Superposition principle for the simultaneous optimization of collective responses

S. Dong, R. Flores, J. Unger, Q. Su, and R. Grobe

Phys. Rev. E **98**, 012221 — Published 31 July 2018

DOI: 10.1103/PhysRevE.98.012221

Superposition principle for the simultaneous optimization of collective responses

S. Dong[1,2], R. Flores[1], J. Unger[1], Q. Su[1] and R. Grobe[1]

(1) Intense Laser Physics Theory Unit and Department of Physics
Illinois State University, Normal, IL 61790-4560, USA

(2) Key Laboratory for Laser Plasmas, School of Physics and Astronomy, Collaborative Innovation Center of IFSA, Shanghai Jiao Tong University, Shanghai 200240, China

We study the dynamics of sets of independent systems, all of which are coupled to the same time-dependent external force. Using optimal control theory, we compute the most efficient temporal pulse shape for this force that can maximize simultaneously the collective response of these systems. This response can be a weighted sum of all amplitudes at the final interaction time. Remarkably, it turns out that for certain systems this optimal force for the collective response can be related to the individual forces that would optimize each system separately. We illustrate this superposition principle for the simultaneous optimization of collective responses with numerical and also analytical solutions for sets of damped linear and non-linear oscillators. We also apply this principle to predict the optimal temporal profile of a laser pulse that can maximize the final macroscopic polarization (total dipole moment) of a set of quantum mechanical two-level atoms.

1. Introduction

There are numerous examples in science, technology and engineering, where certain desired dynamical variables need to be optimized [1-8]. In physics, for example, the very fundamental laws of nature, such as Newton's law, the Schrödinger, Dirac or Maxwell equations, are usually given by differential equations that specify those solutions that optimize a certain action. More specifically, the area of coherent quantum control has recently received a lot of interest [9]. Here the goal is to tailor external laser pulse shapes to control the final outcome of complex chemical reactions or to create or break a particular bond in a molecule [10,11]. Naively, one would expect that lasers that are tuned to particular resonances would be ideal. However, due to internal energy conversions based on vibrational relaxations often the entire molecule was heated. Most recent experiments have exploited rather sophisticated adaptive laser pulse-shaping techniques including closed-loop learning that also permits the independent shaping of the laser's polarization [12], amplitude as well as phase.

From a theoretical point of view a powerful approach is the optimal control theory, where a certain objective function is optimized under suitable constraints. As this is a very intensively studied area, there are numerous reviews available. In some situations, a finite number of variables needs to be optimized that can often be accomplished via straightforward conjugate gradient-like search algorithms. Computationally more demanding are infinite dimensional problems, where, for example, an entire function of time needs to be computed, equivalent to an infinite-dimensional optimization with constraints. What is common to most problems, however, is the fact that dynamical variables that characterize a *single* system are usually considered.

While it is possible in many cases to calculate the optimum signal, however, it is much more difficult to obtain an intuition for the best signal. It is one of the purposes of this article to obtain an intuition or qualitative guidance about the properties of the optimum signal and not to introduce a new computational methodology for a particular system. In particular, in this note, we examine dynamical situations where a *collective* response of several (and possible infinitely many) independent systems needs to be optimized. In particular, in this note, we examine dynamical situations where a *collective* response of several (and possible infinitely many) independent systems needs to be optimized. The key question studied here is whether one can learn something from the corresponding optimal control characteristic for each system individually.

In this article, we will show, that for those situations, where the coupling of the external field can be modeled in an additive way in the equations of motion, the optimum force field that

maximizes the collective response of all systems can be computed from a suitable superposition of the individual optimum pulse shapes, where the weight factors depend on the temporal derivative of the individual fields. While the superposition principle for the simultaneous optimization of collective responses (or abbreviated "SPSO") is not exactly valid for non-linear systems or those where the control field is not coupled as inhomogeneous source terms, it can still provide a remarkably accurate guidance to predict the optimal control field for the collective response from the individual optimizers for each sub-system.

The goal of simultaneously optimizing several degrees of freedom has been studied for single optically driven multi-level systems. For example, in the special case where the energy levels are cleanly separated from each other, the system can be controlled by sequentially addressing the pairs of levels [13]. The optimum pulse can then be accomplished by a sequence of tailored pulses, each of which then addresses only a prescribed pair of levels. A control based on frequency discrimination is not always possible, e.g., it is not suitable for systems with equally or almost equally spaced or degenerate energy levels [14].

The article is organized as follows, in Section 2 we derive the general theoretical framework of optimal control theory applied to an infinite set of general dynamical systems, all of which are coupled to the same external force. We introduce the superposition principle for the simultaneous optimization of collective responses. In Section 3 we illustrate this principle with analytical solutions for sets of damped harmonic oscillators. In Section 4 we show that it can serve as a qualitative guidance to estimate the optimal force for sets of non-linear oscillators. In Section 5 we examine the optimal laser field to maximize the macroscopic polarization of two-level systems. We conclude in Section 6 with critical questions and speculations for future projects.

2. Optimal control theory for collective response of N dynamical systems

2.1 General framework and its numerical solution technique

In this work, we examine the collective response of N dynamical systems, all of which are coupled to the same external force field, denoted by U(t). The ultimate goal is to examine the properties of the optimal control field $U_{opt}(t)$ that can maximize a certain collective response of all systems at a final time T. In the concrete examples, examined below, this collective response will be the sum of all final amplitudes of harmonic as well as anharmonic oscillators, or the macroscopic dipole response of a collection of two-level atoms. We will show that for some systems it is possible to construct this collective response exactly from the knowledge about the optimal control

fields associated with every single system, denoted for the n-th system by $U_{opt,n}(t)$. We denote this the "superposition principle for simultaneous optimization" (SPSO). It turns out that for those systems for which this principle is violated (such as anharmonic oscillators or two-level systems), the SPSO can nevertheless provide a good guidance for the true optimal field $U_{opt}(t)$.

In order to avoid non-sensible optimal force fields that are infinite, we restrict the total "energy" of the external force field to $E = \int_0^T dt\, U(t)^2$. We denote the dynamical variables of the n-th system by the state vector $\mathbf{Y}_n(t)$ and require these "amplitudes" to fulfill the state-equation $d\mathbf{Y}_n/dt = \mathbf{K}_n[\mathbf{Y}_n(t),U(t)]$ with a known initial condition $Y_n(t=0)$. As a collective variable, which we like to maximize after an interaction time T, we could choose the combination $\Sigma_n \Sigma_i w_{i,n} Y_{i,n}(T)$, where the given parameters $w_{i,n}$ would characterize the relative relevance we like to assign to the i-th state variable of the n-th system. In Appendix A we review the application of the variational principles of optimal control theory to this set of systems and derive the following set of equations that the optimal control field $U_{opt}(t)$ has to satisfy.

$$d\mathbf{Y}_n/dt = \mathbf{K}_n[\mathbf{Y}_n(t),U(t)] \qquad \text{with } \mathbf{Y}_n(0) = \mathbf{Y}_n(t=0) \tag{2.1}$$

$$d\boldsymbol{\lambda}_n/dt = -\boldsymbol{\lambda}_n\, \partial\mathbf{K}_n/\partial\mathbf{Y}_n \qquad \text{with } \boldsymbol{\lambda}_n(T) = \mathbf{w}_n \tag{2.2}$$

$$U(t) = (2\lambda_0)^{-1} \Sigma_n\, \boldsymbol{\lambda}_n(t)\, \partial\mathbf{K}_n(\mathbf{Y}_n,U)/\partial U \tag{2.3}$$

The first equation reproduces the state equation for the amplitudes, the second one is for the Lagrange multiplier functions and needs to be solved reversely in time as its final values are known, $\boldsymbol{\lambda}_n(T) = \mathbf{w}_n$. As the functional form of the Lagrange functions $\boldsymbol{\lambda}_n(t)$ depend on the collective U(t), the last equation is in general a complicated transcendental equation that the optimum field has to satisfy. The single Lagrange parameter λ_0 needs to be chosen such that the optimal control field satisfies $E = \int_0^T dt\, U(t)^2$. While Eqs. (2.1) and (2.2) seem to be decoupled at first, the fact that the optimum function $U_{opt}(t)$ in Eq. (2.3) depends on each system's solution, effectively couples all of these equations for different systems.

If only a single system is coupled to the field, such that only $\Sigma_i w_{i,n} Y_{i,n}(T)$ needs to be optimized, then the optimizer has to fulfill $U_{opt,n}(t) = (2\lambda_{0,n})^{-1} \boldsymbol{\lambda}_n(t)\, \partial\mathbf{K}_n(\mathbf{Y}_n,U)/\partial U$. However, as the solutions $\boldsymbol{\lambda}_n(t)$ in this particular expression are different from the solutions $\boldsymbol{\lambda}_n(t)$ in Eq. (2.3),

there is unfortunately in general no relationship between $U_{opt}(t)$ and the individual optimizers $U_{opt,n}(t)$ for each system separately.

