PCCP

Efficient evaluation of electrostatic potential with computerized optimized code†

Jun Zhang ✶[a] and Tian Lu [b]

The evaluation of molecular electrostatic potential (ESP) is a performance bottleneck for many computational chemical tasks like restrained ESP charge fitting or quantum mechanics/molecular mechanics simulations. In this paper, an efficient algorithm for the evaluation of ESP is proposed. It regroups the expression in terms of primitive Gaussian type orbitals (GTOs) with identical angular momentum types and nuclei centers. Each term is calculated using a computerized optimized code. This algorithm was integrated into the wavefunction analysis program Multiwfn and was tested on several large systems. In the cases of dopamine and remdesivir, the performance of this algorithm was comparable to or better than some popular state-of-the-art codes. For meta1−organic framework-5, where the number of GTOs and ESP points is 4840 and 259 262, respectively, our code could finish the evaluation in 1874 seconds on ordinary hardware. It also exhibits good parallelization scaling. The source code of this algorithm is freely available and can become a useful tool for computational chemists.

1 Introduction

A molecule always generates an electrostatic potential (ESP) $\varphi(\mathbf{C})$ at point $\mathbf{C}$ in space. It is a physical observable that can be measured using, for example, high-resolution X-ray diffraction,[1] electron–electron double resonance[2,3] or spin probes.[4–6] Due to the long-range nature of electrostatic interactions, molecules can perceive others through this potential even at nanometer scaling; thus it is an important quantity for understanding physicochemical properties[7] and biological phenomena.[8] The molecular surface ESP distribution has been used to study large systems like binding free energies of anthracycline antibiotics to nucleic acids,[9] silica nanoparticle–water interface properties,[5] and ionic liquid cluster structures.[10] ESP also plays a role in exploring chemical reactivities.[11] For example, by examining ESP in transition states formed by a palladium catalyst and its substrates, it was confirmed that an acetate ligand rather than a phosphate group participates in the key catalytic step, helping to interpret the enantioselectivity of the catalyst.[12] Theoretically, several atomic and molecular properties can be expressed in terms of ESP.[7] In the development of classic force fields, spatial distribution of ESP has been extensively used to fit various kinds of atomic charges for different purposes, some examples include restrained ESP

(RESP),[13] Austin model 1 population charges with bond charge corrections (AM1-BCC),[14,15] RESP2,[16] and some others.[17–20]

Computational chemists can benefit a lot from efficient calculations of ESP. For large biomolecules modelled in an implicit dielectric medium, ESP can be determined from charge distribution by numerically solving the Poisson–Boltzmann equation[21,22] or its approximate form like the generalized Born model.[23] This class of methods is not discussed in this paper.

For small molecules (less than 1000 atoms) in the gas phase, ESP can be calculated from first principles, *i.e.*, using wavefunctions obtained from quantum chemical methods and the definition of ESP:

$$\varphi(\mathbf{C}) = \varphi_N(\mathbf{C}) + \varphi_e(\mathbf{C}) \tag{1}$$

$$\varphi_N(\mathbf{C}) = \sum_A \frac{Z_A}{|\mathbf{C} - \mathbf{A}|} \tag{2}$$

$$\varphi_e(\mathbf{C}) = -\sum_p n_p \int d\mathbf{r} \frac{|\phi_p(\mathbf{r})|^2}{|\mathbf{C} - \mathbf{r}|} \tag{3}$$

where Z_A and $\mathbf{A}$ are the nuclear charge and position of nuclear A, respectively; ϕ_p and n_p are the molecular orbital and its occupation number, respectively. n_p can be a real number in [0, 2], and thus the wavefunction can be a restricted or an unrestricted and a self-consistent field (SCF) or a correlated one. Note that by applying ∇_C^2 to eqn (1) and using $\nabla^2 \frac{1}{r} = -4\pi\delta(\mathbf{r})$, we arrive at Poisson equation:

$$\nabla_C^2 \varphi = -4\pi\rho(\mathbf{C}) \tag{4}$$

[a] *Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, People's Republic of China. E-mail: zhangjun@szbl.ac.cn*

[b] *Beijing Kein Research Center for Natural Sciences, Beijing 100022, People's Republic of China. Web: http://www.keinsci.com*

