Genetic algorithm with migration on topology conserving maps for optimal control of quantum systems

Bjarne Amstrup[2], Gábor J. Tóth[1], Gábor Szabó[3], Herschel Rabitz[4], and András Lőrincz[1]

[1] *Department of Photophysics, Institute of Isotopes of the Hungarian Academy of Sciences, P.O. Box 77, H-1525 Budapest, Hungary*

[2] *Chemistry Department B, Technical University of Denmark, DTU–206, DK–2800 Lyngby, Denmark*

[3] *Department of Optics and Quantum Electronics, József Attila University Dóm tér 9, H-6720 Szeged, Hungary*

[4] *Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*

(August 4, 1995)

Abstract

The laboratory implementation of molecular optimal control has to overcome the problem caused by the changing environmental parameters, such as the temperature of the laser rod, the resonator parameters, the mechanical parameters of the laboratory equipment and other dependent parameters such as the time delay between pulses or the pulse amplitudes. In this paper a solution is proposed: instead of trying to set the parameter(s) with very high precision, their changes are monitored and the control is adjusted to the current values. The optimization in the laboratory can then be run at several values of the parameter(s) with an extended genetic algorithm (GA) which is tailored to such parametric optimization. The extended GA does not presuppose, but can take advantage, and in fact, explores whether the mapping from the parameter(s) to optimal control field is continuous. Then the op-

timization for the different values of the parameter(s) is done cooperatively which reduces the optimization time. A further advantage of the method is its full adaptiveness, i.e., in the best circumstances no information on the system or laboratory equipment is required, only the success of the control needs to be measured. The method is demonstrated on a model problem: a pump-and-dump type model experiment on CsI.

I. INTRODUCTION

Optimal control of molecular processes with shaped ultrashort laser pulses is the focus of much recent research interest (see, e.g., [1–53]).The problem has been investigated extensively through computer simulations, but only few experimental demonstrations have been done [54,55]. One of the reasons that laboratory realization of molecular optimal control is being cautiously approached is the sensitivity of the control to the optimized field. The first attempts to design the appropriate control fields through computer simulations used a model of the molecule being controlled, and similarly, the experimental conditions were only modeled. The error introduced by the necessarily incomplete modeling and the imperfect realization of the optimal field is difficult to judge and little effort was made in this direction.

This problem can be circumvented by the use of adaptive feedback optimal control (AFOC) [22]. AFOC designs the optimal control field in the laboratory, using the actual laboratory equipment and the molecule itself. The general setup of AFOC is the following: propose a control field, apply that field (in the laboratory) to the molecule, measure the success of control, and use that measure to improve the control field. The use of the laboratory equipment in the optimization process gets rid of most of the problems caused by systematic errors. Moreover, as the method does not use computer simulations it can be much faster than the conventional method which had to solve the Schrödinger equation of the molecule being controlled numerically many times.

The choice of the optimization procedure used in AFOC has proven to be a key question.

The quality of the control is often a very irregular function of the control parameters with a large number of optima [56]. Optimization algorithms looking for the closest local optimum are thus not well suited to AFOC of molecular systems; global search methods should be used.

Despite its obvious advantages AFOC is not an answer to all issues. In particular, the problem posed by instabilities in the laboratory equipment needs special care. Instabilities over different time scales pose different problems. Long time scale instabilities, i.e., aging phenomena, can be compensated by the adaptivity alone. Very fast changes constitute a kind of noise, and with the proper choice of optimization algorithm some noise can be allowed [52]. There is an intermediate region, which is too slow to be treated as noise, but too fast to be compensated by the adaptivity alone. In the present paper we propose an adaptive method which can handle the problem of fast and medium speed drifts, too.

There are many physical parameters which are easier to measure than to set with sufficient precision. A typical example might be the temperature of the laser rod, the resonator parameters, the mechanical parameters of the laboratory equipment and other dependent parameters such as the time delay between pulses or the pulse amplitudes. A way of handling the drifts of such parameters is to measure them and adjust the control to their current values. This requires, however, that the optimal control field be known as a function of the environmental parameters. In practice, this function has to be approximated in some way. The most straightforward is to divide the input space (the space of the environmental variables) into small regions, and use the same control setting within each region (but possibly different in the different regions).

A parametric optimization method based on the above setup must be composed of two sub-systems: one performing the discretization (setting up the regions within which the function is taken to be constant) and one finding the optimal values of the control parameters within the regions. There are a number of algorithms which can perform the discretization or the optimization. To have an efficient and flexible approach we need a discretization algorithm that is adaptive, i.e., that can set up the regions without any *a-priori* knowledge

and can modify them if, possibly due to aging, the distribution of the input points changes. Also, the optimization in the different regions should be done in a cooperative manner. In most cases the optimal control settings are a continuous function of the environmental variables, i.e., in regions close to each other the optimal control settings are likely to be similar. The optimization can be speeded up if the algorithm can detect and exploit such continuity.

In conclusion, we need a parametric optimization procedure which can adaptively discretize the input space and find the optimal control settings, which uses a noise tolerant global search method, and which can explore whether the optimal control field is a continuous function of the environmental parameters and if so, can make use of it. In the following section we describe such a method. This is followed by the description of a model experiment used to demonstrate the algorithm in Sec. III. The results and their discussion are given in Sec. IV, the conclusions are drawn in Sec. V.

II. THE ALGORITHM

There are several optimization methods which have been used to drive AFOC. Recent studies have used, e.g., genetic algorithm (GA, [22]) and simulated annealing (SA, [28]). They are both global search methods but the GA was found to tolerate noise better [52]. The parametric optimization algorithm used in this study is the genetic algorithm with migration (GAM, [57,58]). In the following we briefly describe the two constituent parts of GAM: the GA and the so called Kohonen network [59,60] and the GAM itself. A more detailed description of the algorithm can be found in the Appendix.

The GA [61–63] is a biologically inspired heuristic which maintains a "population": a number of so called individuals, i.e., possible solutions of the control problem. The flow of the algorithm is the following. The individuals are initialized either randomly or with bias if prior knowledge is available. The quality of control coded by each individual is tested by performing it, then the population is updated with the help of so called "genetic operators".

