



HAL
open science

Hamiltonian identification in presence of large control field perturbations

Ying Fu, Herschel Rabitz, Gabriel Turinici

► **To cite this version:**

Ying Fu, Herschel Rabitz, Gabriel Turinici. Hamiltonian identification in presence of large control field perturbations. 2016. hal-01326060v1

HAL Id: hal-01326060

<https://hal.science/hal-01326060v1>

Preprint submitted on 3 Jun 2016 (v1), last revised 15 Oct 2016 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hamiltonian identification in presence of large control field perturbations

Ying FU

*CEREMADE, Université Paris Dauphine,
Place du Marechal de Lattre de Tassigny, PARIS 75016,
FRANCE; e-mail: fu@ceremade.dauphine.fr*

Herschel Rabitz

*Department of Chemistry, Princeton University,
Princeton, NJ 08544, USA; e-mail: hrabitz@princeton.edu*

Gabriel TURINICI

*CEREMADE, Université Paris Dauphine,
Place du Marechal de Lattre de Tassigny, PARIS 75016,
FRANCE; e-mail: Gabriel.Turinici@dauphine.fr and
Institut Universitaire de France*

(Dated: June 2, 2016)

Abstract

Quantum system inversion concerns learning the characteristics of the underlying Hamiltonian by measuring suitable observables from the responses of the system's interaction with members of a set of applied fields. Various aspects of inversion have been confirmed in theoretical, numerical and experimental works. Nevertheless, the presence of noise arising from the applied fields may contaminate the quality of the results. In this circumstance, the observables satisfy probability distributions, but often the noise statistics are unknown. Based on a proposed theoretical framework, we present a procedure to recover both the unknown parts of the Hamiltonian and the unknown noise distribution. The procedure is implemented numerically and seen to perform well for illustrative Gaussian, exponential and bi-modal noise distributions.

Keywords: quantum control, quantum identification

I. INTRODUCTION

The interaction between a quantum system and an applied field can be used to gather information about the system by measuring suitable observables for various incarnations of the field (see [1–3, 7, 8, 10] for some related works). Such an inversion endeavor, seeking information about the system, has to take into account the possibility that the measurements are contaminated by noise, possibly from multiple sources, or that the model does not adequately describe all relevant characteristics of the system. Many considerations arise, including the following list:

- S-1** The quantum system can have an orientation that is unknown resulting in an ensemble of systems which on the average is homogeneous. These circumstances will produce a multiplicative statistical factor mediating the field strength in the Hamiltonian.
- S-2** The number of levels is unknown. Except for spins, the number of levels can generally take on any value, including the prospect of there being a continuum. The nature of the field bandwidth and intensity plays into which aspects of the system dynamics are accessible due to the field.
- S-3** Intense fields can be important in some problems where excited electronic states play a role, which can bring in non-resonant processes (i.e., virtual states, requiring adequate models). Moreover intense fields may lead to models including the nonlinear field coupling coefficients in the Hamiltonian.

The performance of inversion needs to acknowledge the issues above. In addition, there will always be noise in the observations, but assuming that it is random and symmetric (e.g. Gaussian) then signal averaging may reduce its influence. However, noise may accompany the control field with the following circumstances:

- F-1** Different molecules (even if they are oriented) will see different fields as the field generally has a spatial (e.g., a Gaussian) lateral spread.
- F-2** A so-called spatio-temporal "chirp" may occur, where there is some unknown complex spatio-temporal pattern (possibly fluctuating) in the field. When the pattern is not uniform it has to be factored into the model with, possibly nonlinear, additional terms.

F-3 Although the field may be known before entering the sample, optical distortion during propagation can lead to additional field uncertainty in the domain where the data is taken.

F-4 The variations in field are stochastic from shot-to-shot with respect to the frequency dependence of the amplitudes and phases. The origin of this behavior is mainly due to the laser source and possibly jitter in the optics. The implications for the control data due to this occurrence is likely more significant at higher intensities.

Collecting all of these items pose a daunting, but not an impossible challenge to treat. In this work we limit the focus on the randomness coming from the last item **F-4** above. We present the theoretical background and analysis in Section II. The numerical algorithm and simulation results are given in Section III. Final remarks are made in Section IV.

II. THEORETICAL FRAMEWORK

A. The model

We introduce the following notation:

- $\mathbb{L}_{M_1, M_2, \dots, M_m}$ is the Lie algebra spanned by the matrices M_1, M_2, \dots, M_m ;
- For any matrix or vector X we denote by X^* its adjoint (the transpose, complex conjugate);
- \mathcal{H}_N is the set of all Hermitian matrices $\mathcal{H}_N = \{X \in \mathbb{C}^{N \times N} | X^* = X\}$;
- $\Psi(t, H, u(\cdot), \mu, \Psi_0)$ is the solution of the equation (1) below; to simplify the notation, when there is no ambiguity, we denote it as $\Psi(t)$;
- $SU(N)$ is the special unitary group of degree N , which is the group of $N \times N$ unitary matrices with determinant 1;
- $\mathfrak{su}(N)$ is the Lie Algebra of skew-Hermitian matrices (the Lie algebra of $SU(N)$);
- A set of commuting observables $\mathcal{O} = \{O_1, \dots, O_K\}$ (named SCO hereafter) satisfies $[O_k, O_\ell] = 0, \forall k, \ell \in \{1, \dots, K\}$;

- $\mathcal{L}(X)$ is the distribution of the random variable X .