† Electronic supplementary information (ESI) available: Redundancies of a VRR path for (*f*|*d*). The *XYZ* coordinates of dopamine, remdesivir, MOF-5 and ATP. See DOI: 10.1039/d1cp02805g

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 20323–20328 | **20323**

where the charge density ρ is given by

$$\rho(\mathbf{C}) = \sum_A Z_A \delta(\mathbf{C} - \mathbf{A}) - \sum_p n_p |\phi_p(\mathbf{C})|^2 \qquad (5)$$

ESP can be calculated directly from eqn (1) (analytically solving eqn (4)[24,25] is essentially the same way), which involves expensive electron-nuclear attraction integrals (*vide infra*). In many applications, such as RESP fitting, ESP visualization, or quantum mechanics/molecular mechanics (QM/MM) simulations, more than hundreds of thousands of ESP calculations have to be carried out. Therefore, fast calculation of ESP from eqn (1) is critical for efficient and accurate theoretical simulations of large systems.

In this paper, we propose an efficient algorithm for the evaluation of ESP from the Gaussian type orbital (GTO)-based molecular wavefunctions. Briefly, it first reformulates the ESP expression to the one on a primitive GTO basis, and then each type of contribution in the expression is calculated using a code specifically optimized by computer. The new algorithm was tested in several large systems and excellent performance was observed.

2 Method

The molecular orbital ϕ_p is a linear combination of its basis functions χ_μ, *i.e.*, normalized spherical or Cartesian contracted GTOs:

$$\phi_p(\mathbf{r}) = \sum_\mu C_{\mu p} \chi_\mu(\mathbf{r}) \qquad (6)$$

In this paper, only real orbitals are considered (all $C_{\mu p}$ are real), but the generalization to complex orbitals is straightforward.

2.1 Reformulation of the electronic ESP

A straightforward way to evaluate electronic ESP by eqn (3) is contracting the density matrix $\mathbf{D}$ and the electron–nuclear attraction integral matrix $\mathbf{V}(\mathbf{C})$:

$$\varphi_e(\mathbf{C}) = \sum_{\mu\nu} D_{\mu\nu} V_{\mu\nu}(\mathbf{C}) \qquad (7)$$

$$D_{\mu\nu} = \sum_p n_p C_{\mu p} C_{\nu p} \qquad (8)$$

$$V_{\mu\nu}(\mathbf{C}) = -\int d\mathbf{r} \chi_\mu(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{C}|} \chi_\nu(\mathbf{r}) \qquad (9)$$

The bottleneck is the evaluation of $V_{\mu\nu}$. In practice, GTOs are always implemented as shells, which means that for angular momentum a, all of its $2a + 1$ spherical or $(a + 1)(a + 2)/2$ Cartesian components will be included in the basis set. To make maximum use of computational intermediates, integrals involving basis functions of the same shell are always evaluated simultaneously.

This "V-based" algorithm is far from being the most efficient one. A more powerful one, proposed in this paper, is to perform density contraction inside the electron–nuclear attraction integral evaluation subroutine, which is inspired by the

Obara–Saika recurrence relationship (OS RR) of integral evaluation.[26] First, we define the un-normalized Cartesian primitive GTO **a**:

$$\mathbf{a} \equiv (x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} \exp(-\alpha(\mathbf{r} - \mathbf{A})^2) \qquad (10)$$

where α is its exponent, and a_x, a_y, and a_z are its angular momentum components and satisfy $a_x + a_y + a_z = a$. For ease of notation, the electron–nuclear attraction integral over them is denoted by

$$(\mathbf{a}|\mathbf{b}) \equiv -\int d\mathbf{r} \frac{\mathbf{ab}}{|\mathbf{r} - \mathbf{C}|} \qquad (11)$$

Therefore, by expanding χ into **a**, eqn (3) can be reformulated as

$$\varphi_e(\mathbf{C}) = \sum_{\mathbf{ab}} D_{\mathbf{ab}}(\mathbf{a}|\mathbf{b}) \qquad (12)$$

where $D_{\mathbf{ab}}$ is the density matrix element in un-normalized Cartesian primitive GTO basis.

