Expeditious Stochastic Approach for MP2 Energies in Large Electronic Systems

Daniel Neuhauser,*,[†] Eran Rabani,*,[‡] and Roi Baer*,[§]

[†]Department of Chemistry and Biochemistry, University of California, Los Angeles, California, 90095, United States
[‡]School of Chemistry, The Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[§]Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, Hebrew University, Jerusalem 91904, Israel

ABSTRACT: A fast stochastic method for calculating the second order Møller-Plesset (MP2) correction to the correlation energy of large systems of electrons is presented. The approach is based on reducing the exact summation over occupied and unoccupied states to a time-dependent trace formula amenable to stochastic sampling. We demonstrate the abilities of the method to treat systems with thousands of electrons using hydrogen passivated silicon spherical nanocrystals represented on a real space grid, much beyond the capabilities of present day MP2 implementations.

$$E_{MP2} = \sum_{ijab} \frac{(2\langle ab|ij\rangle - \langle ab|ji\rangle)\langle ij|ab\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$$
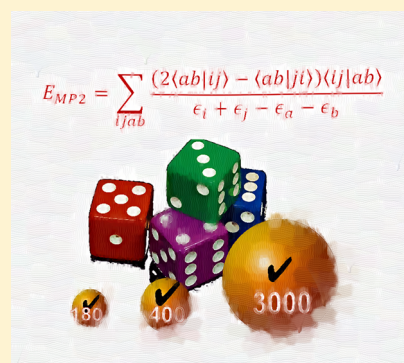
Post Hartree−Fock (HF), fifth rung density functional theory (DFT) and many-body perturbation theory (MBPT) calculations are of importance for the estimation of observables dependent on correlation energy in large systems. Examples include the calculation of cohesion energies and molecular geometries, phonon/vibrational properties in large molecules,[1−11] electron dynamics,[12−15] and quasiparticle energies and gaps in nanocrystals.[16−23] Application of such methods in a straightforward way to large systems of experimental relevance is often hampered by the steep scaling of the computational effort (CPU time and memory) with system size. This is mainly due to the huge number of electron−electron Coulomb integrals needed for the calculation, namely between all pairs of occupied and unoccupied orbitals. The problem is especially critical for grid-based calculations, where the number of unoccupied orbitals easily reaches millions. Thus, accurate first principles post HF/DFT methods cannot treat but the smallest systems, and there is an urgent need for new ideas to overcome the formidable barriers.

Lowering the scaling of such calculations is the topic of recent works, such as localized density fitting methods[24,25] and virtual orbital selection techniques.[26−28] Such approaches lead to the development of linear scaling local second order Møller−Plesset (LMP2) techniques[11,24] which are extremely efficient for low dimensional systems and have been used successfully for small molecules containing $O(100)$ atoms. However, for higher dimensional systems, these methods are too expensive due to the computational cost of a prefactor.

In this letter, we develop an alternative approach that allows treatment of post HF/DFT calculations for large systems and is not particularly sensitive to dimensionality. The basic idea is to use stochastic methods, giving up accuracy (introducing statistical errors) in exchange for efficiency in a controlled and automated way. Ultimately, we aim at systems such as nanocrystals, containing $O(10^4)$ atoms and $O(10^5)$ valence electrons, where one does not need high accuracy in absolute total energy but requires a small average error per electron. As an example of a case of interest, we refer the reader to a recent stochastic correlation method that has been developed to study multiexciton generation (MEG) rates in nanocrystals.[29] The calculation can address questions such as the dependence of the MEG rate on the size and shape of the nanocrystal or the comparison of the MEG rate with phonon emission decay processes with relative accuracies of 1% at reasonable computational effort

Here, we focus on the simplest of correlated electron calculations, namely the MP2 correlation energy, yet the concept we introduce is general and transferable to more challenging correlation calculations. We rely on a novel combination of stochastic and operator techniques allowing MP2 calculations to be performed in grid-based representation with linear scaling effort both in time and storage. The methods we develop here share some ideas from recent work,[29−31] but the present combination and its use for MP2 is new.

