Random tempering of Gaussiantype geminals. II. Molecular systems

S. A. Alexander, H. J. Monkhorst, and K. Szalewicz

Random tempering of Gaussian-type geminals. II. Molecular systems

S. A. Alexander, H. J. Monkhorst, and K. Szalewicz

Quantum Theory Project, Departments of Physics and Chemistry, University of Florida, Gainesville, Florida 32611

We use random tempering formulas to create a basis set of explicitly correlated Gaussian-type geminals for the calculation of the second-order energy of lithium hydride. Like our earlier calculations on the beryllium atom, this technique matches results obtained by the full optimization of all nonlinear parameters but requires considerably less computational effort.

I. INTRODUCTION

In the first paper of this series[1] we showed that random tempering could be used to efficiently produce basis sets of Gaussian-type geminals (GTGs) for atomic systems. Using the beryllium atom as an example, we were able to reduce the computational costs of such calculations by a factor of 100 to 1000 over previous studies. Our algorithm was based on the work of Thakkar and Smith[2] and of Poshusta.[3] Random tempering is a method of shifting the burden of optimization from a large number of parameters to a small number of more sensitive tempering parameters. Since a nonlinear optimization can require at least N^2 function evaluations, where N is the number of nonlinear parameters, any reduction in the number of nonlinear parameters will lead to considerable savings in CPU time. In this paper we examine whether random tempering can also be used to speed up molecular calculations. This generalization, however, is not straightforward. In atomic calculations the centers of each geminal are fixed on the nucleus. In molecular calculations one has the choice of either placing the geminals on the various nuclei or allowing their centers to move. Both of these approaches can be approximated by various random tempered formulas and to test the advantages and disadvantages of each we use the second-order energy of lithium hydride.

In an earlier set of papers[4,5] the second-order energy of the LiH molecule was calculated using a basis set of explicitly correlated Gaussian-type geminals in which all nonlinear parameters (including the geminal positions) were optimized. The accuracy of these calculations was significantly better than any which had been previously reported. Later work, however, by Adamowicz et al.[6,7] produced slightly better results for some second-order pair energies of this molecule (see Table I). Using 82 optimized numerical orbitals, Adamowicz et al. obtained energies for the 12 and 21 pairs which were lower than the energies from GTG calculations. It is interesting that although the total energy of the $K = 40$ GTG calculation is much better than the $K = 20$ calculation both the 12 and 21 pairs are worse. The reason for the poor convergence of the GTG off-diagonal pair energies is partly due to the use of a large strong orthogonality parameter, $\eta = 100$, in the former optimization vs $\eta = 1$ in the latter one. Such a large value was used because of the belief, later disproved,[8] that it was needed for stable coupled pair calculations. Also, the same convergence threshold was used for the diagonal and off-diagonal pairs. Since the off-

diagonal pair energies are small and already quite well converged for $K = 20$, this may have caused the $K = 40$ optimization to complete prematurely. We have carefully reoptimized the $K = 40$ basis set using $\eta = 1$ and obtained substantial improvement in all pairs (see Table I).

II. RANDOM TEMPERING

Our algorithm for producing a basis set of random tempered geminals is a simple one:

Step (1). Choose a tempering formula for each type of nonlinear parameter.

Step (2). Optimize the tempering parameters using a relatively small number of geminals ($K = 50$ or 100), a small SCP basis ($M = 8$), and a small value of the strong orthogonality parameter ($\eta = 1$).

Step (3). Take these optimized parameters and calculate the second-order energy using $M = 20$, $\eta = 0$ and an increasing number of geminals until convergence is reached, the number of geminals becomes intractable or linear dependence occurs.

The choice of the random tempered formula depends to some extent on the system to be investigated. In atomic systems each GTG is a function of three unknown parameters

$$g_k(\mathbf{r}_1, \mathbf{r}_2) = e^{-\alpha_k r_1^2 - \beta_k r_2^2 - \gamma_k r_{12}^2}. \tag{1}$$

TABLE I. Comparison of second-order pair energies for LiH. An internuclear distance of 3.015 bohr is used. Values are in mhartrees with signs reversed.