The original OS RRs for evaluating eqn (11) are some 6-term formulas.[26] A $(\mathbf{a}|\mathbf{b})$ can be derived from some fundamental quantities like Boys functions through many different recurrence paths. A very efficient path, discovered by Head-Gordon and Pople, is the so-called vertical RR (VRR) plus horizontal RR (HRR).[27] The VRR step is to evaluate $(\mathbf{a}|\mathbf{0})$, $(\mathbf{a} + \mathbf{1}|\mathbf{0})$, ..., $(\mathbf{a} + \mathbf{b}|\mathbf{0})$ using

$$(\mathbf{a} + \mathbf{1}_x|\mathbf{0})^N = X_{PA}(\mathbf{a}|\mathbf{0})^N + \frac{1}{2p}a_x(\mathbf{a} - \mathbf{1}_x|\mathbf{0})^N$$
$$- X_{PC}(\mathbf{a}|\mathbf{0})^{N+1} - \frac{1}{2p}a_x(\mathbf{a} - \mathbf{1}_x|\mathbf{0})^{N+1} \qquad (13)$$

and its Y, Z analogues. Here $p = \alpha_a + \alpha_b$ and $\mathbf{P} = (\alpha_a\mathbf{A} + \alpha_b\mathbf{B})/p$. This RR starts from a Boys function $(\mathbf{0}|\mathbf{0})^{a+b}$, the definition and evaluation of which can be found in a previous paper.[28]

The HRR step is to use these integrals to evaluate the target $(\mathbf{a}|\mathbf{b})$ through

$$(\mathbf{a}|\mathbf{b} + \mathbf{1}_x) = (\mathbf{a} + \mathbf{1}_x|\mathbf{b}) + X_{AB}(\mathbf{a}|\mathbf{b}) \qquad (14)$$

and its Y_{AB}, Z_{AB} analogues. Unlike eqn (13), the most important feature of eqn (14) is its independence of GTO exponents. Therefore, for all $(\mathbf{a}|\mathbf{b})$ that have identical angular momentum and $\mathbf{R}_{AB} \equiv \mathbf{A} - \mathbf{B}$ in eqn (12), the contraction with $D_{\mathbf{ab}}$ can be carried out immediately after the VRR step, and then only a single HRR step can be used to calculate the total contribution of all these $(\mathbf{a}|\mathbf{b})$ to $\varphi_e(\mathbf{C})$.

To enable this strategy, the terms in (12) are recombined according to the angular momentum pair ab and the center vector $\mathbf{R}_{AB}$:

$$\varphi_e(\mathbf{C}) = \sum_{\text{integraltype}} \sum_{\mathbf{R}_{AB}} W_{\text{integraltype}}(\mathbf{C}, \mathbf{R}_{AB})$$
$$\equiv \sum_{\mathbf{R}_{AB}} W_{ss}(\mathbf{C}, \mathbf{R}_{AB}) + \sum_{\mathbf{R}_{AB}} W_{ps}(\mathbf{C}, \mathbf{R}_{AB}) + \sum_{\mathbf{R}_{AB}} W_{pp}(\mathbf{C}, \mathbf{R}_{AB})$$
$$+ \sum_{\mathbf{R}_{AB}} W_{ds}(\mathbf{C}, \mathbf{R}_{AB}) + \sum_{\mathbf{R}_{AB}} W_{dp}(\mathbf{C}, \mathbf{R}_{AB}) + \sum_{\mathbf{R}_{AB}} W_{dd}(\mathbf{C}, \mathbf{R}_{AB}) + \cdots$$
$$(15)$$

20324 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 20323–20328

This journal is © the Owner Societies 2021

This is a highly compact representation. First, W_{ab} is always defined for $a \geq b$. This not only reduces the kinds of integral evaluation codes but also takes the integral symmetry into account (*vide infra*). Second, all one-center integrals, *i.e.*, those with $\mathbf{A} = \mathbf{B}$ and thus $\mathbf{R}_{AB} = \mathbf{0}$, can be combined into the same $W_{ab}(\mathbf{C}, \mathbf{0})$. As a result, there will be much less W than V.

