Chattering Phenomenon in Quantum Optimal Control

R. Robin, U. Boscain, M. Sigalotti, D. Sugny

January 3, 2023

Abstract

We present a quantum optimal control problem which exhibits a chattering phenomenon. This is the first instance of such a process in quantum control. Using the Pontryagin Maximum Principle and a general procedure due to V. F. Borisov and M. I. Zelikin, we characterize the local optimal synthesis, which is then globalized by a suitable numerical algorithm. We illustrate the importance of detecting chattering phenomena because of their impact on the efficiency of numerical optimization procedures.

1 Introduction

Consider the experiment in which a ball bounces up and down on the ground. We assume that the impact with the ground is inelastic and that the ball is only subject to the gravity. In the ideal case in which the ball changes its speed instantaneously at each bounce, an infinite number of bounces is performed in the finite time of the process. Chattering thus refers to an observable (here the speed) having very fast discontinuities, which lead in the mathematical limit to an infinite number of jumps in a finite-time interval [1]. This type of process can also be found in quantum physics. Examples are the quantum Zeno effect and dynamical decoupling in which a repeated observation of the system and a periodic sequence of instantaneous pulses prevent, respectively, its time evolution [2] or its coupling with the environment [3]. The possibility of chattering was also established in Optimal Control Theory (OCT). OCT was founded in the sixties by Pontryagin and his co-workers [4], with a rigorous framework to design control protocols for driving a dynamical system from a given initial state into a desired target state, while minimizing energy or other resources. Chattering was found in this field by A.T. Fuller in a planar system [5] [6]. It consists of
Figure 1: Schematic description of the time scale invariance of the optimal control in the chattering process. The optimal solution of the quantum control problem described in this paper is plotted in the top panel. Near the final time $t_f$, the control switches infinitely many times with an asymptotic invariant structure by time dilation as represented on the two lower panels (successive zooms around $t = t_f$ given by the different plot colors).

A control that switches infinitely many times over a finite time period [1]. This observation runs counter to common experience for which control is viewed as a piecewise continuous function, while for chattering, the control lies in a larger class of functions [8]. In Fuller’s example, the number of switchings accumulates with a geometric progression at the final control time. The control law has a time scale invariance near the final time as schematically represented in Fig. 1. At first Fuller’s example was considered a curiosity, but this type of phenomenon is very widespread in optimal control, as rigorously shown few years later by I. Kupka [9]. While optimal trajectories exhibiting chattering have been found in relevant examples from medicine [10] to classical [1] and space dynamics [11], this phenomenon has, to the best of our knowledge, not yet been studied in quantum control [12, 13, 14, 15, 16]. The existence, the role and the ubiquity of this process in quantum systems remain an open question. Such control schemes are interesting from a fundamental point of view even if they turn out to not be feasible in experiments. They may also have a rather severe impact on the numerical search of optimal solutions [11, 17]. It is therefore important to understand why chattering is occurring and how it can be avoided [18]. We propose in this paper to describe this phenomenon in quantum control by studying a simple but fundamental quantum system. We introduce on this key example a systematic procedure to design the optimal control protocol. The impact on the efficiency of optimization procedures is also described.
2 The model system

Let us consider the control of a three-level quantum system described by a pure state belonging to a Hilbert space spanned by the states \(|1\rangle, |2\rangle, \text{and} |3\rangle\). As in a standard STIRAP process \cite{19}, in a suitable rotating frame and in the rotating wave approximation, the dynamics of the system are controlled by two pulses \(\Delta\) and \(u\) that couple, respectively, states \(|1\rangle\) and \(|2\rangle\) and states \(|2\rangle\) and \(|3\rangle\). We assume that \(\Delta\) is a constant and that the only control parameter is \(u(t)\).

The resulting dynamics are

\[
\dot{X} = (\Delta \Omega_3 + u(t) \Omega_1) X,
\]

where \(X\) is a vector of real coordinates \((x_1, x_2, x_3)\) with the condition \(x_1^2 + x_2^2 + x_3^2 = 1\) (see A). The two skew-symmetric matrices \(\Omega_1\) and \(\Omega_3\) are given by

\[
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & -1 \\
1 & 0 & 0
\end{pmatrix},
\begin{pmatrix}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}.
\]

generating respectively the rotations along the \(x_1\)- and \(x_3\)-directions. In the following, the control cannot exceed a certain physical bound, that is, \(|u(t)| \leq u_0\) for some \(u_0 > 0\). A time rescaling results in the multiplication of the two parameters \(\Delta\) and \(u_0\) of the problem by a positive scalar, which leads to the normalization \(u_0 = 1\).

Starting from any state \(X_0\) on the unit sphere, the goal of the control is to steer the system to the state \((0, 0, 1)\), still denoted by \(|3\rangle\), while minimizing the population transfer to the state \(|1\rangle\). This control protocol is interesting in practice if, e.g., the state \(|1\rangle\) is the only state of the system subject to an unwanted relaxation process. Notice that this latter is not modeled by Eq. (1).

The control protocol can be formalized as an optimal control problem by introducing the cost functional \(C = \int_{t_f}^0 x_1^2(t) dt\) to be minimized, where \(t_f\) is the control time, which is not fixed. The existence of a solution in finite time is not obvious and it is a consequence of the general theory developed in \cite{1}, as explained in \cite{D}. We give below an argument showing that, once its existence has been established, such a control has a chattering behaviour.

