

This is the accepted manuscript made available via CHORUS. The article has been published as:

Pareto-front shape in multiobservable quantum control

Qiuyang Sun, Re-Bing Wu, and Herschel Rabitz

Phys. Rev. A **95**, 032319 — Published 17 March 2017

DOI: [10.1103/PhysRevA.95.032319](https://doi.org/10.1103/PhysRevA.95.032319)

The Pareto front shape in multi-observable quantum control

Qiuyang Sun,¹ Re-Bing Wu,² and Herschel Rabitz¹

¹*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*

²*Department of Automation, Tsinghua University and Center for Quantum Information Science and Technology, TNLi, Beijing, 100084, China*

(Dated: February 27, 2017)

Many scenarios in the sciences and engineering require simultaneous optimization of multiple objective functions, which are usually conflicting or competing. In such problems the Pareto front, where none of the individual objectives can be further improved without degrading some others, shows the tradeoff relations between the competing objectives. This paper analyzes the Pareto front shape for the problem of quantum multi-observable control, i.e., optimizing the expectation values of multiple observables in the same quantum system. Analytic and numerical results demonstrate that with two commuting observables the Pareto front is a convex polygon consisting of flat segments only, while with non-commuting observables the Pareto front includes convexly curved segments. We also assess the capability of a weighted-sum method to continuously capture the points along the Pareto front. Illustrative examples with realistic physical conditions are presented, including NMR control experiments on a ^1H - ^{13}C two-spin system with two commuting or non-commuting observables.

PACS numbers: 42.50.Dv, 02.30.Yy

I. INTRODUCTION

Quantum control aims at manipulating the dynamics of quantum systems across a wide range of physical, chemical and biological applications, usually via the implementation of shaped electromagnetic fields [1]. The objective value J , characterizing the degree that the goal of a specific application is achieved by a control, is a function of the control resources. Theoretical analysis of the objective as a function of the control, also known as the *control landscape*, has revealed many interesting properties of single-objective quantum optimal control including optimization of pure state transitions, mixed state preparation and unitary transformation creation [2, 3]. However, in many circumstances we need to consider more than

one criterion simultaneously for assessing the suitability of a control [4–6]. An important example is in the discrimination of similar quantum systems, known as optimal dynamic discrimination (ODD), with applications ranging from biological molecules [7] to homonuclear spins [8], by means of their different dynamics when interacting with a suitably shaped pulse. Multiple ODD experiments are typically performed to aid in the discrimination of the species involved to determine their concentration [9]. Another important case of multi-objective control is time minimization while driving quantum dynamics towards some physical objective [10–12], usually for the purpose of reducing decoherence losses in quantum computation. The tradeoff between the competitive objectives of maximizing a quantum gate fidelity and minimizing the control time has been studied in a number of contexts [13–18]. Finally, in many circumstances competing objectives can arise due to limited control resources (e.g. the pulse structure) with an impact on the level of competitive control [19].

A wide variety of practical problems in many domains inherently involve the simultaneous optimization of multiple objectives depending on the same set of variables [20]. In the situation of conflicting objectives, generally there does not exist a single solution that simultaneously reaches the global optima of all individual objectives. A solution is referred to as *Pareto optimal*, or *nondominated* [21], if none of the objectives can be improved in value without degrading some of the other objective values. Originally introduced in economics, the concept of Pareto optimality has also been applied to wide ranging physical and engineering problems, such as the design of materials [22, 23] or sensor networks [24]. The set of Pareto optimal solutions defines the *Pareto front*, which separates the feasible and infeasible regions in the space of objective values [19].

Many efforts have been made to develop universal algorithms for determining the Pareto front [25–28]. Regarding quantum control applications, experimental identification of the Pareto front has been achieved with stochastic multi-objective optimization algorithms, such as the nondominated sorted genetic algorithm (NSGA-II) [21, 26] and the multi-objective covariance matrix adaptation (MO-CMA) evolutionary algorithm [29, 30], through a process of random mutation, ranking, and selection. These strategies are generally less effective at locating the Pareto front than they are at locating an optimal solution to single objective control problems [31]; locating the Pareto front calls for identifying a family of controls specifying the front, thereby increasing the complexity of the task. In numerical simulations, a variety of deterministic algorithms have been designed for approaching and tracking the Pareto front

[17, 19, 31, 32]. These methods are based on the *scalarization* of a multi-objective problem, i.e., converting the problem to one of single-objective optimization, which can be achieved by some common strategies including (i) assigning different weights to the individual objectives and combining them into a single objective [31, 32], and (ii) selecting one objective as primary while treating the others as constraints [19, 33, 34].

Besides the cases described above, another important multi-objective quantum control problem concerns optimizing the expectation values of multiple observable operators O_m 's in the *same* quantum ensemble, which is known as *multi-observable* control and can be viewed as a generalization of the state preparation control landscape [2] regarding a single observable only. This circumstance may have significance in various scenarios such as selective excitation of multiple vibrational modes while suppressing others [35], or simultaneous optimization of NMR signals corresponding to different nuclei, as discussed in *Example 1* of Sec. II B. Theoretical analysis of this problem was reported in Refs. [31, 32], and various gradient-based deterministic algorithms for identifying and tracking the Pareto front were constructed therein. Although the range of each individual objective J_m can be easily calculated from the initial density matrix ρ_0 and the single observable O_m [2], the conflicting nature of the observables can make the feasible region smaller than the direct product of individual objective value ranges in most cases. The main goal of this paper is to analyze the shape of the feasible region, whose boundary forms the Pareto front. The results provide insights into the tradeoff relations between the individual objectives associated with different observables, and can serve as a guide for designing Pareto control experiments.

The remainder of the paper is organized as follows. Sec. II proves that for a set of mutually commuting observables the feasible region in the objective value space must fit inside a *convex polytope*, and a method is provided for determining its vertices. The analysis is illustrated by three examples with distinct physical circumstances. In Sec. III we introduce a numerical method for identifying the Pareto front by optimizing a weighted-sum objective function, and explain why the method is incapable of continuously capturing the points in the Pareto front in the case of commuting observables. Sec. IV provides analytic and numerical results of the Pareto front for two non-commuting observables in specific examples, revealing the distinction between situations of commuting and non-commuting observables in terms of the Pareto front shape. Nuclear magnetic resonance (NMR) control experiments performed on a two-spin system are reported in Sec. V to demonstrate the Pareto concepts and methods described in the paper, and a

brief conclusion is given in Sec. VI.

II. PARETO FRONT SHAPE WITH MUTUALLY COMMUTING OBSERVABLES

A. Theoretical Analysis

In a large body of quantum optimal control problems, a N -level closed quantum ensemble system is manipulated by externally implemented control fields to optimize the expectation value of an observable O at the target time T . The objective function J can be defined as [2]

$$J = \text{Tr}(\rho_T O) = \text{Tr}(U \rho_0 U^\dagger O). \quad (1)$$

The final density matrix ρ_T of the system is related to the initial state ρ_0 by $\rho_T = U \rho_0 U^\dagger$, with U being a unitary propagator generated by the Hamiltonian involving the controls. Given ρ_0 and O , the objective J can be simply treated as a function of U in what is referred to as a *kinematic* analysis. For simplicity we will assume in the following that the quantum system is fully *controllable*, i.e., any unitary transformation U at time T can be produced by some admissible control fields [36, 37]. Theoretical analysis shows that this assumption is almost surely satisfied by randomly chosen Hamiltonians [38].

As a natural generalization of (1) we seek to simultaneously optimize the expectation values of M distinct observables, $\{O_m\}_{m=1}^M$, in the same quantum system which is assumed to be controllable as stated above. Define a multi-objective optimization problem with a vector-valued objective function \mathbf{J} including the M scalar functions J_m as

$$\mathbf{J} = (J_1, \dots, J_M)^\top, \quad J_m = \text{Tr}(U \rho_0 U^\dagger O_m). \quad (2)$$

When the observables are conflicting, no unitary propagator U can optimize all of the J_m 's simultaneously. The vector space for \mathbf{J} is separated into the feasible and infeasible regions by the Pareto front. This paper will focus on the feasible region shape in two situations: either the set of observables are mutually commuting or not. The main analytic result of the former situation is presented as follows.

