Chemical Physics Letters 527 (2012) 86–88

Contents lists available at SciVerse ScienceDirect

Chemical Physics Letters

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

Reply to Comment

Reply to Comment on 'On the optimal symmetric purification scheme of the one-particle density matrix'

Jaehoon Kim, Yousung Jung *

Graduate School of EEWS and KI for Nanocentury, KAIST, Daejeon 305-701, Republic of Korea

ARTICLE INFO

Article history:
Available online 17 January 2012

ABSTRACT

Rubensson suggested in his Comment that the 9th-order function is an optimal purification degree if one uses a Paterson–Stockmeyer method to evaluate the Holas polynomials, unlike our earlier conclusion that the 5th-order is optimal. Here we show that the Paterson–Stockmeyer factorization to evaluate the 9th-order Holas polynomial is numerically significantly less stable than the 5th-order symmetric form due to the large expansion coefficients involved. When numerical truncation is introduced as is necessary for linear scaling SCF calculations, we show that this numerical error indeed leads to a higher computational cost for the 9th-order purification as compared to the 5th-order function, leaving our previous conclusion unchanged.

Recently, we derived analytically that an optimal scheme for density matrix purification is to use the 5th-order Holas polynomial throughout all iterations [1]. In a Comment by Rubensson [2], he concludes instead that the 9th-order function is optimal if one uses a Paterson–Stockmeyer (PS) polynomial evaluation method for Holas functions. In this Reply, we show that the 9th-order Holas polynomial evaluated using the PS method causes a significant numerical instability due to the large expansion coefficients involved in the PS scheme, when dealing with truncated sparse matrices for linear scaling self-consistent field (SCF) calculations. This numerical error leads to a higher computational cost for the 9th-order function compared to the 5th-order function for real applications. We then demonstrate several ways to reduce numerical errors associated with evaluating the 9th-order Holas polynomial by reducing the magnitudes of the expansion coefficients, but show that the 5th-order polynomial is still a more efficient scheme.

We begin by proving that our Theorem 2 in Ref. [1] still holds when the PS factorization is used to evaluate the purification polynomials $g_n(\mathbf{P}_o)$. That is, if the purification function is required to purify $\mathbf{P}$ at each iteration, the optimal purification scheme is when the polynomials at all iterations are of the same degree. The number of matmuls needed to evaluate $g_n(\mathbf{P}_o)$ using the PS scheme [3] is asymptotically $2\sqrt{2n}$ for large n, roughly $n \geqslant 10$. The cost function would then be approximately proportional to

$$t([n_i]) = \sum_{i=1}^{m} \sqrt{n_i} \tag{1}$$

Again we use the inequality of arithmetic and geometric means,

$$t([n_i]) = \sum_{i=1}^{m} \sqrt{n_i} \geq m \sqrt[2m]{\Pi_{i=1}^{m} n_i} \approx m \sqrt[2m]{k} \tag{2}$$

where $k \sim \Pi_{i=1}^{m} n_i$ is a measure of target accuracy, the exponent of error after m iterations. As before, the equality in Eq. (2) holds true when all n_i are the same, proving that Theorem 2 is still valid even if the PS method is used. In addition, it is not difficult to prove analytically that the optimal purification scheme is then to use the lowest order polynomial throughout all iterations with roughly $n \geqslant 10$. For smaller n before this asymptotic behavior $2\sqrt{2n}$ is reached, however, one needs to test the optimal purification order only numerically.

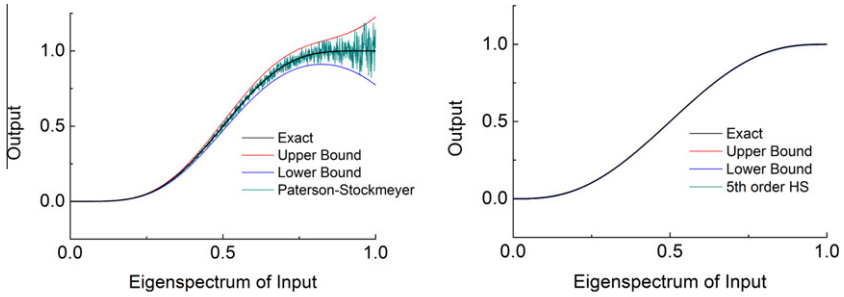
Rubenssen in his Comment demonstrated by numerical experiments for a toy problem that $n = 5$ (the 9th-order Holas function) leads to a faster convergence to a desired step function than $n = 3$ (the 5th-order Holas function) if the PS method of polynomial evaluation [3] is used in a particular sequence as shown in Eq. (3)

$$g_5^{PS}(x) = (126x^2 + (-420 + 540x - 315x^2 + 70x^3)x^3)x^3 \tag{3}$$

When using purification algorithm for linear scaling SCF calculations, the essence is the sparsity of Hamiltonian and density matrices whose multiplications after numerical truncation can be performed much faster than the exact matrix multiplications without truncation. Here, we focus on how these matrix truncation errors accumulate during the evaluation of purification functions depending on the degree of polynomial and its form, and how they affect the total computational cost. Thus, we explored the numeri cal stability of the PS method to evaluate the Holas 9th-order polynomial. To show it in a toy problem, we perturbed every operation

* Corresponding author.
E-mail address: ysjn@kaist.ac.kr (Y. Jung).