3 Application of the Pontryagin Maximum Principle

The existence of chattering can be proved from the Pontryagin Maximum Principle (PMP), which gives a first-order necessary condition for optimality \cite{S}. It can be stated by introducing the Pontryagin Hamiltonian

\[
H_P = P \cdot (\Delta \Omega_3 + u \Omega_1) X + p^0 x_1^2,
\]

where \(X\) is a point on the unit sphere, \(P\) the adjoint state of coordinates \((p_1, p_2, p_3)\), and \(p^0\) a constant equal either to 0 or to \(-\frac{1}{2}\). If \(X(t)\) is an optimal trajectory with corresponding control \(u(t)\), then \(X(t)\) is extremal, namely, there exist \(P(t)\) and \(p^0\) such that \(\dot{X} = \frac{\partial H_P}{\partial P}, \dot{P} = -\frac{\partial H_P}{\partial X}\), and \(H_P(X(t), P(t), u(t), p^0) = \max_{v \in [-1, 1]} H_P(X(t), P(t), v, p^0) = 0\). Moreover, \((P(t), p^0) \neq (0, 0)\) for every \(t\).
and an explicit computation based on the dynamics and on the relation \( \tilde{\Phi}(1) = \Phi(1) - (\Phi(1) \cdot X(1))X(1) \). If \( p^0 = 0 \) (resp. \( p^0 = -\frac{1}{2} \)) the extremal is called abnormal (resp. normal). The condition \( (\tilde{\Phi}(t), p^0) \neq (0, 0) \) (instead of the weaker but more classical one \( (\Phi(t), p^0) \neq (0, 0) \)) reflects the fact that we already know that \( X(t) \) belong to the two-dimensional unit sphere. Actually, \((X(t), \tilde{\Phi}(t)) \) represents an element of the cotangent bundle of the sphere.

The maximization condition of the PMP is solved with the switching function \( \Phi(t) = \bar{P}(t) \cdot \Omega_1 X(t) \). If \( \Phi(t) \) is different from zero then \( H_P \) is maximal when the control, called bang, is a constant control of maximum amplitude, \( u(t) = \text{sign}[\Phi(t)] \). If \( \Phi(t) \neq 0 \) on a time-interval \( [a, b] \), we say that the restriction of the extremal to \( [a, b] \) is a bang arc. When \( \Phi \) is constantly equal to zero on \( (a, b) \) then the restriction of the extremal to \( (a, b) \) is said to be a singular arc. Singular arcs can be characterized as follows. A simple computation shows that \( \Phi = \Delta \bar{P} \cdot \Omega_2 X \), where \( \Omega_2 \) is the generator of the rotations along the \( x_2 \)-axis. If both \( \Phi \) and \( \Phi \) vanish at time \( t \) then either \( \bar{P}(t) = 0 \) or \( \Omega_1 X(t) \) and \( \Omega_2 X(t) \) are linearly dependent, that is \( X(t) \) lies on the equator \( x_3 = 0 \). If \( \bar{P} \) is equal to zero on a time interval then \( p^0 = -\frac{1}{2} \) and, moreover, \( 0 = H_P = -\frac{a}{2} \) leading to \( x_1 = 0 \) on the same interval. Since \( \dot{x}_1 = -\Delta x_2 \), it follows that \( X \) stays at \( [3] \).

We now claim that it is enough to consider normal extremals, i.e., that any optimal trajectory corresponding to an abnormal extremal also corresponds to a normal one. Indeed, if an optimal trajectory reaches \( [3] \) in a finite time \( T \leq t_f \) then, taking \( t_f > t \), the extension \( Y \) of \( X \) on \( [0, t_f] \) that stays in \( [3] \) on the interval \([t, t_f] \) is optimal as well. Moreover, \( Y \) corresponds to a vanishing control for all times in \([T, t_f] \). Let us denote by \( Q \) the adjoint state corresponding to \( Y \), whose existence is guaranteed by the PMP. Then the extremal \( Y \) is singular on \([T, t_f] \) and \( \tilde{Q}_{[T, t_f]} \equiv 0 \), where \( \tilde{Q} \) is defined in analogy to \( \bar{P} \). Such an extremal cannot be abnormal (otherwise \( (\tilde{Q}(t_f), p^0) \) would be zero). We have shown that for any optimal solution reaching \( [3] \) in time \( T \) it can be assumed without loss of generality that the extremal is normal and \( \bar{P}_{[T, t_f]} \equiv 0 \) (this corresponds to replacing \( \bar{P} \) by \( Q_{[0, t_f]} \) and denoting the latter by \( \bar{P} \)).

The last step consists in showing that an optimal trajectory cannot be the finite concatenation of bang arcs in \([0, T] \), where we recall that \( T \) is the time at which \( X \) reaches \([3] \). This fact can be deduced by computing the so-called order of the singular trajectory and using general properties of optimal control theory (see, e.g., [10], [20], [21]). In our specific case we can exhibit an elementary and explicit proof. By contradiction, assume that \( u \) is constantly equal to \(+1\) or to \(-1\) in an interval \([T - \varepsilon, T]\) for some \( \varepsilon > 0 \). Then \( \Phi \) is smooth on \([T - \varepsilon, T]\) and an explicit computation based on the dynamics

\[
\dot{p}_1 = -\Delta p_2 + x_1, \quad \dot{p}_2 = \Delta p_1 - u p_3, \quad \dot{p}_3 = u p_2
\tag{2}
\]

and on the relation \( \tilde{\bar{P}}(T) = 0 \) shows that \( \Phi(T) = \Phi(T) = \Phi^{(2)}(T) = \Phi^{(3)}(T) = 0 \) and \( \Phi^{(4)}(T) = -u \Delta^2 \) (see [3]). Then \( \Phi(t) \) has opposite sign with respect to \( u \) in a small left neighborhood of \( T \), contradicting the fact that \( u(t) = \text{sign}[\Phi(t)] \). Hence the only option to reach \([3] \) is a chattering process in which the control switches infinitely many times in a finite time-interval. Notice that, as a byproduct, we obtain that there are no optimal abnormal extremals reaching the target. Indeed, a simple computation shows that the control corresponding to an abnormal extremal reaching the target at time \( T \) is necessarily bang on an interval \([T - \varepsilon, T]\) for \( \varepsilon > 0 \) small enough (since \( \Phi(T) \) must be zero and \( \Phi(T) \) cannot be zero at the same time).
Having established the chattering behaviour, the next goal is to find the position of the switching points, which form the switching curve. This is not an easy task and an exact analytic expression cannot be derived. Notice however that close to the state $|3\rangle$ the system can be described by the two coordinates $x_1$ and $x_2$ with the dynamics

$$\dot{x}_1 = -\Delta x_2, \quad \dot{x}_2 = \Delta x_1 - u\sqrt{1 - x_1^2 - x_2^2}. \quad (3)$$