2.2 Numerical solution techniques

There are a wide variety of numerical techniques available to find the three solutions $\mathbf{Y}_n(t)$, $\boldsymbol{\lambda}_n(t)$ and U(t), that satisfy simultaneously the set of equations (2.1), (2.2) and (2.3). For most of the specific systems and parameter ranges examined in this work, we found it sufficient to use a simple iteration scheme that lead to convergent solutions. It was based on an initial guess for U(t), denoted by $U^{(0)}(t)$. Using this force field the state equation as well as the co-state equation were solved. The resulting solutions $\mathbf{Y}_n(t)$, $\boldsymbol{\lambda}_n(t)$ were then inserted into the right-hand side of Eq. (2.3). This right hand side was then interpreted as an improved solution for U(t), similar to the numerical strategies employed in most predictor-corrector schemes, denoted by $U^{(1)}(t)$. This new function was then inserted again as an improved guess for $U_{opt}(t)$ into equations (2.1) and (2.2). This scheme was then repeated until the iterates $U^{(n)}(t)$ and $U^{(n+1)}(t)$ became numerically indistinguishable from each other.

To have an independent check of the numerical solutions for $U_{opt}(t)$, we also used steepest descent- and conjugate gradient-based approaches [15]. Here the objective $J[U(t)] = \Sigma_n \Sigma_i w_{i,n} Y_{i,n}(T)$ is solely a function of U(t), however, J depends on $Y_{i,n}(T)$ whose values need to be determined from the solution to the differential equation (2.1) for each given U(t). But there is no need to involve any co-state variables $\lambda_n(t)$ and λ_0. The function U(t) was sampled at M points on a temporal grid, with $t_m=(m-1)/(M-1)T$ and m=1,2, …M, such that the objective J becomes a function of M parameters $U(t_m)\equiv U_m$. For simplicity, we denote the resulting M-dimensional vector $(U_1,U_2,..,U_M) \equiv \mathbf{R}$. The resulting M-dimensional maximization problem required an initial guess for $\mathbf{R}^{(0)}$, which was used to calculate the M-component gradient vector $\text{grad}J = \partial J/\partial \mathbf{R}^{(0)} \equiv (\partial J/\partial U^{(0)}_1, \partial J/\partial U^{(0)}_{2,} .. \partial J/\partial U^{(0)}_M)$. We then performed a line search based on the standard bi-section techniques to find the value for α that would maximize $J(R^{(0)}+\alpha\, \partial J/\partial \mathbf{R}^{(0)})$ along this particular direction in this vector space, i.e. $\partial J(R^{(0)}+\alpha\, \partial J/\partial \mathbf{R}^{(0)})/\partial\alpha = 0$ for $\alpha=\alpha^{(0)}$. The next line search was based on the improved location $R^{(1)}= R^{(0)}+ \alpha^{(0)}\, \partial J/\partial \mathbf{R}^{(0)}$ and a new line search direction

was either given by the new gradient $\partial J/\partial \mathbf{R}^{(1)}$ (steepest ascent method, [15]) or by a combination of the prior and the new gradient (Fletcher-Reeves method, [16]). Due to the fact that the calculation of each component of $\partial J/\partial \mathbf{R}$ required the solution of the differential equation (2.1), this method is very CPU time intensive but leads to a rapid convergence in terms of the number of required line searches.

2.3 The superposition principle for collective responses (SPSO)

We will now show, that if the generator of the time evolution $\mathbf{K}_n$ takes a simpler form, where the state variables $\mathbf{Y}_n$ of the n-th system are mutually coupled by a temporally constant matrix M_n, and the external force field U(t) enters as an additive term, i.e.

$$d\mathbf{Y}_n/dt = M_n \mathbf{Y}_n + \mathbf{V}_n U(t) \tag{2.4}$$

then the optimizing force field of the collective response of all systems can be obtained exactly from the knowledge of the single-system optimizers $U_{opt,n}(t)$. The temporally constant vector $\mathbf{V}_n$ permits us to couple the external force to each amplitude $Y_{i,n}$ differently. The differential equation is inhomogeneous such that the sum of two individual solutions is in general not a solution to the same differential equation. However, due to the special coupling to the external force as a source term, this system can always be solved exactly based on the diagonalization of the matrix M_n.

In Appendix B we derive the superposition principle for the collective responses (SPSO):

$$U_{opt}(t) = (2\lambda_0)^{-1} \Sigma_n a_n U_{opt,n}(t) \tag{2.5}$$

where the time-independent amplitude factors are determined by $a_n \equiv -(M_n^T \mathbf{w}_n)\, \mathbf{V}_n / U'_{opt,n}(T)$. We consider this new principle as the main result of this article. Each amplitude factor a_n depends on the temporal derivative of the optimal control field $U'_{opt,n}(T)$ at the final time t=T and can be easily obtained from the optimal solution $U_{opt,n}(t)$ for each system. These amplitudes a_n depend linearly on the given weight factors $\mathbf{w}_n$ for each amplitude $\mathbf{Y}_n$ as well as the coupling strength $\mathbf{V}_n$ of the external field to each individual state amplitude.

While the magnitude of the overall factor $(2\lambda_0)^{-1}$ depends non-trivially on each individual field $U_{opt,n}(t)$, its sole purpose is to guarantee that the collective optimal control field satisfies $E = \int_0^T dt\ U_{opt}(t)^2$. It can therefore easily be determined at the end after the summation in Eq. (2.5) is calculated.

3. Simultaneous optimization of coupled damped harmonic oscillators

In this section, we consider a special system of the form $d\mathbf{Y}_n/dt = M_n\ \mathbf{Y}_n + \mathbf{V}_n\ U(t)$. The temporally constant vector $\mathbf{V}_n$ permits us to couple the external force to each amplitude $Y_{i,n}$ differently.

We provide fully analytical solutions for $U_{opt,n}(t)$ that allows for a concrete interpretation of the formal amplitude factors a_n in the expression of the superposition law Eq. (2.5). The state equations for the n-th oscillator (of unit mass) are given by $dx_n/dt = p_n$ and $dp_n/dt = -\omega_n^2\ x_n - \gamma_n\ p_n + U(t)$, corresponding to the 2×2 matrix $M_n = \{\{0,1\},\{-\omega_n^2, -\gamma_n\}\}$ and $\mathbf{V}_n = \{0,1\}$. We therefore have the Hamiltonian

$$H \equiv \Sigma_n\ \boldsymbol{\lambda}_n \mathbf{K}_n\ = \Sigma_n\ \lambda_{1,n}\ p_n + \lambda_{2,n}\ [-\omega_n^2\ x_n - \gamma_n\ p_n + U(t)] \tag{3.1}$$

which should not be confused with the Hamiltonian (time-generator) for the evolution of $x_n(t)$ and $p_n(t)$, which exists only for $\gamma_n = 0$. For simplicity, let us assume we try to optimize the sum of all final amplitudes $J'' \equiv \Sigma_n\ x_n(T)$ with equal weight, or equivalently we choose $w_{1,n} = 1$ and $w_{2,n} = 0$. Then the required co-state equations

$$d\lambda_{1,n}/dt\ =\ -\ dH/dx_n\ =\ \omega_n^2\ \lambda_{2,n} \tag{3.2a}$$

$$d\lambda_{2,n}/dt\ =\ -\ dH/dp_n\ =\ -\ \lambda_{1,n} + \gamma_n\ \lambda_{2,n} \tag{3.2b}$$

can be solved analytically with $\lambda_{1,n}(T) = 1$ and $\lambda_{2,n}(T) = 0$, leading to the solution for $\lambda_{2,n}(t)$

$$\lambda_{2,n}(T)\ = -\ 2\ Exp[\gamma_n(t\text{-}T)/2]\ Sinh[(\gamma_n^2 - 4\omega_n^2)^{1/2}(t\text{-}T)/2]\ /\ (\gamma_n^2 - 4\omega_n^2)^{1/2} \tag{3.3}$$

As this solution is directly related to the optimum force field $U_{opt,n}(t)$, we should briefly discuss the time-dependence of this solution. In the non-dissipative limit (γ_n=0) this simplifies to $\lambda_{2,n}(t) = -\mathrm{Sin}[\omega_n(t-T)]/\omega_n$. In the opposite over-damped limit (ω_n=0), we obtain $\lambda_{2,n}(t) = -2\mathrm{Exp}[\gamma_n(t-T)/2]\,\mathrm{Sinh}[\gamma_n(t-T)/2]/\gamma_n$. For $T > \gamma_n^{-1}$ the latter function is nearly constant $\lambda_{2,n}(t) = \gamma_n^{-1}$ before it approaches zero for t=T. Most interestingly, for $\omega_n \approx \gamma_n$, $\lambda_{2,n}(t)$ grows slowly in time to reach its maximum close to t=T, before it decays to zero at t=T. This behavior is intuitively expected as the optimum force field $U_{opt,n}(t)$ required to maximize the final amplitude $x_n(T)$ has only a finite energy E. It is more advantageous to transfer this finite amount close to the final interaction time, as any early excitations are automatically damped out long before the important final time is reached.

