \sum_i (2\langle i\mathbf{k}_i, n\mathbf{k}|i\mathbf{k}_i, n\mathbf{k}\rangle - \langle i\mathbf{k}_i, n\mathbf{k}|n\mathbf{k}, i\mathbf{k}_i\rangle),
$$

(2)

where $\hat{H}_0$ refers to the single-particle component of the many body Hamiltonian.

In this paper, we focus on three-dimensional insulating systems with an indirect gap, i.e., $\varepsilon_{a\mathbf{k}_a} - \varepsilon_{i\mathbf{k}_i} > 0$, $\forall i, a, \mathbf{k}_i, \mathbf{k}_a$. To simplify the analysis, we assume that the orbitals are exact at any $\mathbf{k}$ point and the number of virtual orbitals are truncated to a finite number. In addition, we assume that the exact orbitals and orbital energies in the TDL are smooth and periodic with respect to their momentum vector index $\mathbf{k} \in \Omega^*$. This assumption is a restriction in our current analysis. For systems free of topological obstructions [48,49], these conditions may be replaced by weaker ones using techniques based on Green's functions or Hamiltonians defined in the atomic orbital basis instead of the band basis [50].

A. CCD theory

Based on the reference HF determinant $|\Phi\rangle$, the CCD theory represents the wave function as

$$
|\Psi\rangle = e^{\mathcal{T}}|\Phi\rangle,
$$
$$
\mathcal{T} = \frac{1}{N_{\mathbf{k}}} \sum_{ijab} \sum_{\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a \in \mathcal{K}} T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) a_{a\mathbf{k}_a}^\dagger a_{b\mathbf{k}_b}^\dagger a_{j\mathbf{k}_j} a_{i\mathbf{k}_i},
$$

where $a_{n\mathbf{k}}^\dagger$ and $a_{n\mathbf{k}}$ are creation and annihilation operators, respectively, for $\psi_{n\mathbf{k}}$, $T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ (commonly denoted as $t_{i\mathbf{k}_i, j\mathbf{k}_j}^{a\mathbf{k}_a, b\mathbf{k}_b}$ in the literature) is the *normalized* CCD double amplitude, and $\mathbf{k}_b \in \mathcal{K}$ is uniquely determined using the crystal momentum conservation $\mathbf{k}_i + \mathbf{k}_j - \mathbf{k}_a - \mathbf{k}_b \in \mathbb{L}^*$. The amplitude tensor $T_*^{N_{\mathbf{k}}} = \{T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\}$ is defined as the root of a nonlinear amplitude equation that consists of constant, linear, and quadratic terms.

In practice, the amplitude equation can be solved using a quasi-Newton method [9,51], which is equivalent to applying fixed point iteration to

$$
T = \frac{1}{\varepsilon^{N_{\mathbf{k}}}} \mathcal{A}_{N_{\mathbf{k}}}(T).
$$

(3)

Here, $\varepsilon^{N_{\mathbf{k}}}$ denotes a diagonal operator with entries $\varepsilon_{i\mathbf{k}_i, j\mathbf{k}_j, a\mathbf{k}_a, b\mathbf{k}_b}^{N_{\mathbf{k}}} = \varepsilon_{i\mathbf{k}_i}^{N_{\mathbf{k}}} + \varepsilon_{j\mathbf{k}_j}^{N_{\mathbf{k}}} - \varepsilon_{a\mathbf{k}_a}^{N_{\mathbf{k}}} - \varepsilon_{b\mathbf{k}_b}^{N_{\mathbf{k}}}$, and $1/\varepsilon^{N_{\mathbf{k}}}$ gives the diagonal operator with $1/\varepsilon_{i\mathbf{k}_i, j\mathbf{k}_j, a\mathbf{k}_a, b\mathbf{k}_b}^{N_{\mathbf{k}}}$. The operator $\mathcal{A}_{N_{\mathbf{k}}}(T)$ is referred to as the *ERI-contraction map* (see the definition in Appendix A). It consists of contractions between ERIs and T and does not involve orbital energies. Note that both T and $\mathcal{A}_{N_{\mathbf{k}}}(T)$ are tensors indexed by (i, j, a, b) and $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \in \mathcal{K} \times \mathcal{K} \times \mathcal{K}$. The CCD correlation energy is then defined as

$$
E_{\text{CCD}}^{N_{\mathbf{k}}} = \frac{1}{N_{\mathbf{k}}^3} \sum_{\mathbf{k}_i\mathbf{k}_j\mathbf{k}_a \in \mathcal{K}} \sum_{ijab} W_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)
$$
$$
\coloneqq \mathcal{G}_{N_{\mathbf{k}}}(T_*^{N_{\mathbf{k}}}),
$$

where $W_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ denotes the antisymmetrized ERI $2\langle i\mathbf{k}_i, j\mathbf{k}_j | a\mathbf{k}_a, b\mathbf{k}_b \rangle - \langle i\mathbf{k}_i, j\mathbf{k}_j | b\mathbf{k}_b, a\mathbf{k}_a \rangle$.

In the TDL with $\mathcal{K}$ replaced by Ω^*, $T_*^{N_{\mathbf{k}}} = \{T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\}$ with $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \in \mathcal{K} \times \mathcal{K} \times \mathcal{K}$ should be replaced by $t_* = \{t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\}$, where each $t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ is a function of $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ in $\Omega^* \times \Omega^* \times \Omega^*$. The converged TDL amplitude t_* satisfies a similar amplitude equation

$$
t = \frac{1}{\varepsilon^{\text{TDL}}} \mathcal{A}_{\text{TDL}}(t),
$$

(4)

where $\mathcal{A}_{\text{TDL}}$ can be obtained from $\mathcal{A}_{N_{\mathbf{k}}}$ by taking the limit $(1/N_{\mathbf{k}}) \sum_{\mathbf{k} \in \mathcal{K}} \to (1/|\Omega^*|) \int_{\Omega^*} d\mathbf{k}$ with $N_{\mathbf{k}} \to \infty$. For example, the *four-hole–two-particle* (4h2p) *linear term* in $[\mathcal{A}_{N_{\mathbf{k}}}(T)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ converges in $\mathcal{A}_{\text{TDL}}(t)$ as

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle T_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)$$

$$\xrightarrow{\mathcal{K} \to \Omega^*} \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{k}_k \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle t_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a).$$

$$(5)$$

The image of $\mathcal{A}_{\text{TDL}}$ is a set of functions of $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$ indexed by (i, j, a, b). The CCD correlation energy in the TDL is defined in a similar way as

$$E_{\text{CCD}}^{\text{TDL}} = \frac{1}{|\Omega^*|^3} \int_{\Omega^* \times \Omega^* \times \Omega^*} d\mathbf{k}_i d\mathbf{k}_j d\mathbf{k}_a W_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$$
$$\times t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$$
$$:= \mathcal{G}_{\text{TDL}}(t_*).$$

Applying n steps of fixed point iteration over Eqs. (3) and (4) with zero initial guess (i.e., $t = 0$, $T = 0$), we obtain the CCD(n) amplitude and the CCD(n) energy:

$$E_{\text{CCD}(n)}^{N_{\mathbf{k}}} = \mathcal{G}_{N_{\mathbf{k}}}(T_n^{N_{\mathbf{k}}}) \qquad E_{\text{CCD}(n)}^{\text{TDL}} = \mathcal{G}_{\text{TDL}}(t_n)$$
$$T_m^{N_{\mathbf{k}}} = (\varepsilon^{N_{\mathbf{k}}})^{-1} \mathcal{A}_{N_{\mathbf{k}}}(T_{m-1}^{N_{\mathbf{k}}}) \quad \to \quad t_m = (\varepsilon^{\text{TDL}})^{-1} \mathcal{A}_{\text{TDL}}(t_{m-1})$$
$$T_0^{N_{\mathbf{k}}} = \mathbf{0} \qquad\qquad t_0 = \mathbf{0}.$$

$$(6)$$

CCD(n) is related to the perturbative description of CCD and consists of a subset of finite order perturbation energy terms in the Møller-Plesset perturbation theory. For example, CCD(1) can be identified with MP2, and CCD(2) contains all the terms in MP2 and MP3 and a subset of terms in MP4.

One main result of this paper is the rigorous analysis of the finite-size error in CCD(n) calculation with any fixed $n > 0$. If the fixed point iterations in both the finite and TDL cases converge to the CCD amplitudes as $n \to \infty$, i.e., $T_n^{N_{\mathbf{k}}} \to T_*^{N_{\mathbf{k}}}$ and $t_n \to t_*$ (the technical definition of this convergence is provided in Appendix C), the finite-size error analysis for CCD(n) also applies to the converged CCD calculation. In other words, we analyze the finite-size error in CCD calculation using a perturbative approach based on the analysis on CCD(n).

B. Madelung constant correction

Reference [41] shows that the finite-size errors in CCD(n) and CCD both scale as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ when assuming that exact orbital energies are used in the amplitude equation (3). The same finite-size scaling also appears in Fock exchange energy and occupied orbital energy calculations. One common correction to reduce the finite-size errors in the latter two calculations is to add a Madelung constant shift [6,47,52] to the Ewald kernel. This shift

introduces a correction to all involved ERIs in the calculations as

$$\langle n_1 \mathbf{k}_1, n_2 \mathbf{k}_2 | n_3 \mathbf{k}_3, n_4 \mathbf{k}_4 \rangle - \delta_{n_1 n_3} \delta_{n_2 n_4} \delta_{\mathbf{k}_1 \mathbf{k}_3} \delta_{\mathbf{k}_2 \mathbf{k}_4} N_{\mathbf{k}} \xi.$$

Such a correction is triggered only in ERIs which have fully matched orbital indices, i.e., $n_1 = n_3, n_2 = n_4$, and zero momentum transfer, i.e., $\mathbf{k}_1 = \mathbf{k}_3$. The Madelung constant ξ is defined uniquely by the unit cell and the $\mathbf{k}$-point mesh $\mathcal{K}$ as

$$\xi = \left(\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q} \in \mathcal{K}_{\mathbf{q}}} - \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{q} \right) \sum_{\mathbf{G} \in \mathbb{L}^*}' \frac{4\pi}{|\Omega|} \frac{e^{-\sigma|\mathbf{q}+\mathbf{G}|^2}}{|\mathbf{q} + \mathbf{G}|^2} - \frac{4\pi\sigma}{N_{\mathbf{k}}}$$
$$+ \sum_{\mathbf{R} \in \mathbb{L}_{\mathcal{K}}}' \frac{\text{erfc}(\sigma^{-1/2}|\mathbf{R}|/2)}{|\mathbf{R}|}, \qquad (7)$$

where $\mathcal{K}_{\mathbf{q}}$ is a uniform mesh that is of the same size as $\mathcal{K}$ and contains $\mathbf{q} = \mathbf{0}$ and $\mathbb{L}_{\mathcal{K}}$ is the real-space lattice associated with the reciprocal-space lattice $\mathbf{q} + \mathbf{G}$ with $\mathbf{q} \in \mathcal{K}_{\mathbf{q}}, \mathbf{G} \in \mathbb{L}$. Note that ξ does not vary with respect to $\sigma > 0$ [6] and this parameter σ is commonly tuned to control the lattice cutoffs in the summation over $\mathbb{L}^*$ and $\mathbb{L}_{\mathcal{K}}$ in Eq. (7) when numerically computing ξ.

For finite-size orbital energy calculation in Eq. (2), the Madelung constant correction gives

$$\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}, \xi} = \begin{cases} \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} + \xi & n \text{ is occupied,} \\ \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} & n \text{ is virtual.} \end{cases} \qquad (8)$$

For the ERI contractions in $\mathcal{A}_{N_{\mathbf{k}}}(T)$, this correction should be applied to six linear amplitude terms (see Appendix B 2). For example, the $4h2p$ linear term in Eq. (5) should be modified to

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle T_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)$$
$$- \xi T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a). \qquad (9)$$

Collecting the corrections to all six terms together, $\mathcal{A}_{N_{\mathbf{k}}}(T)$ is modified to

$$\mathcal{A}_{N_{\mathbf{k}}, \xi}(T) = \mathcal{A}_{N_{\mathbf{k}}}(T) + 2\xi T. \qquad (10)$$

In the finite-size CCD(n) and converged CCD calculations, the Madelung constant correction can be applied to the orbital energies, ERI contractions, or both in the amplitude equation (3). As a result, we may have three correction schemes compared to the standard calculation without any correction.

Particularly, applying the Madelung constant correction to both components gives the amplitude equation

$$T = \frac{1}{\varepsilon^{N_{\mathbf{k}}} + 2\xi}[\mathcal{A}_{N_{\mathbf{k}}}(T) + 2\xi T] = \frac{1}{\varepsilon^{N_{\mathbf{k}},\xi}}\mathcal{A}_{N_{\mathbf{k}},\xi}(T). \quad (11)$$

It can be easily verified that this amplitude equation has the same solution as Eq. (3) without any correction. Its CCD solution is, thus, identical to the one without correction. However, the associated CCD(n) calculation differs and can be interpreted as solving the original amplitude equation (3) using a quasi-Newton method that has a 2ξ-diagonal shift to the Jacobian matrix.

III. MAIN STATEMENTS

We start our error analysis with the CCD(n) calculation with a fixed number of iterations $n > 0$ and then generalize the analysis to the fully converged CCD calculation. In CCD(n), the finite-size error is quantified by $E_{\text{CCD}(n)}^{\text{TDL}} - E_{\text{CCD}(n)}^{N_{\mathbf{k}}}$. According to Eq. (6), this error can be traced back to two sources: the difference in energy operators, $\mathcal{G}_{\text{TDL}}$ versus $\mathcal{G}_{N_{\mathbf{k}}}$, and the difference in amplitudes, t_n versus $T_n^{N_{\mathbf{k}}}$. Recall that t_n and $T_n^{N_{\mathbf{k}}}$ are the amplitudes of the system in TDL and finite-size cases, respectively. Let us consider the *evaluation map*, denoted by $\mathcal{M}_{\mathcal{K}}$. This map evaluates a tensor valued function, initially defined on the product space $\Omega^* \times \Omega^* \times \Omega^*$, on a finite-sized grid $\mathcal{K} \times \mathcal{K} \times \mathcal{K}$. Consequently, the values of the TDL amplitude t_n on this finite-size grid, $\mathcal{K} \times \mathcal{K} \times \mathcal{K}$, are given by $\mathcal{M}_{\mathcal{K}}t_n$ and are approximated by the finite-size amplitude $T_n^{N_{\mathbf{k}}}$.

By applying the triangle inequality, we can decompose the finite-size error into two distinct sources: the errors arising from the discretized energy calculation using the exact amplitude $\mathcal{M}_{\mathcal{K}}t_n$ and the errors stemming from the amplitude calculation itself:

$$|E_{\text{CCD}(n)}^{\text{TDL}} - E_{\text{CCD}(n)}^{N_{\mathbf{k}}}|$$
$$\leq |\mathcal{G}_{\text{TDL}}(t_n) - \mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_n)| + |\mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_n) - \mathcal{G}_{N_{\mathbf{k}}}(T_n^{N_{\mathbf{k}}})|$$
$$\leq |\mathcal{G}_{\text{TDL}}(t_n) - \mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_n)|$$
$$+ \frac{C}{N_{\mathbf{k}}^3}\sum_{ijab}\sum_{\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a \in \mathcal{K}}|[\mathcal{M}_{\mathcal{K}}t_n - T_n^{N_{\mathbf{k}}}]_{ijab}(\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a)|. \quad (12)$$

Here, we use the fact that $|W_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)|$ can be upper bounded uniformly by a constant.

To further break down the error in amplitude calculation, we note that t_n and $T_n^{N_{\mathbf{k}}}$ are recursively constructed by Eq. (6) with initial values $t_0 = \mathbf{0}$ and $T_0^{N_{\mathbf{k}}} = \mathbf{0}$. As a result, the error in the CCD(n) amplitude calculation can also be recursively decomposed using the same strategy above as

$$\mathcal{M}_{\mathcal{K}}t_n - T_n^{N_{\mathbf{k}}} = \frac{1}{\varepsilon^{\text{TDL}}}[\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\text{TDL}}(t_{n-1}) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_{n-1})]$$
$$+ \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_{n-1})\left[\frac{1}{\varepsilon^{\text{TDL}}} - \frac{1}{\varepsilon^{N_{\mathbf{k}}}}\right]$$
$$+ \frac{1}{\varepsilon^{N_{\mathbf{k}}}}\left[\mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_{n-1}) - \mathcal{A}_{N_{\mathbf{k}}}(T_{n-1}^{N_{\mathbf{k}}})\right]. \quad (13)$$

The three error terms from this dissection can be interpreted as the errors in ERI contractions using exact CCD($n - 1$) amplitudes, orbital energies, and CCD($n - 1$) amplitude calculation composed with $\mathcal{A}_{N_{\mathbf{k}}}$, respectively. For the last term, it can be shown that entries in $\mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_{n-1}) - \mathcal{A}_{N_{\mathbf{k}}}(T_{n-1}^{N_{\mathbf{k}}})$ have the same scaling with respect to $N_{\mathbf{k}}$ as those in $\mathcal{M}_{\mathcal{K}}t_{n-1} - T_{n-1}^{N_{\mathbf{k}}}$. Replacing the last term by $\mathcal{M}_{\mathcal{K}}t_{n-1} - T_{n-1}^{N_{\mathbf{k}}}$ and applying the dissection recursively, we find that the error in the CCD(n) amplitude calculation is determined by those in the ERI contractions and orbital energies, i.e., the first two terms in Eq. (13).

Overall, the finite-size error of CCD(n) calculation can be decomposed into errors in three basic factors: (i) energy calculation using exact CCD(n) amplitude, (ii) ERI contractions using exact CCD($n - 1$) amplitude, and (iii) orbital energies. This error decomposition is also valid when applying the Madelung constant correction to orbital energies [Eq. (8)] or ERI contractions [Eq. (10)]. By analyzing these three error sources with or without corrections separately, we can obtain the finite-size error estimate for CCD(n) with various correction schemes.

The Madelung constant correction can reduce the finite-size error in orbital energies from $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ to $\mathcal{O}(N_{\mathbf{k}}^{-1})$. This correction is at the HF level. One *main technical result* of this work is to show that the Madelung constant correction can also reduce the finite-size error in *most (but not all) entries* of the ERI contraction [note that $\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\text{TDL}}(t) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t)$ is a tensor] from $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ to $\mathcal{O}(N_{\mathbf{k}}^{-1})$; see Sec. IV B. As a result, when applying the Madelung constant correction to *both* orbital energies *and* ERI contractions in the CCD(n) calculation, the overall finite-size error can be successfully reduced to $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

Theorem 1. In CCD(n) calculation, the finite-size error in the correlation energy scales as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ in each of the following scenarios: (i) there is no finite-size correction, (ii) the Madelung constant correction is applied only to the ERI contraction $\mathcal{A}_{N_{\mathbf{k}}}$, and (iii) the Madelung constant correction is applied only to the orbital energy $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$.

When the Madelung constant correction is applied to both $\mathcal{A}_{N_{\mathbf{k}}}$ and $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$ in the CCD(n) calculation, the overall finite-size error scales as $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

As a special case, the same conclusion applies to MP3 calculations.

Proof. See Appendix B. ∎

Now what happens to the finite-size error of the CCD calculation as $n \to \infty$? For gapless and small-gap systems, it has been observed in practice that the fixed point iteration might not converge or the amplitude equation might have multiple solutions. While we focus on systems with an indirect gap, it is worth noting that, even in such favorable scenarios, the existence and uniqueness of solutions to the CCD amplitude equations in both finite and TDL cases remain an open question and are beyond the scope of this paper.

To study the finite-size error of CCD via the above results on CCD(n), we make additional technical assumptions (see Appendix C) that can guarantee the convergence of CCD(n) to CCD. Under these assumptions, we show that the finite-size scaling of CCD calculation is upper bounded by those of its associated converging CCD(n) calculations. Numerical observations (see Sec. V) further show that this finite-size scaling estimate through CCD(n) is asymptotically sharp for CCD calculation with the Madelung constant correction applied to orbital energies, ERI contractions, or both.

Theorem 2 (informal). Under additional conditions on the convergence of CCD(n) to CCD, the finite-size error of the CCD correlation energy scales as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ in each of the following scenarios: (i) the Madelung constant correction is applied only to the ERI contraction $\mathcal{A}_{N_{\mathbf{k}}}$, and (ii) the Madelung constant correction is applied only to the orbital energy $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$.

When the Madelung constant correction is applied to both $\mathcal{A}_{N_{\mathbf{k}}}$ and $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$ in the CCD calculation, the overall finite-size error scales as $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

Compared to Theorem 1, Theorem 2 does not address the finite-size error of the CCD calculation when no finite-size correction is applied. The natural conclusion from Theorem 1 is that this finite-size scaling should be $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ as well. However, this error estimate is loose and inconsistent with the numerical observations which suggest an inverse volume scaling. To obtain a tight estimate, we now use the observation that the CCD calculation without any correction is equivalent to the CCD calculation with the Madelung constant correction applied to both ERI contractions and orbital energies. Specifically, when applying the Madelung constant corrections Eqs. (8) and (10), the CCD amplitude equation in Eq. (11) can be formulated as

$$(\varepsilon^{N_{\mathbf{k}}} + 2\xi)T = \mathcal{A}_{N_{\mathbf{k}}}(T) + 2\xi T \Leftrightarrow \varepsilon^{N_{\mathbf{k}}} T = \mathcal{A}_{N_{\mathbf{k}}}(T).$$

With $2\xi T$ on both sides canceling each other, this reformulation is exactly reduced to the original amplitude equation (3) without corrections. Therefore, the roots of the two amplitude equations with and without the corrections are the same, and the correlation energy in *converged* CCD calculation without any corrections is the same as that with the Madelung constant correction. In other words, when investigating the finite-size error of CCD calculation without corrections, we should apply Theorem 2 with the Madelung constant correction applied to both $\mathcal{A}_{N_{\mathbf{k}}}$ and $\varepsilon_{n_{\mathbf{k}}}^{N_{\mathbf{k}}}$. This yields a sharp estimate $\mathcal{O}(N_{\mathbf{k}}^{-1})$ and explains the origin of the inverse volume scaling of the finite-size error.