Taking only the dominant terms, the system can be locally approximated as

$$\dot{x}_1 = -\Delta x_2, \quad \dot{x}_2 = -u, \quad (4)$$

which is different from the standard linear approximation. In control literature, system (4) together with the cost $\int_0^t x_1^2(t) \, dt$ (tf free) and the origin $x_1 = x_2 = 0$ as target state, yield the classical Fuller problem, up to the change of coordinates $x_1 \to -x_1/\Delta, \, x_2 \to x_2$. The chattering trajectories of the Fuller model can be described exactly [7] as recalled in [3]. We then deduce the following properties for system (4). The optimal solution is bang-bang with an infinite number of switchings near the origin. The switching curve is defined by $x_1 = \text{sign}[x_2]\Delta x_2^2$, where $\xi = \sqrt{\frac{\sqrt{3}-1}{24}} \simeq 0.44623$. The switching times $t_k$ are given by a geometric progression of the form $\frac{T - t_k}{T - t_{k-1}} = \frac{1}{\alpha}$ with $\alpha = \sqrt{1 + 2\xi} \simeq 4.1301$.

The relation between the local optimal syntheses of two systems differing only by high-order terms (such as (3) and (4)), when one of the two syntheses exhibits chattering, is a subtle question. To this purpose we apply the results of [1], which permit to conclude that a system has a Fuller-like optimal synthesis provided it differs from the Fuller model by terms that are small while applying suitable non-isotropic dilations. On the basis of this study described in [3], we state the following result.

**Proposition 1.** For every point $X_0$ sufficiently close to $|3\rangle$, any optimal solution of (4) connecting $X_0$ to $|3\rangle$ corresponds to a control having infinitely many discontinuities accumulating at the first time $T$ at which $|3\rangle$ is reached (and possibly staying at $|3\rangle$ for larger times). The optimal synthesis is characterized by a switching curve $\Gamma$ passing through $|3\rangle$, whose expression in coordinates $(x_1, x_2)$ is of the form

$$x_1 = \begin{cases} 
\lambda_1(x_2)x_2^2 & \text{if } x_2 > 0, \\
\lambda_0(x_2)x_2^2 & \text{if } x_2 < 0,
\end{cases}$$

where $\lambda_0$ and $\lambda_1$ are $C^1$ functions satisfying $\lambda_0(0) = -\lambda_1(0) = \Delta \xi$. In the coordinates $(x_1, x_2)$ and locally near $(0,0)$ the optimal control is $-1$ below the curve $\Gamma$ and $+1$ above it.

**4 Numerical results**

The optimal synthesis can be computed numerically starting from that of the Fuller model. When we are sufficiently close to $|3\rangle$, we approximate the switching curve of the quantum system by that of its approximation (4). We consider a
Figure 2: Optimal trajectory from $X(0) = (0, 1, 0)$ to $X(t_f) = |3\rangle = (0, 0, 1)$ on the sphere $x_1^2 + x_2^2 + x_3^2 = 1$. The switching curve is plotted in red. The parameter $\Delta$ is set to 10.

point of the latter curve of coordinates $(\xi \Delta x_{20}, x_{20})$, with $x_{20} > 0$ (notice that the same method could be used for $x_{20} < 0$). The third component is obtained from $x_{30} = \sqrt{1 - x_{10}^2 - x_{20}^2}$ and the adjoint state from the condition $P \cdot \Omega_1 \cdot X = 0$ together with $H_P = 0$ [8]. The dynamics of the PMP allows to propagate $X$ and $P$ backward in time, uniquely up to an irrelevant component of $P$ along $X$. Starting from an initial point with $x_{20} > 0$, one integrates the equations taking $u = -1$. When the corresponding switching function vanishes, the control switches from $-1$ to $+1$. Then one goes on by integrating the equations with $u = +1$ up to the next switching time and so on. The stability of this approach is justified in [E] An optimal control law can be obtained for each value of $x_{20}$. Even if the result applied above to characterize the optimal synthesis (see Theorem 1 in [D]) can be used only for initial points $X_0$ close to $|3\rangle$, numerical simulations show that optimal trajectories have the same structure everywhere in the north hemisphere.

The trajectory starting from a given initial state (say $(0, 1, 0)$) can then be determined by a Newton algorithm to estimate the right parameter $x_{20}$ from which the backward propagation arrives at the initial state. The parameter $x_{20}$ is not unique because the forward optimal trajectory has infinitely many switching points near the target. In practice, the choice of $x_{20}$ is dictated by the required precision on the final state. For $\Delta = 10$ and a precision of $10^{-3}$, numerical simulations lead to $x_{20} = 6.9 \times 10^{-4}$. Figure 3 depicts the optimal trajectory on the unit sphere. The corresponding control $u(t)$ and switching function $\Phi(t)$ are displayed in Fig. 3.

The switching curves are reconstructed numerically by varying the small parameter $x_{20}$ and collecting all the switching points of the corresponding trajectories obtained by integrating the extremal flow backwards in time. The result is represented in Fig. 2 and 4. We observe in Fig. 4 that the points $(\pm 1, 0, 0)$ belong to the switching curves of the quantum system, independently
Figure 3: Time evolution of the control $u(t)$ (bold black line) and of the switching function $\Phi(t)$ (blue line) for the optimal trajectory of Fig. 2. The control switches from $\pm 1$ to $\mp 1$ when the switching function changes sign.

of the value of $\Delta$. It can be shown that if this curve goes out of the north hemisphere then it passes through the points ($\pm 1, 0, 0$) (see [23, Section 4.3], [24] and [25]).