The optimization then continues with testing the new "generation", until a sufficiently good solution is found.

The other component of the GAM is the Kohonen network [59]. The Kohonen network contains a number of units (called neurons), arranged on a d dimensional lattice, where d is usually the number of environmental parameters. Each unit stores a reference point from the input space, i.e., the space of the environmental parameter(s). At the beginning the reference points are chosen randomly. During optimization sample points from the input space are presented to the network. When an input is presented to the network it activates the neuron whose reference point lies closest to the input; activation in our case means that the optimization procedure associated with that neuron will be active. At the same time the network adjusts itself: the reference point of each neuron moves closer to the presented point. The amplitude of the move made by the neurons is a decreasing function of their distance from the active neuron.

The GAM contains a Kohonen network performing the discretization of the input space. Each neuron of the network is associated with a GA procedure; the population of each GA procedure is called a "subpopulation" and the word population is reserved for all the individuals from all the subpopulations. The cooperativeness of GAM is expressed by genetic operator migration occasionally exchanging a pair of individuals at neighboring subpopulations.

The migration shows advantageous properties in many respects. Most importantly, a good solution found by a GA procedure can spread by migrating from one site to the other as long as it is successful in the new site; this increases the speed of optimization significantly, both directly and by allowing smaller subpopulations. Another benefit of migration is that it reduces the fluctuation of the optimal solution considerably.

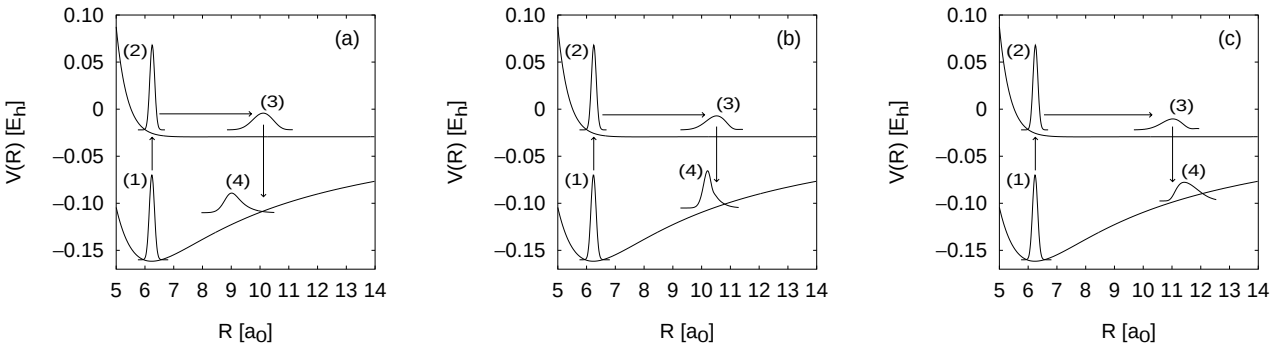

FIG. 1. Upper and ground state potentials of CsI along with a schematic representation of the wave function evolution under the effect of the optimized control fields for the minimal (a), and intermediate (b) and the maximal (c) value of the time delay (τ). Curves 1–4 show the probability density on the ground state before the pump pulse (1), the probability density on the upper surface after the pumping (2) and at the peak of the dumping pulse (3), and probability density on the ground surface after the dumping pulse (4). The arrows show the flow of time, and the packet (4) is at the same total time in all the cases.

III. THE MODEL EXPERIMENT

In the following we apply the above approach to the optimal control of CsI. The control task is a pump-and-dump type experiment: the population is first transferred to the excited dissociative state and after a delay it is dumped back to the electronic ground state. Fig. 1 show the concept of the fragmentation experiment. The molecule is modeled by taking into account two electronic states (BO surfaces).

Although the focus of this paper is the use of parametric optimization it may be worth noting that the above setup easily lends itself to laboratory implementation. The initial excitation can be done with a doubled green (i.e., UV) pulse and the dumping can be achieved by a chirped green pulse. The success of the control can be determined: either of the fragments can be easily detected. Finally, the experiment can be run in a cold molecular beam.

The pump and dump pulses are optimized separately. The pump pulse was optimized as in [45] leading to 95% inversion. The objective of this optimization was

$$I_1 = \int_0^\infty |\psi_u(x,t_1)|^2 dx - \lambda_s \int_{t_0}^{t_1} \mathcal{E}_s^2(t) dt. \tag{1}$$

$\mathcal{E}_s(t)$ was chosen as a sine squared electric field of a given pulse duration, τ_s:

$$\mathcal{E}_s(t) = \mathcal{E}_0^p \sin^2(\pi t/\tau_s) \left[a_1 \cos(\omega_s t) + a_2 \sin(\omega_s t)\right]. \tag{2}$$

The parameters optimized in the SA procedure were $\mathcal{E}_0^p$, a_1, a_2 and ω_s. The pulse duration was chosen to be 20 fs and λ_s was 0.05. The result of the optimization was $\omega_s = 30507\,\mathrm{cm}^{-1}$ and $\mathcal{E}_0^p = 0.0252\,E_h/(ea_0) = 1.296 \times 10^{10}\,\mathrm{V/m}$ corresponding to a maximum intensity $I = \varepsilon_0 c(\mathcal{E}_0^p)^2 = 45\,\mathrm{TW/cm}^2$.

The dumping was optimized by GAM, assuming either the time delay or the amplitude drifting and thus using it as the environmental parameter.

The objective of the optimization was set to

$$I = \int_0^\infty |\psi_g(x,t_f)|^2 dx - \lambda \int_{t_1}^{t_f} \mathcal{E}^2(t) dt. \tag{3}$$

The first term describes the real objective, the second term limits the electric field energy used in the optimization procedure. In the equation ψ_g denotes the electronic ground state wave function and $\mathcal{E}(t)$ is the dump electric field. Parameters t_1 and t_f denote the end of the pumping pulse and the target time, respectively. The balance of the two terms may be tuned with the help of the penalty factor: higher λ values correspond to a "stricter policy" on the use of electric field energy.