Let us consider the following controlled quantum system with time-dependent wavefunction $\Psi(t)$ satisfying the Schrödinger equation:

$$\begin{cases} i\dot{\Psi}(t, H, u(\cdot), \mu, \Psi_0) = (H - u(t)\mu)\Psi(t, H, u(\cdot), \mu, \Psi_0) \\ \Psi(0, H, u(\cdot), \mu, \Psi_0) = \Psi_0, \end{cases} \quad (1)$$

where H is the internal ("free") Hamiltonian and μ the coupling operator between the control $u(t) \in L^1_{loc}(\mathbb{R}_+; \mathbb{R})$ and the system. We work in a finite dimensional framework, therefore $H, \mu \in \mathcal{H}_N$ for some $N \in \mathbb{N}^*$. The free Hamiltonian H is taken as known and the goal is to recover the matrix entries of μ from laboratory measurements of some observables depending on $\Psi(t)$. The control $u(t)$ can be changed in the laboratory to gather sufficient information on the system for the inversion to extract μ .

In the laboratory, the control field is produced by superposition of several frequencies, each with a specific amplitude and phase:

$$u(t) = \sigma(t) \sum_{\alpha \neq \beta} A_{\alpha\beta} \sin(\omega_{\alpha\beta}t + \theta_{\alpha\beta}), \quad (2)$$

where $\sigma(t)$ is a Gaussian envelope in time and $\omega_{\alpha\beta} = E_\beta - E_\alpha$ is the transition frequency between the eigenvalues E_α and E_β of H . The amplitudes $A_{\alpha\beta}$ and the phases $\theta_{\alpha\beta}$ are the control parameters. In practice, the control fields cannot be produced perfectly. When we repeat nominally the same experiment, there can be a random shift in the amplitudes, which means for each experiment, there is a multiplicative noise factor on the amplitudes. Accordingly the perturbation is modeled by a random multiplicative factor Y acting on the control i.e., $u(t)$ is replaced by $Y \cdot u(t)$. This model has the assumption that each $A_{\alpha\beta}$ has the same shift; in addition we do not treat here possible noise appearing in the phases $\theta_{\alpha\beta}$. In principle, the same methodology may apply to noise appearing in the phases; however at this time adapted ensemble controllability results are not available and this requires working with a high-dimensional probability distribution (one dimension for each independent noise in a phase). As a final remark, noise in the amplitudes and phases does not imply that the origin is in the pulse shaper but from the driving laser and represented in the fashion above.

The perturbation Y takes values in $\mathcal{V} = \{y_\ell, \ell \leq L\} \subset \mathbb{R}$. We denote $\xi_\ell = \mathbb{P}(Y = y_\ell)$,

where the column vector ξ with entries ξ_1, \dots, ξ_L is a probability distribution on \mathcal{V} . The numerical values of the possible perturbations y_ℓ are known but their occurrence probabilities ξ_ℓ are unknown and thus Ψ is a random variable, as are all measurements depending on Ψ . Repeating the control experiment several times will yield the distribution of the measurements.

The measurements are of the form $\langle O\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0) \rangle$ with $O \in \mathcal{H}_N$ being a member of a list of possible measurable operators. Often, only one observable operator is readily available, but the experiment can be repeated many times, including with distinct chosen fields. Generally for two observables $O_1, O_2 \in \mathcal{H}_N$ no information is available on the joint distribution of the values $\langle O_1\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0) \rangle$ and $\langle O_2\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0) \rangle$.

B. Theoretical result

The following theorem proves that under certain assumptions on the system and the SCO \mathcal{O} , if we obtain the same distributions for all observables in \mathcal{O} and for all controls, then the dipole moment μ and the probabilities $(\xi_\ell)_{\ell=1}^L$ can be identified up to some multiplicative phases.

Remark II.1. *As the noise (with unknown distribution) Y multiplies the unknown dipole μ , when the couple (Y, μ) is a solution, any couple $(Y/\lambda, \lambda\mu)$ is a solution too; thus it is only possible to obtain Y and μ up to a multiplicative factor (here λ). If, for instance, some additional information on Y or μ is known the constant λ can be set accordingly. As such, we will suppose from now on that for at least one transition an absorption intensity measurement can be performed which provides the value of $|\mu_{k,\ell}|^2$ for some given k, ℓ (see [4, Chapter XIII, Section C.3.b Fermi's Golden Rule, p. 1299]). This is recalled in assumption **Hyp-D** below.*

Theorem II.2. *Let $H, \mu_1, \mu_2 \in \mathcal{H}_N$, H diagonal, $\mu_1 \neq 0, \mu_2 \neq 0$, Y_1, Y_2 two random variables with values in the same set \mathcal{V} , at least one of which is non-null, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$ some initial states and denote for $a = 1, 2$ and $u \in L_{loc}^1(\mathbb{R}_+, \mathbb{R})$: $\Psi_a(t, u) = \Psi(t, H, u(\cdot), \mu_a, \Psi_0^a)$. Let \mathcal{O} be a (non-trivial) SCO. We suppose that $N \geq 3$ and:*

Hyp-A $\mathbb{L}_{iH, i\mu_1} = \mathbb{L}_{iH, i\mu_2} = \mathfrak{su}(N)$;

Hyp-B $\text{tr}(H) = \text{tr}(\mu_1) = \text{tr}(\mu_2) = 0$;

Hyp-C the eigenvalues of H are all of multiplicity one.

Hyp-D $|(\mu_1)_{k,\ell}|^2 = |(\mu_2)_{k,\ell}|^2 \neq 0$ for some fixed k, ℓ .