Theorem 1. In the multi-observable control problem of Eq. (2), if the observables $\{O_m\}_{m=1}^M$ are mutually commuting, then the feasible region of the objective function \mathbf{J} fits inside a *convex polytope* and includes all of its vertices and edges.

Proof. Due to the mutual commutativity of the observables, they can be expressed as diagonal matrices in the same appropriate basis $\{|j\rangle\}_{j=1}^N$. We can further set the initial density matrix ρ_0 also to be diagonal in this basis without loss of generality, guaranteed by the assumed controllability. Let the diagonal entries of ρ_0 and O_m be $\rho_0^{(j)}$ and $O_m^{(j)}$, $j = 1, \dots, N$, respectively.

The diagonal entries of $\rho_T = U\rho_0U^\dagger$ with an arbitrary unitary matrix U are given by

$$\rho_T^{(j)} = \sum_{k=1}^N |U_{jk}|^2 \rho_0^{(k)}. \quad (3)$$

where U_{jk} is the (j, k) -th entry of U . Define a $N \times N$ matrix P by $P_{jk} := |U_{jk}|^2 \geq 0$, which can be physically interpreted as the transition probability from the basis state $|k\rangle$ to $|j\rangle$ under the unitary transformation U . We denote the vectors of diagonal entries of ρ_0 and ρ_T as \mathbf{x}_0 and \mathbf{x}_T , e.g., $\mathbf{x}_0 = (\rho_0^{(1)}, \dots, \rho_0^{(N)})^\top$. Note that for a given ρ_0 , \mathbf{x}_T is a linear function of P that $\mathbf{x}_T = P\mathbf{x}_0$. The matrix variable P is at least subject to $2N$ linear equality constraints resulting from the unitarity of U such that each row and column of P sums to unity, i.e.,

$$\begin{aligned} \sum_{j=1}^N P_{jk} &= 1, \quad \forall k = 1, \dots, N, \\ \sum_{k=1}^N P_{jk} &= 1, \quad \forall j = 1, \dots, N, \end{aligned} \quad (4)$$

as well as N^2 linear inequality constraints:

$$P_{jk} \geq 0, \quad \forall j, k = 1, \dots, N. \quad (5)$$

The set of all matrices satisfying (4) and (5) forms a convex polytope in \mathbb{R}^{N^2} , known as the *Birkhoff polytope* B_N , which has $N!$ vertices given by the permutation matrices [39]. Note that the set of P matrices that can be physically realized by some unitary U is only a subset of B_N ; nevertheless, all of its vertices are included.

Since \mathbf{x}_T is a linear mapping of P , the feasible region of \mathbf{x}_T fits inside a convex polytope in \mathbb{R}^N whose vertices are the permutations of \mathbf{x}_0 . Each edge of this N -polytope connects two vertices that differ by swapping two adjacent values in the sorted sequence of \mathbf{x}_0 [40], and all points on the edge can be realized by some U . To illustrate this, consider the example of $\mathbf{x}_0 = (0.5, 0.3, 0.2)^\top$: The edge connecting two adjacent vertices $\mathbf{x}_T^* = (0.3, 0.5, 0.2)^\top$ and $\mathbf{x}_T^{**} = (0.2, 0.5, 0.3)^\top$, parametrized as $\lambda\mathbf{x}_T^* + (1 - \lambda)\mathbf{x}_T^{**}$

($0 < \lambda < 1$), can be realized by

$$U = \begin{pmatrix} 0 & \sqrt{\lambda} & \sqrt{1-\lambda} \\ 1 & 0 & 0 \\ 0 & -\sqrt{1-\lambda} & \sqrt{\lambda} \end{pmatrix}.$$

Likewise, unitary matrices U corresponding to all the other edges also exist. Therefore, the feasible region of \mathbf{x}_T includes all the vertices and edges of the N -polytope.

To analyze the feasible region of the vector objective function \mathbf{J} in (2), we encapsulate the eigenvalues of the M commuting observables O_m into a single $M \times N$ matrix \mathbf{O} by $\mathbf{O}_{mj} = O_m^{(j)}$. Thus, \mathbf{J} can be expressed as a function of the diagonal entries of $U\rho_0 U^\dagger$, i.e.,

$$\mathbf{J}(U) = \mathbf{O}\mathbf{x}_T(U). \quad (6)$$

Since $\mathbf{J}(U)$ is a linear mapping of $\mathbf{x}_T(U)$, we know that the feasible region of the multi-observable problem in (2) also fits inside a polytope and includes its vertices and edges. *Q.E.D.*

In two special cases we can draw stronger conclusions about the structure of the feasible region and the Pareto front:

Theorem 2. In the problem of Eq. (2) with only two commuting observables, the Pareto front has the shape of a convex polygon.

This conclusion holds because the Pareto front for two objectives is at most one dimensional, which corresponds to the polytope edges in Theorem 1. Note that Theorem 2 does not imply that interior of the polygon has to be filled up by the feasible region.

Theorem 3. In the problem of Eq. (2) with ρ_0 being a pure state, the feasible region has the shape of a convex polytope (including its interior).

Proof. The pure initial state can be expressed as $\rho_0 = |\psi_0\rangle\langle\psi_0|$, where $|\psi_0\rangle$ is a state vector. Due to the controllability, there exists some U that steers $|\psi_0\rangle$ to an arbitrary state vector $|\psi\rangle = \sum_{j=1}^N c_j |j\rangle$, where $\sum_{j=1}^N |c_j|^2 = 1$. Thus, the diagonal entries of $\rho_T = U\rho_0 U^\dagger = |\psi\rangle\langle\psi|$ are given by $\rho_T^{(j)} = |c_j|^2$, and only need to satisfy the constraints that $\sum_j \rho_T^{(j)} = 1$ and $\rho_T^{(j)} \geq 0$. The feasible region of the vector \mathbf{x}_T introduced in the proof of Theorem 1 now reduces to a polytope with N vertices (whose interior is filled up), and the feasible region of $\mathbf{J} = \mathbf{O}\mathbf{x}_T$ is a polytope as well. *Q.E.D.*

According to the analysis above, the individual objective values J_m at the vertices of the feasible region

of \mathbf{J} must satisfy the general form

$$J_m = \text{Tr}(\Pi \rho_0 \Pi^\dagger O_m), \quad m = 1, \dots, M, \quad (7)$$

with Π being a permutation matrix. At the vertices, the final density matrix $\rho_T = \Pi \rho_0 \Pi^\dagger$, which is still diagonal, commutes with each observable O_m and thus corresponds to critical points of each individual objective J_m simultaneously [2]. Note that a point satisfying the form of Eq. (7) does not necessarily turn out to be a vertex; it could also lie inside the polytope region or on the boundary between two vertices.

B. Examples for commuting observables

The consequences of our analysis will be illustrated by the three examples below. The theoretically predicted feasible regions in all examples were confirmed by the following numerical tests. Large numbers of random unitary matrices U were generated by $U = e^{iA}$, where the random Hermitian matrix A was constructed from N^2 real variables, N of them arising from A_{ii} ($1 \leq i \leq N$), and $N(N-1)/2$ from $\Re(A_{ij})$ and $\Im(A_{ij})$ ($1 \leq i < j \leq N$) each. The N^2 variables were independently drawn from the identical uniform distribution in the interval $[0, 2\pi]$. The objective value vector $\mathbf{J}(U)$ was then calculated with the random U and gave a feasible sampling point. All the points were found to fall within the theoretically predicted polytope regions.