To facilitate the successful optimization the possible control pulses should be coded in a way that is (i) experimentally realizable and (ii) compact, i.e., a small number of parameters suffices. Recent studies [28,64,65] suggest that the chirp expansion is such a coding scheme. The chirp expression of an electric field is

$$\mathcal{E}(t) = \Re\left\{ A\sqrt{\frac{\tau}{2}} \int \exp\left[i\omega t - \frac{\tau}{4}(\omega-\omega_0)^2 - iz \sum_{k=0}^{\infty} \frac{\beta^{(k)}}{k!}(\omega-\omega_0)^k \right] d\omega \right\}. \tag{4}$$

The above expression reflects a possible experimental realization of the pulse: it can be achieved by propagating a Gaussian pulse of the form

$$\mathcal{E}_0^d(t) = \Re\left\{A\exp(-t^2/\tau + i\omega_0 t)\right\}, \tag{5}$$

through a medium with a frequency dependent propagation factor β (e.g., an optical fiber) of length z. In both equations $\Re$ denotes the real part of the complex number. The chirp expansion, in principle, is a full expansion, i.e., if one can adjust the Taylor coefficients $\beta^{(k)}$ of the propagation factor as desired then any pulse can be realized. Present day experimental possibilities limit the expansion to the first few terms, but some examples suggest that even three terms give substantial flexibility [28]. The zeroth and first derivatives of the propagation factor (i.e., $\beta^{(0)}$ and $\beta^{(1)}$) give a phase factor and a group delay, the second ($\beta^{(2)}$) causes a linear shift in the central frequency and the third ($\beta^{(3)}$) gives rise to some beating in the pulse. The optimization was run to find the optimal values for the central frequency (ω_0), the second and third order chirp ($z\beta^{(2)}$ and $z\beta^{(3)}$), and the amplitude (A) or the time delay ($z\beta^{(1)}$) of the pulse, as the zeroth derivative gives only a phase factor. It should be noted that chirp control can be easily made adaptive through on-line control, since the elements one controls are the positions and directions of prisms and gratings [66].

It should be understood that there is nothing special about the chirp expansion that makes it especially suitable for GAM. There are other schemes, such as pulse shaping with a many-element liquid crystal phase modulator [67], which are just as good as the chirp expansion in laboratory AFOC. However, they involve the optimization of many more parameters and present day computational power is insufficient for the simulations: even the studies reported here took several days on an IBM-RS550 machine.

IV. RESULTS AND DISCUSSION

The time dependent Schrödinger equation was solved numerically by a split operator (SPO) method [68,69] that uses the grid method and fast Fourier transformation

TABLE I. Parameters of potentials and grid used in the simulation of CsI dynamics [70].

Ground potential	Excited potential	Other parameters
$V_g = ae^{-\beta r} - \frac{E_h}{r-r_1}$	$V_u = e^{-d(r-r_e)}E_h - \frac{C_6}{(r-r_2)^6} + V_\infty$	Dipole moment: $\mu = 0.3\,ea_0$
$a = 131.569\,E_h$	$d = 2.571\,02\,a_0^{-1}$	Mass of Cs atom: $m_{\rm Cs} = 132.909\,{\rm amu}$
$\beta = 1.370\,42\,a_0^{-1}$	$r_e = 4.20\,a_0$	Mass of I atom: $m_{\rm I} = 126.904\,{\rm amu}$
$r_1 = 0.893\,631\,a_0$	$C_6 = 10.0\,E_h a_0^6$	Number of grid points: $N = 256 - 1024$
	$r_2 = 1.889\,72\,a_0$	Spacing of grid points: $\Delta r = 0.016\,5\,a_0$
	$V_\infty = -0.029\,399\,21\,E_h$	

for computing the Laplacian operator. All calculations were carried out using atomic units and are mostly reported in terms of $E_h = 219\,474.64\,{\rm cm}^{-1}$, $\hbar/E_h = 0.024\,188\,8\,{\rm fs}$, $a_0 = 5.291\,77 \times 10^{-11}\,{\rm m}$ and $E_h/(ea_0) = 5.142\,21 \times 10^{11}\,{\rm V/m}$.

Parameters for the CsI BO surfaces and the numerical algorithm are given in table I. The initial ground vibrational state for the CsI dynamics was found by the method of Kosloff and Tal-Ezer [71] through integration of the time dependent Schrödinger equation in imaginary time.

The GAM was used to perform the optimization of the CsI molecule with either the time delay ($z\beta^{(1)}$) or the amplitude (A) taken as drifting. In the following we give the outcome of the optimization and a discussion of the results. The values of the parameters characterizing the GAM are shown in Table II, and the penalty factor for the electric field (λ) was 0.01.

First the time delay was taken as drifting. The GAM was used to optimize the amplitude, the central frequency, and the second and third order chirp, and the result is plotted in Fig. 2. The optimization was conducted with and without the migration present. It is well pronounced in the figures that the migration leads to enhanced performance. With migration the algorithm dumped practically the total population; less than 4% of the population remained on the excited surface. In the same number of trials without migration, even the best control solution failed to dump 15% of the population. It should be noted that the

TABLE II. Parameters concerning GAM and discretization. The exact meaning of the parameters are given in the Appendix. If two values are given separated by a slash then the first is for the time delay, the second for the amplitude discretized optimization.