Corollary 3. Under the same additional conditions as in Theorem 2, the finite-size error of the CCD correlation energy without finite-size correction scales as $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

We provide the proof of Theorem 2 in Appendix C, and Corollary 3 follows directly from Theorem 2 and the reasonings above.

IV. KEY IDEAS

As demonstrated in the previous section, the finite-size errors in CCD(n) and CCD calculations can be reduced to the errors in three simpler basic calculations: energy calculation using exact amplitudes, ERI contraction using exact amplitudes, and orbital energies. A key observation is that the finite-size errors in all three calculations can be interpreted and analyzed from a numerical quadrature perspective.

For a function g over a hypercube V, we denote a (generalized) trapezoidal rule using a uniform mesh $\mathcal{X}$ of V and its quadrature error as

$$\mathcal{Q}_V(g, \mathcal{X}) = \frac{|V|}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x}),$$

$$\mathcal{E}_V(g, \mathcal{X}) = \int_V d\mathbf{x} g(\mathbf{x}) - \mathcal{Q}_V(g, \mathcal{X}).$$

Under the assumption of exact orbitals at any $\mathbf{k}$ point in finite-size calculation, the errors in (i) orbital energy, (ii) energy, and (iii) ERI contraction $\mathcal{A}_{N_{\mathbf{k}}}$ can be, respectively, formulated by their definitions as

$$\varepsilon_{n\mathbf{k}}^{\text{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} = \frac{1}{|\Omega^*|} \mathcal{E}_{\Omega^*} \left(\sum_i W_{inin}(\mathbf{k}_i, \mathbf{k}, \mathbf{k}_i), \mathcal{K} \right),$$

$$\mathcal{G}_{\text{TDL}}(t) - \mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}} t) = \frac{1}{|\Omega^*|^3} \mathcal{E}_{\Omega^* \times \Omega^* \times \Omega^*} \left(\sum_{ijab} W_{ijab} t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a), \mathcal{K} \times \mathcal{K} \times \mathcal{K} \right),$$

$$\mathcal{M}_{\mathcal{K}} \mathcal{A}_{\text{TDL}}(t) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}} t) = \frac{1}{|\Omega^*|} \mathcal{E}_{\Omega^*} \left(\sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle t_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a), \mathcal{K} \right) + \cdots. \quad (14)$$

For ERI contraction in the last line, the error is denoted by a tensor with indices $(i, j, a, b, \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$, and the error for the $4h2p$ linear term Eq. (5) is detailed with $\mathbf{k}_k$ being the integration variable. Meanwhile, terms not shown account for errors from computing other linear and quadratic amplitude terms. From a numerical quadrature perspective, all three finite-sized calculations approximate the corresponding integrals in the TDL using the trapezoidal rule and a finite mesh $\mathcal{K}$ to discretize Ω^*. Therefore, their finite-size errors can be estimated systematically by quadrature error analysis.

In general, the quadrature error associated with a trapezoidal rule is influenced by the integrand's smoothness and boundary conditions. If we take Δh as the mesh size along each dimension, the quadrature error for a smooth integrand is typically of order $\mathcal{O}(\Delta h^2)$. Interestingly, if the integrand is also periodic, the error diminishes much more rapidly than it does in a nonperiodic scenario. The decay rate is faster than any finite power of Δh, showcasing a superalgebraic decay [53]. However, for an integrand periodic but marked by singularities, its quadrature error tends to taper off at a slower rate.

For quadrature errors in Eq. (14), the involved integrands are all periodic across their integration domains. However, many of these integrands have singularities within the domains, affecting the scaling of their quadrature errors. As demonstrated next, these integrands include point singularities arising from both ERIs and amplitudes, resulting in low-order power-law decay of the corresponding quadrature errors as $N_{\mathbf{k}}$ increases.

A. Singularity structure and quadrature error estimate

All the quadrature errors in Eq. (14) have integrands comprising of either ERIs or contractions between ERIs and exact CCD(n) amplitudes. The integration variables are momentum vectors sampled in the Brillouin zone Ω^*. Consequently, understanding the singularity structure of ERIs and exact CCD(n) amplitudes with respect to their momentum vector indices is crucial in comprehending the singularity structure of these integrands and ultimately estimating the quadrature error in the three basic calculations.

First, consider a generic ERI $\langle n_1\mathbf{k}_1, n_2\mathbf{k}_2 | n_3\mathbf{k}_3, n_4\mathbf{k}_4 \rangle$ with fixed band indices (n_1, n_2, n_3, n_4) and treat it as a function of $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{q} = \mathbf{k}_3 - \mathbf{k}_1$ in Ω^*. By its definition, the ERI can be separated as

$$\langle n_1\mathbf{k}_1, n_2\mathbf{k}_2 | n_3\mathbf{k}_3, n_4\mathbf{k}_4 \rangle$$
$$= \frac{4\pi}{|\Omega|} \frac{\hat{\varrho}_{n_1\mathbf{k}_1, n_3(\mathbf{k}_1+\mathbf{q})}(\mathbf{0}) \hat{\varrho}_{n_2\mathbf{k}_2, n_4(\mathbf{k}_2-\mathbf{q})}(\mathbf{0})}{|\mathbf{q}|^2}$$
$$+ \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \setminus \{\mathbf{0}\}} \frac{\hat{\rho}_{n_1\mathbf{k}_1, n_3(\mathbf{k}_1+\mathbf{q})}(\mathbf{G}) \hat{\rho}_{n_2\mathbf{k}_2, n_4(\mathbf{k}_2-\mathbf{q})}(-G)}{|\mathbf{q}+\mathbf{G}|^2}. \quad (15)$$

Since we assume all orbitals $\psi_{n\mathbf{k}}$ periodic and smooth with respect to $\mathbf{k} \in \Omega^*$, this ERI has a point singularity at $\mathbf{q} = \mathbf{0}$ only due to the first fraction term. Specifically, this term is of fractional form $f(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q})/|\mathbf{q}|^2$ with a smooth numerator f. As can be verified by direct calculation, such a fraction term can have its point singularity at $\mathbf{q} = \mathbf{0}$ described by the following general concept called algebraic singularity.

Definition 4 (informal). A function $f(\mathbf{x})$ has *algebraic singularity of order* $\gamma \in \mathbb{R}$ at $\mathbf{x}_0 \in \mathbb{R}^d$ if every lth-order derivative near $\mathbf{x}_0$ is bounded asymptotically by $|\mathbf{x} - \mathbf{x}_0|^{\gamma-l}$, i.e.,

$$\left| \frac{\partial^{\boldsymbol{\alpha}}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}} f(\mathbf{x}) \right| \leq C|\mathbf{x} - \mathbf{x}_0|^{\gamma-|\boldsymbol{\alpha}|}, \quad \forall \, \boldsymbol{\alpha} \geq 0,$$

where $\boldsymbol{\alpha}$ denotes a non-negative d-dimensional derivative multi-index. For brevity, f is said to be singular at $\mathbf{x}_0$ with the order of γ. See Definition 8 in Appendix B for the rigorous mathematical definition.

A representative example of such a singular function of order γ is $g(\mathbf{x})/|\mathbf{x}|^2$, where $g(\mathbf{x})$ is smooth and scales as $\mathcal{O}(|\mathbf{x}|^{2+\gamma})$ near $\mathbf{x} = \mathbf{0}$. Using orbital orthogonality, the generic ERI exhibits singularities at $\mathbf{q} = \mathbf{0}$ with the order of 0 when band indices mismatch ($n_1 \neq n_3$, $n_2 \neq n_4$), -1 when they partially match ($n_1 = n_3, n_2 \neq n_4$ or $n_1 \neq n_3$, $n_2 = n_4$), and -2 when they fully match ($n_1 = n_3, n_2 = n_4$).

We can now characterize the singularity structure of integrand $\sum_i W_{inin}(\mathbf{k}_i, \mathbf{k}, \mathbf{k}_i)$ defined by orbital energy $\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}}$ in Eq. (14). Fixing $(n, \mathbf{k})$, the leading singularity in each $W_{inin}(\mathbf{k}_i, \mathbf{k}, \mathbf{k}_i)$ (as a function of $\mathbf{k}_i$) comes from the exchange term $\langle i\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, i\mathbf{k}_i \rangle$ which has algebraic singularity at $\mathbf{k}_i = \mathbf{k}$ of order 0 when $i \neq n$ and -2 when $i = n$. In computing $\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}}$, the overall integrand is, thus, singular at $\mathbf{k}_i = \mathbf{k}$ of order 0 for a virtual band n and -2 for an occupied band n.

For such periodic functions with one point of algebraic singularity, the conventional textbook analysis of the trapezoidal rule is overly pessimistic. A key technical tool in this work is Lemma 5 below, which provides a rigorous and sharp quadrature error estimate linking its error scaling to the singularity order. (See Lemma 22 in Appendix D for a more general statement.)

Lemma 5. Let $f(\mathbf{x})$ be periodic with respect to $V = [-\frac{1}{2}, \frac{1}{2}]^d$ and smooth everywhere except at $\mathbf{x} = \mathbf{0}$ with singularity order $\gamma \geq -d + 1$. At $\mathbf{x} = \mathbf{0}$, $f(\mathbf{x})$ is set to 0. The quadrature error of a trapezoidal rule using an m^d-sized uniform mesh $\mathcal{X}$ that contains $\mathbf{x} = \mathbf{0}$ can be estimated as

$$|\mathcal{E}_V(f, \mathcal{X})| \leq Cm^{-(d+\gamma)}.$$

If $f(\mathbf{0})$ is set to an $\mathcal{O}(1)$ value in the calculation, it introduces additional $\mathcal{O}(m^{-d})$ quadrature error.

Combining this error estimate with the integrand singularity structure for orbital energies (with $m = N_{\mathbf{k}}^{1/3}$, $d = 3$, and $\gamma = -2, 0$ for occupied and virtual orbitals, respectively), we obtain the finite-size error estimate for orbital energy calculation as

$$\left|\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}}-\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}\right| \leq C\begin{cases} N_{\mathbf{k}}^{-1/3} & n \text{ is occupied}, \\ N_{\mathbf{k}}^{-1} & n \text{ is virtual}. \end{cases} \quad (16)$$

To adapt the above approach for examining energy and ERI-contraction calculations with exact amplitude, we need to characterize the singularity structure of the exact CCD(n) amplitude entries. First, we note that the exact CCD(1) amplitude is just the MP2 amplitude

$$\left\{\langle a\mathbf{k}_a, b\mathbf{k}_b|i\mathbf{k}_i, j\mathbf{k}_j\rangle(\varepsilon_{i\mathbf{k}_i, j\mathbf{k}_j, a\mathbf{k}_a, b\mathbf{k}_b}^{\mathrm{TDL}})^{-1}\right\}.$$

As a function of $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$, each CCD(1) amplitude entry indexed by (i, j, a, b) has the same singularity structure as the included ERI term, which is singular at $\mathbf{k}_a - \mathbf{k}_i = \mathbf{0}$ of order 0. It turns out that the exact CCD(n) amplitude with any $n > 0$ all has the same singularity structure as $\langle a\mathbf{k}_a, b\mathbf{k}_b|i\mathbf{k}_i, j\mathbf{k}_j\rangle$.

Lemma 6 (singularity structure of the amplitude, Lemma 4 in Ref. [41]). In CCD(n) calculation with $n > 0$, each entry of the exact double amplitude $t_n = \{[t_n]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\}$ belongs to the following function space:

$$\mathbb{T}(\Omega^*) = \{f(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\colon f \text{ is periodic with respect to } \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \Omega^*,$$
$$f \text{ is smooth everywhere except at } \mathbf{k}_a = \mathbf{k}_i \text{ with algebraic singularity of order } 0,$$
$$f \text{ is smooth with respect to } \mathbf{k}_i, \mathbf{k}_j \text{ at the singularity } \mathbf{k}_a = \mathbf{k}_i\}.$$

Combining the above singularity structure characterizations of ERIs and exact CCD(n) amplitudes, we are able to analyze the integrands for energy and ERI-contraction calculations in Eq. (14). Take the CCD(n) exchange energy term as an example whose finite-size error can be formulated as

$$-\frac{1}{|\Omega^*|^3}\mathcal{E}_{\Omega^*\times\Omega^*\times\Omega^*}\left(\sum_{ijab}\langle i\mathbf{k}_i, j\mathbf{k}_j|b\mathbf{k}_b, a\mathbf{k}_a\rangle t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a), \mathcal{K}\times\mathcal{K}\times\mathcal{K}\right),$$

with integration variables $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$. For each set of (i, j, a, b), both the associated ERI and exact amplitude in the integrand exhibit an algebraic singularity of order 0 at $\mathbf{k}_a - \mathbf{k}_j = \mathbf{0}$ and $\mathbf{k}_a - \mathbf{k}_i = \mathbf{0}$, respectively. For such a product of two functions with algebraic singularities, we also provide a rigorous quadrature error estimate similar to Lemma 5. Specifically, for all the integrands in the energy and ERI-contraction calculations, their quadrature errors are determined by the most singular product components in the integrand. The quadrature error still scales as $m^{-(d+\gamma)}$ similar to Lemma 5 but with γ denoting the minimum algebraic singularity order of all ERIs and exact amplitudes.

For the exchange term above, all involved ERIs and exact amplitudes have point singularities of order 0 and similar for the direct term. Thus, we can get the finite-size error estimate for the energy calculation using exact CCD(n) amplitude as

$$|\mathcal{G}_{\mathrm{TDL}}(t) - \mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t)| \leq CN_{\mathbf{k}}^{-1}. \quad (17)$$

Similar analysis is also applicable to ERI contractions using exact amplitudes. The key distinction lies in the fact that the ERI contractions involve integrands formed by ERIs with stronger singularities, such as those defined by particle-particle or hole-hole diagrams. A prominent example is the $4h2p$ linear term in Eq. (5), where the involved ERI $\langle k\mathbf{k}_k, l\mathbf{k}_l|i\mathbf{k}_i, j\mathbf{k}_j\rangle$ exhibits an algebraic singularity of order -2 at $\mathbf{k}_k = \mathbf{k}_i$ when $k = i$ and $l = j$. Consequently, as per the above analysis, the finite-size error in the $4h2p$ linear term calculation alone scales as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$. This error turns out to dominate the overall finite-size error in the ERI-contraction calculation, and we have

$$\left|[\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\mathrm{TDL}}(t) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\right|$$
$$\leq CN_{\mathbf{k}}^{-1/3}, \quad \forall i, j, a, b, \quad \forall \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}. \quad (18)$$

B. Madelung constant correction as a quadrature error reduction method

To reduce the quadrature error for a singular integrand, one common numerical quadrature technique is the singularity subtraction method [45]. Essentially, this method involves constructing an auxiliary function h that possesses the same leading singularity as the integrand g. The integral is then approximated as

$$\int_V d\mathbf{x} g(\mathbf{x}) \approx \frac{|V|}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}(g-h)(\mathbf{x}) + \int_V d\mathbf{x} h(\mathbf{x})$$
$$= \frac{|V|}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}g(\mathbf{x}) + \left(\int_V d\mathbf{x} - \frac{|V|}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}\right)h(\mathbf{x}). \quad (19)$$

This approximation consists of the numerical quadrature of $g - h$ and the exact integral of h (which can be computed analytically or numerically with high precision). It is also equivalent to adding a correction $\mathcal{E}_V(h, \mathcal{X})$ to the numerical quadrature of g. By this correction, the quadrature error changes from $\mathcal{E}_V(g, \mathcal{X})$ to $\mathcal{E}_V(g - h, \mathcal{X})$. Since h removes the leading singularity of g in the subtraction $g - h$, $\mathcal{E}_V(g - h, \mathcal{X})$ can be asymptotically smaller than $\mathcal{E}_V(g, \mathcal{X})$.

The Madelung constant defined in Eq. (7) can be reformulated using an arbitrary $\sigma > 0$ as

$$\xi = -\frac{1}{|\Omega^*|}\left(\int_{\Omega^*} d\mathbf{q} - \frac{|\Omega^*|}{N_\mathbf{k}}\sum_{\mathbf{q}\in\mathcal{K}_\mathbf{q}}\right)\sum_{\mathbf{G}\in\mathbb{L}^*}\frac{4\pi}{|\Omega|}\frac{e^{-\sigma|\mathbf{q}+\mathbf{G}|^2}}{|\mathbf{q}+\mathbf{G}|^2}$$
$$+ \mathcal{O}(N_\mathbf{k}^{-1}). \tag{20}$$

Compared to Eq. (19), this representation connects ξ to the singularity subtraction correction defined by an auxiliary function

$$h_\sigma(\mathbf{q}) = \sum_{\mathbf{G}\in\mathbb{L}^*}\frac{4\pi}{|\Omega|}\frac{e^{-\sigma|\mathbf{q}+\mathbf{G}|^2}}{|\mathbf{q}+\mathbf{G}|^2}.$$

The effectiveness of the Madelung constant correction to reduce the finite-size error can be rigorously explained by this connection.

Taking the occupied orbital energy $\varepsilon_{n\mathbf{k}}$ as an example, its exchange portion leads to the dominant finite-size error, and the associated Madelung constant correction modifies the calculation as (with a change of variable $\mathbf{k}_i \to \mathbf{k} - \mathbf{q}$)

$$\frac{1}{N_\mathbf{k}}\sum_{\mathbf{q}\in\mathcal{K}_\mathbf{q}}\left(\sum_i\langle i(\mathbf{k}-\mathbf{q}), n\mathbf{k}|n\mathbf{k}i(\mathbf{k}-\mathbf{q})\rangle\right) - \xi.$$

Comparing this calculation with Eqs. (19) and (20), the Madelung constant correction exactly uses the auxiliary function $h_\sigma(\mathbf{q})$ to remove the leading singularity (i.e., $4\pi/|\Omega||\mathbf{q}|^{-2}$) of the target integrand and, thus, reduces the associated finite-size error asymptotically to $\mathcal{O}(N_\mathbf{k}^{-1})$.

One major technical contribution of this paper is to rigorously prove the effectiveness of the Madelung constant correction for reducing the finite-size error in the ERI-contraction calculations following the same singularity subtraction interpretation. For instance, consider the $4h2p$ linear term calculation in Eq. (5) with any fixed entry index $(i, j, a, b, \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$. Using the change of variable $\mathbf{k}_k \to \mathbf{k}_i - \mathbf{q}$, this term with the Madelung constant correction can be detailed as

$$\frac{1}{N_\mathbf{k}}\sum_{\mathbf{q}\in\mathcal{K}_\mathbf{q}}\left(\sum_{kl}\langle k(\mathbf{k}_i-\mathbf{q}), l(\mathbf{k}_j+\mathbf{q})|i\mathbf{k}_i, j\mathbf{k}_j\rangle t_{klab}(\mathbf{k}_i-\mathbf{q}, \mathbf{k}_j+\mathbf{q}, \mathbf{k}_a)\right) - \xi t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a). \tag{21}$$

The leading singularity of the integrand comes from the product with $(k, l) = (i, j)$. In this product, the ERI is singular at $\mathbf{q} = \mathbf{0}$ of order -2, and the amplitude is singular at $\mathbf{q} = \mathbf{k}_i - \mathbf{k}_a$ of order 0. The correction in Eq. (21) defines a singularity subtraction with an auxiliary function $h_\sigma(\mathbf{q})t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$. This auxiliary function shares exactly the same leading singularity as the integrand at $\mathbf{q} = \mathbf{0}$ due to the ERI term, i.e.,

$$\frac{4\pi}{|\Omega|}\frac{1}{|\mathbf{q}|^2}t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a).$$

Similar to the orbital energy analysis, the finite-size error in this $4h2p$ linear entry can be reduced to $\mathcal{O}(N_\mathbf{k}^{-1})$. However, the key difference here is that this error reduction is the case only for *most but not all* the amplitude entries. The exception is when the amplitude singularity $\mathbf{q} = \mathbf{k}_a - \mathbf{k}_i$ is close or equal to the ERI singularity $\mathbf{q} = \mathbf{0}$, i.e., when computing an ERI-contraction entry whose momentum vector indices $\mathbf{k}_i$ and $\mathbf{k}_a$ are close or identical. In the worst case when $\mathbf{k}_i = \mathbf{k}_a$, the two singularities overlap and the finite-size error of such a $4h2p$ linear entry can be shown to

be still of scale $\mathcal{O}(N_\mathbf{k}^{-1/3})$. Similar analysis also applies to other terms in the ERI-contraction calculation.

In summary, the Madelung constant correction does not uniformly reduce the finite-size errors in the ERI-contraction tensor. This is the case for the $4h2p$ linear term and also for other terms in the ERI contraction. More precisely, we have the following technical error estimate (see Lemma 11 in Appendix B for the general statement and proof).

