J Intell Robot Syst (2018) 92:85–106
https://doi.org/10.1007/s10846-017-0736-x

An Analysis of Policies from Stochastic Linear Quadratic Gaussian in Robotics Problems with State- and Control-Dependent Noise

Hugo Carlos[1] · **Jean-Bernard Hayet**[1] · **Rafael Murrieta-Cid**[1]

Received: 19 May 2016 / Accepted: 2 November 2017/ Published online: 20 November 2017

Abstract Among all the mathematical frameworks used in the control and robotics communities to handle uncertainties, the stochastic variants of optimal control frameworks are appealing, in particular because of the existence of efficient tools to solve them computationally, such as dynamic programming. However, in many cases, because of their formulation as a classical optimization problem, it may be difficult to ponder the expected solutions for a given choice of the objective function to minimize. In this paper, we perform an in-depth analysis of the behavior of the policies obtained from solving Stochastic Linear Quadratic Gaussian problems, thinking in particular in robot motion planning applications. To perform this analysis, we assume simplified linear systems perturbed by Gaussian noise, with state-dependend and control-dependent components, and objective functions summing up control-related and state-related costs. We provide (1) useful bounds for understanding the effect of the objective function parameters, (2) insights on what the expected paths of system should be and (3) results on the optimal choice of the planning horizon.

Keywords Motion Planning under uncertainties · Linear quadratic regulator · Markov decision process

✉ Hugo Carlos
hucarlos@cimat.mx

Jean-Bernard Hayet
jbhayet@cimat.mx

Rafael Murrieta-Cid
murrieta@cimat.mx

[1] Centro de Investigación en Matemáticas (CIMAT), Guanajuato, México

1 Introduction

Dealing with *uncertainty* has always been an important issue in mobile robotics [2, 5, 7–9, 11, 14]. After the pioneering work of [9], where the authors have introduced a probabilistic framework for formally modeling spatial uncertainty, many approaches have been proposed to generate motion strategies for a given robotic task while taking into account uncertainties. A common and important task in robotics is navigation, where a robot must reach a given goal configuration even under uncertainty on states, controls and observations [5, 8, 14].

One common class of solutions for this kind of problems is given by optimal control, which has been the focus of attention in several scientific communities over the years, from robotics to control. Numerous algorithms and methodologies have been developed in fields such as biomechanics, robotics, economics, etc. based on optimal control to solve complex problems [4]. The mathematical tools it provides are well adapted to determining the optimal policies a system should follow, even under uncertainties on the effects of the control and/or on the state estimate.

Stochastic variants of dynamic programming and optimal control [1, 3–5, 8] have been recurrent tools used to generate motion strategies in the form of policies, that map states to controls. Previous works in the fields of robotics, control and artificial intelligence [2, 12, 14] have adapted the dynamic programming framework to partially observable Markov decision processes (POMDP) [6, 11], in order to solve optimal control problems and simultaneously handle uncertainty on states or controls. These algorithms have been applied, for example, in [12], where the authors propose an approach to find the optimal controls for non-linear systems under the Markov decision process (MDP)

framework. In particular they apply this algorithm to biomechanical models. In robotics for instance, one can find in the literature different algorithms for complex robotic systems [13, 14]. In these papers the main idea is to extend the results of [12] for a partially observable Markov decision process (POMDP) and apply it to practical robotic problems.

However, the complexity of these algorithms makes it difficult to assess quantitatively the solution when the inherent system conditions change. Assume for example, that one can predict the variations in the solutions for differents levels of noise. It would then be possible to make decisions on other aspects of the system or even of the admissibility of the solution. Also, because of their formulation as a classical optimization problem, it may be difficult to ponder the expected solutions for a given choice of the objective function to minimize.

In this work, we present the conditions for the optimal cost-to-go function and optimal control parameters to converge to finite values when the horizon increases. This implies that the cost associated to the optimal motion policy is bounded by a finite value, when considering a bounded domain for the initial state. Therefore, one knows whether or not the resulting cost is affordable. We introduce a quantitative and qualitative analysis of the expected solutions to the Linear Quadratic Gaussian (LQG) problem with control-dependent noise. In particular, we analyze how the system trajectory changes when one varies the weights of the immediate costs, control costs and costs associated to the final configuration, thus understanding the effect of these parameters in the objective function.

Finally, it is important to stress that dynamic programming and related methods deliver the global optimum (minimum) for a specific horizon. However, such a minimum might be improved with another planning horizon. In this work, we determine the optimal planning horizon K for the overall cost used to assess the motion policy, under control-dependent uncertainty. The optimal horizon K can be of finite nature, or be reached asymptotically when $K \to \infty$. Our analysis allows us to determine from the initial conditions of the system, which of these two scenarios will occur. In practice, this means that it is feasible to determine, for any system configuration, whether it is possible to find a finite horizon with an optimal that guarantees the lowest overall cost.

2 LQG with State- and Control-Dependent Noise

In this section, we review the general formulation of the LQG problem (Section 2.1) and the general form of its solution (Section 2.2) when state- and noise-dependent noise add up as the result of the application of a control.

2.1 Problem Statement and Notations

Let $\mathbf{x}_k \in \mathbb{R}^n$ be the state of a linear stochastic system at time index k, controlled through a control vector $\mathbf{u}_k \in \mathbb{R}^m$.