Initial width of Gaussian (σ_i)	1.0
Final width of Gaussian (σ_f)	0.001
Initial learning rate (ε_i)	0.1
Final learning rate (ε_f)	0.1
Number of discretization steps (S_d)	40 000
Number of best individuals offspring (α)	2.5
Probability of cross over (p_x)	0.7
Probability of mutation (p_{mut})	0.01/0.001
Probability of migration (p_{mig})	0.01
Number of subpopulations (N_s)	20
Number of individuals per subpopulation	5
FWHM of pulse in equation (5) ($\sqrt{4 \ln 2}\,\tau$) [fs]	20
Range in which the time delay ($z\beta^{(1)}$) varied [fs]	330–410
Range in which the amplitude (A) varied [$E_h/(ea_0)$]	0.0065–0.0323/0.0155–0.0233
Range in which the second order chirp ($z\beta^{(2)}$) varied [$(\hbar/E_h)^2$]	$\pm 10^6$
Range in which the third order chirp ($z\beta^{(3)}$) varied [$(\hbar/E_h)^3$]	$\pm 3 \times 10^8$
Range in which the central frequency (ω_0) varied [$E_h/\hbar$]	0.0665–0.0735

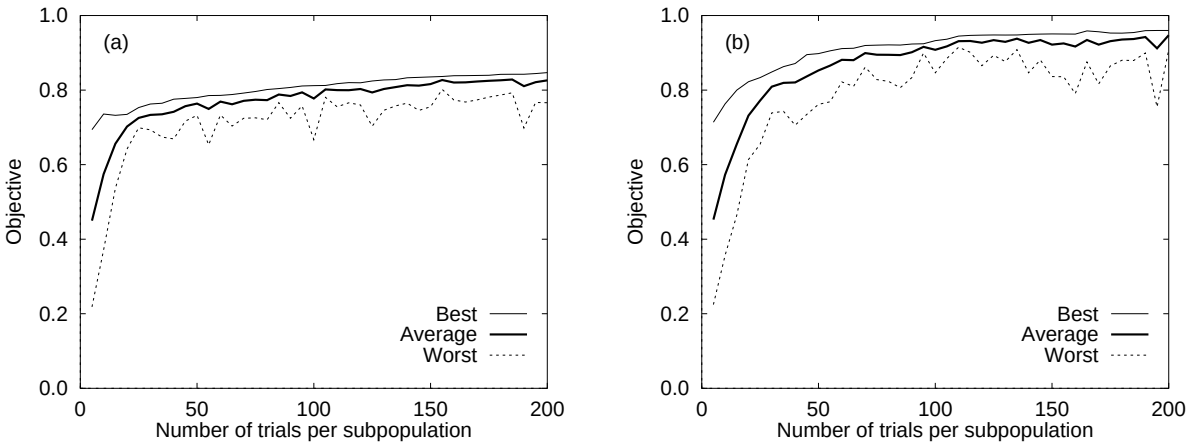

FIG. 2. The performance of the optimization with the time delay taken as drifting, without (a) and with (b) migration. The figures show the best, average and worst performance obtained in each generation averaged along the subpopulations. The parameters being optimized were the amplitude, the second and third order chirp and the central frequency. The effect of migration is well pronounced in the figures.

population remaining on the excited surface can be further decreased by allowing larger amplitudes.

A typical optimal pulse is shown in Fig. 3 along with its Husimi transform. The Husimi transform of the electric field is defined as

$$G(\omega, t) \propto \int_{-\infty}^{\infty} e^{-i\omega(t'-t)} e^{-\gamma(t'-t)^2} \mathcal{E}(t') dt', \tag{6}$$

where γ characterizes the fineness of the transformation, here it is chosen as 2×10^{-6} au. Both plots reveal the significant presence of chirp. The central frequency of the pulse is increasing in time (blue chirp), despite the fact that the difference between the potential surfaces is decreasing with increasing nuclear separation distance. This is characteristic of the strong field optimization of population inversion: the strong pulse keeps the upper and lower state wave functions in a narrow region for efficient population transfer. For weak fields the chirp is red [45]. This effect, known as laser induced avoided crossing, has been extensively studied [72].

A note should be made on the parameters of the GAM. The number of individuals in

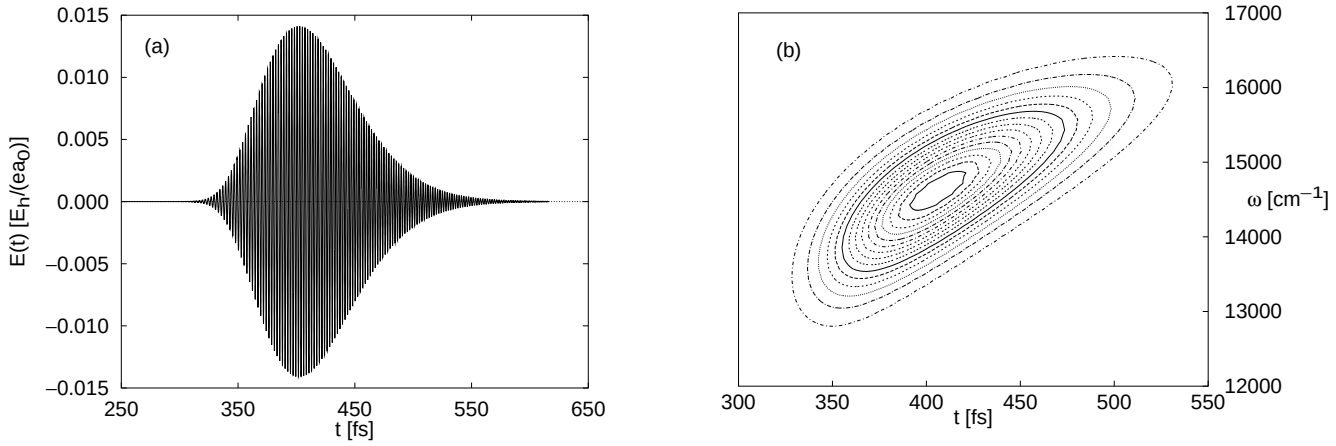

FIG. 3. A typical optimal pulse (a) and its Husimi transform (b).

each subpopulation was fairly low (5). This was possible because the problem domain was relatively easy for the GA. To compensate for this low number of individuals a high level of mutation ($p_{mut} = 0.01$) was used. In general larger subpopulations must be used, which in turn allows a smaller level of mutation to be applied. In practice all these parameters should be tuned to give the best performance.

