Higher order decompositions of ordered operator exponentials

2010 J. Phys. A: Math. Theor. 43 065203

(http://iopscience.iop.org/1751-8121/43/6/065203)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **43** (2010) 065203 (20pp) doi:10.1088/1751-8113/43/6/065203

Higher order decompositions of ordered operator exponentials

Nathan Wiebe[1], Dominic Berry[2,3], Peter Høyer[1,4] and Barry C Sanders[1]

[1] Institute for Quantum Information Science, University of Calgary, Alberta T2N 1N4, Canada
[2] Institute for Quantum Computing, University of Waterloo, Ontario N2 L 3G1, Canada
[3] Centre for Quantum Computer Technology, Macquarie University, Sydney, NSW 2109, Australia
[4] Department of Computer Science, University of Calgary, Alberta T2N 1N4, Canada

Received 24 July 2009, in final form 8 December 2009
Published 22 January 2010
Online at stacks.iop.org/JPhysA/43/065203

Abstract

We present a decomposition scheme based on Lie–Trotter–Suzuki product formulae to approximate an ordered operator exponential with a product of ordinary operator exponentials. We show, using a counterexample, that Lie–Trotter–Suzuki approximations may be of a lower order than expected when applied to problems that have singularities or discontinuous derivatives of appropriate order. To address this problem, we present a set of criteria that is sufficient for the validity of these approximations, prove convergence and provide upper bounds on the approximation error. This work may shed light on why related product formulae fail to be as accurate as expected when applied to Coulomb potentials.

PACS numbers: 02.30.Tb, 02.60.Lj

1. Introduction

Ordered operator exponentials arise frequently in physics and applied mathematics [1–4] because they appear in the solution of systems of first-order linear differential equations. The general form is

$$\partial_\lambda \boldsymbol{v} = A(\lambda)\boldsymbol{v}(\lambda), \tag{1}$$

where λ is real, $\boldsymbol{v}$ is a complex vector and $A(\lambda)$ is a linear operator. An important example where equation (1) arises is the Schrödinger equation, in which case the linear operator $A = -\mathrm{i}H$, where H is the Hamiltonian. The generic solution to (1) can be formally expressed as $U(\eta,\mu)\boldsymbol{v}(\mu)$, where $U(\eta,\mu)$ satisfies the following differential equation:

$$\partial_\eta U(\eta,\mu) = A(\eta)U(\eta,\mu) \qquad \text{with} \quad U(\mu,\mu) = \mathbb{1}. \tag{2}$$

The operator U is called an ordered operator exponential if $A(\lambda)$ is λ-dependent. If U is known, then the corresponding differential equation (1) can be solved for any initial condition.

Finding an exact analytical solution to either (1) or (2) is not possible in general, so approximations are needed. Common methods to approximate the solution to these equations include Runge–Kutta methods, Magnus expansions [5, 6] and product formulae. Product formulae assume that a decomposition of $A(\lambda)$ is known such that

$$A(\lambda) = \sum_{j=1}^{m} A_j(\lambda), \tag{3}$$

where m is the number of operators in the decomposition, and each $A_j(\lambda)$ is chosen so that it can be easily exponentiated for every $\lambda \in [\mu, \eta]$. Product formulae then approximate the ordered operator exponential U with a product of ordinary operator exponentials of $A_j(\lambda)$:

$$U(\mu + \Delta\lambda, \mu) \approx \prod_{p=1}^{N} \mathrm{e}^{A_{j_p}(\lambda_p)\Delta\lambda_p}, \tag{4}$$

where $\Delta\lambda = \eta - \mu$ is the interval, $\Delta\lambda_p$ is a real number proportional to $\Delta\lambda$ and N is the number of terms used in the product. Examples of commonly used product formulae include the Trotter formula [7], symplectic integrators such as the Forest–Ruth formula [8, 9] and the Lie–Trotter–Suzuki formula [10–12].

Product formulae have many advantages over alternative approximations. One advantage is that product formulae can manifestly preserve certain symmetries of U. For example, in many applications in quantum mechanics, $U(\eta, \mu)$ is a unitary operator. Under these circumstances a product formula approximates U with a product of unitary operators, which implies that up to roundoff error (4) is unitary, as opposed to Runge–Kutta methods which will not approximate U with a unitary operator. Another advantage is that no commutators or integrals need to be computed in order to use product formulae, unlike the Magnus formula. Product formulae often involve only sparse matrix vector multiplication, and therefore provide algorithms that are easily parallelized.

Two hurdles must be overcome when devising a product formula. First, the sequence of exponentials must account for the λ-dependence of the A_j. Second, because A_j does not necessarily commute with $A_{j'}$, it is non-trivial to decompose U into a product of exponentials such that (4) is valid, even without λ-dependence.

These two issues can be approached separately or holistically. A situation where it may be preferable to view them separately occurs when the ordering is trivial, such as when $A(\lambda)$ is constant. Product formulae have found a host of applications under this assumption [16–20]. However, if $A(\lambda)$ is not constant, it is simpler to use a single approximation to handle the decomposition as well as the ordering. One such approximation is the Lie–Trotter–Suzuki formula for ordered operator exponentials [10–12].

The Lie–Trotter–Suzuki formulae are a family of recursively generated product formula approximations for ordered or ordinary operator exponentials of $A(\lambda)$ [10–12]. These formulae are generated by applying an approximation building formula that can generate higher order approximations out of lower order formulae. The Lie–Trotter–Suzuki formulae are found by applying this method iteratively $k-1$ times to the Lie–Trotter formula to yield an approximation that is an $O(\Delta\lambda^{2k+1})$ Trotter-like approximation to $U(\mu + \Delta\lambda, \mu)$. This is significant because it is difficult to generalize many other product formulae, such as certain symplectic integrators [13], to arbitrarily high order.

These formulae are particularly promising for quantum computer simulations of quantum systems because Lie–Trotter–Suzuki approximations provide a sequence of unitary operations that accurately approximates the time-evolution operator for the simulated system [19]. Since a quantum computer can directly implement unitary operations on an input state, this approximation can provide a program that a quantum computer can follow to simulate a given time-dependent quantum system. Lie–Trotter–Suzuki formulae have also been widely used to approximate the ordered operator exponentials that appear in molecular dynamics [21–23], calculate reaction rates [24], solve problems in celestial dynamics [25] and solve time-dependent Maxwell equations [26].

There are two unresolved problems surrounding the use of Lie–Trotter–Suzuki formulae. One problem is that the previous work does not examine what conditions are needed to ensure that the formulae provide approximations of the desired order. It would appear from the previous work that arbitrarily high-order approximations are obtained for any $A(\lambda)$, but as we show below, that is not the case. The other problem is that, although upper bounds on the error have been proven in the case where $A(\lambda)$ is constant and anti-Hermitian [19], no bounds have been proven in the general case. The lack of error bounds for this approximation limits the utility of the Lie–Trotter–Suzuki formula in situations where the approximation error cannot be directly measured (such as in quantum computing).

In this work, we find that if the operator or its derivatives change discontinuously, or if the operator contains singularities (such as in the Coulomb potential), then the Lie–Trotter–Suzuki formulae may not provide approximations of the desired order. We show that arbitrarily high-order approximations can be generated if the exponent is sufficiently smooth, provide error bounds for the approximation and find sufficient conditions for convergence of these approximations. In contrast, for the case where $A(\lambda)$ is not sufficiently smooth, we provide an example where arbitrarily high-order approximations are not obtained. These observations may also shed light onto why other sixth-order splitting methods, that are based on Suzuki's idea, fail to provide the expected accuracy when applied to Coulomb potentials [27].

In section 2 we give the background for Trotter product formulae in detail. In section 3 we review Suzuki's decomposition methods, and provide our form of Suzuki's recursive method. In section 4 we introduce our terminology and present our main result. Then we rigorously prove the scaling of the error in section 5, and place an upper bound on the error in section 6. We then use the error bounds in section 7 to find the appropriate order of the integrator to use.

2. Trotter formulae

Typically there are two different scenarios that may be considered. First, one may consider a short interval $\Delta\lambda$; the goal is then to obtain an error that decreases rapidly as $\Delta\lambda \to 0$. Alternatively, the interval $\Delta\lambda$ may be long, and the goal is to obtain an approximation to within a certain error with as few exponentials as possible. For example, given λ-independent operators A and B,

$$\mathrm{e}^{\Delta\lambda(A+B)} = \mathrm{e}^{\Delta\lambda A}\,\mathrm{e}^{\Delta\lambda B} + O(\Delta\lambda^2) \tag{5}$$

holds. This gives an accurate approximation for small $\Delta\lambda$. For large $\Delta\lambda$, we may use equation (5) to derive the Trotter formula

$$\mathrm{e}^{\Delta\lambda(A+B)} = (\mathrm{e}^{\Delta\lambda A/n}\,\mathrm{e}^{\Delta\lambda B/n})^n + O(\Delta\lambda^2/n). \tag{6}$$

This error scaling is obtained because the error for an interval of length $\Delta\lambda/n$ is $O((\Delta\lambda/n)^2)$. Taking the power of n then gives n times this error if the norm of $\exp[(A+B)\Delta\lambda]$ is at most 1 for any $\Delta\lambda > 0$, resulting in the error shown in equation (6). To obtain a given error ϵ, the value of n must then scale as $O(\Delta\lambda^2/\epsilon)$. The goal is to make the value of n needed to achieve a given accuracy as small as possible (n is proportional to the total number of exponentials).

More generally, for a sum of an arbitrary number of operators A_j, similar formulae give the same scaling. To obtain better scaling, one can use a different product of exponentials.

The Lie–Trotter–Suzuki product formulae [12] replace the product for short $\Delta\lambda$ with another that gives error scaling as $O(\Delta\lambda^{p+1})$.

It can be seen that splitting large $\Delta\lambda$ into n intervals as in equation (6) yields an error scaling as $O(\Delta\lambda^{p+1}/n^p)$ if the norm of U is at most 1. It may at first appear that this gives *worse* results for large $\Delta\lambda$ due to the higher power. In fact, there is an advantage due to the fact that a higher power of n is obtained. The value of n required to achieve a given error then scales as $O(\Delta\lambda^{1+1/p}/\epsilon^{1/p})$. Therefore, for large $\Delta\lambda$, increasing p gives scaling of N that is close to linear in $\Delta\lambda$.

