Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials

Siu A. Chin \textsuperscript{a)1} and C. R. Chen

Center for Theoretical Physics, Department of Physics, Texas A&M University, College Station, Texas 77843

(Received 6 March 2002; accepted 24 April 2002)

We show that the method of factorizing the evolution operator to fourth order with purely positive coefficients, in conjunction with Suzuki’s method of implementing time-ordering of operators, produces a new class of powerful algorithms for solving the Schrödinger equation with time-dependent potentials. When applied to the Walker–Preston model of a diatomic molecule in a strong laser field, these algorithms can have fourth order error coefficients that are three orders of magnitude smaller than the Forest–Ruth algorithm using the same number of fast Fourier transforms. Compared to the second order split-operator method, some of these algorithms can achieve comparable convergent accuracy at step sizes 50 times as large. Moreover, we show that these algorithms belong to a one-parameter family of algorithms, and that the parameter can be further optimized for specific applications. © 2002 American Institute of Physics.

[DOI: 10.1063/1.1485725]

I. INTRODUCTION

Recently, we derived a new class of fourth order algorithms for solving both classical\textsuperscript{1,2} and quantum dynamical\textsuperscript{3} problems. These algorithms are based on factorizing the evolution operator $e^{(T+V)}$ to fourth order with purely positive coefficients and require knowing the gradient of the force in the classical case and the gradient of the potential in the quantum case. The resulting algorithms are symplectic or unitary, respectively. While positive coefficients are absolutely necessary for simulating the diffusion process in Monte Carlo algorithms,\textsuperscript{4,5} or doing imaginary time projections,\textsuperscript{6} they are not essential in quantum or classical algorithms. Nevertheless, we have shown that this class of gradient symplectic algorithms is far superior to existing fourth order algorithms with negative coefficients.\textsuperscript{1–3} In this work, using Suzuki’s method\textsuperscript{7} of implementing operator time ordering, we produce a class of even more effective algorithms for solving quantum dynamical problems with explicit time-dependent potentials. Despite the vast amount of literature on this subject,\textsuperscript{8–18} we believe our work has initiated a new direction in algorithm development. In the past, one has labored assiduously to avoid higher order commutators. Here, we show that their inclusion can yield algorithms of great efficiency. Our algorithm 4A, discussed below, is the fastest fourth order algorithm known, needing only four fast Fourier transforms (FFTs) per iteration. This is only twice the computational effort of the second order split-operator method,\textsuperscript{8,14,18} but the algorithm can converge at time step sizes an order of magnitude larger. Our optimized algorithms, compared on an equal effort basis, have fourth order error coefficients that are three orders of magnitude smaller than the Forest–Ruth algorithm\textsuperscript{19} and are a factor of 30 smaller than McLachlan’s algorithm;\textsuperscript{20} both are fourth order algorithms with negative coefficients.

While these gradient symplectic algorithms are very efficient when the potential gradient is known analytically, they remain equally effective when the gradient is obtained numerically.\textsuperscript{6} These algorithms are of particular interest in solving the time-dependent Schrödinger equations in a large three-dimensional (3D) mesh. In 3D, even for a modest grid size of $(256)^3$, the number of mesh points already exceed 16 million. If the wave function array is double precision and complex, its storage alone would have required 268 MB. For such a large number of grid points, any vector–matrix multiplication would be prohibitively expensive and must be avoided. For our algorithms, the most costly computational step is just the use of FFT. The unitary character and the large step acceptance of these algorithms make them ideal for doing long time quantum simulations.

The key problem in solving the Schrödinger equation with time-dependent potentials is the time ordering of operators. This problem is solved in various ways in the literature by transforming it into a classical problem,\textsuperscript{14,15} treating time as another “spatial coordinate,”\textsuperscript{12,13,16} introducing auxiliary variables,\textsuperscript{17} etc. Most end up with some sort of time derivative operator, but none has the simplicity of Suzuki’s method\textsuperscript{7} of directly implementing time ordering via a forward time derivative operator. No time integration is necessary. Since this work is less accessible, but is of special relevance to the operator factorization approach of deriving algorithms, we summarize it in some detail in Sec. II. In Sec. III, we apply Suzuki’s method and derive four gradient sym-
plectic algorithms for solving the time-dependent Schrödinger equation. In Sec. IV, we use these algorithms to solve the Walker–Preston model\(^{21}\) and compare their convergent properties with existing algorithms. In Sec. V, we derive one-parameter families of these algorithms and show that they can be further optimized for specific applications. Section VI summarizes our conclusions.

