Fully efficient time-parallelized quantum optimal control algorithm

M. K. Riahi,[1] J. Salomon,[2] S. J. Glaser,[3] and D. Sugny[4,5,*]

[1]*Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 07102, USA*

[2]*CEREMADE, Université Paris Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France*

[3]*Department of Chemistry, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany*

[4]*Laboratoire Interdisciplinaire Carnot de Bourgogne (ICB), UMR 5209 CNRS-Université de Bourgogne, 9 Avenue A. Savary, BP 47 870, F-21078 Dijon Cedex, France*

[5]*Institute for Advanced Study, Technische Universität München, Lichtenbergstrasse 2 a, D-85748 Garching, Germany*

(Received 28 February 2016; published 13 April 2016)

We present a time-parallelization method that enables one to accelerate the computation of quantum optimal control algorithms. We show that this approach is approximately fully efficient when based on a gradient method as optimization solver: the computational time is approximately divided by the number of available processors. The control of spin systems, molecular orientation, and Bose-Einstein condensates are used as illustrative examples to highlight the wide range of applications of this numerical scheme.

DOI: 10.1103/PhysRevA.93.043410

I. INTRODUCTION

The general goal of quantum control is to actively manipulate dynamical processes at the atomic or molecular scale [1,2]. In recent years, the advances in quantum control have emerged through the introduction of appropriate and powerful tools coming from mathematical control theory and by the use of sophisticated experimental techniques to shape the corresponding control fields [3–6]. In this framework, different numerical optimal control algorithms [7–9] have been developed and applied to a large variety of quantum systems. Optimal control was used in physical chemistry in order to steer chemical reactions [3], but also for spin systems [10,11] with applications in nuclear magnetic resonance [7,12–16] and magnetic resonance imaging [17–19]. Recently, optimal control has attracted attention in view of applications to quantum information processing, for example as a tool to implement high-fidelity quantum gates in minimum time [4,20,21]. Generally, algorithms can also be designed to account for experimental imperfections or constraints related to a specific material or device [4]. The possibility of including such constraints renders optimal control theory more useful in view of experimental applications and helps bridge the gap between control theory and control experiments.

The standard numerical optimal control algorithms based on an iterative procedure compute the control fields through many time propagations of the state of the system, which can be prohibitive for systems of large dimensions in terms of computational time. This numerical limit can be bypassed by making use of parallel computing [22–25]. In the case where the computational time is divided by the number of computers, the method is said to be fully efficient. This full efficiency can be viewed as the physical limit in terms of performance of a parallel algorithm. While in applied mathematics different techniques have been developed using space or time decomposition [22,24], very little has been done in quantum mechanics [25,26]. Note that quantum control computations can also be speeded up by the parallelization of matrix exponential algorithms [27,28] and by parallelizing density operator time evolutions using minimal sets of pure states [29].

This paper is not aimed at proposing a new optimization approach, but rather at describing and studying a general framework, namely, the intermediate state method (ISM), introduced in [24], which uses a time parallelization to speed up the computation of optimal control fields. We investigate the efficiency of ISM on three benchmark quantum control problems, ranging from the control of coupled spin systems and the control of molecular orientation to the control of Bose-Einstein condensates. As a byproduct, we show under which conditions ISM can be made fully efficient.

The paper is organized as follows. Section II is dedicated to the description of the time-parallelization method. The numerical schemes involved in this algorithm are defined in Sec. III. Numerical results on the control of spin systems, molecular orientation, and Bose-Einstein condensates are presented in Sec. IV. Conclusion and prospective views are given in Sec. V.

II. THE TIME-PARALLELIZATION METHOD

We first introduce the optimal control problem and we derive the corresponding optimality conditions. We consider pure quantum states and we assume that the time evolution is coherent. Note that the formalism can be easily extended to mixed states or to the control of evolution operators [30]. The control process is aimed at maximizing the transfer of population onto a target state, but modification of the algorithms in view of optimizing the expectation value of an observable is straightforward. The dynamics of the quantum system is governed by the Hamiltonian H. The initial and target states are denoted by $|\psi_i\rangle$ and $|\psi_f\rangle$, respectively, and the general state of the system at time t is denoted by $|\psi(t)\rangle$. The dynamics of the quantum system is governed by the Schrödinger equation,

$$i\,\partial_t|\psi(t)\rangle = H(u(t))|\psi(t)\rangle, \tag{1}$$

where $u(t)$ is the field to be determined. The control time T is fixed. The objective of the control problem is to maximize the

*dominique.sugny@u-bourgogne.fr

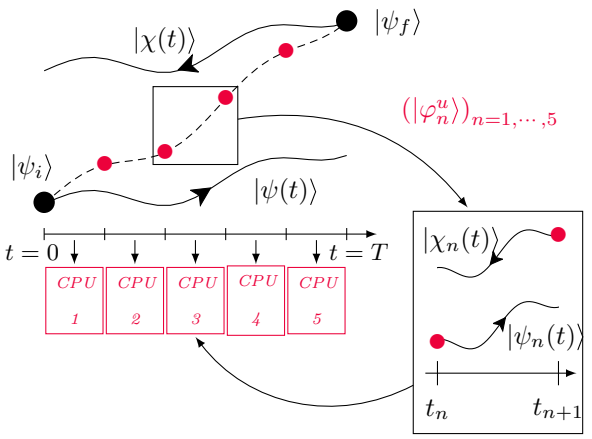

FIG. 1. Schematic description of the intermediate state method (see the text for details).

figure of merit $\mathcal{J}$,

$$\mathcal{J}[u] = \mathrm{Re}[\langle\psi(T)|\psi_f\rangle] - \frac{\alpha}{2}\int_0^T u(t)^2 dt,$$

with α being a positive parameter which expresses the relative weight between the projection onto the target state and the energy of the control field. A necessary condition to ensure the optimality of u is given by the cancellation of the gradient of $\mathcal{J}$ with respect to u [7]:

$$\nabla\mathcal{J}[u](t) = -\alpha u(t) + \mathrm{Im}[\langle\chi(t)|\partial_{u(t)}H(u(t))|\psi(t)\rangle] = 0, \tag{2}$$

where $|\chi(t)\rangle$ is the adjoint state that satisfies

$$i\,\partial_t|\chi(t)\rangle = H(u(t))|\chi(t)\rangle, \tag{3}$$

with the final condition $|\chi(t=T)\rangle = |\psi_f\rangle$.