According to Eqs. (A14) and (A15), and using $\partial \mathbf{K}_n(\mathbf{Y}_n,U)/\partial U = \mathbf{V}_n = \{0,1\}$, the resulting optimizing force for the n-th oscillator is directly proportional to $\lambda_{2,n}(t)$,

$$U_{opt,n}(t) = (2\lambda_{0,n})^{-1}\,\lambda_{2,n}(t) \tag{3.4}$$

with the energy normalization factor $(2\lambda_{0,n}) = E^{-1/2}\{\int dt\,\lambda_{2,n}(t)^2\}^{1/2}$. This integral can be evaluated analytically leading to the expression

$$\int dt\,\lambda_{2,n}(t)^2 = \mathrm{Exp}(-\gamma_n T)\,\{4\omega_n^2+\nu_n^2\,\mathrm{Exp}(\gamma_n T) - \gamma_n\,[\gamma_n\,\mathrm{Cosh}(\nu_n T) +\nu_n\,\mathrm{Sinh}(\nu_n T)]\}/\,(2\omega_n^2\gamma_n\,\nu_n^2) \tag{3.5}$$

where we abbreviate $\nu_n \equiv (\gamma_n^2-4\omega_n^2)^{1/2}$.

As, according to Eq. (B.3), these normalization factors $(2\lambda_{0,n})$ are also equal to the weights a_n of the individual optimal forces with regard to the optimal control for the collective response of all oscillators, we obtain

$$a_n = E^{-1/2}\int dt\,\lambda_{2,n}(t)^2 \tag{3.6}$$

This complicated expression based on Eq. (3.5) takes an easier form in the limit for γ_n=0

$$a_n = E^{-1/2}\{[2\omega_n T- \mathrm{Sin}(2\omega_n T)]/(4\omega_n^3)\}^{1/2} \tag{3.7}$$

Equivalently, the same amplitude factor a_n also could be determined from the form $-(M_n^T \mathbf{w}_n)\, \mathbf{V}_n / U'_{opt,n}(T)$. Here we can use $M_n^T \mathbf{w}_n = \{\{0, -\omega_n^2\}, \{1, -\gamma_n\}\}\{1,0\} = \{0, 1\}$, and $\mathbf{V}_n = \{0,1\}$ and using Eq. (3.2b), $d\lambda_{2,n}(t)/dt|_{t=T} = -1$, this leads to $U'_{opt,n}(T) = -(2\lambda_{0,n})^{-1}$ and therefore obtain for $\gamma_n=0$ the same expression. In the limit of large frequencies ω_n, or, equivalently in the limit of sufficiently long interaction times ($T>>2\pi/\omega_n$), the weight factors a_n decrease with increasing frequency ω_n as $a_n \approx E^{-1/2}\, T^{1/2}/(2^{1/2}\omega_n)$. In the opposite limit of very short interaction times, ($T<<2\pi/\omega_n$), the weight factors lose their frequency dependence, $a_n \approx E^{-1/2}\, T^3$. This means, that in this case, the functional form of the collective optimizer is simply given by $U_{opt}(t) \approx (2\lambda_0)^{-1}\, E^{-1/2}\, T^3\, \Sigma_n\, U_{opt,n}(t)$.

We will close this subsection with a brief numerical example of the time-dependence of $U_{opt}(t)$ for a continuum of oscillators. For simplicity, we assume that the distribution of the frequencies ω_n is uniform between ω_{min} and $\omega_{max} = \infty$. In this special case, the summation $U_{opt}(t) = (2\lambda_0)^{-1}\, \Sigma_n\, a_n\, U_{opt,n}(t) = -(2\lambda_0)^{-1}\, \Sigma_n\, \mathrm{Sin}[\omega_n(t-T)]/\omega_n$ can be expressed as an integral

$$U_{opt}(t) = -N \int_{\omega_{min}}^{\infty} d\omega\, \mathrm{Sin}[\omega(t-T)]/\omega \qquad (3.8)$$

where the normalization factor N guarantees that the total energy of $U_{opt}(t)$ is E. The pulse shape of the optimal collective control field depends on the relationship between ω_{min} and the interaction time T. For $\omega_{min} = 0$, the control field is constant $U_{opt}(t) = (E/T)^{1/2}$. For slightly larger values of ω_{min}, $U_{opt}(t)$ grows linearly to its maximum value close to t=T. For an even larger ω_{min}, $U_{opt}(t)$ becomes oscillatory with frequency ω_{min} and an exponentially increasing amplitude. In the extreme limit where ω_{min} is very large, $U_{opt}(t)$ approaches a function that basically vanishes for all times, except close to the final time, $U_{opt}(t) = (E)^{1/2}\, \delta(t-T)^{1/2}$. We display this trend for a few frequencies in Figure 1.

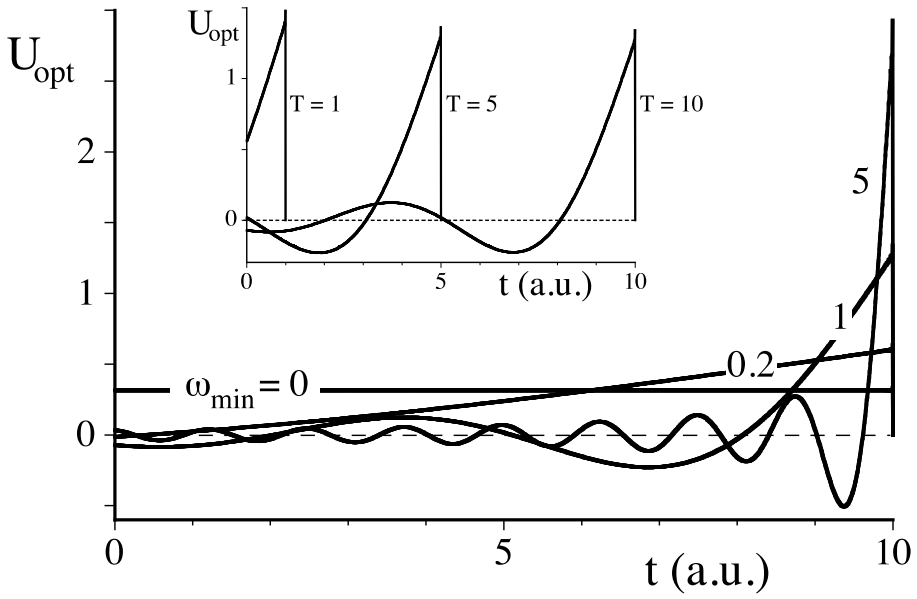

Figure 1 The temporal pulse shape of the optimal force field $U_{opt}(t)$ that simultaneously maximizes sum of all final amplitudes $x_n(T)$ of a continuous set of harmonic oscillators with a uniform distribution of frequencies ω_n between ω_{min} and infinity. Inset: The optimal force field $U_{opt}(t)$ for ω_{min} =1 for three interaction times T.

A similar trend can also be observed for a fixed lowest frequency ω_{min} as a function of the interaction time T. In the inset of the Figure, we show the transition for a linearly growing $U_{opt}(t)$ for small T to one that is non-zero only close to the final interaction time T.

For a closely related example of a physical situation where an external force field needs to be optimized we refer the reader to the works by Glasgow et al. [17-19], where an optimal field was constructed that, for a given excitation level of a passive dielectric material, would minimize the necessary energy density that has to be deposited in the medium. While the main general conclusions were model independent, the linear susceptibility of the medium was modelled by multiple Lorenz oscillators with characteristic resonance frequencies, oscillator strengths and damping rates. The authors showed that there are many admissible field histories that lead to the same state in the medium and optimized the particular shape that would minimize the unavoidable energy loss. The early time behavior of the optimum creation field was characterized by a dc portion, an exponential growth and a spike. These three particular temporal features shown in [18] are in an interesting qualitative agreement with those graphed in Figure 1.

Reference [18] also embodies a similar linear superposition principle that is discussed in the present work. However, in contrast to our work, where a prescribed optimum state at a finite time

was the objective, the prior work addresses an infinite-duration optimal excitation pulse for which even analytical solutions can be found. The authors also provided several theorems applicable to linear as well as nonlinear dielectric media to describe the general properties of optimal fields with regard to extracting as well as infusing energy from and into dielectrics.

4. The SPSO for anharmonic oscillators

In this Section, we will examine a state equation that is highly non-linear in $\mathbf{Y}_n(t)$, but the external force U(t) is still coupled in an additive manner as a source term to the dynamics. Many oscillator systems in atomic or molecular physics are modeled by the Morse, Lennard-Jones or other semi-empirical potential functions, all of which are harmonic oscillators close to equilibrium. Therefore, as non-linearities occur here only for very large excitations, it is obvious that the SPSO would naturally apply as well for lower excitations of those systems. Here we would expect discrepancies to arise only if the amplitudes are driven far from equilibrium into regions where the re-storing forces show deviations from Hooke's law.