### II. OPERATOR DECOMPOSITION OF ORDERED EXPONENTIALS

For \( H(t) \) a time-dependent operator, the evolution equation,

\[
\frac{\partial}{\partial t} \psi(t) = H(t) \psi(t),
\]

has the operator solution

\[
\psi(t + \Delta t) = T \exp \left( \int_t^{t+\Delta t} H(s) \, ds \right) \psi(t).
\]

The time-ordered exponential not only has the conventional expansion,

\[
T \exp \left( \int_t^{t+\Delta t} H(s) \, ds \right) = 1 + \int_t^{t+\Delta t} H(s_1) \, ds_1 + \int_t^{t+\Delta t} \int_s^{t+\Delta t} H(s_1) H(s_2) \, ds_1 \, ds_2 + \cdots,
\]

but also the more intuitive interpretation

\[
T \exp \left( \int_t^{t+\Delta t} H(s) \, ds \right) = \lim_{n \to \infty} T \left( e^{(\Delta t/n) H(t + (n+1) \Delta t/n)} \right)^n,
\]

\[
= \lim_{n \to \infty} e^{(\Delta t/n) H(t + \Delta t)} e^{(\Delta t/n) H(t + 2 \Delta t/n)} \cdots e^{(\Delta t/n) H(t + n \Delta t/n)} \times e^{(\Delta t/n) H(t + \Delta t/n)}.
\]

There are many ways of solving the time-ordering problem. For this work on operator factorization, we prefer Suzuki’s method,\(^7\) which directly implements time ordering without any additional formalism\(^{12}\) or auxiliary variables.\(^{17}\)

Let \( D \) denote the forward time derivative operator,

\[
D = \frac{\partial}{\partial t},
\]

such that for any two time-dependent functions \( F(t) \) and \( G(t) \),

\[
F(t) e^{\Delta t D} G(t) = F(t + \Delta t) G(t).
\]

Suzuki’s method\(^7\) proved that

\[
T \exp \left( \int_t^{t+\Delta t} H(s) \, ds \right) = \exp \left[ \Delta t (H(t) + D) \right].
\]

Using the more intuitive definition of time ordering (4), and invoking Trotter’s formula, one proof only requires two lines:

\[
\exp \Delta t [H(t) + D] = \lim_{n \to \infty} (e^{(\Delta t/n) H(t)} e^{(\Delta t/n) D})^n,
\]

\[
= \lim_{n \to \infty} e^{(\Delta t/n) H(t + \Delta t)} \cdots e^{(\Delta t/n) H(t + 2 \Delta t/n)} \times e^{(\Delta t/n) H(t + \Delta t/n)},
\]

where property (6) has been applied repeatedly and accumulatively.

For the widely applicable case of \( H(t) = T + V(t) \), where only one of the operators is explicitly dependent on time, the short time evolution of (2) can be written using (7) as

\[
\psi(t + \Delta t) = e^{\Delta t [\tilde{T} + V(t)]} \psi(t),
\]

which is just like the time-independent case but with an effective \( \tilde{T} = T + D \). This suggests a two-step approach for deriving time-dependent algorithms. First, decompose \( e^{\Delta t [\tilde{T} + V(t)]} \) terms of \( e^{\Delta t \tilde{T}} \) and \( e^{\Delta t V(t)} \) using any factorization scheme applicable in the time-independent case. Next, since \([T, D] = 0\), factorize exactly

\[
e^{\Delta t \tilde{T}} = e^{\Delta t D} e^{\Delta t T},
\]

incorporate all time-dependent requirements by applying (6). For example, a second order factorization of (9) gives

\[
T^{(2)} = e^{\frac{1}{2} \Delta t \tilde{T}} e^{\Delta t V(t)} e^{\frac{1}{2} \Delta t \tilde{T}}
\]

\[
= e^{\frac{1}{2} \Delta t D} e^{\frac{1}{2} \Delta t T} e^{\Delta t V(t)} e^{\frac{1}{2} \Delta t D} e^{\frac{1}{2} \Delta t T}
\]

\[
= e^{\frac{1}{2} \Delta t D} e^{\Delta t V(t + \Delta t/2)} e^{\frac{1}{2} \Delta t T},
\]

which is the well known midpoint algorithm for time-dependent problems. The other second order factorization gives the alternative second order algorithm,

\[
T^{(2)} = e^{\frac{1}{2} \Delta t V(t)} e^{\Delta t \tilde{T}} e^{\frac{1}{2} \Delta t V(t)}
\]

\[
= e^{\frac{1}{2} \Delta t D} e^{\Delta t T} e^{\frac{1}{2} \Delta t V(t)}
\]

\[
= e^{\frac{1}{2} \Delta t V(t + \Delta t/2)} e^{\Delta t V(t)}.
\]

Thus, for \( H(t) = T + V(t) \), the effect of time ordering is to increment the time dependence of each potential operator \( V(t) \) by the sum of the time steps of all the \( T \) operators to its right.