Similar considerations hold for the case of ordered exponentials (i.e. with λ-dependence). Huyghebaert and De Raedt showed how to generalize the Trotter formula to apply to ordered operator exponentials [15]. Their formula has a decomposition error that is $O(\Delta\lambda^2)$, but requires that the integrals of $A(\lambda)$ and $B(\lambda)$ are known. Subsequently, Suzuki developed a method to achieve an error that scales as $O(\Delta\lambda^{p+1})$ for some ordered exponentials [10], and does not require the integrals of $A(\lambda)$ and $B(\lambda)$ to be known. We find that, in contrast to the λ-independent case, it is not necessarily possible to obtain scaling as $O(\Delta\lambda^{p+1})$ for arbitrarily large p. It is possible if derivatives of all orders exist. If there are higher order derivatives that do not exist, then it is still possible to use Suzuki's method to obtain error scaling as $O(\Delta\lambda^{p+1})$ for some values of p, but the maximum value of p for which this scaling can be proven depends on what orders of derivatives exist.

3. Suzuki decompositions

In this section, we explain Suzuki decompositions in more detail. In general, decompositions are of the form

$$\tilde{U}(\mu + \Delta\lambda, \mu) = \prod_{i=1}^{N} \mathrm{e}^{A_{j_i}(\lambda_i)\Delta\lambda_i}, \tag{7}$$

where $\tilde{U}(\mu + \Delta\lambda, \mu)$ is a decomposition of $U(\mu + \Delta\lambda, \mu)$. There are several types of decompositions, but the type that we focus on in this paper is symmetric decompositions because all Suzuki decompositions are symmetric.

Definition 1. *The operator $\tilde{U}(\mu + \Delta\lambda, \mu)$ is a symmetric decomposition of the operator $U(\mu + \Delta\lambda, \mu)$ if $\tilde{U}(\mu + \Delta\lambda, \mu)$ is a decomposition of $U(\mu + \Delta\lambda, \mu)$ and $\tilde{U}(\mu + \Delta\lambda, \mu) = [\tilde{U}(\mu, \mu + \Delta\lambda)]^{-1}$.*

An important method for generating symmetric decompositions is due to Suzuki [10], which we call Suzuki's recursive method. Furthermore, we call any decomposition formula that is found using this method a Suzuki decomposition.

Suzuki's recursive method takes a symmetric decomposition formula $U_p(\mu + \Delta\lambda, \mu)$, which approximates an ordered operator exponential $U(\mu + \Delta\lambda, \mu)$ with an approximation error that is at most proportional to $\Delta\lambda^{2p+1}$ as input, and outputs a symmetric approximation formula $U_{p+1}(\mu + \Delta\lambda, \mu)$ which will often have an approximation error that is proportional to $\Delta\lambda^{2p+3}$. The approximation $U_{p+1}(\mu + \Delta\lambda, \mu)$ is found using the recursion relation

$$\begin{aligned} U_{p+1}(\mu + \Delta\lambda, \mu) \equiv\, & U_p(\mu + \Delta\lambda, \mu + [1 - s_p]\Delta\lambda) U_p(\mu + [1 - s_p]\Delta\lambda, \mu + [1 - 2s_p]\Delta\lambda) \\ & \times U_p(\mu + [1 - 2s_p]\Delta\lambda, \mu + 2s_p\Delta\lambda) U_p(\mu + 2s_p\Delta\lambda, \mu + s_p\Delta\lambda) \\ & \times U_p(\mu + s_p\Delta\lambda, \mu) \end{aligned} \tag{8}$$

with $s_p \equiv (4 - 4^{1/(2p+1)})^{-1}$.

Suzuki's recursive method does not actually approximate $U(\mu + \Delta\lambda, \mu)$, but rather it builds a higher order approximation formula out of a lower order one. Therefore, this method can only be used to approximate $U(\mu + \Delta\lambda, \mu)$ if it is seeded with an appropriate initial approximation. A convenient approximation formula based on Suzuki's recursive method is the kth-order Lie–Trotter–Suzuki product formula which is defined as follows.

Definition 2. *The* kth-*order Lie–Trotter–Suzuki product formula for the operator* $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$ *and the interval* $[\mu, \mu + \Delta\lambda]$ *is defined to be* $U_k(\mu + \Delta\lambda, \mu)$*, which is found by using*

$$U_1(\mu + \Delta\lambda, \mu) \equiv \left(\prod_{j=1}^{m} \exp(A_j(\mu + \Delta\lambda/2)\Delta\lambda/2)\right)\left(\prod_{j=m}^{1} \exp(A_j(\mu + \Delta\lambda/2)\Delta\lambda/2)\right) \quad (9)$$

as an initial approximation and by applying Suzuki's recursive method to it $k - 1$ *times.*

Based on Suzuki's analysis [10], U_k should have an approximation error that is proportional to $\Delta\lambda^{2k+1}$. Hence, if $\Delta\lambda$ is sufficiently small, then the formula should be highly accurate. One might think that it would be advantageous to increase k without limit, in order to obtain increasingly accurate approximation formulae. However this is not the case because the number of terms in the formula increases exponentially with k. The best value of k to use can be expected to depend on the desired accuracy, as well as a range of other parameters [19].

4. Sufficiency criterion for decomposition

Suzuki's recursive method is a powerful technique for generating high-order decomposition formulae for ordered operator exponentials. The kth-order Lie–Trotter–Suzuki product formula, in particular, seems to be well suited for approximating ordered operator exponentials that appear in quantum mechanics and in other fields; furthermore, it appears that these formulae should be applicable to approximating the ordered exponentials of any finite-dimensional operator $A(\lambda)$. However, it turns out that Suzuki's recursive method does not always generate a higher order decomposition formula from a lower order one.

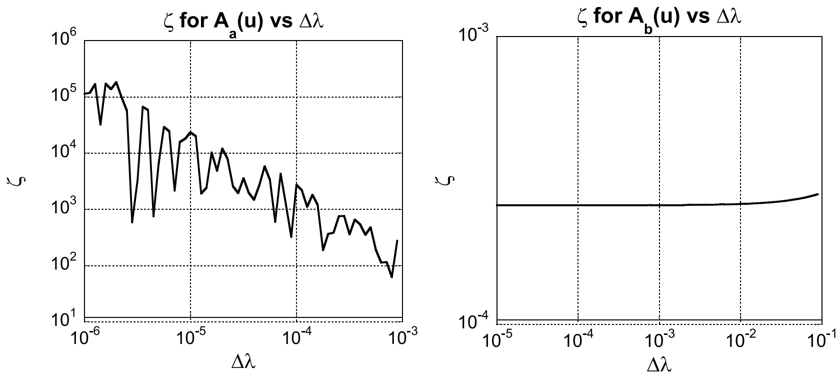
We show this using the example of the operator $A_a(\lambda) = \lambda^3 \sin(1/\lambda)\mathbb{1}$. For this operator, the second-order Lie–Trotter–Suzuki product formula is not an approximation whose error as measured by the 2-norm is $O(\Delta\lambda^5)$. In figure 1 we see that the error is proportional to $\Delta\lambda^4$ for the operator $A_a(\lambda)$, rather than $\Delta\lambda^5$ as predicted by Suzuki's analysis. In contrast, we do observe this scaling for the analytic operator $A_b(\lambda) = \cos(\lambda)\mathbb{1}$. This shows that the second-order Lie–Trotter–Suzuki product formula is not as accurate as may be expected for some non-analytic operators.

Our analysis will show that this discrepancy arises from the fact that $A_a(\lambda) = \lambda^3 \sin(1/\lambda)\mathbb{1}$ is not smooth enough for the second-order Lie–Trotter–Suzuki formula to have an error which is $O(\Delta\lambda^5)$. In the subsequent discussion we will need to classify the smoothness of the operators that arise in decompositions. We use the smoothness criteria $2k$-smooth and Λ-$2k$-smooth, which we define below.

Definition 3. *The set of operators* $\{A_j : j = 1, \ldots, m\}$ *is* P-*smooth on the interval* $[\mu, \eta]$ *if, for each* A_j*, the quantity* $\left\|\partial_\lambda^P A_j(\lambda)\right\|$ *is finite on the interval* $[\mu, \eta]$.

Here, and throughout this paper, we define $\|\cdot\|$ to be the 2-norm. Also, if $\{A_j\}$ is P-smooth for every positive integer P, we call $\{A_j\}$ ∞-smooth.

This condition is not precise enough for all of our purposes. For our error bounds we need to introduce the more precise condition of Λ-P-smoothness. This condition is useful



(*a*) (*b*)

Figure 1. This is a plot of $\zeta = \|U(\Delta\lambda, 0) - U_2(\Delta\lambda, 0)\|_2/\Delta\lambda^5$ for $A_a = \lambda^3 \sin(1/\lambda)\mathbb{1}$ in (*a*) and $A_b = \cos(\lambda)\mathbb{1}$ in (*b*). The error in (*a*) is proportional to $\Delta\lambda^4$ as opposed to the $O(\Delta\lambda)^5$ scaling predicted for Suzuki's corresponding decomposition. The error in (*b*) is proportional to $\Delta\lambda^5$ as expected for that Suzuki decomposition.

because it guarantees that if the set $\{A_j\}$ is Λ-P-smooth and $p \leqslant P$, then $\|A^{(p)}(\lambda)\| \leqslant \Lambda^p$. This property allows us to write our error bounds in a form that does not contain any of the derivatives of $A(\lambda)$ individually but rather in terms of Λ, which provides an upper bound on the magnitude of any of these derivatives. We formally define this condition below.

Definition 4. *The set of operators* $\{A_j : j = 1, \ldots, m\}$ *is* Λ-P-*smooth on the interval* $[\mu, \eta]$ *if* $\{A_j\}$ *is* P-*smooth and* $\Lambda \geqslant \left(\sum_{j=1}^m \left\|A_j^{(p)}(\lambda)\right\|\right)^{1/(p+1)}$ *for all* $\lambda \in [\mu, \mu + \Delta\lambda]$ *and* $p \in \{1, 2, \ldots, P\}$.