Example 1. Consider a coupled two-spin system I - S with I and S denoting two discernible spins-1/2. We use the operators I_a and S_a ($a = x, y, z$) to represent the angular momenta of each spin in the x, y, z directions, defined as $I_a := (\sigma_a/2) \otimes \mathbb{I}_2$ and $S_a := \mathbb{I}_2 \otimes (\sigma_a/2)$, where σ_a is a Pauli matrix and \mathbb{I}_2 is the 2×2 identity matrix. In conventional NMR experiments, the thermal equilibrium state of the system can be well approximated by $\rho_0 = I_z + rS_z$ (i.e., the traceless portion of the density matrix), with $0 < |r| < 1$ corresponding to the different gyromagnetic ratio of spins I and S [41]. For the purpose of simultaneously optimizing the signal intensities of both spins in their corresponding NMR spectra in a single measurement, we can consider the dual-observable optimization of $J_m(U) = \text{Tr}(U \rho_0 U^\dagger O_m)$, $m = 1, 2$, with $O_1 = I_x$ and $O_2 = S_x$. By finding all the $4! = 24$ permutations of the nondegenerate ρ_0 (The complexity for n -spin systems grows rapidly scaling as $2^n!$), we can calculate the simultaneous critical points of J_1 and J_2 with Eq. (7) and thus construct the Pareto front, as plotted in Fig. 1 with

$r = 0.25$. The Pareto front of this problem is an octagon with eight vertices at $(J_1 = \pm 1, J_2 = \pm r)$ and $(J_1 = \pm r, J_2 = \pm 1)$, and the other critical points fall inside or on the boundary of the octagon. The Pareto optimal points between two adjacent vertices are non-critical for at least one of J_1 and J_2 , representing situations that the two objectives may be improved individually, but not simultaneously. When J_1 reaches its global optima at ± 1 , the other objective J_2 can only take values within the range $[-|r|, |r|]$, and *vice versa*. Further improvement of J_2 can be achieved via coherence transfer [41], but the value of J_1 will inevitably deviate from its optimum at the same time; e.g., the maximization of J_1 and J_2 obeys a linear tradeoff relation that $J_1 + J_2 \leq 1 + |r|$. Experimental verification of the analysis in this example is given in Sec. V.

Example 2. The orbitals of the electron in a hydrogen atom can be described by three quantum numbers (ignoring electron spin), the principal quantum number $n = 1, 2, \dots$, the angular momentum quantum number $l = 0, 1, \dots, n-1$, and the magnetic quantum number $m_l = -l, \dots, l$. The orthonormal basis set $|n, l, m_l\rangle$ contains eigenstates of two commuting observables: (i) the orbital angular momentum squared, L^2 , and (ii) the angular momentum in z direction, L_z . In units where $\hbar = 1$, we have

$$L^2|n, l, m_l\rangle = l(l+1)|n, l, m_l\rangle, \quad L_z|n, l, m_l\rangle = m_l|n, l, m_l\rangle. \quad (8)$$

Consider dual-observable optimization of $O_1 = L^2$ and $O_2 = L_z$ with a pure initial state $\rho_0 = |1, 0, 0\rangle\langle 1, 0, 0|$, where $|1, 0, 0\rangle$ is the atomic orbital with $n = 1$ and $l = m_l = 0$. In this infinite dimensional Hilbert space the expectation values of both observables are unbounded from above. We assume that the quantum bound states of the electron are fully controllable, which requires adequate polarization freedom of the light fields interacting with the atom. In the objective space, the feasible region has vertices at $(J_1 = l(l+1), J_2 = \pm l)$, i.e., $(0, 0)$, $(2, \pm 1)$, $(6, \pm 2)$, $(12, \pm 3)$, \dots , as depicted in Fig. 2. By contrast, if we treat the electron as a classical particle, the boundary of the feasible region should simply be the parabola $\langle L^2 \rangle = \langle L_z \rangle^2$, which is also shown in the figure. As $l \rightarrow \infty$ which requires the electron to have a high quantum number n (almost ionized), the slope of the quantum Pareto front will become closer to its classical counterpart. This example shows that quantization can break the smoothness of a Pareto front for multi-observable control, a phenomenon rooted in the fundamental distinction between quantum and classical descriptions of the world.

Example 3. In a four-level system, we define a three-observable optimization problem by $\rho_0 = |1\rangle\langle 1|$ and $O_m = |m\rangle\langle m|$, $m = 1, 2, 3$. The objectives J_1 , J_2 and J_3 represent the populations of levels $|1\rangle$, $|2\rangle$ and $|3\rangle$ at the final state ρ_T , respectively. The feasible region in the objective space (J_1, J_2, J_3) is a tetrahedron given by the inequalities

$$J_1 + J_2 + J_3 \leq 1, \quad J_1, J_2, J_3 \geq 0, \quad (9)$$

since the populations of each level must be non-negative. However, if we decrease the system dimensionality by removing the level $|4\rangle$ while leaving ρ_0 and O_m 's unchanged, the feasible region will possess three vertices only and become a two-dimensional triangle given by

$$J_1 + J_2 + J_3 = 1, \quad J_1, J_2, J_3 \geq 0. \quad (10)$$

Since a linear combination of the observables, $O_1 + O_2 + O_3$, equals the identity matrix in this specific case, the individual objectives are related by an extra equality that $J_1 + J_2 + J_3 = \text{Tr}(\rho_T) = 1$, resulting in a two-dimensional feasible region in the three-dimensional objective space (see Fig. 3). From this example we see that for the cases of mutually commuting observables, the dimensionality of the feasible region is upper bounded by a quantity we call “effective rank” of the observable set $\{O_m\}_{m=1}^M$, defined as $(\text{Rank}\{O_1, \dots, O_M, \mathbb{I}_N\} - 1)$ in a N -level system.

III. NUMERICAL METHODS FOR IDENTIFYING THE PARETO FRONT

Generally, the feasible region of a multi-objective problem can be generated numerically with various Pareto optimization algorithms [31, 32]. This section will start from the derivation of a classic weighted-sum algorithm, and then discuss its drawback when applied to the case of mutually commuting observables as described in Sec. II.

For the multi-observable optimization problem in Eq. (2) we introduce \mathbf{u} to represent a generic control. The vector or function \mathbf{u} is defined in a linear space, which generates the unitary propagator U transforming the initial state ρ_0 to ρ_T and gives the objective values J_m . For example, \mathbf{u} could be a function of time defined on the space $L^2[0, T]$, serving as the control field. Denote the gradient of an individual objective J_m with respect to \mathbf{u} by ∇J_m , then the variation of J_m caused by $\delta\mathbf{u}$ is

$$\delta J_m = \nabla J_m \cdot \delta\mathbf{u}, \quad m = 1, \dots, M. \quad (11)$$

Suppose \mathbf{u}^* is a Pareto optimal control in a smooth segment of the Pareto front. Let the *normal unit vector* to the Pareto front at the point \mathbf{u}^* , pointing from the feasible to the infeasible region, be $\mathbf{k} = (k_1, \dots, k_M)^\top$. By the definition of a Pareto front, any allowed first-order variation of the objective value vector, $\delta \mathbf{J} = (\delta J_1, \dots, \delta J_M)^\top$, at \mathbf{u}^* must satisfy

$$\mathbf{k} \cdot \delta \mathbf{J} = \sum_{m=1}^M k_m \delta J_m = \left(\sum_{m=1}^M k_m \nabla J_m \right) \cdot \delta \mathbf{u} \leq 0, \quad (12)$$

otherwise the infeasible region would be encroached. In order to guarantee Eq. (12) for any $\delta \mathbf{u}$, the gradients ∇J_m 's have to satisfy

$$\sum_{m=1}^M k_m \nabla J_m = \mathbf{0}, \quad (13)$$

i.e., at the Pareto front the gradients of the individual objectives are linearly dependent. This first-order necessary condition for a Pareto optimum is well known, and it has been derived in different manners [42, 43].

The condition enables us to scalarize the vector-valued multi-objective optimization problem (2) for identifying its Pareto optima. Define a new objective function \mathcal{J} as a linear weighted sum of J_m 's with a set of weights $\mathbf{w} = (w_1, \dots, w_M)^\top$,

$$\mathcal{J} = \mathbf{w} \cdot \mathbf{J} = \sum_{m=1}^M w_m J_m, \quad (14)$$

then a Pareto optimum satisfying Eq. (13) will also have a zero gradient on \mathcal{J} if \mathbf{w} is parallel to \mathbf{k} . Therefore, a Pareto optimal point of the original problem may be located by finding the maximum of \mathcal{J} , in which the J_m 's to be maximized (minimized) should be assigned positive (negative) weights. In the framework of multi-observable control (2), \mathcal{J} can also be viewed as the expectation value of a new single observable $\mathcal{O} := \sum_{m=1}^M w_m O_m$,

$$\mathcal{J} = \sum_{m=1}^M w_m \text{Tr}(\rho_T O_m) = \text{Tr} \left[\rho_T \left(\sum_{m=1}^M w_m O_m \right) \right] = \text{Tr}(\rho_T \mathcal{O}). \quad (15)$$

Intuitively, by optimizing \mathcal{J} with continuously adjusted weights one should be capable of capturing different Pareto optimal points and thus reconstruct the entire Pareto front. However, it turns out that for mutually commuting observables, the Pareto front cannot be continuously sampled by this weighted-sum method. As shown in Sec. II, the Pareto front consists of *flat* segments. The normal vector \mathbf{k} points at a constant direction within each segment, and changes abruptly across the edges of the polytope region.