We now present the ISM. A schematic description is displayed in Fig. 1. The main idea consists in considering a combination of the trajectories followed by $|\psi(t)\rangle$ and $|\chi(t)\rangle$ [31]. Given $N \in \mathbb{N}$, we decompose the interval $[0,T]$ into a partition of subintervals $[0,T] = \cup_{n=0}^{N-1}[t_n,t_{n+1}]$, with $0 = t_0 < \cdots < t_N = T$. The parallelization strategy is based on this decomposition. We consider an arbitrary control u and we introduce the sequence $|\varphi^u\rangle = (|\varphi_n^u\rangle)_{n=0,\ldots,N}$ that interpolates the state and adjoint state trajectories at time t_n as follows:

$$|\varphi_n^u\rangle = \frac{T-t_n}{T}|\psi(t_n)\rangle + \frac{t_n}{T}|\chi(t_n)\rangle, \tag{4}$$

where $|\psi(t)\rangle$ and $|\chi(t)\rangle$ are defined by Eqs. (1) and (3), respectively. Note that $|\varphi^u\rangle$ does not sample any usual dynamics, e.g., $|\varphi^u\rangle$ does not correspond to a solution of Eq. (1). Its initial and final states are $|\varphi^u(0)\rangle = |\psi_0\rangle$ and $|\varphi^u(T)\rangle = |\psi_f\rangle$, respectively. The choice of intermediate states made in Eq. (4) is crucial to demonstrate Theorem 1 below [24].

We then introduce in each subinterval the optimal control problem $\max_{u_n}\mathcal{J}_n[|\varphi^u\rangle,u_n]$ defined by the maximization of the subfunctional,

$$\mathcal{J}_n[u_n,|\varphi^u\rangle] = -\frac{1}{2}\big\||\psi_n(t_{n+1})\rangle - |\varphi_{n+1}^u\rangle\big\|^2 - \frac{\alpha_n}{2}\int_{t_n}^{t_{n+1}} u_n(t)^2 dt,$$

with $0 \leqslant n \leqslant N-1$. In this problem, the state $|\psi_n\rangle$ is defined on $[t_n,t_{n+1}]$ by

$$i\,\partial_t|\psi_n(t)\rangle = H(u(t))|\psi_n(t)\rangle, \tag{5}$$

starting from $|\psi_n(t_n)\rangle = |\varphi_n^u\rangle$. The penalization coefficient is defined by $\alpha_n = \frac{t_{n+1}-t_n}{T}\alpha$. Since

$$\mathcal{J}_n[u_n,|\varphi^u\rangle] = -\frac{1}{2}|||\psi_n(t_{n+1})\rangle||^2 - \frac{1}{2}\big|\big||\varphi_{n+1}^u\rangle\big|\big|^2 + \mathrm{Re}\big[\langle\psi_n(t_{n+1})|\varphi_{n+1}^u\rangle\big] - \frac{\alpha_n}{2}\int_{t_n}^{t_{n+1}} u_n(t)^2 dt,$$

maximizing $\mathcal{J}_n$ with respect to u is equivalent to maximizing a figure of merit of the form $\mathrm{Re}[\langle\psi_n(t_{n+1})|\varphi_{n+1}^u\rangle] - \frac{\alpha_n}{2}\int_{t_n}^{t_{n+1}} u_n(t)^2 dt$. In this way, each subproblem has the same structure as the initial one.

We now review some properties of the time decomposition in order to establish the relation with the original optimal control problem. Given an arbitrary trajectory $|\varphi(t)\rangle$, we define an auxiliary figure of merit,

$$\mathcal{J}_\parallel[u,|\varphi\rangle] = \sum_{n=0}^{N-1}\beta_n\mathcal{J}_n[u_n,|\varphi\rangle],$$

with $\beta_n = \frac{T}{t_{n+1}-t_n}$. A first relation between $\mathcal{J}_\parallel$ and $\mathcal{J}$ is given in Theorem 1 (see Ref. [24] for the proof).

Theorem 1. Given an arbitrary control u, we have

$$|\varphi^u\rangle = \mathrm{argmax}_{|\varphi\rangle}(\mathcal{J}_\parallel[u,|\varphi\rangle]).$$

Moreover, the following relation is satisfied:

$$\mathcal{J}_\parallel[u,|\varphi^u\rangle] = \mathcal{J}[u].$$

As a byproduct, this theorem allows us to compute in parallel $\mathcal{J}[u]$, knowing only the sequence $|\varphi_u\rangle$. A similar relation also holds between the gradients of the functionals, as stated in Theorem 2.

Theorem 2. Given an arbitrary control u, we have

$$\nabla\mathcal{J}[u]_{|[t_n,t_{n+1}]} = \beta_n\nabla\mathcal{J}_n[u_{|[t_n,t_{n+1}]},|\varphi^u\rangle].$$

This result provides a different interpretation of the time-parallelized method since the sequence $|\varphi^u(t_n)\rangle, n = 0,\ldots,N$ of intermediate states enables the decomposition of the computation of the gradient.

Proof. Let us consider a fixed value n, with $0 \leqslant n \leqslant N-1$, $t \in [t_n,t_{n+1}]$, and denote by $|\chi_n(t)\rangle$ and $|\psi_n(t)\rangle$ the trajectories defined by

$$i\,\partial_t|\chi_n(t)\rangle = H(u(t))|\chi_n(t)\rangle$$

and

$$i\,\partial_t|\psi_n(t)\rangle = H(u(t))|\psi_n(t)\rangle,$$

with $|\chi_n(t_{n+1})\rangle = |\varphi_{n+1}^u\rangle$ and $|\psi_n(t_n)\rangle = |\varphi_n^u\rangle$. For $t \in [t_n,t_{n+1}]$, we repeat with $\mathcal{J}_n$ the computation made to derive the gradient of $\mathcal{J}$,

$$\nabla\mathcal{J}_n\big[u_{|[t_n,t_{n+1}]},|\varphi^u\rangle\big](t) = \mathrm{Im}(\langle\chi_n(t)|\partial_{u(t)}H|\psi_n(t)\rangle) - \alpha_n u(t).$$

Using the fact that

$$|\chi_n(t)\rangle = \frac{(T-t_{n+1})}{T}|\psi(t)\rangle + \frac{t_{n+1}}{T}|\chi(t)\rangle$$

and

$$|\psi_n(t)\rangle = \frac{(T-t_n)}{T}|\psi(t)\rangle + \frac{t_n}{T}|\chi(t)\rangle,$$

we arrive at

$$\nabla \mathcal{J}_n\big[u_{|[t_n,t_{n+1}]},|\varphi^u\rangle\big](t) = \frac{1}{\beta_n}\mathrm{Im}[\langle\chi(t)|\partial_{u(t)}H|\psi(t)\rangle] - \frac{\alpha}{\beta_n}u(t),$$

and the result follows. ■

We now give the general structure of the ISM. Let $1 \geqslant \eta > 0$ and $u^{(0)}$ be an initial control field.