The final observation time is denoted T (assumed large enough) and we suppose the following equality of distributions:

$$\mathcal{L}(\langle O\Psi_1(T, uY_1), \Psi_1(T, uY_1) \rangle) = \mathcal{L}(\langle O\Psi_2(T, uY_2), \Psi_2(T, uY_2) \rangle) \quad \forall u \in L^1([0, T]; \mathbb{R}), \quad \forall O \in \mathcal{O}, \quad (3)$$

then for some $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$:

$$\begin{cases} (\mu_1)_{jk} = \pm e^{i(\alpha_j - \alpha_k)} (\mu_2)_{jk}, \quad \forall j, k \leq N, \\ \mathbb{P}(Y_1 = y_\ell) = \mathbb{P}(Y_2 = \pm y_\ell) \quad \forall \ell \leq L. \end{cases} \quad (4)$$

Remark II.3. The \pm signs in equation (4) are due to the hypothesis **Hyp-D** which leaves the sign undetermined; it may also appear when the sign of the noise is ambiguous i.e., when Y_1 and $-Y_1$ have the same distribution.

Remark II.4. The SCO \mathcal{O} may contain just one observable.

Proof. The proof requires the tools introduced in [6] where the additive noise $u(\cdot) + Y$ was considered. In order to keep it simple we only give the main ideas and the modifications with respect to proof presented there. Using Lemma 5.1 in [6], for all $\ell \leq L$ and $O \in \mathcal{O}$, there exists $\kappa_O(\ell) \leq L$ such that

$$\langle O\Psi_1(T, u \cdot y_\ell), \Psi_1(T, u \cdot y_\ell) \rangle = \langle O\Psi_2(T, u \cdot y_{\kappa_O(\ell)}), \Psi_2(T, u \cdot y_{\kappa_O(\ell)}) \rangle. \quad (5)$$

Reasoning as in the proof of theorem 4.1 of [6], this implies that there exists $W \in SU(N)$ diagonal such that $y_{\kappa_O(\ell)} \mu_2 = W y_\ell \mu_1 W^{-1}$. Since at least one of Y_1 or Y_2 is non-null, we can suppose, without loss of generality, that $y_\ell \neq 0$; we deduce the existence of some $\lambda \in \mathbb{R} \setminus \{0\}$ such that

$$\begin{cases} (\mu_1)_{jk} = \lambda e^{i(\alpha_j - \alpha_k)} (\mu_2)_{jk}, \quad \forall j, k \leq N, \\ \mathbb{P}(Y_1 = y_\ell) = \mathbb{P}(Y_2 = \lambda y_\ell) \quad \forall \ell \leq L. \end{cases} \quad (6)$$

However, since μ_1 and μ_2 are fixed for all O and ℓ we obtain that λ is independent of ℓ , which, using assumption **Hyp-D**, gives the conclusion. \square

Remark II.5. *The hypothesis **Hyp-A** is required for identification while **Hyp-B** is rather a convention. Hypothesis **Hyp-C** can be relaxed (as in [6]) if the \mathcal{O} is a Complete Set of Commuting Observables (CSCO).*

III. NUMERICAL RESULTS

A. The algorithm

Several numerical simulations were performed in order to illustrate the theoretical result in Section II. In all cases we simulate the "real" system with Hamiltonian H , dipole moment μ_{real} , noise Y^{real} with distribution $\mathcal{L}(Y^{real}) = \sum_{\ell=1}^L \xi_{\ell}^{real} \delta_{y_{\ell}}$ and observables $\mathcal{O} = \{O_1, \dots, O_K\}$ that correspond to the Hamiltonian H , specified by projections $\{|e_1\rangle\langle e_1|, \dots, |e_N\rangle\langle e_N|\}$. Here $|e_j\rangle$ is the j -th eigenstate of H . Note that here we take $\mathcal{O} = \{O_1, \dots, O_K\}$ to be the entire set of projectors, but one would be enough for the theoretical result to hold, see the Remark II.4.

The measurements provide the distributions of the observables for each control u :

$$\sum_{k=1}^L \xi_k^{real} \delta_{|\langle \Psi(T, H, u, y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2}. \quad (7)$$

The Hamiltonian H is assumed known to high accuracy. In contrast, we suppose that *a priori* we only know the order of magnitude of the dipole moment. This information is useful as an initial guess for the inversion procedure.

To find the dipole moment and the noise distribution of the control amplitudes we minimize the difference between

- (i) the observed distribution calculated, at some suitably large time T , with the current dipole candidate μ and the current noise distribution candidate $\sum_{k=1}^L \xi_k \delta_{y_k}$ and
- (ii) the real observable distribution calculated with the real dipole moment μ_{real} and the real distribution $\sum_{k=1}^L \xi_k^{real} \delta_{y_k}$.