This parameter is important because it represents the most significant timescale in the approximation; furthermore, the failure of the approximation in figure 1 can be understood as a consequence of this timescale becoming infinite due to the divergence of the second derivative.

For example, if $\{A\} = \{\sin(2\lambda)\mathbb{1}\}$, then using definition 4, $\{A\}$ is $2^{2/3}$-2-smooth on the interval $[0, \pi]$ because the largest value $\|A(\lambda)^{(p)}\|^{1/(p+1)}$ takes is $2^{2/3}$ for $p = 0, 1, 2$. It is also 2-2-smooth because $2^{2/3} \leqslant 2$; furthermore, since $\|A(\lambda)^{(p)}\|^{1/(p+1)} < 2$ for all positive integers p, $\{A\}$ is also 2-∞-smooth.

Using this measure of smoothness, we can then state the following theorem, which is also the main theorem in this paper.

Theorem 1. *If the set* $\{A_j\}$ *is* Λ-$2k$-*smooth on the interval* $[\mu, \mu+\Delta\lambda]$, $\epsilon \leqslant (9/10)(5/3)^k \Lambda\Delta\lambda$ *and* $\max_{x>y} \|U(x, y)\| \leqslant 1$, *then a decomposition* $\tilde{U}(\mu+\Delta\lambda, \mu)$ *can be constructed such that* $\|\tilde{U} - U\| \leqslant \epsilon$ *and the number of operator exponentials present in* $\tilde{U}$, *N, satisfies*

$$N \leqslant \left\lceil 3m\Lambda\Delta\lambda k \left(\frac{25}{3}\right)^k \left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)^{1/2k} \right\rceil. \tag{10}$$

We prove theorem 1 in several steps, with details presented in sections 5 and 6. In section 5 we construct the Taylor series for an ordered operator exponential, and use this series to prove that the Lie–Trotter–Suzuki product formula can generate an approximation whose error is $O(\Delta\lambda^{2k+1})$ if $\{A_j\}$ is $2k$-smooth on the interval $[\mu, \mu + \Delta\lambda]$. In section 6 we use the order estimates in section 5 to obtain upper bounds on the error. The result of theorem 1 then follows by counting the number of exponentials needed to make the error bound less than ϵ. Finally in section 7 we show that if k is chosen appropriately, then N scales almost linearly with $\Delta\lambda$ if there exists a value of Λ such that $\{A_j\}$ is Λ-∞-smooth on $[\mu, \mu + \Delta\lambda]$ for every $\Delta\lambda > 0$.

5. Decomposing ordered exponentials

In this section we present a new derivation of Suzuki's recursive method. Our derivation has the advantage that it can be rigorously proven that if $\{A_j\}$ is $2k$-smooth, where $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$, then the kth-order Lie–Trotter–Suzuki product formula will have an error of $O(\Delta\lambda^{2k+1})$. We show this in three steps. We first give an expression for the Taylor series expansion of an ordered exponential $U(\mu + \Delta\lambda, \mu)$. Then, using this expression for the Taylor series, we show in theorem 2 that Suzuki's recursive method can be used to generate approximations to $U(\mu + \Delta\lambda, \mu)$ with an error that is $O(\Delta\lambda^{2k+1})$, provided $\{A_j\}$ is $2k$-smooth. Finally, we show in corollary 1 that, if $\{A_j\}$ is $2k$-smooth, then the kth-order Lie–Trotter–Suzuki product formula has an error that is at most proportional to $\Delta\lambda^{2k+1}$.

It is convenient to expand U in a Taylor series of the form

$$U(\mu + \Delta\lambda, \mu) = \mathbb{1} + T_1(\mu)\Delta\lambda + \frac{T_2(\mu)\Delta\lambda^2}{2!} + \cdots. \tag{11}$$

If $A(\lambda)$ is not analytic, then this Taylor series must be truncated, and the error can be bounded by the following lemma.

Lemma 1. *If the operator $A(\lambda)$ is, for $P \in \mathbb{N}_0$, P times differentiable on the interval $[\mu, \mu + \Delta\lambda] \subset \mathbb{R}$, then*

$$\left\| U(\mu + \Delta\lambda, \mu) - \sum_{p=0}^{P} \frac{(\Delta\lambda)^p T_p}{p!} \right\| \leqslant \frac{\max_{\lambda\in[\mu,\mu+\Delta\lambda]} \|T_{P+1}(\lambda)U(\lambda, \mu)\| \Delta\lambda^{P+1}}{(P+1)!}, \tag{12}$$

where $T_p(\lambda)$ is defined by the recursion relation $T_{p+1}(\lambda) \equiv T_p(\lambda)A(\lambda) + \partial_\lambda T_p(\lambda)$, with $T_0 \equiv \mathbb{1}$ chosen to be the initial condition.

Proof. We first show, for the positive integer $\ell \leqslant P + 1$, that if $\lambda \in [\mu, \mu + \Delta\lambda]$, then

$$\frac{\partial^\ell}{\partial\lambda^\ell} U(\lambda, \mu) = T_\ell(\lambda)U(\lambda, \mu). \tag{13}$$

This equation can be validated by using induction on ℓ. The base case follows by setting $\ell = 0$ in (13). We then demonstrate the induction step by noting that if (13) is true for $\ell \leqslant P$, then

$$\begin{aligned} \frac{\partial^{\ell+1}}{\partial\lambda^{\ell+1}} U(\lambda, \mu) &= \frac{\partial}{\partial\lambda} T_\ell(\lambda)U(\lambda, \mu) \\ &= \left[\frac{\partial}{\partial\lambda} T_\ell(\lambda)\right] U(\lambda, \mu) + T_\ell(\lambda) \left[\frac{\partial}{\partial\lambda} U(\lambda, \mu)\right]. \end{aligned} \tag{14}$$

Since $T_\ell(\lambda)$ contains derivatives of $A(\lambda)$ up to order $\ell - 1$, it follows that $\partial_\lambda T_\ell(\lambda)$ contains derivatives up to order ℓ. Then, since $A(\lambda)$ is P times differentiable, $\partial_\lambda T_\ell(\lambda)$ exists if $\ell \leqslant P$, which implies that $\partial_\lambda^{\ell+1} U(\lambda, \mu)$ exists. We then use the differential equation in (2) to evaluate the derivative of $U(\mu + \Delta\lambda, \mu)$ in (14) and then we use the fact that $T_{\ell+1}(\lambda) = T_\ell(\lambda)A(\lambda) + \partial_\lambda T_\ell(\lambda)$ to find that

$$\frac{\partial^{\ell+1}}{\partial\lambda^{\ell+1}} U(\lambda, \mu) = T_{\ell+1}(\lambda)U(\lambda, \mu). \tag{15}$$

This demonstrates the induction step in our proof of (13). Since we have already shown that (13) is valid for T_0, it is also true for all $T_\ell(\lambda)$ if $\ell \leqslant P+1$ by induction on ℓ.

We then use (13) and Taylor's theorem to conclude that

$$U(\mu+\Delta\lambda,\mu)=\sum_{p=0}^{P}\frac{(\Delta\lambda)^pT_p(\mu)}{p!}+\int_0^{\Delta\lambda}T_{P+1}(\mu+\lambda)U(\mu+\lambda,\mu)\frac{(\Delta\lambda-\lambda)^P}{P!}\,\mathrm{d}\lambda. \tag{16}$$

We rearrange this result and find that

$$\left\|U(\mu+\Delta\lambda,\mu)-\sum_{p=0}^{P}\frac{(\Delta\lambda)^pT_p}{p!}\right\|\leqslant\frac{\max_{\lambda\in[\mu,\mu+\Delta\lambda]}\|T_{P+1}(\lambda)U(\lambda,\mu)\|\Delta\lambda^{P+1}}{(P+1)!}. \tag{17}$$

□

Lemma 1 provides a convenient expression for the terms in the Taylor series of $U(\mu+\Delta\lambda,\mu)$, and it also estimates the error invoked by truncating the series at order P for any $P\in\mathbb{N}_1$. The following theorem uses this lemma to show that Suzuki's recursive method will produce a higher order approximation from a lower order symmetric approximation if $\{A_j\}$ is sufficiently smooth on $[\mu,\mu+\Delta\lambda]$ and $A(\lambda)$ is the sum of all of the elements in the set $\{A_j\}$.

Theorem 2. *If*

(i) $A(\lambda)=\sum_{j=1}^m A_j(\lambda)$ *where the set* $\{A_j\}$ *is, for a fixed* $p\geqslant 1$, $2(p+1)$*-smooth on the interval* $[\mu,\mu+\Delta\lambda]$,

(ii) $U_p(\mu+\Delta\lambda,\mu)$ *is a symmetric approximation formula such that* $\|U_p(\mu+\Delta\lambda,\mu)-U(\mu+\Delta\lambda,\mu)\|\in O(\Delta\lambda^{2p+1})$,

(iii) $U_{p+1}(\mu+\Delta\lambda,\mu)$ *is found by applying Suzuki's recursive method on* $U_p(\mu+\Delta\lambda,\mu)$,

then $U_p(\mu+\Delta\lambda,\mu)$ *satisfies*

$$\|U(\mu+\Delta\lambda,\mu)-U_{p+1}(\mu+\Delta\lambda,\mu)\|\in O(\Delta\lambda^{2p+3}). \tag{18}$$

Proof. In this proof we compare the Taylor series of U to that of U_p and show that choosing s_p appropriately will cause both the terms proportional to $\Delta\lambda^{2p+2}$ and $\Delta\lambda^{2p+3}$ to vanish.

