

## Quantum optimal control with quantum computers: A hybrid algorithm featuring machine learning optimization

Davide Castaldo

*Dipartimento di Scienze Chimiche, Università di Padova, via F. Marzolo 1, I-35131, Padova, Italy*

Marta Rosa\*

*Dipartimento di Scienze Chimiche, Università di Padova, via F. Marzolo 1, I-35131, Padova, Italy  
and Center for Biomolecular Nanotechnologies, Istituto Italiano di Tecnologia, Via Barsanti, I-73010 Arnesano, LE, Italy*

Stefano Corni†

*Dipartimento di Scienze Chimiche, Università di Padova, via F. Marzolo 1, I-35131, Padova, Italy  
and Padua Quantum Technologies Research Center, Università di Padova and Istituto Nanoscienze-CNR,  
via Campi 213/A, 41125 Modena, Italy*



(Received 16 July 2020; accepted 27 January 2021; published 24 February 2021)

We develop a hybrid quantum-classical algorithm to solve an optimal population transfer problem for a molecule subject to a laser pulse. The evolution of the molecular wave function under the laser pulse is simulated on a quantum computer, while the optimal pulse is iteratively shaped via a machine learning (evolutionary) algorithm. A method to encode on the quantum computer the  $n$ -electrons wave function is discussed, the circuits accomplishing its quantum simulation are derived and the scalability in terms of number of operations is discussed. Performance on noisy intermediate-scale quantum devices (IBM Q X2) is provided to assess the current technological gap. Furthermore the hybrid algorithm is tested on a quantum emulator to compare performance of the evolutionary algorithm with standard ones. Our results show that such algorithms are able to outperform the optimization with a downhill simplex method and provide performance comparable to more advanced (but quantum-computer unfriendly) algorithms such as Rabitz's or gradient-based optimization.

DOI: [10.1103/PhysRevA.103.022613](https://doi.org/10.1103/PhysRevA.103.022613)

### I. INTRODUCTION

The first proposals for a new paradigm of computation that would have gone beyond standard computer science date back to the works of Feynman and Deutsch [1,2]. In their seminal contributions for the first time it was argued that quantum mechanics could have been exploited, as much as classical physics, as a background for a new way of computing.

Nowadays quantum computation is an extremely active field owing to the fact that such a paradigm of computation has been proven potentially disrupting and able to realize very challenging computational tasks in a faster and more efficient way than our classical most powerful computers [3]. At the origin of this discipline there is the idea that a quantum system may be, by far, more suitable to simulate the properties of another quantum system [1]; it is equally true that the physical realization of quantum hardware can not be independent of the very precise knowledge of the underlying physics and of the capability to control its dynamics with an impressive spatial and temporal resolution [4,5]. From these considerations naturally comes looking at the development of new quantum technologies and the fundamental research in the field of atomic, molecular, and optical physics [6,7] as two

mutually supporting processes. An example of this philosophy is quantum optimal control theory (QOCT) [8,9] where the knowledge of the system (its assumed Hamiltonian) is sufficient to suitably shape, in both time and space, an external perturbation to achieve a desired response of such system. Such a strategy has been applied in several fields ranging from photochemistry to quantum gate synthesis [10–12]; notably, with regard to the last example, QOCT is generally conceived as a tool to improve quantum devices' performance while we rarely think of a quantum computer (exception made for closed-loop feedback control experiments as in Ref. [13]) as an active party of the optimization process. In this work we put ourselves in this latter perspective by developing a hybrid algorithm which features quantum and classical computation to solve an optimal population transfer problem [14]; the approach we propose aims to complement, in a quantum computing fashion, the experiment proposed in the seminal work of Judson and Rabitz [15]. They demonstrated an optimal control protocol based on the experimental measurement of a property disclosing the molecular wave function evolution in the presence of an external light pulse, coupled to a learning procedure that shapes the control pulse to obtain the desired excitation. In the present work the quantum computer replaces the experimental apparatus, extending the range of application of this approach to those systems for which setting up such an experiment from scratch is not feasible (e.g., molecules undergoing photodamage during the experimental optimization,

\*marta.rosa@unipd.it

†stefano.corni@unipd.it

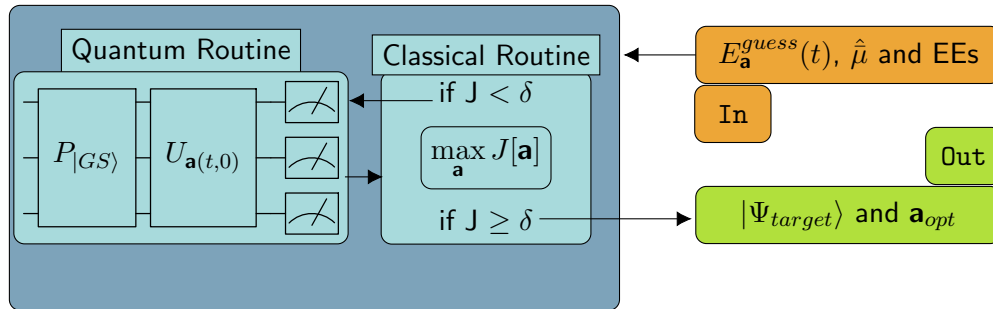


FIG. 1. Schematic diagram for the hybrid algorithm. The input is an initial guess field specified by the set of control parameters  $\mathbf{a}$ , dipole moment matrix  $\hat{\mu}$ , and excitation energies (EEs) of the molecular system of interest. The output is given by the set of control parameter  $\mathbf{a}_{opt}$ , i.e., the optimal field. Parameters  $\mathbf{a}$  in the initial ansatz directly enter in the quantum gate executing the simulation of the molecular dynamics  $U_{\mathbf{a}}(t, 0)$ . Subsequently they are adjusted in a hybrid quantum-classical optimization loop until the functional  $J[\mathbf{a}]$  is above a user-specified threshold. When this loop terminates, the resulting gate sequence [ $P_{|GS\rangle}$ , encoding of the initial state,  $U_{\mathbf{a}}(t, 0)$ , evolution] can also be used to prepare the state of the computer for further uses.

molecules not synthesized yet). The choice of this particular case has been also motivated by the increasing interest in the ultrafast dynamics of electronic states that have been probed experimentally with ultrashort pulses in single-molecule experiments [16]. Moreover, here we intend to provide a further example of the application of quantum computation methods to chemistry. A great effort, to date, has been directed to the development of algorithms for the characterization of properties concerning the electronic structure of molecular systems [17]; although their importance is primary, we hope that this work can demonstrate how new quantum technologies can be exploited to address time-dependent problems to complex systems by combining state-of-the-art quantum chemistry methods with quantum simulation techniques.

In the same spirit of Li *et al.* [18], where the authors focused on the problem of state preparation, we suggest the use of a quantum algorithm as a subroutine external to the classical optimization algorithm needed to evaluate the fitness function. Subsequently the classical optimization algorithm updates the set of control parameters feeding back again the quantum algorithm until convergence is reached (see Fig. 1). A similar approach has been also devised very recently by Magann *et al.* [19]. In particular, they present a thorough discussion of the numerical errors arising in the wave function evolution algorithm, estimate the resources needed with the proposed mapping, and describe possible applications to vibrational and rotational control and to the investigation of light-harvesting complexes. In this work we will focus on different aspects of the investigation with emphasis on the implementation of the routine on noisy devices, assessing the effect of quantum noise at different levels and motivating the choice of a genetic algorithm (GA) as an effective optimization routine.

The choice of the classical optimization algorithm is crucial for the performance of the whole procedure. Due to the quantum nature of the subroutine which computes the value of the fitness function after measurement, the most natural choice is a gradient-free optimizer (although a gradient-based approach, extensively used in the literature [20], will be considered along with the Rabitz algorithm as a reference for our calculations). As shown extensively in the literature for the case of variational quantum eigensolver (VQE) [21,22]

there are various choices for this kind of methods [17,23–26]; due to promising performance of machine learning in the field of multiparameters optimization, we chose to couple the quantum routine to a machine learning global search scheme, notably a genetic algorithm. Similar approaches have been already tested in the context of QOCT, but with a classical (as opposed to the present quantum) evaluation of the fitness function [27,28].

This paper is organized as follows: in Sec. II we present the theoretical tools needed to implement the hybrid algorithm, in particular discussing the main features of QOCT and the quantum simulation of a time-dependent molecular Hamiltonian with emphasis on the choice of the mapping adopted and on the quantum circuit synthesis. Subsequently, after having discussed the computational details in Sec. III, Sec. IV is dedicated to analyzing the numerical results assessing the performance of the mapping on noisy intermediate-scale quantum (NISQ) devices together with the implementation of the hybrid algorithm for the case of the cyanidin molecule; in particular, while providing a comparison between our hybrid solution and the result obtained with state-of-the-art algorithms on classical computers (gradient-based optimization and the algorithm proposed by Zhu *et al.* [14]), we also analyze the performance of genetic algorithms when coupled to different noisy simulated quantum processors. Finally, in the Conclusions we summarize the results obtained and consider possible future perspectives of this work.

## II. THEORY

### A. Proposed algorithm for quantum optimal control with quantum simulation of a time-dependent molecular Hamiltonian

Quantum optimal control theory deals with the evolution of a dynamical quantum system undergoing an external manipulation, able to drive the system itself from a general starting state to a final excited state. The system evolution is described by the time-dependent Schrödinger equation (TDSE), while the external perturbation which allows one to control the dynamics is described adding a term to the isolated system Hamiltonian. In our case the extra term is a suitably shaped laser pulse, able to drive the system from an initial state  $|\psi(0)\rangle = |\psi_0\rangle$  to a final state  $|\psi(T)\rangle = |\psi_T\rangle$  corresponding

to a chosen final time  $t = T$ . The final state should maximize the expectation value of a chosen operator  $\hat{O}$  acting on the system. In this work, we will focus on a particular case, namely, optimal population transfer: given a molecular system in the initial state we look for the shape of the laser pulse of length  $t = T$  able to maximize the population of a specific target excited state. In other words, we want to tune the laser field profile in such a way that it provides the most precise, and experimentally attainable, optical excitation. The optimal field can be obtained maximizing the following unconstrained functional:

$$J[\mathbf{a}] = \langle \psi_{\mathbf{a}}(T) | \hat{O} | \psi_{\mathbf{a}}(T) \rangle - \int_0^T \alpha(t) |E_{\mathbf{a}}(t)|^2 dt \quad (1)$$

Here  $\hat{O} = |\Psi\rangle\langle\Psi|$  is the projector operator on the excited target state  $|\Psi\rangle$ , and the second term is a penalization factor that discourages the optimization algorithm to move towards high fluency fields; the time-dependent factor  $\alpha(t)$  allows one to enforce a given envelope of the laser pulse. Such multiplier can be relevant, e.g., when the optimal problem has to be solved for a particular experimental set up. Finally  $|\psi_{\mathbf{a}}(T)\rangle$  stands for the value of the state vector whose evolution is driven by a field  $\bar{E}_{\mathbf{a}}(t)$ . The vector  $\mathbf{a}$  specifies the set of control parameters.

There are two key elements to be defined within QOCT algorithms: one is the approach to evolve the molecular wave function under the effect of the light pulse so to get  $|\psi_{\mathbf{a}}(T)\rangle$  (and thus  $J$ ) for a given set of control parameters, the other is the optimization algorithm that iteratively leads to the best vector  $\mathbf{a}$ . For the latter algorithm (the ‘‘Classical Routine’’ in Fig. 1), we adapted a machine learning scheme as described in the Methods section. Here we discuss the devised simulation of the evolution of the molecular wave function subject to the laser pulse, which is the part running on the quantum computer (‘‘Quantum Routine’’ in Fig. 1).

Quantum simulation is a branch of quantum computation that aims to compute the evolution of a quantum system by programming the evolution of the state vector of a quantum computer such that they correspond to each other [29]. In other words, let  $U_{QC}(t, 0)$  be the time evolution operator of the quantum computer (QC) and  $U(t, 0)$  the system time evolution operator we are interested in the computation of, and we want to achieve

$$U(t, 0) \begin{array}{c} \xrightarrow{\text{encoding}} \\ \xleftrightarrow{\hspace{1cm}} \\ \xleftarrow{\text{decoding}} \end{array} U_{QC}(t, 0). \quad (2)$$

To this extent the problem of quantum simulation can be thought as a translation problem from the common language of quantum mechanics to the *dialect* of quantum computation.

### 1. Encoding of the $n$ -electrons wave function

In this work we assume that a large enough set of electronic states of the molecule has been precalculated with some suitable quantum chemistry approach. Beside the energies of such states, the method provides also the matrix elements of the interaction operator with the external control. In the present case, since the control is the time-dependent electric field associated with a laser pulse, the required operator is the dipole moment operator  $\bar{\mu}$ .

In this section we describe the method that we have used to map a molecular wave function written in the (approximated) electronic eigenstate basis, e.g., configuration interaction (CI) [30], into the state of the quantum computer.