The optimal solutions found by the GAM are listed in Table III. The effect of migration is well pronounced: solutions, or parts of the solutions are shared by subpopulations corresponding to similar values of the time delay. In some cases very different solutions are found by close subpopulations. This reflects that the problem is easy for the GA as multiple solutions with similar fitness exist.

The optimization was repeated with the amplitude drifting, optimizing the time delay, the central frequency, and the second and third order chirp. The results concerning the role of mutation are similar (see Fig. 4). The overall performance is lower, however. This can be understood by noticing that the inversion favors high amplitude. Some inversion can be achieved with lower amplitude as well, but not as high as with high amplitude (see Fig. 5). The performance plotted in Fig. 4 is the average for all amplitudes which explains its relatively low level.

V. CONCLUSION

The laboratory implementation of optimal control of molecular processes had, and probably still has, to overcome a number of problems posed by the systematic and random errors present in the laboratory. The idea of AFOC answered many of them, most notably the problem of systematic errors and aging. In [52] it was shown that the problem caused by fast random errors can be overcome by using GA. This is explained by the "integrative" nature of the GA: instead of making sharp decisions it gives preferences to certain regions of the search space. The long decision process means that the error is practically averaged out.

TABLE III. The optimal solutions for the time delay discretized optimization.

$z\beta^{(1)}$ [fs]	A $[E_h/(ea_0)]$	$\frac{1}{2}z\beta^{(2)}$ $[10^3\,(\hbar/E_h)^2]$	$\frac{1}{6}z\beta^{(3)}$ $[10^6\,(\hbar/E_h)^3]$	ω_0 $[E_h/\hbar]$	objective
333	0.0304	241	4.1	0.0733	0.82
334	0.0291	241	5.7	0.0733	0.82
342	0.0316	469	−15.1	0.0732	0.90
346	0.0320	469	11.6	0.0735	0.90
347	0.0320	469	10.0	0.0735	0.91
353	0.0316	398	7.3	0.0723	0.90
358	0.0310	398	7.3	0.0731	0.91
362	0.0313	402	26.1	0.0730	0.91
363	0.0313	402	26.1	0.0730	0.91
367	0.0310	402	28.0	0.0712	0.91
374	0.0297	488	−8.8	0.0699	0.89
375	0.0305	488	−8.8	0.0699	0.90
381	0.0321	437	−8.8	0.0686	0.91
386	0.0320	441	−2.9	0.0691	0.92
386	0.0305	382	3.7	0.0707	0.91
393	0.0305	445	28.8	0.0699	0.91
395	0.0305	347	15.1	0.0693	0.91
398	0.0305	367	5.7	0.0698	0.91
406	0.0265	292	8.8	0.0668	0.88
409	0.0297	292	8.8	0.0668	0.90

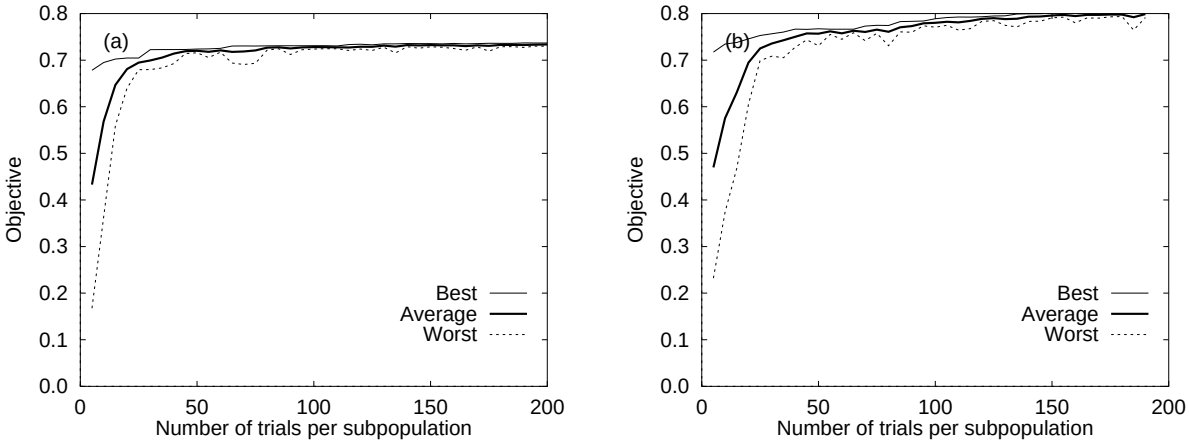

FIG. 4. The performance of the optimization with the amplitude taken as drifting, without (a) and with (b) migration. The figures show the best, average and worst performance obtained in each generation averaged along the subpopulations. The parameters being optimized were the time delay, the second and third order chirp and the central frequency.

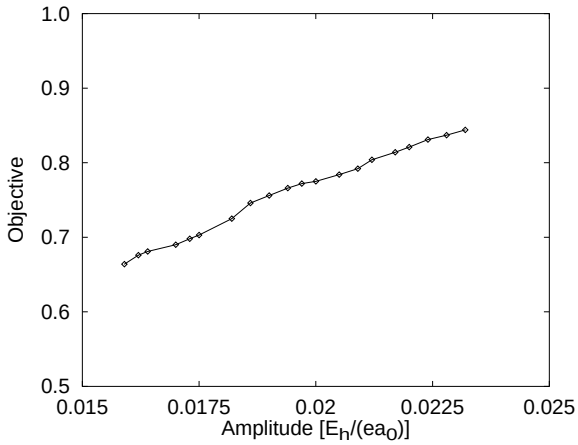

FIG. 5. The objective of the best solution as a function of the amplitude. It is well pronounced that inversion favors the higher amplitudes. The figure shows how the discretization points are placed: the Kohonen algorithm has indeed found a reasonably good distribution of the reference points.

In this paper we went a step further by examining how longer time scale random errors can be overcome. The paper introduced parametric optimization as a possible solution. Instead of trying to set the environmental parameters with high precision we suggested to measure them and use a matching optimal control. To find the optimal control settings for the different values of the environmental parameters parametric optimization is needed.