Expanding the recursive formula in lemma 1, we see that a Taylor polynomial can be constructed for U whose difference from U is $O(\Delta\lambda^{2p+3})$ because $\{A_j\}$ is $2(p+1)$-smooth on $[\mu,\mu+\Delta\lambda]$. A similar polynomial can be constructed for U_p by Taylor expanding each A_j that appears in the exponentials in U_p, and then expanding each of these exponentials. Then, because $\{A_j\}$ is $2(p+1)$-smooth, Taylor's theorem implies that this polynomial can be constructed such that the difference between it and U_p is $O(\Delta\lambda^{2p+3})$. Therefore, since $\|U-U_p\|\in O(\Delta\lambda^{2p+1})$, there exist operators $C(\mu)$ and $E(\mu)$ that are independent of $\Delta\lambda$, such that

$$U_p(\mu+\Delta\lambda,\mu)-U(\mu+\Delta\lambda,\mu)=C(\mu)\Delta\lambda^{2p+1}+E(\mu)\Delta\lambda^{2p+2}+O(\Delta\lambda^{2p+3}). \tag{19}$$

We then use the above equation to write U_{p+1} as

$$\begin{aligned}&(U(\mu+\Delta\lambda,\mu+[1-s_p]\Delta\lambda)+C(\mu+[1-s_p]\Delta\lambda)(s_p\Delta\lambda)^{2p+1}+\cdots)\times\cdots\\&\quad\times(U(\mu+s_p\Delta\lambda,\mu)+C(\mu)(s_p\Delta\lambda)^{2p+1}+\cdots).\end{aligned} \tag{20}$$

Because $\{A_j\}$ is $2(p+1)$-smooth and $A(\lambda)=\sum_{j=1}^m A_j(\lambda)$, it follows that U_{p+1} is differentiable $2(p+1)$ times. Then, since U is differentiable $2(p+1)$ times, it follows from Taylor's theorem that C is differentiable, and hence we can Taylor expand each C in this formula in powers of

$\Delta\lambda$ to lowest order. By doing so, and defining $\tilde{E}(\mu)$ to be the sum of all the terms that are proportional to $\Delta\lambda^{2p+2}$ in this expansion, we find that

$$U_{p+1}(\mu, \mu + \Delta\lambda) = U(\mu + \Delta\lambda, \mu) + \left[4s_p^{2p+1} + [1 - 4s_p]^{2p+1}\right]C(\mu)\Delta\lambda^{2p+1} + \tilde{E}(\mu)\Delta\lambda^{2p+2} + O(\Delta\lambda^{2p+3}). \tag{21}$$

Then we see that if $s_p = (4 - 4^{1/(2p+1)})^{-1}$, then the terms of order $2p+1$ in the above equation vanish. Hence, the error invoked using U_{p+1} instead of U is $O(\Delta\lambda^{2p+2})$ with this choice of s_p.

Next we show that $\tilde{E}(\mu) = 0$ using reasoning that is similar to that used by Suzuki's proof of his recursive method for the case where $A(\lambda)$ is a constant operator [11, 12]. Because U_{p+1} is symmetric, it follows from definition 1 that if we neglect terms of order $\Delta\lambda^{2p+3}$ and higher, then

$$\begin{aligned}\mathbb{1} &= U_{p+1}(\mu, \mu + \Delta\lambda)U_{p+1}(\mu + \Delta\lambda, \mu)\\ &= (U(\mu, \mu + \Delta\lambda) + \tilde{E}(\mu + \Delta\lambda)\Delta\lambda^{2p+2})(U(\mu + \Delta\lambda, \mu) + \tilde{E}(\mu)\Delta\lambda^{2p+2}) + O(\Delta\lambda^{2p+3})\\ &= \mathbb{1} + [U(\mu, \mu + \Delta\lambda)\tilde{E}(\mu) + \tilde{E}(\mu + \Delta\lambda)U(\mu + \Delta\lambda, \mu)]\Delta\lambda^{2p+2} + O(\Delta\lambda^{2p+3}).\end{aligned} \tag{22}$$

This equation is only valid if $[U(\mu, \mu + \Delta\lambda)\tilde{E}(\mu) + \tilde{E}(\mu + \Delta\lambda)U(\mu + \Delta\lambda, \mu)] \in O(\Delta\lambda)$.

We then show that $\tilde{E}$ is zero by taking the limit of the above equation as $\Delta\lambda$ approaches zero. We need to ensure that E is continuous to evaluate this limit. The operator $\tilde{E}$ consists of products of derivatives of elements from the set $\{A_j\}$, and these derivatives are of order at most $2p+1$. Then, since each A_j is differentiable $2p+2$ times, $\tilde{E}$ is differentiable, and hence it is continuous. Then using this fact it follows that

$$\lim_{\Delta\lambda\to 0}[U(\mu, \mu + \Delta\lambda)\tilde{E}(\mu) + \tilde{E}(\mu + \Delta\lambda)U(\mu + \Delta\lambda, \mu)] = 2\tilde{E}(\mu) = 0. \tag{23}$$

This implies that the norm of the difference between U and U_{p+1} is proportional to $\Delta\lambda^{2p+3}$, which concludes our proof of theorem 2. □

We will now show that using the kth-order Lie–Trotter–Suzuki product formula invokes an error that is proportional to $\Delta\lambda^{2k+1}$ if $\{A_j\}$ is $2k$-smooth. It should be noted that although our result in theorem 2 is general, this corollary specifically applies to the Lie–Trotter–Suzuki formula as defined in definition 2. Similar results can be shown for different initial approximations by adapting the techniques used below.

Corollary 1. *Let $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$, where the set $\{A_j\}$ is $2k$-smooth, for $k \geqslant 1$, on the interval $[\mu, \mu + \Delta\lambda]$, and let $U(\mu + \Delta\lambda, \mu)$ be the ordered operator exponential generated by $A(\lambda)$. If $U_k(\mu, \mu + \Delta\lambda)$ is the kth-order Lie–Trotter–Suzuki formula, then*

$$\|U(\mu + \Delta\lambda, \mu) - U_k(\mu + \Delta\lambda, \mu)\| \in O(\Delta\lambda^{2k+1}). \tag{24}$$

Proof. Our proof of the corollary follows from an inductive argument on k. The validity of the base case can be verified by using lemma 1. More specifically, $\{A_j\}$ is $2k$-smooth on $[\mu, \mu + \Delta\lambda]$ and $k \geqslant 1$, so A is at least three times differentiable on that interval. This means that lemma 1 implies

$$U(\mu + \Delta\lambda, \mu) = \mathbb{1} + A(\mu)\Delta\lambda + [A^2(\mu) + A'(\mu)]\Delta\lambda^2/2 + O(\Delta\lambda^3). \tag{25}$$

This expansion is also obtained by Taylor expanding $\exp(A(\mu + \Delta\lambda/2)\Delta\lambda)$ to third order, so

$$\|U(\mu + \Delta\lambda, \mu) - \exp(A(\mu + \Delta\lambda/2)\Delta\lambda)\| \in O(\Delta\lambda)^3. \tag{26}$$

Since $U_1(\mu + \Delta\lambda, \mu)$ is the Lie–Trotter formula for a constant $A(\lambda)$ equal to $A(\mu + \Delta\lambda/2)$, it follows that

$$\|U_1(\mu, \mu + \Delta\lambda) - \exp(A(\mu + \Delta\lambda/2)\Delta\lambda)\| \in O(\Delta\lambda)^3. \tag{27}$$

It follows from the above equations and from the triangle inequality that the norm of the difference between U_1 and U is at most proportional to $\Delta\lambda^3$.

Since we have shown that $U_1(\mu + \Delta\lambda, \mu)$ is a symmetric approximation formula whose error is $O(\Delta\lambda^3)$, it follows from theorem 2 and induction that, if $\{A_j\}$ is $2k$-smooth, then a symmetric approximation formula whose error is $O(\Delta\lambda^{2k+1})$ can be constructed from $U_1(\mu + \Delta\lambda, \mu)$ by applying Suzuki's recursive method to it $k - 1$ times. □

We have shown in this section that if $\{A_j\}$ is $2k$-smooth on the interval $[\mu, \mu + \Delta\lambda]$ and $p \leqslant k$, then Suzuki's recursive method can be used to create a symmetric decomposition whose error is $O(\Delta\lambda)^{2p+1}$ out of a symmetric decomposition whose error is $O(\Delta\lambda)^{2p-1}$. Then we have used this fact to show that the norm of the difference between $U(\mu + \Delta\lambda, \mu)$ and the kth-order Lie–Trotter–Suzuki formula is $O(\Delta\lambda^{2k+1})$. In the following section, we strengthen this result by providing an upper bound on the error invoked by using the kth-order Lie–Trotter–Suzuki product formula.

6. Error bounds and convergence for decomposition

We showed in section 5 that if the kth-order Lie–Trotter–Suzuki product formula is used in the place of the ordered operator exponential of $A(\lambda)$, then an error is incurred that is at most proportional to $\Delta\lambda^{2k+1}$, provided $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$ and the set of operators $\{A_j\}$ is sufficiently smooth. We also showed that a sufficient condition for smoothness of the set $\{A_j\}$ is a condition that we called $2k$-smooth, where this condition is defined in definition 3. In this section, we extend that result by finding upper bounds on the error invoked in using the Lie–Trotter–Suzuki product formula to approximate ordered operator exponentials if $\{A_j\}$ is Λ-$2k$-smooth. Unlike the previous section, here we assume that $\max_{x>y} \|U(x, y)\|$ is at most one. This assumption is important because it ensures that our error bounds are not exponentially large. Our work can be made applicable to the case where this norm is greater than one by normalizing U. We discuss the implications of this in appendix B. Note that, in the particular case that $A(\lambda)$ is anti-Hermitian (e.g. when $A = -\mathrm{i}H$ for Hamiltonian evolution), U is unitary, so $\max_{x>y} \|U(x, y)\| = 1$.

In this section, we first provide an upper bound on the error invoked in using the kth-order Lie–Trotter–Suzuki product formula to approximate the ordered operator exponential $U(\mu + \Delta\lambda, \mu)$ if $\Delta\lambda$ is sufficiently short. We then use this result to provide an upper bound on the error if $\Delta\lambda$ is not short. More specifically, we show that for every $\epsilon > 0$ and $\Delta\lambda > 0$, there exists an integer r such that

$$\left\| U(\mu + \Delta\lambda, \mu) - \prod_{q=1}^{r} U_k\left(\mu + q\Delta\lambda/r, \mu + (q-1)\Delta\lambda/r\right) \right\| \leqslant \epsilon \tag{28}$$

if $\{A_j\}$ is $2k$-smooth on the interval $[\mu, \mu + \Delta\lambda]$. Finally by multiplying the number of exponentials in each kth-order Lie–Trotter–Suzuki product formula by r, we find the number of exponentials used in the product in (28). We then use this result to prove theorem 1.