First, for sake of clarity, we recall the common notation for an arbitrary state of the quantum computer composed of an  $M$ -qubits register given by

$$|\Phi\rangle = \sum_{\beta_1, \dots, \beta_M} c_{\beta_1, \dots, \beta_M} \left( \bigotimes_{i=1}^M |\beta_i\rangle \right), \quad (3)$$

where  $\beta_i$  denotes the state of the single  $i$ th qubit which can be either  $|0_i\rangle$  or  $|1_i\rangle$  with the coefficients of the direct product states  $c_{\beta_1, \dots, \beta_M}$  satisfying the normalization condition  $\sum_{\beta_1, \dots, \beta_M} |c_{\beta_1, \dots, \beta_M}|^2 = 1$ .

Our starting point is the molecular Hamiltonian (i.e., the standard polyelectronic Hamiltonian containing electron-electron and electron-nuclei interactions) in the presence of an external field written as

$$\hat{H}(t) = \sum_k e_k |k\rangle\langle k| - \bar{E}(t) \cdot \sum_{i,j} \bar{\mu}_{ij} |i\rangle\langle j|, \quad (4)$$

where the indexes  $k, i, j$  in the sums run over the possible  $n$ -electrons states approximated as the eigenvectors of the CI matrix;  $e_k$  is the energy of the  $k$ th electronic state,  $\bar{E}(t)$  is a time-dependent electric field associated to the controlling laser pulse, and  $\bar{\mu}_{ij}$  is the transition dipole moment between the states  $|i\rangle$  and  $|j\rangle$  (or the corresponding permanent dipole when  $i = j$ ).

With this approach we can write an arbitrary quantum state of the molecule as a coherent superposition of CI eigenvectors:

$$|\Psi(t)\rangle = \sum_k^N c_k(t) |k\rangle. \quad (5)$$

To find a link with the state of the quantum computer we encode the coefficient  $c_k(t)$  [Eq. (5)] into the bitstring state coefficient of the quantum computer where all the qubits are in the state  $|0\rangle$  but the  $k$ th qubit which is in the state  $|1\rangle$ .

More formally we get the following equivalence:

$$c_k(t) := \langle k | \Psi(t) \rangle \equiv \langle 0_1 \dots 0_{k-1} 1_k 0_{k+1} \dots 0_N | \Phi \rangle, \quad (6)$$

where we recall that  $|\Phi\rangle$  is the state of the quantum computer and  $|\Psi\rangle$  that of the molecule. Similar approaches have been recently considered for the simulation of bosonic Hamiltonians [31–33]; here we extend this mapping to the simulation of many-fermions wave functions. We may notice that this particular choice allows us to encode the coefficients corresponding to the different molecular electronic eigenstates into a particular subset of bitstring states where only one qubit at the time is in the state  $|1\rangle$ . This feature gives us the possibility to make a step further and establish a connection between the set of CI eigenvectors of Eq. (5) and the Fock states of single occupancy for an  $N$ -fermions system. It is important to highlight that such an encoding exploits a *formal* equivalence between a CI eigenvector and a single-occupancy Fock state. Indeed, from a physical point of view we know that we are dealing with very different objects: with the former representing an  $n$ -electrons wave function and the latter a single-body

quantum state. To conclude this parallelism we can think of such a mapping as an interpretation of the CI Hamiltonian [Eq. (4)] as a single-body operator acting on the states of a quasiparticle, the different electronic states being the modes in which the latter can be found.

We can now exploit this equivalence to make use of the techniques that have been developed to perform the quantum simulation of a fermionic Hamiltonian [34,35]. We chose to adapt the well known Jordan-Wigner (JW) mapping [36] to an  $n$ -electrons wave function by means of Eq. (6). We recall that the JW transform allows one to express fermionic operators in terms of Pauli matrices  $\{\sigma_x, \sigma_y, \sigma_z, \mathbb{1}\}$  due to the following expressions:

$$a_j = -\bigotimes_{k=1}^{j-1} \sigma_k^z \sigma_j^-, \quad a_j^\dagger = -\sigma_j^+ \bigotimes_{k=1}^{j-1} \sigma_k^z, \quad (7)$$

where  $a_j$  and  $a_j^\dagger$  are, respectively, the annihilation and creation operators for the  $j$ th particle while  $\sigma_k^z$  and  $\sigma_k^\pm$  are Pauli operators acting on the  $k$ th qubit.

Since we reduced the encoding of the  $n$ -electrons states into the single-occupancy Fock space we can disregard the string of Pauli operators  $\sigma^z$ s leading to a much simpler expression. This simplification would not be possible if we had included in our encoding scheme the representation of a Slater determinant with a bitstring state with more than one qubit in the state  $|1\rangle$ ; indeed, in that case, we should have taken into account permutation properties of the many body quantum states.

Finally, we end up with a simple equivalence between the  $k$ th electronic state and the reference state of the quantum computer (i.e., when all qubits are in  $|0\rangle$ ):

$$\begin{aligned} |k\rangle &\equiv \sigma_k^+ |0_1 \dots 0_{k-1} 0_k 0_{k+1} \dots 0_N\rangle \\ &= |0_1 \dots 0_{k-1} 1_k 0_{k+1} \dots 0_N\rangle. \end{aligned} \quad (8)$$

By substitution of the previous expression in the molecular Hamiltonian [Eq. (4)] we obtain

$$\begin{aligned} \hat{H}(t) &= \sum_k \theta_{kk}(t) (\mathbb{1}_k - \sigma_k^z) + \sum_{p>q} \frac{\alpha_{qp}(t)}{2} [\sigma_q^x \sigma_p^x + \sigma_q^y \sigma_p^y] \\ &= \tilde{H}_z(t) + \tilde{H}_{x,y}(t), \end{aligned} \quad (9)$$

where, in the last equivalence, we use the subscripts to stress the dependence of the different terms on a particular subset of Pauli operators and adopt the following shorthand:

$$\begin{aligned} \theta_{kk}(t) &= e_k - \bar{E}(t) \cdot \bar{\mu}_{kk}, \\ \alpha_{qp}(t) &= \bar{E}(t) \cdot \bar{\mu}_{qp}. \end{aligned} \quad (10)$$

### 2. Trotterization of a time-dependent Hamiltonian

As we can see Eq. (9) is an explicit transposition of the time-dependent Hamiltonian in the  $n$ -electrons basis in terms of one-qubit operators (diagonal elements) and two-qubit operators (off-diagonal elements). We consider now how to derive the circuit representation of the time evolution operator  $U_{QC}(t, 0)$  based of such expression. It is well known [37] that for an Hamiltonian such as the one of Eq. (4) the time evolution operator is given by a time-ordered expansion according

to the equation

$$U(t, 0) = \mathcal{T} e^{-i \int_0^t \hat{H}(t') dt'}, \quad (11)$$

where  $\mathcal{T}$  is the Dyson's time-ordering operator.

To compute the time evolution operator by making use of the Trotter-Suzuki (TS) approximation [38] we discretize the time axis in  $K$  time slots of width  $\Delta t = t_{j+1} - t_j$  and approximate the integration as if the Hamiltonian were constant within the given time interval:

$$U(t, 0) \approx \mathcal{T} e^{-i \sum_{j=0}^K \hat{H}(t_j) \Delta t}. \quad (12)$$

We are now able to decompose the time evolution operator  $U(t, 0)$  as a stepwise time-independent evolution and to apply the TS decomposition to each single-step evolution operator:

$$U(t, 0) = \prod_{j=0}^K U(t_{j+1}, t_j), \quad (13)$$

$$U(t_{j+1}, t_j) \approx e^{-i \tilde{H}_z(t_j) \Delta t} e^{-i \tilde{H}_{x,y}(t) \Delta t}. \quad (14)$$

In the last equation we made use of the first-order TS approximation, but an extension to higher order approximation is straightforward [39]. Notice that in the following we shall refer to this integration scheme for a time-dependent Hamiltonian as Trotter-Suzuki piecewise constant approximation (TS-PCA).

The implementation of the first term of Eq. (14), i.e., the diagonal evolution of the molecular wave function, has been discussed in Ref. [35] consisting in the application of a phase gate on each qubit with the phase specified by  $\theta_{kk}(t_j)$ . Following the same reference the exponential of a string of Pauli operators has been implemented; in this work we considered three different implementations of the circuit accounting for  $e^{-i \tilde{H}_{x,y}(t) \Delta t}$  as summarized in Table I. The second circuit is exactly the implementation of the mapping we presented in the previous section where the string of  $\sigma_z$  (CNOT cascade) has been disregarded. Furthermore we also considered a modification to the previous circuit in which all the gates accounting for the operations containing the exponential of  $\sigma_x$ s and  $\sigma_y$ s are performed separately in order to diminish the single-qubit operations needed to rotate the single-qubit states into the correspondent  $\sigma_x/\sigma_y$  eigenstate basis.

Finally, we highlight that on top of the TS approximation of Eq. (14) the off-diagonal evolution is also approximated by a first-order product formula; in fact the Pauli string operators arising from different pair of states do not, in general, commute. As a consequence we have

$$e^{-i \tilde{H}_{x,y}(t) \Delta t} \approx \prod_{\lambda} \Gamma(\alpha_{\lambda}(t_j)) = \prod_{\lambda} \Gamma_x(\alpha_{\lambda}(t_j)) \Gamma_y(\alpha_{\lambda}(t_j)), \quad (15)$$

where  $\Gamma$  is a shorthand for the boxed circuits (second row of Table I) and the last identity emphasizes the factorization of the circuit in two parts due to the  $\sigma_x$  and  $\sigma_y$  terms.

### 3. Quantum algorithm complexity

A sufficient condition for an algorithm to be told efficient is that each step of its implementation must be accomplished in a number of operation and by requiring an amount of resources which increase polynomially with the system size [40]. Here we consider in more detail the complexity of our



TABLE I. Quantum circuits representing the coupling of different electronic states in the presence of an external perturbation. Here  $q$  and  $p$  correspond to different CI eigenstates, and  $\lambda$  is a shorthand for an arbitrary couple of states. In the first circuit a readaptation of the JW mapping in the  $n$ -electrons basis is shown; notice that with this approach an additional string of  $\sigma_z$ s is present. Dashed lines of the CNOT gate stand for a CNOT cascade between all the qubits with label greater than  $q$  and lower than  $p$ . Dashed boxes refer to the  $\sigma_x$  contribution, and dotted boxes refer to the  $\sigma_y$  part. Notice that the last circuit, unlike the previous ones, refers to the implementation of the whole operator  $e^{-i\tilde{H}_{x,y}(t)\Delta t}$ .

Pauli representation	Circuit
$e^{-\Delta t \frac{i}{2} \alpha_{qp}(t_j) [(\sigma_q^x \sigma_p^x + \sigma_q^y \sigma_p^y) \otimes_{l=q+1}^{p-1} -\sigma_l^z]}$	
$e^{-\Delta t \frac{i}{2} \alpha_{qp}(t_j) (\sigma_q^x \sigma_p^x + \sigma_q^y \sigma_p^y)}$	
$[\prod_{\lambda} \tilde{\Gamma}_x(\alpha_{\lambda}(t_j))] [\prod_{\lambda} \tilde{\Gamma}_y(\alpha_{\lambda}(t_j))]$	

implementation to show that the number of operations scales polynomially with the dimension of the Hilbert space of the molecular  $n$ -electrons states. It is important to notice that, even though the implementation is efficient, the number of Slater determinants required still scales exponentially with the number of electrons and spin orbitals used; as such we will also comment on other options that represent different compromises in terms of quantum hardware budget. We also point out the interesting analysis proposed in Ref. [33] where quantum circuits to switch encoding during the computation are presented and suggested as a possible strategy to improve the ratio between efficiency of the implementation and quantum hardware requirements. Further we will not discuss either the efficiency of the state preparation algorithm for a many-fermions state or the efficiency of the evolution algorithm (Trotter-Suzuki decomposition) as they are both very well known techniques in the area of quantum simulation [34,38]. To this extent, as a first point we consider the qubit efficiency of the encoding scheme; our mapping provides a linear scaling, with respect to the number of qubits, of the Slater determinants' encoding. In fact, as we have pointed out previously, the electronic state  $|k\rangle$  is mapped onto the bitstring state with all the qubits in state  $|0\rangle$  but the  $k$ th qubit. This implies that, considering the Hilbert space resulting from the tensor product of the individual qubits space, we are limiting our interest to a subspace made up of  $Q$  bitstring states. Furthermore, even considering error correction procedures needed for fault-tolerant quantum computation, the qubit scaling remains linear since  $O(Q)$  additional qubits are needed for their implementation [41]. Now, concerning the number of operations needed to simulate a single-step evolution operator [Eq. (14)], we will consider separately the contributions due to the diagonal elements of the Hamiltonian (i.e.,  $e^{-i\tilde{H}_z(t_j)\Delta t}$ )

and the ones arising from the off-diagonal elements (i.e.,  $e^{-i\tilde{H}_{x,y}(t_j)\Delta t}$ ). The former unitary operation needs exactly  $Q$  single-qubit operations to be performed; the latter, as shown in Table I, have been considered in different versions. The first circuit shown in Table I, i.e., the one not considering any constraint on the occupancy of our fictitious Fock states, represents an upper bound for the estimation of gate complexity of this implementation. Due to its hermiticity, this operator has  $Q(Q-1)/2$  independent terms, each of which can be implemented with  $2Q$  two-qubit operations, as discussed in detail in Ref. [42]. Altogether we get a total gate count  $\sim O(Q^3)$ . We can reduce the two-qubit operation complexity of one power if we consider a circuit in which we make explicitly use of the mapping boundary within the single-occupancy Fock space (second row of Table I). In this situation each term of Eq. (15) needs only two CNOT gates, leading to a total gate count of at most  $\sim O(Q^2)$ . Of course the matrix associated with  $\tilde{H}_{x,y}(t_j)$  may be sparse, reducing the scaling or at least the prefactor. It may be worth noticing that these considerations do not take into account that the actual implementation of a two-qubits gate on the hardware could require additional  $O(Q)$  operations, as not all the qubits in the quantum processor interact directly between each other. On the downside, this encoding exploits only a subset of the computational space of the quantum hardware. Specifically, the dimension of such subspace is the same as the number of qubits. Anyway, the idea can be extended by considering mapping to larger subspaces, such as double, triple, etc., excitation vectors (e.g., vectors with all the qubits 0 but two or three or more). The dimension of such space is the binomial coefficient  $\binom{Q}{m}$ , where  $m$  is the excitation order. While using a larger sharing of computational space, such subsets require a larger number of CNOT gates to encode each term of Eq. (15), since in the  $m$ -excitations subspace one

needs strings of  $2m$  Pauli matrices (with  $m \leq Q/2$ ) to specify a particular bitstring state; each term is a sum of  $2^{2m-1}$  of such strings. Depending on the features of the quantum hardware and the problem at hand, the best  $m$ -excitation subspace may be chosen.