Lemma 7 (error in ERI contractions). The finite-size error in the ERI contractions using exact CCD(n) amplitude t_n with the Madelung constant correction in Eq. (10) satisfies

$$\left|[\mathcal{M}_\mathcal{K}\mathcal{A}_{\mathrm{TDL}}(t_n) - \mathcal{A}_{N_\mathbf{k},\xi}(\mathcal{M}_\mathcal{K}t_n)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\right|$$
$$\leq C\begin{cases}\frac{1}{|\mathbf{q}_{ia}|^2}N_\mathbf{k}^{-1} & \mathbf{q}_{ia}\neq\mathbf{0},\\ N_\mathbf{k}^{-1/3} & \mathbf{q}_{ia}=\mathbf{0},\end{cases}$$

where $\mathbf{q}_{ia} = \mathbf{k}_a - \mathbf{k}_i + \mathbf{G}_0$ with $\mathbf{G}_0 \in \mathbb{L}^*$ chosen such that $\mathbf{q}_{ia} \in \Omega^*$.

As a result of this nonuniform error reduction, the maximum entrywise finite-size error in the CCD(n)

amplitude calculation $\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}$ with the Madelung constant correction remains $\mathcal{O}(N_\mathbf{k}^{-1/3})$. However, the *average* entrywise error satisfies the bound

$$\frac{1}{N_\mathbf{k}^3} \sum_{ijab} \sum_{\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}} |[\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \le C N_\mathbf{k}^{-1}.$$

From the error decomposition in Eq. (12), such a refined bound is sufficient for our finite-size error analysis.

Therefore, the Madelung constant corrections to the ERI contractions and the orbital energies effectively reduce the finite-size error in the overall CCD(n) energy calculation to $\mathcal{O}(N_\mathbf{k}^{-1})$.

V. NUMERICAL EXAMPLES

To validate the above theoretical analysis, we conduct CCD and CCD(n) calculations on a 3D periodic system of hydrogen dimers. One hydrogen dimer is positioned at the center of each cubic unit cell with an edge length of 6 bohr in the x direction and a separation distance of 1.8 bohr. For each uniform mesh $\mathcal{K}$, we perform an HF calculation on $\mathcal{K}$ to obtain the orbitals and orbital energies. We then perform CCD and CCD(n) calculations under four distinct settings, each with a different combination of the Madelung constant

correction to the orbital energies and the ERI contractions in the amplitude equation. Specifically, we compute CCD and CCD(n) with both corrections, the correction to the orbital energies alone, the correction to the ERI contractions alone, and without any corrections. All the calculations are carried out using the PYSCF package [46] with a minimal basis set gth-szv.

Figure 1 illustrates the numerical results of the CCD(1), CCD(2), CCD(3), and converged CCD calculations. For CCD(1), which is identical to MP2, we have $T_0 = \mathbf{0}$, and the Madelung constant correction to the ERI contractions has no effect. As a result, the curves for the calculations with and without this correction are identical. For CCD(2) and higher, the four correction settings produce distinct curves. Only the CCD(n) calculation with both corrections exhibits a convergence rate of $\mathcal{O}(N_\mathbf{k}^{-1})$, while the other three calculations have convergence rates of $\mathcal{O}(N_\mathbf{k}^{-1/3})$. This highlights the importance of taking into account the Madelung constant correction to both the orbital energies and ERI contractions in higher-level CCD(n) calculations. As CCD(n) converges to CCD, the difference between the calculation with both corrections and the one without any corrections gradually diminishes to zero, and the finite-size error satisfies the inverse volume scaling.
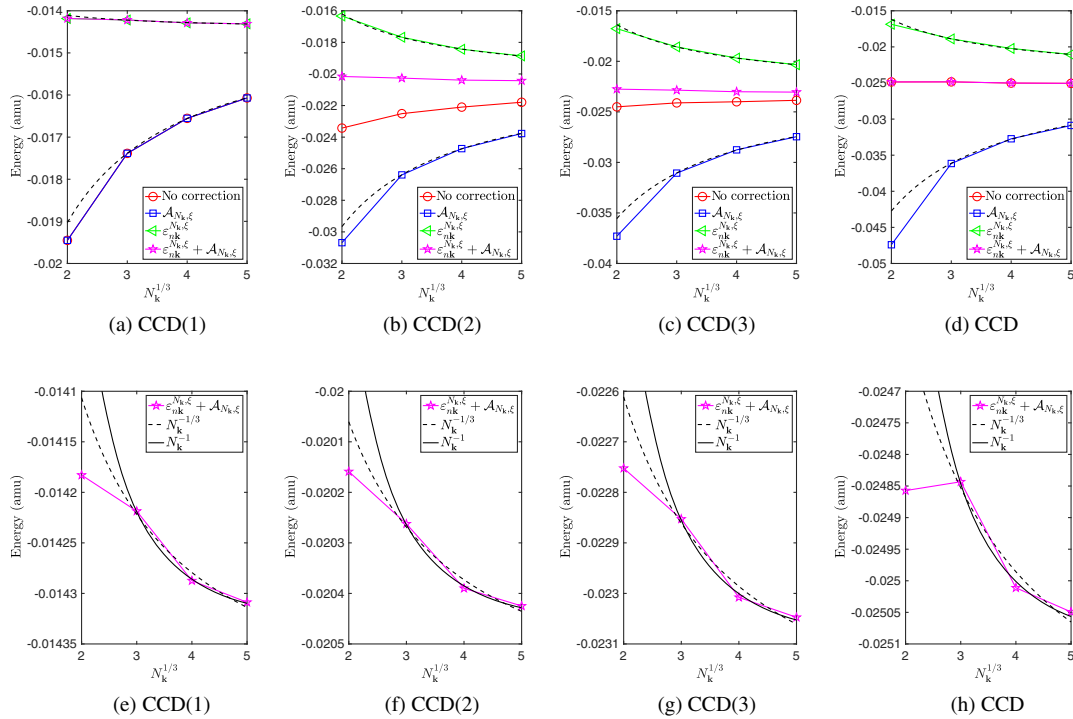

FIG. 1. Convergence of the CCD(1), CCD(2), CCD(3), and CCD correlation energies for a 3D periodic system of hydrogen dimers with increasing $N_\mathbf{k}$. The settings are distinguished by the presence or absence of the Madelung constant correction to the orbital energies ("$\varepsilon_{n\mathbf{k}}^{N_\mathbf{k}, \xi}$") and the ERI contraction ("$\mathcal{A}_{N_\mathbf{k}, \xi}$"). In (a)–(d), the dashed curves show the power-law fitting using $C_0 + C_1 N_\mathbf{k}^{-1/3}$ and data points $N_\mathbf{k}^{1/3} = 3, 4, 5$ for the two cases with partial Madelung constant correction. (e)–(h) plot the curve fittings using $N_\mathbf{k}^{-1/3}$ and $N_\mathbf{k}^{-1}$ over the calculations with correction to both components, numerically corroborating the inverse volume scaling.

VI. DISCUSSION

Recent years have witnessed significant progresses in leveraging wave function methods to study solids. One important driver of this trend is the potential of these methods to provide systematically improvable results. Among the simplest post-HF methods, MP2 has been applied to increasingly large periodic systems along with the development of efficient implementations utilizing density fitting and parallelization techniques [16,23]. Notably, the next-order method, MP3, has also been applied to periodic systems recently for the first time [24]. Though computationally much more expensive, CC theory has also gained momentum in this context. It is being increasingly employed to compute ground state and band structures for a variety of solid materials including insulators and metallic systems [17–19,22,25]. Furthermore, related theories such as equation-of-motion coupled cluster theory, GW method, and algebraic diagrammatic construction theory also start their utilizations in computing excited state properties of solids [20,21,24]. Because of the omnipresence of finite-size effects, as wave function methods continue to be increasingly applied to periodic systems, the importance of developing a thorough understanding of finite-size effects becomes even more pronounced.

In this work, we fill a gap in understanding the finite-size error in periodic coupled cluster calculation for insulating systems. Notably, we reveal an unexpected inverse volume scaling of this error in CCD calculation. This behavior manifests even in the absence of any finite-size correction schemes, owing to an error cancellation. Our findings, together with the methodologies employed in this study, provide valuable insights for practitioners, method developers, and theorists.

For practitioners, when applying computational quantum chemistry methods to periodic systems, reducing finite-size errors using techniques such as power-law extrapolation requires an in-depth understanding of the error scaling. This is particularly important when calculations are constrained to small-sized systems due to the steep increase of the computational cost with respect to the system size and limited resources. Many production-level quantum chemistry packages that support periodic systems use certain finite-size corrections to reduce errors in the HF exchange energy calculation and the HF orbital energies. For instance, the truncated Coulomb correction scheme [54,55] can be applied to insulating systems, so that the finite-size error in orbital energies decays superalgebraically with respect to $N_{\mathbf{k}}$. However, our analysis shows that if this correction is applied only to the orbital energy but without any correction to ERI contractions, the finite-size error of the correlation energy deteriorates from inverse volume scaling to inverse length scaling. This is also a finding consistent with numerical observations.

For method developers, a key aspect of our result is the connection between the Madelung constant correction and

the singularity subtraction method. This relationship serves not just as a crucial element in our theoretical proof but also points toward new methods for further finite-size error reduction. The Madelung constant correction operates as a one-shot correction, while the singularity subtraction method can be systematically improved to reduce the finite-size error. Our analysis and correction schemes can also be extended to more advanced coupled cluster theories such as CCSD, and coupled cluster singles, doubles, and triples (CCSDT). Exploring the finite-size error in lower-dimensional systems, such as MATBG that requires a different form of the Coulomb kernel [4,56], can lead to different finite-size scaling patterns and novel correction schemes. Going beyond finite-size corrections, the singularity structure of the amplitude, as outlined in Lemma 6, may be of independent interest. The singularity structure imposes important analytic constraints that should be taken into account when developing efficient numerical methods for compressing the CC amplitude tensor.

For theorists, several critical questions persist: How can finite-size error analysis be integrated with the study of complete basis limits to tackle basis set dependence? How should the finite-size error behavior in metallic systems be analyzed, especially when the orbital energy difference in the denominator could vanish? Can the scope of finite-size analysis be broadened to include more complex systems like disordered systems and finite-temperature alloys? These questions present a fertile ground for future research.

APPENDIX A: CCD AMPLITUDE EQUATION

In this section, we use a capital letter to denote an index pair consisting of the orbital index and the $\mathbf{k}$ index. For instance, $I = (i, \mathbf{k}_i)$, $J = (j, \mathbf{k}_j)$, $A = (a, \mathbf{k}_a)$, etc. We use $P \in \{I, J, K, L\}$ to refer to occupied orbitals (also known as holes) and $P \in \{A, B, C, D\}$ to refer to unoccupied orbitals (also known as particles). Any summation $\sum_P$ refers to

summing over all occupied or virtual orbital indices p and all momentum vectors $\mathbf{k}_p \in \mathcal{K}$, while the crystal momentum conservation is enforced according to the summand. This notation is used only in this section to simplify the notation and also to connect the equations to those in the molecular case [9] for better readability.

Using a finite mesh $\mathcal{K}$ of size $N_{\mathbf{k}}$, the normalized CCD amplitude $T_*^{N_{\mathbf{k}}} = \{T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\} := \{t_{IJ}^{AB}\}$ with $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}$ is defined as the solution of the amplitude equation

$$
\begin{aligned}
t_{IJ}^{AB} &= \frac{1}{\varepsilon_{IJAB}^{N_{\mathbf{k}}}} [\mathcal{A}_{N_{\mathbf{k}}}(T_*^{N_{\mathbf{k}}})]_{IJAB} \\
&= \frac{1}{\varepsilon_{IJAB}^{N_{\mathbf{k}}}} \Bigg[\langle AB|IJ \rangle + \mathcal{P}\Bigg(\sum_C \kappa_C^A t_{IJ}^{CB} - \sum_K \kappa_I^K t_{KJ}^{AB} \Bigg) + \frac{1}{N_{\mathbf{k}}} \sum_{KL} \chi_{IJ}^{KL} t_{KL}^{AB} + \frac{1}{N_{\mathbf{k}}} \sum_{CD} \chi_{CD}^{AB} t_{IJ}^{CD} \\
&\quad + \mathcal{P}\Bigg(\frac{1}{N_{\mathbf{k}}} \sum_{KC} (2\chi_{IC}^{AK} - \chi_{CI}^{AK}) t_{KJ}^{CB} - \chi_{IC}^{AK} t_{KJ}^{BC} - \chi_{CJ}^{AK} t_{KI}^{BC} \Bigg) \Bigg], \quad \forall I, J, A, B,
\end{aligned}
\tag{A1}
$$

where $\varepsilon_{IJAB}^{N_{\mathbf{k}}} = \varepsilon_{i\mathbf{k}_i}^{N_{\mathbf{k}}} + \varepsilon_{j\mathbf{k}_j}^{N_{\mathbf{k}}} - \varepsilon_{a\mathbf{k}_a}^{N_{\mathbf{k}}} - \varepsilon_{b\mathbf{k}_b}^{N_{\mathbf{k}}}$, $[\mathcal{A}_{N_{\mathbf{k}}}(T)]_{IJAB} = [\mathcal{A}_{N_{\mathbf{k}}}(T)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$, and $\mathcal{P}$ is a permutation operator defined as $\mathcal{P}(\cdots)_{IJ}^{AB} = (\cdots)_{IJ}^{AB} + (\cdots)_{JI}^{BA}$. This reformulation of the CCD amplitude equation is derived from the CCSD amplitude equation in Ref. [30] by removing all the terms related to single amplitudes and normalizing the involved ERIs and amplitudes (which gives the extra $1/N_{\mathbf{k}}$ factor in the equation and the intermediate blocks). The intermediate blocks in the equation are defined as

$$
\begin{aligned}
\kappa_C^A &= -\frac{1}{N_{\mathbf{k}}^2} \sum_{KLD} (2\langle KL|CD \rangle - \langle KL|DC \rangle) t_{KL}^{AD}, \\
\kappa_I^K &= \frac{1}{N_{\mathbf{k}}^2} \sum_{LCD} (2\langle KL|CD \rangle - \langle KL|DC \rangle) t_{IL}^{CD}, \\
\chi_{IJ}^{KL} &= \langle KL|IJ \rangle + \frac{1}{N_{\mathbf{k}}} \sum_{CD} \langle KL|CD \rangle t_{IJ}^{CD}, \\
\chi_{CD}^{AB} &= \langle AB|CD \rangle, \\
\chi_{IC}^{AK} &= \langle AK|IC \rangle + \frac{1}{2N_{\mathbf{k}}} \sum_{LD} (2\langle LK|DC \rangle - \langle LK|CD \rangle) t_{IL}^{AD} - \langle LK|DC \rangle t_{IL}^{DA}, \\
\chi_{CI}^{AK} &= \langle AK|CI \rangle - \frac{1}{2N_{\mathbf{k}}} \sum_{LD} \langle LK|CD \rangle t_{IL}^{DA},
\end{aligned}
$$

and their momentum vector indices also assume the crystal momentum conservation

$$
\begin{aligned}
\kappa_P^Q &\to \mathbf{k}_p - \mathbf{k}_q \in \mathbb{L}^*, \\
\chi_{PQ}^{RS} &\to \mathbf{k}_p + \mathbf{k}_q - \mathbf{k}_r - \mathbf{k}_s \in \mathbb{L}^*.
\end{aligned}
$$

In the TDL, the amplitude equation for the exact amplitude $t_* = \{t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\} := \{t_{IJ}^{AB}\}$ as functions of $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \Omega^*$ can be formulated by letting $\mathcal{K}$ in Eq. (A1) converge to Ω^* as

$$
\begin{aligned}
t_{IJ}^{AB} &= \frac{1}{\varepsilon_{IJAB}^{\mathrm{TDL}}} [\mathcal{A}_{\mathrm{TDL}}(t_*)]_{IJAB} \\
&= \frac{1}{\varepsilon_{IJAB}^{\mathrm{TDL}}} \Bigg[\langle AB|IJ \rangle + \mathcal{P}\Bigg(\sum_C \kappa_C^A t_{IJ}^{CB} - \sum_K \kappa_I^K t_{KJ}^{AB} \Bigg) + \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{k}_k \sum_{kl} \chi_{IJ}^{KL} t_{KL}^{AB} + \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{k}_c \sum_{cd} \chi_{CD}^{AB} t_{IJ}^{CD} \\
&\quad + \mathcal{P}\Bigg(\frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{k}_k \sum_{kc} (2\chi_{IC}^{AK} - \chi_{CI}^{AK}) t_{KJ}^{CB} - \chi_{IC}^{AK} t_{KJ}^{BC} - \chi_{CJ}^{AK} t_{KI}^{BC} \Bigg) \Bigg], \quad \forall I, J, A, B,
\end{aligned}
\tag{A2}
$$

where the intermediate blocks in the TDL are defined as

$$\kappa_C^A = -\frac{1}{|\Omega^*|^2}\int_{\Omega^*\times\Omega^*} d\mathbf{k}_k d\mathbf{k}_l \sum_{kld}(2\langle KL|CD\rangle - \langle KL|DC\rangle)t_{KL}^{AD},$$

$$\kappa_I^K = \frac{1}{|\Omega^*|^2}\int_{\Omega^*\times\Omega^*} d\mathbf{k}_c d\mathbf{k}_d \sum_{lcd}(2\langle KL|CD\rangle - \langle KL|DC\rangle)t_{IL}^{CD},$$

$$\chi_{IJ}^{KL} = \langle KL|IJ\rangle + \frac{1}{|\Omega^*|}\int_{\Omega^*} d\mathbf{k}_c \sum_{cd}\langle KL|CD\rangle t_{IJ}^{CD},$$

$$\chi_{CD}^{AB} = \langle AB|CD\rangle,$$

$$\chi_{IC}^{AK} = \langle AK|IC\rangle + \frac{1}{2|\Omega^*|}\int_{\Omega^*} d\mathbf{k}_l \sum_{ld}(2\langle LK|DC\rangle - \langle LK|CD\rangle)t_{IL}^{AD} - \langle LK|DC\rangle t_{IL}^{DA},$$

$$\chi_{CI}^{AK} = \langle AK|CI\rangle - \frac{1}{2|\Omega^*|}\int_{\Omega^*} d\mathbf{k}_l \sum_{ld}\langle LK|CD\rangle t_{IL}^{DA}.$$

APPENDIX B: PROOF OF THEOREM 1

Restatement of Theorem 1. In CCD(n) calculation, the finite-size error in the correlation energy scales as $\mathcal{O}(N_\mathbf{k}^{-1/3})$ in each of the following scenarios: (i) there is no finite-size correction, (ii) the Madelung constant correction is applied only to the ERI contraction $\mathcal{A}_{N_\mathbf{k}}$, and (iii) the Madelung constant correction is applied only to the orbital energy $\varepsilon_{n_\mathbf{k}}^{N_\mathbf{k}}$.

When the Madelung constant correction is applied to both $\mathcal{A}_{N_\mathbf{k}}$ and $\varepsilon_{n_\mathbf{k}}^{N_\mathbf{k}}$ in the CCD(n) calculation, the overall finite-size error scales as $\mathcal{O}(N_\mathbf{k}^{-1})$.

As a special case, the same conclusion applies to MP3 calculations.

1. Proof outline

The main context of this paper has provided a brief description of the main proof idea. In this proof, we recap some of the equations and concepts discussed before to make it self-contained. The proofs of Theorem 1 for CCD(n) calculations with various types of corrections are based on the error splitting in Eqs. (12) and (13), i.e.,

$$\left|E_{\text{CCD}(n)}^{\text{TDL}} - E_{\text{CCD}(n)}^{N_\mathbf{k}}\right| \leq |\mathcal{G}_{\text{TDL}}(t_n) - \mathcal{G}_{N_\mathbf{k}}(\mathcal{M}_{\mathcal{K}}t_n)|$$
$$+ C\|\mathcal{M}_{\mathcal{K}}t_n - T_n^{N_\mathbf{k}}\|_1, \quad \text{(B1)}$$

$$\mathcal{M}_{\mathcal{K}}t_n - T_n^{N_\mathbf{k}} = \frac{1}{\varepsilon^{\text{TDL}}}[\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\text{TDL}}(t_{n-1}) - \mathcal{A}_{N_\mathbf{k}}(\mathcal{M}_{\mathcal{K}}t_{n-1})]$$
$$+ \mathcal{A}_{N_\mathbf{k}}(\mathcal{M}_{\mathcal{K}}t_{n-1})\left[\frac{1}{\varepsilon^{\text{TDL}}} - \frac{1}{\varepsilon^{N_\mathbf{k}}}\right]$$
$$+ \frac{1}{\varepsilon^{N_\mathbf{k}}}[\mathcal{A}_{N_\mathbf{k}}(\mathcal{M}_{\mathcal{K}}t_{n-1}) - \mathcal{A}_{N_\mathbf{k}}(T_{n-1}^{N_\mathbf{k}})]. \quad \text{(B2)}$$

Note that the amplitude error here is measured in the average norm as

$$\|T\|_1 = \frac{1}{N_\mathbf{k}^3}\sum_{\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a\in\mathcal{K}}\sum_{ijab}|T_{ijab}(\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a)|.$$

The finite-size error in CCD(n) calculation is, thus, decomposed into the error in energy calculation using exact amplitude, the error in ERI contractions, the error in orbital energies, and the error accumulated from previous iteration. The latter three errors together make up of the error in amplitude calculation. The error in energy calculation using exact amplitude is studied previously in Ref. [41]. For completeness, we provide a brief review of the main results below.

a. Brief review of error in energy calculation with exact amplitude

One basic observation is that this error in CCD(n) calculation can be interpreted as the quadrature error of a specific trapezoidal rule as

$$|\mathcal{G}_{\text{TDL}}(t_n) - \mathcal{G}_{N_\mathbf{k}}(\mathcal{M}_{\mathcal{K}}t_n)|$$
$$= \left|\frac{1}{|\Omega^*|^3}\mathcal{E}_{\Omega^*\times\Omega^*\times\Omega^*}\left(\sum_{ijab}(W_{ijab}[t_n]_{ijab}),\mathcal{K}\times\mathcal{K}\times\mathcal{K}\right)\right|. \quad \text{(B3)}$$

Since both W_{ijab} and $[t_n]_{ijab}$ are periodic with respect to $\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a\in\Omega^*$, the asymptotic scaling of this quadrature error depends on the smoothness of these two components that constitute the integrand.