In this work, we consider the following LQG problem associated to this system

$$\min_{\mathbf{u}_{0:K-1}} E_{\xi_{\mathbf{u}_0}, \xi_{\mathbf{x}_0}, \ldots, \xi_{\mathbf{u}_{K-1}}, \xi_{\mathbf{x}_{K-1}}} \left(\sum_{k=0}^{K-1} l_k(\mathbf{x}_k, \mathbf{u}_k) + l_K(\mathbf{x}_K) \right) \tag{1}$$

with:

$$\begin{cases} \mathbf{x}_{k+1} &= A\mathbf{x}_k + B\mathbf{u}_k + \xi_{\mathbf{u}_k} + \xi_{\mathbf{x}_k} \\ \xi_{\mathbf{u}_k} &\sim L_u(\mathbf{u}_k) N(0, I) \\ \xi_{\mathbf{x}_k} &\sim L_x(\mathbf{x}_k) N(0, I) \\ \ell_k(\mathbf{x}, \mathbf{u}) &= \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \frac{1}{2}\mathbf{u}^T R\mathbf{u} \text{ for } 0 \leq k < K \\ \ell_K(\mathbf{x}) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}_f)^T Q_f (\mathbf{x} - \mathbf{x}_f). \end{cases}$$

In these equations, $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are respectively the dynamics and control matrices that characterize the linear stochastic system, $\xi_{\mathbf{u}_k} \in \mathbb{R}^n$, $\xi_{\mathbf{x}_k} \in \mathbb{R}^n$ are independent Gaussian noises that add up to the state at each time step k, and $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{m \times m}$, $Q_f \in \mathbb{R}^{n \times n}$ are three positive semi-definite matrices describing the cost function. Finally, $N(0, I)$ is the normal distribution in $\mathbb{R}^n$ and $L_u(\mathbf{u}_k)$, $L_x(\mathbf{x}_k)$ two matrices described below.

This is a simple but very typical LQG system in robotic problems, encoding the search for an optimal control in the case of a linear stochastic system described by the first equation above, and that can be associated to a Markov Decision Process. Its instantaneous cost function $l_k(\mathbf{x}, \mathbf{u})$ is described in the fourth equation above and includes two terms: One quadratic term in the state $\mathbf{x}$ that should favor solutions driving the system close to the minimal value of the quadratic function, which is the origin in that case, and one quadratic term in the control $\mathbf{u}$ that favors solutions with controls of small amplitude. Finally, the term $l_K(\mathbf{x})$ corresponds to a constraint imposed to the last position of the generated path and is used to drive the system close to some configuration of interest, $\mathbf{x}_f \in \mathbb{R}^n$, at time K. Note that the goal configuration $\mathbf{x}_f$ is not reached precisely, since the robot objective is to minimize the overall cost function, so that in some situations it may result better in terms of the cost, not to reach exactly the goal configuration.

Now, one important characteristic that we will assume is that the uncertainty applying on this system comes from two separate terms $\xi_{\mathbf{u_k}}, \xi_{\mathbf{x}_k}$. The variance of the additive noise $\xi_{\mathbf{u_k}}$ is a function of the control, and the variance of $\xi_{\mathbf{x}_k}$ depends on the state. More specifically, in this analysis,

we choose the following forms for the scaling matrices $L_u(\mathbf{u}_k) \in \mathbb{R}^{n\times n}$ and $L_x(\mathbf{x}_k) \in \mathbb{R}^{n\times n}$:

$$L_u(\mathbf{u}_k)_i = F_{u_i}\mathbf{u}_k \tag{2}$$

$$L_x(\mathbf{x}_k)_i = F_{x_i}\mathbf{x}_k \tag{3}$$

where the index i refers to the i-th column and where we introduce two matrices $F_{u_i} \in \mathbb{R}^{n\times m}$, $F_{x_i} \in \mathbb{R}^{n\times n}$. In other terms, the noise has two independent Gaussian components with individual standard deviations linear in the control and state components, respectively. It is straightforward to verify that the variance of $\xi_{\mathbf{u}_k}$ and $\xi_{\mathbf{x}_k}$ at time k are respectively given by $\sum_{i=1}^{n} F_{u_i}\mathbf{u}_k\mathbf{u}_k^T F_{u_i}^T$ and $\sum_{i=1}^{n} F_{x_i}\mathbf{x}_k\mathbf{x}_k^T F_{x_i}^T$.

Our objective is to understand the behavior of the solutions under varying cost functions parameters.

2.2 General Solution of the System

In this sub-section, we describe the general solutions to problem (1).

First of all, given the nature of this system, one can show that the optimal cost-to-go function at time k (i.e. the function giving the value of the minimal possible cost attainable starting at time k up to time K, by a sequence of optimal controls) is quadratic in terms of the state [4]. Namely, if we denote the optimal cost-to-go function starting from time k, at state $\mathbf{x}_k$, as $v_k^*(\mathbf{x}_k)$, we can write it as

$$v_k^*(\mathbf{x}_k) = s_k + \mathbf{x}_k^T\mathbf{s}_k + \frac{1}{2}\mathbf{x}_k^T S_k\mathbf{x}_k,$$

with s_k a scalar, $\mathbf{s}_k \in \mathbb{R}^n$ a vector, and S_k an $n \times n$ matrix. The matrix S_k can be expressed by the principles of Dynamic Programing (DP) into a stochastic form of the Riccati equations as follows. The control-dependent cost-to-go $v_k(\mathbf{x}_k, \mathbf{u}_k)$ is given by the sum of the control-dependent instantaneous cost plus the expected value of the optimal cost-to-go function after this control is done:

$$\begin{aligned} v_k(\mathbf{x}_k, \mathbf{u}_k) &= \frac{1}{2}\mathbf{x}_k^T Q\mathbf{x}_k + \frac{1}{2}\mathbf{u}_k^T R\mathbf{u}_k \\ &\quad + E_{\xi_{\mathbf{u}_k},\xi_{\mathbf{x}_k}}\left[v_{k+1}^*(A\mathbf{x}_k + B\mathbf{u}_k + \xi_{\mathbf{u}_k} + \xi_{\mathbf{x}_k})\right] \\ &= \frac{1}{2}\mathbf{x}_k^T Q\mathbf{x}_k + \frac{1}{2}\mathbf{u}_k^T R\mathbf{u}_k \\ &\quad + \frac{1}{2}(A\mathbf{x}_k + B\mathbf{u}_k)^T S_{k+1}(A\mathbf{x}_k + B\mathbf{u}_k) \\ &\quad + (A\mathbf{x}_k + B\mathbf{u}_k)^T\mathbf{s}_{k+1} + s_{k+1} \\ &\quad + \frac{1}{2}E_{\xi_{\mathbf{x}_k}}[\xi_{\mathbf{x}_k}^T S_{k+1}\xi_{\mathbf{x}_k}] + \frac{1}{2}E_{\xi_{\mathbf{u}_k}}[\xi_{\mathbf{u}_k}^T S_{k+1}\xi_{\mathbf{u}_k}], \end{aligned} \tag{4}$$