In order to study exclusively non-linear responses, we choose in this section purposely a set of systems that are intrinsically non-linear from the very beginning for any amplitude. We choose a set of quartic oscillators, whose state equations are given by

$$dx_n/dt = p_n \tag{4.1a}$$

$$dp_n/dt = -\kappa_n x_n^3 + U(t) \tag{4.1b}$$

with positive coefficients of non-linearity κ_n. Furthermore, to enhance the importance of the non-linearity, we place the particles initially at $x_n(t=0) = 1$, such that they experience the spatial nonlinearity of the force field from the very beginning. We therefore have the Hamiltonian

$$H \equiv \Sigma_n \boldsymbol{\lambda}_n \mathbf{K}_n = \Sigma_n \lambda_{1,n} p_n + \lambda_{2,n} [-\kappa_n x_n^3 + U(t)] \tag{4.2}$$

To be comparable to the conclusions of the prior Section, we try to optimize again the sum of all final amplitudes $J'' \equiv \Sigma_n x_n(T)$ with equal weight, or equivalently we choose $w_{1,n} = 1$ and $w_{2,n} = 0$. Then the required co-state equations take the form

$$d\lambda_{1,n}/dt = -dH/dx_n = \kappa_n \lambda_{2,n} 3 x_n^2 \quad (4.3a)$$

$$d\lambda_{2,n}/dt = -dH/dp_n = -\lambda_{1,n} \quad (4.3b)$$

with $\lambda_{1,n}(T) = 1$ and $\lambda_{2,n}(T) = 0$. In contrast to the co-state equations (3.2), which were entirely independent of U(t), this set of equations depends implicitly on U(t) due to the coupling with the source term $\kappa_n \lambda_{2,n}(t) 3 x_n(t)^2$. Obviously, this highly non-linear coupling of the state and co-state equations requires numerical approaches to construct $U_{opt,n}(t)$ as well as $U_{opt}(t)$. We will argue below that the absence of an explicit dependence of U(t) in the co-state equation is extremely important for the validity of the superposition principle in the low-E limit for U(t).

In order to test the applicability of the SPSO according to Eq. (2.5), we have to compute the suitable weight factors a_n for our system. Due to the coupling of U(t) as an additive source term, we have here again $\partial \mathbf{K}_n(\mathbf{Y}_n,U)/\partial U = \{0,1\}$, such that Eq. (A14) predicts here the transcendental equation $U(t) = (2\lambda_0)^{-1} \Sigma_n \lambda_{2,n}$ for the collective response, which is also consistent with Eq. (B.1). In contrast to the prior Section, here the co-states depend intrinsically on U(t), i.e., $\lambda_{2,n} = \lambda_{2,n}[t,U(t)]$, such that this expression is an transcendental equation that the optimal $U_{opt}(t)$ has to satisfy. While the equation contains a simple sum of the co-states and suggests some kind of "additivity property", due to its transcendental character, it does not imply any validity of the superposition principle. In fact, it also follows that the co-state solutions are therefore different from the co-state solutions for each system individually, i.e. $\lambda_{2,n}[t,U_{opt}(t)] \neq \lambda_{2,n}[t,U_{opt,n}(t)]$. However, it is nevertheless rather tempting to examine if the superposition principle has at least some relevance for this system.

In order to examine this question, we have computed $U_{opt,1}(t)$ and $U_{opt,2}(t)$ for two systems and compared it with $U_{opt}(t)$ for the collective response. In the special case of two identical systems, i.e. $\kappa_1 = \kappa_2$, it is clear that $U_{opt,1}(t) = U_{opt,2}(t)$ and therefore trivially we have $[U_{opt,1}(t) + U_{opt,2}(t)]/2 = U_{opt}(t)$.

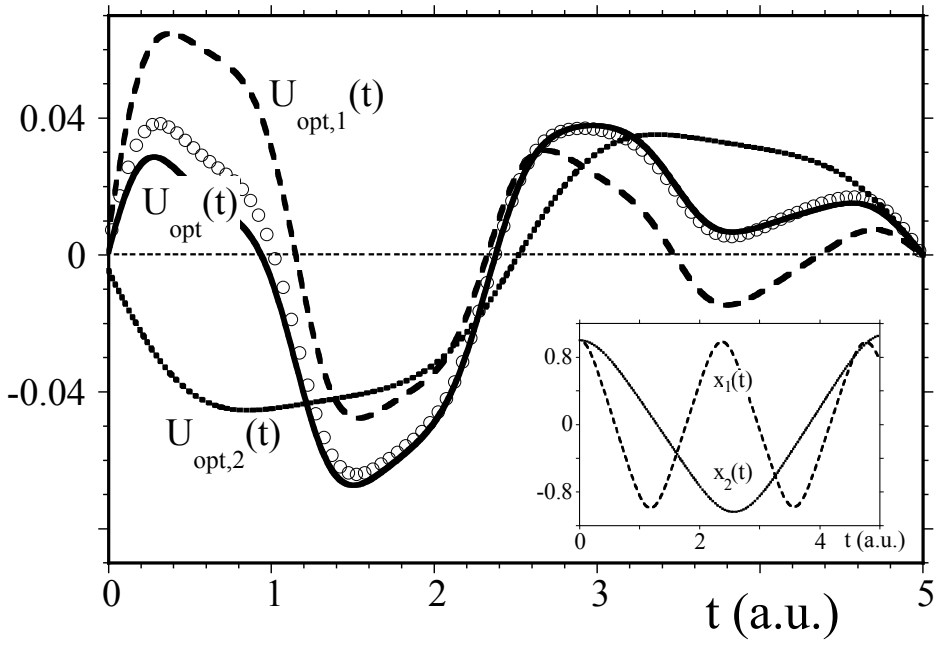

Figure 2 The optimal pulse shapes $U_{opt,1}(t)$ and $U_{opt,2}(t)$ for two quartic oscillators with κ_1=2 and κ_2=10. (dashed and dotted lines). The continuous line is the optimum $U_{opt}(t)$ to maximize simultaneously the sum of the two final amplitudes, J"= $x_1(T)$+ $x_2(T)$. The open circles are the predictions of the SPSO for $U_{opt}(t)$. The inset shows the individual orbits optimized under $U_{opt,1}(t)$ and $U_{opt,2}(t)$. (E=0.005)

Let us discuss a concrete numerical example for the opposite and more interesting case, where the two non-linearities are rather different, say κ_1=2 and κ_2=10. In the inset of Figure 2 we show the time-evolution of the two amplitudes $x_1(t)$ and $x_2(t)$ associated with the two optimal fields $U_{opt,1}(t)$ and $U_{opt,2}(t)$. In the absence of any external field, both particles (starting at $x_{1,2}(t=0) = 1$) experience the potential $V(x) = \kappa_{1,2}/4\ x^4$ and would initially first accelerate to the left and then approach their final amplitudes at time t=T=5, $x_1(T) = 0.610$ and $x_2(T) = 0.942$. In order to maximize the final elongation, the required actions of the two optimal fields are very different. In order to optimize the first oscillator, $U_{opt,1}(t)$ has to deccelerate the particle first [$U_{opt,1}(t)>0$], while the second particle requires apparently an initial boost to the left [$U_{opt,2}(t)<0$] to accelerate it in order to maximize its final elongation. As a result of the time-dependent force fields, the final optimized amplitudes are $x_1(T) = 0.757$ and $x_2(T) = 1.06$. The amount of enhancement of the final amplitudes increases obviously with the available energy E of the control field.

The Figure also shows the optimal control field $U_{opt}(t)$ for the collective response, leading to $x_1(T) = 0.723$ and $x_2(T) = 1.02$, which is slightly less than their optimal values possible under

individual optimization. For comparison, the open circles show that predicted optimal response according to the superposition principle according to Eq. (2.5), i.e.,

$$U_{opt}(t) = (2\lambda_0)^{-1} [a_1 U_{opt,1}(t) + a_2 U_{opt,2}(t)] \tag{4.4}$$

with $a_n \equiv -1/U'_{opt,n}(T)$. The qualitative agreement with the exact $U_{opt}(t)$ is remarkable. As we increase the available energy E of the control field, we found that the agreement deteriorates.