Calculation	Pair 11	Pair[a] 12	Pair[a] 21	Pair 22	Total
$K = 20^{\text{b}}$	39.428	1.292	1.440	29.752	71.912
$K = 40^{\text{c}}$	39.548	1.273	1.424	29.938	72.183
Reference 6[d]	38.246	1.332	1.436	29.239	70.252
Reference 7[e]	38.567	1.335	1.447	29.528	70.877
$K = 40^{\text{f}}$	39.584	1.318	1.463	30.106	72.471

[a] "12" and "21" pairs are triplet and singlet coupled $1s-2s$ electron pairs, respectively.
[b] Reference 4. $M = 30$, $\eta = 0$. Full optimization with $M = 8$, $\eta = 0.5$.
[c] Reference 5. $M = 20$, $\eta = 0$. Full optimization with $M = 8$, $\eta = 100$.
[d] Seventy numerical orbitals used. Individual pair energies furnished by the authors of Ref. 6.
[e] Eighty-two numerical obitals used. Individual pair energies furnished by the authors of Ref. 7.
[f] This work. $M = 20$, $\eta = 0$. Full optimization with $M = 8$, $\eta = 1$.

TABLE II. Random tempered calculations of the 22 pair using geminal centers fixed on nuclei. An internuclear distance of 3.015 bohr is used. Values in mhartrees with signs reversed.

Method	Opt.[a]	$K = 50$ $M = 20$ $\eta = 0$	$K = 100$ $M = 20$ $\eta = 0$	$K = 200$ $M = 20$ $\eta = 0$	$K = 300$ $M = 20$ $\eta = 0$	$K = 400$ $M = 20$ $\eta = 0$	$K = 500$ $M = 20$ $\eta = 0$	$K = 600$ $M = 20$ $\eta = 0$	$K = 700$ $M = 20$ $\eta = 0$
1[b]	4.873	5.064	5.114	5.149	5.156	5.159	⋯	⋯	⋯
2[c]	27.723	28.005	28.145	28.234	28.240	28.242	⋯	⋯	⋯
3[d]	25.677	26.136	26.595	27.004	27.104	27.148	27.187	27.212	27.230
4[e]	29.185	29.277	29.517	29.660	29.743	29.782	29.818	29.834	29.836

[a] These optimizations were done with $K = 50$, $M = 8$, and $\eta = 1$.
[b] Geminal centers fixed on Li and Li.
[c] Geminal centers fixed on H and H.
[d] Geminal centers fixed on Li and H.
[e] Geminal centers fixed on Li + Li, H + H, and Li + H with the ratio of these structures optimized.

Previously we showed[1] that α_k, β_k, and γ_k could be accurately determined with six tempering parameters:

$$\alpha_k = e^{A_1 \langle k,j \rangle + A_2 \langle k,j+1 \rangle},$$
$$\beta_k = e^{B_1 \langle k,j+2 \rangle + B_2 \langle k,j+3 \rangle}, \tag{2}$$
$$\gamma_k = e^{C_1 \langle k,j+4 \rangle + C_2 \langle k,j+5 \rangle},$$

where $\langle k,j \rangle$ is the fractional part of $[k(k+1)/2 \, P(j)^{1/2}]$ and $P(j)$ is the jth prime number in the sequence 2,3,5,7... . In molecular systems each geminal is defined as

$$g_k(\mathbf{r}_1,\mathbf{r}_2) = e^{-\alpha_k r_{1A}^2 - \beta_k r_{2B}^2 - \gamma_k r_{12}^2}, \tag{3}$$

where

$$r_{1A}^2 = |\mathbf{r}_1 - \mathbf{A}_k|^2 \tag{4}$$

and

$$r_{2B}^2 = |\mathbf{r}_2 - \mathbf{B}_k|^2, \tag{5}$$

and where $\mathbf{A}_k$ and $\mathbf{B}_k$ are the two centers of each geminal. Here not only the parameters α_k, β_k, and γ_k must be determined but also $\mathbf{A}_k$ and $\mathbf{B}_k$. Because LiH is a diatomic, only those values which lie along the internuclear axis, A_{xk} and B_{xk}, are needed. In the next section we propose to use Eq. (2) to temper α_k, β_k, and γ_k and to investigate various ways of selecting A_{xk} and B_{xk}.