where the expectation is taken over the possible realizations of $\xi_{\mathbf{u}_k}, \xi_{\mathbf{x}_k}$. Let us analyze the expectation terms. Since for any pair of matrices $X, Q \in \mathbb{R}^{n\times n}$, we have $\mathrm{Tr}[X^T QX] = \sum_i X_i^T QX_i$ where X_i are the columns of X, then,

$$\begin{aligned} E_{\xi_{\mathbf{u}_k}}[\xi_{\mathbf{u}_k}^T S_{k+1}\xi_{\mathbf{u}_k}] &= \mathrm{Tr}(S_{k+1}L_u(\mathbf{u}_k)L_u(\mathbf{u}_k)^T) \\ &= \mathbf{u}_k^T\left(\sum_i F_{u_i}^T S_{k+1}F_{u_i}\right)\mathbf{u}_k. \end{aligned}$$

Similarly for $E_{\xi_{\mathbf{x}_k}}[\xi_{\mathbf{x}_k}^T S_{k+1}\xi_{\mathbf{x}_k}]$ we have:

$$E_{\xi_{\mathbf{x}_k}}[\xi_{\mathbf{x}_k}^T S_{k+1}\xi_{\mathbf{x}_k}] = \mathbf{x}_k^T\left(\sum_i F_{x_i}^T S_{k+1}F_{x_i}\right)\mathbf{x}_k.$$

Replacing the previous two equations into Eq. 4, we have

$$\begin{aligned} v_k(\mathbf{x}_k, \mathbf{u}_k) = &\frac{1}{2}\mathbf{x}_k^T\left(Q + A^T S_{k+1}A + \sum_i F_{x_i}^T S_{k+1}F_{x_i}\right)\mathbf{x}_k \\ &+ \frac{1}{2}\mathbf{u}_k^T\left(R + B^T S_{k+1}B + \sum_i F_{u_i}^T S_{k+1}F_{u_i}\right)\mathbf{u}_k \\ &+ \mathbf{x}_k^T\left(A^T\mathbf{s}_{k+1}\right) + \mathbf{u}_k^T\left(B^T\mathbf{s}_{k+1} + B^T S_{k+1}A\mathbf{x}_k\right) + s_{k+1}. \end{aligned} \tag{5}$$

Differentiating Eq. 5 with respect to $\mathbf{u}_k$ and equating the derivative to zero, we get the optimal control $\mathbf{u}_k^*$ as an affine function of the state $\mathbf{x}_k$

$$\mathbf{u}_k^* = -H_k^{-1}\mathbf{g}_k - H_k^{-1}G_k\mathbf{x}_k, \tag{6}$$

with

$$H_k \triangleq R + B^T S_{k+1}B + \sum_i F_{u_i}^T S_{k+1}F_{u_i}, \tag{7}$$

$$G_k \triangleq B^T S_{k+1}A, \tag{8}$$

$$\mathbf{g}_k \triangleq B^T\mathbf{s}_{k+1}, \tag{9}$$

where $H_k \in \mathbb{R}^{m\times m}$, $G_k \in \mathbb{R}^{m\times n}$, $\mathbf{g}_k \in \mathbb{R}^m$. By replacing $\mathbf{u}_k^*$ into Eq. 5 we get the corresponding Riccati recursive equations that give the coefficients of the quadratic optimal cost-to-go function $v_k^*(\mathbf{x}_k)$,

$$S_k = Q + A^T S_{k+1}A + \sum_i F_{x_i}^T S_{k+1}F_{x_i} - G_k^T H_k^{-1}G_k, \tag{10}$$

$$\mathbf{s}_k = A^T\mathbf{s}_{k+1} - G_k^T H_k^{-1}\mathbf{g}_k, \tag{11}$$

$$s_k = s_{k+1} - \frac{1}{2}\mathbf{g}_k^T H_k^{-1}\mathbf{g}_k. \tag{12}$$

To solve the recursion, we start at K with

$$\begin{aligned} S_K &= Q_f, \\ \mathbf{s}_K &= -Q_f\mathbf{x}_f, \\ s_K &= \frac{1}{2}\mathbf{x}_f^T Q_f\mathbf{x}_f \end{aligned} \tag{13}$$

and apply Eqs. 10, 11, and 12. This is the general form of the solution, as used in several optimal control problems.

3 Behavior of LQG Solutions on a Simplified System

To give an insight to the behavior of the general solutions presented in the previous section, we consider in this section a simplified version of the described system and analyze its solutions in terms of the different parameters involved in the optimal control formulation. The parameters that we analyse are: λ_f, a factor weighting the importance of reaching the goal configuration; λ, a factor weighting the cost of the state; μ, a factor that weights the control cost; σ_x, the standard deviation factor of the state-dependent noise; σ_u, the standard deviation factor of the control-dependent noise (the larger these two parameters are, the more imprecise the robot is). These factors define the costs $\ell_k(\mathbf{x}, \mathbf{u})$ and $\ell_K(\mathbf{x})$ in the simplified system derived from Eq. 1 and we analyze them hereafter.

3.1 Simplified System

To perform our analysis further, we will suppose in the following sections that the state and control dimensions are the same, $m = n$, that $A = B = I_{n\times n}$ for the dynamic system equations, and that $Q = \lambda I_{n\times n}$, $Q_f = \lambda_f I_{n\times n}$, $R = \mu I_{m\times m}$ for the individual terms of the quadratic cost function, where as it was mentioned above, λ, λ_f, μ are scalars that weight the importance of each of the three components of the cost function described in the previous section. The matrix $I_{n\times n}$ refers to the identity matrix in dimension n. When clear from the context, we will denote it simply as I.