For the Schrödinger equation with a time-dependent potential, the wave function evolves forward in a short time \( \Delta t \) by

\[
\psi(\epsilon) = e^{\epsilon [\tilde{T} + V(t)]} \psi(0),
\]

where \( \epsilon = -i \Delta t, \tilde{T} = \tilde{D} + T, \tilde{D} = iD \),

\[
T = -\frac{1}{2\mu} \sum_i \frac{\partial^2}{\partial x_i^2}, \quad \text{and} \quad V(t) = V(x_i, t).
\]
where \(i\). (In this way, algorithms can be directly applied to the classical case with \(\epsilon\) purely real without changes in form.) For conciseness, we will always regard the present time step as time zero. Thus, the two second order algorithms for solving the Schrödinger equation with step size \(\Delta t\) can be denoted simply as

\[
T_{A}^{(2)}(\epsilon) = e^{\frac{1}{2}\epsilon T}e^{\frac{1}{2}\epsilon^* T},
\]

\[
T_{B}^{(2)}(\epsilon) = e^{\frac{1}{2}\epsilon^* V(\epsilon)}e^{\frac{1}{2}\epsilon V(\epsilon)}. \quad \text{(15)}
\]

Since the kinetic energy operator is diagonal in momentum space, the operator \(e^{\epsilon T}\) can be implemented as vector-vector multiplication in Fourier space. Every occurrence of \(e^{\epsilon T}\) requires two FFTs, one direct to Fourier space for kinetic energy multiplication, and one inverse back to real space for potential energy multiplication. To minimize the call for FFTs, one favors algorithms with the fewest occurrences of the kinetic energy operator.

### III. GRADIENT SYMPLECTIC FOURTH ORDER ALGORITHMS

Following our two-step approach, we can transcribe any time-independent factorization algorithm into a time-dependent algorithm. For example, the well known Forest–Ruth (FR) algorithm\(^9\) (also discovered independently by Campostrini and Rossi\(^22\) and by Candy and Rozmus\(^23\)) can now be transcribed to solve time-dependent problems as

\[
T_{FR}^{(4)}(\epsilon) = e^{s\epsilon^* V(\epsilon)}e^{s\epsilon T}e^{s^2 T}e^{s^2 V(\epsilon)}e^{s^4 T}e^{s^4 V(\epsilon)}e^{s^6 T}e^{s^6 V(\epsilon)}, \quad \text{where} \quad s = 2^{1/3},
\]

\[
u_0 = v_3 = \frac{1}{2} - s, \quad v_1 = v_2 = -\frac{1}{2} - s, \quad t_1 = t_3 = \frac{1}{2} - s, \quad t_2 = -\frac{s}{2 - s}, \quad \text{and} \quad a_1 = t_1, \quad a_2 = t_1 + t_2. \quad \text{(17)}
\]

For easy identification, we adopt the convention of labeling the time step coefficients of operators \(T\) and \(V\) by \(t_i\) and \(v_i\), respectively, and denote the intermediate time arguments of \(V\) by coefficients \(a_i\). The coefficient of the first operator on the right will be denoted by \(v_0\) or \(t_0\), followed by paired coefficients of \(v_i\) and \(t_i\) for \(i = 1, 2, \ldots, n\). If one were to decompose \(e^{\epsilon(T + V)}\) only in terms of \(e^{\epsilon T}\) and \(e^{\epsilon V}\), then Forest–Ruth algorithm, is the only fourth order algorithm possible with six FFTs per iteration. An alternative algorithm with \(T\) and \(V\) interchanged, with appropriate modifications of the \(a_i\) coefficients, is also possible, but it would require eight FFTs. Notice that some of the coefficients are negative, which requires backward propagation and evaluating the potential sometime prior to the present. This is a consequence of Suzuki’s no-go theorem,\(^24\) which proved that, beyond second order, \(e^{\epsilon(T + V)}\) cannot be decomposed into finite products of \(e^{\epsilon T}\) and \(e^{\epsilon V}\) with purely positive coefficients \(t_i\) and \(v_i\). Thus without exception, all higher order factorization algorithms heretofore proposed in the literature\(^9\)–\(^11\),\(^15\)–\(^18\) contain negative coefficients.

It is the search for positive coefficient factorizations schemes\(^25\) that led one of us to derive fourth order symplectic algorithms for solving classical\(^1\) and subsequently quantum dynamical\(^3\) problems. To circumvent Suzuki’s no-go theorem, these factorization schemes employ an additional operator,

\[
[V_i(T,V)] = \frac{1}{\mu} \sum_{\tau} \left( \frac{\partial V_i}{\partial x_\tau} \right)^2,
\]

which is just the square of the gradient of the potential.\(^25\) This can be computed analytically or numerically. For brevity, we will transcribe four previously derived gradient algorithms\(^1\) to their time-dependent form. In Sec. IV, we will describe in more detail the one-parameter family of these algorithms.