Other options for the mapping can be considered as well, such as the use of a binary mapping to encode the Slater determinant integer label (sorted, for instance, according to increasing energy ordering) and the bitstring state of the computer. In this case we will have a qubit register dimension of  $\log_2 Q$ . This option has been considered in Ref. [19]; even though more efficient from an information storage standpoint, this mapping requires more complex expressions for the Pauli representation that, in turn, imply a greater number of two-qubit gates. Again, we notice that the choice of the mapping should be done carefully evaluating the balance between circuit depth, classical, and quantum computational resource requirements.

To conclude this discussion we recall that, in the particular application considered in this work, i.e., the quantum simulation of a time-dependent Hamiltonian, the total complexity must take into account also the number of propagation steps  $K$  needed. As a consequence, the final estimation of the two-qubit operations for our circuit is  $\sim O(Q^2 K)$ , still ensuring the efficiency (as defined above) of the implementation.

### B. Optimization of the $J$ functional

Different approaches are available to optimize the  $J$  functional [Eq. (1)] with different pros and cons depending on the system under study and the purpose of the optimization [9]. A notable example is represented by the well-established Rabitz algorithm [10], which shows excellent convergence properties both in vacuum and in the presence of a solvent [43]. Despite its success, this algorithm is intrinsically unsuitable to be used coupled with a quantum simulation algorithm, as it requires the knowledge of the state of the system at each instant of its evolution. This means we should measure the state of the quantum computer at each time step, nullifying the advantages of the quantum procedure. Notwithstanding, in the choice of the final algorithm to be used for our procedure, we have kept the Rabitz algorithm as a reference to assess the quality of our results.

In this work we follow the idea of Li *et al.* [18] working out a hybrid algorithm in which quantum computation accounts for propagating the system wave function under the influence of the external field, while an external classical algorithm optimizes a set of control parameters for the field. The reason behind our choice is that, doing so, we are able to assign to the quantum processor the computationally most demanding part of the algorithm (i.e., the propagation of the wave function and consequently the evaluation of the functional  $J$ ) leaving to the classical computer the task of choosing how to proceed in the control parameter landscape exploration. Nevertheless, exploitation of quantum search algorithms in conjunction with the quantum simulation circuits discussed above (see Table I) could allow a further speedup of this implementation. The last point is still an open problem and could be the subject of a future work.

Differently from Li *et al.* [18], we focus on classical optimization algorithms that do not require the knowledge of the gradient of the objective function with respect to the optimization parameters. As a matter of fact, calculating the gradients adds extra noise to the procedure, as noted for VQE [44]. A first possibility that we considered for the classical optimization procedure is the Nelder-Mead (NM) algorithm [45], which has been applied successfully in the CRAB and DCRA methods [46,47] and has been the choice for various hybrid quantum classical approaches to determine the electronic structure [17]. This algorithm is a direct search method which first constructs a simplex in the multidimensional parameter space, and then modifies the simplex at each iteration under a set of predefined rules moving its centroid towards a minimum (hence the alternative name of downhill simplex method). Even though this method has been proved extremely robust [48], at the same time is affected by a very slow convergence rate, already considering a few tens of control parameters [49], as we will show in the Results section.

From a broader perspective, optimal control is an instance of optimization problems where machine learning (ML) has shown high potential [12]: for this reason, in order to improve the performance of the optimization algorithm, we decided to implement a GA [50], which is a direct search method evaluating multiple copies of the system and combining the best results to create a better generation, hence the name “genetic algorithm.” In our case the algorithm propagates the system under the influence of different fields, and then the fields parameters are modified learning from the best results obtained (see Sec. II). We may notice that both in NM and GA methods a simple and massive parallel implementation is possible, as we are dealing with algorithms which rely on several independent evaluations of the functional  $J$ .

In both these options, we need to find a set of control parameters  $\mathbf{a}$  that explicitly define the shape of the external perturbation (the field) in order to optimize the evolution of our system and the value of the functional  $J$ . To obtain such a parametric dependence we decided to write the laser pulse as

$$\vec{E}_{\mathbf{a}}(t) = \sum_{\alpha} \vec{u}_{\alpha} \left[ a_{0,\alpha} + \sum_j^M a_{j,\alpha} \sin\left(\frac{j\pi t}{T}\right) \right], \quad (16)$$

where  $\vec{u}_{\alpha}$  is a unit vector with direction specified by the index  $\alpha$  running over the three Cartesian components.

Here the field takes the form of a sum over different harmonics with frequency  $\omega_j = \frac{j\pi}{T}$ , where  $T$  is the time duration of the laser pulse. The subscript emphasizes the parametric dependence on the set of amplitudes  $\mathbf{a}$  which are, ultimately, our set of control parameters. We chose to use predetermined  $\omega_j$  values corresponding to Fourier frequencies, which have the advantage of describing a field which starts and ends at zero (provided it has no constant component  $a_0$ ) and thus more easily obtained with an experimental apparatus. Moreover, if a specific frequency range has to be considered due to particular experimental limitations, this is easily accomplished by choosing the terms to include in the summation of Eq. (16). It may be worth noticing that alternative optimization algorithms implemented the optimization of the frequencies as well as of the amplitudes, in order to allow the systems to reach the global optimal solution [46]. In our work the opti-

mization of the amplitudes alone was enough to escape local minima and reach a satisfactory optimal solution, due to the implementation of the evolutionary algorithm. Nevertheless, it would be possible both for the NM and the GA optimizers to include the frequencies in the optimization procedure. In the specific case of the GA it would be a straightforward procedure that would be easily implemented at only the price of increasing the number of optimization parameters.

### *Genetic algorithms to optimize $J$*

Due to the quantum nature of the subroutine which computes the value of the fitness function, we decided to rely on a gradient-free optimizer. One of the most popular and transferable choices for a reliable algorithm in a multidimensional space are genetic algorithms, considered to be part of the broad family of machine learning techniques.

The term genetic algorithms is due to the inspiration that these optimization strategies draw from the process of development of new species that takes place in nature [50]. As the physical performance and the consequent success of an individual within a herd are related to the chromosomes, so the value of a function  $f(\mathbf{x})$  depends on the variables  $\mathbf{x}$ . In this picture the performance of an individual (i.e., the value assumed by the function  $f$ ) is completely specified by its genes (i.e., the value assumed by the set of variables  $\mathbf{x}$ ). In GAs several initial sets of parameters are generated (either by an educated guess or at random) thus forming a population of individuals; the optimization of the function is provided by letting the population evolve, forming a new generation of individuals, in accordance with the foundational rules of the evolution process (note that in our implementation the size of the population is kept constant, nonetheless alternatives with a varying number of individuals are possible). In our case each individual is a vector of amplitudes  $\mathbf{a}$  describing the field used to evolve the wave function of the system in its final state. Generally speaking, the final state is the objective function of the algorithm and can be determined by numerical calculation or experimentally. Each of the individuals is then rated according to its fitness, which describes the success of the individual parameters. From our standpoint it is easy to understand that the fitness function corresponds to the  $J$  functional. The fitness could be determined also from an experimental procedure, measuring any observable effect consequent to the electronic excitation, as charge transfer or light emission.

Concerning the mimicking of the evolution process, three main traits are usually considered in this framework: selection, recombination, and mutation.

(a) *Selection.* An example of biological merit is the transfer of one's own genetic kit: this is most commonly achieved by successful individuals. This behavior can be transferred in the implementation of a GA in different manners [51,52]. We chose to build the next generation on top of the best  $m$  individuals. It is important to notice that the ratio between the selected individual and the size of the population ( $\frac{m}{n}$ ) is a critical parameter: indeed, if a too small number of individuals is selected, important features allowing to drive the optimization to the global minimum may be lost determining a premature convergence. Conversely, if too many individuals influence the next generation the successful traits could not emerge.

This behavior would result in an optimization resembling a random exploration of the parameter landscape.

(b) *Recombination.* Once a subset of the population is selected to give birth to the new generation, the recombination process occurs. Here two individuals are chosen randomly between the selected ones, and their amplitudes are mixed to give a new child vector. There are different flavours of recombination which allow more or less mixing of the two parent individuals. In our algorithm, each amplitude in the final new vector has 0.5 probability to come from one or the other parent. As a consequence, recombining the same two individuals a second time does not give the same child.

(c) *Mutation.* With the last procedure (i.e., recombination) the rules to compute the offsprings are set. Nevertheless, eventually, a mutation can arise changing randomly a gene in a certain individual. In our implementation this phenomenon is performed summing a random number to an individual amplitude. Each random number is extracted from a Gaussian distribution specified by a value of mean  $\mu$  and standard deviation  $\sigma$ . We point out that the mutation is a nondeterministic process happening for each amplitude of each individual with a given probability  $P$ . In this implementation the choice of the three parameters (namely,  $\mu$ ,  $\sigma$ , and  $P$ ) is crucial to guarantee a sufficient exploration of the parameter landscape; in particular higher values of these three parameters lead to a higher chance and entity of the exploration.

Mutation and recombination can be combined in different ways to finally obtain the full population starting from the subset of selected individuals. Parameters for mutation and recombination can change during the evolution depending on the fitness behavior. Generally speaking, recombination favors convergence toward a (local) solution, while mutation, as already mentioned, favors exploration.

The three steps described so far constitute the main blocks of the genetic algorithm we adopted in this work. It is important to notice that a genetic algorithm reaches the global minimum of a multivariate function, as at each iteration it acquires more information about the position of the minimum. The learning loop is actually made by the evolutionary process and it is efficient only when an accurate balance between recombination and mutation is, achieved, avoiding, in this way, to get stuck in local minima or, on the other hand, to lose information due to an excessive mutation rate of the individuals.

Despite the number of choices needed to set up an evolutionary procedure, most of the optimization problems are robust with respect to a large interval of values for the parameters, with the difference between different choices translating mainly in a faster or slower convergence towards the optimal solution; see Appendix B for an account of the GA implementation performance with respect to increasing system complexity. Our final choice was very simple: we recombine the subset of individual to obtain the full population (e.g., 60 starting individuals, 10 selected and recombined randomly to return 60 new individuals), and then we mutate all of them (i.e., each amplitude of each individual is mutated with a given probability). Several different choices could have been adopted to improve the efficiency of the algorithm and implement an automatic behavior; nevertheless our purpose was to find a simple algorithm able to work efficiently combined

TABLE II. Genetic parameters adopted in the numerical applications. Different values have been considered for the complete set of excited state (first column) and for the smaller subset (second column). Likewise different parameters have been adopted to enhance a converging or exploring behavior. In brackets, when needed, values adopted in the hybrid implementation to reduce the computational cost.

	11 states		3 states	
	Exploration	Convergence	Exploration	Convergence
Individuals	70 (80)	70 (50)	40	40
Selected individuals	10	10	10	10
$P_{\text{recombination}}$	1	1	1	1
Mutated individuals	70 (80)	70 (50)	40	40
$P_{\text{mutation}}$	0.2	0.2	0.2	0.2
$\mu$	0	0	0	0
$\sigma$	0.001	0.0001	0.001	0.0001

with a propagation performed on a quantum computers. More details on the choice of the parameters are given in the following section. As shown in Sec. IV, the results obtained were satisfactory.

### III. COMPUTATIONAL DETAILS

All the calculations were performed with OCPy [53], a homemade stand-alone Python software which features all the possibilities we have considered so far, i.e., hybrid or classical propagation coupled to different optimization algorithms, like Rabitz, GA, NM, and Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods.