Moreover, we will assume that the noises $\xi_{\mathbf{u}_k}$ and $\xi_{\mathbf{x}_k}$ have their standard deviations given by the matrices of the following form specifying Eqs. 2 and 3:

$$L_u(\mathbf{u}) = \sigma_u D(\mathbf{u}),\ \ L_x(\mathbf{x}) = \sigma_x D(\mathbf{x}),$$

where $D(\mathbf{a})$ is the diagonal matrix with its diagonal elements being the entries of the vector $\mathbf{a}$. Hence, the corresponding covariance matrices are diagonal with diagonal elements $\sigma_a^2 a_i^2$, i.e. the noise vector $\xi_{\mathbf{u}_k}$ (resp. $\xi_{\mathbf{x}_k}$) has independent components on each direction of the control (resp. the state), with the variance being proportional to the square of the control (resp. state) components. For example:

$$E_{\xi_{\mathbf{u}_k}}[\xi_{\mathbf{u}_k}\xi_{\mathbf{u}_k}^T] = \sigma_u^2 \begin{pmatrix} u_1^2 & 0 & \dots & 0 \\ 0 & u_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & u_n^2 \end{pmatrix}.$$

For the particular case of the control, such a form means that the noise amplitude increases with the absolute value of the control.

3.2 Optimal cost-to-go Function

Under the assumptions presented above, the terms S_k, $\mathbf{s}_k$ and s_k that correspond to the coefficients of the cost-to-go quadratic form can be written as:

$$S_k = \eta_k I, \tag{14}$$

$$\mathbf{s}_k = \rho_{k+1}(\mu + \sigma_u^2\eta_{k+1})\mathbf{s}_{k+1}, \tag{15}$$

$$s_k = s_{k+1} - \frac{1}{2}\rho_{k+1}\mathbf{s}_{k+1}^T\mathbf{s}_{k+1}. \tag{16}$$

where η_k and ρ_k are scalars such that

$$\rho_k \triangleq \frac{1}{\mu + (1+\sigma_u^2)\eta_k}.$$

Note that under the assumptions of this section, we have $S_K = \lambda_f I$, $\mathbf{s}_K = -\lambda_f \mathbf{x}_f$ and $s_K = \frac{\lambda_f}{2}\mathbf{x}_f^T\mathbf{x}_f$.

To derive the equations 14, 15 and 16, we have used the simplified forms of A, B, Q, Q_f, R explained above, and have plugged them into Eqs. 10, 11, 12. Using the aforementioned form of the covariance matrices,

$$\sum_i F_{u_i}^T S_{k+1} F_{u_i} = \sigma_u^2 diag(S_{k+1}),$$

where $diag(M)$ takes a matrix M and outputs a diagonal matrix with the same entries as M on its diagonal. Then, since S_K is a multiple of the identity, an inductive reasoning with Eq. 10 shows that S_k is also a multiple of the identity for any $0 \le k < K$, for each one of the four terms in Eq. 10. i.e. S_k has the form $\eta_k I$. We deduce $H_k = (\mu + (1+\sigma_u^2)\eta_{k+1})I$ and by plugging the derived expressions into Eqs. 11 and 12, we deduce Eqs. 15 and 16.

Let us solve the expressions (14)–(16). Observe that by using the Eq. 9, one can express the scalar η_k recursively from the following expression

$$\begin{cases} \eta_K = \lambda_f \\ \eta_k = \lambda + \eta_{k+1}(1+\sigma_x^2) - \frac{\eta_{k+1}^2}{\mu + \eta_{k+1}(1+\sigma_u^2)}. \end{cases}$$

Similarly, if we solve the recursion in $\mathbf{s}_k$, we get:

$$\begin{aligned} \mathbf{s}_k &= \left(\prod_{i=k+1}^{K} \rho_i(\mu + \sigma_u^2\eta_i)\right)\mathbf{s}_K \\ &\triangleq -\lambda_f \phi_k \mathbf{x}_f. \end{aligned} \tag{17}$$

The initial value of ϕ_k is $\phi_K = 1$. The scalar s_k is resolved in the same way as:

$$s_k = \frac{\lambda_f}{2}\mathbf{x}_f^T\mathbf{x}_f\left(1 - \lambda_f\left(\sum_{i=k+1}^{K}\rho_i\phi_i^2\right)\right).$$

Finally, the optimal cost-to-go function at state $\mathbf{x}_k$ is given by

$$v_k^*(\mathbf{x}_k) = \frac{\lambda_f}{2}\mathbf{x}_f^T\mathbf{x}_f\left(1-\lambda_f\left(\sum_{i=k+1}^{K}\rho_i\phi_i^2\right)\right) - \lambda_f\phi_k\mathbf{x}_k^T\mathbf{x}_f + \frac{\eta_k}{2}\mathbf{x}_k^T\mathbf{x}_k. \quad (18)$$

To understand the behavior of this optimal cost-to-go function and of the corresponding optimal control, observe first that if $1 < k \ll K$ then we get $\phi_k = \prod_{i=k+1}^{K}\frac{\mu+\sigma_u^2\eta_i}{\mu+(1+\sigma_u^2)\eta_i} \approx 0$. This is because we can write ϕ_k as

$$\phi_k = \prod_{i=k+1}^{K}\frac{1}{1+\frac{1}{\frac{\mu}{\eta_i}+\sigma_u^2}} < \left(\frac{1}{1+\frac{1}{\frac{\mu}{\lambda}+\sigma_u^2}}\right)^{K-k}$$

by using the inequality $\eta_i > \lambda$ for $i > 0$. Note that as $K-k$ increases the term $\left(\frac{1}{1+\frac{1}{\frac{\mu}{\lambda}+\sigma_u^2}}\right)^{K-k}$ is smaller. This means that when the horizon is large enough, the linear terms in the beginning of the planning horizon tend to fade off. This has consequences on the associated optimal control, as described below.