The difference is summed over several controls u_1, \dots, u_{N_u} and defined as:

$$\mathcal{J}(\mu, (\xi_k)_{k=1}^L; (u_i)_{i=1}^{N_u}) = \log \left\{ \frac{1}{N_u} \sum_{i=1}^{N_u} \sum_{j=1}^N \mathcal{W}_1 \left[\sum_{k=1}^L \xi_k \delta_{|\langle \Psi(T, H, u_i \cdot y_k, \mu, \Psi_1^0), e_j \rangle|^2}, \sum_{k=1}^L \xi_k^{real} \delta_{|\langle \Psi(T, H, u_i \cdot y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2} \right] \right\}. \quad (8)$$

Here \mathcal{W}_1 represents the 1-Wasserstein (also known as Kantorovich-Rubinstein) distance between two distributions (see page 34-35 in [9]); for two probability distributions Z_1, Z_2 having cumulative distribution functions F_{Z_1} (respectively F_{Z_2}) the distance is:

$$\mathcal{W}_1(Z_1, Z_2) = \int_0^1 |F_{Z_1}^{-1}(x) - F_{Z_2}^{-1}(x)| dx. \quad (9)$$

Other distances could also be used, e.g. \mathcal{W}_2 . We start the optimization with an initial guess μ^0 (see (14)); the distribution ξ^0 is initialized to be uniform. The iteration $n \geq 1$ consists in the following steps:

Algo 1 Randomly choose N_u controls $u_i^n, i = 1, \dots, N_u$;

Algo 2 minimize $\xi \rightarrow \mathcal{J}(\mu^{n-1}, \xi; (u_i)_{i=1}^{N_u})$ and set ξ^n to be a minimizer (in practice a close approximation);

Algo 3 minimize $\mu \rightarrow \mathcal{J}(\mu, \xi^n; (u_i)_{i=1}^{N_u})$ and set μ^n to be a minimizer (in practice a close approximation);

For the step **Algo 2**, denote $a_k^{ij} = |\langle \Psi(T, H, u_i^n \cdot y_k, \mu^n, \Psi_1^0), e_j \rangle|^2$ and $b_k^{ij} = |\langle \Psi(T, H, u_i^n \cdot y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2$. A part of the algorithm is to minimize the error:

$$\log \left[\frac{1}{N_u} \sum_{i=1}^{N_u} \sum_{j=1}^N \mathcal{W}_1 \left(\sum_{k=1}^L \xi_k \delta_{a_k^{ij}}, \sum_{k=1}^L \xi_k^{real} \delta_{b_k^{ij}} \right) \right]. \quad (10)$$

The differential of the distance \mathcal{W}_1 is not trivial to compute (see [9] for a rigorous mathematical treatment); instead, in this step, we use the L^2 distance between the smoothed densities: each Dirac mass is replaced by a Gaussian distribution with small variance (here $\nu^2 = 10^{-4}$): $\delta_{a_k^{ij}}$ is replaced by the normal distribution $\mathcal{N}(a_k^{ij}, \nu^2)$ and $\delta_{b_k^{ij}}$ is replaced by $\mathcal{N}(b_k^{ij}, \nu^2)$. The use of parameter ν does not assume any particular distribution for ξ^{real}

and does not bias towards one, it is only a rapid way to obtain a computable gradient. We compute the $L \times L$ matrix \mathcal{M}^{ij} whose entries $\mathcal{M}_{k,\ell}^{ij} = \frac{1}{\nu \cdot \sqrt{2\pi}} e^{-\frac{(a_k^{ij} - a_\ell^{ij})^2}{2 \cdot \nu^2}}$ account for the density of the distribution $\mathcal{N}(a_\ell^{ij}, \nu^2)$ at the point a_k^{ij} (and a similar matrix \mathcal{M}'^{ij} with $\mathcal{M}'_{k,\ell}{}^{ij} = \frac{1}{\nu \cdot \sqrt{2\pi}} e^{-\frac{(b_k^{ij} - b_\ell^{ij})^2}{2 \cdot \nu^2}}$ for the real distribution). When ν is small the minimizer of the term in (10) is close to the minimizer of $\sum_{i=1}^{N_u} \sum_{j=1}^N \|\mathcal{M}^{ij} \xi - \mathcal{M}'^{ij} \xi^{real}\|^2$ which is given by the formula:

$$\xi^{n,raw} = \left(\sum_{i=1}^{N_u} \sum_{j=1}^N (\mathcal{M}^{ij})^T \mathcal{M}^{ij} \right)^{-1} \left(\sum_{i=1}^{N_u} \sum_{j=1}^N (\mathcal{M}'^{ij})^T \mathcal{M}'^{ij} \right) \xi^{real}. \quad (11)$$

The term $\xi^{n,raw}$ is corrected to be a probability distribution and set to ξ^n :

$$\xi^n = \frac{1}{\sum_{\ell=1}^L |\xi_\ell^{n,raw}|} (|\xi_1^{n,raw}|, \dots, |\xi_L^{n,raw}|). \quad (12)$$

For the step **Algo 3** there is no explicit solution; a classical unconstrained nonlinear optimization algorithm is employed (we used the Gnu Octave procedure "fminunc").

Once the algorithm finished, the remaining overall multiplicative constant λ (see Remark II.1)) is set consistent with the data (see below the examples).

B. Numerical tests: N observables

We consider the 4-level system ($N = 4$) in [5] having:

$$H = \begin{pmatrix} 0.0833 & -0.0038 & -0.0087 & 0.0041 \\ -0.0038 & 0.0647 & 0.0083 & 0.0038 \\ -0.0087 & 0.0083 & 0.0036 & -0.0076 \\ 0.0041 & 0.0038 & -0.0076 & 0.0357 \end{pmatrix}, \quad \mu_{real} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 7 \\ 0 & -1.5 & 7 & 0 \end{pmatrix}. \quad (13)$$