#### A. Quantum chemistry calculations

Geometry optimization of the cyanidin molecule was obtained using the NWChem software [54] at the density functional theory (B3LYP, cc-pVDZ) level of theory. The first 10 electronic excited states were computed with GAMESS [55] at the CI single (CIS) level of theory using a 6-31G\*\* basis set. The suitability of the number of excited states considered to evaluate the optimal dynamics involving a sufficiently large set of electronic states is assumed on the basis of a previous related work [43]. It may be worth noticing that, when QOCT is applied to an actual experimental situation, a more accurate description of the system is needed; the CIS method is usually affected by excitation energies overestimation [56]; nevertheless an accurate description of the molecule electronic structure goes beyond the purpose of the present study.

#### B. Optimization procedure

As previously mentioned, our aim is to maximize the population transfer between the ground state  $|GS\rangle$ , that we set as the initial state  $|\psi_0\rangle$ , and a given excited state, here the first excited state,  $|1\rangle$ . We applied the optimal control procedure considering both a manifold of 11 electronic states (the ground state plus the first 10 excited states) and a smaller subset consisting in the ground state and the first two excited states. For sake of brevity, in the following we will refer to the former as system **Cyan11** and to the latter as system **Cyan3**.

The duration of the laser pulse is set equal to 250 a.u., i.e.,  $\approx 6$  fs. The reason to focus on this kind of perturbation is that a genuinely selective pulse is such if the electron dynamics is fast enough to occur in the manifold of electronic states

unaffected by the consequent nuclear relaxation. In fact, as shown in Ref. [57], a  $\pi$ -pulse (i.e., a light pulse which causes population inversion in a two-level system) is not able to achieve the desired result in such a short time. In all cases in which we consider a classical method to compute the numerical propagation of the wave function we have adopted a Euler propagation method using a time step of 0.01 a.u. (i.e.,  $\approx 2 \times 10^{-4}$  fs); other, more robust, methods are possible (e.g., operator splitting technique in Ref. [57]), but nevertheless the choice of Euler propagator was done to keep the procedure as simple as possible and focus our discussion on the comparison between the classical and the hybrid implementation.

Concerning the parameters choice for the reference algorithms (Rabitz and BFGS), the initial guess field is a sinusoidal laser pulse resonant with the transition of interest ( $|GS\rangle \rightarrow |1\rangle$ ,  $\omega = 0.125$  a.u.) with amplitude set equal to 0.01 a.u. along all the three coordinates. To enforce small values of the field amplitudes we set the penalty factor  $\alpha(t)$  equal to 10 (value in a.u., i.e.,  $e^2 a_0^2 \hbar^{-1} E_h^{-1}$ ) and kept constant throughout the evolution. NM calculations were performed with the initial guess simplex built upon the same guess pulse of the Rabitz algorithm by modifying per each vertex, one at the time, the value of a single component of  $\mathbf{a}$ . More specifically, if the  $i$ th component of the guess field is null it is automatically set to 0.00025 a.u., otherwise it is increased by 5%.

In Table II we show the genetic parameters adopted for our numerical calculations. We decided to alternate optimizations runs favoring exploration of the landscape and convergence runs towards the optimal solution. The only difference between the two is the value of  $\sigma$  in the mutation procedure. A larger sigma means that the random numbers added during mutation span a larger interval, favoring exploration. In the exploration runs  $\sigma$  value is equal to 0.001 a.u., while in the convergence one it is equal to 0.0001 a.u.. Our first and simpler choice was to perform only two runs, the first favoring exploration and the second favoring convergence. Depending on the system, a different number of iteration is needed: regarding **Cyan11** the total number of generations was 100 (50 generations of exploration and 50 generations of convergence), with a population size of 70 individuals. On the other hand, considering the case of **Cyan3**, the complete computation consisted of 30 generations (15 generations of exploration and 15 generations of convergence), with a population size of 40 individuals. Subsequently, we worked



TABLE III. Noise models adopted in the numerical applications. Different error probabilities have been considered in order to assess the behavior of the genetic algorithm to the increase of the noise during the execution of the quantum simulation circuit. In brackets we give the error probability for two-qubit operations. Depending on the different kind of noise used and their intensity, in the paper we refer to the different models with the following shorthand: - model (BF model); depolarizing error on single-qubit gates [SQ depol. (1) or SQ depol. (2), depending on the intensity]; and mixed model when both the previous channels are included.

	BF model	SQ depol. (1)	SQ depol. (2)	Mixed model
Depolarizing channel probability	0 (0)	$1 \times 10^{-5}$ (0)	$5 \times 10^{-5}$ (0)	$5 \times 10^{-5}$ (0)
Bit-flip channel probability	$1 \times 10^{-5}$ ( $1 \times 10^{-5}$ )	0 ( $1 \times 10^{-5}$ )	0 ( $5 \times 10^{-5}$ )	$5 \times 10^{-5}$ ( $5 \times 10^{-5}$ )

in reducing the number of iterations and replicas in order to lower the computational effort when the propagation is performed on the 11 qubits computer. Here we performed six runs made of 10 iterations alternating convergence and exploration until the algorithm achieves a threshold value for the target state population (that we set equal to 0.95). Then we end out optimization with a convergence run made of 40 iterations. Altogether the total number of iterations is still 100. Here there is a further difference between the exploration and convergence run, apart for the value of  $\sigma$ : indeed during the exploration phase a larger population size has been adopted (80 individuals), while the convergence run has been performed with a population size of 50 individuals. This tailoring of the optimization procedure was made specifically for the problem under study (i.e., system and hardware) and does not claim to be general or optimal. It was chosen on the basis of the obtained performance and can be improved or replaced depending on the final purpose, usually exploiting a small number of test runs. Finally, the value of the penalization factor for all the numerical calculations performed with the GA was kept constant and equal to 1 a.u. as in this case to limit the intensity of the field is more effective to choose the amplitude within a fixed interval,  $\pm 0.005$  a.u.

### C. Quantum propagation

Quantum circuits accounting for the propagation were built due to the Python library Qiskit [58]. The time step  $\Delta t$  used for all our simulations is equal to 1 a.u. (i.e.,  $\approx 2 \times 10^{-2}$  fs); see Sec. IV for a more detailed account on the benchmark of the propagation. All the results concerning the hybrid implementation of the optimal control algorithm were obtained using the second circuit of Table I. In all the computations performed on a quantum hardware we used the five-qubit chip IBM Q X2 [59]; analogously all the noisy simulations were carried out including single-qubit measurement errors, thermal relaxation, and depolarizing noise tuned upon experimental parameters of the same quantum processor. Each circuit (for both quantum state tomography and quantum simulation) has been run 2048 times to build relevant statistics.

Implementation of the hybrid algorithm in the presence of a noisy simulation was accomplished considering four different kinds of noise models. The standard approach to describe an error in quantum information science is to model the error event as a transformation of the qubit (or qubits when it is the case) state [41]. These general transformations are usually named quantum channels. The bit-flip channel depicts a situation where for each operation we carry on the qubit register (i.e., one- or two-qubit gate) there is a probability  $p$  for the state of the qubit(s) to be flipped, otherwise (i.e.,

with probability  $1 - p$ ) it is left unchanged. Similarly, the depolarizing channel refers to a transformation of the qubit(s) state to the maximally mixed state with probability  $p$ , else the identity applies. These types of error have been combined to obtain the models summarized in Table III.

## IV. NUMERICAL RESULTS

In this section we provide the results of the numerical applications of the hybrid algorithm (and corresponding classical benchmark) to the cyanidin molecule, see Fig. 2. Cyanidin is an interesting system where the excitation from the ground state to the first excited state  $|1\rangle$  has been proved of fundamental importance for its role in the charge separation process in natural dye-based solar cells [60]. We applied the optimal control procedure to a cyanidin molecule, considering both a manifold of 11 electronic states (the ground state plus the first 10 excited states) and a smaller subset consisting in the ground state and the first two excited states. For sake of brevity, in the following we will refer to the former as system **Cyan11** and to the latter as system **Cyan3**. Showing the results for the system **Cyan3** will allow one not only to emphasize which dynamical features are lost when considering a too small subset of electronic states for the optimal problem, but also to establish a clearer connection with the results of the last section. In fact, to conclude the results section, we provide a performance analysis of the quantum simulation algorithm, referring to the particular case of system **Cyan3**, on IBM Q devices.

### A. Benchmarking the quantum simulation circuit

In Fig. 3 we compare the time evolution (upper panel) of the first three states of cyanidin (for **Cyan11**) under the influence of an ultrashort laser pulse (bottom panel) obtained with different methods. In particular we consider the

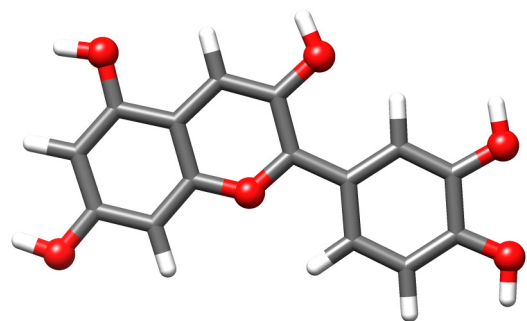


FIG. 2. Optimized structure of the cyanidin molecule. Color code: (gray) red = O, dark gray = C, white = H. Oxygen atoms are highlighted with a ball representation.

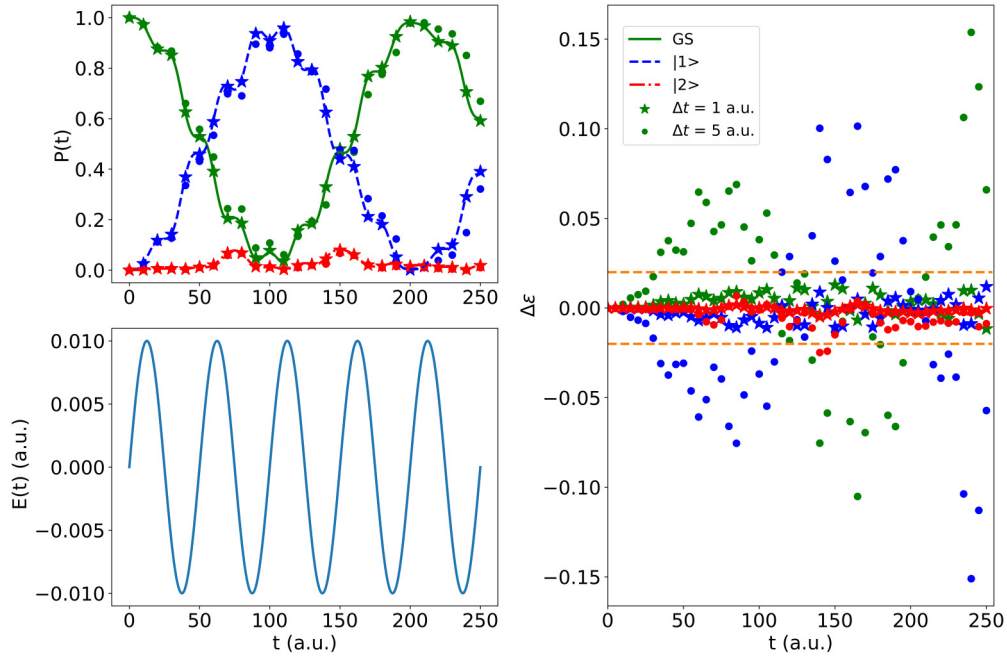


FIG. 3. Upper panel: time evolution of the ground state (green solid line) and first two excited states ( $|1\rangle$ , blue dashed line;  $|2\rangle$ , red dashed-dotted line) of cyanidin. Lines represent the evolution computed with Euler first-order propagation ( $\Delta t = 10^{-2}$  a.u.). Stars ( $\Delta t = 1$  a.u.) and dots ( $\Delta t = 5$  a.u.) are computed simulating the quantum circuit representing the time evolution operator (see Table I); stars are reported every five steps. Bottom panel: laser pulse driving the evolution of the electronic states. Only one component is shown. (Right panel) Absolute deviation  $\Delta\epsilon$  between the classical benchmark algorithm (Euler first-order) and the simulated quantum circuits implementing the TS-PCA algorithm with different time steps. Dashed lines at  $\Delta\epsilon = 0.02$  pose a threshold value for the minimum acceptable accuracy.

dynamics computed integrating the TDSE with a classical algorithm (first-order Euler), represented in figure by the solid line, and the evolution obtained adopting a simulated quantum circuit derived in the previous section based on the TS-PCA algorithm; see Sec. II. As a first consideration we may notice that the TS-PCA algorithm is able to reproduce the dynamics driven by the laser pulse; moreover it may be worth noticing the extreme robustness of the quantum algorithm with respect to the time step choice: our benchmark solution has been obtained with a time step  $\Delta t = 0.01$  a.u. ( $\approx 2 \times 10^{-4}$  fs). On the other hand, the evolution computed simulating the quantum circuit (stars in the figure) represents the result obtained with a time step equal to 1 a.u. while the dots are obtained setting  $\Delta t = 5$  a.u.. In both cases the simulated quantum algorithm recovers the qualitative information of the dynamics. To assess quantitatively the accuracy of the numerical solution we show in the right panel the absolute deviation  $\Delta\epsilon$  of the TS-PCA propagation from the reference values of the classical algorithm using different time steps. A reasonable threshold for  $\Delta\epsilon$  is 0.02 (dashed lines in figure) which is never exceeded when  $\Delta t = 1$  a.u.; conversely we may notice that a higher time step, even though still useful for a qualitative inspection of the system, provides a much higher inaccuracy. For this reason all the results we will consider from now on have been obtained with a time step of 1 a.u.