3.3 Optimal control

Note that from the results presented above, one gets $H_k = \rho_{k+1}I$ and can also write the optimal control as the following affine function of the state:

$$\mathbf{u}_k^* = \lambda_f\rho_{k+1}\phi_{k+1}\mathbf{x}_f - \rho_{k+1}\eta_{k+1}\mathbf{x}_k, \quad (19)$$

which can also be written as:

$$\mathbf{u}_k^* = -\rho_{k+1}\phi_{k+1}\lambda_f(\mathbf{x}_k-\mathbf{x}_f) - \rho_{k+1}(\eta_{k+1}-\phi_{k+1}\lambda_f)\mathbf{x}_k \triangleq \alpha_{k+1}\left[\lambda_f(\mathbf{x}_k-\mathbf{x}_f)\right] + \beta_{k+1}\left[\lambda\mathbf{x}_k\right]. \quad (20)$$

The above equation represents the system optimal control, which is written as a positive linear combination of two gradients: 1) the one of the final cost function $\lambda_f\|\mathbf{x}-\mathbf{x_f}\|^2$ (driving the system close to the final position $\mathbf{x}_f$) and 2) the one of the instantaneous cost function $\lambda\|\mathbf{x}\|^2$ (driving the system close to the origin). Note that the contributions of each term in the optimal control depends on the following scalar values: λ, λ_f, μ, σ_u and σ_x. Also note that α_k and β_k vary with k. There are a few observations that can be done on these two terms. First, when the control time is far from the horizon (i.e., during the first executions of the planned controls $1 < k \ll K$), the first term (depending on ϕ_k) gets smaller, because ϕ_k tends to zero while ρ_k is bounded by $\frac{1}{\mu}$. This is consistent with the idea that, far from the final position, the constraint pulling the robot toward the goal does not apply. In that case, the second term will be close to

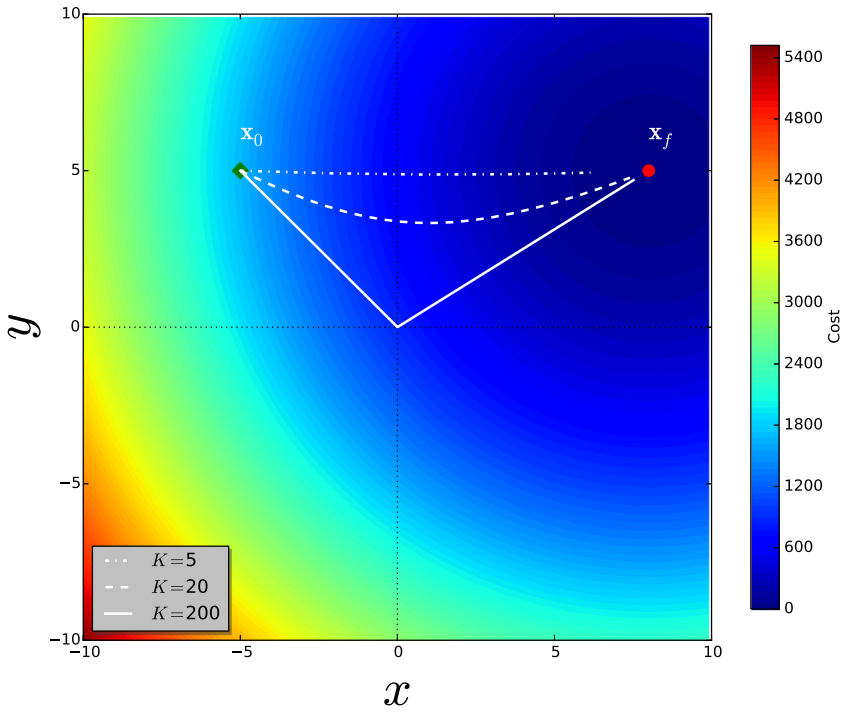

Fig. 1 Paths with optimal controls for different values of the planning horizon, with parameters $\mathbf{x}_0 = (-5, 5)$, $\mathbf{x}_f = (8, 5)$, $\lambda_f = 10$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.01$ and $\mu = 1.0$. The color map for the cost: $\lambda\mathbf{x}^T\mathbf{x} + \lambda_f(\mathbf{x}-\mathbf{x}_f)^T(\mathbf{x}-\mathbf{x}_f)$ identifies the states with high costs

$-\rho_{k+1}\eta_{k+1}\mathbf{x}_k$. Note that this instantaneous cost-related gain satisfies:

$$0 < \rho_{k+1}\eta_{k+1} < \frac{1}{1+\sigma_u^2}.$$

Hence, when $\sigma_u^2 \gg 1$, which means that the control amplitudes induce large values of noise, a conservative behavior is being applied, by reducing the gain.

In Fig. 1, one can observe the behavior described previously for one system, and under several different horizons K. We have chosen the values of $\lambda_f, \sigma_x, \sigma_u, \lambda, \mu$ in such a way that each term could have a similar importance in the total cost. With short horizons, the system is immediately brought to the desired final configuration, along a straight line. But as the planning horizon gets longer, the optimal controls first lead the system near the origin. If the planning horizon is long enough to achieve the condition $\phi_k \approx 0$, the state will remain at the origin for several consecutive time steps. This is because the term $\lambda_f \rho_{k+1}\phi_{k+1}\mathbf{x}_f \approx 0$ and is negligible as compared to $\rho_{k+1}\eta_{k+1}\mathbf{x}_k$, which is precisely the one that guides the system to the origin. Once $\phi_k \approx 0$ and $\mathbf{x}_k \approx 0$, it is important to note that the control keeps the system close to the origin and the cumulative cost tends to remain constant (see Figs. 2 and 3).

To further analyze Eq. 20, let us define $w_k = \frac{\eta_k}{\phi_k \lambda_f}$. Then Eq. 20 can be written

$$\mathbf{u}_k^* = -\rho_{k+1}\phi_{k+1}\lambda_f[(\mathbf{x}_k - \mathbf{x}_f) - (w_{k+1} - 1)\mathbf{x}_k].$$

Now let us examine the term w_k carefully. We have

$$w_k = \frac{\eta_k}{\phi_k \lambda_f} = \frac{\lambda + \eta_{k+1}(1+\sigma_x^2) - \frac{\eta_{k+1}^2}{\mu+(1+\sigma_u^2)\eta_{k+1}}}{\phi_{k+1}\frac{\mu+\sigma_u^2\eta_{k+1}}{\mu+(1+\sigma_u^2)\eta_{k+1}}\lambda_f}.$$