The MP2 energy for a closed shell system, in terms of spatial orbitals, is

$$E_{\mathrm{MP2}} = \sum_{ijab} \frac{(2\langle ab|ij\rangle - \langle ab|ji\rangle)\langle ij|ab\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \tag{1}$$
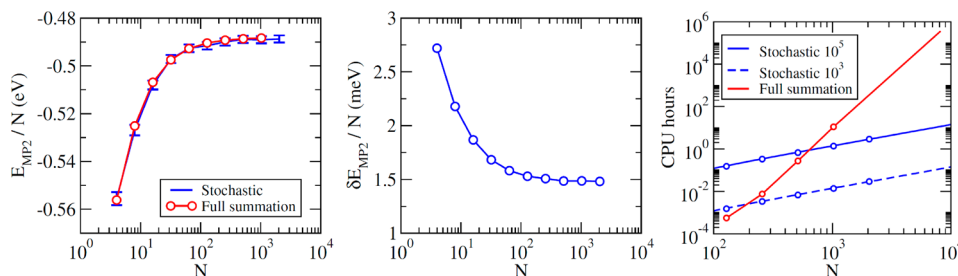
Figure 1. Left panel: Full summation vs a stochastic calculation with $\sim 10^5$ iterations for the E_{MP2} energy per electron for the half-filled model as a function of the number of electrons (sites) N. Middle panel: The MC error bars per electron decrease toward a plateau as the system size grows. Right panel: The CPU times for full summation (scaling as N^5) vs MC averaging ($\sim O(N)$) for the MP2 energy for 10^5 (blue solid) and 10^3 (blue dashed) MC runs. The costs above $N = 1024$ for full-summation and $N = 2048$ for the MC are extrapolated.

Here, $\langle ab|ij \rangle = \iint a(\mathbf{x}_1)^* b(\mathbf{x}_2)^* x_{12}^{-1} i(\mathbf{x}_1) j(\mathbf{x}_2) \, d^3x_1 \, d^3x_2$, ε_i, ε_j (ε_a, ε_b) and $i(\mathbf{x})$, $j(\mathbf{x})$ ($a(\mathbf{x})$, $(b(\mathbf{x}))$ are, respectively, the occupied (virtual) eigenvalues and eigenfunctions of the Fock operator $\hat{H}_0$. When calculated as-is without further processing, we label this formula as "full-summation." For nanocrystals of experimentally relevant sizes, the density of states in the valence and conduction bands reaches several thousands per electron-volt;[29] thus the summation in eq 1 involves more than 10^{12} terms, but the nonlocal nature of the orbitals increases the computations to 10^{15} to 10^{16} operations.

To replace the full summation by random averages, we need to circumvent the state-specific energy denominator. This is done through time integration, using the identity Im $\int_0^\infty e^{i\varepsilon t} \, \text{erfc}(t/\tau) \, dt = \frac{1}{\varepsilon}(1 - e^{-[(\varepsilon\tau)/2]^2})$, where $\text{erfc}(x)$ is the complementary error function. This expression converges to $1/\varepsilon$ once $\tau \gg 1/(2E_g)$ where E_g is the HOMO−LUMO gap (note: it is not recommended to use the Laplace transform[32] instead of the above identity; the advantage of using the erfc is that it is not sensitive to the sign of ε and so can be used for cases where denominators are not necessarily positive preventing instabilities by numerical contamination of the eigenstates). Defining $\phi_j(t) \equiv e^{-i\varepsilon_j t} j(\mathbf{x}) = (e^{-i\hat{H}_0 t} j)(\mathbf{x})$, where $\hat{H}_0$ is the DFT Hamiltonian (and we set $\hbar = 1$), the MP2 correlation energy takes the form:

$$E_{MP2} = \text{Im} \int_0^\infty C(t) \, \text{erfc}\left(\frac{t}{\tau}\right) dt \tag{2}$$

where

$$C(t) \equiv \sum_{ijab} (2\langle ab|ij \rangle - \langle ab|ji \rangle) \times \langle i(t)j(t)|a(t)b(t) \rangle \tag{3}$$