The geometric analysis gives the optimal control protocol with a very high numerical precision. Due to their complexity, quantum control problems are usually solved by numerical optimization algorithms in which the control is assumed to be a piecewise constant function [12]. An intriguing question is therefore to study to what extent such solutions can approximate the chattering phenomenon. The numerical simulations presented below use a direct optimal method with the software BOCOP [25] with a fixed control time computed using our analysis. In Fig. 3, we observe that the chattering process can only be reproduced approximately by numerical optimization. Additional results can be found in [25]. The fineness of the time discretization corresponding here to $t_f/N$, where $N$ is the number of time steps, is a key factor to improve the protocol efficiency and to reproduce the control shape. However, for $N = 400$, we observe an erratic structure of the control. Without a precise understanding of the optimal control strategy, this switching accumulation could be misinterpreted as numerical instabilities or artifacts, while it is due to the very structure of the optimal control. Reasonable efficiencies of the control protocols are achieved for quite small values of $N$. Such sub-optimal strategies could be used to bypass the problem due to chattering in the numerical optimizations.
5 Conclusion

In summary, the analysis of a simple system has shown that chattering can appear in quantum optimal control. Such a process may play a key role in numerical optimization algorithms, as illustrated in the example by the rapid oscillations that occur in the numerical solutions. In accordance with the existing results on the ubiquity of such phenomena in OCT [9], we expect chattering to appear in many further examples, especially for higher-dimensional problems. Actually, one can give sufficient conditions on the relations between the commutators of the uncontrolled and controlled Hamiltonians to guarantee the existence of chattering solutions (see [9] and [1, Chapter 4]). Such conditions render the chattering phenomenon more and more frequent as the dimension grows, and our example shows that no obstacle to their appearance comes from the quantum structure of the control problem. It should be noticed, however, that one cannot deduce from the general conditions in high dimensions the optimality of the chattering trajectories, unlike the two-dimensional system considered in this paper.

Acknowledgment
This research has been partially supported by the ANR project “QUACO” ANR-17-CE40-0007-01 and received funding from the European Union Horizon 2020 research and innovation program under Marie-Sklodowska-Curie Grant No. 765267 (QUASCO). This research has been supported by the ANR-DFG project CoRoMo and the ANR project QuCoBEC ANR-22-CE47-0008-02.
Figure 5: Comparison between the optimal solution of the PMP (bottom right panel) and a numerical control designed by a direct approach for different time steps \(N\).
Figure 6: Schematic representation of the quantum system with the coupling $\Delta$ and $u(t)$ between the states $|1\rangle$ and $|2\rangle$ and $|2\rangle$ and $|3\rangle$ (red arrows). The black and blue dots indicate respectively the initial and the target states. The black arrow represents the relaxation process.

Data availability statement
The data that support the findings of this study are available upon reasonable request from the authors.

A Derivation of the model system

We show in this section how to derive the system used in this study. We consider a three-level quantum system in a $\Lambda$ configuration whose dynamics are governed by the Schrödinger equation. The system is described by a pure state $|\psi(t)\rangle$ which belongs to a three-dimensional Hilbert space spanned by the basis $\{|1\rangle, |2\rangle, |3\rangle\}$. The system is subject to a pump and a Stokes pulses coupling, respectively, states $|1\rangle$ and $|2\rangle$ and states $|2\rangle$ and $|3\rangle$. They are assumed to be on-resonance with the corresponding frequency transitions. There is no direct coupling between levels $|1\rangle$ and $|3\rangle$. A schematic representation of the control problem is given in Fig. 6.

The time evolution of the system is given by the Schrödinger equation

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = H(t)|\psi(t)\rangle,$$

where units such that $\hbar = 1$ are used. In the interaction representation and in the rotating-wave approximation, the Hamiltonian of the system can be written
as

\[ H(t) = \begin{pmatrix} 0 & \Delta & 0 \\ \Delta & 0 & u(t) \\ 0 & u(t) & 0 \end{pmatrix}, \]

where \( \Delta \) and \( u(t) \) are half the Rabi frequencies of the two pulses. We denote by \( c_1, c_2, \) and \( c_3 \) the complex coefficients of the state \( |\psi(t)\rangle \), and we introduce the real coefficients \( x_1, \ldots, x_6 \) defined by \( c_1 = x_1 + ix_4, c_2 = x_5 - ix_2, c_3 = x_3 + ix_6. \)

Since \( H(t) \) is Hermitian, we have that \( \sum_{i=1}^{6} x_i^2 \) is constant. Up to normalization, \( \sum_{i=1}^{6} x_i^2 = 1 \) in accordance to the description of a pure state. Straightforward computations from the Schrödinger equation show that the variables \( x_1, x_2, \) and \( x_3 \) are decoupled from \( x_4, x_5, \) and \( x_6. \) For our purposes it is sufficient to study the dynamics of the first set of variables, which turn out to be

\[
\begin{align*}
\dot{x}_1 &= -\Delta x_2, \\
\dot{x}_2 &= \Delta x_1 - u(t)x_3, \\
\dot{x}_3 &= u(t)x_2.
\end{align*}
\]

In the control problem, we consider an initial state of the dynamics such that \( x_4 = x_5 = x_6 = 0. \) We deduce that \( x_1^2 + x_2^2 + x_3^2 = 1 \) at any time \( t. \)

### B The Fuller model

This paragraph briefly describes the main results that can be established for the classical Fuller model. The interested reader will find the proofs of the different statements in textbooks of mathematical control theory [1, 7].