This expression can be split in three

$$w_k = \frac{\lambda}{\phi_k\lambda_f} + \frac{\eta_{k+1}\sigma_x^2}{\phi_k\lambda_f} + \frac{\eta_{k+1} - \frac{\eta_{k+1}^2}{\mu+(1+\sigma_u^2)\eta_{k+1}}}{\phi_{k+1}\frac{\mu+\sigma_u^2\eta_{k+1}}{\mu+(1+\sigma_u^2)\eta_{k+1}}\lambda_f},$$

which can be transformed into

$$w_k = \frac{\lambda}{\phi_k\lambda_f} + \frac{\eta_{k+1}\sigma_x^2}{\phi_k\lambda_f} + \frac{\eta_{k+1}}{\phi_{k+1}\lambda_f} = \frac{\lambda}{\phi_k\lambda_f} + \frac{\eta_{k+1}\sigma_x^2}{\phi_k\lambda_f} + w_{k+1}.$$

We deduce:

$$w_k = \frac{\lambda}{\lambda_f}\left[\frac{1}{\phi_k} + \frac{1}{\phi_{k+1}} \ldots + \frac{1}{\phi_{K-1}}\right] + \frac{\sigma_x^2}{\lambda_f}\left[\frac{\eta_{k+1}}{\phi_k} + \frac{\eta_{k+2}}{\phi_{k+1}} \ldots + \frac{\eta_K}{\phi_{K-1}}\right] + w_K.$$

Now, as we saw above,

$$\left(\frac{1}{1+\frac{1}{\sigma_u^2}}\right)^{K-k} < \phi_k < \left(\frac{1}{1+\frac{1}{\frac{\mu}{\lambda}+\sigma_u^2}}\right)^{K-k}. \tag{21}$$

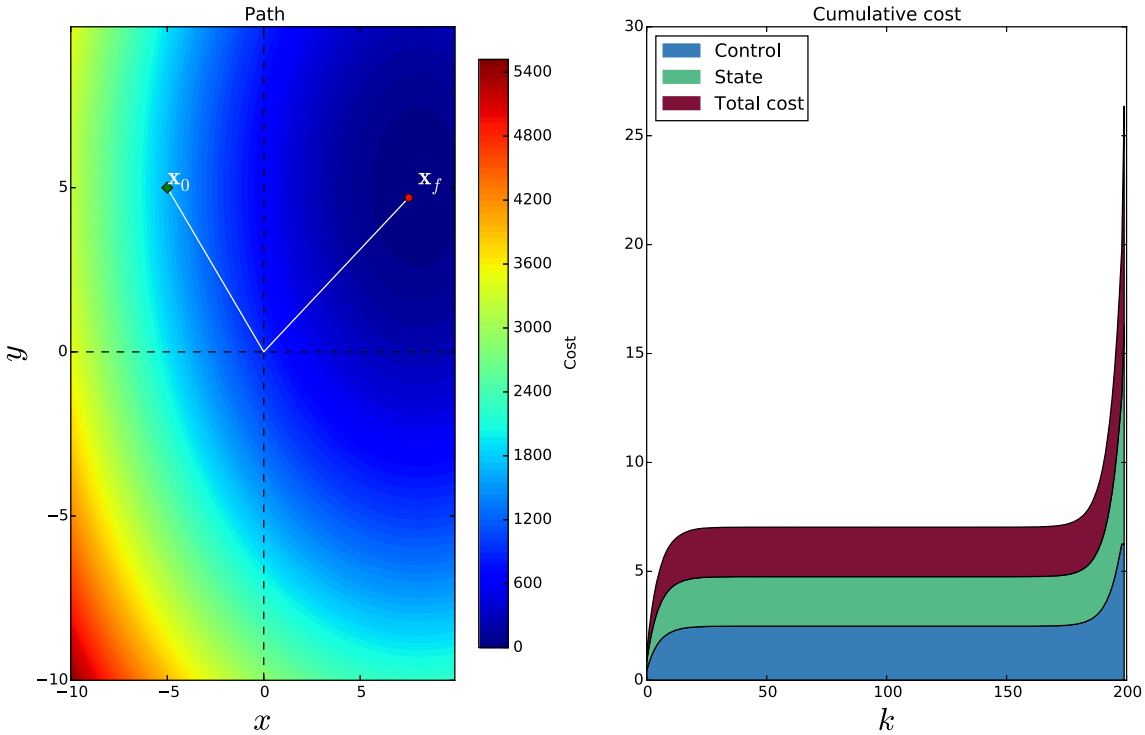

Fig. 2 System behavior, for $K = 200$. The parameters have the following values: $\mathbf{x}_0 = (-5, 5)$, $\mathbf{x}_f = (8, 5)$, $\lambda_f = 10$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.01$ and $\mu = 1.0$. Left: path and color map for the costs: $\lambda\mathbf{x}^T\mathbf{x} + \lambda_f(\mathbf{x} - \mathbf{x}_f)^T(\mathbf{x} - \mathbf{x}_f)$. Right: accumulated cost