We set $N_u = 36$; the controls are defined by formula (2) with amplitudes ($A_{\alpha\beta}$) chosen at random, uniformly in $[0, 0.1 \cdot \frac{\|H\|_{l^\infty}}{\|\mu_{real}\|_{l^\infty}} = 0.0012]$ and the phases $\theta_{\alpha\beta}$ in $[0, 2\pi]$. For the

values (13) the initial guesses are:

$$\mu_g^0 = \begin{pmatrix} 0 & 3.76 & -1.31 & 0 \\ 3.76 & 0 & 3.51 & -1.78 \\ -1.31 & 3.51 & 0 & 6.72 \\ 0 & -1.78 & 6.72 & 0 \end{pmatrix}, \mu_e^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \mu_b^0 = \begin{pmatrix} 0 & 7.48 & -0.51 & 0 \\ 7.48 & 0 & 8.83 & -0.87 \\ -0.51 & 8.83 & 0 & 5.87 \\ 0 & -0.87 & 5.87 & 0 \end{pmatrix} \quad (14)$$

The average relative errors of these initial guesses are 42%, 70% and 50%. Guesses μ_g^0 and μ_b^0 are obtained by multiplying element-wise μ_{real} by uniform random variables in $[0.4, 1.6]$ (i.e., up to 60% average relative error); the term μ_e^0 is obtained by taking roughly the order of magnitude of the entries of μ_{real} : the values smaller or about 1 are taken to be 1 while the others are set to 10. The observables arising from the Hamiltonian H are projections $\{|e_1\rangle\langle e_1|, |e_2\rangle\langle e_2|, |e_3\rangle\langle e_3|, |e_4\rangle\langle e_4|\}$ to eigenstates:

$$\begin{aligned} |e_1\rangle &= \begin{pmatrix} 0.0845 & -0.1313 & 0.9651 & 0.2101 \end{pmatrix}^T, |e_2\rangle = \begin{pmatrix} -0.1305 & -0.0856 & -0.2103 & 0.9651 \end{pmatrix}^T, \\ |e_3\rangle &= \begin{pmatrix} 0.2118 & 0.9647 & 0.0838 & 0.1325 \end{pmatrix}^T, |e_4\rangle = \begin{pmatrix} 0.9649 & -0.2118 & -0.1314 & 0.0830 \end{pmatrix}^T. \end{aligned} \quad (15)$$

In the basis $\{e_1, e_2, e_3, e_4\}$ the Hamiltonian H is diagonal with eigenvalues $E_1 = 0$, $E_2 = 0.0365$, $E_3 = 0.0651$, $E_4 = 0.0857$. We set the final time $T = 3200$ which is about 10 periods of the smallest transition frequency $2\pi/(E_4 - E_3) = 314$ in H .

With respect to Remark II.1 we suppose that the value $|(\mu_{real})_{12}|^2 = 25$ is known.

The support for the distribution Y is known and denoted $[y_m, y_M]$; in the numerical tests we take $y_m = 0.5$ and $y_M = 1.5$. We discretize the set of possible values of the perturbation with $L = 51$ equidistant points $y_\ell = y_m + (\ell - 1) \cdot \frac{y_M - y_m}{L-1}$, $\ell = 1, \dots, L$. The values y_ℓ are supposed known but not the probabilities ξ_ℓ^{real} that define the distribution $\mathcal{L}(Y^{real}) = \sum_{\ell=1}^L \xi_\ell^{real} \delta_{y_\ell}$ of the perturbation Y^{real} . Several distributions are tested; they are constructed by discretizing, truncating and re-normalizing several classical distributions:

- $Y^{real} = Y^g$ being a Gaussian distribution centered at 1 with variance equal to 0.0025:

$$\xi_\ell^{real,g} = \frac{f_\ell^g}{\sum_{k=1}^L f_k^g} \quad \ell = 1, \dots, L \quad \text{with} \quad f_k^g = \frac{1}{\sqrt{0.0025}\sqrt{2\pi}} e^{-\frac{(y_k-1)^2}{2 \cdot 0.0025}}, \quad (16)$$

- $Y^{real} = Y^e$ being a shifted exponential distribution form:

$$\xi_\ell^{real,e} = \frac{f_\ell^e}{\sum_{k=1}^L f_k^e} \quad \ell = 1, \dots, L \text{ with } f_k^e = 5 \cdot e^{-5 \cdot (y_k - y_m)}, \quad (17)$$

- $Y^{real} = Y^b$ being the bi-modal distribution which is the sum of two Gaussian distributions. We choose the first one centered at 0.8 with variance equal to 0.0025 and the second one centered at 1.2 with variance equal to 0.0049:

$$\xi_\ell^{real,b} = \frac{f_\ell^b}{\sum_{k=1}^L f_k^b} \quad \ell = 1, \dots, L \text{ with } f_k^b = \frac{1}{\sqrt{0.0025}\sqrt{2\pi}} e^{-\frac{(y_k - 0.8)^2}{2 \cdot 0.0025}} + \frac{1}{\sqrt{0.0049}\sqrt{2\pi}} e^{-\frac{(y_k - 1.2)^2}{2 \cdot 0.0049}}. \quad (18)$$

The dipole moments converge in 10 iterations. The numerical values are rescaled in order to use that $|(\mu_{real})_{12}|^2 = 25$; same is done for the noise distributions. We obtain μ_g^{10} , μ_e^{10} and μ_b^{10} :

$$\|\mu_g^{10} - \mu_{real}\|_\infty = 5 \cdot 10^{-5}, \quad \mu_g^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99995 & -1.5 \\ -1 & 5.99995 & 0 & 6.99999 \\ 0 & -1.5 & 6.99999 & 0 \end{pmatrix}, \quad (19)$$

$$\|\mu_e^{10} - \mu_{real}\|_\infty = 10^{-4}, \quad \mu_e^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 6.9999 \\ 0 & -1.5 & 6.9999 & 0 \end{pmatrix}, \quad (20)$$

$$\|\mu_b^{10} - \mu_{real}\|_\infty = 6 \cdot 10^{-5}, \quad \mu_b^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99999 & -1.5 \\ -1 & 5.99999 & 0 & 6.99994 \\ 0 & -1.5 & 6.99994 & 0 \end{pmatrix}. \quad (21)$$

The error norm is $\|\cdot\|_\infty$ is the largest, in absolute value, of the error components. See figures 1,2, 3 for the results. The tables I and II present the match of the probability distributions of the observables.