Our upper bound on the error invoked by using a single U_k to approximate the ordered operator exponential $U(\mu + \Delta\lambda, \mu)$ is given in theorem 3. Before stating theorem 3, we first define the following terms. Since the kth-order Lie–Trotter–Suzuki product formula is a product of $2m5^{k-1}$ exponentials, we can express this product as $\prod_{c=1}^{2m5^{k-1}} \exp(A_{j_c}(\mu_c)\Delta\lambda_c)$. We then use this expansion to define the following two useful quantities.

Definition 5. *We define* $q_{c,2k} \equiv \frac{\Delta\lambda_c}{\Delta\lambda}$ *and* $Q_k \equiv \max_c |q_{c,2k}|$.

It can be shown that $Q_1 = 1/2$, and that if $p > 1$ then $Q_p = |1-4s_1|\cdots|1-4s_{p-1}|$. We show in appendix A that for any integer $p > 0$, $Q_p \leqslant 2p/3^p$, implying that Q_p decreases exponentially with p. We use this definition of Q_k in the following theorem, which gives an upper bound on the difference between the ordered operator exponential $U(\mu+\Delta\lambda, \mu)$ and the kth-order Lie–Trotter–Suzuki product formula $U_k(\mu+\Delta\lambda, \mu)$.

Theorem 3. *Let $A(\lambda) = \sum_{j=1}^m A_j(\lambda)$, $\{A_j\}$ be Λ-$2k$-smooth on the interval $[\mu, \mu+\Delta\lambda]$ and $\max_{x>y}\|U(x,y)\| \leqslant 1$. If $2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda \leqslant 1/2$, then*

$$\|U(\mu+\Delta\lambda,\mu) - U_k(\mu+\Delta\lambda,\mu)\| \leqslant 2[3(5)^{k-1}Q_k\Lambda\Delta\lambda]^{2k+1}, \tag{29}$$

where U_k is given in definition 2.

The proof of theorem 3 requires us to first prove two lemmas before we can conclude that the theorem is valid. We now introduce some notation to state these lemmas concisely. Since we have assumed that $\{A_j\}$ is $2k$-smooth, theorem 2 implies that the difference between $U(\mu+\Delta\lambda,\mu)$ and $U_k(\mu+\Delta\lambda,\mu)$ is $O(\Delta\lambda)^{2k+1}$. Then using this fact, we know that we only need to compare the terms of $O(\Delta\lambda^{2k+1})$ to bound the difference between U and U_k. We introduce the following notation to denote only those terms that do not necessarily cancel.

Definition 6. *If the operator $A(\Delta\lambda)$ can be written as $A(\Delta\lambda) = \sum_{p=0}^{2k} A_p\Delta\lambda^p + R(\Delta\lambda)$ where the norm of $R(\Delta\lambda)$ is $O(\Delta\lambda)^{2k+1}$, then we define $\boldsymbol{R}_{2k}[A(\Delta\lambda)]$ to be the norm of $R(\Delta\lambda)$.*

This definition simply means that $\boldsymbol{R}_{2k}$ is the error term for a Taylor expansion to order $2k$. Then using this definition, it follows from the triangle inequality that the norm of the difference between U and U_k is at most

$$\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)] + \boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]. \tag{30}$$

Our proof of theorem 3 then follows from (30) and upper bounds that we place on $\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]$ and $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$. Our bound on $\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]$ follows directly from lemma 1, but the bound on $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$ does not. We will provide the latter upper bound in lemma 3, but first we provide definition 7 and lemma 2.

Definition 7. *Let k be a positive integer and $A(\lambda) = \sum_{j=1}^m A_j(\lambda)$; then $U_k(\mu+\Delta\lambda, \Delta\lambda)$ can be written as a product of the form $\prod_{c=1}^{2m5^{k-1}} \exp(A_{j_c}(\mu_c)\Delta\lambda_c)$. We then define X_p for $p < 2k$ to be*

$$X_p \equiv \sum_{c=1}^{2m5^{k-1}} \left\|A_{j_c}^{(p)}(\mu)\right\| \frac{(\mu_c-\mu)^p}{\Delta\lambda^p}|q_{c,2k}|, \tag{31}$$

and for $p = 2k$, we define X_{2k} to be

$$X_{2k} \equiv \sum_{c=1}^{2m5^{k-1}} \max_{\tau\in[\mu,\mu+\Delta\lambda]} \left\|A_{j_c}^{(2k)}(\tau)\right\| \frac{(\mu_c-\mu)^{2k}}{\Delta\lambda^{2k}}|q_{c,2k}|. \tag{32}$$

Here the quantity $q_{c,2k}$ is given in definition 5.

Then, using this definition, our lemma can be expressed as follows.

Lemma 2. *Let $A(\lambda) = \sum_{j=1}^m A_j(\lambda)$ and let the set $\{A_j\}$ be $2k$-smooth on the interval $[\mu, \mu+\Delta\lambda]$. Then the norm of the difference between $U_k(\mu+\Delta\lambda,\mu)$ and its Taylor series in powers of $\Delta\lambda$ truncated at order $2k$ is bounded above by*

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k}\frac{X_p}{p!}\Delta\lambda^{p+1}\right)\right]. \tag{33}$$

Proof. We begin our proof of lemma 2 by writing U_k as a product of $2m5^{k-1}$ exponentials, and use Taylor's theorem to write U_k as

$$\prod_{c=1}^{2m5^{k-1}} \exp\left[\left(\sum_{p=0}^{2k-1} \frac{A_{j_c}^{(p)}(\mu)(\mu_c-\mu)^p}{p!} + \int_t^{t_c} A_{j_c}^{(2k)}(s)\frac{(\mu_c-s)^{2k-1}}{(2k-1)!}\,\mathrm{d}s\right)\Delta\lambda_c\right]. \tag{34}$$

We introduce the terms $v_c = (\mu_c-\mu)/\Delta\lambda$ and $q_{c,k} = \frac{\Delta\lambda_c}{\Delta\lambda}$, and use them to write $U_k(\mu+\Delta\lambda,\mu)$ as

$$\prod_{c=1}^{2m5^{k-1}} \exp\left[\left(\sum_{p=0}^{2k-1} \frac{A_{j_c}^{(p)}(\mu)v_c^p\Delta\lambda^p}{p!} + G_{j_c}^{(2k)}(\Delta\lambda)\right)|q_{c,k}|\Delta\lambda\right], \tag{35}$$

where we have defined

$$G_{j_c}^{(2k)}(\Delta\lambda) \equiv \int_0^{v_c} A_{j_c}^{(2k)}(\mu+x\Delta\lambda)\frac{(v_c-x)^{2k-1}}{(2k-1)!}(\Delta\lambda)^{2k}\,\mathrm{d}x. \tag{36}$$

We now prove the lemma by placing an upper bound on $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$. We expand this equation in powers of $\Delta\lambda$, while retaining only those terms of order $2k+1$ and higher. As mentioned previously, the lower order terms are irrelevant since theorem 2 guarantees that they cancel.

By expanding the exponentials in (35), taking the norm, using the triangle inequality, upper bounding each of the norms present in the expansion, and collecting terms again, we find that an upper bound on $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$ is

$$\boldsymbol{R}_{2k}\left(\exp\left[\sum_{c=1}^{2m5^{k-1}}\left(\sum_{p=0}^{2k-1} \frac{\|A_{j_c}^{(p)}(\mu)\|v_c^p\Delta\lambda^p}{p!} + G_{j_c}^{(2k)}(\Delta\lambda)\right)|q_{c,k}|\Delta\lambda\right]\right). \tag{37}$$

Note that

$$G_{j_c}^{(2k)}(\Delta\lambda) \leqslant \max_{\tau\in[\mu,\mu+\Delta\lambda]}\left\|A_{j_c}^{(2k)}(\tau)\right\|\frac{(v_c\Delta\lambda)^{2k}}{(2k)!}. \tag{38}$$

This equation can be simplified by substituting the constants X_p into it. These constants are introduced in definition 7. After this substitution our upper bound becomes

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k}\frac{X_p}{p!}\Delta\lambda^{p+1}\right)\right]. \tag{39}$$

□

We use lemma 2 to provide an upper bound on the sum of the norm of all terms in the Taylor expansion of $U_k(\mu+\Delta\lambda,\mu)$ which are of order $2k+1$ or higher. This bound is given in the following lemma.