III. CALCULATIONS

Rather than testing all possible methods of tempering A_{xk} and B_{xk} on all four pairs we decided to concentrate initially on the 22 pair. Because this pair is the most diffuse we believe that it should be the most difficult one to optimize accurately. Any method of tempering which successfully determines this pair will then be applied to all of the other pairs.

One simple method of selecting A_{xk} and B_{xk} is to put them on the atomic centers. For LiH this leads to three possibilities: (1) Li + Li, (2) Li + H, (3) H + H. Although such basis sets are incomplete, they are of potential interest because they require only a relatively small number of nonlinear parameters. To find out how effective such schemes are, we examined a number of possibilities. As Table II shows, the Li + H and H + H calculations yield, respectively, an encouraging 92% and 88% of the full optimization value. To try and improve this result we performed an optimization which contained all three possibilities and which had the ability to find the optimal percentage of each combination. This calculation required 20 nonlinear parameters (6 for each possibility and 2 to optimize the percentages) and gave an energy which was 98% of the full optimization value when an expansion of 700 geminals was used. Although no sign of linear dependence was observed, the computational costs associated with such a large number of geminals is quite high. Clearly to match or improve the full optimization energy a more efficient method of tempering is needed.

One alternative to placing geminals on atoms is to allow the geminal centers to move. For one-dimensional molecules we propose the formula

TABLE III. Random tempered calculations on the 22 pair using moving geminal centers. An internuclear distance of 3.015 bohr is used. Values in mhartrees with signs reversed.

Method	Opt.[a]	$K = 50$ $M = 20$ $\eta = 0$	$K = 100$ $M = 20$ $\eta = 0$	$K = 200$ $M = 20$ $\eta = 0$	$K = 300$ $M = 20$ $\eta = 0$	$K = 400$ $M = 20$ $\eta = 0$	$K = 500$ $M = 20$ $\eta = 0$	$K = 600$ $M = 20$ $\eta = 0$	$K = 700$ $M = 20$ $\eta = 0$
1[b]	29.241	28.986	29.394	29.818	30.030	30.135	30.232	30.260	30.278
2[c]	29.194	29.199	29.658	29.994	30.121	30.160	30.202	30.220	30.245
3[d]	29.156	29.387	29.743	29.978	30.098	30.157	30.197	30.214	30.240
4[e]	29.918	29.767	30.038	30.231	30.324	30.361	30.382	30.390	30.396

[a] These optimizations were done with $K = 50$, $M = 8$, and $\eta = 1$.
[b] Geminals centered around Li and Li.
[c] Geminals centered around H and H.
[d] Geminals centered around Li and H.
[e] Geminals centered around Li + Li, H + H, and Li + H with the ratio of these structures optimized.

TABLE IV. Random tempered calculations of all pairs using moving geminal centers. An internuclear distance of 3.015 bohr is used. Values in mhartrees with signs reversed.

Pair	Opt.[a]	$K = 50$ $M = 20$ $\eta = 0$	$K = 100$ $M = 20$ $\eta = 0$	$K = 200$ $M = 20$ $\eta = 0$	$K = 300$ $M = 20$ $\eta = 0$	$K = 400$ $M = 20$ $\eta = 0$	$K = 500$ $M = 20$ $\eta = 0$	$K = 600$ $M = 20$ $\eta = 0$	$K = 700$ $M = 20$ $\eta = 0$
11	39.458	⋯	39.495	39.548	39.561	39.569	39.578	39.586	39.590
12[b]	1.252	1.246	1.265	1.289	1.303	1.310	1.317	1.321	1.324
21[b]	1.411	1.349	1.377	1.423	1.446	1.454	1.461	1.465	1.471
22	29.918	29.767	30.038	30.231	30.324	30.361	30.382	30.390	30.396
	72.039	⋯	72.175	72.491	72.634	72.694	72.738	72.762	72.781