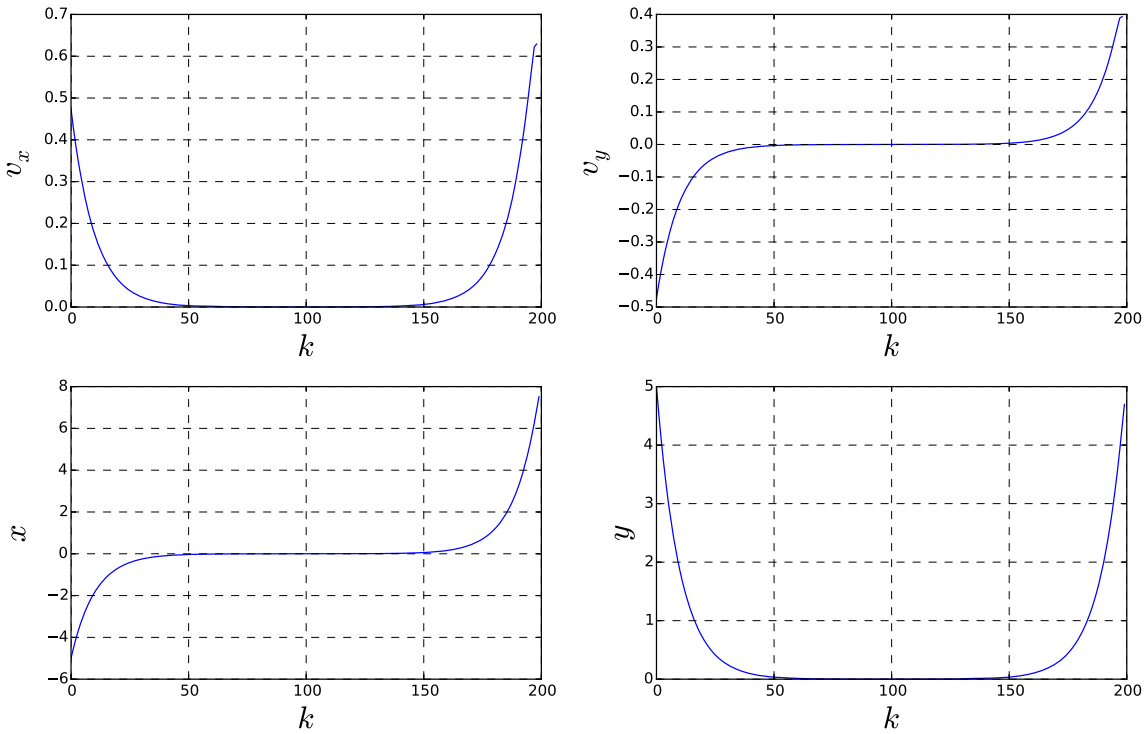

Fig. 3 System behavior, for $K = 200$. The parameters have the following values: $\mathbf{x}_0 = (-5, 5)$, $\mathbf{x}_f = (8, 5)$, $\lambda_f = 10$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.01$ and $\mu = 1.0$. Top: behavior of the optimal controls (in x and y). Bottom: trajectory for x and y

Also, in situations where σ_x is not too large (as we will see in the next Section), for all k, η_k is bounded, i.e. there exists values $\eta^-, \eta^+ > 0$ such that for all k, $\eta^- < \eta_k < \eta^+$. With these elements in mind, and after some developments,

$$\frac{\lambda + \sigma_x^2 \eta^-}{\lambda_f} \left(\frac{\mu}{\lambda} + \sigma_u^2 + 1\right) \left[\left(1 + \frac{1}{\frac{\mu}{\lambda} + \sigma_u^2}\right)^{K-k} - 1\right] < w_k - 1 < \frac{\lambda + \sigma_x^2 \eta^+}{\lambda_f} (\sigma_u^2 + 1) \left[\left(1 + \frac{1}{\sigma_u^2}\right)^{K-k} - 1\right].$$

This last equation shows that w_k grows exponentially with $K - k$, and gives a lower and an upper bound by two exponential functions. This means that, in Eq. 20, the proportion of the goal-driven control (the term in $\mathbf{x}_k - \mathbf{x}_f$) with respect to the second term goes shrinking exponentially as $\frac{1}{w_k - 1}$. In particular, for $k = 0$,

$$\frac{\lambda + \sigma_x^2 \eta^-}{\lambda_f} \left(\frac{\mu}{\lambda} + \sigma_u^2 + 1\right) \left[\left(1 + \frac{1}{\frac{\mu}{\lambda} + \sigma_u^2}\right)^{K} - 1\right] < w_0 - 1 < \frac{\lambda + \sigma_x^2 \eta^+}{\lambda_f} (\sigma_u^2 + 1) \left[\left(1 + \frac{1}{\sigma_u^2}\right)^{K} - 1\right]. \quad (22)$$

3.4 Variations of η_k with k

At this point the reader could note that η_k has a classical formulation as the iteration of a recursive function $\eta_{k+1} = f(\eta_k)$, with

$$f(\eta) = \lambda + \eta(1 + \sigma_x^2) - \frac{\eta^2}{\mu + \eta(1 + \sigma_u^2)}. \quad (23)$$

This equation can also be written as

$$f(\eta) = \lambda + \eta\sigma_x^2 + \frac{\eta(\mu + \eta\sigma_u^2)}{\mu + \eta(1 + \sigma_u^2)}. \quad (24)$$

It should also be noted that its derivative is

$$f'(\eta) = \sigma_x^2 + \frac{(\mu + \eta\sigma_u^2)^2 + \sigma_u^2 \eta^2}{(\mu + \eta(1 + \sigma_u^2))^2},$$

so f is a strictly increasing function. A natural question to answer is whether this function admits a fixed point (i.e., a point of convergence for η_k when considering $K \to \infty$); this point is examined in the next subsection.

3.5 Convergence Properties and Steady State Values for η_k

An interesting question to answer is whether a fixed point η^* exists for f, which has an important implication on the cost related to a motion policy. If a fixed point exists then the cost is bounded by a finite value for states within a bounded

domain and hence one can evaluate whether or not, we can afford to pay it.

A sufficient condition for the existence of η^*, is that f is Lipschitz continuous with a Lipschitz constant inferior to one, i.e. there should exist L such that $\forall \eta \in \mathbb{R}^+ \; |f'(\eta)| \leq L < 1$, which also means there should exist $\epsilon > 0$ such that $\forall \eta \in \mathbb{R}^+ \; |f'(\eta)| \leq 1 - \epsilon$. Note that from the previous subsection, $\eta > 0$ and $f'(\eta) > 0$. With the previous expression for $f'(\eta)$, the sufficient condition translates into

$$\exists \epsilon > 0 \quad \forall \eta \quad \sigma_x^2 + \epsilon \leq 1 - \frac{(\mu + \eta\sigma_u^2)^2 + \sigma_u^2\eta^2}{(\mu + \eta(1+\sigma_u^2))^2}$$

and finally

$$\exists \epsilon > 0 \quad \forall \eta \quad \sigma_x^2 + \epsilon \leq \eta \frac{2\mu + (1+\sigma_u^2)\eta}{(\mu + (1+\sigma_u^2)\eta)^2}. \tag{25}$$