A general ERI $\langle n_1\mathbf{k}_1,n_2\mathbf{k}_2|n_3\mathbf{k}_3,n_4\mathbf{k}_4\rangle$ can be viewed as a function of momentum vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ and its momentum transfer vector $\mathbf{q} = \mathbf{k}_3 - \mathbf{k}_1$. This function is

periodic with respect to each variable in Ω^*, and its definition in Eq. (1) can be decomposed as

$$\frac{4\pi}{|\Omega|}\frac{\hat{\varrho}_{n_1\mathbf{k}_1,n_3(\mathbf{k}_1+\mathbf{q})}(\mathbf{0})\hat{\varrho}_{n_2\mathbf{k}_2,n_4(\mathbf{k}_2-\mathbf{q})}(\mathbf{0})}{|\mathbf{q}|^2}+\frac{4\pi}{|\Omega|}\sum_{\mathbf{G}\in\mathbb{L}^*\backslash\{\mathbf{0}\}}\frac{\cdots}{|\mathbf{q}+\mathbf{G}|^2}.$$

$$(B4)$$

The numerators of all the fractions are smooth with respect to $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{q}$ [note the assumption that $\psi_{n\mathbf{k}}(\mathbf{r})$ is smooth with respect to $\mathbf{k}$]. Therefore, this ERI example is smooth with respect to $\mathbf{k}_1$ and $\mathbf{k}_2$ and has one point singularity in Ω^* with respect to $\mathbf{q}$ at $\mathbf{q}=\mathbf{0}$, which is due to the first fraction term. For any fixed $\mathbf{k}_i$ and $\mathbf{k}_j$, this point singularity can be characterized using the concept of *algebraic singularity of certain orders*.

Definition 8 (algebraic singularity for univariate functions). A function $f(\mathbf{x})$ has algebraic singularity of order $\gamma \in \mathbb{R}$ at $\mathbf{x}_0 \in \mathbb{R}^d$ if there exists $\delta > 0$ such that

$$\left|\frac{\partial^{\boldsymbol{\alpha}}}{\partial\mathbf{x}^{\boldsymbol{\alpha}}}f(\mathbf{x})\right|\leq C_{\boldsymbol{\alpha},\delta}|\mathbf{x}-\mathbf{x}_0|^{\gamma-|\boldsymbol{\alpha}|},\quad\forall\, 0<|\mathbf{x}-\mathbf{x}_0|<\delta,\quad\forall\boldsymbol{\alpha}\geq 0,$$

where the constant $C_{\boldsymbol{\alpha},\delta}$ depends on δ and the non-negative d-dimensional derivative multi-index $\boldsymbol{\alpha}$. For brevity, f is also said to be singular at $\mathbf{x}_0$ of order γ.

The numerator of the first fraction in Eq. (B4) scales as $\mathcal{O}(|\mathbf{q}|^s)$ with $s\in\{0,1,2\}$ near $\mathbf{q}=\mathbf{0}$ using the orbital orthogonality. The value of s depends on the relation between orbital indices (n_1,n_2) and (n_3,n_4). As a result, the algebraic singularity of the ERI example above at $\mathbf{q}=\mathbf{0}$ with any fixed $\mathbf{k}_1$ and $\mathbf{k}_2$ has order in $\{-2,-1,0\}$. In addition, to connect this singularity with varying $\mathbf{k}_1,\mathbf{k}_2\in\Omega^*$, we introduce the algebraic singularity with respect to one variable for a multivariate function.

Definition 9 (algebraic singularity for multivariate functions). A function $f(\mathbf{x},\mathbf{y})$ is smooth with respect to

$\mathbf{y}\in V_Y\subset\mathbb{R}^{d_y}$ for any fixed $\mathbf{x}$ and has algebraic singularity of order γ with respect to $\mathbf{x}$ at $\mathbf{x}_0\in\mathbb{R}^{d_x}$ if there exists $\delta > 0$ such that

$$\left|\frac{\partial^{\boldsymbol{\alpha}}}{\partial\mathbf{x}^{\boldsymbol{\alpha}}}\left(\frac{\partial^{\boldsymbol{\beta}}}{\partial\mathbf{y}^{\boldsymbol{\beta}}}f(\mathbf{x},\mathbf{y})\right)\right|\leq C_{\boldsymbol{\alpha},\boldsymbol{\beta},\delta}|\mathbf{x}-\mathbf{x}_0|^{\gamma-|\boldsymbol{\alpha}|},$$

$$\forall\, 0<|\mathbf{x}-\mathbf{x}_0|<\delta,\quad\forall\,\mathbf{y}\in V_Y,\quad\forall\boldsymbol{\alpha},\boldsymbol{\beta}\geq 0,$$

where constant $C_{\boldsymbol{\alpha},\boldsymbol{\beta},\delta}$ depends on δ, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$. Compared to the univariate case in Definition 4, the key additions are the shared algebraic singularity of partial derivatives over $\mathbf{y}$ at $\mathbf{x}=\mathbf{x}_0$ of order γ and the independence of $C_{\boldsymbol{\alpha},\boldsymbol{\beta},\delta}$ on $\mathbf{y}\in V_Y$.

By this definition, the ERI $\langle n_1\mathbf{k}_1,n_2\mathbf{k}_2|n_3\mathbf{k}_3,n_4\mathbf{k}_4\rangle$ is smooth everywhere with respect to $\mathbf{k}_i,\mathbf{k}_j,\mathbf{q}\in\Omega^*$ except at $\mathbf{q}=\mathbf{0}$, and the singularity order $\gamma\in\{-2,-1,0\}$. Specifically, γ equals to -2, -1, and 0, respectively, when the orbital indices are fully matched, i.e., $n_1=n_3,n_2=n_4$, partially matched, i.e., $n_1=n_3,n_2\neq n_4$ or $n_1\neq n_3,n_2=n_4$, and not matched, i.e., $n_1\neq n_3,n_2\neq n_4$. If treating the ERI example as a function of $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$ instead, we equivalently claim that the function is singular at $\mathbf{k}_1=\mathbf{k}_3$ of order γ.

One key result in Ref. [41] is the singularity structure characterization for the exact CCD(n) amplitude t_n, which is essential for estimating the quadrature error in Eq. (B3). It turns out that each exact amplitude entry $[t_n]_{ijab}(\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a)$ with any $n>0$ has one point of algebraic singularity of order 0 at $\mathbf{k}_a-\mathbf{k}_i=\mathbf{0}$, sharing a similar singularity structure as the ERI $\langle i\mathbf{k}_i,j\mathbf{k}_j|a\mathbf{k}_a,b\mathbf{k}_b\rangle$ or the exact MP2/CCD(1) amplitude entry $(\varepsilon^{\text{TDL}}_{i\mathbf{k}_i,j\mathbf{k}_j,a\mathbf{k}_a,b\mathbf{k}_b})^{-1}\langle a\mathbf{k}_a,b\mathbf{k}_b|i\mathbf{k}_i,j\mathbf{k}_j\rangle$.

Restatement of Lemma 6 (singularity structure of the amplitude, Lemma 4 in Ref. [41]). In CCD(n) calculation with $n>0$, each entry of the exact double amplitude t_n belongs to the following function space:

$$\mathbb{T}(\Omega^*)=\{f(\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a)\colon f\text{ is periodic with respect to }\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a\in\Omega^*,$$

$$f\text{ is smooth everywhere except at }\mathbf{k}_a=\mathbf{k}_i\text{ with algebraic singularity of order 0,}$$

$$f\text{ is smooth with respect to }\mathbf{k}_i,\mathbf{k}_j\text{ at the singularity }\mathbf{k}_a=\mathbf{k}_i\}.$$

Based on the above singularity structures of ERIs and exact amplitudes, the integrand in the energy calculation in Eq. (B3) consists of products of periodic functions where each has one point singularity of order 0. (Recall that W_{ijab} is the antisymmetrized ERI and consists of two ERIs that can be treated separately.) We provide a sharp quadrature error bound for trapezoidal rules over periodic functions in such a product form, and its application to Eq. (B3) gives the following lemma.

Lemma 10 (energy error with exact amplitude, Lemma 5 in Ref. [41]). In CCD(n) calculation with any $n>0$, the finite-size error in the energy calculation using the exact amplitude t_n can be estimated as

$$|\mathcal{G}_{\text{TDL}}(t_n)-\mathcal{G}_{N_\mathbf{k}}(\mathcal{M}_\mathcal{K}t_n)|\leq CN_\mathbf{k}^{-1},$$

where the constant C depends on t_n.

b. Error in amplitude calculation

Based on the error splitting in Eq. (B2), analysis of the error in amplitude calculation is reduced to estimating the two main error terms $\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\mathrm{TDL}}(t) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t)$ and $\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$ and understanding how $\mathcal{A}_{N_{\mathbf{k}}}$ amplifies the amplitude error $\mathcal{M}_{\mathcal{K}}t_{n-1} - T_{n-1}^{N_{\mathbf{k}}}$ from the previous iteration.

Like the energy calculation, the errors in ERI contractions and orbital energies consist of specific quadrature errors. In both cases, we can show that the Madelung constant correction is connected to certain singularity subtraction methods and can significantly reduce the corresponding dominant quadrature errors.

Lemma 11 (error in ERI contractions). In CCD(n) calculation with any $n > 0$, the finite-size error in the entries of the ERI contractions using the exact amplitude t_n without any corrections can be bounded as

$$\left|[\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\mathrm{TDL}}(t_n) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_n)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\right| \leq C N_{\mathbf{k}}^{-1/3}.$$
(B5)

The Madelung constant correction reduces this error to

$$\left|[\mathcal{M}_{\mathcal{K}}\mathcal{A}_{\mathrm{TDL}}(t_n) - \mathcal{A}_{N_{\mathbf{k}},\xi}(\mathcal{M}_{\mathcal{K}}t_n)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)\right|$$
$$\leq C \begin{cases} \frac{1}{|\mathbf{q}|^2}N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}, \end{cases}$$
(B6)

where $\mathbf{q} = \mathbf{k}_a - \mathbf{k}_i + \mathbf{G}_0$ with $\mathbf{G}_0 \in \mathbb{L}^*$ chosen such that $\mathbf{q} \in \Omega^*$. In both cases, the constant C depends on t_n but not on the entry index (i, j, a, b) and $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \in \mathcal{K} \times \mathcal{K} \times \mathcal{K}$.

Remark 12. The prefactor $1/|\mathbf{q}|^2$ in the above estimate is important when $\mathbf{q} \in \mathcal{K}_{\mathbf{q}}$ is near the origin. For example, if $\mathbf{k}_i, \mathbf{k}_a \in \mathcal{K}$ are adjacent to each other, $|\mathbf{q}|$ is of scale $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ and the estimate in Lemma 11 suggests an error bound of scale $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$.

Lemma 13 (error in orbital energies). The finite-size error in orbital energies without any corrections is bounded as

$$\left|\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}\right| \leq C \begin{cases} N_{\mathbf{k}}^{-1/3} & n \text{ is occupied}, \\ N_{\mathbf{k}}^{-1} & n \text{ is virtual}. \end{cases}$$

The Madelung constant correction reduces this error to

$$\left|\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}},\xi}\right| \leq C N_{\mathbf{k}}^{-1}.$$

In both cases, the constant C is independent of the entry index $n, \mathbf{k} \in \mathcal{K}$.

Note that there are three multipliers for the three error terms in the amplitude error splitting Eq. (B2). These prefactors are bounded by constants independent of $N_{\mathbf{k}}$.

First, the orbital energy difference in Eq. (B2) satisfies $|\varepsilon_{i\mathbf{k}_i, j\mathbf{k}_j, a\mathbf{k}_a, b\mathbf{k}_b}^{\mathrm{TDL}}| \geq 2\varepsilon_g$ by the assumed indirect gap $\varepsilon_{a\mathbf{k}_a}^{\mathrm{TDL}} - \varepsilon_{i\mathbf{k}_i}^{\mathrm{TDL}} \geq \varepsilon_g > 0$. Second, since $\mathcal{A}_{N_{\mathbf{k}}}(T)$ consists of constant, linear, and quadratic terms of T, a straightforward estimate shows (e.g., using Lemma 7 in Ref. [41]) that

$$\max_{ijab, \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}} |[\mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}}t_n)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)|$$
$$\leq C\max_{ijab}\|[t_n]_{ijab}\|_{L^\infty(\Omega^* \times \Omega^* \times \Omega^*)}^2,$$

where the L^∞-norm of $[t_n]_{ijab}$ is finite according to Lemma 6.

Based on the above estimates of these multipliers, we have that the summation of the first two error terms in Eq. (B2) is dominated by the error in the ERI contractions and the orbital energies discussed in Lemmas 11 and 13. If applying the Madelung constant correction to both orbital energies and ERI contractions, the summation of the first two error terms has its entries bounded asymptotically the same as Eq. (B6). Otherwise, its entries are bounded asymptotically as in Eq. (B5). Lastly, for the third error term in Eq. (13), the application of $\mathcal{A}_{N_{\mathbf{k}}}/\mathcal{A}_{N_{\mathbf{k}},\xi}$ can be proved to maintain the entrywise error scaling obtained in Lemma 11.

Lemma 14. Consider two arbitrary bounded amplitude tensors $T, S \in \mathbb{C}^{n_{\mathrm{occ}} \times n_{\mathrm{occ}} \times n_{\mathrm{vir}} \times n_{\mathrm{vir}} \times N_{\mathbf{k}} \times N_{\mathbf{k}} \times N_{\mathbf{k}}}$ and assume their entrywise upper bound independent of $N_{\mathbf{k}}$. If the difference between T and S satisfies an estimate similar to that in Eq. (B5), i.e.,

$$|[T - S]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq C N_{\mathbf{k}}^{-1/3},$$

the ERI-contraction map without any corrections $\mathcal{A}_{N_{\mathbf{k}}}$ satisfies

$$|[\mathcal{A}_{N_{\mathbf{k}}}(T) - \mathcal{A}_{N_{\mathbf{k}}}(S)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq C N_{\mathbf{k}}^{-1/3}.$$

If the difference between T and S satisfies an estimate similar to that in Eq. (B6), i.e.,

$$|[T - S]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq C \begin{cases} \frac{1}{|\mathbf{q}|^2}N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}, \end{cases}$$

the ERI-contraction map with Madelung constant correction $\mathcal{A}_{N_{\mathbf{k}},\xi}$ satisfies

$$|[\mathcal{A}_{N_{\mathbf{k}},\xi}(T) - \mathcal{A}_{N_{\mathbf{k}},\xi}(S)]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq C \begin{cases} \frac{1}{|\mathbf{q}|^2}N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}. \end{cases}$$

Combining the estimates of the three error terms in Eq. (13) and the initial condition $\mathcal{M}_{\mathcal{K}}t_0 = T_0^{N_{\mathbf{k}}} = \mathbf{0}$ in both cases with and without Madelung constant correction, we

can obtain the entrywise estimate of the error in the CCD(n) amplitude calculation recursively. First, for the CCD(n) calculation without any corrections or with partial Madelung constant corrections to either orbital energies or ERI contractions, we have

$$|[\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \le C N_\mathbf{k}^{-1/3}. \quad \text{(B7)}$$

Accordingly, the average entrywise error can be estimated as

$$
\begin{aligned}
&\|\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}\|_1 \\
&= \frac{1}{N_\mathbf{k}^3} \sum_{\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}} \sum_{ijab} |[\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \\
&\le C N_\mathbf{k}^{-1/3}.
\end{aligned}
$$

Second, for the CCD(n) calculation with the Madelung constant correction to both orbital energies and ERI contractions, we have

$$|[\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \le C \begin{cases} \frac{1}{|\mathbf{q}|^2} N_\mathbf{k}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_\mathbf{k}^{-1/3} & \mathbf{q} = \mathbf{0}. \end{cases} \quad \text{(B8)}$$

Accordingly, the average entrywise error can be estimated as

$$
\begin{aligned}
&\|\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}\|_1 \\
&= \frac{1}{N_\mathbf{k}^3} \sum_{\mathbf{k}_i, \mathbf{k}_j \in \mathcal{K}} \sum_{\mathbf{q} \in \mathcal{K}_\mathbf{q}} \sum_{ijab} |[\mathcal{M}_\mathcal{K} t_n - T_n^{N_\mathbf{k}}]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_i + \mathbf{q})| \\
&\le C \frac{1}{N_\mathbf{k}} \sum_{\mathbf{q} \in \mathcal{K}_\mathbf{q} \setminus \{\mathbf{0}\}} \frac{1}{|\mathbf{q}|^2} N_\mathbf{k}^{-1} + C \frac{1}{N_\mathbf{k}} N_\mathbf{k}^{-1/3} \\
&\le C N_\mathbf{k}^{-1}.
\end{aligned}
$$

Plugging the above estimate and the error estimate for energy calculation in Lemma 10 into Eq. (12) then finishes the proof of Theorem 1.

Remark 15. In our previous work [41], we use the maximum entrywise norm to characterize the finite-size error in the amplitude calculation, which loosens the average entrywise norm used in the error splitting in Eq. (B1) as

$$\|T\|_1 \le \|T\|_\infty := \max_{ijab, \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}} |T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)|.$$

Without the Madelung constant correction, the maximum norm suffices since all entries in the amplitude error are of the same scale, as shown in Eq. (B7). However, this norm is no longer sufficient for the calculation with the Madelung constant correction, bounded in Eq. (B8). Because the maximum entrywise error is now of scale $\mathcal{O}(N_\mathbf{k}^{-1/3})$, while most entries are of scale $\mathcal{O}(N_\mathbf{k}^{-1})$. In this case, the average entrywise norm provides a necessary and tighter estimate of amplitude error.

2. Proof of Lemma 11: Error in ERI contractions

According to the singularity structure of exact CCD(n) amplitude in Lemma 6, we consider the ERI contraction using an arbitrary exact amplitude $t \in \mathbb{T}(\Omega^*)^{n_\text{occ} \times n_\text{occ} \times n_\text{vir} \times n_\text{vir}}$. Fixing a set of entry index (i, j, a, b) and $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \in \mathcal{K} \times \mathcal{K} \times \mathcal{K}$, the error in the indexed ERI-contraction entry can be detailed as [by comparing Eqs. (A1) and (A2)]

$$
\begin{aligned}
[\mathcal{M}_\mathcal{K} \mathcal{A}_\text{TDL}(t) - \mathcal{A}_{N_\mathbf{k}, \xi}(\mathcal{M}_\mathcal{K} t)]_{ijab, \mathbf{k}_i \mathbf{k}_j \mathbf{k}_a} &= \frac{1}{|\Omega^*|} \left[\mathcal{E}_{\Omega^*} \left(\sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle t_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a), \mathcal{K} \right) + |\Omega^*| \xi t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \right] \\
&\quad + \frac{1}{|\Omega^*|^2} \mathcal{E}_{\Omega^* \times \Omega^*} \left(\sum_{klcd} \langle k\mathbf{k}_k, l\mathbf{k}_l | c\mathbf{k}_c, d\mathbf{k}_d \rangle t_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c) t_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a), \mathcal{K} \times \mathcal{K} \right) \\
&\quad + \cdots, \quad \text{(B9)}
\end{aligned}
$$

where the constant terms cancel with each other, the first term is the error in the $4h2p$ linear term calculation with the Madelung constant correction, and the second term is the error in the $4h2p$ quadratic term calculation. The neglected ones are the errors in remaining linear and quadratic term calculations, which can all be similarly formulated as quadrature errors of specific trapezoidal rules.

In the error analysis for CCD calculation without corrections in ERI contractions [41], it has been shown

that without the Madelung constant correction the error entry in the ERI contractions is uniformly bounded as

$$\left| [\mathcal{M}_\mathcal{K} \mathcal{A}_\text{TDL}(t) - \mathcal{A}_{N_\mathbf{k}}(\mathcal{M}_\mathcal{K} t)]_{ijab, \mathbf{k}_i \mathbf{k}_j \mathbf{k}_a} \right| \le C N_\mathbf{k}^{-1/3}. \quad \text{(B10)}$$

More specifically, all the quadratic terms and part of the linear terms that contain only ERIs with mismatched orbital indices contribute at most $\mathcal{O}(N_\mathbf{k}^{-1})$ errors in both Eqs. (B9)

and (B10). The dominant error in Eq. (B10) comes from the calculation of the following six linear terms:

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle t_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a),$$

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_c \in \mathcal{K}} \sum_{cd} \langle a\mathbf{k}_a, b\mathbf{k}_b | c\mathbf{k}_c, d\mathbf{k}_d \rangle t_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c),$$

$$-\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kc} \langle a\mathbf{k}_a, k\mathbf{k}_k | c\mathbf{k}_c, i\mathbf{k}_i \rangle t_{kjcb}(\mathbf{k}_k, \mathbf{k}_j, \mathbf{k}_c),$$

$$-\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kc} \langle b\mathbf{k}_b, k\mathbf{k}_k | c\mathbf{k}_c, j\mathbf{k}_j \rangle t_{kica}(\mathbf{k}_k, \mathbf{k}_i, \mathbf{k}_c),$$

$$-\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kc} \langle a\mathbf{k}_a, k\mathbf{k}_k | c\mathbf{k}_c, j\mathbf{k}_j \rangle t_{kibc}(\mathbf{k}_k, \mathbf{k}_i, \mathbf{k}_b),$$

$$-\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kc} \langle b\mathbf{k}_b, k\mathbf{k}_k | c\mathbf{k}_c, i\mathbf{k}_i \rangle t_{kjac}(\mathbf{k}_k, \mathbf{k}_j, \mathbf{k}_a),$$

which contain ERIs with fully or partially matched orbital indices. The Madelung constant correction in Eq. (B9) is only triggered in these six terms. In this proof, we focus on the error estimate for the $4h2p$ linear term (the first term above) with the correction, and similar analysis can be applied to all the other five terms.