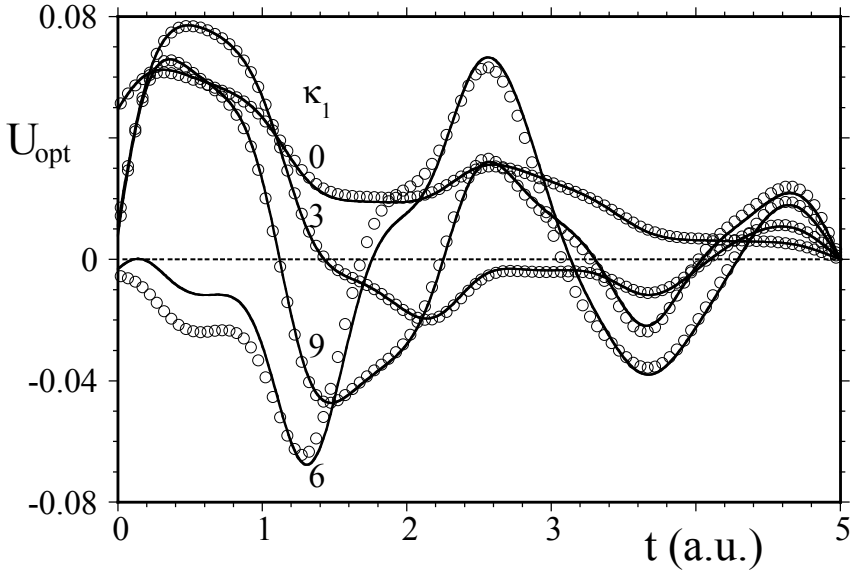

Figure 3 Comparison of the exact $U_{opt}(t)$ for the collective response of two nonlinear oscillators (with four nonlinearities κ_1 =0, 3, 6 and 9 and κ_2=10) and the corresponding prediction according to the SPSO denoted by the open circles. (E=0.005).

To provide a more general idea about the quality of the SPSO for a wider range of non-linearities, we display in Figure 3 a comparison of the exact collective optimizing force $U_{opt}(t)$ and the (approximate) predictions based on the SPSO, calculated from the individual optimizers $U_{opt,1}(t)$ and $U_{opt,2}(t)$ and their derivatives at t=T according to Eq. (4.4). We have kept the nonlinearity κ_2 of the first oscillator constant (=10) and chose selected values for κ_1 in the range from zero to 10. Due to the large degree of non-linearity, the forces $U_{opt}(t)$ depend extremely sensitively on κ_1. The qualitative agreement of the predicted and exact optimizers over the entire parameter range is in view of the large non-linearities rather remarkable. While for κ_1 very close to 10 the agreement is expected, as the two systems become identical, the optimal trajectories for κ_1=9 and κ_2 =10 and the resulting individual optimizers $U_{opt,1}(t)$ and $U_{opt,2}(t)$ are entirely different.

The surprising applicability of the SPSO for this highly nonlinear system (for which analytical solutions do not exist) can also be suggested analytically if the pulse energy E is sufficiently small. We can view $U_{opt}(t)$ as a perturbation in the state equations and can formally insert this solution for $x_n(t)$ [obtained for U(t)=0 in Eqs. (4.1)] into the co-state equations (4.3), despite the fact that this non-linear solution is not analytically known. As Eqs. (4.3) do not contain U(t) explicitly, the co-state equations therefore no longer depend on U(t) even implicitly. As a result, the Lagrange functions $\lambda_{1,n}(t)$ and $\lambda_{2,n}(t)$ that determine the final $U_{opt,n}(t)$ make the SPSO exact in this case.

In summary, while the SPSO is not expected at all to have any meaning for (intrinsically) non-linear oscillators, it can still serve as a surprisingly accurate guidance to predict $U_{opt}(t)$ from $U_{opt,n}(t)$ for two special cases. In the first case, the systems are dynamically similar ($\kappa_n \approx \kappa_m$) and, more importantly, in the second case the SPSO is valid for all systems (independently of the degree of non-linearity) where the energy of the optimal force is small, and the set of non-linear systems can be controlled perturbatively.

5. The SPSO for the macroscopic dipole moment of a set of driven two-level atoms

In contrast to the two prior systems discussed in Sections 3 and 4, we examine here the SPSO for a system where the external control field is not coupled as a source term. Here the state equations follow the general form $d\mathbf{Y}_n/dt = L_n[U(t)]\ \mathbf{Y}_n$. This set is linear as the generator $L_n[U(t)]$ is a U(t)-dependent matrix and the sum of two individual solutions is automatically also solution to the same differential equation. However, despite this additivity, analytical solutions for $\mathbf{Y}_n(t)$ are possible only in very special situations, as the two generator matrices $L_n[U(t_j)]$ and $L_n[U(t_k)]$ do not commute in general at different times t_j and t_k.

More specifically, we will now examine the validity of the SPSO for a quantum mechanical system. We consider a set of two-level atoms, each of which is coupled to the same external field U(t). For a nicely written review on optimal control theory for quantum systems and on the optimization of a single two-level system in the rotating wave or the perturbative approximation in particular, see [9]. The Schrödinger-Hamiltonian (in scaled or atomic units) is given by

$$H_{QM} \equiv \Sigma_n \left[g_n \,|g;n\rangle\langle g;n| + e_n \,|e;n\rangle\langle e;n| + U(t)\ |e;n\rangle\langle g;n| + U(t)\,|g;n\rangle\langle e;n| \right] \qquad (5.1)$$

and the collective state is a product of the superpositions of the ground and excited state of the n-th atom, $|\Psi(t)\rangle = \Pi_n \, (C_{g,n}(t) \, |g;n\rangle + C_{e,n}(t) \, |e;n\rangle)$. The time-dependent amplitudes follow from the Schrödinger equation $i \; \partial|\Psi(t)\rangle / \partial t = H_{QM} \, |\Psi(t)\rangle$ as

$$i \; d \, C_{g,n}(t) \, /dt = \; g_n \, C_{g,n}(t) + U(t) \, C_{e,n}(t) \tag{5.2a}$$

$$i \; d \, C_{e,n}(t) \, /dt = \; e_n \, C_{e,n}(t) + U(t) \, C_{g,n}(t) \tag{5.2b}$$

In order to convert the complex amplitudes into real variables, we can introduce the Bloch vector variables as

$$S_{1,n}(t) \; \equiv \; C_{g,n}(t) C_{e,n}(t)^* + C_{g,n}(t)^* C_{e,n}(t) \tag{5.3a}$$

$$S_{2,n}(t) \; \equiv -i \, [C_{g,n}(t) C_{e,n}(t)^* - C_{g,n}(t)^* C_{e,n}(t)] \tag{5.3b}$$

$$S_{3,n}(t) \; \equiv |C_{e,n}(t)|^2 - |C_{g,n}(t)|^2 \tag{5.3c}$$

These variables also permit us to include phenomenologically various dissipative terms. For example, the collisional broadening can lead to an atomic dipole's dephasing rate γ_n. The longitudinal decay rate is denoted by Γ_n [17,18]. If we abbreviate the energy-level spacing by $\omega_n \equiv e_n - g_n$, the equations of motion become:

$$d \, S_{1,n}(t)/dt = - \omega_n \, S_{2,n}(t) - \gamma_n \, S_{1,n}(t) \tag{5.4a}$$

$$d \, S_{2,n}(t)/dt = \; \omega_n \, S_{1,n}(t) - 2 \, S_{3,n}(t) \, U(t) - \gamma_n \, S_{2,n}(t) \tag{5.4b}$$

$$d \, S_{3,n}(t)/dt = \; 2 \, S_{2,n}(t) \, U(t) - \Gamma_n \, [S_{3,n}(t) + 1] \tag{5.4c}$$

to be solved with $S_{1,n}(t=0) = S_{2,n}(t=0) = 0$ and $S_{3,n}(t=0) = -1$, meaning that all atoms are initially in their ground state. Note the non-additive coupling of the external field, due to the terms $S_{3,n}(t) \, U(t)$ and $S_{2,n}(t) \, U(t)$.

The goal would be here to construct the properties of an optimum laser pulse $U_{opt}(t)$, such that the final macroscopic dipole moment after the pulse $\Sigma_n \, S_{1,n}(T)$ is maximal. The resulting Hamiltonian of optimal control theory is given by

$$H = \Sigma_n \{\lambda_{1,n}(-\omega_n S_{2,n} - \gamma_n S_{1,n}) + \lambda_{2,n}(\omega_n S_{1,n} - 2S_{3,n}U - \gamma_n S_{2,n}) + \lambda_{3,n}(2S_{2,n}U - \Gamma_n [S_{3,n}+1])\} \tag{5.5}$$

which leads to the following co-state equations for the Lagrange functions $\lambda_{i,n}$

$$d\,\lambda_{1,n}(t)/dt = \gamma_n \lambda_{1,n}(t) - \omega_n \lambda_{2,n}(t) \tag{5.6a}$$

$$d\,\lambda_{2,n}(t)/dt = \omega_n \lambda_{1,n}(t) + \gamma_n \lambda_{2,n}(t) - 2\,U(t)\,\lambda_{3,n}(t) \tag{5.6b}$$

$$d\,\lambda_{3,n}(t)/dt = 2\,U(t)\,\lambda_{2,n}(t) - 2\,U(t)\,\lambda_{3,n}(t) + \Gamma_n \lambda_{3,n}(t) \tag{5.6c}$$

to be solved with $\lambda_{1,n}(T) = 1$ and $\lambda_{2,n}(T) = \lambda_{3,n}(T) = 0$. Coincidentally, the state and co-state equations are remarkable similar. While this set of equations is formally decoupled from the state equations, due to the terms $U\,\lambda_{2,n}$ and $U\,\lambda_{3,n}$ they do not permit any relationship between the solutions $\lambda_{1,n}(t)$ obtained for $U(t) = U_{opt,n}(t)$ and $U(t) = U_{opt}(t)$, which would be required by the SPSO to be valid exactly.