If the vector of weights \mathbf{w} for a chosen objective function \mathcal{J} is not parallel to the normal vector of any flat segment, then \mathcal{J} can only be optimized at the locations where multiple segments join together and the normal vector is undefined. As \mathbf{w} varies continuously, the gradient optimization trajectory of \mathcal{J} will converge to one of the polytope vertices for most choices of \mathbf{w} , as experimentally demonstrated in Sec. V. This behavior agrees with the finding that the weighted-sum method succeeds in obtaining points from all parts of a Pareto front only when it is *convex*, but fails to work when it is flat or concave [44, 45], as illustrated in Fig. 4.

IV. PARETO FRONT SHAPE WITH NON-COMMUTING OBSERVABLES

This section will discuss the Pareto front shape in the more general and complex cases involving non-commuting observables. The theorems in Sec. II A no longer applies because the observables cannot be simultaneously diagonalized in the same basis, and the feasible region does not have to be a polytope. No closed-form expression for the Pareto front is found except in some special problems. However, with some numerical methods, such as random sampling in the unitary space of U and the weighted-sum algorithm in Eq. (15), we can still sketch the Pareto front of a particular problem. For simplicity we deal with two non-commuting observables here, which is sufficient for the purpose of revealing the fundamental distinction of commuting and non-commuting observables in Pareto optimization.

The Pareto front shape with non-commuting observables is first illustrated within a two-level system. A two-observable optimization problem is defined as:

$$\rho_0 = (\mathbb{I}_2 + \sigma_z)/2, \quad O_1 = \sigma_x, \quad O_2(\theta) = \cos \theta \sigma_x + \sin \theta \sigma_y, \quad (16)$$

where σ_x , σ_y and σ_z are the Pauli matrices. The observables O_1 and O_2 are non-commuting except when $\sin \theta = 0$. In the circumstance of a single spin-1/2, this problem can be viewed as optimization of the angular momenta in two distinct orientations separated by an angle θ . Note that any two-observable problem in a two-level system can be converted to a form like Eq. (16) after some rescaling and frame rotation. The final state ρ_T after the control process corresponds to a point on the Bloch sphere, so

$$[\text{Tr}(\rho_T \sigma_x)]^2 + [\text{Tr}(\rho_T \sigma_y)]^2 \leq 1. \quad (17)$$

Representing $\text{Tr}(\rho_T \sigma_x)$ and $\text{Tr}(\rho_T \sigma_y)$ in terms of J_1 and J_2 (i.e., corresponding to O_1 and O_2 , respec-

tively), we obtain

$$J_1^2 - 2 \cos \theta J_1 J_2 + J_2^2 \leq \sin^2 \theta. \quad (18)$$

This inequality gives a feasible region whose boundary is an ellipse in the objective space (J_1, J_2) , the eccentricity of which depends on θ . In the extremal cases that $\sin \theta = 0$ or $\cos \theta = 0$ the ellipse converts to a line segment $(-1 \leq J_1 \leq 1, J_2 = \pm J_1)$ or a circle $(J_1^2 + J_2^2 = 1)$, respectively. The feasible region shapes at three representative values of θ , 0, $\pi/4$ and $\pi/2$, are plotted in Fig. 5, which have been verified numerically with random samplings of U . For such convexly curved segments as a Pareto front, the normal vector direction changes continuously as we move on the Pareto front. This property enables us to systematically sample the Pareto front by optimizing weighted-sum objective functions in Eq. (15) with varying weights, as opposed to the case of flat Pareto front segments discussed in Sec. III.

Generally in higher-dimensional quantum systems, the feasible region with two non-commuting observables cannot be simplified to a form like Eq. (18), and the Pareto front is not necessarily an ellipse. In all the examples we numerically tested with random ρ_0 and non-commuting observables, the Pareto fronts contained convex or flat segments only, but no concave ones. In some situations the Pareto front may even be a mixture of both flat and curved segments, and a three-level illustrative example is given as follows:

$$\rho_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad O_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad O_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (19)$$

The Pareto front of this problem, as shown in Fig. 6, was generated by random samplings followed by gradient optimization of weighted-sum functions in Eq. (15). It consists of two line segments, \overline{AB} and \overline{BC} , and an arc \widehat{AC} . Point B located at the global maximum of both J_1 and J_2 forms the only nondifferentiable point in the Pareto front, where the normal vector changes abruptly from $(0, 1)^\top$ to $(1, 0)^\top$. To understand this behavior we recall that a Pareto optimum of the multi-observable problem must also be a critical point of the weighted-sum function in Eq. (15) with some choice of \mathbf{w} . The condition is equivalent to that [2]

$$\left[\rho_T, \sum_{m=1}^M w_m O_m \right] = 0, \quad (20)$$

i.e., ρ_T commutes with a weighted sum of the single observables for some choice of w_m 's. In the two-observable case, if a Pareto optimal ρ_T commutes with both O_1 and O_2 , the coefficients w_1 and w_2 can be arbitrary. Thus, the Pareto front will have no unique normal vector at that point, possibly resulting in nondifferentiability. For commuting O_1 and O_2 there always exist such ρ_T 's that commute with both observables, which are responsible for the vertices of the feasible region polytope. For non-commuting observables the existence is not guaranteed; e.g., no Hermitian matrix commutes with both σ_x and σ_y except the identity or zero matrix. In the three-level example of Eq. (19) O_1 and O_2 are non-commuting; nevertheless, there still exists an accessible ρ_T (i.e., $\rho_T = \rho_0$) that commutes with both O_1 and O_2 , resulting in the nondifferentiable vertex B on the Pareto front.

V. EXPERIMENTAL DEMONSTRATIONS

In this section we demonstrate the theoretical analysis on the shape of Pareto front by ^1H and ^{13}C NMR experiments performed on a 800 MHz spectrometer. We use ^{13}C -labeled chloroform ($^{13}\text{CHCl}_3$) as the sample, and denote the two coupled nuclei ^1H and ^{13}C as spin I and S , respectively. Two radio frequency pulses with an equal duration of $T = 5\text{ms}$ irradiate the sample simultaneously as the control resources, whose carrier frequencies are resonant with ^1H or ^{13}C . The amplitudes and phases of the control pulses are modulated in the time domain for optimally driving the dynamics of the two-spin system in the desired manner. The detailed experimental setup is similar to that used in our previous works [41, 46, 47].

The initial density matrix of the system is taken as $\rho_0 = I_z + 0.25S_z$, which is proportional to the traceless portion of the thermal equilibrium state. Two cases of multi-observable control are considered: (i) two commuting observables, $O_1 = I_x$ and $O_2 = S_x$ (c.f., *Example 1* in Sec. II B); and (ii) two non-commuting observables, $O_1 = I_x$ and $O_2 = I_y$. The expectation values of these observables can be characterized by the integrated areas of the doublet peak ascribed to ^1H or ^{13}C in the corresponding NMR spectrum of that isotope. In case (i) the signals of the two isotopes are acquired simultaneously by the dual detectors and then processed individually, while in case (ii) the two observables, both associated with the spin angular momentum of ^1H but in two orthogonal orientations (denoted by x and y), are successively measured by ^1H NMR with identical control pulses and detector phases differing by 90° .