### B. QOCT: Classical implementation

Figure 4 shows the results for the optimal problem applied to cyanidin molecule with a fully classical implementation.

Here our aim is to compare the efficiency of the different approaches considered in this work. For this reason we will show only the results obtained for system **Cyan11**, as similar considerations apply for the simpler case of system **Cyan3**. Values for the functional  $J$  [see Eq. (1)], target state population and field fluency are shown. As a first comment, we can see that the NM algorithm is completely outperformed by GA, BFGS, and Rabitz algorithms; in fact, we cannot appreciate a substantial improvement for any of the values reported. Actually, an additional order of magnitude of iterations are needed to achieve results of comparable quality to that of the other two algorithms. It is important to stress that this result is general, regardless of system specifics, and well established within the optimization community [49].

On the other hand the performance of the GA, even with our very simple strategy, is satisfactory and competitive with respect to the reference algorithms, as in all cases the optimization is capable of achieving an almost complete population transfer within the same number of iterations. Regarding the value of the  $J$  functional, it is important to notice that, as mentioned in Sec. II B, we have adopted a different strategy for penalizing high fluency in the reference algorithms and in the GA. Concerning BFGS and Rabitz's algorithm, we used  $\alpha(t) = 10$  a.u.; for the latter, we included no  $\alpha(t)$  in  $J$  but imposed an external maximum threshold for the absolute value of the amplitudes. Hence we see the discrepancy between Fig. 4(a), where GA shows a higher value of  $J$ , and Figs. 4(b) and 4(c), where Rabitz or BFGS results are equal or better than the ones obtained with the GA. Particularly for the field integrals values, it is clear that the

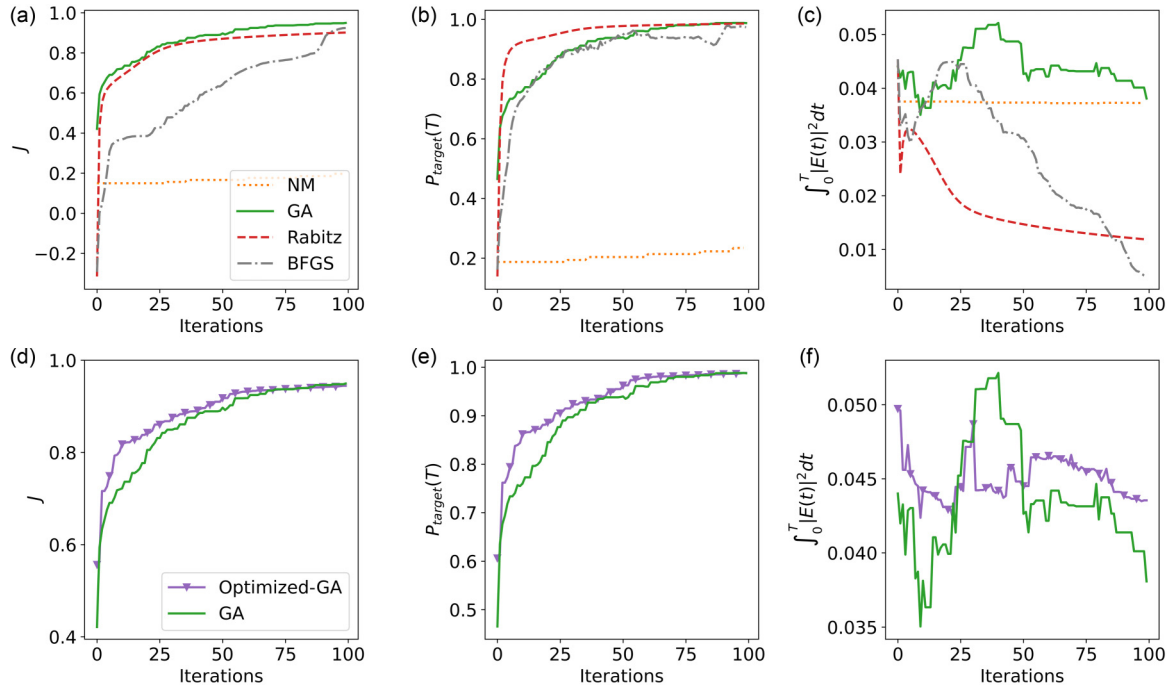


FIG. 4. Optimal control applied to cyanidin with different approaches (fully classical implementation). Plots of  $J$  (a, d), target state (11) population (b, e), and field fluency (c, f). Bottom row plots show a comparison between two different strategies for the implementation of the GA; green solid line refers to the two-step strategy for balancing convergence and exploration; purple triangles on top of the solid line refer to the less computationally demanding scheme which alternates multiple times greater or lesser exploration of the optimization landscape. All plots refer to system **Cyan11**.

best result is achieved by the BFGS algorithm. This last point should be interpreted considering that we are comparing a general purpose gradient-free algorithm (GA) with a gradient-based optimization or with an algorithm whose fundamental working equations are problem-dependent: such an accuracy could likely be achieved, within the framework of GAs, with a finer tuning of the parameters, which, again, is a problem-dependent task, or by increasing the number of iterations. An in-depth analysis of the resulting electric field is presented in Appendix C.

The bottom row plots of Fig. 4 show a comparison of the GA results obtained with two different strategies to balance convergence and exploration. The green solid lines in the upper and lower panels refer to the same results, obtained with the simpler two-step exploration and convergence procedure; purple solid line shows the results for the same control problem but with the computationally less demanding scheme which alternates several times exploration and convergence (see Sec. II B for details). This second choice will be adopted in our hybrid implementation to reduce the cost of simulating an 11-qubit device: as we can see they are almost equivalent considering both performance and final results.

### C. QOCT: Hybrid implementation using classical quantum emulator software

The aim of this section is to present the results of the hybrid implementations for both system **Cyan11** and system **Cyan3**. In particular Fig. 5 shows values for  $J$ , population of the target state and field fluency as a function of the iterations; upper row panels refer to system **Cyan3**, middle row panels

to system **Cyan11**, and bottom row panels show a comparison between the hybrid and classical implementation adopting GA as optimization strategy. Again we can see that also in the case of a simpler problem, i.e., optimal population transfer for system **Cyan3**, the NM algorithm displays a slow growth rate. Conversely GA, BFGS, and Rabitz algorithms are able to reach a complete population transfer in fewer iterations. Regarding field fluency values of Fig. 5(g) we can notice a smaller discrepancy [with respect to what shown in Fig. 5(h)] between GA and the reference algorithms that can be ascribed to the fewer control parameters [Eq. (16)] used than in the case of system **Cyan11**; see Appendix A for a discussion on the improvement of GA performances of field fluency values. It may be also worth noticing that in the case of this smaller system the BFGS and Rabitz algorithm have an equivalent performance also regarding field fluency values. Concerning the results obtained for system **Cyan11** we can notice that all the considerations made in the previous section still hold. Finally, in Fig. 5 (bottom panel), we can appreciate the equivalence (in as much as we compare two results of a nondeterministic algorithm like the genetic one) between the classical and hybrid implementation as a further proof of the reliability of the quantum simulation algorithm. An analysis of the resulting optimized fields in terms of frequencies is given in Appendix C.

### D. QOCT: Hybrid implementation simulating a noisy device

Here we discuss the results regarding the solution of the control problem for **Cyan3** in the presence of different noise models; see Fig. 6. As mentioned in Sec. III, we considered

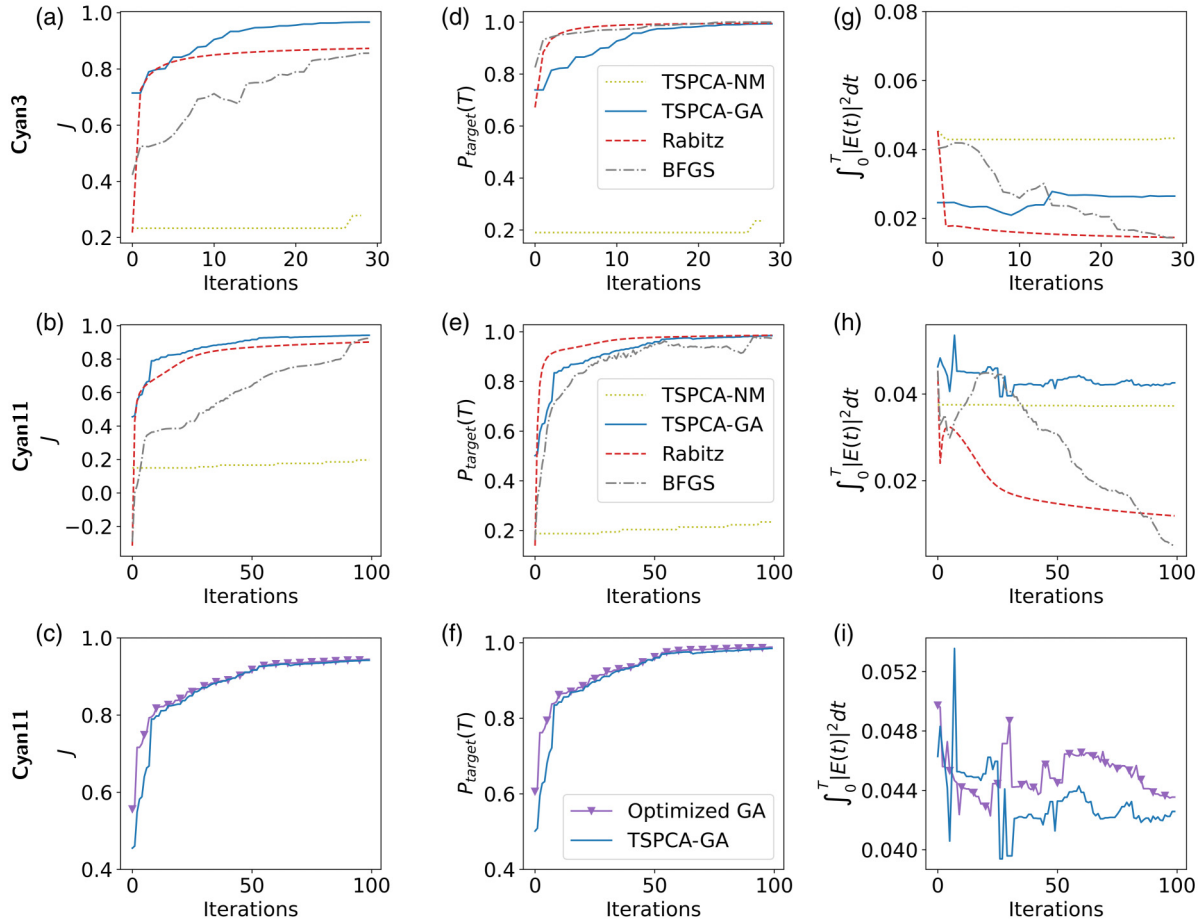


FIG. 5. Optimal control applied to cyanidin with different approaches. Plots of  $J$  (a–c), target state ( $|1\rangle$ ) population (d–f), and field fluency (g–i). Color code: purple (triangles), classical implementation coupled to the GA optimizer; gray (dashed-dotted line), classical implementation coupled to the BFGS optimizer; blue (solid line), Hybrid implementation coupled to the GA optimizer; yellow (dotted line), hybrid implementation coupled to the NM optimizer; red (dashed line), Rabitz algorithm. Upper row plots show the results for system **Cyan3**. Notice that all the results shown for the GA are obtained with the optimized scheme assessed in the previous section (see Fig. 4).

four different noise models to assess the robustness of the GA with respect to a stochastic behavior of the function due to device imperfections. Single-qubit error probability ranges

from  $1 \times 10^{-5}$  (BF Model) to  $1 \times 10^{-4}$  (mixed model) per gate execution, while two-qubit operations reach a maximum error probability of  $5 \times 10^{-5}$  (see Table III). Even though

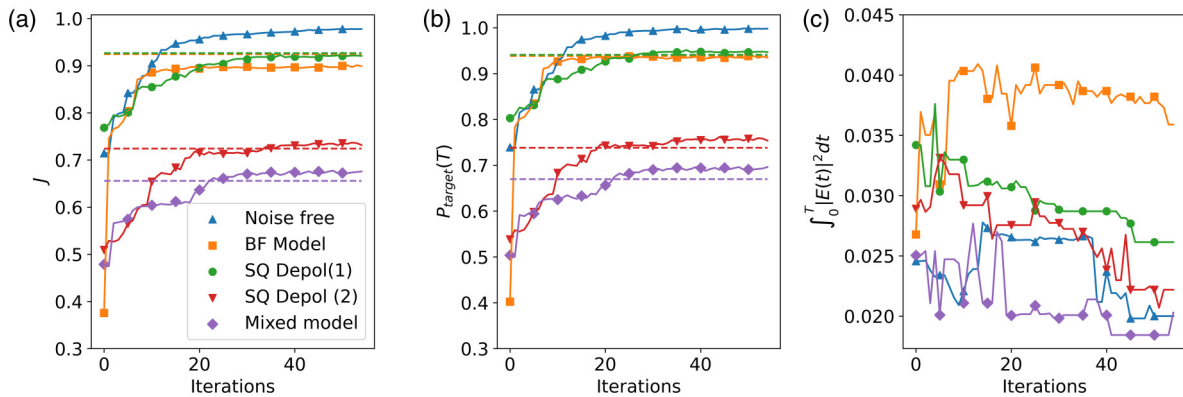


FIG. 6. Optimal control with a noisy simulation of the quantum device. Plots of  $J$  (a), target state ( $|1\rangle$ ) population (b), and field fluency (c). Different noise models (see Table III) have been used to assess the robustness of the GA with respect to different error probabilities in the function evaluation. Legend: upward blue triangles, noise free; orange squares, BF model; green dots, SQ depol. (1); red downward triangles, SQ depol. (2); purple diamonds, mixed model. Dashed lines represent a reference value obtained with the optimal solution given by the BFGS optimization in the presence of the corresponding noise model. All plots refer to system **Cyan3**.