[a] Each pair was optimized with $K = 50$, $M = 8$, and $\eta = 1$ except for the 11 pair which was done with $K = 100$, $M = 8$, and $\eta = 1$.
[b] See Table I, footnote a.

$$A_{xk} = A_0 + [X1\langle k,j+6\rangle + X2\langle k,j+7\rangle]/\alpha_k^{X3},$$

$$B_{xk} = B_0 + [Y1\langle k,j+8\rangle + Y2\langle k,j+9\rangle]/\beta_k^{Y3}. \qquad (6)$$

Here A_0 and B_0 are nuclear positions and $X1$, $X2$, $X3$, $Y1$, $Y2$, and $Y3$ are additional nonlinear parameters. This formula is based on the observation that fully optimized molecular geminals tend to be located somewhat closely to the nuclei and that this distance is inversely proportional to the exponent of the geminal, i.e., $A_{kx} \sim 1/\alpha_k$, etc. Because the values of the geminal positions can become quite large due to random fluctuations, we restrict all positions to lie within 2 bohr of either atom. This helps to reduce the occurrence of linear dependence. In Table III we present the results of this method in its simplest form. Placing the centers of the geminals around a single combination (Li + Li, Li + H, or H + H) requires 12 nonlinear parameters [6 for Eq. (2) and 6 for Eq. (6)]. Unlike the fixed geminal calculations each of these combinations gives roughly the same energy and the $K = 700$ expansion reproduces 99% of the fully optimized energy. As before we can test the most general version of this technique by combining all three possibilities and the ability to optimize their percentages (for a total of 38 nonlinear parameters). The energy from this calculation is better than even the fully optimized result.

As Table IV shows, this method also gives good results for the other pairs. Because of its highly localized nature, the 11 pair converged rather slowly but a reoptimization with 100 geminals solved this problem. The final second-order energy is better than the $K = 40$ full optimization by 0.3 mhartrees. This energy is substantially better (~ 2.0 mhartrees) than the values of Adamowicz et al. although their 12 pair is slightly lower (0.011 mhartrees) than ours. Although it is difficult to predict what limiting value for each pair energy should be, the pattern of convergence in Table IV suggests that our pair energies are accurate to about 0.01 mhartree so a difference of this order can be expected. By reoptimizing the 12 pair with 100 geminals we think we could match or exceed their value. A comparison of Tables I and IV also suggests that the method of Adamowicz et al.[6,7] converges nonuniformly since the off-diagonal pairs appear to be more converged than the diagonal ones. The final optimization parameters for each of our pairs are given in Table V.

IV. SORTING

Due to the random nature of our tempering algorithm, a large number of geminals are necessary to obtain a saturated second-order energy. This is because the optimization phase places only some of the geminals close to the physically important regions of the pair function. Other geminals are placed in unimportant regions which contribute little to the energy. Particularly when expanding the number of geminals, many unimportant types can be added. In Ref. 1 we showed that after the desired second-order energy has been obtained, the expanded set of random tempered geminals could be sorted according to their increasing contributions to the energy. This method produced relatively compact basis sets. Below we describe a more effective version of our earlier algorithm. To begin we first find the geminal which has the lowest individual second-order energy. Using this geminal we then find the next geminal whose combination with the first geminal gives the best second-order energy. This procedure continues, adding one geminal at a time to the last group obtained, until the best set of 75 geminals are found (roughly 10% of the total number of geminals). Next we calculate as before the pair energies for all subsets such that the L th subset contains the L first geminals of the full basis set for $L = 1,2,...K$. The geminals are then reordered such that those geminals which lower the energy most are moved to the top of the basis set. The whole process is then repeated several times (typically six). In subsequent steps, however, the sum of the increments of the previous steps is taken into account. We found that this combination of procedures gives a much better results than either one would by itself. The first part serves to filter the most important geminals to the top where the second part mixes them so as to group strongly coupled pairs of geminals. The results of this computation compare very favorably with the full optimization basis (see Table VI).