Assuming that there are K exponent pairs in W_{ab}, the total number of terms in W_{ab} is therefore KN_{ab} where $N_{ab} = (a + 1)(a + 2)/2 \times (b + 1)(b + 2)/2$. For example, an explicit expression for W_{fd} is

$$
\begin{aligned}
W_{\text{fd}}(\mathbf{C}, \mathbf{R}_{AB}) &= D'_{11}(f_{x^3}|d_{x^2})^{(1)} + D'_{12}(f_{x^3}|d_{xy})^{(1)} \\
&\quad + \cdots + D'_{1,60}(f_{z^3}|d_{z^2})^{(1)} + \cdots + D'_{K1}(f_{x^3}|d_{x^3})^{(K)} \\
&\quad + D'_{K2}(f_{x^3}|d_{xy})^{(K)} + \cdots + D'_{K,60}(f_{z^3}|d_{z^2})^{(K)} \\
&\equiv W_{f_{x^3}d_{x^2}} + W_{f_{x^3}d_{xy}} + \cdots + W_{f_{z^3}d_{z^2}}
\end{aligned}
$$

$$(16)$$

In eqn (16), all of $(f|d)^{(k)}$ with the same k have identical exponent pairs. Note that $\mathbf{D}'$ has taken the integral symmetry into account. For example, for different $\mathbf{a}$ and $\mathbf{b}$, $D_{\mathbf{ab}}(\mathbf{a}|\mathbf{b}) + D_{\mathbf{ba}}(\mathbf{b}|\mathbf{a})$ in eqn (12) is reduced to $2D_{\mathbf{ab}}(\mathbf{a}|\mathbf{b})$, and thus the corresponding element of $\mathbf{D}'$ is $2D_{\mathbf{ab}}$. In the last line of eqn (16), all terms having identical angular momentum components have been collected together like, $W_{f_{x^3}d_{x^2}} \equiv \sum_{k=1}^{K} D'_{k1}(f_{x^3}|d_{x^2})^{(k)}$ *etc.* This algorithm is referred to as a "W-based" one.

2.2 Evaluation of W_{ab}

The essential part of this W-based algorithm is the efficient evaluation of W_{ab}. First, VRR is used to calculate $(\mathbf{a}|\mathbf{0})^{(k)}$, $(\mathbf{a} + \mathbf{1}|\mathbf{0})^{(k)}, \ldots, (\mathbf{a} + \mathbf{b}|\mathbf{0})^{(k)}$. Then, they are contracted with $\mathbf{D}'$ to form $W_{a0}, \ldots, W_{a+b,0}$. Finally, HRR is applied to obtain W_{ab}.

The efficient implementation of VRR and HRR (eqn (13) and (14)) is not straightforward since there are many recurrence paths arriving at the same target integral. As pointed out by one of us[28] and others[29] in the case of electron repulsion integrals, on different recurrence paths, the number of intermediates necessary for obtaining target integrals is different. This also holds for electron-nuclear attraction integrals. For example, to calculate $(f|d)$, the VRR step aims to calculate $(f|s)$, $(g|s)$, and $(h|s)$ starting from $(s|s)^N$ where $N = 5, 4, \ldots$, and 0, the number of all integrals being 126. However, in a possible path, there are 20 integrals that are never involved to obtain the target integrals using eqn (13); thus, they are redundant and can be removed from the evaluation code. For example, in the shell pair $(f|s)^2$, $(f_{x2y}|s)^2$, $(f_{x2z}|s)^2$, $(f_{xyz}|s)^2$, and $(f_{y2z}|s)^2$ are redundant (see ESI†). Therefore, the VRR step can be optimized by searching a path where the largest number of redundancies can be removed. For the contraction and HRR step, unlike the case of integral evaluation of contracted GTOs, different angular momentum components in W_{ab} are contracted with different coefficients. Thus, to eliminate redundant calculations, one should carry

out only the necessary contractions for each component. For example, in the case of Eqn (16), the contraction and HRR procedure of $W_{f_{x^3}d_{xy}}$ is:

$$
\begin{aligned}
W_{f_{x^3}s} &= \sum_{k=1}^{K} D'_{k2}(f_{x^3}|s)^{(k)} \\
W_{g_{x^4}s} &= \sum_{k=1}^{K} D'_{k2}(g_{x^4}|s)^{(k)} \\
W_{g_{x^3y}s} &= \sum_{k=1}^{K} D'_{k2}(g_{x^3y}|s)^{(k)} \\
W_{h_{x^4y}s} &= \sum_{k=1}^{K} D'_{k2}(h_{x^4y}|s)^{(k)}
\end{aligned}
$$

$$(17)$$

$$
\begin{aligned}
W_{f_{x^3}p_y} &= W_{g_{x^3y}s} + Y_{AB}W_{f_{x^3}s} \\
W_{g_{x^4}p_y} &= W_{h_{x^4y}s} + Y_{AB}W_{g_{x^4}s} \\
W_{f_{x^3}d_{xy}} &= W_{g_{x^4}p_y} + X_{AB}W_{f_{x^3}p_y}
\end{aligned}
$$

$$(18)$$

The other 59 components can be calculated in the same way. Summing them up gives W_{fd}.