The GAM is a parametric optimization method. We have shown that it can be successfully used for molecular control applications. The demonstration used a pump-and-dump type model experiment on CsI. The results show the feasibility of such parametric optimization and demonstrate well the speed of GAM.

This paper suggests that the noise and medium speed drifting parameters inevitably present in the laboratory can be overcome by the GAM procedure. At the same time both components of GAM, the Kohonen network and the GA subsystems retain adaptivity and thus overcome long time scale changes, i.e., the aging of the laboratory components.

ACKNOWLEDGMENT

The authors acknowledge the support of The Danish Natural Research Science Council, the OTKA Grant of the Hungarian Academy of Sciences (OTKA 1890/1991), the US-Hungarian Joint Fund Grant (#168a) and the Office of Naval Research.

APPENDIX

In Sec. II we gave a brief description of the GAM. In the following we describe the algorithm in more detail.

As already described in Sec. II, the GA maintains a "population": a number of so called individuals, i.e., possible solutions of the control problem. The individuals are tested by evaluating the control coded by them, and using this knowledge the population is updated with the help of the so called genetic operators.

In most cases the GA works on bit strings as they facilitate a straightforward expression of the genetic operators. The bit strings have to be mapped onto the problem domain one is using. In function optimization, or in optimal control, where the goal is to find the best values of some parameters, it is usually done by cutting the string into as many parts as the number of variables, and then interpreting each part as a fixed point number. For example, if we have two variables, both in the range [0; 1], then the bit string "0010101101001010" would be interpreted as (0.168, 0.290).

When forming a new generation a mating pool is filled with individuals from the previous generation which are manipulated by the genetic operators; the individuals thus formed constitute the members of the new generation. Operator selection is used first to fill the pool. It serves to give individuals, and through them regions of the search space, of better quality more weight in the search. It is implemented by choosing individuals into the pool in proportion to their fitness. Then cross over is applied to combine partial solutions contained in different individuals to produce fitter individuals than either of the "parents". It can be implemented by cutting the pair of individuals at the same position and then interchanging the half strings (this is called single-point cross over). Generally cross over is not applied to all individuals: pairs of individuals undergo cross over with probability p_x. Mutation is used to maintain a level of diversity in the population by flipping each bit with probability p_{mut}.

Let us give an example how the genetic operators work. Suppose we want to cross over strings "0010101101001010" and "0101001001010010". First a cutting site is chosen: let us suppose that it is the fourth position. The result of the cross over is then a new pair of strings, namely "0010001001010010" and "0101101101001010". Mutating these strings we get, supposing that positions 2 and 11 in the first string and position 5 in the second string were chosen for flipping, is "0110001001110010" and "0101001101001010".

The fitness that the selection operator uses can either be the measured success of the control or some function thereof. In practice it is usual to use a linear function of the measured value. The parameters of the function are chosen so that the expected number of

offspring of the best individual is at most a given number (α), usually two or three. This so called linear scaling is applied in the present model.

The Kohonen network contains a number of units r (called neurons in this context), $r = 1,\ 2,\ \ldots,\ N_s$. Each unit stores a reference point $\boldsymbol{x}_r$ from the input space, i.e., the space of the environmental parameter(s). There is a metric ρ defined among the neurons, which is usually a d-dimensional Euclidean metric. The algorithm is most efficient if d is set to the number of uncorrelated environmental parameters, but it is also possible to set it to the total number of environmental parameters. The neurons are located on the lattice points of a d-dimensional lattice of unit lattice constant. The metric ρ is used to define a nearest neighbor relationship: two neurons are nearest neighbors if their distance is unity.

The network works as follows. At the beginning the reference points are chosen randomly. During optimization sample points from the input space are presented to the network. When an input $\boldsymbol{x}$ is presented to the network it activates the neuron whose reference point lies closest to the input; activation in our case means that the optimization procedure associated with that neuron will be active. At the same time the network adjusts itself: the reference point of *each* neuron moves closer to the presented point. The amplitude of the move made by neuron r is a decreasing function of its distance from the active neuron a:

$$\boldsymbol{x}_r \to \boldsymbol{x}_r + \epsilon G[\rho(r,a)](\boldsymbol{x} - \boldsymbol{x}_a), \tag{7}$$

where ϵ is the overall learning rate and G carries the distance dependence. In our case G is a Gaussian function:

$$G(x) = \exp\left[-\frac{x^2}{2\sigma^2}\right], \tag{8}$$

where σ characterizes the cooperativeness of the network, i.e., it determines the range of the neighbor training. Both ϵ and σ can depend on step number S (the number of input vectors presented):

$$\eta(S) = \begin{cases} \eta_i \left[\frac{\eta_f}{\eta_i}\right]^{S/S_d} & \text{if } S \le S_d \\ \eta_f & \text{otherwise,} \end{cases} \tag{9}$$

where η denotes either ϵ or σ, and S_d is the number of steps needed to train the network. Note that the network remains adaptive forever: the speed of adaptivity is decreased only to η_f.

The above learning rule assures that the discretization is optimal and topology conserving. Optimality means that the distribution of the reference points follows the distribution of the samples presented to the network. Topology conserving means that neurons close to each other with metric ρ will have reference points that are also close to each other. This is important to make the optimization cooperative. Without conserving the topology it would be difficult to know which optimization procedures correspond to similar values of the environmental parameters.

We can now proceed with the description of the GAM. As explained in Sec. II it contains a Kohonen network performing the discretization of the input space and each neuron of the network is associated with a GA procedure. The migration operator, which expresses the cooperativeness in GAM, is used when all of the GA procedures have filled their mating pools. It is applied to each neighboring pair of subpopulations, exchanging each individual in one pool with its corresponding pair in the other one with probability p_{mig}. After processing each pair of subpopulations the reproduction cycle continues in the usual way.