Denote the ERI-amplitude product in the $4h2p$ linear term with orbital indices (k, l) as

$$F^{kl}(\mathbf{q}_1) = \langle k(\mathbf{k}_i - \mathbf{q}_1), l(\mathbf{k}_j + \mathbf{q}_1) | i\mathbf{k}_i, j\mathbf{k}_j \rangle$$
$$\times t_{klab}(\mathbf{k}_i - \mathbf{q}_1, \mathbf{k}_j + \mathbf{q}_1, \mathbf{k}_a)$$
$$= H_{\text{eri}}^{kl}(\mathbf{q}_1) H_{\text{amp}}^{kl}(\mathbf{q}_1),$$

where $\mathbf{q}_1 = \mathbf{k}_i - \mathbf{k}_k$ is the momentum transfer vector of the ERI. The $4h2p$ linear term calculation with the Madelung constant correction using a finite mesh $\mathcal{K}$ can be reformulated as

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} F^{kl}(\mathbf{k}_i - \mathbf{k}_k) - \xi H_{\text{amp}}^{ij}(\mathbf{0})$$

$$= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}}} \sum_{kl} F^{kl}(\mathbf{q}_1) - \xi H_{\text{amp}}^{ij}(\mathbf{0}),$$

using the change of variable $\mathbf{k}_k \to \mathbf{k}_i - \mathbf{q}_1$ and the periodicity of $F^{kl}(\mathbf{q}_1)$. The error of this calculation compared to its TDL value, i.e., the first error term in Eq. (B9), can be written as

$$\frac{1}{|\Omega^*|} \left(\sum_{kl} \mathcal{E}_{\Omega^*}(F^{kl}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}) + |\Omega^*| \xi H_{\text{amp}}^{ij}(\mathbf{0}) \right).$$

Previously, in Ref. [41], the quadrature error for $F^{kl}(\mathbf{q}_1)$ with varying (k, l) is estimated as

$$|\mathcal{E}_{\Omega^*}(F^{kl}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}})| \leq C \begin{cases} N_{\mathbf{k}}^{-1} & k \neq i, l \neq j, \\ N_{\mathbf{k}}^{-2/3} & k = i, l \neq j \text{ or } k \neq i, l = j, \quad \forall i, j, a, b, \quad \forall \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}. \\ N_{\mathbf{k}}^{-1/3} & k = i, l = j, \end{cases} \tag{B11}$$

As demonstrated next, in the case of partially matched orbital indices (e.g., $k = i, l \neq j$), this error estimate turns out to be loose when $\mathbf{k}_i \neq \mathbf{k}_a$ and can be further improved as

$$|\mathcal{E}_{\Omega^*}(F^{il}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}})| \leq C \begin{cases} \frac{1}{|\mathbf{q}|} N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-2/3} & \mathbf{q} = \mathbf{0}, \end{cases} \quad \forall i, j, a, b, \quad \forall \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}, \tag{B12}$$

where $\mathbf{q} = \mathbf{k}_a - \mathbf{k}_i + \mathbf{G}_0$ with $\mathbf{G}_0 \in \mathbb{L}^*$ chosen to make $\mathbf{q} \in \Omega^*$. In the case of fully matched orbital indices ($k = i, l = j$), the Madelung constant correction is triggered in the ERI evaluation and can help remove the leading quadrature error when $\mathbf{k}_i \neq \mathbf{k}_a$ as

$$|\mathcal{E}_{\Omega^*}(F^{ij}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}) + |\Omega^*| \xi H_{\text{amp}}^{ij}(\mathbf{0})| \leq C \begin{cases} \frac{1}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}, \end{cases} \quad \forall i, j, a, b, \quad \forall \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}. \tag{B13}$$

In the following discussion, we prove these two error estimates Eqs. (B12) and (B13) respectively at $\mathbf{k}_i \neq \mathbf{k}_a$.

a. Error estimate for linear terms with partially matched orbital indices

Consider a fixed set of (i, j, a, b) and $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}$ with $\mathbf{k}_i \neq \mathbf{k}_a$. Assume $k = i$ and $l \neq j$. Our target is a sharper estimate of $\mathcal{E}_{\Omega^*}(F^{il}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}})$. Since $\mathcal{K}_{\mathbf{q}}$ and Ω^* are both inversely symmetric over $\mathbf{q}_1 = \mathbf{0}$, the quadrature error can be symmetrized as

$$\mathcal{E}_{\Omega^*}(F^{il}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}) = \frac{1}{2}\mathcal{E}_{\Omega^*}(H_{\text{eri}}^{il}(\mathbf{q}_1)H_{\text{amp}}^{il}(\mathbf{q}_1)$$
$$+ H_{\text{eri}}^{il}(-\mathbf{q}_1)H_{\text{amp}}^{il}(-\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}).$$

This symmetrized integrand can be further decomposed into two terms:

$$[(H_{\text{eri}}^{il}(\mathbf{q}_1) + H_{\text{eri}}^{il}(-\mathbf{q}_1))H_{\text{amp}}^{il}(\mathbf{q}_1)]$$
$$- [H_{\text{eri}}^{il}(-\mathbf{q}_1)(H_{\text{amp}}^{il}(\mathbf{q}_1) - H_{\text{amp}}^{il}(-\mathbf{q}_1))]. \quad \text{(B14)}$$

For the first term in Eq. (B14), we note that the ERI term can be detailed as

$$H_{\text{eri}}^{il}(\mathbf{q}_1) = \frac{4\pi}{|\Omega|} \frac{\hat{\varrho}_{i(\mathbf{k}_i - \mathbf{q}_1), i\mathbf{k}_i}(\mathbf{0})\hat{\varrho}_{l(\mathbf{k}_j + \mathbf{q}_1), j\mathbf{k}_j}(\mathbf{0})}{|\mathbf{q}_1|^2}$$
$$+ \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \backslash \{\mathbf{0}\}} \frac{\cdots}{|\mathbf{q}_1 + \mathbf{G}|^2}.$$

The ERI nonsmoothness with $\mathbf{q}_1 \in \Omega^*$ comes from the first fractional term whose numerator is smooth and scales as $\mathcal{O}(|\mathbf{q}_1|)$ near $\mathbf{q}_1 = \mathbf{0}$ using orbital orthogonality. It can be verified directly that $H_{\text{eri}}^{il}(\mathbf{q}_1)$ is singular at $\mathbf{q}_1 = \mathbf{0}$ of order -1, and after symmetrization $H_{\text{eri}}^{il}(\mathbf{q}_1) + H_{\text{eri}}^{il}(-\mathbf{q}_1)$ is singular at $\mathbf{q}_1 = \mathbf{0}$ of order 0. Meanwhile, the amplitude $H_{\text{amp}}^{il}(\mathbf{q}_1)$ is smooth everywhere in Ω^* except at $\mathbf{q}_1 = -\mathbf{q}$ of order 0 according to Lemma 6. An error estimate lemma in Ref. [41] (restated as Lemma 24 in Appendix D) provides a quadrature error estimate for periodic functions in such a product form and can show that

$$|\mathcal{E}_{\Omega^*}((H_{\text{eri}}^{il}(\mathbf{q}_1) + H_{\text{eri}}^{il}(-\mathbf{q}_1))H_{\text{amp}}^{il}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}})| \leq C N_{\mathbf{k}}^{-1},$$

where constant C is independent of i, j, a, b and $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}$.

For the second term in Eq. (B14), direct application of Lemma 24 leads to error estimate of scale $\mathcal{O}(N_{\mathbf{k}}^{-2/3})$, since $H_{\text{eri}}^{il}(-\mathbf{q}_1)$ is singular at $\mathbf{q}_1 = \mathbf{0}$ of order -1. However, note that $H_{\text{amp}}^{il}(\mathbf{q}_1)$ is smooth at $\mathbf{q}_1 = \mathbf{0}$ [recall that $H_{\text{amp}}^{il}(\mathbf{q}_1)$ is singular only at $\mathbf{q}_1 = -\mathbf{q} \neq \mathbf{0}$], and, thus, the subtraction $H_{\text{amp}}^{il}(\mathbf{q}_1) - H_{\text{amp}}^{il}(-\mathbf{q}_1)$ scales as $\mathcal{O}(|\mathbf{q}_1|)$ near $\mathbf{q}_1 = \mathbf{0}$. Multiplication by this extra $\mathcal{O}(|\mathbf{q}_1|)$ term improves the algebraic singularity of $H_{\text{eri}}^{il}(\mathbf{q}_1)$ at $\mathbf{q}_1 = \mathbf{0}$, and the overall product can be shown to be singular at $\mathbf{q}_1 = \mathbf{0}$ of order 0.

Intuitively, this improved algebraic singularity at $\mathbf{q}_1 = \mathbf{0}$ can lead to asymptotically smaller quadrature errors. To rigorously justify this statement, we generalize Lemma 24 to estimate the quadrature error for this special case.

Lemma 16. Let $f(\mathbf{x}) = f_1(\mathbf{x})f_2(\mathbf{x})$, where $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are periodic with respect to $V = [-\frac{1}{2}, \frac{1}{2}]^d$ and
   (i) $f_1(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_1 = \mathbf{0}$ of order $\gamma \leq -1$,
   (ii) $f_2(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_2 \neq \mathbf{0}$ of order 0,
   (iii) $\partial_{\mathbf{x}}^{\alpha} f_2(\mathbf{0}) = 0$ for any derivative order $|\alpha| \leq s$.
Assume $\gamma > -d$ for $f(\mathbf{x})$ to be integrable in V and $\gamma + s + 1 \leq 0$ so the leading algebraic singularity of $f(\mathbf{x})$ is at $\mathbf{x} = \mathbf{0}$. Consider an m^d-sized uniform mesh $\mathcal{X}$ in V. Assume that $\mathcal{X}$ satisfies that $\mathbf{z}_1, \mathbf{z}_2$ are either on the mesh or $\Theta(m^{-1})$ away from any mesh points and m is sufficiently large that $|\mathbf{z}_1 - \mathbf{z}_2| = \Omega(m^{-1})$. At $\mathbf{x} = \mathbf{z}_1$ and $\mathbf{x} = \mathbf{z}_2$, $f(\mathbf{x})$ is set to 0. The trapezoidal rule using $\mathcal{X}$ has quadrature error

$$|\mathcal{E}_V(f, \mathcal{X})| \leq C\mathcal{H}_{V, \mathbf{z}_1}^{d+1}(f_1)\mathcal{H}_{V, \mathbf{z}_2}^{d+1}(f_2)|\mathbf{z}_1 - \mathbf{z}_2|^{-(s+1)}$$
$$\times m^{-(d+\gamma+s+1)}.$$

Remark 17. The two factors $\mathcal{H}_{V, \mathbf{z}_1}^{d+1}(f_1)$ and $\mathcal{H}_{V, \mathbf{z}_2}^{d+1}(f_2)$ characterize the algebraic singularities of the two functions, and their exact definition can be found in Appendix D. Proof of Lemma 16 is provided in Appendix D 2.

In order to utlize this result, we further decompose the second term in Eq. (B14) into

$$[H_{\text{eri}}^{il}(-\mathbf{q}_1)(H_{\text{amp}}^{il}(\mathbf{q}_1) - H_{\text{amp}}^{il}(\mathbf{0}))]$$
$$+ [H_{\text{eri}}^{il}(-\mathbf{q}_1)(H_{\text{amp}}^{il}(\mathbf{0}) - H_{\text{amp}}^{il}(-\mathbf{q}_1))].$$

Applying Lemma 16 to both terms with $\gamma = -1$ and $s = 0$ gives

$$|\mathcal{E}_{\Omega^*}(H_{\text{eri}}^{il}(-\mathbf{q}_1)(H_{\text{amp}}^{il}(\mathbf{q}_1) - H_{\text{amp}}^{il}(-\mathbf{q}_1)), \mathcal{K}_{\mathbf{q}})| \leq C\frac{1}{|\mathbf{q}|}N_{\mathbf{k}}^{-1},$$

where constant C is independent of i, j, a, b and $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}$. Combining the above quadrature error estimates for the two terms in Eq. (B14), we prove a tighter error estimate at $\mathbf{q} \neq \mathbf{0}$ shown in Eq. (B12) while the previous result in Eq. (B11) at $\mathbf{q} = \mathbf{0}$ still holds.

b. Error estimate for linear terms with fully matched orbital indices

Let $k = i$, $l = j$ and $\mathbf{k}_i \neq \mathbf{k}_a$ and consider the corresponding calculation in the $4h2p$ linear term. The Madelung constant correction is applied in the ERI evaluation at $\mathbf{k}_k = \mathbf{k}_i$ or, equivalently, at $\mathbf{q}_1 = \mathbf{0}$, and the calculation can be written as

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}}} F^{ij}(\mathbf{q}_1) - \xi H^{ij}_{\mathrm{amp}}(\mathbf{0})$$

$$= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}}} \left(H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{0}) \right)$$

$$+ \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{q}_1 h_\sigma(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{0}) + \mathcal{O}(N_{\mathbf{k}}^{-1}),$$

which uses the expansion of the Madelung constant with an arbitrary fixed parameter $\sigma > 0$ in Eq. (20). Note that the prefactor in the $\mathcal{O}(N_{\mathbf{k}}^{-1})$ remainder term above depends only on σ. The right-hand side of the above reformation is equivalent to a singularity subtraction method that decomposes the original integrand into two terms:

$$F^{ij}(\mathbf{q}_1) = H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1)$$

$$= \left(H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{0}) \right)$$

$$+ h_\sigma(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{0}),$$

and then computes the numerical quadrature of the first term and the exact integral of the second term. The Madelung-corrected calculation, thus, has quadrature error only from the first term as

$$\frac{1}{|\Omega^*|} \mathcal{E}_{\Omega^*}(F^{ij}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}) + \xi H^{ij}_{\mathrm{amp}}(\mathbf{0})$$

$$= \frac{1}{|\Omega^*|} \mathcal{E}_{\Omega^*} \left(H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1) H^{ij}_{\mathrm{amp}}(\mathbf{0}), \mathcal{K}_{\mathbf{q}} \right)$$

$$+ \mathcal{O}(N_{\mathbf{k}}^{-1}).$$

Following a similar approach as in the partially matched case, the effective integrand after the Madelung constant correction above can be split into two terms as

$$\left[(H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1)) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) \right]$$

$$+ [h_\sigma(\mathbf{q}_1)(H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - H^{ij}_{\mathrm{amp}}(\mathbf{0}))]. \qquad \text{(B15)}$$

For the first term in Eq. (B15), the subtraction part can be detailed as

$$H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1)$$

$$= \frac{4\pi}{|\Omega|} \frac{\hat{\varrho}_{i(\mathbf{k}_i - \mathbf{q}_1), i\mathbf{k}_i}(\mathbf{0}) \hat{\varrho}_{j(\mathbf{k}_j + \mathbf{q}_1), j\mathbf{k}_j}(\mathbf{0}) - e^{-\sigma|\mathbf{q}_1|^2}}{|\mathbf{q}_1|^2}$$

$$+ \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \backslash \{\mathbf{0}\}} \frac{\cdots}{|\mathbf{q}_1 + \mathbf{G}|^2},$$

where the numerator of the first fraction scales as $\mathcal{O}(|\mathbf{q}_1|)$ near $\mathbf{q}_1 = \mathbf{0}$. This subtraction is, thus, singular at $\mathbf{q}_1 = \mathbf{0}$ of

order -1 and shares a similar form as the ERI with partially matched orbital indices. Using the earlier inversion symmetry analysis for partially matched case, the quadrature error of the first term in Eq. (B15) when $\mathbf{k}_i \neq \mathbf{k}_a$ can be estimated as

$$\left| \mathcal{E}_{\Omega^*}((H^{ij}_{\mathrm{eri}}(\mathbf{q}_1) - h_\sigma(\mathbf{q}_1)) H^{ij}_{\mathrm{amp}}(\mathbf{q}_1), \mathcal{K}_{\mathbf{q}}) \right| \leq C \frac{1}{|\mathbf{q}|} N_{\mathbf{k}}^{-1}.$$

For the second term in Eq. (B15), we exploit the inversion symmetry of $\mathcal{K}_{\mathbf{q}}$ and Ω^* again, and its quadrature error equals to that of its symmetrized version as

$$\frac{1}{2} h_\sigma(\mathbf{q}_1)(H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - 2H^{ij}_{\mathrm{amp}}(\mathbf{0}) + H^{ij}_{\mathrm{amp}}(-\mathbf{q}_1)).$$

This formula uses $h_\sigma(\mathbf{q}_1) = h_\sigma(-\mathbf{q}_1)$ after symmetrization. Note that $h_\sigma(\mathbf{q}_1)$ is singular at $\mathbf{q}_1 = \mathbf{0}$ of order -2, while the term in the parentheses scales as $\mathcal{O}(|\mathbf{q}_1|^2)$ near $\mathbf{q}_1 = \mathbf{0}$ using the smoothness of $H^{ij}_{\mathrm{amp}}(\mathbf{q}_1)$ at $\mathbf{q}_1 = \mathbf{0}$. Therefore, the overall function above is singular at $\mathbf{q}_1 = \mathbf{0}$ of order 0. To fit the integrand form in Lemma 16, we further decompose the symmetrized integrand above into

$$[h_\sigma(\mathbf{q}_1)(H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - H^{ij}_{\mathrm{amp}}(\mathbf{0}) - \mathbf{q}_1 \cdot \nabla H^{ij}_{\mathrm{amp}}(\mathbf{0}))]$$

$$+ [h_\sigma(\mathbf{q}_1)(H^{ij}_{\mathrm{amp}}(-\mathbf{q}_1) - H^{ij}_{\mathrm{amp}}(\mathbf{0}) - (-\mathbf{q}_1) \cdot \nabla H^{ij}_{\mathrm{amp}}(\mathbf{0}))].$$

Applying Lemma 16 to these two terms separately with $\gamma = -2$ and $s = 1$ gives

$$\left| \mathcal{E}_{\Omega^*} \left(\frac{1}{2} h_\sigma(\mathbf{q}_1)(H^{ij}_{\mathrm{amp}}(\mathbf{q}_1) - 2H^{ij}_{\mathrm{amp}}(\mathbf{0}) + H^{ij}_{\mathrm{amp}}(-\mathbf{q}_1)), \mathcal{K}_{\mathbf{q}} \right) \right|$$

$$\leq C \frac{1}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1}.$$

Combining the above quadrature error estimates for the two terms in Eq. (B15), we prove a tighter error estimate at $\mathbf{q} \neq \mathbf{0}$ shown in Eq. (B13) while the previous result in Eq. (B11) at $\mathbf{q} = \mathbf{0}$ still holds.

The same error estimate can be obtained for all the six linear terms that trigger the Madelung constant correction, and the remaining linear and quadratic terms contribute at most $\mathcal{O}(N_{\mathbf{k}}^{-1})$ quadrature error. Gathering the error estimates for all these terms together and plugging into Eq. (B9), we finish the proof.