A code generator was developed for building efficient evaluation codes. For every integral type from W_{ss} to W_{hh}, the code generator looked for an optimal recurrence path for the VRR and HRR steps and generated codes to carry out operations like eqn (17) and (18). Note that since K is quite large, eqn (17) can be performed using highly efficient dot product subroutines in modern linear algebra libraries. Eqn (18) only involves a few variables; thus these codes enjoy high cache hits. Therefore, the generated code is highly efficient both theoretically and technically.

3 Performance

The Fortran source code implementing the W-based algorithm is freely available from http://zhjun-sci.com/software-libreta-download.php. It has been integrated into the latest version of the wavefunction analysis program Multiwfn.[30] The performance was tested in several large systems.

3.1 Calculation of RESP charges

RESP is a commonly used atomic charge in the field of molecular dynamics. In the process of calculating RESP charges, the most time-consuming part is the evaluation of ESP at a large number of fitting points distributed around the molecular van der Waals surface. To demonstrate the efficiency of the new ESP evaluation code and compare it with other available ones, we selected dopamine and remdesivir, which are representatives of small organic molecules and large drug molecules, respectively. A wide variety of popular basis sets were taken into account. Wavefunctions were produced at the B3LYP[31] level of theory with Gaussian16.[32] Note that only the size of the basis set affects the computational cost of evaluating ESP. The test results are collectively shown in Table 1.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 20323–20328 | **20325**

Table 1 Wall clock time (in seconds) consumed for evaluating the ESP stage during calculation of RESP charges[a]

System	Basis set	N_{GTO}[b]	Multiwfn (new)[c]	Multiwfn (old)[d]	orca_vpot[e]	"code1"
Dopamine	def2-SVP	352	3	51	5	2
	aug-cc-pVDZ	594	6	170	18	5
$H_{11}C_8NO_2$	6-311G**	440	4	80	8	3
22 atoms	6-311++G**	495	5	113	12	3
13 815 points	def2-TZVP	649	7	259	16	6
	def2-QZVP	1430	31	2106	91	31
Remdesivir	def2-SVP	1307	21	609	75	20
$H_{35}C_{27}N_6O_8P$	def2-SVPD	1730	30	1383	201	47
77 atoms	6-31+G*	1508	29	807	153	23
31 180 points	6-311G**	1638	35	1045	104	31
	def2-TZVP	2440	57	3520	226	77

[a] Calculations for dopamine and remdesivir were performed on an Intel i7-10870H notebook CPU (8 cores) under a Windows 10 64 bit system and a dual Intel Xeon E5-2696v3 server CPU (totally 36 cores) under a CentOS 8 64 bit system, respectively. The "points" indicate the number of points for which the ESP will be calculated. [b] Number of primitive GTOs. [c] Multiwfn with new ESP code in this work. [d] Multiwfn with the old ESP code was developed in 2011. [e] orca_vpot in ORCA 4.2.1 was used in the current test.

For dopamine, our calculations were performed on an ordinary notebook CPU Intel i7-10870H (8 physical cores). There are in total 13 815 fitting points to calculate ESP. Table 1 suggests that for all the considered basis sets, our new ESP code embedded in Multiwfn is faster than the previous un-optimized one by one order of magnitude; for the largest basis set def2-QZVP,[33] the new code is even 70 times faster. Note that the old ESP evaluation code in Multiwfn was developed in 2011 based on an inefficient integral algorithm.[34] Clearly, the new ESP code makes the evaluation of RESP charges for small systems fairly easy on a personal computer even if a very high-quality basis set is employed. The well-known and highly efficient ORCA quantum chemistry package[35] contains a utility named orca_vpot, which was developed specifically for evaluating ESP for a given set of coordinates. From Table 1, one can see that although orca_vpot is significantly faster than the old ESP code in Multiwfn, its speed is conspicuously lower than the new ESP code in this work, especially under large basis sets. "code1" is a utility from a commercial computational software package for calculating real space functions and it also has the capability of evaluating ESP at user-provided points. Table 1 implies that "code1" has comparable efficiency with our new ESP code. However, it is worth noting that the algorithm proposed in this paper is fully primitive GTO-based and does not utilize the contraction of the basis function for additional optimization; thus our code has a strong compatibility and universality. For example, users are allowed to use WFN and WFX formats as input. Both are popular formats in the field of wavefunction analysis and do not contain information about the basis function definition. In contrast, when using "code1", users are limited to the format specific for that commercial software.