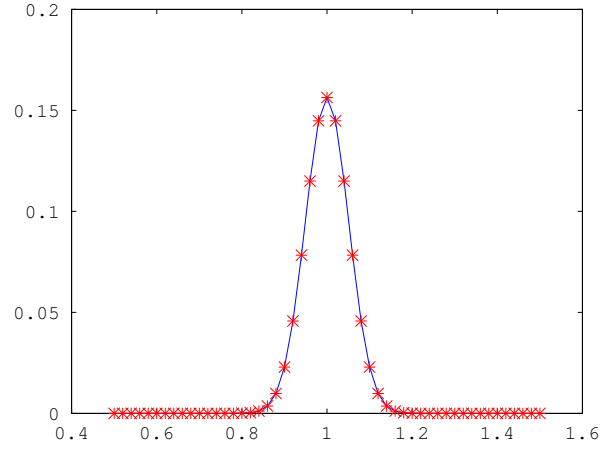


FIG. 1. Identification of the Gaussian distribution (16). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is $5 \cdot 10^{-5}$, see equation (19).

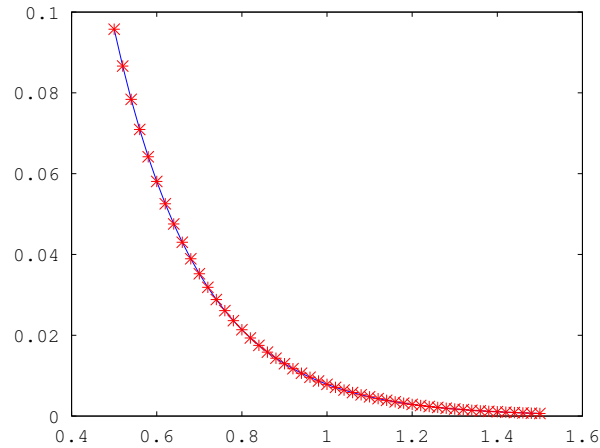


FIG. 2. Identification of the exponential distribution (17). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is 10^{-4} , see equation (20).

C. A single measured observable

We use the same system as in the Section III B except that here we consider the extreme case when only one observable is available as function of the control field. The observable is the projection $|e_3\rangle\langle e_3|$ to the third eigenstate.

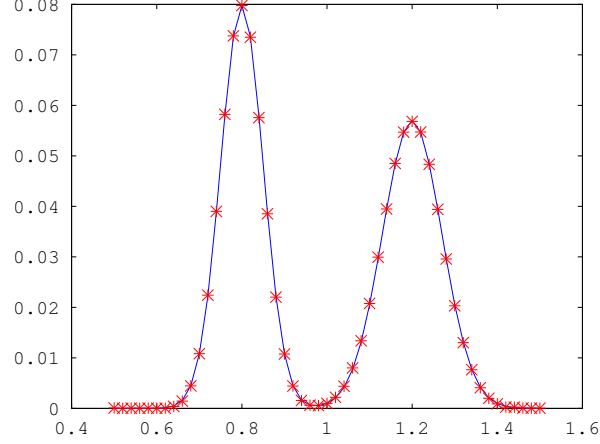


FIG. 3. Identification of the bi-modal distribution (18). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is $6 \cdot 10^{-5}$, see equation (21).

Iteration $n = 1$							
Real distribution				Initial guess			
$O \setminus u$	u_1^1	u_2^1	u_3^1	$O \setminus u$	u_1^1	u_2^1	u_3^1
O_1							
O_2							
O_3							
O_4							

TABLE I. The match of the observation distributions for the bi-modal distribution (18). We plot the histograms at the start of the inversion algorithm i.e., iteration step $n = 1$. In the left sub-table are the histograms of the observations $\sum_{k=1}^L \xi_k^{real} \delta_{|\langle \Psi(T, H, u_1^1, y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2}$ with the real distribution; in the right sub-table the histograms correspond to the initial guess which is the uniform distribution. Each column corresponds to a control field, here only the first 3 control fields u_1^1 , u_2^1 and u_3^1 are shown. Each line corresponds to a specific observable in the SCO set \mathcal{O} . The initial guess is seen to be a poor approximation, as the histograms in left and right sub-tables differ substantially.