However, there are two important limiting cases, where the SPSO becomes valid. These two cases can be realized when either the damping Γ_n is sufficiently large, the pulse energy E (or equivalently the total interaction time T is less than the Rabi period of the atom) is short. Then we can assume that the inversion $S_{3,n}(t)$ differs only slightly from its initial value such that $S_{3,n}(t) = -1 + \varepsilon(t)$. Consequently, the last equation can be decoupled from $S_{1,n}(t)/dt$ and $S_{2,n}(t)/dt$, leading to

$$d\,S_{1,n}(t)/dt = -\omega_n S_{2,n}(t) - \gamma_n S_{1,n}(t) \tag{5.7a}$$

$$d\,S_{2,n}(t)/dt = \omega_n S_{1,n}(t) - \gamma_n S_{2,n}(t) + 2\,U(t) \tag{5.7b}$$

In order to test the reliability of this approximation with regard to the optimization scheme, we have calculated numerically the final polarization $S_{1,n}(T)$ for the n-th atom for the optimal laser pulse with a fixed total pulse energy $E = \int dt\, U(t)^2$. We compare the predictions according to the exact theory (5.6) with the approximation (5.7) as a function of the pulse energy E for $\omega_n = 2\pi$ and T=11.3.

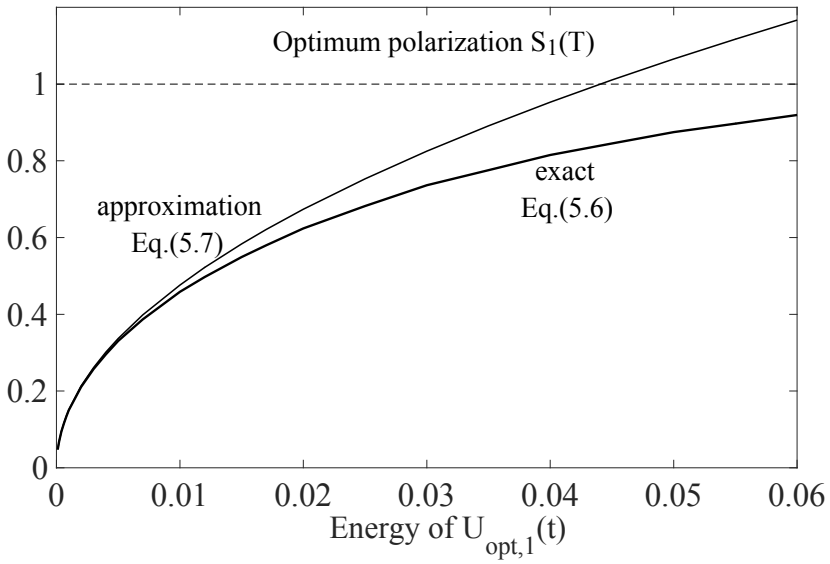

Figure 4 The final polarization $S_1(T)$ for the optimal pulse-shape for a driven two-level atom as a function of the fixed energy of the pulse. ($T=11.3$, $\omega_1=1/(2\pi)$, $\gamma_n=0$)

We find that the approximation (5.7) is surprisingly valid up to pulse energies E for which the polarization is almost 50% of its maximum total value [$S_1(T) = 1$]

Note that in this weak-field approximation the external laser field is coupled to the equations of motion in an *additive* manner. This means automatically that the superposition principle derived in Section 2 applies exactly if we optimize the collective macroscopic polarization at the end of the laser pulse, given by $S_1(t) \equiv \Sigma_n S_{1,n}(t)$. In this limit, the Hamiltonian takes the form $H = \Sigma_n \{\lambda_{1,n}(-\omega_n S_{2,n} - \gamma_n S_{1,n}) + \lambda_{2,n}(\omega_n S_{1,n} + 2U - \gamma_n S_{2,n})\}$ such that the co-state equations become

$$d\,\lambda_{1,n}(t)/dt = -\omega_n\,\lambda_{2,n}(t) + \gamma_n\,\lambda_{1,n}(t) \tag{5.8a}$$

$$d\,\lambda_{2,n}(t)/dt = \omega_n\,\lambda_{1,n}(t) + \gamma_n\,\lambda_{2,n}(t) \tag{5.8b}$$

which leads to the solution for $\lambda_{2,n}(t) = \mathrm{Exp}[\gamma_n(t-T)]\,\mathrm{Sin}[\omega_n(t-T)]$. Applying Eq. (A14) to construct the collective optimal control $U_{opt}(t)$ to maximize $S_1(t) \equiv \Sigma_n S_{1,n}(t)$ then takes the form $U_{opt}(t) = (2\lambda_0)^{-1} \Sigma_n\, 2\,\lambda_{2,n}(t)$, such that we have

$$U_{opt}(t) = (\lambda_0)^{-1} \Sigma_n\, \mathrm{Exp}[\gamma_n(t-T)]\,\mathrm{Sin}[\omega_n(t-T)] \tag{5.9}$$

If we assume an infinite ensemble of atoms with uniformly distributed transition frequencies ω_n, between ω_{min} and ω_{max} such that $\Sigma_n \rightarrow$ const. $\int d\omega$, and assume the same damping constants $\gamma_n \equiv \gamma$, the optimum pulse shape becomes

$$U_{opt}(t) = N \, Exp[\gamma(t-T)] \, \{Cos[\omega_{min}(t-T)] - Cos[\omega_{max}(t-T)]\}/(t-T) \qquad (5.10)$$

where the normalization constant guarantees that $\int dt \, U_{opt}(t)^2 = E$.

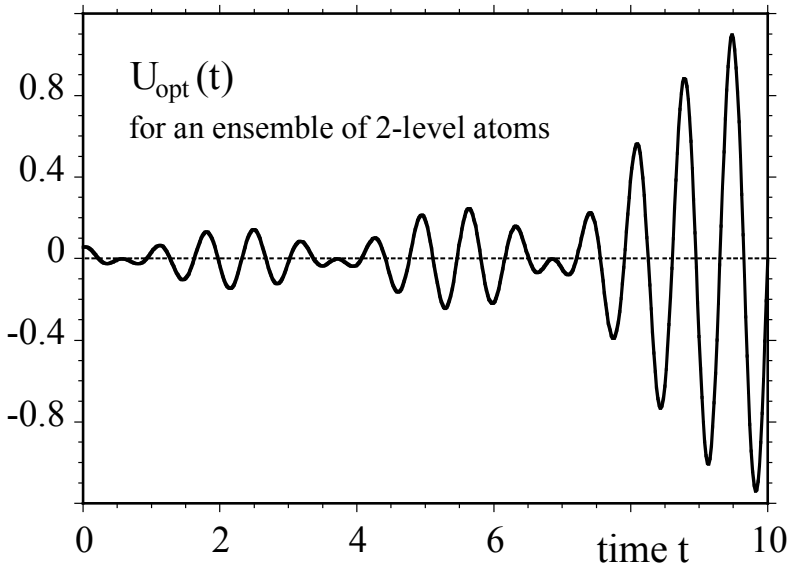

Figure 5 The optimal field $U_{opt}(t)$ (in units of $E^{1/2}$) to maximize the final macroscopic dipole moment of a distribution of two-level atoms with transition frequencies ω in the range $8 < \omega < 10$ and equal damping $\gamma = 0$.

In Figure 5 we provide a typical example for $U_{opt}(t)$ for $\gamma = 0$. As the damping is increased the oscillations at early times vanish and more of the pulses energy is shifted towards the final time T, approaching a single peaked-distribution close to t=T.

4. Summary and outlook

In summary, we have pointed out a superposition principle for the simultaneous optimization for the final collective response of a set of dynamical systems that permits us to predict the optimal control field $U_{opt}(t)$ from a weighted superposition of the optimal force fields $U_{opt,n}(t)$ of every single sub-system separately. The weights are inversely proportional to the temporal derivative at

the final time, $U'_{opt,n}(T)$. It is interesting that this principle can be exact for systems for which a sum of individual solutions to the state equations is not a solution, but it is never exactly valid for those systems whose solutions are additive. We have illustrated the SPSO for several systems and suggested that even for those systems, where it is not strictly valid (non-linear oscillators or two-level systems), it can provide under certain situations (small total energy of the control field or short interaction times) a surprisingly accurate guidance to predict the optimal control field $U_{opt}(t)$.