3. Proof of Lemma 13: Error in orbital energies

In the TDL, the orbital energy $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$ with any fixed n and $\mathbf{k} \in \mathcal{K}$ converges to

$$\varepsilon_{n\mathbf{k}}^{\mathrm{TDL}} = \langle n\mathbf{k}|\hat{H}_0|n\mathbf{k}\rangle + \frac{1}{|\Omega^*|} \int_{\Omega^*} d\mathbf{k}_i \sum_i (2\langle i\mathbf{k}_i, n\mathbf{k}|i\mathbf{k}_i, n\mathbf{k}\rangle$$

$$- \langle i\mathbf{k}_i, n\mathbf{k}|n\mathbf{k}, i\mathbf{k}_i\rangle).$$

Comparing this exact orbital energy with its finite-size calculation in Eq. (2), the finite-size error without any corrections can be written as

$$\varepsilon_{n\mathbf{k}}^{\text{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} = \frac{2}{|\Omega^*|} \mathcal{E}_{\Omega^*}\left(\sum_i \langle i\mathbf{k}_i, n\mathbf{k} | i\mathbf{k}_i, n\mathbf{k} \rangle, \mathcal{K} \right)$$
$$- \frac{1}{|\Omega^*|} \mathcal{E}_{\Omega^*}\left(\sum_i \langle i\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, i\mathbf{k}_i \rangle, \mathcal{K} \right). \quad (B16)$$

For the first quadrature error in Eq. (B16), i.e., the finite-size error in the direct term, the ERI with each i can be specified as

$$\langle i\mathbf{k}_i, n\mathbf{k} | i\mathbf{k}_i, n\mathbf{k} \rangle = \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \setminus \{\mathbf{0}\}} \frac{\hat{\rho}_{i\mathbf{k}_i, i\mathbf{k}_i}(\mathbf{G}) \hat{\rho}_{n\mathbf{k}, n\mathbf{k}}(-\mathbf{G})}{|\mathbf{G}|^2}.$$

Note that the momentum transfer vector of this ERI is always zero and the singular fraction term in this ERI is set to 0 by definition. As a result, the integrand is periodic and smooth with respect to $\mathbf{k}_i \in \Omega^*$. Therefore, the quadrature error of the direct term calculation, thus, decays super-algebraically according to Lemma 20 as

$$\left| \mathcal{E}_{\Omega^*}\left(\sum_i \langle i\mathbf{k}_i, n\mathbf{k} | i\mathbf{k}_i, n\mathbf{k} \rangle, \mathcal{K} \right) \right| \le C_l N_{\mathbf{k}}^{-l}, \quad \forall l > 0.$$

For the second quadrature error in Eq. (B16), i.e., the finite-size error in the exchange term, the ERI with each i can be written as

$$\langle i\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, i\mathbf{k}_i \rangle = \frac{4\pi}{|\Omega|} \frac{\hat{\rho}_{i\mathbf{k}_i, n\mathbf{k}}(\mathbf{0}) \hat{\rho}_{n\mathbf{k}, i\mathbf{k}_i}(\mathbf{0})}{|\mathbf{q}|^2}$$
$$+ \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \setminus \{\mathbf{0}\}} \frac{\hat{\rho}_{i\mathbf{k}_i, i\mathbf{k}_i}(\mathbf{G}) \hat{\rho}_{n\mathbf{k}, n\mathbf{k}}(-\mathbf{G})}{|\mathbf{q} + \mathbf{G}|^2},$$

which is singular at $\mathbf{q} := \mathbf{k} - \mathbf{k}_i = \mathbf{0}$ of order -2 when $n = i$ and 0 otherwise. An error estimate lemma in Ref. [41] (restated as Lemma 22 in Appendix D) gives a tight quadrature error estimate for such periodic functions with one point of algebraic singularity, and its application to the above integrand gives

$$|\mathcal{E}_{\Omega^*}(\langle i\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, i\mathbf{k}_i \rangle, \mathcal{K})| \le C \begin{cases} N_{\mathbf{k}}^{-1/3} & n = i, \\ N_{\mathbf{k}}^{-1} & n \ne i. \end{cases}$$

Combining the estimates of the two error terms in Eq. (B16), we obtain the overall finite-size error estimate for orbital energies without any corrections as

$$\left| \varepsilon_{n\mathbf{k}}^{\text{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}} \right| \le C \begin{cases} N_{\mathbf{k}}^{-1/3} & n \text{ is occupied}, \\ N_{\mathbf{k}}^{-1} & n \text{ is virtual}. \end{cases}$$

From the above analysis, the dominant finite-size error in an occupied orbital energy lies in the calculation of the exchange term with $i = n$. In the Madelung-corrected orbital energy $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}, \xi}$, the correction is applied to this exchange term as (ignoring the prefactor $-|\Omega^*|^{-1}$)

$$\mathcal{Q}_{\Omega^*}(\langle n\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, n\mathbf{k}_i \rangle, \mathcal{K}) \to \mathcal{Q}_{\Omega^*}(\langle n\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, n\mathbf{k}_i \rangle, \mathcal{K})$$
$$- |\Omega^*|\xi.$$

Applying the change of variable $\mathbf{k}_i \to \mathbf{k} - \mathbf{q}$ and using the periodicity of the ERI with respect to $\mathbf{k}_i$, this corrected calculation can be reformulated as

$$\mathcal{Q}_{\Omega^*}(\langle n\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, n\mathbf{k}_i \rangle, \mathcal{K}) - |\Omega^*|\xi$$
$$= \mathcal{Q}_{\Omega^*}(\langle n(\mathbf{k} - \mathbf{q}), n\mathbf{k} | n\mathbf{k}, n(\mathbf{k} - \mathbf{q}) \rangle - h_\sigma(\mathbf{q}), \mathcal{K}_{\mathbf{q}})$$
$$+ \int_{\Omega^*} d\mathbf{q}_1 h_\sigma(\mathbf{q}_1) + \mathcal{O}(N_{\mathbf{k}}^{-1}),$$

using the expansion of the Madelung constant with any fixed $\sigma > 0$ in Eq. (2). The quadrature error after the correction can be written as

$$\mathcal{E}_{\Omega^*}(\langle n\mathbf{k}_i, n\mathbf{k} | n\mathbf{k}, n\mathbf{k}_i \rangle, \mathcal{K}) + |\Omega^*|\xi$$
$$= \mathcal{E}_{\Omega^*}(\langle n(\mathbf{k} - \mathbf{q}), n\mathbf{k} | n\mathbf{k}, n(\mathbf{k} - \mathbf{q}) \rangle - h_\sigma(\mathbf{q}), \mathcal{K}_{\mathbf{q}})$$
$$+ \mathcal{O}(N_{\mathbf{k}}^{-1}).$$

The effective integrand above after the correction can be detailed as

$$\langle n(\mathbf{k} - \mathbf{q}), n\mathbf{k} | n\mathbf{k}, n(\mathbf{k} - \mathbf{q}) \rangle - h_\sigma(\mathbf{q})$$
$$= \frac{4\pi}{|\Omega|} \frac{|\hat{\varrho}_{n(\mathbf{k} - \mathbf{q}), n\mathbf{k}}(\mathbf{0})|^2 - e^{-\sigma|\mathbf{q}|^2}}{|\mathbf{q}|^2} + \frac{4\pi}{|\Omega|} \sum_{\mathbf{G} \in \mathbb{L}^* \setminus \{\mathbf{0}\}} \frac{\cdots}{|\mathbf{q} + \mathbf{G}|^2}.$$

The integrand singularity comes from the first fractional term and is of order -1, since the ERI and $h_\sigma(\mathbf{q})$ share the same leading singular term. Similar to the analysis in the ERI contraction, we can then combine this singularity subtraction with the inverse symmetry of Ω^* and $\mathcal{K}_{\mathbf{q}}$ to show that

$$|\mathcal{E}_{\Omega^*}(\langle n(\mathbf{k} - \mathbf{q}), n\mathbf{k} | n\mathbf{k}, n(\mathbf{k} - \mathbf{q}) \rangle - h_\sigma(\mathbf{q}), \mathcal{K}_{\mathbf{q}})| \le C N_{\mathbf{k}}^{-1}.$$

Combining this estimate with the above estimate for all the remaining direct and exchange terms in $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}, \xi}$, we obtain the finite-size error estimate for orbital energies with the Madelung constant correction as

$$\left| \varepsilon_{n\mathbf{k}}^{\text{TDL}} - \varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}, \xi} \right| \le C N_{\mathbf{k}}^{-1}.$$

4. Proof of Lemma 14: Error from previous iteration

Fixing a set of entry index (i, j, a, b) and $(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \in \mathcal{K} \times \mathcal{K} \times \mathcal{K}$, the ERI-contraction entry $[\mathcal{A}_{N_{\mathbf{k}},\xi}(T)]_{ijab,\mathbf{k}_i\mathbf{k}_j\mathbf{k}_a}$ can be detailed as

$$[\mathcal{A}_{N_{\mathbf{k}},\xi}(T)]_{ijab}(\mathbf{k}_i\mathbf{k}_j\mathbf{k}_a) = \langle a\mathbf{k}_a, b\mathbf{k}_b|i\mathbf{k}_i, j\mathbf{k}_j\rangle + \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l|i\mathbf{k}_i, j\mathbf{k}_j\rangle T_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)$$

$$+ \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}_k\mathbf{k}_c \in \mathcal{K}} \sum_{klcd} \langle k\mathbf{k}_k, l\mathbf{k}_l|c\mathbf{k}_c, d\mathbf{k}_d\rangle T_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c) T_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)$$

$$+ \cdots + 2\xi T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a),$$

where the neglected terms are all the other linear and quadratic terms and the Madelung constant corrections to different terms are collected together at the end.

In the subtraction $\mathcal{A}_{N_{\mathbf{k}},\xi}(T) - \mathcal{A}_{N_{\mathbf{k}},\xi}(S)$, the constant terms cancel each other. The subtraction between the two Madelung constant correction terms can be estimated directly as

$$|2\xi[T - S]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq CN_{\mathbf{k}}^{-1/3} \begin{cases} \frac{1}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1} & \mathbf{q} \neq \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}, \end{cases}$$

using the fact that $\xi = \mathcal{O}(N_{\mathbf{k}}^{-1/3})$.

The subtraction between the two $4h2p$ linear terms can be formulated and bounded as

$$\left| \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l|i\mathbf{k}_i, j\mathbf{k}_j\rangle[T - S]_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a) \right|$$

$$\leq \left(\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}\backslash\{\mathbf{k}_i,\mathbf{k}_a\}} + \frac{1}{N_{\mathbf{k}}}\delta_{\mathbf{k}_k,\mathbf{k}_i} + \frac{1}{N_{\mathbf{k}}}\delta_{\mathbf{k}_k,\mathbf{k}_a} \right) \left| \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l|i\mathbf{k}_i, j\mathbf{k}_j\rangle[T - S]_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a) \right|. \quad \text{(B17)}$$

For the term with $\mathbf{k}_k = \mathbf{k}_i$ (i.e., $\delta_{\mathbf{k}_k,\mathbf{k}_i}$), we have

$$\frac{1}{N_{\mathbf{k}}} \left| \sum_{kl} \langle k\mathbf{k}_i, l\mathbf{k}_j|i\mathbf{k}_i, j\mathbf{k}_j\rangle[T - S]_{klab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a) \right| \leq \frac{1}{N_{\mathbf{k}}} \sum_{kl} CN_{\mathbf{k}}^{-1/3} \leq CN_{\mathbf{k}}^{-4/3},$$

where the ERI definition at zero momentum transfer skips the singular fraction and is $\mathcal{O}(1)$, and according to the assumption on $T - S$ it always holds that

$$\max_{ijab,\mathbf{k}_i,\mathbf{k}_j,\mathbf{k}_a \in \mathcal{K}} |[T - S]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)| \leq CN_{\mathbf{k}}^{-1/3}. \quad \text{(B18)}$$

For the term with $\mathbf{k}_k = \mathbf{k}_a$ (i.e., $\delta_{\mathbf{k}_k,\mathbf{k}_a}$), its estimate is the same as the term above when $\mathbf{k}_a = \mathbf{k}_i$. When $\mathbf{k}_a \neq \mathbf{k}_i \in \mathcal{K}$, we have $|\mathbf{q}| \geq CN_{\mathbf{k}}^{-1/3}$ and

$$\frac{1}{N_{\mathbf{k}}} \left| \sum_{kl} \langle k\mathbf{k}_a, l\mathbf{k}_b|i\mathbf{k}_i, j\mathbf{k}_j\rangle[T - S]_{klab}(\mathbf{k}_a, \mathbf{k}_b, \mathbf{k}_a) \right| \leq \frac{1}{N_{\mathbf{k}}} \sum_{kl} \frac{C}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1/3} \leq C\frac{1}{|\mathbf{q}|} N_{\mathbf{k}}^{-1}.$$

For the first summation term in Eq. (B17), we introduce the change of variable $\mathbf{k}_k \to \mathbf{k}_i - \mathbf{q}_1$ and write it as

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}}\backslash\{\mathbf{0},-\mathbf{q}\}} \left| \sum_{kl} \langle k(\mathbf{k}_i - \mathbf{q}_1), l(\mathbf{k}_j + \mathbf{q}_1)|i\mathbf{k}_i, j\mathbf{k}_j\rangle[T - S]_{klab}(\mathbf{k}_i - \mathbf{q}_1, \mathbf{k}_j + \mathbf{q}_1, \mathbf{k}_a) \right|.$$

When $\mathbf{q} = \mathbf{0}$, this term can be further bounded using Eq. (B18) by

$$(*) \le \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}} \backslash \{\mathbf{0}\}} \frac{C}{|\mathbf{q}_1|^2} N_{\mathbf{k}}^{-1/3} \le C \int_{\Omega^*} d\mathbf{q}_1 \frac{1}{|\mathbf{q}_1|^2} N_{\mathbf{k}}^{-1/3} \le C N_{\mathbf{k}}^{-1/3}.$$

When $\mathbf{q} \ne \mathbf{0}$, this term can be further bounded as

$$(*) \le C \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}} \backslash \{\mathbf{0}, -\mathbf{q}\}} \frac{1}{|\mathbf{q}_1|^2} \sum_{\mathbf{G} \in \mathbb{L}_0^*} \frac{1}{|\mathbf{q}_1 + \mathbf{q} + \mathbf{G}|^2} N_{\mathbf{k}}^{-1}$$

$$\le C N_{\mathbf{k}}^{-1} \int_{\Omega^*} d\mathbf{q}_1 \sum_{\mathbf{G} \in \mathbb{L}_0^*} \frac{1}{|\mathbf{q}_1|^2} \frac{1}{|\mathbf{q}_1 + \mathbf{q} + \mathbf{G}|^2}$$

$$\le C N_{\mathbf{k}}^{-1} \sum_{\mathbf{G} \in \mathbb{L}_0^*} \frac{1}{|\mathbf{q} + \mathbf{G}|^2} \le C \frac{1}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1},$$

where $\mathbb{L}_0^*$ denotes the set of 27 lattice vectors in $\mathbb{L}^*$ around the origin. The first inequality is based on the lemma assumption on $T - S$ that

$$\left| [T - S]_{klab}(\mathbf{k}_i - \mathbf{q}_1, \mathbf{k}_j + \mathbf{q}_1, \mathbf{k}_a) \right| \le C \frac{1}{|\mathbf{q}_1 + \mathbf{q} + \mathbf{G}_0|^2} N_{\mathbf{k}}^{-1} \le C \sum_{\mathbf{G} \in \mathbb{L}_0^*} \frac{1}{|\mathbf{q}_1 + \mathbf{q} + \mathbf{G}|^2} N_{\mathbf{k}}^{-1},$$

where $\mathbf{G}_0 \in \mathbb{L}_0^*$ is the unique lattice vector that makes $\mathbf{q}_1 + \mathbf{q} + \mathbf{G}_0 \in \Omega^*$. The third inequality can be obtained from the nonsmoothness characterization of function

$$f(\mathbf{z}) = \int_{\Omega^*} d\mathbf{q}_1 \frac{1}{|\mathbf{q}_1|^2} \frac{1}{|\mathbf{q}_1 + \mathbf{z}|^2}.$$

Using Lemma 11 in Ref. [41], $f(\mathbf{z})$ is singular only at $\mathbf{z} = \mathbf{0}$ of order -2, and its value at $\mathbf{z} \in \Omega^* \backslash \{\mathbf{0}\}$ is bounded by $C|\mathbf{z}|^{-2}$.

Based on the above estimates of the first term in Eq. (B17), we obtain the estimate of the error accumulation in the $4h2p$ linear term calculation as

$$\left| \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kl} \langle k\mathbf{k}_k, l\mathbf{k}_l | i\mathbf{k}_i, j\mathbf{k}_j \rangle [T - S]_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a) \right| \le C \begin{cases} \frac{1}{|\mathbf{q}|^2} N_{\mathbf{k}}^{-1} & \mathbf{q} \ne \mathbf{0}, \\ N_{\mathbf{k}}^{-1/3} & \mathbf{q} = \mathbf{0}. \end{cases}$$

The same analysis can be applied to all the similar linear terms that contain ERIs with matched orbital indices. For other linear terms, the analysis can be done similarly, and they all contribute at most $\mathcal{O}(N_{\mathbf{k}}^{-1})$ error to the overall subtraction. Taking the subtraction between the following $3h3p$ linear terms as an example, it can be formulated and bounded as

$$\left| \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K}} \sum_{kc} \langle a\mathbf{k}_a, k\mathbf{k}_k | i\mathbf{k}_i, c\mathbf{k}_c \rangle [T - S]_{kjbc}(\mathbf{k}_k, \mathbf{k}_j, \mathbf{k}_b) \right|$$

$$\le C \left(\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}_k \in \mathcal{K} \backslash \{\mathbf{k}_b\}} \sum_{kc} + \frac{1}{N_{\mathbf{k}}} \delta_{\mathbf{k}_k, \mathbf{k}_b} \right) |[T - S]_{kjbc}(\mathbf{k}_k, \mathbf{k}_j, \mathbf{k}_b)|$$

$$\le C \left(\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{q}_1 \in \mathcal{K}_{\mathbf{q}} \backslash \{\mathbf{0}\}} \frac{1}{|\mathbf{q}_1|^2} N_{\mathbf{k}}^{-1} + \frac{1}{N_{\mathbf{k}}} N_{\mathbf{k}}^{-1/3} \right) \le C N_{\mathbf{k}}^{-1}.$$

For the subtraction between quadratic terms, we consider the $4h2p$ quadratic term as an example. The subtraction between the two $4h2p$ quadratic terms can be formulated and bounded as

$$\left| \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}} \sum_{\mathbf{k}_k \mathbf{k}_c \in \mathcal{K}} \sum_{klcd} \langle k\mathbf{k}_k, l\mathbf{k}_l | c\mathbf{k}_c, d\mathbf{k}_d \rangle (T_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c) T_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a) - S_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c) S_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)) \right|$$

$$\leq C \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}} \sum_{\mathbf{k}_k \mathbf{k}_c \in \mathcal{K}} \sum_{klcd} |T_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c)[T - S]_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a) + [T - S]_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c) S_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)|$$

$$\leq C \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}} \sum_{\mathbf{k}_k \mathbf{k}_c \in \mathcal{K}} \sum_{klcd} |[T - S]_{klab}(\mathbf{k}_k, \mathbf{k}_l, \mathbf{k}_a)| + C \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}} \sum_{\mathbf{k}_k \mathbf{k}_c \in \mathcal{K}} \sum_{klcd} |[T - S]_{ijcd}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_c)|$$

$$\leq C N_{\mathbf{k}}^{-1}.$$

Similar analysis can be done to all the remaining quadratic terms, and they all contribute at most $\mathcal{O}(N_{\mathbf{k}}^{-1})$ error to the overall subtraction. Gathering all the estimates above together, we finish the proof.

APPENDIX C: PROOF OF THEOREM 2

To guarantee the convergence and control the regularity of CCD(n) calculations with $n \to \infty$, we introduce additional technical assumptions similar to those in Ref. [41]. One key difference is that Ref. [41] measures the error in the amplitude calculation using the maximum entrywise norm. When Madelung constant correction is used, the error of the amplitude should be measured by the average entrywise norm (related to the L^1 norm) instead, which is denoted by

$$\|T\|_1 = \frac{1}{N_{\mathbf{k}}^3} \sum_{\mathbf{k}} \sum_{\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a \in \mathcal{K}} \sum_{ijab} |T_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)|,$$

$$\|t\|_1 = \frac{1}{|\Omega^*|^3} \int_{\Omega^* \times \Omega^* \times \Omega^*} d\mathbf{k}_i d\mathbf{k}_j d\mathbf{k}_a \sum_{ijab} |t_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)|,$$

where T and t denote a generic amplitude computed using mesh $\mathcal{K}$ and in the TDL, respectively.

Assume that the CCD amplitude equations using a sufficiently fine mesh $\mathcal{K}$ (with or without the Madelung constant correction) and in the TDL have unique solutions and denote the solutions as $T_*^{N_{\mathbf{k}}}$ and t_*, respectively. Convergence of the CCD(n) amplitudes to the CCD amplitude is defined in the $\|\cdot\|_1$-norm sense as

$$\lim_{n \to \infty} \|T_n^{N_{\mathbf{k}}} - T_*^{N_{\mathbf{k}}}\|_1 = 0 \quad \text{and} \quad \lim_{n \to \infty} \|t_n - t_*\|_1 = 0.$$

We impose a sufficient condition that guarantees the convergence of fixed point iterations by requiring the target mapping to be contractive in a domain that contains both the solution point and the initial guess.

Assumption 18. For $N_{\mathbf{k}}$ sufficiently large, the following statements hold.

(1) The exact CCD(n) amplitude t_n converges to the CCD amplitude t_* pointwisely as $n \to \infty$, i.e.,

$$\lim_{n \to \infty} [t_n]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a)$$
$$= [t_*]_{ijab}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a), \quad \forall i, j, a, b, \mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_a. \quad \text{(C1)}$$

(2) $(\varepsilon^{N_{\mathbf{k}}})^{-1} \mathcal{A}_{N_{\mathbf{k}}}$ is a contraction map in some domain $\mathbb{B}_{N_{\mathbf{k}}} \subset \mathbb{C}^{n_{\mathrm{occ}} \times n_{\mathrm{occ}} \times n_{\mathrm{vir}} \times n_{\mathrm{vir}} \times N_{\mathbf{k}} \times N_{\mathbf{k}} \times N_{\mathbf{k}}}$ that contains $T_*^{N_{\mathbf{k}}}$ and the initial guess 0, i.e.,

$$(\varepsilon^{N_{\mathbf{k}}})^{-1} \mathcal{A}_{N_{\mathbf{k}}}(T) \in \mathbb{B}_{N_{\mathbf{k}}}, \quad \forall T \in \mathbb{B}_{N_{\mathbf{k}}},$$
$$\|(\varepsilon^{N_{\mathbf{k}}})^{-1}(\mathcal{A}_{N_{\mathbf{k}}}(T) - \mathcal{A}_{N_{\mathbf{k}}}(S))\|_1$$
$$\leq L \|T - S\|_1, \quad \forall T, S \in \mathbb{B}_{N_{\mathbf{k}}}, \quad \text{(C2)}$$

with a Lipschitz constant $L < 1$.