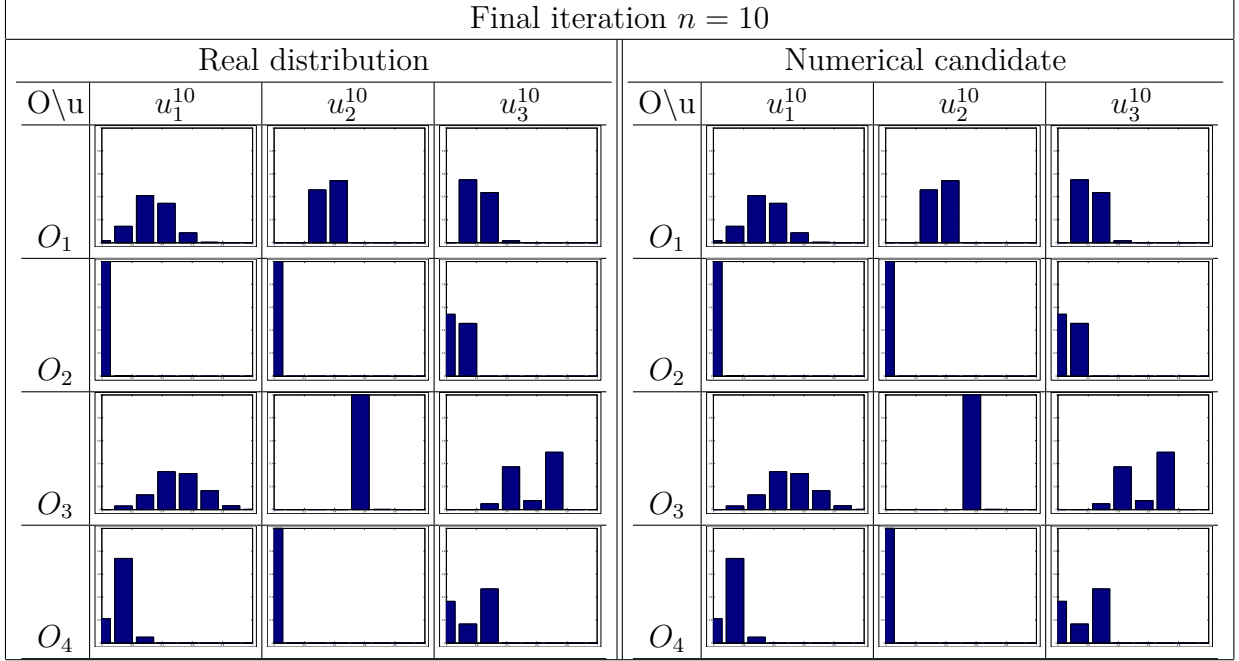


TABLE II. Converged result (iteration $n = 10$) from initial guess Table I. The identification works well as the left and right sub-tables match. As explained, the control fields are chosen randomly at each iteration and in particular the controls chosen at iteration $n = 1$ and $n = 10$ are not the same (otherwise the histograms corresponding to observations with the real noise distribution would be the same as in Table I left sub-table). Here the results for u_1^{10} , u_2^{10} and u_3^{10} are displayed.

The initial guess is:

$$\mu_b^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \quad (22)$$

The coefficient $\mu_{2,1} = \mu_{1,2} = 5$ is fixed, as previously stated, but now it is treated as a constraint by the algorithm (which will thus only optimize the other coefficients). The algorithm minimizes the difference:

$$\mathcal{J}(\mu, (\xi_k)_{k=1}^L; (u_i)_{i=1}^{N_u}) = \log \left\{ \frac{1}{N_u} \sum_{i=1}^{N_u} \mathcal{W}_1 \left[\sum_{k=1}^L \xi_k \delta_{|\langle \Psi(T, H, u_i; y_k, \mu, \Psi_1^0), e_3 \rangle|^2}, \right. \right. \\ \left. \left. \sum_{k=1}^L \xi_k^{real} \delta_{|\langle \Psi(T, H, u_i; y_k, \mu_{real}, \Psi_1^0), e_3 \rangle|^2} \right] \right\}. \quad (23)$$

The distribution Y^{real} tested is the bi-modal model. The algorithm converges and after 5,10

and 15 iterations, the dipole moments we obtain respectively have an L^2 error of 0.17246, 0.03736 and $4.5652 \cdot 10^{-4}$ respectively. In order to test the robustness of the algorithm with respect to other error norms, we use the L^2 error norm $\|\cdot\|_{L^2}$ (the square root of the sum of squares of the components). The convergence is slower than in Section III B when we treat $\mu_{2,1} = \mu_{1,2} = 5$ as a constraint.

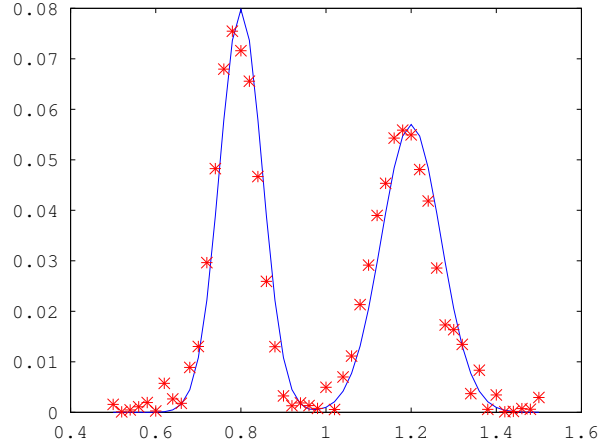


FIG. 4. Identification of the bi-modal distribution (18) after 5 iterations. The real distribution is in blue, the numerical result in red. The noise distribution starts to have the same qualitative features as ξ^{real} ; the L^2 error on the dipole moment is 0.17246.

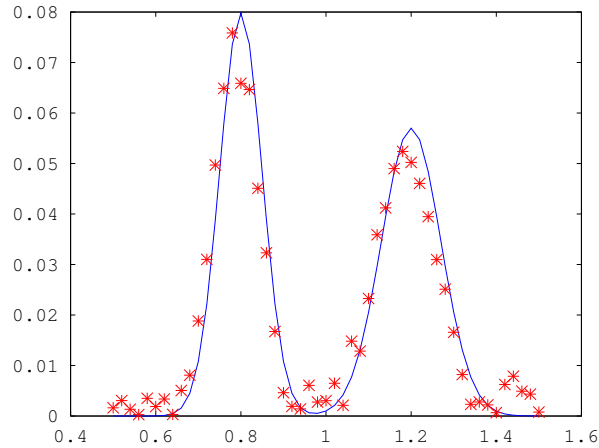


FIG. 5. Identification of the bi-modal distribution (18) after 10 iterations. The real distribution is in blue, the numerical result in red. The noise distribution starts to converge; the L^2 error on the dipole moment is 0.03736.