Our algorithm \(4A\)\(^3\) (see also Ref. 26), when applied to the time-dependent case of (13), gives

\[
T_{A}^{(4)}(\epsilon) = e^{s^2 V(\epsilon)}e^{s^4 T}e^{s^4 V(\epsilon)}e^{s^6 T}e^{s^6 V(\epsilon)}e^{s^8 T}e^{s^8 V(\epsilon)}e^{s^10 V(\epsilon)}, \quad \text{with} \quad \tilde{V} \text{ defined by}
\]

\[
\tilde{V}(t) = V(t) + \frac{1}{\Delta t} \epsilon^2 [V(t),[T,T,V(t)]] = V(t) + \frac{1}{\Delta t} \epsilon^2 [V(t),[T,V(t)]]. \quad \text{(20)}
\]

Note that this is a crucial simplification, because \([V(t),[D,V(t)]] = 0\!\!10\). Thus the addition of the forward time derivative operator \(D\) causes no additional complication to the gradient term. (This is very fortunate, because if one were required to keep the alternative double commutator \([T,[V(t),T]]\), this commutator would have given rise to complicated new terms that involve \(\partial \tilde{V}(t)/\partial t\) and \(\partial^2 \tilde{V}(t)/\partial t^2\)\!) With the introduction of the gradient term, algorithm \(4A\) can achieve fourth order accuracy with only four FFTs. Note also that in order to minimize evaluation of the gradient term, the double commutator has been placed at the center. This choice seems obvious, but there is intrinsic freedom to redistribute the commutator term among the three potential operators without affecting the fourth order convergence of the algorithm. This is true for all algorithms described below. We will assume that this redistribution can be done when necessary.

Similarly algorithm \(4B\) can be transcribed as

\[
T_{B}^{(4)}(\epsilon) = e^{s^2 V(\epsilon)}e^{s^4 T}e^{s^4 V(\epsilon)}e^{s^6 T}e^{s^6 V(\epsilon)}e^{s^8 T}e^{s^8 V(\epsilon)}e^{s^10 V(\epsilon)}, \quad \text{(22)}
\]

where

\[
t_0 = t_2 = \frac{1}{2} \left( 1 - \frac{1}{\sqrt{3}} \right), \quad t_1 = \frac{1}{\sqrt{3}}, \quad a_1 = \frac{1}{2} \left( 1 - \frac{1}{\sqrt{3}} \right), \quad a_2 = \frac{1}{2} \left( 1 + \frac{1}{\sqrt{3}} \right), \quad \text{and with} \quad \tilde{V} \text{ given by}
\]

\[
\tilde{V}(t) = V(t) + \frac{1}{\Delta t} (2 - \sqrt{3}) \epsilon^2 [V(t),[T,V(t)]] = V(t) + \frac{1}{\Delta t} \epsilon^2 [V(t),[T,V(t)]]. \quad \text{(24)}
\]

The time-dependent forms of algorithms \(4C\) and \(4D\) are, respectively.
where lem has been used by many authors\textsuperscript{13,14,17,18} to test their algorithm, diatomic molecule in a strong laser field. Since this problem is \textsuperscript{13} two-dimensional Hamiltonian, we developed new time-dependent algorithms. The model is defined by the one-dimensional Hamiltonian (in atomic units),

\[ H = -\frac{1}{2\mu} \frac{\partial^2}{\partial x^2} + V(x,t), \]

\[ V(x,t) = V_0(1 - e^{-\alpha t})^2 + Ax \cos(\omega t), \]

where \( \mu = 1745, \) \( V_0 = 0.2251, \) \( \alpha = 1.1741, \) \( A = 0.011025, \) and \( \omega = 0.01787. \) The wave function is initially chosen to be the Morse oscillator ground state, which has corresponding ground state energy \( E_0 = (\omega_0/2)[1 - \omega_0/8V_0] \textsuperscript{19} \) with \( \omega_0 = \alpha \sqrt{2V_0/\mu}. \) In conformity with the above authors, we discretize the wave function \( \psi(x,t) \) using 64 uniform points at \( x_k = -0.8 + k\Delta x, \) with \( \Delta x = 0.08. \) The natural time scale defined by the laser oscillation frequency is \( \tau = 2\pi/\omega = 351.6. \)