Lemma 3. *Let $A(\lambda) = \sum_{j=1}^m A_j(\lambda)$, where the set $\{A_j\}$ is Λ-$2k$-smooth on the interval $[\mu,\mu+\Delta\lambda]$, and let $2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda \leqslant \frac{1}{2}$. Then the norm of the difference between $U_k(\mu+\Delta\lambda,\mu)$ and its Taylor series in $\Delta\lambda$ truncated at order $2k$ is upper bounded by*

$$\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)] \leqslant 2(2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda)^{2k+1}. \tag{40}$$

Proof. To simplify the following discussion, we introduce Γ_{2k}, defined by

$$\Gamma_{2k} \equiv \max_{p=0,\ldots,2k} X_p^{1/(p+1)}. \tag{41}$$

We first find an upper bound on Γ_{2k}. Now, by definition 7,

$$X_p \leqslant \sum_{c=1}^{2m5^{k-1}} |q_{c,k}| \left(\frac{t_c - t}{\Delta\lambda}\right)^p \max_{\tau\in[\mu,\mu+\Delta\lambda]} \left\| A_{j_c}^{(p)}(\tau) \right\|. \tag{42}$$

It follows from definition 5 that $q_{c,k} \leqslant Q_k$ for all c. Also, for a kth-order Lie–Trotter–Suzuki product formula, $\frac{t_c-t}{\Delta\lambda} \leqslant 1$. Using these two bounds in the above inequality, and using the fact that each element of $\{A_j\}$ occurs $2(5^{k-1})$ times in U_k, we obtain

$$X_p \leqslant 2(5)^{k-1} Q_k \sum_{j=1}^{m} \max_{\tau\in[\mu,\mu+\Delta\lambda]} \left\| A_j^{(p)}(\tau) \right\|. \tag{43}$$

Since we assume that $\{A_j\}$ is Λ-$2k$-smooth, $\sum_{j=1}^m \left\| A_j^{(p)}(\tau) \right\| \leqslant \Lambda^{p+1}$, so $X_p^{1/(p+1)} \leqslant (2(5)^{k-1}Q_k)^{1/(p+1)}\Lambda$. We also show in appendix A that $Q_k \geqslant \frac{1}{2}\frac{1}{3^{k-1}}$, implying that $2(5)^{k-1}Q_k \geqslant 1$, and thus that $X_p^{1/(p+1)} \leqslant 2(5)^{k-1}Q_k\Lambda$. Since this upper bound holds for all $0 \leqslant p \leqslant 2k$, we conclude that

$$\Gamma_{2k} \leqslant 2(5)^{k-1} Q_k \Lambda. \tag{44}$$

To show the main inequality in lemma 3, we expand the formula

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k} \frac{X_p}{p!} \Delta\lambda^{p+1}\right)\right], \tag{45}$$

in powers of $\Delta\lambda$. We then use the upper bounds $\frac{1}{p!} \leqslant \frac{(\sqrt{2})^{p+1}}{p+1}$ and $X_p \leqslant \Gamma_{2k}^{p+1}$ to simplify the expanded products. Finally we rewrite the resulting expansion as an exponential to find that

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k} \frac{X_p}{p!} \Delta\lambda^{p+1}\right)\right] \leqslant \boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k} \frac{(\sqrt{2}\Gamma_{2k}\Delta\lambda)^{p+1}}{p+1}\right)\right]. \tag{46}$$

Since X_p is a positive number, then so is Γ_{2k}, and hence the above expression is upper bounded by

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{\infty} \frac{(\sqrt{2}\Gamma_{2k}\Delta\lambda)^{p+1}}{p+1}\right)\right]. \tag{47}$$

Using the Taylor expansion of $\ln(1-x)$, we rewrite this as

$$\boldsymbol{R}_{2k}[\exp(-\ln(1-\sqrt{2}\Gamma_{2k}\Delta\lambda))] = \boldsymbol{R}_{2k}\left[\frac{1}{1-\sqrt{2}\Gamma_{2k}\Delta\lambda}\right] = \sum_{p=2k+1}^{\infty} (\sqrt{2}\Gamma_{2k}\Delta\lambda)^p. \tag{48}$$

Provided $\sqrt{2}\Gamma_{2k}\Delta\lambda \leqslant \frac{1}{2}$, this is upper bounded by $2(\sqrt{2}\Gamma_{2k}\Delta\lambda)^{2k+1}$. Using inequality (44) in equation (48) then gives

$$\boldsymbol{R}_{2k}\left[\exp\left(\sum_{p=0}^{2k} \frac{X_p}{p!} \Delta\lambda^{p+1}\right)\right] \leqslant 2(2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda)^{2k+1}. \tag{49}$$

The lemma follows by applying lemma 2. □

Now that we have proven lemma 3 we have an upper bound on $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$. We now use this upper bound to prove theorem 3.

Proof of Theorem 3. Our proof of theorem 3 begins by using the fact that

$$\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)-U_k(\mu+\Delta\lambda,\mu)]\leqslant \boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]+\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]. \quad (50)$$

We then place an upper bound on $\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]$ using lemma 1. Using the notation of lemma 1, we write the Taylor series of $U(\mu+\Delta\lambda,\mu)$ as $\sum_p T_p(\mu)\Delta\lambda^p/p!$. We then use the assumption that $\|U(\mu+\Delta\lambda,\mu)\|$ is less than one, to show from lemma 1 that $\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]$ is at most

$$\frac{\max_{\lambda\in[\mu,\mu+\Delta\lambda]}\|T_{2k+1}(\lambda)\|\Delta\lambda^{2k+1}}{(2k+1)!}. \quad (51)$$

Then using the recursive relations in lemma 1, it follows that T_{2k+1} can be written as a sum of $(2k+1)!$ terms that are each products of A and its derivatives. Then since $\{A_j : j=1,\ldots,m\}$ is Λ-$2k$-smooth and $A=\sum_{j=1}^m A_j$, it follows from definition 4 that $\|A^{(p)}(\lambda)\|\leqslant\Lambda^p$ for all λ in the interval $[\mu,\mu+\Delta\lambda]$. Then it can be verified that each term in $T_{2k+1}(\lambda)$ must have a norm that is less than Λ^{2k+1}. Therefore, it follows that $\|T_{2k+1}(\lambda)\|\leqslant(2k+1)!\Lambda^{2k+1}$ and hence

$$\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]\leqslant(\Lambda\Delta\lambda)^{2k+1}. \quad (52)$$

Using (50) and lemma 3, we see that if $2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda\leqslant 1/2$, then an upper bound on the sum of $\boldsymbol{R}_{2k}[U(\mu+\Delta\lambda,\mu)]$ and $\boldsymbol{R}_{2k}[U_k(\mu+\Delta\lambda,\mu)]$ is

$$(\Lambda\Delta\lambda)^{2k+1}+2[2\sqrt{2}(5)^{k-1}Q_k\Lambda\Delta\lambda]^{2k+1}. \quad (53)$$

We then replace this upper bound with the following simpler upper bound

$$2[3(5)^{k-1}Q_k\Lambda\Delta\lambda]^{2k+1}. \quad (54)$$

This is the claim in theorem 3, and hence we have proven the theorem. □

The error bound in theorem 3 is vital to our remaining work because it provides us with an upper bound on the error invoked by approximating an ordered operator exponential by $U_k(\mu+\Delta\lambda,\mu)$ if $\Delta\lambda$ is short. We will now show a method to devise accurate approximations to the ordered operator exponential $U(\mu+\Delta\lambda,\mu)$, even if $\Delta\lambda$ is not short. Our approach is similar to that used by Berry *et al* in [19] and that used by Suzuki in [14]; we split the ordered exponential into a product of ordinary exponentials, each of which has a short duration. To do so, we need to present a method to relate the error invoked by using one U_k to the error invoked by using a product of them. This result is provided in the following lemma.

Lemma 4. *If $\|A_p-B_p\|\leqslant\delta/P$, where δ is a positive number less than $1/2$, and $\|A_p\|\leqslant 1$ for every $p\in\{1,2,\ldots,P\}$, then $\left\|\prod_{p=1}^P A_p-\prod_{p=1}^P B_p\right\|\leqslant 2\delta$.*

Proof. Our proof begins by assuming that there exists some integer q such that

$$\left\|\prod_{p=1}^q A_p-\prod_{p=1}^q B_p\right\|\leqslant\frac{q\delta\left(1+\frac{\delta}{P}\right)^{q-1}}{P}. \quad (55)$$

We then prove lemma 4 by using induction on q. The proof of the base case follows from $\|A_p-B_p\|\leqslant\delta/P$. We then begin to prove the induction step by noting that from $\|A_p-B_p\|\leqslant\delta/P$ there exists an operator C with a norm at most one, such that $B_{q+1}=A_{q+1}+(\delta/P)C$. Then, by making this substitution and using the triangle inequality, it follows that

$$\left\|\prod_{p=1}^{q+1}A_p-\prod_{p=1}^{q+1}B_p\right\|\leqslant\left\|A_{q+1}\left(\prod_{p=1}^{q}A_p-\prod_{p=1}^{q}B_p\right)\right\|+(\delta/P)\left\|C\prod_{p=1}^{q}B_p\right\|. \quad (56)$$

Then, because $\|A_p\| \leqslant 1$ and $\|B_p\| \leqslant 1+\delta/P$, it can be verified using our induction hypothesis that the left-hand side of equation (56) is bounded above by

$$\frac{(q+1)\delta\left(1+\frac{\delta}{P}\right)^q}{P}. \tag{57}$$

This proves our induction step, and so it follows that $\left\| \prod_{p=1}^{P} A_p - \prod_{p=1}^{P} B_p \right\| \leqslant \delta(1+\delta/P)^{P-1}$ by using induction on q until $q = P$. The proof of the lemma then follows from the fact that if $\delta \leqslant 1/2$, then $(1+\delta/P)^{P-1} \leqslant 2$. □

Using lemma 4, we can now place an upper bound on the error for decompositions with longer $\Delta\lambda$.

Lemma 5. *If*

(i) $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$ *is* $2k$*-smooth on the interval* $[\mu, \mu+\Delta\lambda]$,
(ii) $\max_{x>y} \|U(x,y)\| \leqslant 1$
(iii) $\epsilon \leqslant 3Q_k(5)^{k-1}\Lambda\Delta\lambda$, *where* Q_k *is given in definition 6,*
(iv) r *is a positive integer greater than*

$$\frac{2(3Q_k(5)^{k-1}\Lambda\Delta\lambda)^{1+1/2k}}{\epsilon^{1/2k}}, \tag{58}$$

then we obtain that

$$\left\| U(\mu+\Delta\lambda, \mu) - \prod_{q=1}^{r} U_k(\mu+q\Delta\lambda/r, \mu+(q-1)\Delta\lambda/r) \right\| \leqslant \epsilon, \tag{59}$$

where U_k *is the* kth*-order Lie–Trotter–Suzuki product formula, which we introduced in definition 2.*

Proof. We find using the bound $\epsilon \leqslant 3Q_k(5)^{k-1}\Lambda\Delta\lambda$ and (58) that $3Q_k(5)^{k-1}\Lambda\Delta\lambda/r \leqslant 1/2$. Hence, we can use theorem 3 to obtain, for each $q = 1, \ldots, r$,

$$\begin{aligned}\|U(\mu+q\Delta\lambda/r, \mu+(q-1)\Delta\lambda/r) - U_k(\mu+q\Delta\lambda/r, \mu+(q-1)\Delta\lambda/r)\| \\ \leqslant 2(3Q_k(5)^{k-1}\Lambda\Delta\lambda/r)^{2k+1}.\end{aligned} \tag{60}$$