We now replace the summation using a stochastic averaging procedure. We introduce four random wave functions:

$$\eta(\mathbf{x}) \equiv \sum_i \eta_i i(\mathbf{x}), \quad \chi(\mathbf{x}) \equiv \sum_j \chi_j j(\mathbf{x}),$$

$$\xi(\mathbf{x}) \equiv \sum_a \xi_a a(\mathbf{x}), \quad \zeta(\mathbf{x}) \equiv \sum_b \zeta_b b(\mathbf{x}) \tag{4}$$

where η_i etc. are independent complex random numbers, normalized so that $[\eta_i^* \eta_{i'}] = \delta_{ii'}$ (and similarly for the χ's, ζ's, and ξ's) and $[\eta_i^* \chi_j] = [\eta_i^* \xi_a] = ... = 0$. Here, $[...]$ denotes the average over the random variables. With these, $C(t)$ is given by

$$C(t) = [(2\langle \xi\zeta|\eta\chi \rangle - \langle \xi\zeta|\chi\eta \rangle) \times \langle \eta(t)\chi(t)|\xi(t)\zeta(t) \rangle] \tag{5}$$

We consider in this work cases where the operation of $\hat{H}_0$ scales linearly with system size. This is true in semiempirical and basis-set calculations, where $\hat{H}_0$ is the sparse Fockian matrix (for sufficiently large systems), and it is true in DFT using grid representations (based on local/semilocal and hybrid exchange correlation potentials) where MP2 type calculations can be used in the context of double hybrid corrections.[23]

Naïve construction of random functions through eqs 4 requires the calculation and storage of all possible eigenstates, a prohibitive task for large grid systems. Alternatively, one can perform this task using operator techniques. We first choose a function with random values at each grid point $\mathbf{x}_n$, e.g., for 3D systems,

$$\eta_0(\mathbf{x}_n) = \frac{e^{i\theta_n}}{h^{3/2}} \tag{6}$$

where θ_n is a random sampled uniformly from $[0,2\pi]$ and h is the grid spacing. A similar procedure applies for constructing the other three random functions χ_0, ζ_0, and ξ_0. The random occupied-space function η is now obtained by applying a "purification operator," $\eta(\mathbf{x}) = \langle \mathbf{x}|\theta(\mu - \hat{H}_0)|\eta_0 \rangle$, where $\theta(E) \equiv \text{erfc}(-E/T)/2$ and μ is the chemical potential (placed in the middle of the HOMO−LUMO gap). $T \approx 1/4E_g$ is an artificial temperature, and E_g is the HOMO−LUMO gap. To apply $\theta(E)$, we use an iterative Chebyshev expansion for the step function:[33]

$$\eta(\mathbf{x}) = \langle \mathbf{x}|\theta(\mu - \hat{H}_0)|\eta_0 \rangle = \sum_K^{k=0} \gamma_k \eta_k(\mathbf{x}) \tag{7}$$

where

$$\eta_1 = \hat{H}^N \eta_0, \quad \eta_k = 2\hat{H}^N \eta_{k-1} - \eta_{k-2} \tag{8}$$

and $\hat{H}^N = (\hat{H}_0 - \bar{H})/(\Delta H)$ is a normalized Hamiltonian[34] where $\bar{H}$ and ΔH are the average and half-width of the spectrum of $\hat{H}_0$; the normalization is required as the Chebyshev expansion applies for operators with spectra between -1 and $+1$.[34] The coefficients γ_k are obtained using a numerical Fourier transform of the periodic function $f(\alpha) = \theta(\mu - (\bar{H} + \Delta H \cos \alpha))$. We repeat eq 7 also for χ, ζ, and ξ; however, for the latter two, since they must be projected into the virtual space, we subtract the "occupied" projected part from the original function to obtain the "virtual" projected part, namely,

$$\zeta \leftarrow \zeta_0 - \zeta, \xi \leftarrow \xi_0 - \xi \tag{9}$$

The number of required Chebyshev steps in the purification stage can be estimated to be $K \approx 30 + (10\Delta H)/T$.