The SPSO opens the door to many interesting future studies, out of which we sketch here one. A research area where the SPSO would find an obvious and direct application is the predicted laser-induced decay of the quantum field theoretical vacuum state. Here it is predicted that the photon energy of an extremely focused laser pulse can be converted to the creation of electron-positron pairs [22]. Very recently, it was suggested by Kohlfürst et al. [23,24] and Hebenstreit and Fillion-Gourdeau [25] that optimal control theory could be applied to determine the time-dependence of the optimal electromagnetic field configuration that would lead to the largest number of created electron-positron pairs. Due to the enormous requirement on CPU time, most studies so far in this research have focused on a finite -dimensional optimization, where typically some phases, amplitudes or spatial length scales [26] where optimized. However, if the spatial inhomogeneity of electromagnetic field is neglected, which is sometimes non-trivial [27], then the interaction of the time-dependent electric field with the vacuum state can be exactly mapped onto an infinite set of uncoupled two-level systems, whose energy separation exceeds twice the rest mass energy of the electron, $\omega_{min}>2mc^2$. The resulting theoretical framework is then nearly identical to the one examined in Section 5, where the SPSO was shown to be ideally suited to provide accurate predictions for short interaction times.

Acknowledgements

SD and RF would like to thank ILP for the nice hospitality during their visits to Illinois State. RF acknowledges the support by the exchange program of Illinois State University and ITESM of Mexico. This work has been supported by the NSF, NSFC (#11529402, #11721091), Research Corporation, and the National Basic Research Program of China (2015CB859700). SD acknowledges the China Schlarship Council program.

Appendix A Optimal control theory for collective response

We consider N independent dynamical systems, each of which is characterized by I degrees of freedom, represented by the I-dimensional vector $\mathbf{Y}_n(t) \equiv \{Y_{1,n}(t), Y_{2,n}(t), \ldots, Y_{I,n}(t)\}$. We also assume that each vector has to satisfy the corresponding equations:

$$d\mathbf{Y}_n/dt = \mathbf{K_n}[\mathbf{Y}_n, U(t)] \tag{A1}$$

where the I-component vector $\mathbf{K_n}$ can be a non-linear function of $\mathbf{Y}_n$ and the external force U(t). We assume that we know the initial values $\mathbf{Y}_n(t=0)$.

The goal is to construct the best possible pulse-shape U(t) such that a certain collective quantity of all N systems, such as a certain weighted superposition of the final amplitudes at a given time T, becomes maximal. In other words, we might want to optimize the objective J"

$$J'' \equiv \Sigma_n \Sigma_i w_{i,n} Y_{i,n}(T) \tag{A2}$$

where we can choose specific factors $w_{i,n}$ to possibly weight the contributions of each amplitude $Y_{i,n}$ of the n-th system in a desired way.

In order to limit the possible range of forces U(t), we consider only those with a given "energy" E, i.e., U(t) has to fulfill the constraint $E = \int dt\, U(t)^2$, where the integration limits on all temporal integrals are from now on from t=0 to t=T. We can introduce the single Lagrange multiplier λ_0 to account for this constraint and obtain a new objective J'

$$J' = \Sigma_n \Sigma_i w_{i,n} Y_{i,n}(T) + \lambda_0 [E - \int dt\, U(t)^2] \tag{A3}$$

Furthermore, we introduce a set of Lagrange-functions $\boldsymbol{\lambda}_n(t)$ to account for the equations of motions (A1) as constraints. As each system has several degrees of freedom, the notation $\boldsymbol{\lambda}_n(t)$ refers to a I-component vector with the individual components denoted by $\lambda_{i,n}(t)$. We therefore we arrive at the objective J, which is a function of λ_0 and a functional of U(t), $\mathbf{Y}_n(t)$ and $\boldsymbol{\lambda}_n(t)$.

$$J = \Sigma_n \left\{\Sigma_i w_{i,n} Y_{i,n}(T) + \int dt \left[\boldsymbol{\lambda}_n (\mathbf{K}_n - d\mathbf{Y}_n/dt)\right\} + \lambda_0 [E - \int dt\, U(t)^2\right] \quad (A4)$$

It turns out that the notation can be simplified if we define a collective Hamiltonian according to

$$H \equiv \Sigma_n \boldsymbol{\lambda}_n \mathbf{K}_n = \Sigma_n \Sigma_i \lambda_{i,n} K_{i,n} \quad (A5)$$

then the objective functional reduces to

$$J = \Sigma_n \left\{\Sigma_i w_{i,n} Y_{i,n}(T) - \int dt\, \boldsymbol{\lambda}_n d\mathbf{Y}_n/dt\right\} + \lambda_0 [E - \int dt\, U(t)^2] + \int dt\, H \quad (A6)$$

If we compute the variation of the objective δJ, we obtain

$$\delta J = \Sigma_n \left\{\Sigma_i w_{i,n} \delta Y_{i,n}(T) - \int dt \left[\delta\boldsymbol{\lambda}_n d\mathbf{Y}_n/dt + \boldsymbol{\lambda}_n \delta[d\mathbf{Y}_n/dt]\right]\right\}$$
$$+ \delta\lambda_0 [E - \int dt\, U(t)^2] - \lambda_0 \int dt\, 2U\delta U] + \int dt\, \delta H \quad (A7)$$

After integration the terms containing $\delta[d\mathbf{Y}_n/dt]$ by parts, and using $\delta\mathbf{Y}_n(t=0)=0$, due to the fixed initial conditions, we write $-\boldsymbol{\lambda}_n \delta\mathbf{Y}_n = -\Sigma_i \lambda_{i,n} \delta Y_{i,n}$ and obtain

$$\delta J = \Sigma_n \left\{\Sigma_i[w_{i,n} - \lambda_{i,n}(T)]\delta Y_{i,n}(T) - \int dt \left[\delta\boldsymbol{\lambda}_n d\mathbf{Y}_n/dt - d\boldsymbol{\lambda}_n/dt\, \delta\mathbf{Y}_n\right]\right\}$$
$$+ \delta\lambda_0 [E - \int dt\, U(t)^2] - \lambda_0 \int dt\, 2U\delta U] + \int dt\, \delta H \quad (A8)$$

Next, the variation of H amounts to

$$\delta H = (\partial H/\partial U)\, \delta U + \Sigma_n \left\{(\partial H/\partial \mathbf{Y}_n)\, \delta\mathbf{Y}_n + (\partial H/\partial\boldsymbol{\lambda}_n)\, \delta\boldsymbol{\lambda}_n\right\} \quad (A9)$$

Using the specific definition of H, the partial derivatives simplify to $\partial H/\partial\boldsymbol{\lambda}_n = \mathbf{K}_n$. If we want that the variation δJ vanishes for the optimal solutions, we have to require that each co-factor of each variation (δU, $\delta\lambda_0$, $\delta\boldsymbol{\lambda}_n$, $\delta\mathbf{Y}_n$, and $\delta\mathbf{Y}_n(T)$) vanishes, i.e., we obtain,

$$-\lambda_0\, 2U + \partial H/\partial U = 0 \tag{A10}$$

$$E - \int dt\, U(t)^2 = 0 \tag{A11}$$

$$d\mathbf{Y}_n/dt = \partial H/\partial \boldsymbol{\lambda}_n = \mathbf{K}_n \qquad \text{with } \mathbf{Y}_n(0) = \mathbf{Y}_n(t=0) \tag{A12}$$

$$d\boldsymbol{\lambda}_n/dt = -\partial H/\partial \mathbf{Y}_n = -\boldsymbol{\lambda}_n\, \partial \mathbf{K}_n/\partial \mathbf{Y}_n \qquad \text{with } \boldsymbol{\lambda}_n(T) = \mathbf{w}_n \tag{A13}$$

Eq. (A10) is the main control equation and can be rewritten as $U(t) = (2\lambda_0)^{-1}\, \partial H/\partial U$. Together with Eq. (A5) it leads to

$$U(t) = (2\lambda_0)^{-1} \Sigma_n\, \boldsymbol{\lambda}_n\, \partial \mathbf{K}_n(\mathbf{Y}_n, U)/\partial U \tag{A14}$$

This is in general a very complicated and possibly even transcendental equation that U(t) has to fulfill. This is especially true as each system can be coupled to the force U(t) in a functionally different way. The single Lagrange parameter $(2\lambda_0)^{-1}$ needs to be chosen such that Eq. (A11) is satisfied, i.e., $E - \int dt\, [(2\lambda_0)^{-1} \Sigma_n\, \boldsymbol{\lambda}_n\, \partial \mathbf{K}_n(\mathbf{Y}_n, U)/\partial U]^2 = 0$. Therefore, we have the energy normalization factor

$$(2\lambda_0)^{-1} = E^{1/2} \{\int dt\, [\Sigma_n\, \boldsymbol{\lambda}_n\, \partial \mathbf{K}_n(\mathbf{V}_n, U)/\partial U]^2\}^{-1/2} \tag{A15}$$

Appendix B The superposition principle for collective responses

Let us now examine special cases where the right-hand side of the equation of motion (A12) is given by simple matrices. The system we consider is $d\mathbf{Y}_n/dt = M_n\,\mathbf{Y}_n + \mathbf{V}_n\,U(t)$. The temporally constant vector $\mathbf{V}_n$ permits us to couple the external force to each amplitude $Y_{i,n}$ differently. Here the matrix M_n is time-independent, but the differential equation is inhomogeneous such that the sum of two individual solutions is in general not a solution to the same differential equation. However, due to the special coupling to the external force as a source term, this system can always be solved exactly based on the diagonalization of the matrix M.