(3) The exact CCD(n) amplitude t_n and the domain $\mathbb{B}_{N_{\mathbf{k}}}$ above satisfy that

$$\mathcal{M}_{\mathcal{K}} t_n \in \mathbb{B}_{N_{\mathbf{k}}}, \quad \forall n > 0. \quad \text{(C3)}$$

(4) For all the amplitudes $\{t_n\}$, there exists a constant C such that

$$\|\mathcal{M}_{\mathcal{K}} \mathcal{A}_{\mathrm{TDL}}(t_n) - \mathcal{A}_{N_{\mathbf{k}}, \xi}(\mathcal{M}_{\mathcal{K}} t_n)\|_1$$
$$\leq C N_{\mathbf{k}}^{-1}, \quad \forall n > 0,$$
$$\|\mathcal{M}_{\mathcal{K}} \mathcal{A}_{\mathrm{TDL}}(t_n) - \mathcal{A}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}} t_n)\|_1$$
$$\leq C N_{\mathbf{k}}^{-1/3}, \quad \forall n > 0. \quad \text{(C4)}$$

Note that, when we consider the finite-size calculation with Madelung constant correction, the components $\varepsilon^{N_{\mathbf{k}}}$ or $\mathcal{A}_{N_{\mathbf{k}}}$ in the second assumption need to be changed to $\varepsilon^{N_{\mathbf{k}}, \xi}$ or $\mathcal{A}_{N_{\mathbf{k}}, \xi}$ accordingly.

Remark 19. The second assumption guarantees that $\{T_n^{N_{\mathbf{k}}}\}$ lies in $\mathbb{B}_{N_{\mathbf{k}}}$ and converges to $T_*^{N_{\mathbf{k}}}$. For the third

assumption, Theorem 1 shows that with each fixed $n > 0$ the amplitude $T_n^{N_{\mathbf{k}}}$ converges to t_n in the sense of

$$\lim_{N_{\mathbf{k}} \to \infty} \|\mathcal{M}_{\mathcal{K}} t_n - T_n^{N_{\mathbf{k}}}\|_1 = 0,$$

suggesting that $\{T_n^{N_{\mathbf{k}}}\} \subset \mathbb{B}_{N_{\mathbf{k}}}$ converges to t_n with $\mathcal{K} \to \Omega^*$. Therefore, $\mathbb{B}_{N_{\mathbf{k}}}$ and $\mathcal{M}_{\mathcal{K}} t_n$ should be related to each other, which leads to the third assumption.

For the last assumption, we note that the two related error estimates in Lemma 11 have the prefactor C dependent on the amplitude t_n. The assumption here is stronger in the sense that C needs to be independent of t_n.

Rigorous statement of Theorem 2. Under Assumption 18, the finite-size error of the CCD correlation energy scales as $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ in each of the following scenarios: (i) the Madelung constant correction is applied only to the ERI contraction $\mathcal{A}_{N_{\mathbf{k}}}$ and (ii) the Madelung constant correction is applied only to the orbital energy $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$. When the Madelung constant correction is applied to both $\mathcal{A}_{N_{\mathbf{k}}}$

and $\varepsilon_{n\mathbf{k}}^{N_{\mathbf{k}}}$ in the CCD calculation, the overall finite-size error scales as $\mathcal{O}(N_{\mathbf{k}}^{-1})$.

Proof. The finite-size error in the CCD energy calculation with or without the Madelung constant correction can be estimated as

$$\begin{aligned}
\left| E_{\mathrm{CCD}}^{\mathrm{TDL}} - E_{\mathrm{CCD}}^{N_{\mathbf{k}}} \right| &= |\mathcal{G}_{\mathrm{TDL}}(t_*) - \mathcal{G}_{N_{\mathbf{k}}}(T_*^{N_{\mathbf{k}}})| \\
&\leq |\mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}} t_*) - \mathcal{G}_{N_{\mathbf{k}}}(T_*^{N_{\mathbf{k}}})| \\
&\quad + |\mathcal{G}_{\mathrm{TDL}}(t_*) - \mathcal{G}_{N_{\mathbf{k}}}(\mathcal{M}_{\mathcal{K}} t_*)| \\
&\leq C\|\mathcal{M}_{\mathcal{K}} t_* - T_*^{N_{\mathbf{k}}}\|_1 + C N_{\mathbf{k}}^{-1}, \quad (\mathrm{C5})
\end{aligned}$$

where the last inequality uses the boundedness of $|W_{ijab}|$ in $\mathcal{G}_{N_{\mathbf{k}}}$ and Lemma 10. To first estimate the above amplitude error when the Madelung constant correction is applied to both orbital energy and ERI contractions, we consider the error splitting Eq. (B2) for the amplitude calculation at the nth fixed point iteration as

$$\begin{aligned}
\|\mathcal{M}_{\mathcal{K}} t_n - T_n^{N_{\mathbf{k}}}\|_1 &\leq \left\| \frac{1}{\varepsilon^{\mathrm{TDL}}} [\mathcal{M}_{\mathcal{K}} \mathcal{A}_{\mathrm{TDL}}(t_{n-1}) - \mathcal{A}_{N_{\mathbf{k}},\xi}(\mathcal{M}_{\mathcal{K}} t_{n-1})] \right\|_1 \\
&\quad + \left\| \left[\frac{1}{\varepsilon^{\mathrm{TDL}}} - \frac{1}{\varepsilon^{N_{\mathbf{k}},\xi}} \right] \mathcal{A}_{N_{\mathbf{k}},\xi}(\mathcal{M}_{\mathcal{K}} t_{n-1}) \right\|_1 + \left\| \frac{1}{\varepsilon^{N_{\mathbf{k}},\xi}} [\mathcal{A}_{N_{\mathbf{k}},\xi}(\mathcal{M}_{\mathcal{K}} t_{n-1}) - \mathcal{A}_{N_{\mathbf{k}},\xi}(T_{n-1}^{N_{\mathbf{k}}})] \right\|_1 \\
&\leq C N_{\mathbf{k}}^{-1} + L \left\| \mathcal{M}_{\mathcal{K}} t_{n-1} - T_{n-1}^{N_{\mathbf{k}}} \right\|_1,
\end{aligned}$$

where the last estimate uses the assumption in Eq. (C4) for the first term and the assumptions of contraction maps in Eq. (C2) and $\mathcal{M}_{\mathcal{K}} t_{n-1} \in \mathbb{B}_{N_{\mathbf{k}}}$ in Eq. (C3) for the third term. Since the initial guesses in the finite and the TDL cases satisfy $\|\mathcal{M}_{\mathcal{K}} t_0 - T_0^{N_{\mathbf{k}}}\|_1 = 0$, we can recursively derive that

$$\|\mathcal{M}_{\mathcal{K}} t_n - T_n^{N_{\mathbf{k}}}\|_1 \leq C \frac{1 - L^n}{1 - L} N_{\mathbf{k}}^{-1},$$

and, thus, the first assumption Eq. (C1) gives

$$\|\mathcal{M}_{\mathcal{K}} t_* - T_*^{N_{\mathbf{k}}}\|_1 = \lim_{n \to \infty} \|\mathcal{M}_{\mathcal{K}} t_n - T_n^{N_{\mathbf{k}}}\|_1 \leq C N_{\mathbf{k}}^{-1}.$$

Plugging this estimate into Eq. (C5) then finishes the proof for the scenario when the Madelung constant correction is applied to both orbital energies and ERI contractions.

For the two scenarios with partial Madelung constant correction, a similar analysis as above gives

$$\|\mathcal{M}_{\mathcal{K}} t_n - T_n^{N_{\mathbf{k}}}\|_1 \leq C N_{\mathbf{k}}^{-1/3} + L \left\| \mathcal{M}_{\mathcal{K}} t_{n-1} - T_{n-1}^{N_{\mathbf{k}}} \right\|_1,$$

where the dominant $C N_{\mathbf{k}}^{-1/3}$ error comes from uncorrected orbital energy or uncorrected ERI-contraction term. Recursively, we can obtain the amplitude error estimate as

$$\|\mathcal{M}_{\mathcal{K}} t_* - T_*^{N_{\mathbf{k}}}\|_1 \leq C N_{\mathbf{k}}^{-1/3},$$

which finishes the proof for the two scenarios with partial correction. ∎

APPENDIX D: QUADRATURE ERROR ESTIMATE FOR PERIODIC FUNCTIONS WITH ALGEBRAIC SINGULARITY

This section presents a collection of lemmas that provide quadrature error estimates for trapezoidal rules over periodic functions with specific algebraic singularities, which are used in this paper. Most of the lemmas are proven in Ref. [41] and are restated here for completeness. In addition, we introduce and prove a new quadrature error estimate that is critical in describing the efficacy of the Madelung constant correction and inverse symmetry in reducing the quadrature error in orbital energies and ERI contractions.

The lemmas presented in this section provide the asymptotic scaling of the quadrature errors and also a quantitative relationship between the prefactors in the estimate and the algebraic singularities of the integrand. In addition to the singularity order, we further quantitatively characterize the algebraic singularity as follows. For a univariate function $f(\mathbf{x})$ that is smooth everywhere in V except at $\mathbf{x} = \mathbf{x}_0$ with algebraic singularity of order γ, we define a constant

$$\mathcal{H}^l_{V,\mathbf{x}_0}(f) = \min\left\{ C : |\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} f(\mathbf{x})| \leq C|\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|}, \ \forall \ |\boldsymbol{\alpha}| \leq l, \ \forall \ \mathbf{x} \in V \backslash \{\mathbf{x}_0\} \right\}$$
$$= \max_{|\boldsymbol{\alpha}| \leq l} \left\| (\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} f(\mathbf{x})) / |\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|} \right\|_{L^\infty(V)}.$$

For a multivariate function $f(\mathbf{x}, \mathbf{y})$ that is smooth everywhere in $V_X \times V_Y$ except at $\mathbf{x} = \mathbf{x}_0$ with algebraic singularity of order γ, we define a constant

$$\mathcal{H}^l_{V_X \times V_Y,(\mathbf{x}_0,\cdot)}(f) = \min\left\{ C : |\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} \partial^{\boldsymbol{\beta}}_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})| \leq C|\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|}, \ \forall |\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \leq l, \ \forall \mathbf{x} \in V_X \backslash \{\mathbf{x}_0\}, \mathbf{y} \in V_Y \right\}$$
$$= \max_{|\boldsymbol{\alpha}| \leq l} \left\| (\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} \partial^{\boldsymbol{\beta}}_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})) / |\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|} \right\|_{L^\infty(V \times V)},$$

where "$\cdot$" in the subscript "$(\mathbf{x}_0, \cdot)$" is a placeholder to indicate the smooth variable. Using these two quantities, we have the following function estimates that are extensively used in this section:

$$|\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} f(\mathbf{x})| \leq \mathcal{H}^l_{V,\mathbf{x}_0}(f)|\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|}, \quad \forall l \geq |\boldsymbol{\alpha}|, \quad \forall \mathbf{x} \in V \backslash \{\mathbf{x}_0\},$$
$$|\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} \partial^{\boldsymbol{\beta}}_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})| \leq \mathcal{H}^l_{V_X \times V_Y,(\mathbf{x}_0,\cdot)}(f)|\mathbf{x} - \mathbf{x}_0|^{\gamma - |\boldsymbol{\alpha}|}, \quad \forall l \geq |\boldsymbol{\alpha}|, |\boldsymbol{\beta}|, \quad \forall \mathbf{x} \in V_X \backslash \{\mathbf{x}_0\}, \mathbf{y} \in V_Y.$$

1. Existing results from Ref. [41]

Lemma 20. Let $f(\mathbf{x})$ be smooth and periodic in $V = [-\frac{1}{2}, \frac{1}{2}]^d$. The quadrature error of a trapezoidal rule using an m^d-sized uniform mesh $\mathcal{X}$ in V decays super-algebraically as

$$|\mathcal{E}_V(f, \mathcal{X})| \leq C_l m^{-l}, \quad \forall \ l > 0.$$

Remark 21. If we replace $f(\mathbf{x})$ by $f(\mathbf{x}, \mathbf{y})$ defined in $V \times V_Y$ which is smooth and periodic with respect to $\mathbf{x}$ for each $\mathbf{y} \in V_Y$ and satisfies $\sup_{\mathbf{x} \in V, \mathbf{y} \in V_Y} |\partial^{\boldsymbol{\alpha}}_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})| < \infty$ for any $\boldsymbol{\alpha} \geq 0$, Lemma 20 can be generalized as

$$|\mathcal{E}_V(f(\cdot, \mathbf{y}), \mathcal{X})| \leq C_l m^{-l}, \quad \forall \ l > 0, \quad \forall \ \mathbf{y} \in V_Y,$$

where constant C_l is independent of $\mathbf{y} \in V_Y$.

Lemma 22. Let $f(\mathbf{x})$ be periodic with respect to $V = [-\frac{1}{2}, \frac{1}{2}]^d$ and smooth everywhere except at $\mathbf{x} = \mathbf{0}$ of order $\gamma \geq -d + 1$. At $\mathbf{x} = \mathbf{0}$, $f(\mathbf{x})$ is set to 0. The quadrature error of a trapezoidal rule using an m^d-sized uniform mesh $\mathcal{X}$ that contains $\mathbf{x} = \mathbf{0}$ can be estimated as

$$|\mathcal{E}_V(f, \mathcal{X})| \leq C\mathcal{H}^{d + \max(1, \gamma)}_{V,\mathbf{0}}(f)m^{-(d+\gamma)}.$$

If $f(\mathbf{0})$ is set to an $\mathcal{O}(1)$ value in the calculation, it introduces an additional $\mathcal{O}(m^{-d})$ quadrature error.

Remark 23. If we replace $f(\mathbf{x})$ by $f(\mathbf{x}, \mathbf{y})$ defined in $V \times V_Y$ which is smooth everywhere in $V \times V_Y$ except at $\mathbf{x} = \mathbf{0}$ of order γ, Lemma 22 can be generalized to

$$|\mathcal{E}_V(f(\cdot, \mathbf{y}), \mathcal{X})| \leq C\mathcal{H}^{d + \max(1, \gamma)}_{V \times V_Y,(\mathbf{0},\cdot)}(f)m^{-(d+\gamma)}, \quad \forall \ \mathbf{y} \in V_Y,$$

where the prefactor applies uniformly across all $\mathbf{y} \in V_Y$.

Lemma 24. Let $f(\mathbf{x}) = f_1(\mathbf{x})f_2(\mathbf{x})$, where $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are periodic with respect to $V = [-\frac{1}{2}, \frac{1}{2}]^d$ and
  (i) $f_1(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_1 = \mathbf{0}$ of order $\gamma \leq 0$ and
  (ii) $f_2(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_2 \neq \mathbf{0}$ of order 0.

Consider an m^d-sized uniform mesh $\mathcal{X}$ in V. Assume that $\mathcal{X}$ satisfies that $\mathbf{z}_1$, $\mathbf{z}_2$ are either on the mesh or $\Theta(m^{-1})$ away from any mesh points and m is sufficiently large that $|\mathbf{z}_1 - \mathbf{z}_2| = \Omega(m^{-1})$. At $\mathbf{x} = \mathbf{z}_1$ and $\mathbf{x} = \mathbf{z}_2$, $f(\mathbf{x})$ is set to 0. The trapezoidal rule using $\mathcal{X}$ has quadrature error

$$|\mathcal{E}_V(f, \mathcal{X})| \leq C\mathcal{H}^{d+1}_{V,\mathbf{z}_1}(f_1)\mathcal{H}^{d+1}_{V,\mathbf{z}_2}(f_2)m^{-(d+\gamma)}.$$

If $f(\mathbf{z}_1)$ and $f(\mathbf{z}_2)$ are set to arbitrary $\mathcal{O}(1)$ values, it introduces an additional $\mathcal{O}(m^{-d})$ quadrature error.

Remark 25. If we replace $f_i(\mathbf{x})$ with $i = 1, 2$ by $f_i(\mathbf{x}, \mathbf{y})$ defined in $V \times V_Y$ which is smooth everywhere in $V \times V_Y$ except at $\mathbf{x} = \mathbf{z}_i$ of order γ and 0, respectively, Lemma 24 can be generalized to

$$|\mathcal{E}_V(f_1(\cdot, \mathbf{y}) f_2(\cdot, \mathbf{y}), \mathcal{X})|$$
$$\leq C \mathcal{H}^{d+1}_{V \times V_Y, (\mathbf{0}, \cdot)}(f_1) \mathcal{H}^{d+1}_{V \times V_Y, (\mathbf{0}, \cdot)}(f_2) m^{-(d+\gamma)}, \quad \forall \mathbf{y} \in V_Y,$$

where the prefactor applies uniformly across all $\mathbf{y} \in V_Y$.

2. A new quadrature error estimate

Here, we prove Lemma 16, which is used in the quadrature error estimate of the ERI contractions with the Madelung constant correction in Appendix B 2. This lemma is a generalization of the existing result in Lemma 24 with an additional condition.

Restatement of Lemma 16. Let $f(\mathbf{x}) = f_1(\mathbf{x}) f_2(\mathbf{x})$, where $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are periodic with respect to $V = [-\frac{1}{2}, \frac{1}{2}]^d$ and

(i) $f_1(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_1 = \mathbf{0}$ of order $\gamma \leq -1$,

(ii) $f_2(\mathbf{x})$ is smooth everywhere except at $\mathbf{x} = \mathbf{z}_2 \neq \mathbf{0}$ of order 0, and

(iii) $\partial_\mathbf{x}^\alpha f_2(\mathbf{0}) = 0$ for any derivative order $|\alpha| \leq s$.

Assume $\gamma > -d$ for $f(\mathbf{x})$ to be integrable in V and $\gamma + s + 1 \leq 0$ so the leading algebraic singularity of $f(\mathbf{x})$ is at $\mathbf{x} = \mathbf{0}$. Consider an m^d-sized uniform mesh $\mathcal{X}$ in V. Assume that $\mathcal{X}$ satisfies that $\mathbf{z}_1$, $\mathbf{z}_2$ are either on the mesh or $\Theta(m^{-1})$ away from any mesh points and m is sufficiently large that $|\mathbf{z}_1 - \mathbf{z}_2| = \Omega(m^{-1})$. At $\mathbf{x} = \mathbf{z}_1$ and $\mathbf{x} = \mathbf{z}_2$, $f(\mathbf{x})$ is set to 0. The trapezoidal rule using $\mathcal{X}$ has quadrature error

$$|\mathcal{E}_V(f, \mathcal{X})| \leq C \mathcal{H}^{d+1}_{V, \mathbf{z}_1}(f_1) \mathcal{H}^{d+1}_{V, \mathbf{z}_2}(f_2) |\mathbf{z}_1 - \mathbf{z}_2|^{-(s+1)}$$
$$\times m^{-(d+\gamma+s+1)}.$$

Proof. For $\mathbf{z}_1 = \mathbf{0}$ and any $\mathbf{z}_2 \in V$, we can introduce a proper translation $f(\mathbf{x}) \to f(\mathbf{x} - \mathbf{x}_0)$ to move both the singular points $\mathbf{z}_1$ and $\mathbf{z}_2$ to the smaller cube $[-\frac{1}{4}, \frac{1}{4}]^d$ in V. The target quadrature error can be correspondingly reformulated as

$$\mathcal{E}_V(f(\cdot), \mathcal{X}) = \mathcal{E}_{V+\mathbf{x}_0}(f(\cdot - \mathbf{x}_0), \mathcal{X} - \mathbf{x}_0)$$
$$= \mathcal{E}_V(f(\cdot - \mathbf{x}_0), \mathcal{X} - \mathbf{x}_0),$$

which is the quadrature error of the translated function $f(\mathbf{x} - \mathbf{x}_0)$. Without loss of generality, we assume such a translation has been applied to $f(\mathbf{x})$ and $\mathcal{X}$ and both singular points $\mathbf{z}_1$ and $\mathbf{z}_2$ lie in $[-\frac{1}{4}, \frac{1}{4}]^d$.

Define a cutoff function $\psi \in \mathbb{C}_c^\infty(\mathbb{R}^n)$ satisfying

$$\psi(\mathbf{x}) = \begin{cases} 1, & |\mathbf{x}| < \frac{1}{4}, \\ 0, & |\mathbf{x}| > \frac{1}{2}. \end{cases}$$

Denote the distance between the two singular points as $\delta \mathbf{z} = |\mathbf{z}_2 - \mathbf{z}_1|$ and define two local cutoff functions that isolate the two singular points as

$$\psi_{\delta \mathbf{z}, 1}(\mathbf{x}) = \psi\left(\frac{\mathbf{x} - \mathbf{z}_1}{\delta \mathbf{z}}\right), \qquad \psi_{\delta \mathbf{z}, 2}(\mathbf{x}) = \psi\left(\frac{\mathbf{x} - \mathbf{z}_2}{\delta \mathbf{z}}\right),$$

whose supports are both inside V. The target integrand can be split into three parts as

$$f(\mathbf{x}) = f(\mathbf{x}) \psi_{\delta \mathbf{z}, 1}(\mathbf{x}) + f(\mathbf{x}) \psi_{\delta \mathbf{z}, 2}(\mathbf{x})$$
$$+ f(\mathbf{x})(1 - \psi_{\delta \mathbf{z}, 1}(\mathbf{x}) - \psi_{\delta \mathbf{z}, 1}(\mathbf{x})).$$

All three terms satisfy the periodic boundary condition on ∂V. The first term is smooth everywhere except at $\mathbf{x} = \mathbf{z}_1$ of order $\gamma + s + 1$, the second term is smooth everywhere except at $\mathbf{x} = \mathbf{z}_2$ of order 0, and the last term is smooth everywhere. Application of Lemmas 20 and 22 to these terms suggests that

$$|\mathcal{E}_V(f \psi_{\delta \mathbf{z}, 1}, \mathcal{X})| \leq C \mathcal{H}^{d+1}_{V, \mathbf{z}_1}(f \psi_{\delta \mathbf{z}, 1}) m^{-(d+\gamma+s+1)}, \quad \text{(D1)}$$

$$|\mathcal{E}_V(f \psi_{\delta \mathbf{z}, 2}, \mathcal{X})| \leq C \mathcal{H}^{d+1}_{V, \mathbf{z}_2}(f \psi_{\delta \mathbf{z}, 2}) m^{-d}, \quad \text{(D2)}$$

$$|\mathcal{E}_V(f(1 - \psi_{\delta \mathbf{z}, 1} - \psi_{\delta \mathbf{z}, 2}), \mathcal{X})|$$
$$\leq C_{|\alpha|} \left(\int_V d\mathbf{x} |\partial_\mathbf{x}^\alpha f(1 - \psi_{\delta \mathbf{z}, 1} - \psi_{\delta \mathbf{z}, 2})|\right) m^{-|\alpha|}, \quad \forall |\alpha| > d.$$

$$\text{(D3)}$$

Note that the last estimate means a superalgebraic decaying error and constant $C_{|\alpha|}$ depends only on V and $|\alpha|$ (see Eq. (H.6) in Ref. [41] for the derivation of this detailed prefactor for Lemma 20). Since $\gamma + s + 1 \leq 0$, the three estimates together prove that the quadrature error of $f(\mathbf{x})$ scales asymptotically as $m^{-(d+\gamma+s+1)}$ as $m \to \infty$. In order to describe the extreme case where the two singular points are only $\mathcal{O}(m^{-1})$ away from each other for a given mesh $\mathcal{X}$, i.e., $\delta \mathbf{z} = \mathcal{O}(m^{-1})$, we provide a precise description of the three prefactors in the estimates above using $\delta \mathbf{z}$.