TABLE IV. Target state population  $P_{|1\rangle}(T)$  at time  $T$  for system **Cyan3**. Evolution driven by the optimal field obtained as result of the control problem (see Fig. 6) under noise-free conditions.

	BF model	SQ depol. (1)	SQ depol. (2)	Mixed model
$P_{ 1\rangle}(T)$	0.9980	0.9968	0.9977	0.9984

our models assume much longer coherence times than the current ones (we have disregarded thermal relaxation errors), very recent contributions suggest that such models will be realistic for quantum processors available within a few years [61]. As we can see in Fig. 6 we reported values of the  $J$  functional [Fig. 6(a)], target state population [Fig. 6(b)], and field fluency [Fig. 6(c)]. As the noise increases we can see how the optimization reaches convergence at lower values of the functional. Nevertheless it is likely that, depending on the function evaluation accuracy, the maximum value of  $J$  reachable with the optimization decreases while not necessarily affecting the quality of the optimal result. To better understand this observation we reported [dashed lines in Figs. 6(a) and 6(b)], for each different noise model, the value obtained by executing the circuit which simulates the evolution driven by the optimal pulse found with the BFGS algorithm in Fig. 5. As we can see, the GA approaches  $J$  values very similar to the ones obtained in noise-free conditions with a gradient-based optimization. We may notice that in some cases (namely, red downward triangles and purple diamonds) the field obtained by performing the optimization in the presence of noise is performing slightly better than the optimal field found by the BFGS optimization in noise free conditions; we suppose that this behavior can be related to the fact that the field found with a noisy optimization drives the quantum computer wave function through states more resilient towards depolarizing or bit-flip errors. Further studies will be dedicated to address this point. On the other hand Table IV provides the complementary data for this analysis: target state population at time  $T$  obtained with the optimal field corresponding to different noisy optimizations. As we can see GA show excellent noise resiliency by providing fields that completely populate the target state. This feature could be provided by two concurrent factors: first, the algorithm is forced to retain the best solution found within a set of individuals (the functional grows regardless of the correctness of the propagation) and, second, each individual represents an independent propagation (i.e., more chances of obtaining physically meaningful information when the errors are not too big).

### E. Results on IBM Q hardware

Up to now we presented results obtained on classical hardware, disregarding the effect of any noise in the quantum propagation or considering mild effects that do not yet correspond to the current technological situation. In this section we focus on the last issue. Due to the early stage of current quantum computer development, we were able to perform noisy simulations and actual computations on the IBM Q X2 hardware [59] only for the system **Cyan3**. We have limited our analysis to this simpler system as enlarging the elec-

tronic states subspace quickly prevents from distinguishing the characteristics of the different circuits implemented; in fact, due to the complexity of the investigated dynamics, the noise associated to the number of two-qubit operations needed for a system larger than **Cyan3** would determine a complete suppression of the simulated wave function evolution. To quantify the effect of noise, we compute the fidelity between the density matrix reconstructed via quantum state tomography (QST) and a reference state obtained simulating the circuit in noise-free conditions. The choice of the hardware (among the available IBM Q devices on cloud) has been dictated by the high interqubit connectivity required by the simulated Hamiltonian; indeed, in order to minimize the number of additional gates needed to implement the circuits of Table I (which assumes an all-to-all connectivity) in the actual computers, the device with the highest number of connections has to be preferred.

As a general consideration we observe, as expected, that as the number of gate counts increases (i.e., the depth of the circuit increases) we obtain a result farther from the classical benchmark; more explicitly, as mentioned in Sec. II, the number of operations grows linearly with the propagation steps [Figs. 7(b) and 7(d)] thus determining lower fidelity values [Figs. 7(a)–7(c)]. Our aim is to present these results as example to assess the technological gap that still has to be covered to exploit in this perspective our simulation algorithm. Clearly, as the depth of the circuit depends on the time duration of the pulse  $T$ , a profitable use of a quantum hardware with this algorithm is not likely at the present time; indeed, as we can see, after 20 steps of the propagation in all cases the state of the computer reaches a value of  $\mathcal{F}$  equal to the value of thermal state that, according to Ref. [62] is  $\frac{1}{N}$  with  $N$  equal to the dimension of the Hilbert state of the quantum processor. Nevertheless these plots allow us to understand the importance of an accurate synthesis of the circuit as, exploiting the particular connection of our mapping with the single-occupancy Fock space (second circuit in Table I) we were able to obtain significant improvements. Indeed, we can see that the readaptation of the Jordan-Wigner mapping on the electronic configuration space (upper circuit in Table I) has a worse performance both in terms of fidelity and number of operations. To further improve the quality of the results we considered equivalent circuits obtained with an optimization routine (named transpilation) implemented in Qiskit, which allows one to reduce gate counts when a given circuit is mapped on to the topology of a particular hardware [63]; optimal compilation is a complex task that aims to minimize the performance loss due to the mapping on a limited connectivity hardware by exploiting gate cancellation and permutation rules to find the best virtual-to-physical qubit mapping and logical-to-native circuit mapping [64]. As shown in the bottom panels of Fig. 7 even though these improvements are significant when larger propagation steps are considered, we were not able to reach better fidelity values. We note that our reordered circuit (bottom circuit in Table I) does not get any significant improvement by this procedure. As a further assessment, in Appendix D we discuss the probability distributions obtained measuring the computational basis states population, that confirm the considerations outlined here from the fidelity plots.

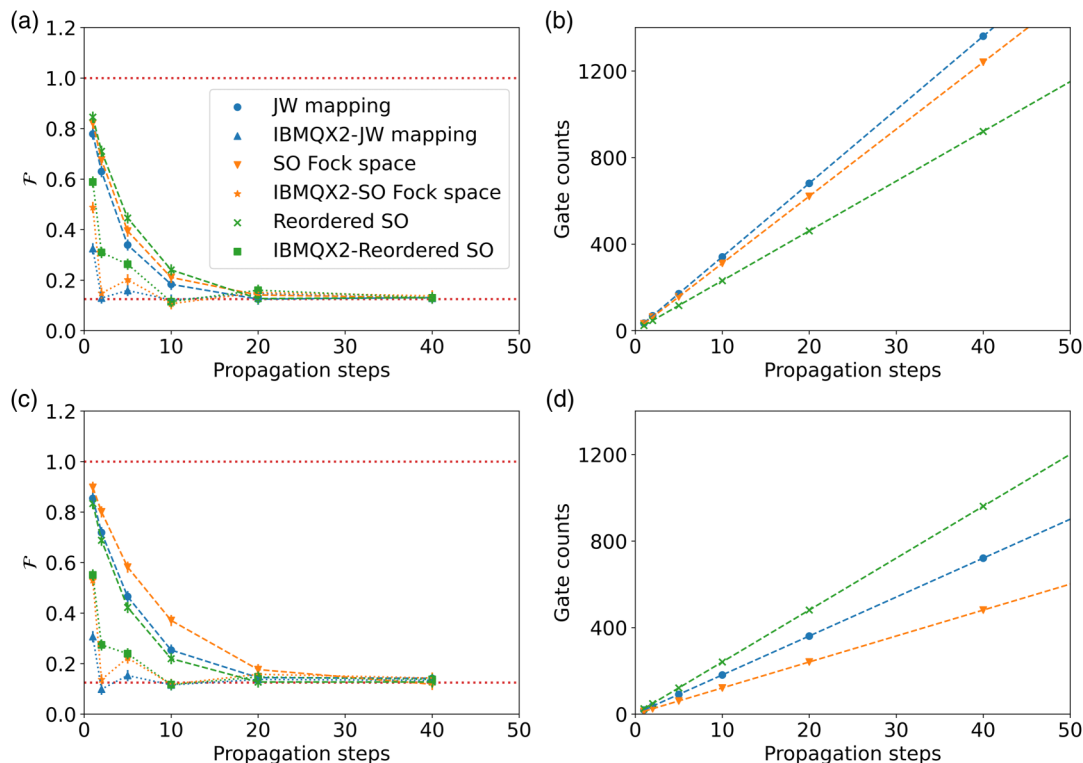


FIG. 7. (a, c) Fidelity of quantum state vs propagation steps for noisy simulations of system **Cyan3** (dashed lines) and actual experiments on IBM Q X2 (dotted lines). Panel (a) refers to the circuits in Table I, and panel (c) shows the results for the equivalent circuits after transpilation. Color code: blue (dots and upward triangles), Jordan-Wigner mapping readaptation in the  $n$ -electrons basis; orange (downward triangles and stars), single-occupancy Fock space mapping; green (crosses and squares), reordered single-occupancy Fock space mapping. (b, d) Total gate count vs propagation steps; same color code is used. Error bars represent the standard error of the quantum state tomography procedure with number of measurements  $N$  for each component of the density matrix  $N = 2048$  assuming Gaussian statistics.

## V. CONCLUSIONS

In this work we presented a hybrid approach featuring quantum and classical computation to solve optimal control problems in molecular systems. More thoroughly we adopted a quantum-based routine for the evaluation of the cost function: to accomplish this task we proposed an efficient mapping of a multideterminantal wave function into the bitstring state of a quantum processor. A possible advantage of this construction, which is not based on physical principles, is that we can choose the criterion with which to assign the correspondence between the two sets of states thus exploiting hardware-specific information to improve the quality of the result. We will seek to explore these aspects in future works. As a case study we focused about the optimal population transfer between molecular electronic states of the cyanidin molecule. We compared the results of this problem both for different algorithms and for different choices of the classical optimization routine showing that machine learning approaches (here genetic algorithms), even without a fine tuning, can be competitive with respect to the well-established and robust Rabitz's algorithm. The extension to optimization of other chemical relevant processes, such as photochemistry, may leverage on the ongoing development of excited state nuclear dynamics performed on quantum hardware [32].

Moreover we discussed the practicalities of this implementation. From this standpoint we considered two different sets

of states in which to follow the electronic dynamics, namely, a three-level system and a bigger set including up to the tenth excited state. The importance of discussing the results on a simpler system is manifold: the current fragility of quantum processors, even if equipped with a few dozen qubits, is put to the test not only by the implementation of very deep circuits, but also by maintaining coherence between qubits very far apart. Following the same idea in the last part of this work we have reported some tests with simple noise models and on IBM Q hardware to assess quantitatively the technological gap which has to be covered in order to exploit this quantum algorithm for many-electrons wave function evolution.

Efficient leverage of quantum properties of matter is an extremely challenging task calling for a common effort of the scientific community. For its part computer science, sustained by the increasing computer power, has provided chemists and physicists with the tools of machine learning which enable us to extract information and find hidden patterns in data in an unprecedented way. With the advent of first quantum computers the development of quantum simulation algorithms must be explored to get the best out of both sides.

## ACKNOWLEDGMENTS

D.C. thanks Prof. Ida-Marie Høyvik for useful discussions on cyanidin and the related quantum chemistry calculations.

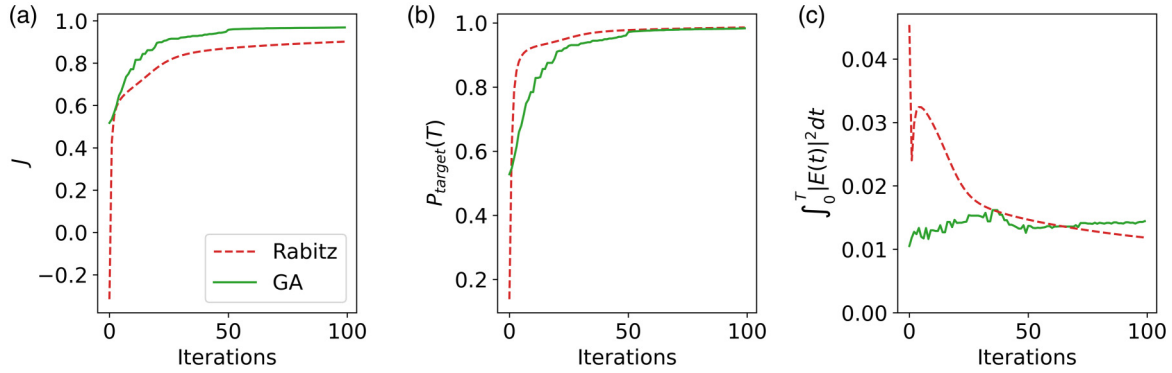


FIG. 8. Optimal control applied to system **Cyan11**. Plots of (a)  $J$ , (b) target state ( $|1\rangle$ ) population, and (c) field fluency. Red dashed line represents the reference result obtained with Rabitz's algorithm; green solid line shows the result obtained with the GA with a lower bound on amplitude values.