V. CONCLUSIONS

We have shown that random tempering can generate systematic sets of molecular geminals with a relatively small computational effort. We have also shown that placing the centers of the geminals only on nuclei leads at best to slow convergence. Much better results are obtained if the centers

TABLE V. Random tempered optimized parameters. The value in parentheses gives the value of the associated prime number in $\langle k,j \rangle$.

	Pair 11	Pair[a] 12	Pair[a] 21	Pair 22	
(Li + H combination) α and A_{xk} parameters					
A 1	1.299 88	− 1.546 30	− 2.173 35	− 0.610 10	(2)
A 2	6.523 88	− 2.569 11	− 1.959 06	− 2.796 29	(3)
X 1	1.367 83	0.370 90	1.234 48	− 0.253 30	(67)
X 2	− 0.331 73	0.681 30	2.306 12	3.821 21	(71)
X 3	31.278 54	0.608 31	0.229 59	0.032 73	
β and B_{xk} parameters					
B 1	0.250 17	− 0.977 24	− 0.516 00	− 4.000 62	(5)
B 2	− 0.881 51	1.428 46	− 0.047 12	1.285 16	(7)
Y 1	0.266 46	− 4.887 00	− 2.923 70	0.084 97	(73)
Y 2	− 3.456 49	− 0.116 78	− 0.557 77	− 0.160 91	(79)
Y 3	0.307 23	0.000 02	0.000 27	0.007 33	
γ parameters					
C 1	− 3.215 94	− 0.394 02	− 5.558 05	− 5.918 45	(11)
C 2	6.463 28	− 5.812 57	− 1.179 48	− 9.415 02	(13)
(Li + Li combination) α and A_{xk} parameters					
A 1	2.319 83	1.741 74	2.582 74	2.407 98	(17)
A 2	− 2.378 38	0.864 11	0.256 24	− 1.197 85	(19)
X 1	0.006 95	18.516 65	0.370 87	− 4.330 50	(83)
X 2	0.002 17	19.134 08	− 1.186 56	1.934 16	(89)
X 3	3.320 38	21.875 41	1.045 81	6.463 04	
β and B_{xk} parameters					
B 1	− 2.040 97	− 1.103 91	2.037 17	− 0.012 23	(23)
B 2	4.822 45	− 22.130 50	− 3.536 42	− 3.895 54	(29)
Y 1	0.131 21	− 8.659 04	− 2.953 54	− 5.890 50	(97)
Y 2	− 0.047 83	8.044 84	1.786 89	4.913 84	(101)
Y 3	3.132 75	0.000 34	0.000 04	0.000 01	
γ parameters					
C 1	− 15.677 73	− 26.510 33	26.727 66	25.825 31	(31)
C 2	13.101 72	− 33.006 65	− 31.614 36	− 49.774 83	(37)
(H + H combination)[b] α and A_{xk} parameters					
A 1	⋯	− 3.922 88	− 3.620 33	− 8.986 20	(41)
A 2	⋯	1.321 91	2.011 86	0.825 71	(43)
X 1	⋯	0.347 48	0.699 87	− 0.069 86	(103)
X 2	⋯	− 3.216 11	− 3.453 07	− 0.041 56	(107)
X 3	⋯	1.793 37	2.442 42	1.028 10	
β and B_{xk} parameters					
B 1	⋯	− 4.080 08	− 5.572 21	− 5.395 56	(47)
B 2	⋯	1.676 71	1.591 10	3.228 63	(53)
Y 1	⋯	0.722 45	0.521 49	− 0.037 19	(109)
Y 2	⋯	− 0.165 45	− 1.619 57	− 0.062 81	(113)
Y 3	⋯	0.797 43	0.947 16	1.193 73	
γ parameters					
C 1	⋯	3.019 53	2.399 57	4.314 57	(59)
C 2	⋯	− 4.376 77	− 1.577 67	− 3.884 94	(61)
Fraction of Li + H					
	0.177 50	0.687 72	0.726 42	0.558 84	
Fraction of Li + Li					
	0.822 50	0.183 65	0.090 49	0.215 49	

[a] See Table I, footnote a.
[b] The 11 pair has the fraction of H + H set to zero by the optimization so the parameters of this type have been omitted.