For a control that steers the quantum system from ρ_0 to ρ_T , the two objectives $J_m = \text{Tr}(\rho_T O_m)$, $m = 1, 2$, directly read from the NMR spectra, are given in arbitrary units. Utilizing a gradient algorithm and starting from the same initial control, we maximized the weighted-sum objective function

$$\mathcal{J}(w_1) = w_1 J_1 + (1 - w_1) J_2, \quad 0 \leq w_1 \leq 1, \quad (21)$$

along its own gradient to approach the Pareto front, and attempted to locate different points on the front by varying the weight w_1 from 0 to 1. The maximum values of J_1 and J_2 under the experimental condition were first determined by setting w_1 to 1 or 0 and then normalized to 1, as the theoretical maximum values of J_1 or J_2 . The evolution of individual objective values (J_1, J_2) in the optimization processes is displayed in Fig. 7.

The theoretical result for the Pareto front of case (i) is given in Fig. 1, which has eight vertices. The trajectories in Fig. 7(a) with a relatively small weight of J_1 ($w_1 = 0.3, 0.4$) converged toward the vertex $A(J_1 = 0.25, J_2 = 1)$, the global maximum of J_2 but not J_1 , while the trajectory with a larger weight of J_1 ($w_1 = 0.6$) converged to another vertex $B(J_1 = 1, J_2 = 0.25)$ at the global maximum of J_1 . We also verified from simulation that any trajectory with $0 < w_1 < 0.5$ or $0.5 < w_1 < 1$ should also converge to vertex A or B , respectively, since the line segment \overline{AB} has a constant normal direction of $(1, 1)^\top$. A trajectory with $w_1 = 0.5$ could stop *somewhere* between A and B , in principle, because all the points in the line segment give an identical \mathcal{J} value, which equals its global maximum. Therefore, the weighted-sum method will be incapable of continuously sampling the Pareto front with commuting observables I_x and S_x . In case (ii), the feasible region is predicted to be the disc $J_1^2 + J_2^2 \leq 1$. The five search trajectories in Fig. 7(b) converged to different points in the Pareto front, moving gradually along the arc from the maximum of J_1 to the maximum of J_2 as w_1 decreases from 1 to 0. The normal vector at each identified Pareto optimal point is roughly parallel to the vector of weights $\mathbf{w} = (w_1, 1 - w_1)^\top$ as expected. The collective experimental results are in agreement with our theoretical predictions on the Pareto front shape of multi-observable optimal control with commuting and non-commuting observables, and the effectiveness of the weighted-sum algorithm.

VI. CONCLUSIONS

This paper deals with the generic problem of quantum multi-observable control, which seeks to simultaneously optimize the expectation values of M distinct and usually conflicting observables in the same quantum ensemble with an initial state of ρ_0 , steered by a single unitary transformation U . The individual objective functions are defined by $J_m(U) = \text{Tr}(U\rho_0U^\dagger O_m)$, $m = 1, \dots, M$. We particularly focus on the tradeoff relations between these competitive objectives described by the shape of Pareto front, given by the feasible region boundary in the space of objective values. We proved that for two commuting observables, the Pareto front in the objective value space (J_1, J_2) has the shape of a convex polygon consisting of flat segments. When non-commuting observables are involved, convexly curved segments will emerge in the Pareto front while flat segments can also be present in certain cases. These conclusions on Pareto front shapes should apply to any quantum control circumstances with unitary system dynamics. When the full controllability condition is violated, the realized feasible region may shrink to a reachable subset. The results here provide a new perspective for understanding multi-objective optimization in quantum control, and reveal an interesting distinction between commuting and non-commuting observables as a feature of quantum mechanics.

We also discussed the capability of some numerical methods in Pareto front identification, especially the strategy of optimizing a scalar objective function defined as a linear weighted sum of the individual objectives. With the weighted-sum method one can continuously capture the points along a convex Pareto front segment by varying the relative weights assigned to the J_m 's, but for flat segments the optimization search will almost always converge to a vertex of the feasible region, as demonstrated in NMR control experiments performed on a ^1H - ^{13}C two-spin system.

Acknowledgments

Q.S. acknowledges support from the NSF Grant No. CHE-1464569, the Princeton Plasma Science and Technology program, and the John Templeton Foundation. H.R. acknowledges support from the DOE Grant No. DE-FG02-02ER15344. R.B.W. acknowledges support from NSFC Grants No. 61374091 and

No. 61134008. We acknowledge Dr. István Pelczer for providing technical support on NMR facilities.

- [1] Control of quantum phenomena: past, present and future, C. Brif, R. Chakrabarti, H. Rabitz, *New J. Phys.* 12, 075008 (2010)
- [2] Why do effective quantum controls appear easy to find? T.-S. Ho, H. Rabitz, *J. Photochem. Photobiol. A: Chem.* 180, 226 (2006)
- [3] Quantum control landscapes, R. Chakrabarti and H. Rabitz, *Int. Rev. Phys. Chem.* 26, 671-735 (2007)
- [4] Multi-objective optimization on alkali dimers, S. M. Weber, F. Sauer, M. Plewicky, A. Merli, L. Wöste, and A. Lindinger, *J. Mod. Opt.* 54, 2659-2666 (2007)
- [5] Multi-objective genetic algorithm optimization of 2D- and 3D-Pareto fronts for vibrational quantum processes, C. Gollub and R. de Vivie-Riedle, *New J. Phys.* 11, 013019 (2009)
- [6] Topology of classical molecular optimal control landscapes for multi-target objectives, C. Joe-Wong, T.-S. Ho, H. Rabitz, and R. Wu, *J. Chem. Phys.* 142, 154115 (2015)
- [7] Quantum control of tightly competitive product channels, M. Roth, L. Guyon, J. Roslund, V. Boutou, F. Courvoisier, J.-P. Wolf, and H. Rabitz, *Phys. Rev. Lett.* 102, 253001 (2009)
- [8] Minimum-time selective control of homonuclear spins, T.-M. Zhang, R.-B. Wu, F.-H. Zhang, T.-J. Tarn, and G.-L. Long, *IEEE Trans. Contr. Syst. Technol.* 23, 2018-2025 (2015)
- [9] Photonic reagents for concentration measurements of fluorescent proteins with overlapping spectra, A. Goun, D. I. Bondar, A. O. Er, Z. Quine, and H. Rabitz, *Sci. Rep.* 6, 25827 (2016)
- [10] Time-optimal quantum evolution, A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, *Phys. Rev. Lett.* 96, 060503 (2006)
- [11] Minimum time optimal synthesis for two level quantum systems, F. Albertini and D. D'Alessandro, *J. Math. Phys.* 56, 012106 (2015)
- [12] Quantum speed limit and optimal control of many-boson dynamics, I. Brouzos, A. I. Streltsov, A. Negretti, R. S. Said, T. Caneva, S. Montangero, and T. Calarco, *Phys. Rev. A* 92, 062110 (2015)
- [13] Optimal control-based efficient synthesis of building blocks of quantum algorithms: A perspective from network complexity towards time complexity, T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, *Phys. Rev. A* 72, 042331 (2005)
- [14] Time-optimal quantum evolution, A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, *Phys. Rev. Lett.* 96, 060503 (2006)