The Fuller model is an optimal control problem in \( \mathbb{R}^2. \) The dynamics of the state \( (x, y) \) are governed by the differential equations

\[
\begin{align*}
\dot{x} &= y, \\
\dot{y} &= u,
\end{align*}
\]

where the control \( u = u(t) \) satisfies the constraint \( u(t) \in [-1, 1]. \) Starting from the state \( (x_0, y_0), \) the goal of the control is to reach the origin \( (0, 0), \) while minimizing the cost functional \( J = \int_0^{t_f} x^2(t) dt, \) in which the control time \( t_f \) is not fixed.

The existence of optimal solutions to the problem above is not obvious, and can be proved by showing that the Hamilton–Jacobi–Bellman equation satisfied by the value function has a classical solution (see [7, Theorem 5.1.1 and Example 5.1.2]).

The optimal trajectories have the discrete symmetry

\( (x(t), y(t), u(t)) \mapsto (-x(t), -y(t), -u(t)), \)

and a scaling symmetry defined by the family of transformations

\( (x(t), y(t), u(t)) \mapsto (x_{\lambda}(t) = \lambda^2 x(t/\lambda), y_{\lambda}(t) = \lambda y(t/\lambda), u_{\lambda}(t) = u(t/\lambda)), \)

where \( \lambda \) is a positive parameter. This means that if \( (x(t), y(t), u(t)) \) is an optimal solution then \( (x_{\lambda}(t), y_{\lambda}(t), u_{\lambda}(t)) \) is also a solution, with initial state \( (\lambda^2 x_0, \lambda y_0) \)
and cost $\mathcal{J}_3 = \lambda^5 \mathcal{J}$. We deduce that if $(\bar{x}, \bar{y})$ is a switching point for an optimal trajectory, i.e., the corresponding control goes from $\pm 1$ to $\mp 1$ when the trajectory crosses $(\bar{x}, \bar{y})$, then the optimal synthesis has a switching curve of equation $x = -\xi \text{sign}[y] y^2$, where $\xi$ is a positive constant such that the curve passes through $(\bar{x}, \bar{y})$.

The second step of the analysis consists in applying the Pontryagin Maximum Principle, which is a necessary condition for optimality [3]. The Pontryagin Hamiltonian can be expressed as

$$H_P = p_x y + p_y u + p^0 x^2,$$

where $(p_x, p_y)$ is the adjoint state and $p^0$ is a constant multiplier equal either to 0 or to $-\frac{1}{2}$. If $(x(t), y(t))$ is an optimal trajectory with corresponding control $u(t)$ then there exist $(p_x(t), p_y(t))$ and $p_0$ such that $(p_x(t), p_y(t), p^0) \neq (0, 0, 0)$, $\bar{x} = \frac{\partial H_P}{\partial p_x}, \bar{y} = \frac{\partial H_P}{\partial p_y}, \bar{p}_x = -\frac{\partial H_P}{\partial x}, \bar{p}_y = -\frac{\partial H_P}{\partial y}$, and

$$H_P(x(t), y(t), p_x(t), p_y(t), u(t), p^0) = \max_{v \in [-1, 1]} H_P(x(t), y(t), p_x(t), p_y(t), v, p^0) = 0.$$

The switching function $\Phi$ is given by $\Phi(t) = p_y(t)$. As in the case studied in Section [3], one can restrict the analysis to normal extremals. Integrating system [5] from the state $(-\xi y_0^2, y_0)$ with $y_0 > 0$ and $u = +1$, and imposing that the corresponding switching function vanishes at a point $(\xi y_1^2, y_1)$ with $y_1 > 0$, we obtain that $\xi$ is solution of a polynomial equation of order 4, given by

$$\xi = \sqrt{\frac{\sqrt{33} - 1}{24}} \approx 0.44623. \quad (6)$$

The optimal trajectory has the following symmetries: denoting by $t_k$ the $k$-th switching time, by $t^\text{Full}_f$ the minimum time at which the optimal trajectory reaches $(0, 0)$, and introducing the parameter $\alpha = \sqrt{\frac{1 + \sqrt{2}}{1 - \sqrt{2}}} \approx 4.13016$, we have

$$t^\text{Full}_f - t_{k+1} = \frac{t^\text{Full}_f - t_k}{\alpha}, \quad \frac{x(t_{k+1})}{x(t_k)} = -\frac{1}{\alpha^2}, \quad \frac{y(t_{k+1})}{y(t_k)} = -\frac{1}{\alpha}.$$

Finally, starting from the state $(-\xi y_0^2, y_0)$, it can be shown that $t^\text{Full}_f = \frac{1 + \sqrt{2}}{\alpha^2} y_0$. An example of optimal trajectory is plotted in Fig. [7].

Figure [8] compares the results obtained for the quantum system and its approximation (Eq. (4) in the main text), which coincides with the Fuller model up to a linear rescaling of the first coordinate. In Fig. [8] we project the optimal trajectory of the quantum system onto the $x_1, x_2$- plane. This comparison highlights that the two solutions are very close to each other near the target state.

C Properties of the switching function for the three-level quantum system

We focus in this paragraph on the switching function $\Phi$ for the three-level quantum system. Using the Hamiltonian equations for the adjoint state in the normal
Figure 7: (top) Optimal trajectory for the Fuller model (black solid line). The switching curves are plotted in red. (bottom) Zoom of the top panel near the origin.
Figure 8: Plot onto the \((x_1, x_2)\)-plane of the optimal trajectory of the quantum control problem (solid black line). The dashed line depicts the solution of the approximation (4) given in the main text. The red and blue curves represent respectively the switching curves for the quantum and Fuller systems. The bottom panel is a zoom of the top one.
case (Eq. (2) in the main text), one computes that \( \Phi \) and its time derivatives can be expressed as
\[
\Phi = p_3 x_2 - p_2 x_3, \\
\dot{\Phi} = \Delta(x_1 p_3 - x_3 p_1), \\
\ddot{\Phi} = -\Delta^2 \Phi + \Delta u(x_1 p_2 - x_2 p_1) - \Delta x_1 x_3.
\]