For the large remdesivir system, calculations were carried out on a mainstream computing server equipped with dual Intel E5-2696v3 CPUs (totally 36 physical cores). ESP at 31180 fitting points needs to be evaluated for calculating RESP charges of remdesivir. From Table 1, it can be seen that even using a high-quality 3-zeta basis set def2-TZVP[33] RESP charges can be readily obtained by means of our new ESP code in less

than one minute. In fact this calculation is also fully feasible on a personal computer: the cost on an Intel i7-10870H CPU is 255 seconds. Table 1 also shows that for the case of remdesivir on the 36-core CPU, the speed of the new ESP code still outperforms the old one in Multiwfn significantly, and it is even faster than orca_vpot by about 10 times under some basis sets. Although "code1" is also highly efficient, it is found to be slower than our new code under some basis sets such as def2-SVPD[36] and def2-TZVP.

3.2 Quantitative analysis of ESP over the van der Waals surface

Quantitative analysis of ESP over molecular surfaces plays an important role in, for example, predicting the condensed phase properties of molecules,[37,38] analyzing the probability of cocrystal formation,[39] and estimating molecular effective polarity.[40] Multiwfn is able to realize such analysis based on an improved marching tetrahedron algorithm[41] and it has been widely employed by users. After completing this calculation by Multiwfn, a PDB file recording the vertices can be exported to plot ESP colored van der Waals surfaces in VMD.[42] This kind of map is quite popular in the analysis of potential intermolecular interactions and preferential reaction sites.[43–45] Most of the computational overhead of this kind of analysis is for the evaluation of ESP at vertices over electron density isosurfaces. Therefore, the calculation efficiency of ESP directly determines the maximal size of the system that can be investigated by this important analysis.

To demonstrate that our new ESP code can be applied to very large systems, the quantitative analysis of ESP is applied on a truncated repeat unit of metal–organic framework-5 (MOF-5) crystals. After saturating the dangling C–C bonds by methyl groups, the chemical composition is $H_{120}C_{144}O_{104}Zn_{32}$, corresponding to as many as 400 atoms. The wavefunction used for ESP analysis was generated at the GFN2-xTB level of theory,[46] which is a very fast and robust semi-empirical method combined with a minimal contracted valence basis set augmented by polarization functions for most non-hydrogen atoms. The resulting wavefunction consists of 4840 GTOs. Using a relatively fine grid spacing of 0.25 bohr for the analysis, the number of
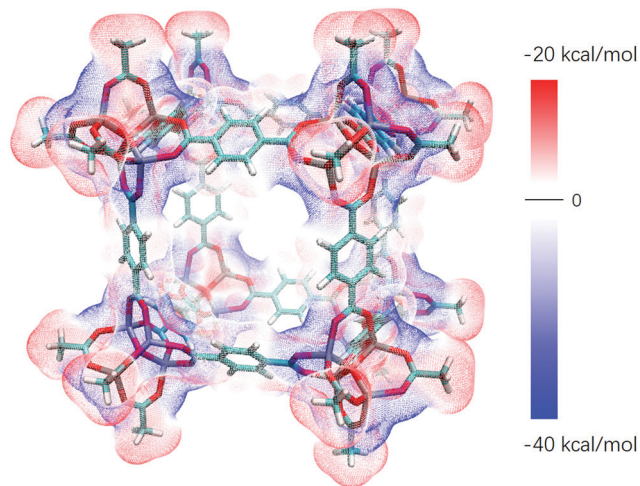
20326 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 20323–20328

This journal is © the Owner Societies 2021

Fig. 1 Truncated repeat unit of MOF-5. ESP is colored according to the color bar on the van der Waals surface (ρ = 0.001 a.u.) of the system. GFN2-xTB theory was used to generate the wavefunction.

vertices to evaluate ESP is as high as 259 262. Even in the case of such a large number of GTOs and points, on our 36-core server, the ESP evaluation only took 1874 seconds, being about half an hour. Then the ESP colored van der Waals surface map of MOF-5 can be readily rendered by VMD based on the output file of Multiwfn, see Fig. 1. This example shows that our new code enables the analysis of ESP to reach a considerable scale.