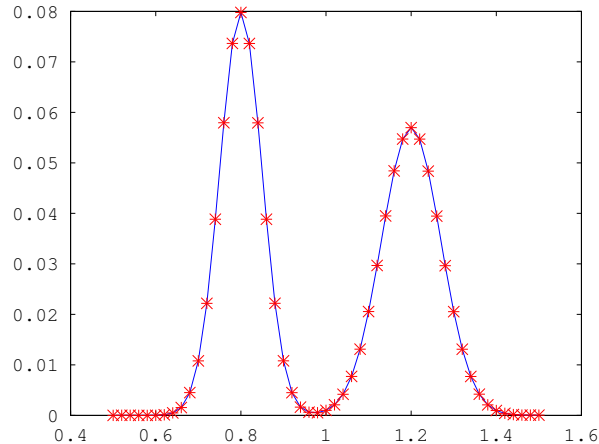


FIG. 6. Identification of the bi-modal distribution (18) after 15 iterations. The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the L^2 error on the dipole moment is $4.5652 \cdot 10^{-4}$.

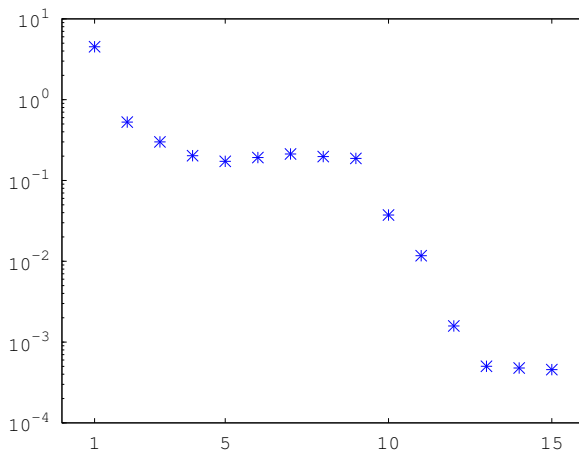


FIG. 7. The L^2 error $\|\mu^k - \mu^{real}\|_{L^2}$ for $k = 1, \dots, 15$ iterations.

IV. CONCLUSION

The inversion of the dipole moment has been considered in a model where the noise coming from the laser source is non-negligible and of unknown distribution. The model considered here has noise acting multiplicatively on the control intensity and is the same for all frequency components.

First, we proved theoretically that if one can measure repeatedly (at least) one observable for many control fields, the set of probability distributions of this observable is enough to

recover both the dipole and the noise distribution. Then, a numerical algorithm based on the Wasserstein distance between the probability distributions was proposed and seen to perform well for several different, non-perturbative, noise distributions and initial guesses for the dipole. As subject for future study is the question of how to treat noise of possibly distinct character reflected in each of the amplitudes and phases. In addition, as listed in the Introduction, there are several further issues to consider in future work for creating a realistic algorithm for quantum system data inversion.

ACKNOWLEDGMENTS

Y.F. and G.T. were partially supported by the "Agence Nationale de la Recherche" (ANR) Projet Blanc EMAQS number ANR-2011-BS01-017-01. H.R. acknowledges support from the U.S. National Science Foundation CHE-1464569.

-
- [1] Lucie Baudouin and Alberto Mercado. An inverse problem for Schrodinger equations with discontinuous main coefficient. *Applicable Analysis*, 87(10-11):1145–1165, 2008.
 - [2] S. Bonnabel, M. Mirrahimi, and P. Rouchon. Observer-based Hamiltonian identification for quantum systems. *Automatica*, 45(5):1144 – 1155, 2009.
 - [3] Constantin Brif, Raj Chakrabarti, and Herschel Rabitz. Control of quantum phenomena: past, present and future. *New Journal of Physics*, 12(7):075008, 2010.
 - [4] Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloë. *Quantum Mechanics, volume 2*. Wiley, VCH, reprint of the 2nd edition, 2005.
 - [5] Ashley Donovan and Herschel Rabitz. Exploring the Hamiltonian inversion landscape. *Phys. Chem. Chem. Phys.*, 16:15615–15622, 2014.
 - [6] Ying Fu and Gabriel Turinici. Quantum Hamiltonian and dipole moment identification in presence of large control perturbations. <https://hal.archives-ouvertes.fr/hal-01068969> (submitted), September 2014.
 - [7] J. M. Geremia and H. Rabitz. Optimal Hamiltonian identification: The synthesis of quantum optimal control and quantum inversion. *The Journal of Chemical Physics*, 118(12):5369–5382, 2003.

- [8] Y. Maday and J. Salomon. A greedy algorithm for the identification of quantum systems. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 375–379, Dec 2009.
- [9] Cédric Villani. *Topics in optimal transportation*. Graduate studies in mathematics. American Mathematical Society, cop., Providence (R.I.), 2003.
- [10] Wusheng Zhu and Herschel Rabitz. Potential surfaces from the inversion of time dependent probability density data. *The Journal of Chemical Physics*, 111(2):472–480, 1999.