Algorithm 1.

(1) Set Err $= 1$, $k = 0$.

(2) While Err $> \eta$, do the following:

(a) Compute on $[0,T]$ the trajectories $|\psi^{(k)}(t)\rangle$ and $|\chi^{(k)}(t)\rangle$ associated with $u^{(k)}$ according to Eqs. (1) and (3).

(b) Compute $|\varphi^{(k)}(t)\rangle = |\varphi^{u^{(k)}}(t)\rangle$ according to Eq. (4).

(c) On each subinterval $[t_n,t_{n+1}]$, compute in parallel an approximation of the solution $u_n^{(k+1)}$ of the problem $\max_{u_n} \mathcal{J}_n[u_n,|\varphi_k\rangle]$.

(d) Define $u^{(k+1)}$ as the concatenation of the controls $u_n^{(k+1)}$, $n = 1,\ldots,N-1$.

(e) Set Err $= \sum_{n=0}^{N-1}\int_{T_n}^{T_{n+1}} \|\nabla \mathcal{J}_n[u^{(k+1)}_{|[t_n,t_{n+1}]}](t)\|dt$.

(f) Set $k = k+1$.

Step (2a) contradicts the parallelization paradigm, since it requires a sequential solving of an evolution equation on the full interval $[0,T]$. We will see how this problem can be solved in Sec. IV. However, note that the most time-consuming step, namely, step (2c), is achieved in parallel.

III. DESCRIPTION OF THE NUMERICAL METHODS USED IN THE PARALLELIZATION

Different schemes can be used to implement the time-parallelized algorithm in practice. This requires two ingredients: a numerical scheme to solve approximately the evolution equations of steps (2a) and (2c), and an optimization procedure for the subproblem of step (2c). Here, we give some details about the used numerical methods and we explain how the full efficiency can be approached in the case of quantum systems of sufficiently small dimensions.

In the different numerical examples, we consider two numerical solvers for the Schrödinger equation (1): a Crank-Nicholson scheme and a second-order Strang operator splitting. Such solvers can be described through an equidistant time-discretization grid $t_a = t_0 < t_1 < \cdots < t_J = t_b$ of an interval $[t_a,t_b]$. The time step is denoted by $\tau = (t_b - t_a)/(J-1)$ for some $J \in \mathbb{N}$. For each time grid point t_j, we introduce the state $|\psi_j\rangle$ and the control u_j, which are some approximations of the exact state $|\psi(t_j)\rangle$ and the exact control field $u(\frac{t_j+t_{j+1}}{2})$.

The Crank-Nicholson algorithm is based on the following recursive relation:

$$\frac{i}{\tau}(|\psi_{j+1}\rangle - |\psi_j\rangle) = \frac{H(u_j)}{2}(|\psi_{j+1}\rangle + |\psi_j\rangle), \tag{6}$$

which can be rewritten in a more compact form as

$$(Id + L_j)|\psi_{j+1}\rangle = \big(Id - L_j\big)|\psi_j\rangle\,, \tag{7}$$

where Id is the identity operator and $L_n := i\frac{\tau}{2}H(u_n(t))$.

The second-order Strang operator splitting is rather used in the case of infinite-dimensional systems. Indeed, this method is particularly relevant when the Hamiltonian includes a differential operator. We consider, for example, the case $H(u(t)) = -\Delta + V(u(t),x)$, where Δ denotes the Laplace operator and $V(u(t)) = V(u(t),x)$ is a scalar potential. In this case, Strang's method gives rise to the iteration

$$|\psi_{j+1}\rangle = \exp\left(-\frac{i\tau}{2}\Delta\right)\exp[-i\tau V(u_j)]\exp\left(-\frac{i\tau}{2}\Delta\right)|\psi_j\rangle. \tag{8}$$

In Eq. (8), each product can be determined very quickly since the operator $V(u_j)$ is diagonal in the physical space, while Δ is generally diagonal in the Fourier space, and the change of basis can be achieved efficiently by fast Fourier transform.

These schemes provide a second-order approximation with respect to time, which leads to an accurate approximation of the trajectory $|\psi(t)\rangle$. In addition, both propagators automatically preserve the normalization of the wave function, which is very interesting to avoid nonphysical solutions. A specific advantage of these solvers is that they allow an exact differentiation with respect to the control in the discrete setting, in the case of scalar control for the Strang solver (8) and in any case for the Crank-Nicholson solver (6).

We now explain how the full efficiency can be reached with the parallelization algorithm. Both solvers lead to a linear relation between the initial and the final states of the system of the form

$$|\psi_J\rangle = M(u)|\psi_0\rangle.$$

As an example, for the Crank-Nicholson solver, we have

$$M(u) = \Pi_{j=0}^{J-1}(Id + L_j)^{-1}(Id - L_j).$$

The matrix $M(u)$ can be computed in parallel during the propagation of Eq. (7). Knowing the state $|\psi_0\rangle$, this matrix enables one to compute in one matrix-vector product $|\psi_J\rangle$. As a consequence, this propagator assembling technique allows us to avoid the sequential solving on the full interval $[0,T]$ in step (2a) of Algorithm 1. More precisely, assume for example that at iteration k of Algorithm 1 and on each subinterval, a matrix $M_n(u_n^{(k)})$ is computed and transmitted to the main processor. The computations of the sequences $|\psi(t_n)\rangle$ and $|\chi(t_n)\rangle$, which are required to define the intermediate states $|\varphi^{(k)}(t_n)\rangle$, can be achieved in $2N$ matrix-vector products. Due to storage and communications issues of the matrices, note that this approach can only be used for quantum systems of small dimensions.