TABLE VI. Comparison of sorted second-order pair energies for LiH with fully optimized values. An internuclear distance of 3.015 bohr is used. Values in mhartrees with signs reversed.

K	Pair 11	Pair[a] 12	Pair[a] 21	Pair 22	Total
Random tempered optimized + sorting[b]					
10	38.208	1.080	1.207	27.936	68.431
20	39.201	1.242	1.329	29.526	71.298
30	39.391	1.275	1.384	29.963	72.013
40	39.498	1.286	1.411	30.104	72.299
50	39.520	1.296	1.425	30.187	72.428
60	39.534	1.300	1.438	30.239	72.511
70	39.547	1.305	1.444	30.279	72.575
80	39.553	1.309	1.448	30.304	72.614
90	39.558	1.311	1.451	30.318	72.638
100	39.561	1.312	1.454	30.329	72.656
110	39.564	1.314	1.456	30.340	72.674
120	39.568	1.315	1.458	30.347	72.688
130	39.569	1.316	1.459	30.353	72.697
140	39.572	1.317	1.460	30.360	72.709
150	39.575	1.317	1.461	30.365	72.718
200	39.581	1.320	1.466	30.376	72.743
300	39.587	1.323	1.469	30.388	72.767
400	39.589	1.324	1.471	30.393	72.777
500	39.590	1.324	1.471	30.395	72.780
600	39.590	1.324	1.471	30.396	72.781
700	39.590	1.324	1.471	30.396	72.781
Fully optimized[c]					
20	39.428	1.292	1.440	29.752	71.912
40	39.584	1.318	1.463	30.106	72.471

[a] See Table I, footnote a.
[b] These pairs were optimized with $K = 50$, $M = 8$, and $\eta = 1$ (except for the 11 pair which was optimized with $K = 100$, $M = 8$, and $\eta = 1$) then expanded to $K = 700$, $M = 20$, and $\eta = 0$.
[c] The $K = 20$ calculation is described in Ref. 4 and the $K = 40$ calculation is described in this work.

of the geminal are allowed to move. Like our earlier calculations on the beryllium atom those geminals which contribute most significantly to the energy can be efficiently sorted from the others. Although it has been successfully applied to lithium hydride, the straightforward application of our tempering algorithm to general polyatomic systems may become too cumbersome. As the number of nuclei increase, the number of nonlinear parameters will grow at an even faster rate. Should this become a problem, a simple version of our algorithm (e.g., our 12 parameter scheme) could still be used to provide a good initial guess for a full optimization. As the entries in Tables II and III show, a relatively small number of tempering parameters can accurately estimate the important regions of a pair function. We believe, however, that future improvements to the tempering process itself can increase the efficiency of this method by placing more points in the important regions of the pair function and can reduce the number of nonlinear parameters required for molecular calculations. Work on this project is in progress.

ACNOWLEDGMENTS

[1]S. A. Alexander, H. J. Monkhorst, and K. Szalewicz, J. Chem. Phys. 85, 5821 (1986).
[2]A. J. Thakkar and V. H. Smith, Jr., Phys. Rev. A 15,1,16, 2143 (1977).
[3]R. D. Poshusta, Int. J. Quantum Chem. Symp. 13, 59 (1979).
[4]K. Szalewicz, B. Jeziorski, H. J. Monkhorst, and J. G. Zabolitzky, J. Chem. Phys. 78, 1420 (1983).
[5]K. Szalewicz, B. Jeziorski, H. J. Monkhorst, and J. G. Zabolitzky, J. Chem. Phys. 79, 5543 (1983).
[6]L. Adamowicz, R. J. Bartlett, and E. A. McCullough, Jr., Phys. Rev. Lett. 54, 426 (1985)
[7]L. Adamowicz and R. J. Bartlett, in Applied Quantum Chemistry, edited by V. H. Smith et al. (Reidel, New York, 1986), p. 111.
[8]K. Szalewicz, J. G. Zabolitzky, B. Jeziorski, and H. J. Monkhorst, J. Chem. Phys. 81, 2723 (1984).