Consider now the time propagation of each random orbital. The total propagation time is $t_{max} \approx 3\tau$, determined by the parameter $\tau \gg E_g^{-1}$ governing the exponential damping of $C(t)$. Each random orbital is evolved using a Chebyshev series of the propagator[35] $e^{-i\hat{H}_0 t}$, and the number of terms in the series is $\sim 30 + t_{max}\Delta H$. Alternatively, one can use a split operator to carry out the propagation.

Summarizing the algorithm:

1. Construct the random complex functions, $\eta_0(\boldsymbol{x})$, $\chi_0(\boldsymbol{x})$, $\zeta_0(\boldsymbol{x})$, and $\xi_0(\boldsymbol{x})$ on the grid (eq 4).
2. Purify these functions to yield random occupied $\eta(\boldsymbol{x})$ and $\chi(\boldsymbol{x})$ (eqs 7 and 8) and virtual $\zeta(\boldsymbol{x})$ and $\xi(\boldsymbol{x})$ (eq 9) state functions.
3. Propagate the purified functions to time t_{max}. At each time step along the propagation, evaluate the two-electron integrals and the MP2 correlation function given by eq 5, and carry out the time integration in eq 2.
4. Repeat steps 1−3 to average over the random functions.

To study the performance of this method, we first consider a Pariser−Parr−Pople Hamiltonian of a one-dimensional system of N sites and N electrons at half-filling, described in ref 36. In the limit $N \rightarrow \infty$, this model has a gap $E_g = 0.11E_h$ and energy span $\Delta H = 0.3E_h$. For the calculation, we took (in atomic units) $\tau = 20$ and $t_{max} = 60$.

Figure 1 (left) shows the MP2 energies per site as a function of the number of sites (electrons) from a tiny ($N = 4$) to a large system ($N = 2048$). There is excellent agreement between the full-summation and stochastic results, and the scale of the deviation, per particle, is far better than chemical accuracy. The middle panel of Figure 1 shows that the statistical error per electron decreases toward a plateau as the number of electrons increases. The asymptotic value of the error is proportional to the square root of the number of MC iterations and can be easily controlled. The right panel of Figure 1 shows the cost in CPU hours (measured per a regular 2.5 GHZ i7 Intel processor) of the full-summation and the stochastic methods (costs beyond 2048 sites for the stochastic method and 1024 for the full-summation are extrapolated based on the methods' scaling). We also show results for lower accuracy (10^3) MC runs. The calculations are very fast and still yield chemical accuracies of about 10−30 meV per electron. The stochastic method is already faster than the direct MP2 calculation for 600 sites at high accuracy and is 2 orders of magnitude faster for 2000 electrons/sites.

Next, we apply the stochastic method to hydrogen passivated spherical silicon nanocrystals (NCs) of several sizes, reaching systems with over 3000 electrons. The different systems are described in Table 1. We use a semiempirical pseudopotential model to construct $\hat{H}_0$ as described in ref 37.

Table 1. Parameters for the Hydrogen Passivated Silicon Nanocrystals (NC)[a]

N_{Si}	N_H	N_e	D (nm)	grid	E_g (eV)
1	4	8		8^3	10.7
35	36	176	1.3	32^3	3.9
87	76	424	1.6	48^3	3.2
353	196	1608	2.4	64^3	2.2
705	300	3120	3.0	72^3	2.0

[a]Shown, the number of silicon (N_{Si}) and hydrogen atoms (N_H), the number of electrons (N_e), the NC diameter (D), the number of grid points, and the occupied-virtual energy gap E_g.