We conclude this section by presenting a way to derive the gradient of time-discretized figures of merit of the form

$$\mathcal{J}_\tau[u] = \mathrm{Re}\langle\psi_J|\psi_f\rangle - \frac{\alpha}{2}\tau\sum_{j=0}^{J-1}u_j^2.$$

We consider the case of a Crank-Nicholson solver, but similar computations can be made for Strang's solver. We introduce the functional $Ł_\tau$, defined by

$$Ł_\tau[u,|\psi\rangle,|\chi\rangle] = \mathcal{J}_\tau[u] + \mathrm{Re}\left(\sum_{j=0}^{J-1}\langle\chi_j|Id + L_j|\psi_{j+1}\rangle - \langle\chi_j|Id - L_j|\psi_j\rangle\right).$$

Since L_j is anti-Hermitian, differentiating $Ł_\tau$ with respect to $|\psi\rangle$ gives rise to the discrete adjoint evolution equation

$$(Id - L_{j-1})|\chi_{j-1}\rangle = (Id + L_j)|\chi_j\rangle,$$

with the final condition $(Id - L_{J-1})|\chi_{J-1}\rangle = |\psi_J\rangle$. To derive the gradient of $\mathcal{J}_\tau[u]$, it remains to differentiate L_τ with respect to u, which leads to the jth entry of the gradient of $\mathcal{J}_\tau$,

$$(\nabla \mathcal{J}_\tau[u])_j = \alpha dt u_j + \frac{i\tau}{2}\langle \chi_j|\partial_u H|\psi_{j+1} + \psi_j\rangle.$$

In the sequel, we use this result to implement a constant step gradient method: the approximation of the solution of the subproblem in step (2d) is computed by iterating on ℓ in the formula

$$u^{\ell+1} = u^\ell - \rho \nabla \mathcal{J}_\tau(u^\ell), \tag{9}$$

for some $\rho > 0$.

Other optimization methods such as pseudo- or quasi-Newton approaches can be used to perform step (2c).

IV. NUMERICAL RESULTS

This section is dedicated to some numerical results obtained with the ISM, used with the schemes presented in Sec. III. The efficiency of this approach is illustrated on three benchmark examples in quantum control [3,4], namely, the control of a system of coupled spins, the control of molecular orientation, and the control of a Bose-Einstein condensate whose dynamics is governed by the Gross-Pitaevskii equation.

A. Control of a system of five coupled spins

Here, we consider the control of a system of coupled spin-1/2 particles. Since the principles of control in nuclear magnetic resonance are already described in Refs. [10,11,18], only a brief account will be given here in order to introduce the used model. We investigate the control of a system of coupled spins by means of different magnetic fields acting as local controls on each spin. Each field only acts on one spin and does not interact with the others, i.e., the spins are assumed to be selectively addressable.

We introduce a system of five coupled spins [32,33], the evolution of which is described by the following Hamiltonian:

$$H = H_0 + \sum_{k=1}^{5} \left[u_x^k(t) I_x^{(k)} + u_y^k(t) I_y^{(k)}\right], \tag{10}$$

where the operators $I_x^{(k)}$ and $I_y^{(k)}$ are, up to a factor, Pauli matrices which only act on the kth spin:

$$I_x := \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}, \quad I_y := \begin{pmatrix} 0 & -i/2 \\ i/2 & 0 \end{pmatrix},$$
$$I_z := \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}.$$

We assume that the free-evolution Hamiltonian H_0 is associated with a topology [32,33] defined by

$$H_0 = 2\pi\left(J_{12} I_z^{(1)} I_z^{(2)} + J_{13} I_z^{(1)} I_z^{(3)} + J_{23} I_z^{(2)} I_z^{(3)} + J_{25} I_z^{(2)} I_z^{(5)} + J_{34} I_z^{(3)} I_z^{(4)}\right).$$

Note that this model system is valid in heteronuclear spin systems if the coupling strength between the spins is small with respect to the frequency shifts [10,11]. The coupling constant between the spins is taken to be uniform and equal to $J_p = 140$. For the numerical simulations, we move to the density matrix formalism, with $I_x^{(1)}$ and $I_x^{(5)}$ as initial and final states, respectively. The control time is fixed to $T = J_p/10$. The parameter α is set to 0.

For the time parallelization, we consider a uniform grid so that $t_{n+1} - t_n = T/N$, for $n = 0, \ldots, N-1$, and we compare the results for different values of N. The time discretization is done by the Crank-Nicholson method of Eq. (6) with the time step $T/2^{15}$. In step (2c), one iteration of the constant step gradient descent method [see Eq. (9)] is used, with $\rho = 10^4$. As a result of Theorem 2, the values obtained after a given number of iterations are the same for all values of N. In this way, the method is almost fully efficient. It is actually equivalent to a standard gradient method, except that the computation of the gradient is done in parallel. The computational effort is therefore exactly divided by the number of processors, and the full efficiency is only limited by the memory usage and also by the communication between processors required by the update of the intermediate states in steps (2a) and (2b) of the algorithm. Figure 2 displays the figure of merit with respect to the parallel computational time.

In order to more precisely evaluate the efficiency of the algorithm, we give some details about the speedup of the numerical computations. Numerical simulations are implemented with MATLAB, where the parallelization is realized using the open-source library MatlabMPI [34]. The tests have been carried out on a shared memory machine under a LINUX system with a core of Intel(R) Xeon(R) CPU type (@ 2.90 GHz with 198 Giga byte shared memory). The parallel computation uses N processors where, as above, N is the number of subintervals of the time-domain decomposition. In

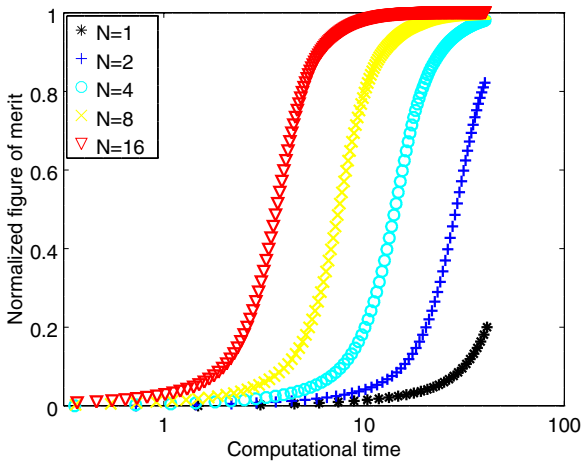

FIG. 2. Evolution of the normalized figure of merit (dimensionless) at each iteration and for different values of N (the number of processors) with respect to computational time (wall-clock time) in the case of the control of a spin system. As stated in Theorem 2, the values obtained after a given number of iterations are the same for all values of N. Logarithmic scale is used in the x axis. The computational time is expressed in seconds.