Moreover, on a segment where \( \Phi \neq 0 \) and, therefore, \( u \) is constantly equal to +1 or -1, one has
\[
\Phi^{(3)} = -(\Delta^2 + 1) \dot{\Phi} - 2\Delta u x_1 x_2 + \Delta^2 x_2 x_3, \\
\Phi^{(4)} = - (\Delta^2 + 1) \ddot{\Phi} + \Delta (\Delta^2 + 2) x_1 x_3 + \Delta^2 u(3x_2^2 - 2x_1^2 - x_3^2).
\]

Note that, if \( x_2 \neq 0 \) and \( \Phi = 0 \), then \( p_1 = \frac{x_1}{x_2} p_2 \) and hence
\[
\dot{\Phi} = \Delta x_2 \left( x_1 p_2 - x_2 p_1 \right).
\]

Notice also that, since \( H_P = 0 \) and using again that \( \Phi = 0 \), we have
\[
x_1 p_2 - x_2 p_1 = \frac{x_2^2}{2} \geq 0.
\]

Hence, for \( x_2 x_3 \neq 0 \) and \( x_1 \neq 0 \), the derivative \( \dot{\Phi} \) has the same sign as \( x_2 x_3 \), which means that switches only occur from -1 to +1 if \( x_2 x_3 > 0 \) and from +1 to -1 if \( x_2 x_3 < 0 \). If \( x_2 \neq 0 \) and \( x_1 = 0 \) then \( \dot{\Phi} = 0 \), which implies that \( p_1 = 0 \). Therefore \( \dot{\Phi} = 0 \) and
\[
\Phi^{(3)} = \Delta^2 x_2 x_3.
\]

It follows that everywhere on the set \( \{ x \mid x_2 x_3 > 0 \} \) switches only occur from -1 to +1, while on \( \{ x \mid x_2 x_3 < 0 \} \) they only occur from +1 to -1. This result proves a statement of the main text, i.e. if the switching curve goes out of the north hemisphere then it passes through the points \((\pm 1, 0, 0)\). This property does not depend on the parameter \( \Delta \).

**D A sufficient condition for chattering**

In this section we present an adaptation of the results presented in Chapter 3 of the book [1] by Zelikin and Borisov, concerning sufficient conditions for the appearance of Fuller-like chattering in a two-dimensional optimal synthesis.

We consider a control system of the form
\[
\begin{aligned}
\dot{x} &= \Delta y + \phi^x_1(x, y) + u \phi^x_2(x, y), \\
\dot{y} &= u + \phi^y_1(x, y) + u \phi^y_2(x, y).
\end{aligned}
\]

(7)

We also suppose that \( \phi^x_i, \phi^y_i \) are smooth and small in the following sense: denoting by \( g_\kappa \) the anisotropic dilatation \( g_\kappa(x, y) = (\kappa^2 x, \kappa y) \), one has
\[
\limsup_{\kappa \to 0^+} \frac{|\phi^x_i(g_\kappa(x, y))|}{\kappa} < \infty, \quad \limsup_{\kappa \to 0^+} \frac{|\phi^y_i(g_\kappa(x, y))|}{\kappa^2} < \infty,
\]

(8)
for every \((x, y) \in \mathbb{R}^2\). Then, the optimal control problem

\[
\begin{align*}
\int_0^T x^2(t)dt & \longrightarrow \min \\
T > 0 \text{ free, } \quad u & \in L^\infty([0, T], [-1, 1]) \\
t & \mapsto (x(t), y(t)) & \in W^{1, \infty}([0, T], \mathbb{R}^2)
\end{align*}
\]

satisfies the following properties.

Theorem 1. For every \((x_0, y_0)\) in a sufficiently small neighbourhood of \((0, 0)\), Problem (9) admits a solution. Such a solution reaches the fixed point \((0, 0)\) in finite time and its corresponding control has infinitely many discontinuities accumulating at the final time. Moreover, there is no other solution of (9) up to prolongation by a constant trajectory at \((0, 0)\). The optimal synthesis is characterized by a switching curve of the form

\[
\Gamma = \begin{cases} 
    x = \lambda_1(y)y^2, & y > 0, \\
    x = \lambda_0(y)y^2, & y < 0,
\end{cases}
\]

where \(\lambda_0\) and \(\lambda_1\) are \(C^1\) function satisfying \(\lambda_0(0) = -\lambda_1(0) = \xi \Delta\), where \(\xi\) is as in (7). The optimal control is \(-1\) above \(\Gamma\) and \(+1\) below it.

All the proof material is in the book [1] by M. I. Zelikin and V. F. Borisov. Nevertheless, since this theorem is not explicitly stated as such, we propose to retrace the ingredients of the proof and point out the relevant statements in [1].

Proof. First, let us highlight the fact that the existence of a solution of (9) is not obvious. Indeed, it is a priori possible that the cost decreases by reaching the target later in time. As a consequence we are lacking compactness for existence of the solution of the problem with free final time. Existence of optimal trajectories can be obtained once an extremal synthesis has been constructed, using a field-of-extremals argument ([1] Theorem 3.3).