In Fig. 1, the percentage errors of the total energy, \( 100|E(1000\tau) - E_{\text{cont}}/E_{\text{cont}}, \) is plotted as a function of the step size \( \Delta t \) used in the calculation. \( E_{\text{cont}} = 5.029155E_0 \) is the converged value at small step sizes. The symbols plotted indicate calculated results. The lines are monomial fits to either \( (\Delta t)^2 \) or \( (\Delta t)^4. \) \( SO \) is our gradient symplectic algorithm using the second order split-operator algorithm (16). RS3 is Gray and Verosky’s best convergent results.\textsuperscript{14} They used a third order algorithm, but the error is degraded by the Magnus approximation to second order. Both can be fitted well by a single quadratic \( b_3(\Delta t)^2, \) with coefficients \( b_{SO} = 1.1 \) and \( b_{RS3} = 0.67. \) These are plotted as lines running through the data. The convergent range for both is below \( \Delta t < \tau/100 = 3.52. \) Algorithm RS3 requires three times as much effort as SO. For the same amount of computational effort, one can run SO three times at \( \Delta t/3 \) and gain a reduction of \( (1/3)^2 \) in its error coefficient. This projected convergence curve \( b_{SO}(\tau/3)(\Delta t)^2 \) is plotted as a broken line and labeled SO’. For the same effort (six FFTs) one can also run the Forest–Ruth algorithm and obtain results, shown by closed triangles, with a fitted line \( 9.7 \times 10^{-4}(\Delta t)^4. \) This comparison of equal effort clearly demonstrates the greater efficiency of fourth order algorithms. For errors in the range of 0.1%–0.01%, the FRs step size can be four to six times as large as that of SO’. This ratio would increase if greater accuracy is required. Algorithm 4A, which requires only 2/3 the effort of FR, can be fitted by \( 4.8 \times 10^{-5}(\Delta t)^4. \) This fit is plotted as a dotted line, barely visible above the zero error line over the range of the plot.

To discuss the convergence of our gradient symplectic algorithms, we greatly expanded the plotting range in Fig. 2. Here, we plot directly the convergence of \( E(1000\tau)/E_0, \) the symbols plotted indicate calculated results. We retained SO and FR for comparison. The SO result is fitted well by \( E_{\text{cont}}/E_0 + 0.054(\Delta t)^2. \) Algorithms 4C and 4D yielded indiscernible fits.

### Table I. Equal computational effort comparison of all fourth order algorithms discussed in this work. All algorithms except Forest–Ruth (FR) and McLachlan (M) are our new time-dependent algorithms. \( N_i \) is the number of FFTs per execution for each algorithm, \( d_i \) is the corresponding fourth order error coefficient, \( \delta_{eq} \) is the equal-effort fourth order error coefficient normalized to the FR value, and \( \tau_{eq} \) is the effective-time step size relative to the FR value, e.g., for the same amount of effort, algorithm 4ACB can achieve the same convergence error as FR at a time step 8.5 times as large.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>N(_i)</th>
<th>M</th>
<th>4A</th>
<th>4B</th>
<th>4C</th>
<th>4D</th>
<th>4ACB</th>
<th>4BDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>(d_i)</td>
<td>5.0 \times 10^{-5}</td>
<td>1.3 \times 10^{-7}</td>
<td>2.4 \times 10^{-7}</td>
<td>0.5 \times 10^{-7}</td>
<td>1.0 \times 10^{-7}</td>
<td>1.0 \times 10^{-7}</td>
<td>3.0 \times 10^{-9}</td>
<td>9.5 \times 10^{-9}</td>
</tr>
<tr>
<td>(\delta_{eq})</td>
<td>1</td>
<td>8.2 \times 10^{-3}</td>
<td>0.95 \times 10^{-3}</td>
<td>1.0 \times 10^{-3}</td>
<td>6.3 \times 10^{-3}</td>
<td>2.0 \times 10^{-3}</td>
<td>1.9 \times 10^{-4}</td>
<td>1.9 \times 10^{-4}</td>
</tr>
<tr>
<td>(\tau_{eq})</td>
<td>1</td>
<td>3.3</td>
<td>5.7</td>
<td>5.6</td>
<td>3.5</td>
<td>3.5</td>
<td>4.7</td>
<td>8.5</td>
</tr>
</tbody>
</table>