In the lower panel of Figure 2, we show the MP2 energy for the silicon NCs, which grow with the size of the NC. The full
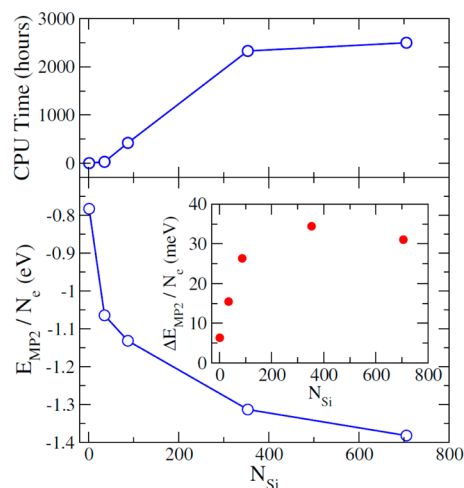
Figure 2. MP2 calculation data for hydrogen-passivated silicon nanoclusters (NCs) with N_{Si} silicon atoms. Top panel: CPU time for computing the MP2 correlation energy to a statistical error of 30 meV. Bottom panel: The MP2 energy per electron as a function of NC size. Inset: the statistical error for 2500 MC iterations as a function of NC size.

summation MP2 results can only be carried out for the smallest system, where we find that 20 000 MC iterations are sufficient to converge the results within an error smaller than 1 meV.

The top panel of Figure 2 shows the total CPU time for a calculation that yields a statistical error of 30 meV per electron. Of course, the error can be reduced by running longer in time, and for the largest NC an error of 1% requires 10^4 CPU hours. It is seen that CPU times grow by a factor of ~ 5 when going from the 87 to the 705 silicon atom clusters. There are two distinct regimes for the scaling which are influenced by several factors: (a) The decrease of E_g with NC size (almost a factor of 2 between the smallest and largest NCs studied) leads to an increase in the Chebyshev expansion by the same factor. (b) The statistical error grows when going from the small sized to 353 NCs, requiring more MC iterations. The difference in CPU time from $Si_{353}H_{196}$ to $Si_{705}H_{300}$ is rather small (sub-linear), since the decrease in E_g is rather small and this is also correlated with a decrease in the statistical errors. We note in passing that for even larger NCs (above 3 nm), E_g becomes nearly system size independent, and both the Chebyshev length and statistical errors saturate (the latter can also reduce with system size). Indication of this is seen in the inset of Figure 2, where the dependence of the statistical errors on system size (for 2500 iterations) is shown.

Summarizing, we have developed a stochastic method for calculating the MP2 correlation energies for systems of unprecedented sizes. The method was shown capable of addressing silicon nanocrystals with over 3000 electrons. Comparing to linear scaling local MP2,[11,24] the present method is considerably less dependent on dimensionality since it is not highly sensitive to density matrix localization, as long as the zero order Hamiltonian can be operated in linear scaling. A second advantage of the method is its low memory requirements, involving only four orbitals, while linear scaling local MP2 will still require a huge amount of memory in calculations for three-dimensional dense systems. Another advantage is that

the stochastic method is naturally parallelizable where each processor is an independent sampler and propagator and due to the low memory requirements can be implemented on graphic processing units.

There are several ways to extend the proposed approach. Higher order MPs, e.g., MP3 or MP4, seem within reach using similar ideas: one needs to replace energy denominators by time integrations as done here. This has the potential to allow efficient calculation of high order perturbation energy calculations. Another interesting feature is the extension to finite temperatures, which is straightforward to implement within the present formalism. Finally, we plan to investigate whether for very large sizes a localization scheme could reduce the error per electron as the system gets larger.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: roi.baer@huji.ac.il; dxn@chem.ucla.edu; eran.rabani@gmail.com.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Marsman, M.; Gruneis, A.; Paier, J.; Kresse, G. *J. Chem. Phys.* **2009**, *130*, 184103.