0009-2614/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.cplett.2012.01.015



(a) 9th-order purification function (b) 5th-order purification function

Fig. 1. Accumulation of numerical errors associated with (a) the evaluation of 9th-order Holas purification function using the Paterson–Stockmeyer method, Eq. (3), and (b) the evaluation of 5th-order function using the symmetry property. To introduce numerical noise, every operation was perturbed by 10^{-3}. The upper and lower bounds are calculated as twice the standard deviation of errors for each point.

of addition and multiplication by adding a uniform random number in the interval $[-10^{-3}, 10^{-3}]$, during the evaluation of $g_5^{\mathrm{PS}}(x)$, Eq. (3).

As shown in Fig. 1, the evaluation of $g_5^{\mathrm{PS}}(x)$ using the PS method is unstable, and errors get amplified at around one. One can expand $g_5^{\mathrm{PS}}(x)$ with the power of $(x-a)^n$ for some a to reduce errors, but only to a limited extent. As in Fig. 1b, the 5th-order Holas function, however, does not show such instability. Introducing (more realistic) smaller perturbations in the interval $[-10^{-9}, 10^{-9}]$ also shows the same behavior: average error for the PS method is approximately 10^{-7} while for the 5th-order function the error is about 10^{-9}.

The reason for this numerical instability in the PS form of 9th-order Holas polynomial is due to the large coefficients involved. When the density matrix is idempotent, $g_5^{\mathrm{PS}}(P)$ can be written as $(126\mathbf{P} + (-125\mathbf{P}')\mathbf{P}^3)\mathbf{P}^3$, where we emphasized $\mathbf{P}'$ such that it has a different error from $\mathbf{P}$. Therefore the calculation results may contain numerical errors that can be 100 times larger than the true density matrix. Since the sparse matrix multiplication truncates all elements below a certain threshold (thresh) value as in linear scaling SCF calculations, one cannot purify density matrix with errors below 100 × thresh. Even if no truncation is used, since the matrix multiplications or additions have an intrinsic numerical error, the magnitudes of errors accumulated using Eq. (3) would be larger than those using the functional forms with smaller coefficients as described below.

There are several ways to remove or at least reduce these large coefficients in Eq. (3). One method is to use the particle–hole expansion (PH) of Mazziotti [4]. Defining h as in Eq. (4a), the 9th-order Holas function can be written as Eq. (4b)

$$x(1-x) = h \tag{4a}$$

$$g_5^{\mathrm{PH}}(x) = x + (2x-1)[h + 3h^2 + 10(h + 3.5h^2)h^2] \tag{4b}$$

The $g_5^{\mathrm{PH}}(x)$ in this form can be evaluated using four matmuls, the same number as in the SP scheme but with much smaller coefficients. Alternatively, one can use the symmetry property of $g_5(x)$ as in Holas [5] or Eqs. (5a) and (5b). We denote it as the Holas symmetric (HS) scheme in this Letter. The number of matmuls to evaluate $g_5^{\mathrm{HS}}(x)$ using this symmetry property is five. One can also combine the Holas Symmetric formula and the Paterson–Stockmeyer method in a hybrid fashion, denoted as $g_5^{\mathrm{HS/PS}}(x)$, where the PS scheme is used to evaluate the terms in square bracket in Eq. (5b) to reduce the number of matmuls to four.

$$\left(x - \frac{1}{2}\right) = D \tag{5a}$$

$$g_5^{\mathrm{HS/PS}} = \frac{1}{2} + D\left[\frac{315}{129} - \frac{105}{8}D^2 + \left(\frac{189}{c} - 90D^2 + 70(D^2)^2\right)(D^2)^2)\right] \tag{5b}$$

To investigate how numerical truncation errors evolve in practical calculations of molecular systems using these various polynomial evaluation methods and how they affect the total cost, we implemented these algorithms (PS, PH, HS, and HS/PS) in Q-CHEM [6] and performed the BLYP/STO-3G calculations for a $C_{60}H_{122}$ alkane chain. The results are summarized in Table 1. The convergence criterion is set by the maximum absolute element of $P_n - P_{n-1}$ to be smaller than 100 × thresh, where thresh = 10^{-8} or 10^{-11} was used for truncating matrix elements. Under this measure, the 5th-order function in a symmetric Holas form indeed shows a much faster convergence (lower cost) to the density matrix of desired accuracy than the 9th-order Holas polynomial with various evaluation methods. On the other hand, using the Paterson–Stockmeyer scheme, the error in density matrix never reaches below 100 × thresh, which is consistent with the large coefficients

Table 1
The number of matrix multiplications in the last SCF cycle for a linear alkane chain $C_{60}H_{122}$ at the BLYP/STO-3G level. The convergence criterion is set by the maximum absolute element of $\mathbf{P}_n - \mathbf{P}_{n-1}$ to be smaller than 100 × thresh, where thresh = 10^{-8} or 10^{-11} was used as a threshold for truncating small matrix elements.