In application to molecular control this form of migration can be used if the changes of the environmental parameters are of moderate time scale. If their value can change from one experiment to the other then there is no guarantee that the individual tested is taken from the correct subpopulation. This last problem can be solved by measuring the values of the environmental parameters during the experiment and if they correspond to a different GA procedure then discarding the result of the testing. However, the same measurement can be used as the driving source of a new form of migration. This migration will be activated when the measurement before and during the experiment correspond to a different subpopulation. The individual is then exchanged with an individual from the subpopulation corresponding to the measurement during the experiment. This ensures that the result of the experiment is not wasted and that the individual will get an appropriate fitness value.

There is a correspondence between the two forms of migration. The latter form, supposing only one environmental parameter with a Gaussian noise of σ^2 variance, corresponds to the former with

$$p_{mig} = 1 - \mathrm{Erf}\left(\frac{\Delta}{\sqrt{2}\sigma}\right) \approx \sqrt{\frac{2}{\pi}}\frac{\sigma}{\Delta}\exp\left(-\frac{\Delta^2}{2\sigma^2}\right), \tag{10}$$

where 2Δ is the size of the discretization regions, and the approximation assumes that σ/Δ (i.e., p_{mig}) is small. If the migration level caused by the noise is insufficient then the two forms of migration can be combined; if it is too high, then some of the experiments have to be discarded.

REFERENCES

[1] Tannor, D. J.; Rice, S. A. *J. Chem. Phys.* **1985**, *83*, 5013.

[2] Tannor, D. J.; Kosloff, R.; Rice, S. A. *J. Chem. Phys.* **1986**, *85*, 5805.

[3] Shapiro, M.; Brumer, P. *J. Chem. Phys.* **1986**, *84*, 4103.

[4] Asaro, C.; Brumer, P.; Shapiro, M. *Phys. Rev. Lett.* **1988**, *60*, 1634.

[5] Peirce, A. P.; Dahleh, M. A.; Rabitz, H. *Phys. Rev. A* **1988**, *37*, 4950.

[6] Shi, S.; Rabitz, H. *Chem. Phys.* **1989**, *139*, 185.

[7] Kosloff, R.; Rice, S. A.; Gaspard, P.; Tersigni, S.; Tannor, D. J. *Chem. Phys.* **1989**, *139*, 201.

[8] Dahleh, M. A.; Peirce, A. P.; Rabitz, H. *Phys. Rev. A* **1990**, *42*, 1065.

[9] Schwieters, C. D.; Beumee, J. G. B.; Rabitz, H. *J. Opt. Soc. Am. B* **1990**, *7*, 1736.

[10] Shi, S.; Rabitz, H. *J. Chem. Phys.* **1990**, *92*, 364.

[11] Shi, S.; Rabitz, H. *J. Chem. Phys.* **1990**, *92*, 2927.

[12] Tersigni, S. H.; Gaspard, P.; Rice, S. A. *J. Chem. Phys.* **1990**, *93*, 1670.

[13] Judson, R. S.; Lehmann, K. K.; Rabitz, H.; Warren, W. S. *J. Mol. Struct.* **1990**, *223*, 425.

[14] Paramonov, G. K. *Chem. Phys. Lett.* **1990**, *169*, 573.

[15] Amstrup, B.; Carlson, R.; Matro, A.; Rice, S. A. *J. Phys. Chem.* **1991**, *95*, 8019.

[16] Gross, P.; Neuhauser, D.; Rabitz, H. *J. Chem. Phys.* **1991**, *94*, 1158.

[17] Schwieters, C. D.; Rabitz, H. *Phys. Rev. A* **1991**, *44*, 5224.

[18] Tannor, D. J.; Jin, Y. Design of femtosecond pulse sequences to control photochemical products. In *Mode selective chemistry*, pp 333–345, Dordrecht, 1991. Kluwer.

[19] Beumee, J. G. B.; Rabitz, H. *J. Chem. Phys.* **1992**, *97*, 1353.

[20] Brumer, P.; Shapiro, M. *Ann. Rev. Phys. Chem.* **1992**, *43*, 257.

[21] Gross, P.; Neuhauser, D.; Rabitz, H. *J. Chem. Phys.* **1992**, *96*, 2834.

[22] Judson, R. S.; Rabitz, H. *Phys. Rev. Lett.* **1992**, *68*, 1500.

[23] Kosloff, R.; Hammerich, A. D.; Tannor, D. *Phys. Rev. Lett.* **1992**, *69*, 2172.

[24] Shapiro, M.; Brumer, P. *J. Chem. Phys.* **1992**, *97*, 6259.

[25] Just, B.; Manz, J.; Trisca, I. *Chem. Phys. Lett.* **1992**, *193*, 423.

[26] Combariza, J. E.; Görtler, S.; Just, B.; Manz, J. *Chem. Phys. Lett.* **1992**, *195*, 393.

[27] Amstrup, B.; Lőrincz, A.; Rice, S. A. *J. Phys. Chem.* **1993**, *97*, 6175.

[28] Amstrup, B.; Doll, J. D.; Sauerbrey, R. A.; Szabó, G.; Lőrincz, A. *Phys. Rev. A* **1993**, *48*, 3830.

[29] Chelkowski, S.; Bandrauk, A. D. *J. Chem. Phys.* **1993**, *99*, 4279.

[30] Goswami, D.; Warren, W. S. *J. Chem. Phys.* **1993**, *99*, 4509.

[31] Gross, P.; Neuhauser, D.; Rabitz, H. *J. Chem. Phys.* **1993**, *98*, 4557.

[32] Gross, P.; Singh, H.; Rabitz, H.; Mease, K.; Huang, G. M. *Phys. Rev. A* **1993**, *47*, 4593.

[33] Jakubetz, W.; Kades, E.; Manz, J. *J. Phys. Chem.* **1993**, *97*, 12609.

[34] Kohler, B.; Krause, J. L.; Raksi, F.; Rose-Petruck, C.; Whitnell, R. M.; Wilson, K. R.; Yakovlev, V. V.; Yan, Y.; Mukamel, S. *J. Phys. Chem.* **1993**, *97*, 12602.