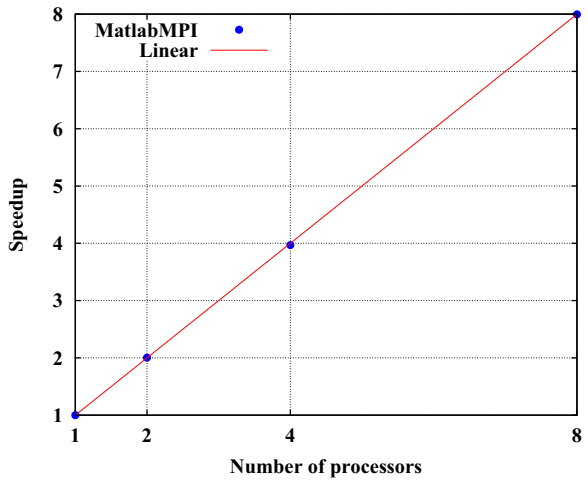

FIG. 3. Speedup $S(1.74,N)$ (dimensionless) of the parallel implementation (y axis) with respect to the number of processors N (x axis). The blue dots indicate the speedup achieved with MatlabMPI for $\varepsilon = 1.74$ (see Table I for details). The red solid line corresponds to a linear evolution of the speedup as a function of N.

Fig. 3, parallel numerical performances are compared with the sequential performance, which is obtained when a single processor is used to treat the whole time domain. Given $\varepsilon > 0$, we introduce the quantities $S(\varepsilon,N) := t(\varepsilon,1)/t(\varepsilon,N)$ and $\mathrm{Eff}(\varepsilon,N) := 10^2[S(\varepsilon,N)/N]$, as the parallel speedup and the efficiency, respectively, where $t(\varepsilon,N)$ denotes the computational time (with N processors) necessary to reach a value $\mathcal{J}[u^k]$ such that $\mathcal{J}[u^\infty] - \mathcal{J}[u^k] < \varepsilon$, where u^∞ is the value of the sequence u^k obtained at the numerical convergence. Figure 3 displays results about the speedup of the parallel implementation.

We observe that the algorithm behaves as expected when increasing the number of processors. Despite the use of input-output (IO) data files to ensure the communication between CPUs (as required by MatlabMPI), the ISM achieves a linear scalability. A profiling of the parallel computing is reported in Fig. 4, where we present the time spent to achieve the communications for the master and one slave processors during 20 iterations of the optimization process. As expected, the speedup is independent on the value of ε. This point is clearly exhibited in Table I. The communications in MatlabMPI are of the point-to-point type through IO files; a for-loop is therefore necessary to cover all sender and receiver

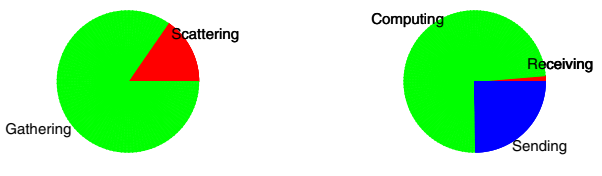

FIG. 4. MatlabMPI profiling of the tasks performed by the master and a slave processors with $N = 4$ and for 20 iterations of the optimization process. The proportions are computed separately regarding the wall-clock timing of a given processor.

TABLE I. Efficiency $\mathrm{Eff}(\varepsilon,N)$ of the parallel MathlabMPI implementation. Note the full efficiency performance through several snapshots in the parallel-running computation.

N processors	1	2	4	8
$\mathrm{Eff}(0.32,N)$	100%	100.5%	99.7 %	100.9%
$\mathrm{Eff}(0.89,N)$	100%	100.2 %	99.2 %	99.9%
$\mathrm{Eff}(1.74,N)$	100%	100.2%	99.2%	99.9%

processors. In this view, a slave processor waits for its turn in order to be able to read the message from the master processor. On the contrary, the printing message addressed to the master processor is done on slave processors and hence is nonblocking.

Figure 4 shows that in the case $N = 4$, the slave processor mainly works on parallel computing. Its communication part is shared between sending and receiving data. The sending part is longer due to the amount of data to be treated, which consist not only of partial control but also of propagator matrices.

The full efficiency of the ISM is clearly shown in the numerical simulations; see Table I. Note that in some cases, the efficiency is greater than 100% because the full problem requires more memory, and thus spends more time in hardware storage processes. On the contrary, the parallel computing uses a smaller amount of data. Also, depending on the processor architecture, the computational time shall behave nonlinearly with respect to the size of the data. Similar linear and superlinear speedup have been observed in Ref. [34] with MatlabMPI.

B. Optimal control of molecular orientation

In a second series of numerical tests, we consider the control of molecular orientation by THz laser fields. Molecular orientation [35,36] is nowadays a well-established topic both from the experimental [37,38] and theoretical points of views [39–43]. Different optimal control analyses have been made on this quantum system [44–47].

Here, we consider the control of a linear polar molecule, HCN, by a linearly polarized THz laser field $E(t)$. We assume that the molecule is in its ground vibronic state and described by a rigid rotor. In this case, the Hamiltonian of the system can be written as

$$H(t) = BJ^2 - \mu_0 \cos\theta E(t) - \frac{E(t)^2}{2}[(\alpha_\parallel - \alpha_\perp)\cos^2\theta + \alpha_\perp],$$

where B is the rotational constant, J^2 is the angular momentum operator, μ_0 is the permanent dipolar moment, $\alpha_\parallel$ and $\alpha_\perp$ are the dipole polarizability components parallel and perpendicular to the molecular axis, respectively, and θ is the angle between the molecular axis and the polarization direction of the electric field. At zero temperature, the dynamics of the system is ruled by the following differential equation:

$$i\frac{\partial|\psi(t)\rangle}{\partial t} = H(t)|\psi(t)\rangle,$$

where the initial state at $t = 0$ is $|0,0\rangle$, in the basis of the spherical harmonics $\{|j,m\rangle, j \geqslant 0, -j \leqslant m \leqslant j\}$. Numerical values are taken to be $B = 6.6376 \times 10^6$, $\mu_0 = 1.1413$, $\alpha_\parallel =$

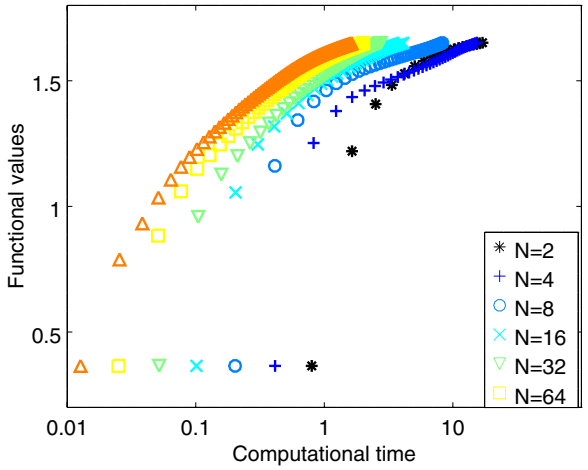

FIG. 5. Functional values at each iteration and for various values of N with respect to computational time for the control of molecular orientation by a monotonic algorithm. Logarithmic scale is used in the x axis. The functional values are dimensionless and the computational time is expressed in seconds.