The Hamiltonian given by the PMP for Problem (9) is

\[
H_P(x, y, p_x, p_y, p_0, u) = p_x (\Delta y + \phi_1^y(x, y) + u\phi_2^y(x, y)) + p_y (u + \phi_1^x(x, y) + u\phi_2^x(x, y)) + p_0 x^2.
\]

Hence, the adjoint equations for normal extremals \((p_0 = -1/2)\) are

\[
\begin{align*}
\dot{p}_x &= x - p_x \partial_x (\phi_1^y + u\phi_2^y) - p_y \partial_y (\phi_1^x + u\phi_2^x), \\
\dot{p}_y &= -\Delta p_x - p_x \partial_y (\phi_1^y + u\phi_2^y) - p_y \partial_y (\phi_1^x + u\phi_2^x). \tag{11}
\end{align*}
\]

If we define \((z_1, z_2, z_3, z_4) = ((p_y + p_x \phi_2^y + p_y \phi_2^y)/\Delta^2, -p_x/\Delta, -x/\Delta, -y)\), we get

\[
\begin{align*}
\dot{z}_1 &= z_2 + f_1(z, u), \\
\dot{z}_2 &= z_3 + f_2(z, u), \\
\dot{z}_3 &= z_4 + f_3(z, u), \\
\dot{z}_4 &= u = \text{sign}[z_1]. \tag{12}
\end{align*}
\]

Introducing \(g_\kappa(z_1, z_2, z_3, z_4) = (\kappa^4 z_1, \kappa^3 z_2, \kappa^2 z_3, \kappa z_4)\), we have the smallness property

\[
\limsup_{\kappa \to 0} \frac{|f_\kappa(g_\kappa(z), u)|}{\kappa^{3-i}} < \infty \quad \text{for } 1 \leq i \leq 4, \tag{13}
\]
which generalises [5]. As a consequence, this system is of the form given in Equation (3.5) of [1]. The smallness condition [13] allows to apply [1][Proposition 4.1], which guarantees, under a rank condition discussed below, that the conclusions of Theorem 3.3 in [1] hold true.

The study of this dynamical system around \( z = 0 \) is performed in Chapter 3 of [1] thanks to a blow-up procedure. The (degenerate) hyperbolicity of the fixed point \( z = 0 \) is established, as well as the existence of a two-dimensional invariant contracting manifold \( \Sigma \) corresponding to the trajectories converging to \( z = 0 \) [1, Section 3.4 - 3.8]. This manifold is given by the trajectories of the non-smooth system (12) switching on curves of the form

\[
\hat{\Gamma}^0 = \{ (0, \mu_0(\kappa)\kappa^3, \lambda_0(\kappa)\kappa^2, \kappa) \mid \kappa < 0 \},
\]

\[
\hat{\Gamma}^1 = \{ (0, \mu_1(\kappa)\kappa^3, \lambda_1(\kappa)\kappa^2, \kappa) \mid \kappa > 0 \},
\]

where \( \mu_i, \lambda_i \) are smooth functions satisfying (cf. [1, Lemma 3.3])

\[
\lambda_0(0) \in \left( -\frac{\Delta}{2}, 0 \right), \lambda_1(0) \in \left( 0, \frac{\Delta}{2} \right), \mu_0(0) = \frac{1}{2} \lambda_0(0)^2, \mu_1(0) = \frac{1}{2} \lambda_1(0)^2.
\]

Besides, since \( \lambda_0(0) \) and \( \lambda_1(0) \) are solutions of the same polynomial system (cf. Eq. (3.13) in [1]) as in the case of the following Fuller dynamics

\[
\dot{x} = \Delta y, \\
\dot{y} = u,
\]

we get \( \lambda_0(0) = -\lambda_1(0) = \xi \Delta \).

Let us now check that the projection \( \pi : (z_1, z_2, z_3, z_4) \mapsto (z_3, z_4) \) restricted to \( \Sigma \) is a \( C^1 \) mapping with a Jacobian matrix of maximal rank on \( \Sigma \setminus \hat{\Gamma} \), as required in [1, Theorem 3.3]. Let us denote by \( \mathcal{F}(\kappa, t) \) the solution of System (12) at time \( t \) with initial condition parameterized by \( \kappa \) as in (14) (an analogous argument applying for the points of \( \Sigma \) parameterized by the trajectories of System (12) starting from \( \hat{\Gamma}^1 \)). Let \( t \) be such that \( \mathcal{F}(\kappa, t) \notin \hat{\Gamma}^0 \cup \hat{\Gamma}^1 \) for all \( 0 < s < t \). Then \( \mathcal{F}(\kappa, s) \) is given by the solution of (12) with \( u = -1 \) and \( \pi(\mathcal{F}(\kappa, s)) \) is the solution of System (11) with constant control \( u = -1 \) and initial condition \( (x_0, y_0) = (\lambda_0(\kappa)\kappa^2, \kappa) \). Therefore, \( (\kappa, s) \mapsto \pi(\mathcal{F}(\kappa, s)) \) is smooth. Let us check that \( \partial_{\kappa} \pi(\mathcal{F}(\kappa, t)) \) and \( \partial_t \pi(\mathcal{F}(\kappa, t)) \) are linearly independent. Using the fact that the flow of (11) is a diffeomorphism, it is enough to check that the two vectors

\[
\partial_{\kappa} \pi(\mathcal{F}(\kappa, 0)) = (\lambda_0'(\kappa)\kappa^2 + 2\lambda_0(\kappa)\kappa, 1)
\]

and

\[
\partial_t \pi(\mathcal{F}(\kappa, 0)) = (\kappa + f_3(z, -1), -1 + f_4(z, -1)),
\]

are linearly independent. Using (16), this is the case for \( \kappa \) small enough. We can thus apply [1][Theorem 3.3], which allows to conclude the proof of Theorem 1.

### E Numerical optimization procedure

#### E.1 Backward stability of the optimal synthesis

For the quantum system, we point out that the optimal trajectory can be computed numerically starting from that of the Fuller model. When we are sufficiently close to \( |3\rangle \), we approximate the switching curve of the quantum system
by that of its approximation. As described in the main text, we then integrate backwards in time the dynamics of the state and of the adjoint state. We stress that it makes sense to integrate backwards the extremal flow since, according to the results in [1], the map \( \Upsilon \) associating with a switching point the subsequent one is hyperbolic and has \( \Gamma \) as stable manifold towards \( |3\rangle \). As a result, the \( k \)-th iteration of the inverse of \( \Upsilon \) tends towards \( \Gamma \) as \( k \) grows.