![FIG. 1. Percentage of error at \( t = 1000\tau \) for various algorithms in solving the Walker–Preston model. The plotted symbols are results of calculations. The lines are second or fourth order fits. SO is the split-operator algorithm (16), SR3 is one of Gray and Verosky’s algorithms, FR is the Forest–Ruth algorithm (17), and 4A is the gradient algorithm (20). SO’ is an equal computational effort version of SO for comparing with SR3 and FR. See the text for details.](image-url)
tistinguishable results, despite the fact that 4D uses only six FFTs, two less than 4C. Obviously one should not use algo-
rithm 4C in the present case; we only included it here for completeness. Since the FR algorithm is known to have
rather large errors, we have also implemented McLachlan’s fourth order algorithm\(^{20}\) and obtained results, labeled M.
This algorithm requires eight FFTs per iteration and is the best algorithm tested by Sanz-Serna and Portillo.\(^{17}\) Both FR
and M are examples of fourth order algorithms with negative coefficients. The step size convergence of all fourth order
algorithms can be very fitted well by \(E_{\text{con}}/E_0 + d_i (\Delta t)^4\). These are plotted as lines going through data points. The
coefficient \(d_i\) for each algorithm is listed in Table I. Since the computational effort per time step is proportional to the
number of FFTs, \(N_i\), the error per unit effort, taking into account variation of the step size with effort, would be propor-
tional to \(d_i/N_i^4\), i.e., this is a measure of error of equal computational effort for all fourth order algorithms. Again,
for example, algorithm 4A only requires half the number of FFTs as McLachlan’s algorithm. At a given \(\Delta t\) in run-
ing McLachlan’s algorithm, one can execute algorithm 4A twice at \(\Delta t/2\) and reduce its error by a factor of \((1/2)^4\). Thus
de spite the appearance it has in Fig. 2, algorithm 4A actually has a much smaller error per unit effort than McLachlan’s
algorithm. We normalize this equal effort error to the value of FR and define the normalized, equal effort error coeffi-
cient as \(\delta_{\text{eq}} = d_i/d_{\text{FR}} (N_i/N_{\text{FR}})^4\). This value for each algo-
rithm is listed in Table I. In this equal effort comparasion, despite their different appearances in Fig. 2, algorithms 4A
and 4B are similar in efficiency. Thus, excluding 4C, our gradient algorithms are roughly a factor of 10\(^3\) smaller in
error than FR and a factor of 3–8 smaller than McLachlan’s algorithm. An even more useful measure is to consider \(\tau_{\text{eff}}
= \delta_{\text{eq}}^{-1/4}\), which would give the time step size relative to FR for achieving the same error with equal effort. This is also
given in Table I. From examining Figs. 1 and 2, it certainly seems reasonable that our algorithms can converge at time
steps nearly five times as large as those of FR and on the order of 30 times as large as those of SO\(^{'}\). We will discuss
further optimized algorithms 4ACB and 4BDA in Sec. V.

V. ONE-PARAMETER FAMILY OF ALGORITHMS

Algorithms 4A and 4C are special cases of the more general algorithm,
\[
T^{(4)}_{\text{ACB}}(\epsilon) = e^{v_3^2 T} e^{V(\epsilon)} e^{v_2^2 T} e^{V(\epsilon)} e^{v_1^2 T} e^{V(\epsilon)} e^{v_0 T} e^{v(0)}.
\]
(29)

By use of a MATHEMATICA program\(^{4,5}\) to symbolically com-
bine exponentials of operators, it is easy to determine these positive coefficients and obtain a fourth order algorithm. For
the above form of the algorithm, there is one free parameter, which we will take to be \(t_0\). Given \(t_0\), the rest of the coef-
cients are
\[
t_1 = t_2 = \frac{1}{2} - t_0, \quad t_3 = t_0, \quad v_1 = v_3 = \frac{1}{6} \left[ \frac{1}{(1 - 2t_0)^2} \right],
\]
(30)
\[
v_2 = 1 - (v_1 + v_3),
\]
where \(\bar{V}\) here is given by
\[
\bar{V}(t) = V(t) + \frac{u_0}{v_2} \epsilon^2 [V(t), [T, V(t)]]
\]
(31)
with \(a_1 = t_0, \quad a_2 = \frac{2}{3}, \quad a_3 = 1 - t_0, \quad \text{and}
\]
\[
u_0 = \frac{1}{12} \left[ \frac{1 - 1 - 2t_0}{1 - 1 - 2t_0} + \frac{1}{6(1 - 2t_0)^2} \right].
\]
(32)