[35] Krause, J. L.; Whitnell, R. M.; Wilson, K. R.; Yan, Y.; Mukamel, S. *J. Chem. Phys.* **1993**, *99*, 6562.

[36] Manz, J.; Paramonov, G. K. *J. Phys. Chem.* **1993**, *97*, 12625.

[37] Shapiro, M.; Brumer, P. *Chem. Phys. Lett.* **1993**, *208*, 193.

[38] Shen, L.; Shi, S.; Rabitz, H. *J. Phys. Chem.* **1993**, *97*, 12114.

[39] Shen, L.; Shi, S.; Rabitz, H.; Lin, C.; Littman, M.; Heritage, J. P.; Weiner, A. M. *J. Chem. Phys.* **1993**, *98*, 7792.

[40] Somlói, J.; Kazakov, V. A.; Tannor, D. J. *Chem. Phys.* **1993**, *172*, 85.

[41] Szakács, T.; Lucza, T.; Lőrincz, A. *Surface Science* **1993**, *296*, 251.

[42] Szakács, T.; Somlói, J.; Lőrincz, A. *Chem. Phys.* **1993**, *172*, 1.

[43] Yan, Y.; Gillilan, R. E.; Whitnell, R. M.; Wilson, K. R.; Mukamel, S. *J. Phys. Chem.* **1993**, *97*, 2320.

[44] Amstrup, B.; Rice, S. A. *Chem. Phys. Lett.* **1994**, *225*, 1.

[45] Amstrup, B.; Szabó, G.; Sauerbrey, R. A.; Lőrincz, A. *Chem. Phys.* **1994**, *188*, 87.

[46] Gross, P.; Bairagi, D. B.; Mishra, M. K.; Rabitz, H. *Chem. Phys. Lett.* **1994**, *223*, 263.

[47] Kaluža, M.; Muckerman, J. T.; Gross, P.; Rabitz, H. *J. Chem. Phys.* **1994**, *100*, 4211.

[48] Michaels, A.; Schwieters, C.; Rabitz, H. *J. Phys. Chem.* **1994**, *98*, 2508.

[49] Shen, H.; Dussault, J.-P.; Bandrauk, A. D. *Chem. Phys. Lett.* **1994**, *221*, 498.

[50] Shen, L.; Rabitz, H. *J. Chem. Phys.* **1994**, *100*, 4811.

[51] Szakács, T.; Amstrup, B.; Gross, P.; Kosloff, R.; Rabitz, H.; Lőrincz, A. *Phys. Rev. A* **1994**, *50*, 2540.

[52] J.Tóth, G.; Lőrincz, A.; Rabitz, H. *J. Chem. Phys.* **1994**, *101*, 3715.

[53] Zhang, H.; Rabitz, H. *Phys. Rev. A* **1994**, *49*, 2241.

[54] Potter, E. D.; Herek, J. L.; Pedersen, S.; Liu, Q.; Zewail, A. H. *Nature* **1992**, *355*, 66.

[55] Herek, J. L.; Materny, A.; Zewail, A. H. *Chem. Phys. Lett.* **1994**, *228*, 15.

[56] Demiralp, M.; Rabitz, H. *Phys. Rev. A* **1993**, *47*, 831.

[57] J.Tóth, G.; Lőrincz, A. Genetic Algorithm with migration on topology conserving maps. In Gielen, S.; Kappen, B., Ed., *Proceedings of the International Conference on Artificial Neural Networks*, pp 605–608, New York, 1993. Springer-Verlag.

[58] J.Tóth, G.; Lőrincz, A. *Neural Networks World, submitted* **1994**.

[59] Kohonen, T. *Self-organization and associative memory*; Springer-Verlag: New York, 1984.

[60] Hertz, J.; Krogh, A.; Palmer, R. G. *Introduction to the theory of neural computation*; Addison-Wesley: Redwood,CA, 1991.

[61] Goldberg, D. E. *Genetic Algorithms in search, optimization and machine learning*; Addison Wesley: Reading, MA, 1989.

[62] Beasley, D.; Bull, D. R.; Martin, R. R. *University Computing* **1993**, *15*, 58. (A postscript version of this manuscript is available from ENCORE sites, e.g. by anonymous ftp from ftp.Germany.EU.net in the directory /pub/research/softcomp/EC/GA/papers/over93.ps.gz).

[63] Beasley, D.; Bull, D. R.; Martin, R. R. *University Computing* **1993**, *15*, 170. (A postscript version of this manuscript is available from ENCORE sites, e.g. by anonymous ftp from ftp.Germany.EU.net in the directory /pub/research/softcomp/EC/GA/papers/over93-2.ps.gz).

[64] Janszky, J.; Kobayashi, T.; Vinogradov, A. *Opt. Comm.* **1990**, *76*, 30.

[65] Ruhman, S.; Kosloff, R. *J. Opt. Soc. Am. B* **1990**, *7*, 1748.

[66] Gogolák, Z.; Bor, Zs.; Szabó, G. *Exp. Tech. Phys.* **1991**, *39*, 323.

[67] Weiner, A. M.; Leaird, D. E.; Patel, J. S.; Wullert, II, J. R. *IEEE J. Quant. Elect.* **1992**, *28*, 908.

[68] Leforestier, C.; Bisseling, R. H.; Cerjan, C.; Feit, M. D.; Friesner, R.; Guldberg, A.; Hammerich, A.; Jolicard, G.; Karrlein, W.; Meyer, H.-D.; Lipkin, N.; Roncero, O.; Kosloff, R. *J. Comput. Phys.* **1991**, *94*, 59.

[69] Feit, M. D.; Fleck, Jr., J. A.; Steiger, A. *J. Comput. Phys.* **1982**, *47*, 412.

[70] Ruhman, S.; Kosloff, R. *J. Opt. Soc. Am. B* **1990**, *7*, 1748.

[71] Kosloff, R.; Tal-Ezer, H. *Chem. Phys. Lett.* **1986**, *127*, 223.

[72] Aubanel, E. E.; Bandrauk, A. D. *J. Phys. Chem.* **1993**, *97*, 12620.