We then rewrite this bound as

$$2(3Q_k(5)^{k-1}\Lambda\Delta\lambda/r)^{2k+1} = \frac{2}{r}\frac{(3Q_k(5)^{k-1}\Lambda\Delta\lambda)^{2k+1}}{r^{2k}}. \tag{61}$$

Then from (58) we can see that, because $r \geqslant 2(3Q_k(5)^{k-1}\Lambda\Delta\lambda)^{1+1/2k}/\epsilon^{1/2k}$, it follows that

$$2(3Q_k(5)^{k-1}\Lambda\Delta\lambda/r)^{2k+1} \leqslant \frac{2}{r}\left(\frac{\epsilon}{2^{2k}}\right). \tag{62}$$

Then, since $k \geqslant 1$, it follows that

$$\|U(\mu+(q-1)\Delta\lambda/r, \mu+q\Delta\lambda/r) - U_k(\mu+(q-1)\Delta\lambda/r, \mu+q\Delta\lambda/r)\| \leqslant \frac{\epsilon}{2r}. \tag{63}$$

Then since both ϵ and $\max_{x>y} \|U(x,y)\|$ are less than one, the result of this lemma follows from lemma 4. □

Lemma 5 shows that if the maximum value of the norm of U is one, then a product of kth-order Lie–Trotter–Suzuki formulae converges to U as r increases. Furthermore, we can also use this result to prove theorem 1 by using the value of r from this lemma and multiplying it by the number of exponentials in each U_k to find a bound on the number of exponentials that are needed to approximate $U(\mu + \Delta\lambda, \mu)$. This proof is presented below.

Proof of Theorem 1. It can be verified from the definition of $U_k(\mu + \Delta\lambda, \mu)$ in theorem 2 that there are at most $2m5^{k-1}$ exponentials in each U_k, and since at most r different U_k are needed to approximate U within an error of ϵ, then if the conditions of lemma 5 are satisfied, it follows from the lemma that the number of exponentials used to decompose $U(\mu + \Delta\lambda, \mu)$ is at most

$$N \leqslant 2m5^{k-1}r \leqslant 2m5^{k-1}\left\lceil \frac{2(3(5)^{k-1}\Lambda Q_k\Delta\lambda)^{1+1/2k}}{\epsilon^{1/2k}}\right\rceil. \tag{64}$$

We then use the fact from appendix A that $Q_k \leqslant 2k/3^k$, and the fact that $k^{1/2k} < 1.3$, to show that

$$N \leqslant \left\lceil 3m\Lambda\Delta\lambda k\left(\frac{25}{3}\right)^k\left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)^{1/2k}\right\rceil. \tag{65}$$

It then follows from our bounds on Q_k that the requirements on ϵ in lemma 5 are satisfied if $\epsilon \leqslant (9/10)(5/3)^k\Lambda\Delta\lambda$, which then completes our proof of the theorem. □

Theorem 1 provides an upper bound on the number of exponentials that are needed to decompose an ordered operator exponential using a product of kth-order Lie–Trotter–Suzuki product formula, while guaranteeing that the approximation error is at most ϵ. In the following section we present a formula that provides a reasonable value of k, for a particular set of values for ϵ, Λ and $\Delta\lambda$. Furthermore, we show that if $\{A_j\}$ is Λ-∞-smooth and that formula for k is used, then the number of exponentials used scales almost linearly with $\Delta\lambda$.

7. Almost linear scaling

Reference [19] shows that there exist operator exponentials that, when decomposed into a sequence of N exponentials, require that N scale at least linearly with $\Delta\lambda$ for large $\Delta\lambda$. This implies that any decomposition method that does not use any special properties of the operator being exponentiated will also require that N scale at least linearly with $\Delta\lambda$. We now show that if there exists a Λ such that the set of operators $\{A_j\}$ is Λ-∞-smooth on $[\mu, \mu + \Delta\lambda]$ for every $\Delta\lambda > 0$, then we can choose k such that N scales almost linearly in $\Delta\lambda$. Specifically, we show that $N/\Delta\lambda$ is sub-polynomial in $\Delta\lambda$, i.e. that $\lim_{\Delta\lambda\to\infty}\frac{N}{\Delta\lambda^{1+d}} = 0$ for all constants $d > 0$, provided that $\max_{x>y}\|U(x, y)\| \leqslant 1$.

Theorem 1 implies that

$$N \leqslant \left\lceil 3m\Lambda\Delta\lambda k\left(\frac{25}{3}\right)^k\left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)^{1/2k}\right\rceil. \tag{66}$$

We set

$$k_0 = \left\lceil\sqrt{\frac{1}{2}\log_{25/3}\left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)}\right\rceil, \tag{67}$$

so that $\left(\frac{25}{3}\right)^{k_0} \geqslant \left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)^{1/2k_0}$. Then

$$\frac{N}{\Delta\lambda} \leqslant 3m\Lambda k_0\left(\frac{25}{3}\right)^{2k_0} = 3m\Lambda k_0\exp\left(k_0 2\ln\left(\frac{25}{3}\right)\right), \tag{68}$$

which is sub-polynomial in $\Delta\lambda$, though not poly-logarithmic in $\Delta\lambda$. In conclusion, if $\{A_j\}$ is Λ-∞-smooth and we choose $k = k_0$, then the number of exponentials needed to decompose an ordered operator exponential of A using the kth-order Lie–Trotter–Suzuki formula scales almost linearly in $\Delta\lambda$.

This choice of k_0 will cause N to scale nearly linearly with $\Delta\lambda$ if $\{A_j\}$ is ∞-smooth; however, if $\{A_j\}$ is only $2P$-smooth for some positive integer P, then we do not expect this because theorem 1 cannot be used if $k_0 > P$. Hence, a reasonable choice of k_0 is

$$k_0 = \min\left\{P, \left\lceil \sqrt{\frac{1}{2}\log_{25/3}\left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)}\right\rceil\right\}. \tag{69}$$

The choice of k_0 in (69) does not allow for near-linear scaling of N with $\Delta\lambda$, but it does cause N to be proportional to $\Delta\lambda^{1+1/(2P)}$ in the limit of large $\Delta\lambda$, and causes N to have the same scaling with $\Delta\lambda$ that a Λ-∞-smooth $\{A_j\}$ would have if $\Delta\lambda$ is sufficiently short.

In this section we require that $\max_{x>y}\|U(x,y)\| \leqslant 1$ for this near-linear scaling result to hold, but if this inequality does not hold then U can be normalized to ensure that it does, so it may seem that this result is more general than we claim. However, we note in appendix B that the un-normalized error in the decomposition of the un-normalized U can vary exponentially with $\Delta\lambda$. As a result, we can only guarantee that the value of N needed to ensure that $\|U - \tilde{U}\| \leqslant \epsilon$ can be chosen to scale near linearly with $\Delta\lambda$ if $\max_{x>y}\|U(x,y)\| \leqslant 1$.

8. Conclusions

We have shown how to use high-order integrators to accurately approximate ordered operator exponentials, shown what order of integrator is possible based on the smoothness of the operator and explicitly bounded the error. Our method is based on the recursive approximation building technique developed by Suzuki [10]. In contrast to Suzuki's work on ordered exponentials, we do not use a time-displacement operator. The time-displacement operator is problematic because it is a different type of operator than that for which Suzuki integrators were originally developed. It is therefore unclear from Suzuki's analysis what conditions are needed to ensure validity of the results. Our results show that Suzuki's approach is only applicable if the operator is sufficiently smooth, and we have presented an example where Suzuki's approach breaks down. We have rigorously shown what conditions are sufficient to ensure a given order of the integrator, to address such cases. In addition, we have placed upper bounds on the error, thus showing how the error scales in all relevant quantities.

If the operator to be exponentiated is $A(\lambda) = \sum_{j=1}^{m} A_j(\lambda)$, and the norm of the $2k$th derivative of each $A_j(\lambda)$ is bounded at every point in the interval $[\mu, \mu+\Delta\lambda]$, then our method can be used to build an approximation formula for the ordered exponential of $A(\lambda)$ with an error that is at most proportional to $\Delta\lambda^{2k+1}$. If this condition is not satisfied, then our method can fail as we show in figure 1. This failure results from the derivatives of $A(\lambda)$ becoming large near $\lambda = 0$. This suggests that the reason why sixth-order formulae have been observed to fail when applied to the Coulomb problem [27] (with λ taking the role of position) may be that the derivatives of $A(\lambda)$ become significant near the singularity even if the singularity is outside the interval $[\mu, \mu+\Delta\lambda]$.

We have also shown that, if the above conditions are met, $\max_{x>y}\|U(x,y)\| \leqslant 1$ and ϵ is sufficiently small, then the ordered operator exponential of $A(\lambda)$ may be approximated with total error of at most ϵ using

$$N \leqslant \left\lceil 3m\Lambda\Delta\lambda k\left(\frac{25}{3}\right)^{k}\left(\frac{\Lambda\Delta\lambda}{\epsilon}\right)^{1/2k}\right\rceil \tag{70}$$

exponentials. If the $\{A_j\}$ are smooth, with the higher order derivatives suitably bounded, then the value of k may be taken to be arbitrarily large, yielding a number of exponentials that scales nearly linearly with $\Delta\lambda$. It has been shown that sub-linear scaling with $\Delta\lambda$ is impossible for a generic $A(\lambda)$ [19, 28], so in this sense our scheme is nearly optimal for analytic $A(\lambda)$.

One extension of this work will be to develop methods of adapting the step sizes. This is likely to provide substantial improvements in some cases because the step sizes used depend on the upper bounds on the norms of the derivatives. For example, in cases where there is a divergence in the derivatives (as in figure 1), it would be useful to reduce the step size closer to the divergence. It may also be useful to adapt the order of the integrator in such cases because the derivatives are bounded away from the singular point, enabling higher order integrators to be used.