- [15] Time-optimal unitary operations, A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, *Phys. Rev. A* 75, 042308 (2007)
- [16] The quantum speed limit of optimal controlled phasegates for trapped neutral atoms, M. H. Goerz, T. Calarco, and C. P. Koch, *J. Phys. B: At. Mol. Opt. Phys.* 44, 154011 (2011)
- [17] Exploring the tradeoff between fidelity and time optimal control of quantum unitary transformations, K. W. M. Tibbetts, C. Brif, M. D. Grace, A. Donovan, D. L. Hocker, T.-S. Ho, R.-B. Wu, and H. Rabitz, *Phys. Rev. A* 86, 062309 (2012)
- [18] Time optimal quantum control of two-qubit systems, B. Li, Z.-H. Yu, S.-M. Fei, and X.-Q. Li-Jost, *Sci. China Phys. Mech. Astron.* 56, 2116-2121 (2013)
- [19] Exploring the capabilities of quantum optimal dynamic discrimination, V. Beltrani, P. Ghosh, and H. Rabitz, *J. Chem. Phys.* 130, 164112 (2009)
- [20] Multicriteria optimization in engineering and in the sciences, W. Stadler, ed., New York, Plenum Press (1988)
- [21] Multi-objective genetic approach for optimal control of photoinduced processes, L. Bonacina, J. Extermann, A. Rondi, V. Boutou, and J.-P. Wolf, *Phys. Rev. A* 76, 023408 (2007)
- [22] Pareto-optimal alloys, T. Bligaard, G. H. Jóhannesson, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, *Appl. Phys. Lett.* 83, 4527-4529 (2003)
- [23] Ranking the stars: a refined Pareto approach to computational materials design, K. Lejaeghere, S. Cottenier, and V. V. Speybroeck, *Phys. Rev. Lett.* 111, 075501 (2013)
- [24] Design of an FBG sensor network based on Pareto multi-objective optimization, H. Jiang, J. Chen, T. Liu, and H. Fu, *IEEE Photon. Technol. Lett.* 25, 1450-1453 (2013)
- [25] Multi-objective evolutionary algorithms: a comparative case study and the strength Pareto approach, E. Zitzler and L. Thiele, *IEEE Trans. Evol. Comput.* 3, 257-271 (1999)
- [26] A fast and elitist multi-objective genetic algorithm: NSGA-II, K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, *IEEE Trans. Evol. Comput.* 6, 182-197 (2002)
- [27] Multi-objective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II, H. Li and Q. Zhang, *IEEE Trans. Evol. Comput.* 13, 257-271 (2009)
- [28] Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems, I. Das and J. E. Dennis, *SIAM J. Optim.* 8, 631-657 (1998)
- [29] Quantum control experiments as a testbed for evolutionary multi-objective algorithms, O. M. Shir, J. Roslund, Z. Leghtas, H. Rabitz, *Genet Program Evolvable Mach* 13, 445 (2012)
- [30] Multi-objective adaptive feedback control of two-photon absorption coupled with propagation through a

- dispersive medium, F. O. Laforge, J. Roslund, O. M. Shir, and H. Rabitz, Phys. Rev. A 84, 013401 (2011)
- [31] Quantum Pareto optimal control, R. Chakrabarti, R. Wu, and H. Rabitz, Phys. Rev. A 78, 033414 (2008)
- [32] Quantum multi-observable control, R. Chakrabarti, R. Wu, and H. Rabitz, Phys. Rev. A 77, 063425 (2008)
- [33] Multiple-objective problems: Pareto-optimal solutions by method of proper equality constraints, J. G. Lin, IEEE Trans. Autom. Control 21, 641-650 (1976)
- [34] Proper inequality constraints and maximization of index vectors, J. G. Lin, J. Optim. Theory Appl. 21, 505-521 (1977)
- [35] Coherent control of multiple vibrational excitations for optimal detection, S. D. McGrane, R. J. Scharff, M. Greenfield, and D. S. Moore, New J. Phys. 11, 105047 (2009)
- [36] On the controllability of quantum-mechanical systems, G. M. Huang, T. J. Tarn and J. W. Clark, J. Math. Phys. 24, 2608-2618 (1983)
- [37] Complete controllability of quantum systems, S. G. Schirmer, H. Fu, and A. I. Solomon, Phys. Rev. A 63, 063410 (2001)
- [38] Modeling and control of quantum systems: an introduction, C. Altafini and F. Ticozzi, IEEE Trans. Automat. Contr. 57, 1898 (2012)
- [39] Tres observaciones sobre el algebra lineal [Three observations on linear algebra], G. Birkhoff, Univ. Nac. Tucumán. Revista A 5, 147-151 (1946)
- [40] Analyse algebrigue dun scrutin, G. Th. Guilbaud and P. Rosenstiehl, Mathématiques et sciences humaines 4, 9-33 (1963)
- [41] Experimental observation of saddle points over the quantum control landscape of a two-spin system, Q. Sun, I. Pelczer, G. Riviello, R.-B. Wu, and H. Rabitz, Phys. Rev. A 91, 043412 (2015)
- [42] Nonlinear Programming, H. Kuhn and A. Tucker, Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California, Berkeley, California, 481 (1951)
- [43] Generalized homotopy approach to multi-objective optimization, C. Hillermeier, J. Optim. Theory Appl. 110, 557 (2001)
- [44] Multicriteria truss optimization, J. Koski, in Ref. [20].
- [45] A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems, I. Das and J. E. Dennis, Struct. Optim. 14, 63-69 (1997)
- [46] Experimental exploration over a quantum control landscape through nuclear magnetic resonance, Q. Sun, I. Pelczer, G. Riviello, R.-B. Wu, and H. Rabitz, Phys. Rev. A 89, 033413 (2014)
- [47] Identifying and avoiding singularity-induced local traps over control landscapes of spin chain systems, Q.

Sun, I. Pelczer, G. Riviello, R.-B. Wu, and H. Rabitz, *Phys. Chem. Chem. Phys.* 17, 29714-29722 (2015)

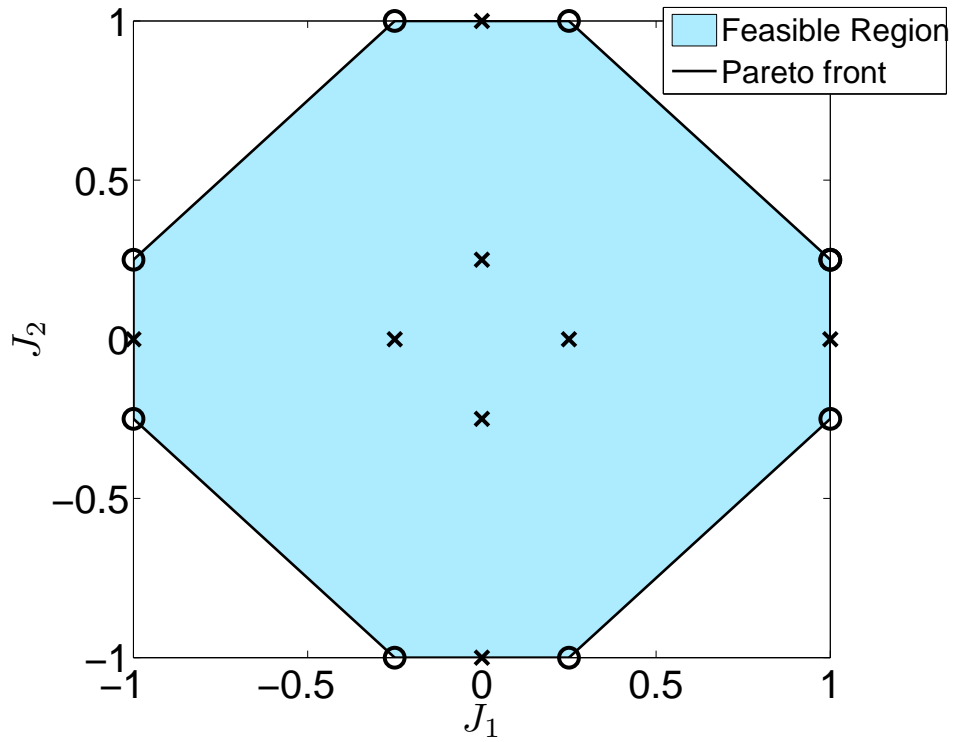


FIG. 1: (Color online) The Pareto front of *Example 1*, a dual-observable problem defined in a two-spin system I - S , which has the shape of an octagon. The initial density matrix is chosen as $\rho_0 = I_z + 0.25S_z$, and the two *commuting* observables are $O_1 = I_x$ and $O_2 = S_x$. The points marked by \circ and \times are determined by permuting the diagonal entries of ρ_0 and calculating the corresponding J_1 and J_2 values; the former points are vertices of the octagon while the latter ones are not.