For \(t_0 = 0\), one has algorithm 4A. For \(t_0 = \frac{1}{2} \approx 0.167\), one recovers algorithm 4C. Thus one can change continuously from algorithm 4A to algorithm 4C, and beyond, by “dial-
ing” the parameter \(0 \leq t_0 \leq \frac{1}{2}(1 - 1/\sqrt{3}) = 0.21\). At the upper limit of \(t_0 \approx \frac{1}{2}(1 - 1/\sqrt{3})\), \(v_3 = 0\), and the algorithm becomes a variant of algorithm 4B, where the commutator term re-
ained at the center rather than being equally distributed between the two potential operators at positions \(v_1\) and \(v_3\).
We shall refer to this algorithm as 4B’\(^{'}\). Algorithm 4B’\(^{'}\) can be continuously transformed to 4B by redistributing the commuta-
tor term from the center to both sides. For example, one can multiply the central commutator term by a factor of
\(1 - \alpha\) and add \(\alpha t_0\) times the commutator term to each poten-
tial operator on the side. As \(\alpha\) ranges from 0 to 1, algorithm 4B’\(^{'}\) is continuously transformed into 4B. When \(\alpha\) reaches 1,
the central commutator disappears and the number of FFTs collapses from eight to six. Thus both algorithms 4A and 4B
are singular end points of this general algorithm with discon-
tinuous changes in the number of FFTs.

Algorithms 4B and 4D are special cases of the general algorithm,
\[
T^{(4)}_{\text{BDA}}(\epsilon) = e^{v_3^2 T} e^{V(\epsilon)} e^{v_2^2 T} e^{V(\epsilon)} e^{v_1^2 T} e^{V(\epsilon)} e^{v_0 T} e^{v(0)}
\]
all requiring six FFTs. Here, we choose \(t_1\) as the free param-
ter. The coefficients are then
Algorithm 4B corresponds to setting \( v_0 = 0 \) and selecting the smaller of the two quadratic solutions, \( t_0 = \frac{1}{2} (1 - \sqrt{3}) \). When \( t_1 = 1/3 \), we observe a variant of algorithm 4D, which we will denote 4D'. Again, one can transform 4D to 4D' continuously by distributing the commutator term at positions \( v_1 \) and \( v_2 \) to \( v_0 \) and \( v_3 \). However, in order to keep the number of gradient terms to a minimum, we will not bother with this refinement here. For positive coefficients, we must have \( \frac{1}{2} (1 - \sqrt{3}) \leq t_1 \leq \frac{1}{2} \). At the upper limit of \( t_1 = 1/2 \), \( t_2 = 0 \), and the algorithm collapses back to algorithm 4A. Thus, there are two continuous families of algorithms, the C type that requires eight FFTs and the D type that requires six. They are joined at both ends by algorithms 4A and 4B with discontinuous changes in the number of FFTs. Generally, there are two continuous families of algorithms, the C type that requires eight FFTs and the D type that requires six. They are joined at both ends by algorithms 4A and 4B with discontinuous changes in the number of FFTs. (Recently, Omelyan, Mryglod, and Folk considered the one-parameter factorization schemes from 4A to 4C and from 4B to 4D separately, without realizing that they can be further extended and joined into one.)

Figure 2 shows that the fourth order convergence error is negative for algorithms 4A and 4B, and positive for 4C and 4D. Since algorithm 4A can be continuously changed to 4C, and algorithm 4B can be transformed into 4D, there must be parameter values such that this fourth order error can be made to vanish. This immediately suggests a simple strategy for further optimization: for each application, one can first minimize the error for a short time run with respect to \( t_0 \) in \( T_{ACB}(4) \) or \( t_1 \) in \( T_{BDA}(4) \). One can then use the algorithm at that value and at a large time step to do long time simulations. Since the error term is generally time dependent, there is no guarantee that, when it is small for one time, it will remain small for all times. For the important case of periodic time dependence, one hopes to reduce the bound within which the error can fluctuate. This seems to work for the present case.

From Fig. 2, as the algorithm is changed continuously from 4A to 4C and back to 4B, one should observe two zero error crossings. In varying \( t_0 \) in \( T_{ACB}(4) \), we only see one. This suggests that the second crossing is when \( 4B' \rightarrow 4B \), which we did not consider in this work. The one zero crossing we observed is at \( t_0 = 0.144 \), precisely between 4A(\( t_0 = 0 \)) and 4C (\( t_0 = \frac{1}{2} \)). A short run (\( t = 100 \tau \)) to determine the optimal value for \( t_0 \) is shown by closed symbols with solid fitting lines in Fig. 3. For \( t_0 = 0.142 \) and 0.146, the convergence errors can be fitted well by fourth order lines as shown. At the value of \( t_0 = 0.144 \), the fourth order coefficient is an order of magnitude smaller. We deem it sufficient to determine the zero-crossing parameter value to three-digit accuracy. The resulting long run is shown as 4ACB in Fig. 2. As shown in Table I, \( \delta_{eq} \) in this case is reduced by an order of magnitude and \( \tau_{eff} \) is lengthened to 8.5. The convergence error curve for 4ACB remained very flat upon going from \( t = 0 \) to 1000 \( \tau \).