Acknowledgments

NW would like to thank A Rezakhani for many helpful discussions. We would also like to acknowledge the following agencies that have generously funded this research, the MITACS research network, the Canadian Institute for Advanced Research (CIFAR) of which PH is a Scholar and BCS is an Associate, the Natural Sciences and Engineering Research Council of Canada, the Informatics Circle of Research Excellence and the Australian Research Council.

Appendix A. Derivation of bounds on Q_k

When proving bounds on the error introduced by our decomposition of an ordered exponential in section 6, we use the quantities Q_k defined by

$$Q_k = \tfrac{1}{2}\max\{s_1, |1-4s_1|\}\max\{s_2, |1-4s_2|\}\cdots\max\{s_{k-1}, |1-4s_{k-1}|\} \tag{A.1}$$

for $k > 1$ and $Q_1 = \frac{1}{2}$, where

$$s_k = \frac{1}{4-4^{1/(2k+1)}} \tag{A.2}$$

for all $k \geqslant 1$. We now show that Q_k decreases exponentially in k,

$$\frac{3}{2}\frac{1}{3^k} \leqslant Q_k \leqslant \frac{2k}{3^k}. \tag{A.3}$$

The lower bound follows directly by noting that $s_k \geqslant \frac{1}{3}$ for all $k \geqslant 1$.

Set $a = 2\ln(4)$, which is approximately 2.7726. Using that $-x \geqslant \ln(1-x)$ for $0 \leqslant x < 1$, we then have that for $k \geqslant 1$,

$$\frac{-1}{2k+1}\ln(4) = \frac{-a}{2(2k+1)} \geqslant \ln\left(1-\frac{a}{2(2k+1)}\right). \tag{A.4}$$

Taking exponentials on either side yields

$$4^{-1/(2k+1)} \geqslant 1 - \frac{a}{2(2k+1)}. \tag{A.5}$$

Multiplying by 4 and subtracting 1 on either side gives,

$$4^{2k/(2k+1)} - 1 \geqslant 3\left(1-\frac{2a}{3(2k+1)}\right), \tag{A.6}$$

and taking reciprocals then yields

$$4s_k - 1 = \frac{1}{4^{2k/(2k+1)}-1} \leqslant \frac{1}{3}\frac{k+\frac{1}{2}}{k+\frac{1}{2}-\frac{a}{3}} \leqslant \frac{1}{3}\frac{k+\frac{1}{2}}{k+\frac{1}{2}-1}. \tag{A.7}$$

Noting that $4s_k - 1 \geqslant s_k$ since $s_k \geqslant \frac{1}{3}$, and using that $s_1 \leqslant \frac{2}{3}$, we conclude that

$$Q_k \leqslant \frac{1}{2}\frac{2}{3}\frac{1}{3^{k-2}}\frac{k-\frac{1}{2}}{\frac{3}{2}} \leqslant \frac{2k}{3^k} \tag{A.8}$$

for $k \geqslant 2$. By inspection the inequality $Q_k \leqslant \frac{2k}{3^k}$ also holds for $k = 1$.

Appendix B. Norms larger than 1

In this work we have restricted the norm of $U(\eta, \mu)$ to not exceed 1. This means that the eigenvalues of $A(\lambda)$ can have no positive real part. In the case where they do, then the analysis can be performed in the following way. Simply define the new operators

$$A'(\lambda) \equiv A(\lambda) - \kappa(\lambda)\mathbb{1}, \tag{B.1}$$

$$A'_j(\lambda) \equiv A_j(\lambda) - (\kappa(\lambda)/m)\mathbb{1} \tag{B.2}$$

for some $\kappa(\lambda)$ such that the eigenvalues of $A'(\lambda)$ have no positive real part. Note that the prime does not denote a derivative here. Then the result we have given in theorem 1 will hold for A' and $\{A'_j\}$ (provided we also define Λ in terms of these operators). The difference between A and A' simply corresponds to a normalization factor, i.e.

$$U(\eta, \mu) = U'(\eta, \mu)\, \mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x}. \tag{B.3}$$

We can simply use the Lie–Trotter–Suzuki formula to approximate $U'(\eta, \mu)$, which gives

$$\begin{aligned} U(\eta, \mu) &\approx \mathrm{e}^{\int_\mu^\eta \kappa(x)\mathrm{d}x} \prod_{i=1}^{N} \exp(A'_{j_i}(\Delta\lambda_i)\Delta\lambda_i) \\ &= K \prod_{i=1}^{N} \exp(A_{j_i}(\lambda_i)\Delta\lambda_i), \end{aligned} \tag{B.4}$$

where K is a normalization correction

$$K = \mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x - \sum_{i=1}^{N} \kappa(\lambda_i)\Delta\lambda_i}. \tag{B.5}$$

Thus, the same series of exponentials can be used, except for a normalization factor. There is a difference in the final error that can be obtained because

$$\left\| U'(\eta, \mu) - \prod_{i=1}^{N} \exp(A'_{j_i}(\lambda_i)\Delta\lambda_i) \right\| \leqslant \epsilon \tag{B.6}$$

implies that

$$\left\| U(\eta, \mu) - K \prod_{i=1}^{N} \exp(A_{j_i}(\lambda_i)\Delta\lambda_i) \right\| \leqslant \mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x}\epsilon. \tag{B.7}$$

It might be imagined that the *relative* error can be kept below ϵ with similar scaling of N. That is, that $\mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x}$ can be replaced with $\|U(\eta, \mu)\|$ in (B.7). Unfortunately, that is not the case. The reason is that, due to the submultiplicativity of the operator norm, $\|U(\eta, \mu)\|$ can be much smaller than $\mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x}$.

For example, consider the case where $A(\lambda)$ is initially σ_z (the Pauli operator) over an interval $\Delta\lambda/2$; then is $-\sigma_z$ over another interval $\Delta\lambda/2$. Then $U(\eta, \mu) = \mathbb{1}$, and has norm 1, but $\mathrm{e}^{\int_\mu^\eta \kappa(x)\,\mathrm{d}x} = \mathrm{e}^{\Delta\lambda}$. A small error in between the two intervals of length $\Delta\lambda/2$ can then yield a large relative error in the final result. For example, consider the error $E = \mathrm{e}^{\mathrm{i}\delta\sigma_y}$. That yields a final result

$$U(\eta, \mu + \Delta\lambda/2)EU(\mu + \Delta\lambda/2, \mu) = \begin{bmatrix} \cos\delta & \mathrm{e}^{-\Delta\lambda}\sin\delta \\ -\mathrm{e}^{\Delta\lambda}\sin\delta & \cos\delta \end{bmatrix}. \tag{B.8}$$

The error in this result scales as $\mathrm{e}^{\Delta\lambda}$, despite the final norm being small for $U(\eta, \mu)$.

With the possibility that the norm of $U(\eta, \mu)$ exceeds 1, our approach need not give scaling for N that is close to linear in $\Delta\lambda$. In the lower bound on N in theorem 1, the $(1/\epsilon)^{1/2k}$ will be replaced with

$$(1/\epsilon)^{1/2k}\,\mathrm{e}^{\frac{1}{2k}\int_\mu^\eta \kappa(x)\,\mathrm{d}x}. \tag{B.9}$$

To prevent this term scaling exponentially in $\Delta\lambda$, one would need to take k proportional to $\Delta\lambda$. However, this would result in $(25/3)^k$ scaling exponentially in $\Delta\lambda$. As a result, it does not appear to be possible to obtain sub-exponential scaling if there is no bound on the norm of $U(\eta, \mu)$.

References

[1] Suzuki M 2003 *J. Stat. Phys.* **110** 945–56
[2] De Raedt H, Michielsen K, Hams A, Miyashita S and Saito K 2002 *Eur. Phys. J.* B **27** 15–28
[3] Lloyd S 2000 *Phys. Rev.* A **62** 022108
[4] Wilson K 1974 *Phys. Rev.* D **10** 2445–59
[5] Magnus W 1954 *Commun. Pure Appl. Math.* **7** 649–73
[6] Baye D, Goldstein G and Capel P 2003 *Phys. Lett.* A **317** 337–42
[7] Chin S A 2002 *J. Chem. Phys.* **117** 1409–15
[8] Trotter H 1959 *Proc. Am. Math. Soc.* **10** 545–51
[9] Forest E and Ruth D 1990 *Physica* D **43** 105–17
[10] Suzuki M 1993 *Proc. Japan Acad.* B **69** 161–6
[11] Suzuki M 1990 *Phys. Lett.* A **146** 319–23
[12] Suzuki M 1991 *J. Math. Phys.* **32** 400–7
[13] Yoshida H 1990 *Phys. Lett.* A **150** 262–8
[14] Suzuki M 1994 *Commun. Math. Phys.* **163** 491–508
[15] Huyghebaert H and De Raedt H 1990 *J. Phys. A: Math. Gen.* **23** 5777–93
[16] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[17] Forbert H A and Chin S A 2001 *Phys. Rev.* E **63** 016703
[18] Forbert H A and Chin S A 2001 *Phys. Rev.* B **63** 144518
[19] Berry D W, Ahokas G, Cleve R and Sanders B C 2007 *Commun. Math. Phys.* **270** 359–71
[20] Brown K, Harrow A and Chuang I L 2004 *Phys. Rev.* A **70** 052318
[21] Ricci A and Ciccotti G 2003 *Mol. Phys.* **101** 1927–31
[22] Khalili M, Liwo A, Jagielska A and Scheraga H A 2005 *J. Phys. Chem.* B **109** 13798–810
[23] De Fabritiisa G, Serrano M, Espanõl P and Coveney P V 2006 *Physica* A **361** 429-40
[24] Sergi A and Kapral R 2003 *J. Chem. Phys.* **119** 12776
[25] Chin S A 1997 *Phys. Lett.* A **226** 344–8
[26] Kole J, M Figge M and De Raedt H 2002 *New Unconditionally Stable Algorithms to Solve the Time-Dependent Maxwell Equations* (*Lecture Notes in Computer Science*) (Berlin: Springer) pp 803–12
[27] Chin S A 2006 *J. Chem. Phys.* **124** 054106
Chin S A 2006 *J. Chem. Phys.* **270** 359–71
[28] Childs A M and Kothari R 2009 Limitations on the simulation of non-sparse Hamiltonians arXiv:0908.4398