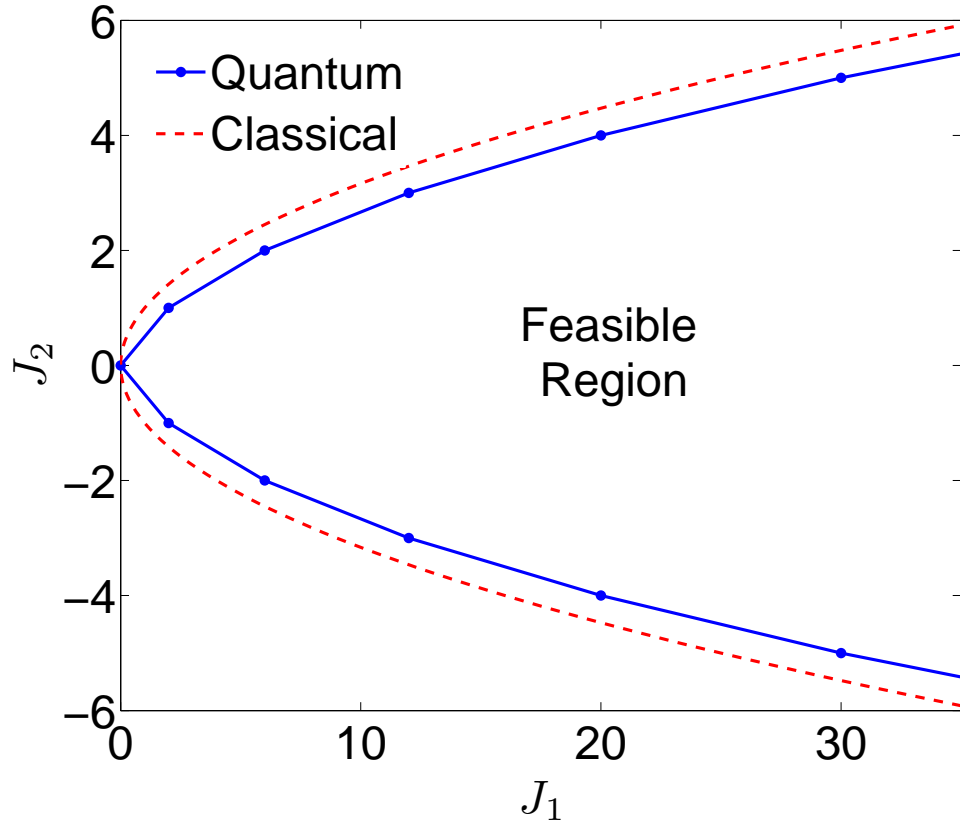


FIG. 2: (Color online) The Pareto front of *Example 2* on manipulating the orbital angular momentum of the electron in a hydrogen atom. A pure initial state $\rho_0 = |1, 0, 0\rangle\langle 1, 0, 0|$ is chosen, while the two *commuting* observables are $O_1 = L^2$ and $O_2 = L_z$. Both quantum and classical results of the Pareto front are provided to show the distinction caused by quantization.

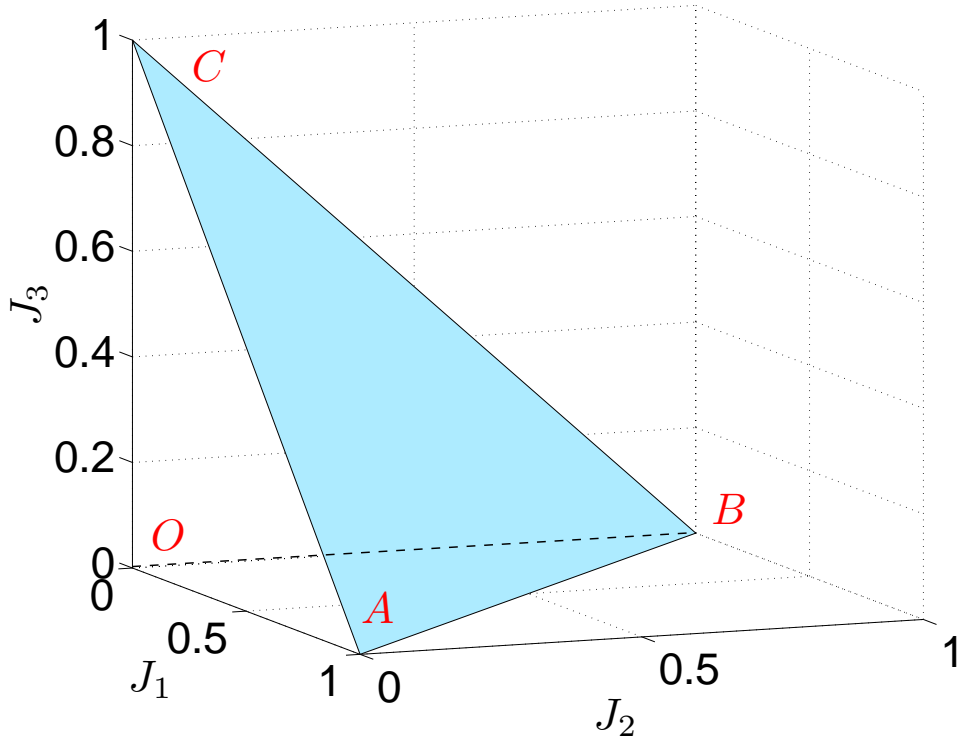


FIG. 3: (Color online) The feasible region of *Example 3*, defined by $\rho_0 = |1\rangle\langle 1|$ and three mutually *commuting* observables $O_m = |m\rangle\langle m|$, $m = 1, 2, 3$. If the system only has the three levels $|1\rangle$, $|2\rangle$ and $|3\rangle$, the feasible region will be the colored triangle ABC . Once a fourth level (or more) is added in, the feasible region will be extended to the tetrahedron $OABC$.

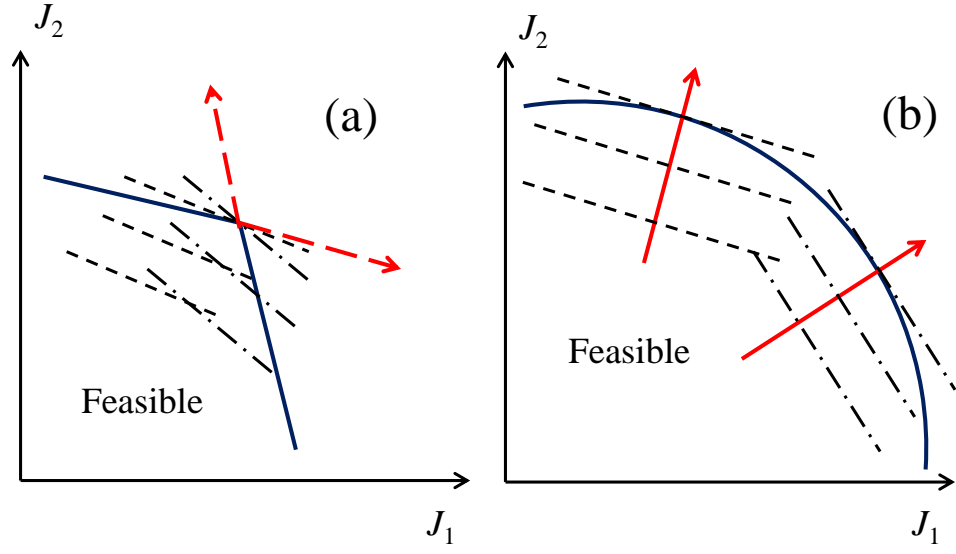


FIG. 4: (Color online) Searching for the Pareto front by optimizing a weighted-sum objective $\mathcal{J} = w_1 J_1 + w_2 J_2$. The dashed lines depict level sets with constant \mathcal{J} values at different choices of $\mathbf{w} = (w_1, w_2)$, and in optimizing \mathcal{J} along its own gradient we move across the level sets until stopping at a level set tangent (or nearly so, in practice) to the feasible region boundary. In (b) where the Pareto front (solid line) is smooth and convex, the search will stop where the normal vector (an arrow) is parallel to \mathbf{w} , while in (a) where the Pareto front consists of flat segments, the search will always stop at the vertex if \mathbf{w} is aligned with any direction between the two dashed arrows.