For the prefactor in Eq. (D1), it is defined by

$$\mathcal{H}^l_{V, \mathbf{z}_1}(f \psi_{\delta \mathbf{z}, 1})$$
$$= \max_{|\alpha| \leq l} \left\| (\partial_\mathbf{x}^\alpha f(\mathbf{x}) \psi_{\delta \mathbf{z}, 1}(\mathbf{x})) / |\mathbf{x} - \mathbf{z}_1|^{\gamma + 1 - |\alpha|} \right\|_{L^\infty(V)},$$
$$\forall l > 0.$$

For any derivative order $|\alpha| \leq l$, we have

$$|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}}(f_1(\mathbf{x})f_2(\mathbf{x})\psi_{\delta\mathbf{z},1}(\mathbf{x}))| \leq C \sum_{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\boldsymbol{\alpha}} |\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_1}f_1(\mathbf{x})||\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})||\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_3}\psi_{\delta\mathbf{z},1}(\mathbf{x})|$$

$$\leq C \sum_{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\boldsymbol{\alpha}} \mathcal{H}_{V,\mathbf{z}_1}^l(f_1)|\mathbf{x}-\mathbf{z}_1|^{\gamma-|\boldsymbol{\alpha}_1|}|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})|\delta\mathbf{z}^{-|\boldsymbol{\alpha}_3|}\delta_{|\mathbf{x}-\mathbf{z}_1|<\frac{1}{2}\delta\mathbf{z}}$$

$$\leq C \sum_{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\boldsymbol{\alpha}} \mathcal{H}_{V,\mathbf{z}_1}^l(f_1)|\mathbf{x}-\mathbf{z}_1|^{\gamma-|\boldsymbol{\alpha}_1|-|\boldsymbol{\alpha}_3|}|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})|\delta_{|\mathbf{x}-\mathbf{z}_1|<\frac{1}{2}\delta\mathbf{z}}. \tag{D4}$$

The last inequality uses $|\mathbf{x}-\mathbf{z}_1| \leq \frac{1}{2}\delta\mathbf{z}$ by noting that $\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_3}\psi_{\delta\mathbf{z},1}(\mathbf{x})$ is zero when $|\mathbf{x}-\mathbf{z}_1| > \frac{1}{2}\delta\mathbf{z}$. Next, we estimate $|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})|$ in the ball $|\mathbf{x}-\mathbf{z}_1| \leq \frac{1}{2}\delta\mathbf{z}$. When $|\boldsymbol{\alpha}_2| \leq s$, we have

$$|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})| = \left| \sum_{|\boldsymbol{\beta}|=s+1-|\boldsymbol{\alpha}_2|} \frac{s+1-|\boldsymbol{\alpha}_2|}{\boldsymbol{\beta}!}(\mathbf{x}-\mathbf{z}_1)^{\boldsymbol{\beta}} \int_0^1 (1-t)^{s-|\boldsymbol{\alpha}_2|}\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2+\boldsymbol{\beta}}f_2(\mathbf{z}_1+t(\mathbf{x}-\mathbf{z}_1))dt \right|$$

$$\leq C|\mathbf{x}-\mathbf{z}_1|^{s+1-|\boldsymbol{\alpha}_2|} \sup_{|\mathbf{x}-\mathbf{z}_1|\leq\frac{1}{2}\delta\mathbf{z},|\boldsymbol{\beta}|=s+1} |\partial_{\mathbf{x}}^{\boldsymbol{\beta}}f_2(\mathbf{x})|$$

$$\leq C|\mathbf{x}-\mathbf{z}_1|^{s+1-|\boldsymbol{\alpha}_2|}\mathcal{H}_{V,\mathbf{z}_2}^{s+1}(f_2)\delta\mathbf{z}^{-(s+1)}.$$

The first equality applies the Taylor expansion of $\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})$ at $\mathbf{x}=\mathbf{z}_1$ with the assumption that $\partial_{\mathbf{x}}^{\boldsymbol{\beta}}f_2(\mathbf{z}_1)=\mathbf{0}$ for any $|\boldsymbol{\beta}|\leq s$. The last inequality uses the algebraic singularity of $f_2(\mathbf{x})$ at $\mathbf{x}=\mathbf{z}_2$. When $|\boldsymbol{\alpha}_2| > s$, we have

$$|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})| \leq \sup_{|\mathbf{x}-\mathbf{z}_1|<\frac{1}{2}\delta\mathbf{z}} |\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})| \leq C\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-|\boldsymbol{\alpha}_2|} \leq C\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)|\mathbf{x}-\mathbf{z}_1|^{s+1-|\boldsymbol{\alpha}_2|}\delta\mathbf{z}^{-(s+1)}.$$

Plugging the estimates of $|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})|$ above into Eq. (D4), we obtain

$$|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}}(f_1(\mathbf{x})f_2(\mathbf{x})\psi_{\delta\mathbf{z},1}(\mathbf{x}))| \leq C\mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-(s+1)}|\mathbf{x}-\mathbf{z}_1|^{\gamma+s+1-|\boldsymbol{\alpha}|}, \qquad l \geq d+1,$$

which suggests that

$$\mathcal{H}_{V,\mathbf{z}_1}^l(f\psi_{\delta\mathbf{z},1}) \leq C\mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-(s+1)}, \quad \forall \ l \geq d+1.$$

Similar analysis can also be applied to the prefactor in Eq. (D2) to obtain

$$\mathcal{H}_{V,\mathbf{z}_2}^l(f\psi_{\delta\mathbf{z},2}) \leq C\mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-(s+1)}, \quad \forall \ l \geq d+1.$$

For the estimate of the third term in Eq. (D3) with any $|\boldsymbol{\alpha}|=l\geq d+1$, the prefactor can be bounded as

$$\int_V d\mathbf{x}|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}}f(\mathbf{x})(1-\psi_{\delta\mathbf{z},1}(\mathbf{x})-\psi_{\delta\mathbf{z},2}(\mathbf{x}))|$$

$$\leq C \sum_{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\boldsymbol{\alpha}} \int_V d\mathbf{x}|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_1}f_1(\mathbf{x})||\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_2}f_2(\mathbf{x})||\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_3}(1-\psi_{\delta\mathbf{z},1}(\mathbf{x})-\psi_{\delta\mathbf{z},2}(\mathbf{x}))|$$

$$\leq C\mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2) \sum_{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\boldsymbol{\alpha}} \int_V d\mathbf{x}|\mathbf{x}-\mathbf{z}_1|^{\gamma-|\boldsymbol{\alpha}_1|}|\mathbf{x}-\mathbf{z}_2|^{-|\boldsymbol{\alpha}_2|}|\partial_{\mathbf{x}}^{\boldsymbol{\alpha}_3}(1-\psi_{\delta\mathbf{z},1}(\mathbf{x})-\psi_{\delta\mathbf{z},2}(\mathbf{x}))|.$$

For the integral with each set of $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$, we consider two cases.
(i) $\boldsymbol{\alpha}_3 = \mathbf{0}$. The integral can be bounded by

$$(*) \leq \int_{V\setminus(B(\mathbf{z}_1,\frac{1}{4}\delta\mathbf{z})\cup B(\mathbf{z}_2,\frac{1}{4}\delta\mathbf{z}))} d\mathbf{x}|\mathbf{x}-\mathbf{z}_1|^{\gamma-|\boldsymbol{\alpha}_1|}|\mathbf{x}-\mathbf{z}_2|^{|\boldsymbol{\alpha}_1|-|\boldsymbol{\alpha}|} \leq C(1+\delta\mathbf{z}^{\gamma+d-|\boldsymbol{\alpha}|}),$$

using the Hölder-inequality technique developed in the proof of Lemma 24 in Ref. [41].

(ii) $\boldsymbol{\alpha}_3 > \mathbf{0}$. The integral can be bounded by

$$(*) \le C \int_{(B(\mathbf{z}_1,\frac{1}{2}\delta\mathbf{z})\backslash B(\mathbf{z}_1,\frac{1}{4}\delta\mathbf{z}))\cup(B(\mathbf{z}_2,\frac{1}{2}\delta\mathbf{z})\backslash B(\mathbf{z}_2,\frac{1}{4}\delta\mathbf{z}))} d\mathbf{x}|\mathbf{x}-\mathbf{z}_1|^{\gamma-|\boldsymbol{\alpha}_1|}|\mathbf{x}-\mathbf{z}_2|^{-|\boldsymbol{\alpha}_2|}\delta\mathbf{z}^{-|\boldsymbol{\alpha}_3|} \le C\delta\mathbf{z}^{\gamma+d-|\boldsymbol{\alpha}|}.$$

Collecting these two terms together, the error estimate in Eq. (D3) with any $|\boldsymbol{\alpha}| = l \ge d + 1$ can be further bounded as

$$\begin{aligned}
|\mathcal{E}_V(f(1-\psi_{\delta\mathbf{z},1}-\psi_{\delta\mathbf{z},2}),\mathcal{X})| &\le C_l \mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{\gamma+d-l}m^{-l} \\
&\le C_l \mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-(s+1)}m^{-(\gamma+s+1+d-l)}m^{-l} \\
&\le C_l \mathcal{H}_{V,\mathbf{z}_1}^l(f_1)\mathcal{H}_{V,\mathbf{z}_2}^l(f_2)\delta\mathbf{z}^{-(s+1)}m^{-(\gamma+s+1+d)},
\end{aligned}$$

where the second inequality uses the assumption $\delta\mathbf{z} = \Omega(m^{-1})$ and $\gamma + s + 1 + d - l \le 0$. Gathering the above prefactor descriptions for Eqs. (D1)–(D3), we obtain

$$|\mathcal{E}_V(f,\mathcal{X})| \le C\mathcal{H}_{V,\mathbf{z}_1}^{d+1}(f_1)\mathcal{H}_{V,\mathbf{z}_2}^{d+1}(f_2)\delta\mathbf{z}^{-(s+1)}m^{-(\gamma+s+1+d)},$$

which finishes the proof. ∎

[1] R. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, England, 2008).

[2] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons* (Cambridge University Press, Cambridge, England, 2016).

[3] R. Bistritzer and A. H. MacDonald, *Moiré bands in twisted double-layer graphene*, Proc. Natl. Acad. Sci. U.S.A. **108**, 12233 (2011).

[4] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, *Ground state and hidden symmetry of magic-angle graphene at even integer filling*, Phys. Rev. X **10**, 031034 (2020).

[5] G. Makov and M. C. Payne, *Periodic boundary conditions in ab initio calculations*, Phys. Rev. B **51**, 4014 (1995).

[6] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, *Finite-size effects and Coulomb interactions in quantum Monte Carlo calculations for homogeneous systems with periodic boundary conditions*, Phys. Rev. B **53**, 1814 (1996).

[7] T. Gruber, K. Liao, T. Tsatsoulis, F. Hummel, and A. Grüneis, *Applying the coupled-cluster ansatz to solids and surfaces in the thermodynamic limit*, Phys. Rev. X **8**, 021043 (2018).

[8] J. Hermann, R. A. DiStasio Jr, and A. Tkatchenko, *First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications*, Chem. Rev. **117**, 4714 (2017).

[9] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, Cambridge, Engaldn, 2009).

[10] M. Marsman, A. Grüneis, J. Paier, and G. Kresse, *Second-order Møller-Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set*, J. Chem. Phys. **130**, 184103 (2009).

[11] A. Grüneis, M. Marsman, and G. Kresse, *Second-order Møller-Plesset perturbation theory applied to extended systems. II. Structural and energetic properties*, J. Chem. Phys. **133**, 074107 (2010).

[12] T. Gruber and A. Grüneis, *Ab initio calculations of carbon and boron nitride allotropes and their structural phase transitions using periodic coupled cluster theory*, Phys. Rev. B **98**, 134108 (2018).

[13] A. Grüneis, *A coupled cluster and Møller-Plesset perturbation theory study of the pressure induced phase transition in the LiH crystal*, J. Chem. Phys. **143**, 102817 (2015).

[14] C. Müller and B. Paulus, *Wavefunction-based electron correlation methods for solids*, Phys. Chem. Chem. Phys. **14**, 7605 (2012).

[15] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, *Towards an exact description of electronic wavefunctions in real solids*, Nature (London) **493**, 365 (2013).

[16] T. Schäfer, B. Ramberger, and G. Kresse, *Quartic scaling MP2 for solids: A highly parallelized algorithm in the plane wave basis*, J. Chem. Phys. **146**, 104101 (2017).

[17] J. McClain, Q. Sun, G. K. L. Chan, and T. C. Berkelbach, *Gaussian-based coupled-cluster theory for the ground-state and band structure of solids*, J. Chem. Theory Comput. **13**, 1209 (2017).

[18] I. Y. Zhang and A. Grüneis, *Coupled cluster theory in materials science*, Front. Mater. **6**, 123 (2019).

[19] Y. Gao, Q. Sun, J. M. Yu, M. Motta, J. McClain, A. F. White, A. J. Minnich, and Garnet Kin-Lic Chan, *Electronic structure of bulk manganese oxide and nickel oxide from coupled cluster theory*, Phys. Rev. B **101**, 165138 (2020).

[20] X. Wang and T. C. Berkelbach, *Excitons in solids from periodic equation-of-motion coupled-cluster theory*, J. Chem. Theory Comput. **16**, 3095 (2020).

[21] T. Zhu and G. K.-L. Chan, *All-electron Gaussian-based G_0W_0 for valence and core excitation energies of periodic systems*, J. Chem. Theory Comput. **17**, 727 (2021).

[22] V. A. Neufeld, H.-Z. Ye, and T. C. Berkelbach, Ground-state properties of metallic solids from *ab initio* coupled-cluster theory, J. Phys. Chem. Lett. **13**, 7497 (2022).

[23] S. J. Bintrim, T. C. Berkelbach, and H.-Z. Ye, *Integral-direct Hartree-Fock and Møller-Plesset perturbation theory for periodic systems with density fitting: Application to the benzene crystal*, J. Chem. Theory Comput. **18**, 5374 (2022).

[24] S. Banerjee and A. Y. Sokolov, *Non-Dyson algebraic diagrammatic construction theory for charged excitations in solids*, J. Chem. Theory Comput. **18**, 5337 (2022).

[25] V. A. Neufeld and T. C. Berkelbach, *Highly accurate electronic structure of metallic solids from coupled-cluster theory with nonperturbative triple excitations*, Phys. Rev. Lett. **131**, 186402 (2023).

[26] F. Coester and H. Kümmel, *Short-range correlations in nuclear wave functions*, Nucl. Phys. **17**, 477 (1960).

[27] H. Kümmel, *Origins of the coupled cluster method*, Theor. Chim. Acta **80**, 81 (1991).

[28] J. Čížek, *On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods*, J. Chem. Phys. **45**, 4256 (1966).

[29] H. J. Monkhorst and J. D. Pack, *Special points for Brillouin-zone integrations*, Phys. Rev. B **13**, 5188 (1976).

[30] S. Hirata, R. Podeszwa, M. Tobita, and R. J. Bartlett, *Coupled-cluster singles and doubles for extended systems*, J. Chem. Phys. **120**, 2581 (2004).

[31] X. Xing, X. Li, and L. Lin, *Unified analysis of finite-size error for periodic Hartree-Fock and second order Møller-Plesset perturbation theory*, Math. Comput. **93**, 679 (2024).

[32] X. Xing, X. Li, and L. Lin, *Staggered mesh method for correlation energy calculations of solids: Second-order Møller-Plesset perturbation theory*, J. Chem. Theory Comput. **17**, 4733 (2021).

[33] X. Xing and L. Lin, *Staggered mesh method for correlation energy calculations of solids: Random phase approximation in direct ring coupled cluster doubles and adiabatic connection formalisms*, J. Chem. Theory Comput. **18**, 763 (2022).

[34] K. Liao and A. Grüneis, *Communication: Finite size correction in periodic coupled cluster theory calculations of solids*, J. Chem. Phys. **145**, 141102 (2016).

[35] T. N. Mihm, B. Yang, and J. J. Shepherd, *Power laws used to extrapolate the coupled cluster correlation energy to the thermodynamic limit*, J. Chem. Theory Comput. **17**, 2752 (2021).

[36] T. N. Mihm, L. Weiler, and J. J. Shepherd, *How the exchange energy can affect the power laws used to extrapolate the coupled cluster correlation energy to the thermodynamic limit*, J. Chem. Theory Comput. **19**, 1686 (2023).

[37] T. N. Mihm, A. R. McIsaac, and J. J. Shepherd, *An optimized twist angle to find the twist-averaged correlation energy applied to the uniform electron gas*, J. Chem. Phys. **150**, 191101 (2019).

[38] T. N. Mihm, W. Z. Van Benschoten, and J. J. Shepherd, *Accelerating convergence to the thermodynamic limit with twist angle selection applied to methods beyond many-body perturbation theory*, J. Chem. Phys. **154**, 024113 (2021).

[39] L. Weiler, T. N. Mihm, and J. J. Shepherd, *Machine learning for a finite size correction in periodic coupled cluster theory calculations*, J. Chem. Phys. **156**, 204109 (2022).

[40] T. Schäfer, W. Z. Van Benschoten, J. J. Shepherd, and A. Grüneis, *Sampling the reciprocal Coulomb potential in finite and anisotropic cells*, J. Chem. Phys. **160**, 051101 (2024).

[41] X. Xing and L. Lin, *Finite-size effects in periodic coupled cluster calculations*, J. Comput. Phys. **500**, 112755 (2024).

[42] D. J. Thouless, *Stability conditions and nuclear rotations in the Hartree-Fock theory*, Nucl. Phys. **21**, 225 (1960).

[43] S. R. Gwaltney and M. Head-Gordon, *A second-order perturbative correction to the coupled-cluster singles and doubles method: CCSD (2)*, J. Chem. Phys. **115**, 2014 (2001).

[44] F. Izzo, O. Runborg, and R. Tsai, *High-order corrected trapezoidal rules for a class of singular integrals*, Adv. Comput. Math. **49**, 60 (2023).

[45] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration* (Dover, New York, 2007).

[46] 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, *PYSCF: The Python-based simulations of chemistry framework*, Comput. Mol. Sci. **8**, e1340 (2018).

[47] S. Chiesa, D. M. Ceperley, R. M. Martin, and M. Holzmann, *Finite-size error in many-body simulations with long-range interactions*, Phys. Rev. Lett. **97**, 076404 (2006).

[48] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, *Exponential localization of Wannier functions in insulators*, Phys. Rev. Lett. **98**, 046402 (2007).

[49] D. Monaco, G. Panati, A. Pisante, and S. Teufel, *Optimal decay of Wannier functions in Chern and quantum Hall insulators*, Commun. Math. Phys. **359**, 61 (2018).

[50] G. E. Scuseria and P. Y. Ayala, *Linear scaling coupled cluster and perturbation theories in the atomic orbital basis*, J. Chem. Phys. **111**, 8330 (1999).

[51] R. Schneider, *Analysis of the projected coupled cluster method in electronic structure calculation*, Numer. Math. **113**, 433 (2009).

[52] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, *Finite-size errors in continuum quantum Monte Carlo calculations*, Phys. Rev. B **78**, 125106 (2008).

[53] L. N. Trefethen and J. Weideman, *The exponentially convergent trapezoidal rule*, SIAM Rev. **56**, 385 (2014).

[54] R. Sundararaman and T. A. Arias, *Regularization of the Coulomb singularity in exact exchange by Wigner-Seitz truncated interactions: Towards chemical accuracy in nontrivial systems*, Phys. Rev. B **87**, 165122 (2013).

[55] J. Spencer and A. Alavi, *Efficient calculation of the exact exchange energy in periodic systems using a truncated Coulomb potential*, Phys. Rev. B **77**, 193110 (2008).

[56] S. Ismail-Beigi, *Truncation of periodic image interactions for confined systems*, Phys. Rev. B **73**, 233103 (2006).