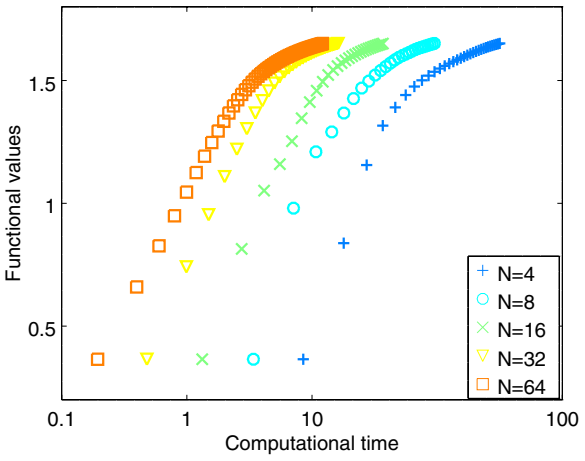

FIG. 6. Same as Fig. 5, but for a Newton solver.

20.055, and $\alpha_\perp = 8.638$, in a.u. We refer the reader to Ref. [44] for details on the numerical implementation of this control problem. The α parameter is chosen as a time-dependent function of the form $10^5(\frac{t-T/2}{T/2})^6 + 10^4$ in order to design a control field which is experimentally relevant [44]. The target state is the eigenvector of the observable $\cos\theta$ with the maximum eigenvalue in the subspace such that $j \leqslant 4$ [41,48].

Having investigated the implementation issues in Sec. IV A, here we focus on the efficiency $\mathrm{Eff}^\star(\varepsilon, N)$ achieved when neglecting the time associated with IO communications. As a consequence, the results hereafter do not depend on the used computer and software. We study the efficiency of the parallelization method in the cases where the optimization solver consists of one step of either a monotonic algorithm or a Newton method. We start with a simulation using a monotonic algorithm (see [8,49] for details about this method). Given a target value ε of the figure of merit, we measure the computational time necessary to obtain it. In the numerical computation, we use $\varepsilon = 0.3$, while the optimum has been numerically estimated as 0.2909 [44]. The values of the figure of merit are plotted in Fig. 5.

In this case, the full efficiency is not obtained, as reported in Table II. This point can be qualitatively understood as follows. The monotonic schemes are intrinsically sequential. In [31], it is shown that they can be interpreted in terms of a tracking procedure. At each time step, the control field is computed in order to decrease the distance between the current and the adjoint states. This feature is perturbed by the decomposition associated with the time parallelization, and may explain why the full efficiency can hardly be obtained with these schemes.

In the different numerical simulations, we observe that this method seems to be significantly more efficient than gradient descent solvers. The wall-clock computational time is a decreasing function of N, so that solving benefits from large parallelization. However, $\mathrm{Eff}^\star(\varepsilon, N)$ does not appear to be a monotonic function of N. The analysis of this point is beyond the scope of this paper.

We repeat this test with one iteration of the Newton method as optimization solver. More precisely, we implement a matrix-free version of the algorithm that updates the control by means of a generalized minimum residual (GMRES) routine [50,51]. In this case, we observe that this approach actually enables one to obtain convergence; the algorithm does not converge for $N = 1$ and $N = 2$, but converges for larger values of N. The values of the figure of merit are plotted in Fig. 6. Since the algorithm does not converge for $N = 1$, $t(\varepsilon, 1)$ and $\mathrm{Eff}^\star$ are not defined. In this case, we consider the quantity $Nt(\varepsilon, N)$ to measure the efficiency of the process. The results are presented in Table III.

TABLE II. Efficiency of the ISM for various values of N in the case of the control of molecular orientation with a monotonic algorithm.

N	$\mathrm{Eff}^\star(\varepsilon, N)$
1	100%
2	55.2%
4	51.2%
8	51.5%
16	38.5%
32	26.4%
64	16.3%

TABLE III. Equivalent sequential time of the ISM in the case of a Newton solver for the control of molecular orientation.

N	$Nt(\varepsilon, N)$
1	
2	
4	1264.763737
8	759.976361
16	589.424517
32	516.603943
64	774.557304

C. Optimal control of Bose-Einstein condensates

The last example investigated in this work deals with the optimal control of Bose-Einstein condensates [52]. This subject has been extensively studied in the past few years [53–59]. Following Ref. [53], we consider the control of a condensate in magnetic microtraps whose dynamics is ruled by the Gross-Pitaevskii equation,

$$\frac{\partial}{\partial t}|\psi(x,t)\rangle = [H_0 + V(x,\lambda(t))]|\psi(x,t)\rangle, \tag{11}$$

where $|\psi(x,t)\rangle$ is the state of the system, $H_0 = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + \kappa\langle\psi(x,t)|\psi(x,t)\rangle$, λ is the radio-frequency control field, and κ is a positive coupling constant. The potential V is defined by

$$V(x,\lambda) = \begin{cases} \frac{1}{2}\left(|x| - \frac{\lambda d}{2}\right)^2 & \text{for } |x| > \frac{\lambda d}{4} \\ \frac{1}{2}\left(\frac{(\lambda d)^2}{8} - x^2\right) & \text{otherwise,} \end{cases}$$

with $d > 0$. Unitless parameters are used here. We refer the reader to Ref. [53] for details on the model system. The figure of merit associated with this control problem is

$$\mathcal{J}[\lambda] = \mathrm{Re}[\langle\psi(T)|\psi_f\rangle]. \tag{12}$$

The final and initial states of the control problem are, respectively, the ground states of the Hamiltonians $H_0 + V(x,0)$ and $H_0 + V(x,1)$. The parameter α is set to 0.