3.3 Parallel efficiency

To examine the parallel efficiency of our new ESP code, we calculated RESP charges for adenosine triphosphate (ATP, $H_{16}C_1N_5O_{13}P_3$) at the B3LYP/6-311+G(2d,p) level of theory using different numbers of CPU cores (N_{core}). Note that the parallelization was realized over the loop of fitting points of Multiwfn using OpenMP, while the subroutines for evaluating ESP are still serial. The wall clock time consumed during ESP evaluation is shown in Fig. 2. It can be seen that the parallel efficiency
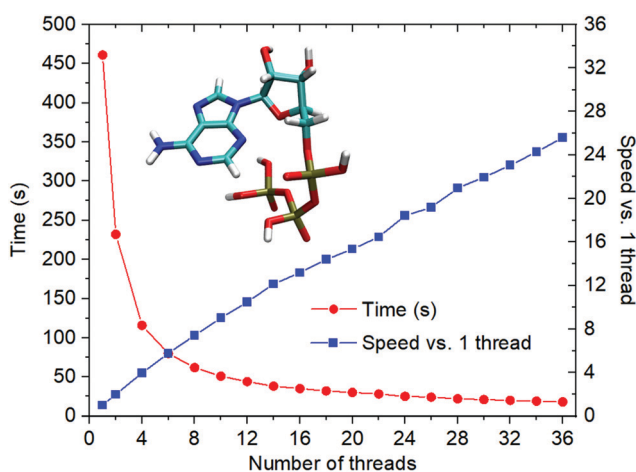

Fig. 2 Wall clock time consumed during evaluation of ESP in calculating RESP charges for ATP. The wavefunction was produced at the B3LYP/6-311+G(2d,p) level. ESP at 26 446 fitting points were calculated.

is satisfactory. When $N_{core} \leq 10$, the increase of speed is almost perfectly linear with N_{core}. Even for a large N_{core}, such as 36, the calculation speed can reach 26 times that of a single core. Therefore, using Multiwfn based on the new ESP evaluation code for ESP analysis can make full use of the computing power of the server CPU with a large number of cores.

4 Conclusions

In this paper, an efficient ESP evaluation algorithm is proposed. It transforms the electronic ESP expression into compact terms of primitive GTO basis and uses computerized optimized code for carrying out the necessary recurrence relationships. Tests of our code revealed that the new algorithm enables Multiwfn to perform expensive wavefunction analyses on ordinary software in acceptable time periods. For some large systems (say remdesivir with more than 1700 GTOs), this computerized optimized code even outperforms some popular state-of-the-art codes by up to 10 times. Our new algorithm has successfully demonstrated its applications in solving realistic chemical problems. The source code is freely available from one author's website (http://zhjun-sci.com/software-libreta-download.php) or from Multiwfn to benefit computational chemists. There is still room to improve the evaluation efficiency, like Schwarz screening, explicit SIMD instructions, or intermediate sharing between different points. These will be considered in future studies.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 A. Volkov, C. Gatti, Y. Abramov and P. Coppens, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2000, **56**, 252–258.

2 Y. Shin and W. Hubbell, *Biophys. J.*, 1992, **61**, 1443–1453.

3 J. L. Hecht, B. Honig, Y.-K. Shin and W. L. Hubbell, *J. Phys. Chem.*, 1995, **99**, 7782–7786.

4 M. A. Voinov, C. T. Scheid, I. A. Kirilyuk, D. G. Trofimov and A. I. Smirnov, *J. Phys. Chem. B*, 2017, **121**, 2443–2453.

5 V. Perelygin, M. A. Voinov, A. Marek, E. Ou, J. Krim, D. Brenner, T. I. Smirnova and A. I. Smirnov, *J. Phys. Chem. C*, 2019, **123**, 29972–29985.