In the case of \( T_{BDA}(4) \), algorithm 4D' is sufficiently dissimilar to 4D that the convergence error remained negative as 4B' is morphed to 4D' at \( t_1 = \frac{1}{2} \). However, as \( t_1 \) increases above \( \frac{1}{2} \), the convergence error moves upward and turns positive. Since \( T_{BDA}(4) \) ends up at 4A, the error must cross zero again on its way down. The first crossing is at \( t_1 = 0.35 \). This is shown in Fig. 3 by the open symbol connected by fitted, broken lines. The second crossing is near \( t_1 = 0.46 \), but we do not bother to show this result since it is similar. The long run results using \( t_1 = 0.35 \) in \( T_{BDA}(4) \) are shown in Fig. 2 as 4BDA. As listed in Table I, its equal effort error and effective time step factor are virtually identical to those of 4ACB. In both cases, without any additional computational effort, we have reduced the error by a factor of 5 and lengthened the time step size by a factor of 1.5. Both algorithms can achieve \( 10^{-4} \) accuracy in total energy at a time step size 50 times as large as that of \( \tau_{eff} \).

A key advantage of these factorization algorithms is that one can obtain analytically their energy error terms by use of MATHEMATICA. For fourth order gradient algorithms, there are four nonvanishing error operators in the Hamiltonian:

\[
\Delta E^{(4)} = e_1^2[VTVTV] + e_2^2[TTVV] + e_3^2[VTTTV] + e_4^2[TTTTV],
\]

where \( e_1 \) are coefficient functions depending on \( t_0 \) or \( t_1 \), and \( [ABC] = [A,[B,(C)]] \), etc. The zero-crossing values of \( t_{0,1} \) in \( T_{ACB}(4) \) and \( T_{BDA}(4) \), closely match the crossing points of the first two error coefficients, \( e_1 = e_2 \). For 4ACB and 4BDA, the predicted zero-crossing values from solving \( e_1 = e_2 = 0 \) numerically are \( t_{0,1} = 0.142 \) and 0.350, respectively, in excellent agreement with empirical values. It seems that of the four error operators, only two are dominant.
and they are opposite in sign. It should be noted that these crossing points have nothing to do with the minimum of \( \sum_{i=1}^{4} e_i^2 \). These algorithms are not optimized by simply minimizing the sum of squares of error coefficients. Further discussions on the above one-parameter family of gradient algorithms will be presented in a separate publication.

VI. CONCLUSIONS

In this work, we have derived a new class of fourth order algorithms for solving the time-dependent Schrödinger equation with explicit time-dependent potentials. This class of algorithms is characterized by having only positive factorization coefficients and can achieve great efficiency by knowing the gradient of the potential. For the Walker–Preston model, on an equal effort basis, the convergence errors of these algorithms can be 5000 and 30 times smaller than negative-coefficient algorithms such as the Forest–Ruth and McLachlan algorithms, respectively. Since we have demonstrated similar efficiency gain in 3D when propagating the Schrödinger equation in imaginary time using similar algorithms, we have every reason to expect the efficiency demonstrated here can be carried over to more realistic 3D calculations.

Our one-parameter family of gradient algorithms illustrates the power of the operator factorization approach for solving any evolution equation. The capability to tailor a specific algorithm for a particular application reflects the great versatility of our method. In solving the Walker–Preston model, by tailoring the parameter value to this specific problem, we can further reduce the convergent error by a factor of 5 without additional effort. Our work is encouraging for further systematic study of these and other parameterized families of algorithms and their optimization. In particular, one should explore the freedom in redistributing the commutator term upon going from 4B to 4B’ and from 4D to 4D’. We have not done so here in order to concentrate on algorithms with the minimum number of gradient evaluations. With the development of these powerful gradient algorithms for solving the Schrödinger equation directly, we can see no practical advantage in solving the corresponding classical problem using classical methods. This is in agreement with Sanz-Serna and Portillo’s earlier assessment.

This work spurs further interest in finding six and higher order factorization algorithms with purely positive coefficients. Despite intense effort, we have yet to find a sixth order factorization scheme with positive coefficients. It is most likely that one does not exist, even with the inclusion of the double commutator term. In a recent work the same conclusion is reached, but it offered no proof either. Such proof, if it exists, would make these fourth order gradient algorithms unique. There would be no higher order generalizations. This would raise many other interesting questions, such as what other operators are necessary for a higher order positive time step factorization, what would be the optimal sixth order algorithm with minimal of negative coefficients, etc. All such investigations would deepen our understanding of the operator factorization method of solving evolution equations.

ACKNOWLEDGMENT

This work was supported, in part, by the National Science Foundation, Grant No. PHY-0100839 to one of the authors (S.A.C.).