The Hamiltonian reduces here to $H= \Sigma_n\ \boldsymbol{\lambda}_n\,[M_n\,\mathbf{Y}_n + \mathbf{V}_n\,U(t)]$, such that the control equation (A14) simplifies significantly to

$$U(t) = (2\lambda_0)^{-1}\,\Sigma_n\,\boldsymbol{\lambda}_n(t)\,\mathbf{V}_n \qquad (B.1)$$

In this case, the co-state equation also simplifies significantly. If we rewrite the products in terms of its components, $\boldsymbol{\lambda}_n\,M_n\,\mathbf{Y}_n = \Sigma_i\,\lambda_{i,n}\,\Sigma_j\,M_{i,j,n}\,\mathbf{Y}_{j,n}$, then $\partial H/\partial Y_{k,n}$ leads to $\partial H/\partial Y_{k,n} = \Sigma_i\,\lambda_{i,n}\,M_{i,k,n} = \Sigma_i\,\lambda_{i,n}\,M^{T}_{k,i,n}$, where the superscript T denotes the transposed matrix. In other words, we obtain $\partial H/\partial\mathbf{Y}_n = M_n^{\,T}\,\boldsymbol{\lambda}_n$, such that the resulting co-state equation $d\boldsymbol{\lambda}_n/dt = -\,M_n^{\,T}\,\boldsymbol{\lambda}_n$ (with $\boldsymbol{\lambda}_n(T) = \mathbf{w}_n$) no longer depends on U(t). In other words, the transcendental equation (A14) for the optimum force field becomes therefore a solution solely in terms of the relevant Lagrange functions $\boldsymbol{\lambda}_n(t)$ and $\mathbf{V}_n$. This observation has significant implications with regard to the ultimate goal of predicting the optimum force U(t) for the collective response for all N systems from the N optimum forces associated with each system separately, denoted by $U_{opt,n}(t)$ and given by

$$U_{opt,n}(t) = (2\lambda_{0,n})^{-1}\,\boldsymbol{\lambda}_n(t)\,\mathbf{V}_n \qquad (B.2)$$

The collective optimum function $U_{opt}(t)$ can therefore be rewritten as a linear superposition of the individual optimizers $U_{opt,n}(t)$ with individual weight factors $(2\lambda_{0,n}) \equiv a_n$.

$$U_{opt}(t) \;=\; (2\lambda_0)^{-1}\,\Sigma_n\,a_n\,U_{opt,n}(t) \qquad (B.3)$$

As a side issue, we remark that the normalization $E = \int dt\, U_{opt,n}(t)^2 = \int dt\, [(2\lambda_{0,n})^{-1} \boldsymbol{\lambda}_n \mathbf{V}_n]^2$ leads to $(2\lambda_{0,n}) = E^{-1/2} [\int dt\, (\boldsymbol{\lambda}_n \mathbf{V}_n)^2]^{1/2}$.

As the Lagrange functions $\boldsymbol{\lambda}_n(t)$ are just mathematical auxiliary functions and cannot be measured directly, we have to show how the weight factors $a_n \equiv (2\lambda_{0,n})$ are related to the observed $U_{opt,n}(t)$. The time derivative of $U_{opt,n}(t)$ evaluated at the final time T is given by

$$\begin{aligned} dU_{opt,n}(t)/dt|_{t=T} &\equiv U'_{opt,n}(T) \\ &= (2\lambda_{0,n})^{-1} (d\boldsymbol{\lambda}_n(t)/dt|_{t=T})\, \mathbf{V}_n \\ &= -(2\lambda_{0,n})^{-1} (M_n{}^T \boldsymbol{\lambda}_n|_{t=T})\, \mathbf{V}_n \\ &= -(2\lambda_{0,n})^{-1} (M_n{}^T \mathbf{w}_n)\, \mathbf{V}_n \end{aligned} \qquad \text{(B.4)}$$

such that the weight factors can be constructed via $a_n = (2\lambda_{0,n}) = -(M_n{}^T \mathbf{w}_n)\, \mathbf{V}_n / U'_{opt,n}(T)$. We therefore arrived at the final form of the superposition principle for collective responses. As one might expect, the a_n depend linearly on the weight factors $w_{i,n}$ associated with each amplitude $Y_{i,n}$ for the n-th system. It is interesting to note that once the optimum field $U_{opt,n}(t)$ is measured for each system, also its temporal derivative at the end point T needs to be calculated in order to predict the optimum force field for the collective response. The larger the slope of the individual optimizer $U_{opt,n}(t)$ is at the final moment in time, the less $U_{opt,n}(t)$ will contribute to the optimizer of the collective system.

References

[1] D.E. Kirk,"Optimal control theory: an introduction" (Prentice-Hall, Englewood Cliffs, 1970).

[2] R. Dechter, "Constraint processing" (Morgan Kaufmann, Burlington, 2003).

[3] D.S. Naidu, "Optimal control systems" (CRC Press, Baca Raton, LA, 2003).

[4] J.J. Leader, "Numerical analysis and scientific computation" (Addison Wesley, 2004).

[5] R. Battiti, M. Brunato and F. Mascia, "Reactive search and intelligent optimization" (Springer Verlag, Heidelberg, 2008).

[6] W. Sun and Y.X. Yuan, "Optimization theory and methods: nonlinear programming" (Springer Verlag, Heidelberg, 2010).

[7] A. Gosavi, "Simulation-based optimization" (Springer Verlag, Heidelberg, 2015).

[8] D. Bertsekas, "Dynamic programming and optimal control" (Athena Scientific, 2017).

[9] J. Werschnik and E.K.U Gross, J. Phys. B 40, R175 (2007).

[10] A.G. Butkovskiy and Y.I. Samoilenko, "Control of Quantum Mechanical Processes and Systems" (Kluwer Academic, Dordrecht, 1990).

[11] J. Zabczyk, "Mathematical Control Theory" (Birkhauser, Basel, Switzerland, 1990).

[12] T. Brixner, G. Krampert, T. Pfeifer, R. Selle, G. Gerber, M. Wollenhaupt, O. Graefe, C. Horn, D. Liese and T. Baumert, Phys. Rev. Lett. 92, 208301 (2004).

[13] V. Ramakrishna, R. Ober, X. Sun, O. Steuernagel, J. Botina and H. Rabitz, Phys. Rev. A 61, 032106 (2000).

[14] S.G. Schirmer, H. Fu and A.I. Solomon, Phys. Rev. A 63, 063410 (2001).

[15] W.H. Press, S.A. Teukolsky, W.T.Vetterlin and B.P.Flannery, "Numerical Recipes: The Art of Scientific Computing" (Cambridge University Press, New York, 1992).

[16] R. Fletcher and C.M. Reeves, Comp. J. 7, 149 (1964).

[17] S. Glasgow, M. Meistrup, J. Peatross and M. Ware, Phys. Rev. E 75, 016616 (2007).

[18] S. Glasgow and M. Ware, Phys. Rev. A 80, 043817 (2009).

[19] S. Glasgow, J. Corson and C. Verhaaren, Phys. Rev. E 82, 011115 (2010).

[20] L. Allen and J.H. Eberly, "Optical Resonance and Two-Level Atoms" (John Wiley & Sons, Inc., New York, 1975).

[21] P. Milonni and J.H. Eberly, "Lasers" (Wiley, New York, 1988).

[22] For a recent comprehensive review, see, B.S. Xie, Z.L. Li, S. Tang, Matter and Radiation at Extremes 2, 225 (2017).

[23] C. Kohlfürst, M. Mitter, G. von Winckel, F. Hebenstreit and R. Alkofer, Phys. Rev. D 88, 045028 (2013).

[24] C. Kohlfürst, Master's thesis, Graz University (2012), arXiv:1212.0880.

[25] F. Hebenstreit and F. Fillion-Gourdeau, Phys. Lett. B 739, 189 (2014).

[26] S.S. Dong, M. Chen, Q. Su and R. Grobe, Phys. Rev. A 96, 032120 (2017).

[27] Q.Z. Lv, S. Dong, Y.T. Li, Z.M. Sheng, Q. Su and R. Grobe, Phys. Rev. A 97, 022515 (2018).