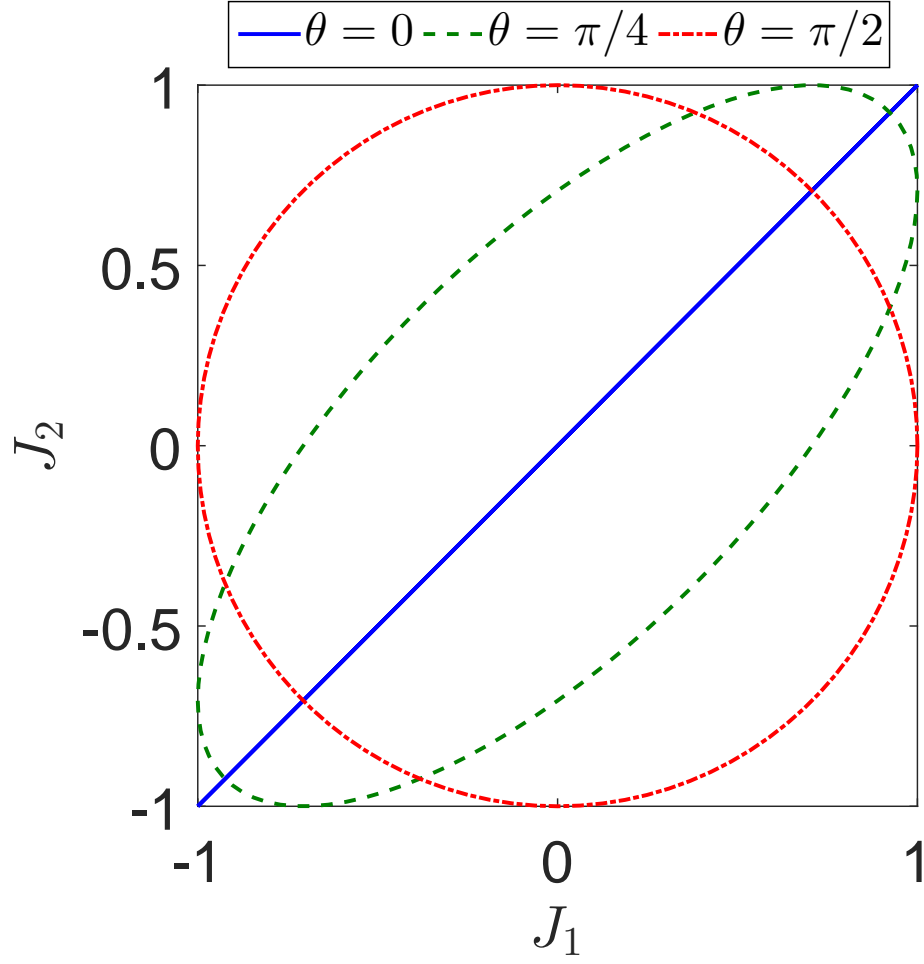


FIG. 5: (Color online) The Pareto front of the two-observable problem in (16). The observables are *non-commuting* except when $\sin \theta = 0$. As the parameter θ varies, the Pareto front may take on the extreme shapes of a circle, an ellipse or a line segment.

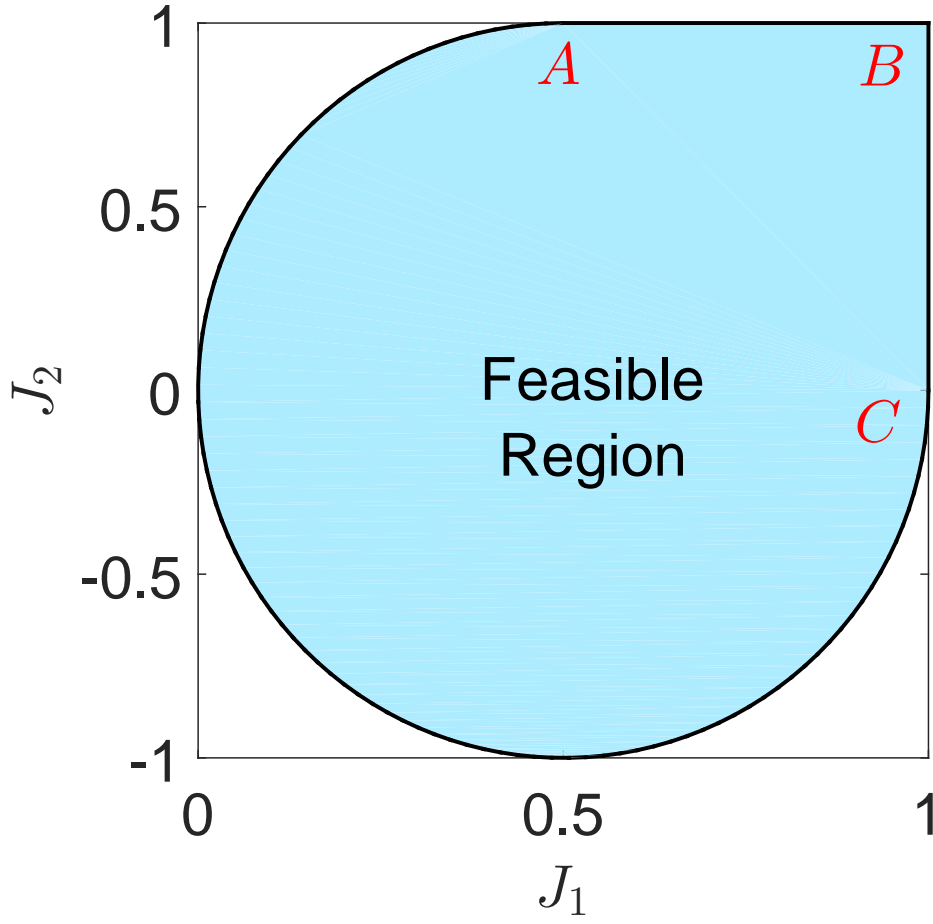


FIG. 6: (Color online) The Pareto front of the three-level example in Eq. (19) is a mixture of both flat and convexly curved segments. The nondifferentiable vertex B corresponds to critical points of both individual objectives J_1 and J_2 , and the Pareto front does not have a unique normal direction at point B .

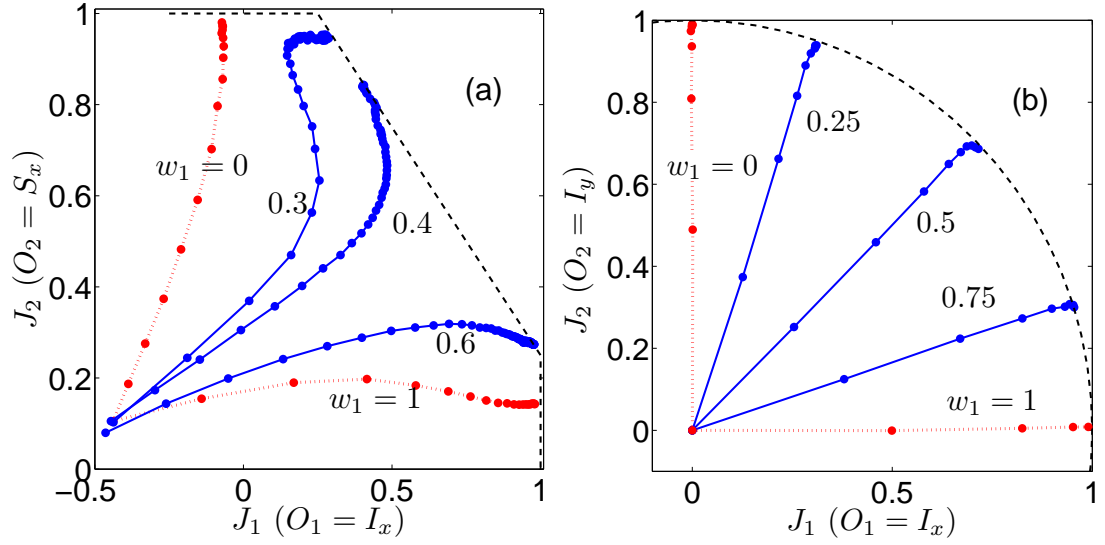


FIG. 7: (Color online) Experimental NMR illustrations of the theoretical Pareto principles. Dual-observable quantum control experiment in the two-spin system of $^{13}\text{CHCl}_3$ ($I = ^1\text{H}$, $S = ^{13}\text{C}$): the evolution of (J_1, J_2) during the optimization of a weighted-sum objective function $\mathcal{J}(w_1) = w_1 J_1 + (1 - w_1) J_2$. (a) $O_1 = I_x$ and $O_2 = S_x$ (commuting); (b) $O_1 = I_x$ and $O_2 = I_y$ (non-commuting). The dashed lines depict a portion of the theoretical Pareto fronts in the two cases. Red dotted and blue solid lines show optimization processes of an individual objective (J_1 or J_2) and combinations of both, respectively, with each point representing an experimental iteration.