6 E. G. Kovaleva, L. S. Molochnikov, D. Tambasova, A. Marek, M. Chestnut, V. A. Osipova, D. O. Antonov, I. A. Kirilyuk and A. I. Smirnov, *J. Membr. Sci.*, 2020, **604**, 118084.

7 P. Politzer and J. S. Murray, *Theor. Chem. Acc.*, 2002, **108**, 134–142.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 20323–20328 | **20327**

8 P. Politzer, P. R. Laurence and K. Jayasuriya, *Environ. Health Perspect.*, 1985, **61**, 191–202.

9 M. Baginski, F. Fogolari and J. M. Briggs, *J. Mol. Biol.*, 1997, **274**, 253–267.

10 J. Zhang, E. T. Baxter, M.-T. Nguyen, V. Prabhakaran, R. Rousseau, G. E. Johnson and V.-A. Glezakou, *J. Phys. Chem. Lett.*, 2020, **11**, 6844–6851.

11 J. S. Murray and P. Politzer, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 153–163.

12 J. Zhang, *Org. Biomol. Chem.*, 2018, **16**, 8064–8071.

13 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

14 A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2000, **21**, 132–146.

15 A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.

16 M. Schauperl, P. S. Nerenberg, H. Jang, L.-P. Wang, C. I. Bayly, D. L. Mobley and M. K. Gilson, *Chem. Commun.*, 2020, **3**, 1–11.

17 R. H. Henchman and J. W. Essex, *J. Comput. Chem.*, 1999, **20**, 483–498.

18 Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo and T. Lee, *et al.*, *J. Comput. Chem.*, 2003, **24**, 1999–2012.

19 C. Campañá, B. Mussard and T. K. Woo, *J. Chem. Theory Comput.*, 2009, **5**, 2866–2878.

20 T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput.*, 2010, **6**, 2455–2468.

21 N. A. Baker, *Rev. Comput. Chem.*, 2005, **21**, 349.

22 F. Dong, B. Olsen and N. A. Baker, *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*, Academic Press, 2008, vol. 84, pp. 843–870.

23 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.

24 S. Srebrenik, H. Weinstein and R. Pauncz, *Chem. Phys. Lett.*, 1973, **20**, 419–423.

25 O. Mó and M. Yáñez, *Theor. Chim. Acta*, 1978, **47**, 263–273.

26 S. Obara and A. Saika, *J. Chem. Phys.*, 1988, **89**, 1540–1559.

27 M. Head-Gordon and J. A. Pople, *J. Chem. Phys.*, 1988, **89**, 5777–5786.

28 J. Zhang, *J. Chem. Theory Comput.*, 2018, **14**, 572–587.

29 B. G. Johnson, P. M. W. Gill and J. A. Pople, *Int. J. Quantum Chem.*, 1991, **40**, 809–827.

30 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.

31 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

32 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revision A.03*, 2016, Gaussian Inc., Wallingford CT.

33 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.

34 D. B. Cook, *Handbook of Computational Quantum Chemistry*, Dover Publications, Inc., 2005.

35 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**, e1327.

36 D. Rappoport and F. Furche, *J. Chem. Phys.*, 2010, **133**, 134105.

37 J. S. Murray and P. Politzer, *THEOCHEM*, 1998, **425**, 107–114.

38 J. S. Murray, F. Abu-Awwad and P. Politzer, *J. Phys. Chem. A*, 1999, **103**, 1853–1856.

39 D. Musumeci, C. A. Hunter, R. Prohens, S. Scuderi and J. F. McCabe, *Chem. Sci.*, 2011, **2**, 883–890.

40 Z. Liu, T. Lu and Q. Chen, *Carbon*, 2021, **171**, 514–523.

41 T. Lu and F. Chen, *J. Mol. Graphics Modell.*, 2012, **38**, 314–323.

42 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.

43 S. Manzetti and T. Lu, *J. Phys. Org. Chem.*, 2013, **26**, 473–483.

44 S. Manzetti, T. Lu, H. Behzadi, M. D. Estrafili, H.-L. Thi Le and H. Vach, *RSC Adv.*, 2015, **5**, 78192–78208.

45 T. Lu and S. Manzetti, *Struct. Chem.*, 2014, **25**, 1521–1533.

46 C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.