Due to the nonlinearity of the model system, the preceding approach has to be adapted. In this way, we no longer consider the sequence of states $|\varphi_n^u\rangle$ [see Eq. (4)], but we split up these intermediate states into two sets. The sequence of initial states is taken on the trajectory $|\psi(t)\rangle$, i.e., defined by $|\psi(t_n)\rangle$, $n = 0,\ldots,N$, while the sequence of target states is taken on the adjoint trajectory $|\chi(t)\rangle$, i.e., defined by $|\chi(t_n)\rangle$, $n = 0,\ldots,N$. The maximization problem in step (2c) of Algorithm 1 is therefore replaced by the maximization of the subfunctional,

$$\mathcal{J}_n[u_n,|\psi^u\rangle] = \mathrm{Re}[\langle\psi_n(t_{n+1})|\chi(t_{n+1})\rangle], \tag{13}$$

with $0 \leqslant n \leqslant N-1$. In this problem, the state $|\psi_n\rangle$ is defined on $[t_n,t_{n+1}]$ by Eq. (5), but starting from $|\psi_n(t=t_n)\rangle = |\psi(t_n)\rangle$. The rest of the procedure remains unchanged.

This modification does not dramatically affect the computational time since these trajectories are not computed sequentially, but in parallel. We then use the propagator assembling technique presented in Sec. IV to compute the sequences of initial and final states. The numerical values of the parameters are set to $T = 8$, $\kappa = 1$, and $d = 10$. The space domain we consider is $[-10,10]$. For the space discretization, we consider a uniform grid composed of 50 points. The time discretization is achieved with a time grid of 2^9 points, and we use Strang's splitting (8) to compute the trajectories. The optimization solver in step (2c) consists of one iteration of the constant step gradient descent method [see Eq. (9)], with $\rho = 10^{-1}$. The results are presented in Fig. 7.

We observe that the full efficiency is not only reached, but even overtaken, as confirmed in Table IV. This observation is certainly a consequence of the nonlinear setting and can be interpreted as follows. As can be seen in Sec. IV A, linear dynamics enables one to obtain full efficiency. Applying the parallelization method on such a problem with a control time T and with two subintervals (and two processors) requires a time

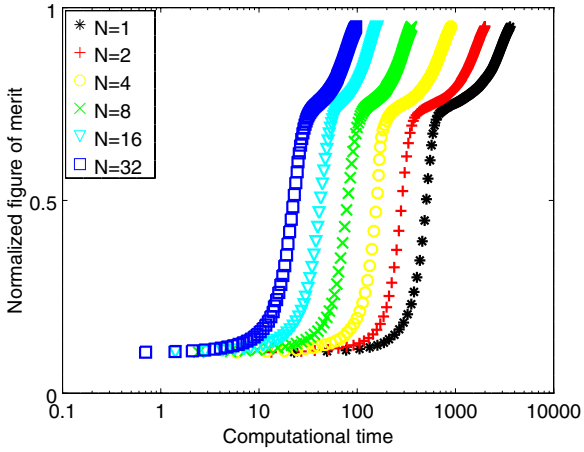

FIG. 7. Evolution of the normalized figure of merit at each iteration and for various values of N with respect to computational time (expressed in seconds) for the control of a Bose-Einstein condensate.

for solving the problem of the order of $T/2$. As a consequence, the complexity of the optimization problem appears to be a linear function of T. This dependence is more complicated in the nonlinear case. The numerical results show, in our example, that the complexity behaves as a sublinear function of T. In this way, the subcontrol problems are simpler not only because of the size reduction induced by the time decomposition, but also because of the dynamics itself. This point will be explored in a forthcoming work.

V. CONCLUSION AND PERSPECTIVES

In this work, we have investigated the numerical efficiency of a time-parallelized optimal control algorithm on standard quantum control problems, extending from the manipulation of spin systems and molecular orientation to the control of Bose-Einstein condensates. We have shown that the full efficiency can be reached in the case of a linear dynamics optimized by means of gradient methods. On the contrary, full efficiency is not achieved when using monotonic algorithms and Newton solvers. In the case of a Newton method, the parallelization setting reduces the length of time intervals where the solver is used, and makes the subproblems easier to solve. Such a property is also observed in the case of nonlinear dynamics, as shown with the example of Bose-Einstein condensates. The results of this work can be viewed as an important step forward for the implementation of parallelization methods in

TABLE IV. Efficiency of the ISM for various values of N in the case of the control of a Bose-Einstein condensate.

N	$\mathrm{Eff}^\star(\varepsilon,N)$
1	100%
2	90%
4	99.7%
8	126.3%
16	141.8%
32	116.65%

quantum optimal control algorithms. Their use will become a prerequisite in the near future to simulate quantum systems of increasing complexity.

ACKNOWLEDGMENTS

S.J.G. acknowledges support from the DFG (Grant No. Gl 203/7-1), SFB 631, and the BMBF FKZ 01EZ114 project. D.S. and S.J.G. acknowledge support from the ANR-DFG research program Explosys (ANR-14-CE35-0013-01; DFG-Gl 203/9-1). J.S. was partially supported by the Agence Nationale de la Recherche (ANR), Projet Blanc EMAQS No. ANR-2011-BS01-017-01 and Projet Blanc CINE-PARA No. ANR-15-CE23-0019-01. This work has been done with the support of the Technische Universität München Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under Grant Agreement No. 291763.