We acknowledge use of IBM Q Experience for this work. We also acknowledge funding from the ERC under the grant ERC-CoG-681285 TAME-Plasmons.

#### APPENDIX A

The optimization algorithms presented so far adopt different strategies to control the field fluency, which we want to keep as small as possible: Rabitz, BFGS, and NM algorithms include in  $J$  a scaling factor  $\alpha$ , which controls both the shape and the fluency of the field. The equation for  $J$  is the same in the case of the genetic algorithm, nevertheless the algorithm itself enforces a limit to the value of the amplitude during their random generation and the mutation procedure. As a consequence, the scaling factor in  $J$  can still control the shape of the field, but is quite useless in controlling the fluency. In Fig. 4 we decided to show results obtained with  $\alpha = 1$  for the genetic algorithm and  $\alpha = 10$  for Rabitz, with the purpose of not giving the false impression that the genetic algorithm was performing better than the Rabitz one. In Fig. 8 we show again the same comparison where the limit on the absolute value on the amplitudes in the genetic algorithm calculation is 0.0025 instead than 0.005. As it is possible to see, the result for the field is much more similar to the Rabitz one, without losing accuracy on the result on populations.

#### APPENDIX B

We performed some tests to understand how the very simple GA we implemented performs depending on the complexity of the system under analysis. Our interest was to assess which factors influence positively or negatively the results obtained.

To increase the complexity of the system, we decided to modify three different factors: the number of molecular excited states included in the calculation, the density of such states, and the number of parameters to optimize with the genetic algorithm. Concerning this third point, the genetic algorithm is built to optimize the amplitudes of the field in Eq. (16), the number of amplitudes is three times the number of frequencies included in the summation, which in our calculation for cyanidin were chosen in an automatic way:

the frequencies are such that the last one is higher than the last energy state. For the calculations performed in the main text, where we included 10 excited states, we have 12 frequencies plus the one corresponding to  $i = 0$ , which means 39 amplitudes as parameters for the genetic algorithm. In the calculations performed in this section, the number of individuals is kept fixed at 70, and the calculations were done with the same parameters used in Fig. 5.

Figure 9 shows the results obtained on three different version of cyanidin system. In the first we included in the calculation 20 excited states instead than 10. Because of the automatic way in which the frequencies are chosen, this corresponds to 15 frequencies, i.e., 45 amplitudes to be optimized instead than 39. Then we scaled the 20 states in such a way to have the same energy range than in the 10-state calculation. In particular, we kept fixed the energies of the ground state and of the first excited state, while we scaled the others in such a way that the 20th excited state has the same energy of the 10th. This allows us to check what happens when the energy levels are closer the ones to the other and possibly other paths open up to go from the ground state to the target state. In this case the energy range is the same that the one for the 10-state calculations, and the amplitudes to be optimized are again 39. Finally, to further increase the number of parameters to optimize, we performed an optimization calculation on the same 10-state cyanidin system but asking for 24 Fourier frequencies instead than the standard 12, as usual plus the one corresponding to  $i = 0$ . This means this time there are 75 parameters for the algorithm to optimize. In Fig. 9 the cyan plot shows the results obtained on cyanidin including in the calculation 20 excited states instead than 10. Comparing the results obtained with the ones in Fig. 5, we can see that we have a decrease in the performance, with the value of the target state population after the 100 iteration, which is 0.955 instead than 0.989. For this reason we tried to improve this result adding 50 more convergence iterations, and we were able to obtain a final population of 0.972, with similar results on the values of  $J$  and field fluency. As expected, more iterations, possibly focusing on both convergence and exploration, are able to solve the problem of a greater complexity.

We then compared this result on the 20-state system with the other two variations of the cyanidin system: in the case of

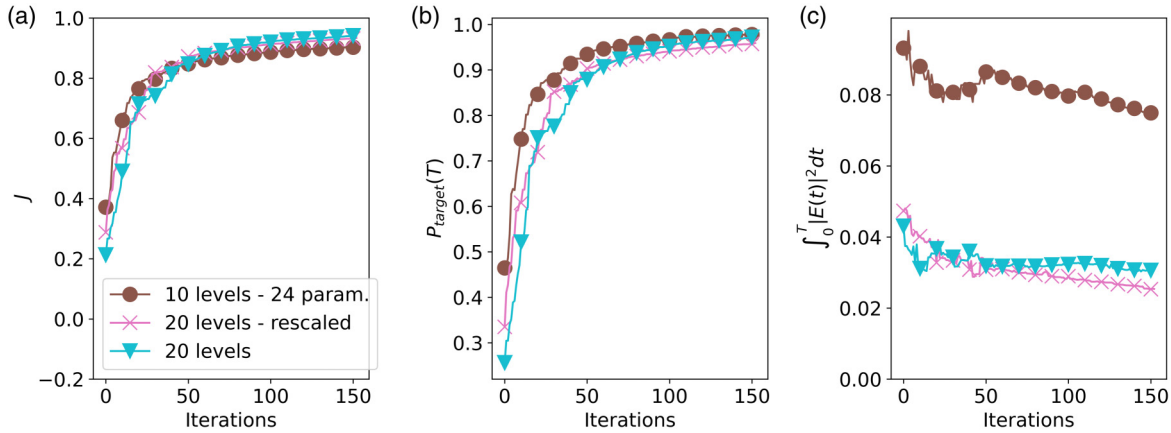


FIG. 9. Optimal control problem solution for the cyanidin molecule. Plots of (a)  $J$ , (b) target state ( $|1\rangle$ ) population, and (c) field fluency. Color code: cyanidin (20 levels), cyan triangles; cyanidin (20 levels with scaled energy spectrum), pink crosses; cyanidin (10 levels with 24 optimization parameters), brown dots.

20 scaled states (solid pink line), we have a slightly worst result that in the not scaled case, even if the difference can easily be a consequence of performance fluctuations in different runs of genetic algorithm. The third case (solid brown line), the one with 25 frequencies for 10 excited states, performs similarly to the others.

Finally, as our results has been all obtained on cyanidin, we tested our method on  $\beta$ -carotene, a molecule with a linear structure very different from the cyanidin aromatic one. We included in the description 20 energy states, which result in an energy range similar to the one of cyanidin. Figure 10 shows how the results obtained with the genetic algorithm are very similar in term of performances with respect to the one obtained on the 20-state cyanidin system.

### APPENDIX C

Here we present an analysis of the optimal control pulse obtained with the approach described in the main text. In Fig. 11 we report the optimal result, i.e., the laser pulse after the optimization and its Fourier transform, obtained with the hybrid implementation using a GA optimizer for the control problem referring to system **Cyan11** [upper row, Figs. 11(a) and 11(b)] and to system **Cyan3** (bottom row, Figs. 11(d)

and 11(e)]; later we also will discuss Figs. 11(c)–11(f) reporting the evolution of the field frequency distribution as a function of the iterations for the  $x$  component of the field. We reported in both cases only the plots for the  $x$  and  $y$  components because, due to the planarity of our system, the transition dipole moment along the  $z$  direction is almost null. As we already mentioned in describing the purpose of optimal control, the optimal field not only realizes the direct transition from the ground to the target state, but must also discourage subsequent and competitive transitions to other excited states [43,57]. This means that, even if the direct transition from the ground state to the target state corresponds to  $\omega = 0.126$  a.u., other frequencies are present in the optimal field. Looking at the plots, it is evident how increasing the complexity of the system leads to an increase in the complexity of the optimal field. **Cyan11** system has a larger number of electronic states with respect to **Cyan3**, which also means a greater number of possible path to reach the optimal solution. Indeed, we can see that the frequency distributions reveal some contributions which are even more intense than the one corresponding to the direct transition from the ground state to the target state ( $\omega = 0.126$  a.u.), to be ascribed either to multistep paths among different states or to processes of suppression of competitive transitions. This is an established result of solving

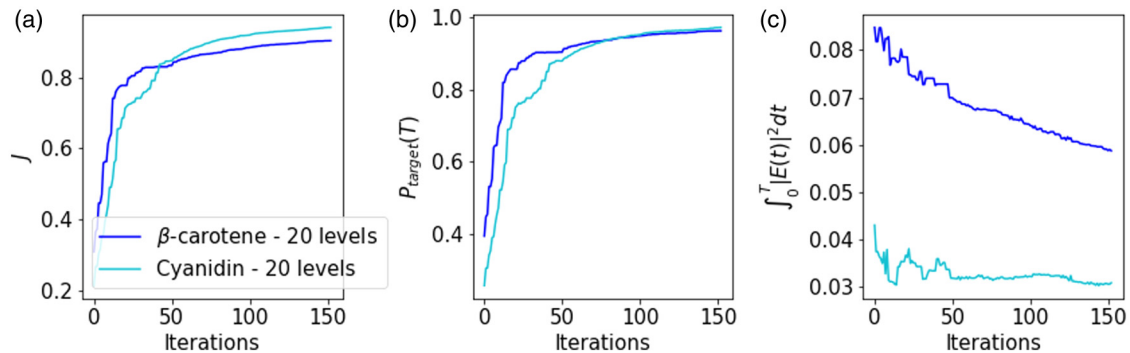


FIG. 10. Comparison of the optimal control problem solution between cyanidin and  $\beta$ -carotene. Plots of  $J$  (a), target state ( $|1\rangle$ ) population (b) and field fluency (c). Color code: cyanidin, solid cyan (light gray) line;  $\beta$ -carotene, solid dark blue (gray) line.



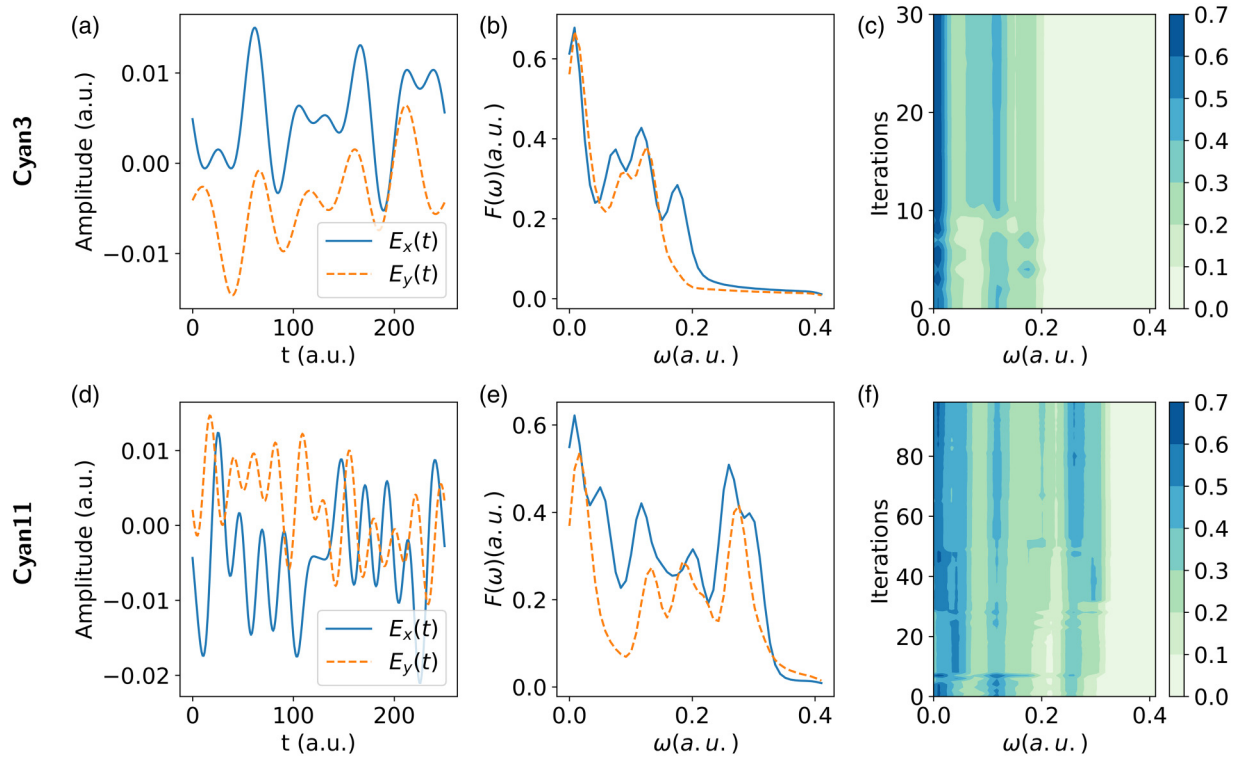


FIG. 11. Results of the optimal control problem solved with the TS-PCA propagation coupled with a GA. Upper row (a–c): optimal pulse (a), frequency distribution (b), and evolution of the frequency distribution (c) during the optimization for system **Cyan3**. Bottom row (d–f): optimal pulse (d), frequency distribution (e), and evolution of the frequency distribution (f) during the optimization for system **Cyan11**. Color maps show the frequency distributions of the  $x$  component of the fields.

optimal control problems which the genetic algorithm is able to reproduce. Nevertheless, there is also a second reason for the larger number of frequencies in the larger system which is a consequence of how the field is represented in the case of the GA optimizer: in the case of **Cyan11** the number of harmonics, and consequently amplitudes, which are summed up to describe the field is larger with respect to system **Cyan3**. In principle *useless* amplitudes, i.e., the ones which do not have a direct effect on the final value of the population in the target state, should be minimized and their value should be set to zero. In practice the GA optimizer is more efficient in identifying, and optimizing, amplitudes which have a direct effect on the result (i.e., amplitudes which favor or discourage the transition).