(2) Reimers, J. R. *Computational Methods for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*; Wiley: Hoboken, NJ, 2011.

(3) Casassa, S.; Halo, M.; Maschio, L.; et al. *Theor. Chem. Acc.* **2007**, *117*, 781.

(4) Feller, D.; Jordan, K. D. *J. Phys. Chem. A* **2000**, *104*, 9971.

(5) Furche, F. *J. Chem. Phys.* **2008**, *129*, 114105.

(6) Hellgren, M.; Rohr, D. R.; Gross, E. K. U. *J. Chem. Phys.* **2012**, *136*, 034106.

(7) Kozuch, S.; Martin, J. M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104.

(8) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *131*, 174105.

(9) Zhang, Y.; Xu, X.; Goddard, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4963.

(10) Sharkas, K.; Toulouse, J.; Savin, A. *J. Chem. Phys.* **2011**, *134*, 064113.

(11) Schutz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691.

(12) Rabani, E.; Baer, R. *Nano Lett.* **2008**, *8*, 4488.

(13) Jaeger, H. M.; Fischer, S.; Prezhdo, O. V. *J. Chem. Phys.* **2012**, *136*, 064701.

(14) Reboredo, F. A.; Franceschetti, A.; Zunger, A. *Phys. Rev. B* **2000**, *61*, 13073.

(15) Wang, X. Y.; Ren, X. F.; Kahen, K.; et al. *Nature* **2009**, *459*, 686.

(16) Ogut, S.; Chelikowsky, J. R.; Louie, S. G. *Phys. Rev. Lett.* **1997**, *79*, 1770.

(17) Berman, O.; Mukamel, S. *Phys. Rev. A* **2003**, *67*, 042503.

(18) Blase, X.; Attaccalite, C.; Olevano, V. *Phys. Rev. B* **2011**, *83*, 115103.

(19) Hybertsen, M. S.; Louie, S. G. *Phys. Rev. B* **1986**, *34*, 5390.

(20) Kümmel, S.; Kronik, L. *Rev. Mod. Phys.* **2008**, *80*, 3.

(21) Samsonidze, G.; Jain, M.; Deslippe, J.; et al. *Phys. Rev. Lett.* **2011**, *107*, 186404.

(22) Onida, G.; Reining, L.; Rubio, A. *Rev. Mod. Phys.* **2002**, *74*, 601.

(23) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(24) Werner, H. J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149.

(25) Schutz, M.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349.

(26) Kurashige, Y.; Yang, J.; Chan, G. K. L.; Manby, F. R. *J. Chem. Phys.* **2012**, *136*, 124106.

(27) Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K. L. *J. Chem. Phys.* **2011**, *134*, 044123.

(28) Hohenstein, E. G.; Parrish, R. M.; Martinez, T. J. *J. Chem. Phys.* **2012**, *137*, 044103.

(29) Baer, R.; Rabani, E. *Nano Lett.* **2012**, *12*, 2123.

(30) Baer, R.; Seideman, T.; Ilani, S.; Neuhauser, D. *J. Chem. Phys.* **2004**, *120*, 3387.

(31) Baer, R.; Neuhauser, D. *J. Chem. Phys.* **2012**, *137*, 051103.

(32) Almlöf, J. *Chem. Phys. Lett.* **1991**, *181*, 319.

(33) Baer, R.; Head-Gordon, M. *J. Chem. Phys.* **1997**, *107*, 10003.

(34) Kosloff, R. *J. Phys. Chem.* **1988**, *92*, 2087.

(35) Tal-Ezer, H.; Kosloff, R. *J. Chem. Phys.* **1984**, *81*, 3967.

(36) Adhikari, S.; Baer, R. *J. Chem. Phys.* **2001**, *115*, 11.

(37) Wang, L. W.; Zunger, A. *J. Phys. Chem.* **1994**, *98*, 2158.