[1] D. D'Alessandro, *Introduction to Quantum Control and Dynamics*, Applied Mathematics and Nonlinear Science Series (Chapman and Hall, Boca Raton, FL, 2008).
[2] P. Brumer and M. Shapiro, *Principles and Applications of the Quantum Control of Molecular Processes* (Wiley Interscience, New York, 2003).
[3] C. Brif, R. Chakrabarti, and H. Rabitz, New J. Phys. **12**, 075008 (2010).
[4] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Kockenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen *et al.*, Eur. Phys. J. D **69**, 79 (2015).
[5] C. Altafini and F. Ticozzi, IEEE Trans. Automat. Control **57**, 1898 (2012).
[6] D. Dong and I. A. Petersen, IET Control Theory A **4**, 2651 (2010).
[7] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).
[8] D. Reich, M. Ndong, and C. P. Koch, J. Chem. Phys. **136**, 104103 (2012).
[9] P. Doria, T. Calarco, and S. Montangero, Phys. Rev. Lett. **106**, 190501 (2011).
[10] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (Wiley, New York, 2008).
[11] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon, Oxford, 1987), Vol. 14.
[12] E. Assémat, M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Phys. Rev. A **82**, 013415 (2010).
[13] M. Lapert, J. Salomon, and D. Sugny, Phys. Rev. A **85**, 033406 (2012).
[14] M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Phys. Rev. Lett. **104**, 083001 (2010).
[15] N. Khaneja, R. Brockett, and S. J. Glaser, Phys. Rev. A **63**, 032308 (2001).
[16] Y. Zhang, M. Lapert, D. Sugny, M. Braun, and S. J. Glaser, J. Chem. Phys. **134**, 054103 (2011).
[17] S. Conolly, D. Nishimura, and A. Macovski, IEEE Trans. Med. Imaging **5**, 106 (1986).
[18] M. A. Bernstein, K. F. King, and Zhou, *Handbook of MRI Pulse Sequences* (Elsevier, London, 2004).
[19] M. Lapert, Y. Zhang, M. Janich, S. J. Glaser, and D. Sugny, Sci. Rep. **2**, 589 (2012).
[20] A. Garon, S. J. Glaser, and D. Sugny, Phys. Rev. A **88**, 043422 (2013).
[21] W.-C. Chu and C. D. Lin, Phys. Rev. A **85**, 013409 (2012).
[22] A. Toselli and O. B. Widlund, *Domain Decomposition Methods - Algorithms and Theory*, Springer Series in Computational Mathematics (Springer, New York, 2005), Vol. 34.
[23] G. Horton, Commun. Appl. Num. Methods **8**, 585 (1992).
[24] Y. Maday, J. Salomon, and G. Turinici, SIAM J. Numer. Anal. **45**, 2468 (2007).
[25] H. J. Hogben, M. Krzystyniak, G.T.P. Charnock, P. J. Hore, and I. Kuprov, J. Magn. Reson. **208**, 179 (2011).
[26] D.L. Goodwin, and I. Kuprov, arXiv:1510.02420.
[27] T. Gradl, A. K. Spörl, T. Huckle, S. J. Glaser, and T. Schulte-Herbrüggen, Proc. EUROPAR 2006, Parallel Proc. Lect. Notes Comput. Sci. **4128**, 751 (2006).
[28] T. Auckenthaler, M. Bader, T. Huckle, A. Spörl, and K. Waldherr, Parallel Comput. Arch. **36**, 359 (2010).
[29] T. E. Skinner and S. J. Glaser, Phys. Rev. A **66**, 032112 (2002).
[30] J. P. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003).
[31] J. Salomon and G. Turinici, J. Chem. Phys. **124**, 074102 (2006).
[32] R. Marx, A. F. Fahmy, J. M. Myers, W. Bermel, and S. J. Glaser, Phys. Rev. A **62**, 012310 (2000).
[33] R. Marx, N. Pomplum, W. Bermel, H. Zeiger, F. Engelke, A. F. Fahmy, and S. J. Glaser, Magn. Reson. Chem. **53**, 442 (2015).
[34] J. Kepner and S. Ahalt, J. Parallel Distrib. Comput. **64**, 997 (2004).
[35] H. Stapelfeldt and T. Seideman, Rev. Mod. Phys. **75**, 543 (2003).
[36] T. Seideman and E. Hamilton, Adv. At. Mol. Opt. Phys. **52**, 289 (2006).
[37] S. De, I. Znakovskaya, D. Ray, F. Anis, N. G. Johnson, I. A. Bocharova, M. Magrakvelidze, B. D. Esry, C. L. Cocke, I. V. Litvinyuk, and M. F. King, Phys. Rev. Lett. **103**, 153002 (2009).
[38] S. Fleischer, Y. Zhou, R. W. Field, and K. A. Nelson, Phys. Rev. Lett. **107**, 163603 (2011).
[39] C. M. Dion, A. Keller, and O. Atabek, Eur. Phys. J. D **14**, 249 (2001).
[40] I. S. Averbukh and R. Arvieu, Phys. Rev. Lett. **87**, 163601 (2001).
[41] D. Daems, S. Guérin, D. Sugny, and H. R. Jauslin, Phys. Rev. Lett. **94**, 153003 (2005).
[42] M. Lapert and D. Sugny, Phys. Rev. A **85**, 063418 (2012).
[43] R. Tehini and D. Sugny, Phys. Rev. A **77**, 023407 (2008).
[44] J. Salomon, C. M. Dion, and G. Turinici, J. Chem. Phys. **123**, 144310 (2005).
[45] M. Lapert, R. Tehini, G. Turinici, and D. Sugny, Phys. Rev. A **78**, 023408 (2008).
[46] M. Yoshida and Y. Ohtsuki, Phys. Rev. A **90**, 013415 (2014).
[47] K. Nakajima, H. Abe, and Y. Ohtsuki, J. Phys. Chem. A **116**, 11219 (2012).
[48] D. Sugny, A. Keller, O. Atabek, D. Daems, C. M. Dion, S. Guérin, and H. R. Jauslin, Phys. Rev. A **71**, 063402 (2005).
[49] J. Salomon, ESAIM: M2AN **41**, 77 (2007).
[50] R. Barret, M. Berry, T. F. Chan, J. Demmel, J. M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, J. CharlKepner, and S. Ahalt,

Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods (SIAM, Philadelphia, 1994).

[51] Y. Saad and M. H. Schultz, SIAM J. Sci. Stat. Comput. **7**, 856 (1986).

[52] R. Folman, P. Krüger, J. Schmiedmyaer, J. Denschlag, and C. Henkel, Adv. At. Mol. Opt. Phys. **48**, 263 (2002).

[53] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, Phys. Rev. A **75**, 023602 (2007).

[54] A. Borzi and U. Hohenester, SIAM J. Sci. Comput. **30**, 441 (2008).

[55] R. Bücker, T. Berrada, S. van Frank, J.-F. Schaff, T. Schumm, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, J. Phys. B **46**, 104012 (2013).

[56] G. Jäger, D. M. Reich, M. H. Goerz, C. P. Koch, and U. Hohenester, Phys. Rev. A **90**, 033628 (2014).

[57] J. F. Mennemann, D. Matthes, R.-M. Weishäupl, and T. Langen, New J. Phys. **17**, 113027 (2015).

[58] M. Lapert, G. Ferrini, and D. Sugny, Phys. Rev. A **85**, 023611 (2012).

[59] J.-F. Schaff, X.-L. Song, P. Capuzzi, P. Vignolo, and G. Labeyrie, Eur. Phys. Lett. **93**, 23001 (2013).