Polynomial order	Multiplication method	Number of matmuls per iteration	Total number of matmuls (thresh = 10^{-8})	Total number of matmuls (thresh = 10^{-11})
3rd-order	McWeeny	2	28	28
5th-order	Holas symmetric	3	24	27
	Kim–Jung (optimized)[a]	3	22	22
9th-order	Holas symmetric (HS)	5	35	35
	Kim–Jung (KJ)	5	divergent	divergent
	Paterson–Stockmeyer (PS)	4	divergent	divergent
	Holas symmetric + Paterson–Stockmeyer (HS/PS)	4	28	28
	Particle–Hole (PH)	4	28	28

[a] Generalized optimized purification function is used [7].

in Eq. (3) of the order of 100. One can tighten the truncation threshold (i.e., less aggressive truncation) or/and loosen the convergence criterion of purification ($1000 \times$ thresh, for example) to make sure that the PS scheme is also converged, but any of the latter changes will undoubtedly increase the computational cost to achieve a given target accuracy.

Other 9th-order polynomial evaluation methods that reduce the magnitudes of the coefficients, such as Particle–Hole, Holas symmetric, and the hybrid of Holas symmetric/Paterson–Stockmeyer method, are numerically more stable and thus do converge with cost comparable to the 5th-order purification. Since the 9th-order HS and HS/PS methods still have relatively larger coefficients than the symmetric form of the 5th-order function, the converged density matrix using the 9th-order HS and HS/PS methods have larger errors than the 5th-order HS method. The fact that using the 5th-order function converges slightly faster (by 1–4 matmuls) than the 9th-order function is also related to the flexibility of using the 5th-order function (in multiples of three matmuls) than using the 9th-order function (in multiples of four matmuls).

In the above examples, we introduced the truncation after each matrix multiplication (truncation scheme I), but there is another way to truncate, namely, truncation only after each purification (truncation scheme II). As Rubensson offered some discussions in his Comment, truncation after each matrix multiplication (introducing more zeros) will save computational cost further than the truncation scheme II but at the expense of potentially larger truncation error which is also related to the numerical stability of polynomial evaluation methods. We therefore additionally experimented the truncation scheme II.

We calculated the absolute energy of $C_{60}H_{122}$ at BLYP/STO-3G using the 9th-order PS vs. 5th-order Holas methods and introducing truncation scheme II. For a given target accuracy (the same converged absolute energy), we find that the total number of matmuls required for the converged density matrix using the 9th-order PS method is 32 while that using the 5th-order Holas is 30. Incidentally, the total number of matmuls using the 5th-order Holas is 30 using both truncation schemes I and II. Therefore, the 5th-order Holas is more efficient than the 9th-order PS method for alkane chains under both truncation schemes I and II. We also counted the total number of actual floating point operations (FLOPs) between nonzero elements, and still found that the 5th-order Holas takes 1.2 times less FLOPs than the 9th-order PS using the truncation scheme II.

In summary, we have shown that the Paterson–Stockmeyer method of polynomial evaluation for the 9th-order Holas function is numerically significantly less stable than the 5th-order polynomial due to the large coefficients involved. We have confirmed in our numerical experiments for a real molecule that, when numerical truncation is introduced as is essential for linear scaling SCF calculations, the latter fact indeed leads to a higher computational cost for the 9th-order purification function to achieve a given target accuracy as compared to the 5th-order function. We note [7] that this 5th-order Holas purification function can also be generalized to nonpurifying functions to accelerate the convergence even further by removing the stability conditions in Holas [5] that the functions have fixed points and vanishing derivatives at 0 and 1.

Acknowledgements

We are pleased to acknowledge the support of Basic Science Research (2010-0023018) and WCU (R-31-2008-000-10055-0) programs funded by the Ministry of Education, Science, and Technology of Korea.

References

[1] J. Kim, Y. Jung, Chem. Phys. Lett. 511 (2011) 159.
[2] E.H. Rubensson, J. Chem. Theory Comput. 7 (2011) 1233.
[3] M.S. Paterson, L.J. Stockmeyer, SIAM J. Comput. 2 (1973) 60.
[4] D. Mazziotti, Phys. Rev. E 68 (2003) 1.
[5] A. Holas, Chem. Phys. Lett. 340 (2001) 552.
[6] Y. Shao et al., Phys. Chem. Chemical Phys.: PCCP 8 (2006) 3172.
[7] J. Kim, Y. Jung, J. Chem. Theory Comput. 7 (2011) 3853.