As a consequence, it is possible that some of the smaller amplitudes are a consequence of an incomplete minimization. Nevertheless, the second term in the  $J$  functional (i.e., the integral accounting for the filed fluency) works exactly in the way of minimizing those *useless* amplitudes, which means that if needed, a finer tuning of the optimization algorithm would lead to a further improved result. Now we discuss Figs. 11(c)–11(f), which show the frequency distribution of the fields as a function of the iterations for direction  $x$ . With regard to **Cyan3** system [Fig. 11(c)] we can see that all along the optimization the main contribution to the field is given by a low-lying frequency component, which, as we are dealing with a very simple system, we can easily assign to the frequency resonant with the transition between the first two

excited states; in this case it is easy to rationalize the role of this frequency component as it allows us to bring back to the target state all the population which has been mistransferred to the second excited state. On the other hand, in the case of system **Cyan11**, the heat map of Fig. 11(f) allows us to appreciate even more the main information that we have extracted from the frequency distribution of the optimal result: to achieve high controllability within a manifold of electronic states in ultrashort timescales, alternative paths including transitions are important as much as the contribution due to the frequency resonant with the transition of interest. This general statement is in agreement with other analysis carried out in literature [43,57]. An in-depth analysis of the population dynamics and the fluxes among excited states is beyond the purpose of this work and will be carried out in a future study.

#### APPENDIX D

Here we show additional results for the quantum simulation performed with the mapping discussed in Sec. II A. Probability distributions obtained measuring the computational basis states population (Fig. 12) confirm the considerations outlined discussing the results of the fidelity plots (main text). We can appreciate how the presence of noise drives the evolution of the quantum computer out of the subspace of the bitstring states with only one qubit in state  $|1\rangle$  until it reaches thermalization with all the states equally populated.

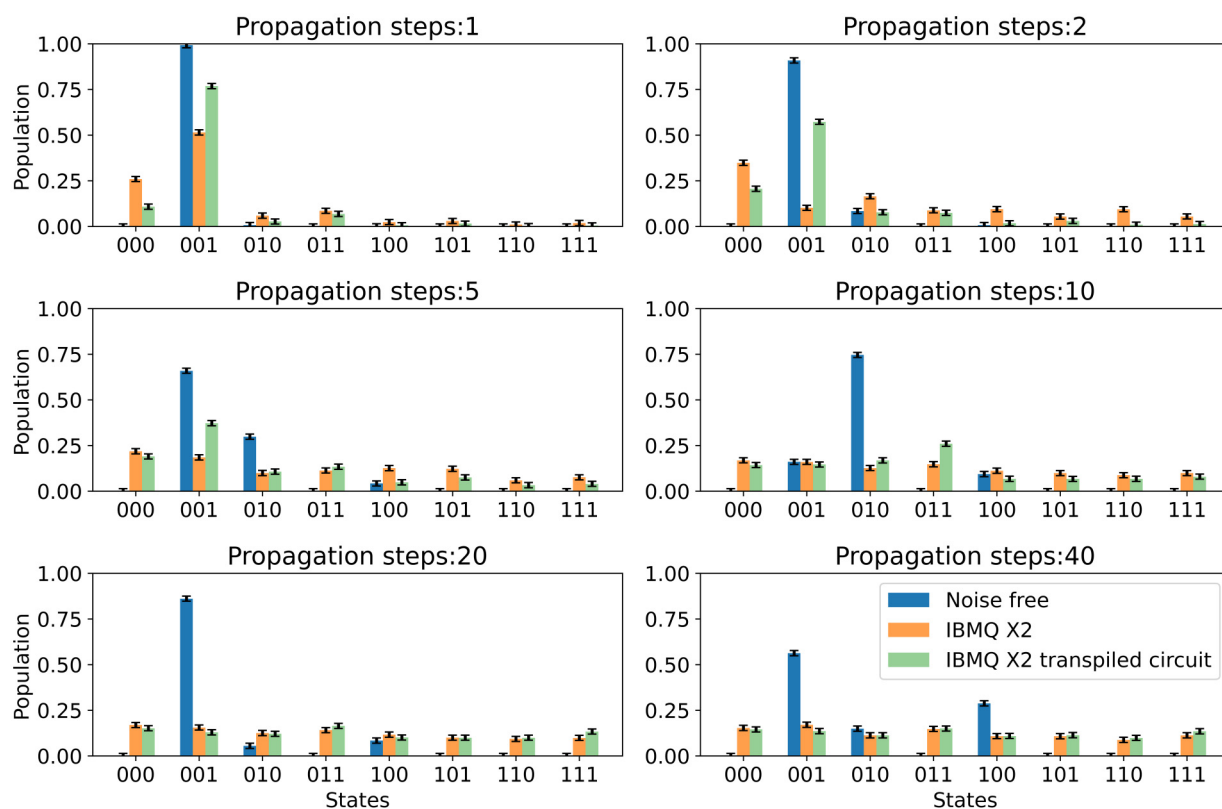


FIG. 12. Quantum simulation of system **Cyan3**. Experimental probability distributions obtained for the quantum simulation circuit (second row of Table I) on the IBM Q X2 hardware, light orange (light gray) or pale green (pale gray) with a transpiled circuit. Each panel shows the results obtained performing different propagation steps, in blue (gray) the reference noise-free distribution. We recall that at  $t = 0$  the molecule is in the ground state  $|001\rangle$ . Error bars represent the standard error computed assuming Gaussian statistics for 2048 repetitions.

- [1] R. P. Feynman, *Int. J. Theor. Phys* **21**, 467 (1982).
- [2] D. Deutsch, *Proc. R. Soc. Ser. A* **400**, 97 (1985).
- [3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, *Nature (London)* **574**, 505 (2019).
- [4] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, *Nature (London)* **536**, 63 (2016).
- [5] J. M. Gambetta, J. M. Chow, and M. Steffen, *npj Quantum Inf.* **3**, 2 (2017).
- [6] M. A. Castellanos, A. Dodin, and A. P. Willard, *Phys. Chem. Chem. Phys.* **22**, 3048 (2020).
- [7] A. Blais, S. M. Girvin, and W. D. Oliver, *Nat. Phys.* **16**, 247 (2020).
- [8] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, *Science* **288**, 824 (2000).
- [9] C. Brif, R. Chakrabarti, and H. Rabitz, *New J. Phys.* **12**, 075008 (2010).
- [10] W. Zhu, J. Botina, and H. Rabitz, *J. Chem. Phys.* **108**, 1953 (1998).
- [11] J. L. Herek, W. Wohlleben, R. J. Cogdell, D. Zeidler, and M. Motzkus, *Nature (London)* **417**, 533 (2002).
- [12] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, *npj Quantum Inf.* **6**, 6 (2020).
- [13] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, *Phys. Rev. Lett.* **109**, 240502 (2012).
- [14] L. Zhu, V. Kleiman, X. Li, S. P. Lu, K. Trentelman, and R. J. Gordon, *Science* **270**, 77 (1995).
- [15] R. S. Judson and H. Rabitz, *Phys. Rev. Lett.* **68**, 1500 (1992).
- [16] N. Accanto, P. M. De Roque, M. Galvan-Sosa, S. Christodoulou, I. Moreels, and N. F. Van Hulst, *Light Sci. Appl.* **6**, e16239 (2017).
- [17] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya *et al.*, *Chem. Rev.* **119**, 10856 (2019).
- [18] J. Li, X. Yang, X. Peng, and C.-P. Sun, *Phys. Rev. Lett.* **118**, 150503 (2017).
- [19] A. B. Magann, M. D. Grace, H. A. Rabitz, and M. Sarovar, *arXiv:2002.12497* (2020).
- [20] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *J. Magn. Reson.* **172**, 296 (2005).
- [21] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *Nat. Commun.* **5**, 4213 (2014).
- [22] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *New J. Phys.* **18**, 023023 (2016).
- [23] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, *Phys. Rev. A* **95**, 020501(R) (2017).
- [24] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, *Nature (London)* **549**, 242 (2017).

- [25] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, *Phys. Rev. X* **8**, 011021 (2018).
- [26] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, *Quantum Sci. Technol.* **3**, 030503 (2018).
- [27] X. Chu and S.-I. Chu, *Phys. Rev. A* **64**, 021403(R) (2001).
- [28] E. Zahedinejad, S. Schirmer, and B. C. Sanders, *Phys. Rev. A* **90**, 032310 (2014).
- [29] I. M. Georgescu, S. Ashhab, and F. Nori, *Rev. Mod. Phys.* **86**, 153 (2014).
- [30] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, New York, 2014).
- [31] S. McArdle, A. Mayorov, X. Shan, S. Benjamin, and X. Yuan, *Chem. Sci.* **10**, 5725 (2019).
- [32] P. J. Ollitrault, G. Mazzola, and I. Tavernelli, *Phys. Rev. Lett.* **125**, 260511 (2020).
- [33] N. P. D. Sawaya, T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik, and G. G. Guerreschi, *npj Quantum Inf.* **6**, 49 (2020).
- [34] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, *Phys. Rev. A* **65**, 042323 (2002).
- [35] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, *Mol. Phys.* **109**, 735 (2011).
- [36] P. Jordan and E. P. Wigner, *Z. Phys.* **47**, 631 (1928).
- [37] S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
- [38] S. Lloyd, *Science* **273**, 1073 (1996).
- [39] N. Wiebe, D. Berry, P. Høyer, and B. C. Sanders, *J. Phys. A: Math. Theor.* **43**, 065203 (2010).
- [40] M. Sipser, *Introduction to the Theory of Computation* (Cengage Learning, Boston, 2012).
- [41] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
- [42] E. Ovrum and M. Hjorth-Jensen, [arXiv:0705.1928](https://arxiv.org/abs/0705.1928) (2007).
- [43] M. Rosa, G. Gil, S. Corni, and R. Cammi, *J. Chem. Phys.* **151**, 194109 (2019).
- [44] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [45] J. A. Nelder and R. Mead, *Comput. J.* **7**, 308 (1965).
- [46] T. Caneva, T. Calarco, and S. Montangero, *Phys. Rev. A* **84**, 022326 (2011).
- [47] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, *Phys. Rev. A* **92**, 062343 (2015).
- [48] J. Nocedal and S. Wright, *Numerical Optimization* (Springer Science & Business Media, New York, 2006).
- [49] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, *Phys. Rev. Lett.* **120**, 150401 (2018).
- [50] H. Schwefel, *Evolution and Optimum Seeking* (Wiley, New York, 1995).
- [51] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley Professional, Reading, MA, 1989).
- [52] D. Zeidler, S. Frey, K.-L. Kompa, and M. Motzkus, *Phys. Rev. A* **64**, 023420 (2001).
- [53] <https://github.com/martarosa/OCpy/releases/tag/release-quantum-0.1>.
- [54] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus *et al.*, *Comput. Phys. Commun.* **181**, 1477 (2010).
- [55] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.
- [56] A. Dreuw and M. Head-Gordon, *Chem. Rev.* **105**, 4009 (2005).
- [57] T. Klamroth, *J. Chem. Phys.* **124**, 144310 (2006).
- [58] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. Cabrera-Hernandez, J. Carballo-Franquis, A. Chen, C. Chen *et al.*, Qiskit: An Open-Source Framework for Quantum Computing (2019).
- [59] IBM Q team 5-qubit backend: IBM Q 5 Yorktown backend specification V1.1.0, <https://quantum-computing.ibm.com> (2018).
- [60] P. Ekanayake, M. R. R. Kooh, N. Kumara, A. Lim, M. I. Petra, N. Y. Voo, and C. M. Lim, *Chem. Phys. Lett.* **585**, 121 (2013).
- [61] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Günlük, T. Itoko, N. Kanazawa *et al.*, [arXiv:2008.08571](https://arxiv.org/abs/2008.08571) (2020).
- [62] K. Życzkowski and H.-J. Sommers, *Phys. Rev. A* **71**, 032313 (2005).
- [63] G. Li, Y. Ding, and Y. Xie, in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems* (Association for Computing Machinery, New York, 2020), p. 1031.
- [64] P. Murali, J. M. Baker, A. Javadi-Abhari, F. T. Chong, and M. Martonosi, in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems* (Association for Computing Machinery, New York, 2019), p. 1015.