### E.2 A direct optimization method

We apply in this paragraph a direct optimization method to design numerically the solution of the optimal control problem. We use the open access optimal solver BOCOP [25]. A direct optimization approach is a procedure in which the state and the time are discretized in time, transforming the initial optimal control problem into a nonlinear constrained optimization problem. In this optimization protocol, the optimal control problem is slightly modified. The goal is to minimize the cost 
\[
C = \int_0^{t_f} x_1(t) dt,
\]
while reaching a state as close as possible of the target \(|3\rangle\) in a fixed time \(t_f\) (note that the algorithm does not converge if the time is free). In order to have the fairest comparison possible between the two approaches, we estimate as follows the time of the optimal solution presented in the main text. We assume that the control time is the sum of the time \(\tau\) taken by the quantum system to go from \((0, 1, 0)\) to a point with first two components equal to \((\xi, x_2, 0)\), \(x_2 > 0\) small, and the time \(t_{Ful}^{\text{ful}}\) corresponding to the initial condition \((\xi, x_2^{\text{ful}}, 0)\) for the Fuller problem.

For \(x_2 = 6.9 \times 10^{-4}\) we obtain \(\tau = 2.5890\) and \(t_{Ful}^{\text{ful}} = 0.0011\), which leads to \(t_f = 2.5901\). We set \(t_f\) to 2.59 in the numerical optimization procedure. The time subdivision is regular and given by the number of time steps \(N\), going from \(N = 50\) to \(N = 400\). The other optimization parameters are set to their default or recommended values in BOCOP.

In addition of the controls displayed in Fig. 5 of the main text, we report in Fig. 9 additional numerical results about the distance to the target and the cost \(C\) with respect to \(N\). When \(N\) becomes larger, the numerical optimal process seems to converge towards the optimal solution both in terms of final state and cost. For quite small values of \(N\), we observe that the efficiency of the control is already reasonable. We point out that such sub-optimal strategy could be a possible option to bypass the chattering phenomenon in the optimization procedure.

### References

[1] Zelikin M. I. and Borisov V. F. 1994 *Theory of Chattering Control* (Systems and Control: Foundations and Applications, Boston, MA, USA: Birkhauser)

[2] Misra B. and Sudarshan E. C. G. 1977 *J. Math. Phys.* 18, 756

[3] Viola L., Knill E. and Lloyd S. 1999 *Phys. Rev. Lett.* 82, 2417

[4] Pontryagin L. et al 1974 *Mathematical theory of optimal processes* (Mir, Moscow)
Figure 9: Evolution of the distance to the target (top) and of the cost $C$ (bottom) as a function of the number of time steps (crosses). In the bottom panel, the horizontal solid line (in red) indicates the efficiency of the optimal solution. The solid black line has been added in the two panels to help the visualisation.
[5] Fuller A. T. 1960 *Relay control systems optimized for various performance criteria*, Proc. First World Congress IFAC, Moscow, page 510.

[6] Borisov V. F. 2000 J. Math. Sci. **100**, 2311

[7] Schättler H. and Ledzewicz U. 2012 *Geometric optimal control* (Springer, volume 38 of Interdisciplinary Applied Mathematics, New York)

[8] Boscain U., Sigalotti M. and Sugny D. 2021 *PRX Quantum* **2**, 030203

[9] Kupka I. 1990 *The ubiquity of Fuller’s phenomenon* in Nonlinear Controllability and Optimal Control (Monograph Textbooks Pure and Applied Mathematics, New York, NY, USA: Dekker **133**, 313)

[10] Schättler H. and Ledzewicz U. 2009 *Control and Cybernetics* **38**, 1501

[11] Zu J., Trélat E. and Cerf M. 2016 Discrete Continuous Dynam. Syst. B **21**, 1347

[12] Glaser S. J., Boscain U., Calarco T., Koch C., Kockenberger W., Kosloff R., Kuprov I., Luy B., Schirmer S., Schulte-Herbrüggen T., Sugny D. and Wilhelm F. 2015 *Eur. Phys. J. D* **69**, 279

[13] Brif C., Chakrabarti R. and Rabitz H. 2010 *New J. Phys.* **12**, 075008

[14] Koch C.P., Boscain U., Calarco T., Dirr G., Filipp S., Glaser S. J., Kosloff R., Montangero S., Schulte-Herbrüggen T., Sugny D. and Wilhelm F. K. 2022 *EPJ Quantum Technol.* **9**, 19

[15] Koch C. P., Lemeshko M. and Sugny D. 2019 *Rev. Mod. Phys.* **91**, 035005

[16] D’Alessandro, D. 2008 *Introduction to Quantum Control and Dynamics* (Chapman and Hall, Boca Raton, FL)

[17] Bryson A. E. and Ho Y. C. 1975 *Applied Optimal Control: Optimization, Estimation, and Control* (Taylor and Francis, Philadelphia)

[18] Caponigro M., Ghezzi R., Piccoli B. and Trélat E. 2018 IEEE Transaction on Automatic Control **63**, 2046

[19] Vitanov N. V., Rangelov A. A., Shore B. W. and Bergmann K. 2017 Rev. Mod. Phys. **89**, 015006

[20] Lewis, R. M. 1980 *SIAM J. Control Optim.* **18**, 21

[21] Lokutsievskii, L. V. 2014 *Sb. Math.* **205**, 432

[22] Agrachev A., Barilari D. and Boscain U. 2020 *A Comprehensive Introduction to sub-Riemannian Geometry* (Cambridge University Press, volume 181 of Cambridge Studies in Advanced Mathematics, Cambridge)

[23] Boscain U. and Chitour Y. 2005 *SIAM J. Control Optim.* **44**, 111

[24] Boscain U. and Piccoli B. 2004 *Optimal syntheses for control systems on 2-D manifolds* (vol. 43 of Mathématiques and Applications (Berlin), Springer-Verlag, Berlin)
[25] Bonnans F., Grelard V. and Martinon P. 2011 *Bocop, the optimal control solver, open source toolbox for optimal control problems*; URL http://bocop.org.