Proposition 1 *When $\sigma_x > 0$, a sufficient condition for the sequence of η_k's to converge to a finite value when k goes to infinity is*

$$0 < \sigma_x^2 < \lambda \frac{2\mu + (1+\sigma_u^2)\lambda}{(\mu + (1+\sigma_u^2)\lambda)^2}. \tag{26}$$

Proof First, note that $\eta_k > 0$ for all $k \geq 0$. Also, from Eq. 24, one can show that $\eta_k > \lambda$ for all $k > 0$. Also, the function on the right term of condition (25) is increasing on $\mathbb{R}^+$. Since for all $k > 0$, $\eta_k > \lambda$ then the right term is always superior to its value in λ, $\lambda \frac{2\mu+(1+\sigma_u^2)\lambda}{(\mu+(1+\sigma_u^2)\lambda)^2}$. Note that $\lambda \frac{2\mu+(1+\sigma_u^2)\lambda}{(\mu+(1+\sigma_u^2)\lambda)^2} < \frac{\mu^2+2\mu(1+\sigma_u^2)\lambda+(1+\sigma_u^2)^2\lambda^2}{(\mu+(1+\sigma_u^2)\lambda)^2} = 1$ so that $1 > \lambda \frac{2\mu+(1+\sigma_u^2)\lambda}{(\mu+(1+\sigma_u^2)\lambda)^2} > \sigma_x^2$. This implies $|f'(\eta)| \leq 1 - \epsilon$ where $\epsilon = \lambda \frac{2\mu+(1+\sigma_u^2)\lambda}{(\mu+(1+\sigma_u^2)\lambda)^2} - \sigma_x^2$. The assumption of this proposition ensures the convergence of the recursion $\eta_{k+1} = f(\eta_k)$ to a unique value η^*. □

Corollary 1 *When $\sigma_x = 0$ (no additive noise depending on the state), the sequence of η_k's converges to a finite value when k goes to infinity.*

Proof In that case, we have $|f'(\eta)| \leq 1 - \epsilon$ where $\epsilon = \lambda \frac{2\mu+(1+\sigma_u^2)\lambda}{(\mu+(1+\sigma_u^2)\lambda)^2} < 1$, which ensures convergence of the iterative procedure. □

Note that the sufficient condition of Proposition 1 also implies:

$$0 < \sigma_x^2 < \frac{1}{(1+\sigma_u^2)}. \tag{27}$$

From now on, we will suppose that the condition enounced in Proposition 1 is satisfied. If the fixed point of f exists then it satisfies the second degree equation in η induced by the recursion, so it has to be $\eta = \eta^*$ with:

$$\eta^* = \frac{(\lambda(1+\sigma_u^2) + \sigma_x^2\mu) + \sqrt{(\lambda(1+\sigma_u^2) + \sigma_x^2\mu)^2 + 4\lambda\mu(1 - \sigma_x^2(1+\sigma_u^2))}}{2(1 - \sigma_x^2(1+\sigma_u^2))}. \tag{28}$$

Note that from the comments above, $\sigma_x^2(1+\sigma_u^2) < 1$, hence both the numerator and denominator in the expression above are positive. Also, we have bounded values for σ_x^2. Observe that if σ_u^2 becomes larger, then the upper bound value for σ_x^2 decreases and may tend to zero.

Figure 4 shows the growing trend of η^* as σ_x approaches its upper limit. This means that, when considering an inifinite horizon, the quadratic coefficient of the cost function is unbounded while σ_x is increased.

In addition, because f is increasing, there are only two cases of evolutions of the sequence η_k, provided that the previous sufficient condition is satisfied. Recall that the first value of the sequence is $\eta_K = \lambda_f$, i.e. the term associated to the quadratic cost at the final position. Two cases are possible: (1) for low values of the final quadratic coefficient λ_f ($\lambda_f < \eta^*$), the cost-to-go quadratic coefficient η_k goes increasing up to convergence; (2) for large values of

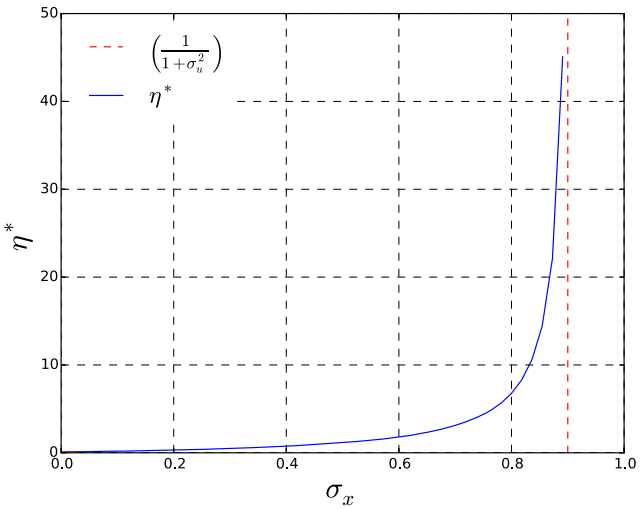

Fig. 4 Behavior of η^* for increasing values of σ_x^2 with $\sigma_u^2 = 0.1$

the final quadratic coefficient λ_f ($\lambda_f > \eta^*$), the values of η_k decrease over time. In each of these cases, the interval of the possible values of η is bounded according to:

$$\begin{cases} \eta \in [\lambda_f, \eta^*] \text{ if } \lambda > -\sigma_x^2 \lambda_f + \frac{\lambda_f^2}{\mu + (1+\sigma_u^2)\lambda_f} \text{ (case 1)}, \\ \eta \in [\eta^*, \lambda_f] \text{ otherwise (case 2)}. \end{cases} \tag{29}$$

Note that we have here the bounds η^-, η^+ mentionned above in Eq. 22. Until now we only provided sufficient conditions for the existence of the fixed point. If the fixed point exists, then it is also important to estimate its rate of convergence. We want the minimum number of iterations k^* needed to get:

$$||\eta_{k^*} - \eta^*|| \leq \tau,$$

where τ is a given bound. Using the Contraction Mapping Theorem, we can write:

$$||\eta_{k^*} - \eta^*|| \leq L^{k^*} ||\lambda_f - \eta^*|| \leq \tau.$$

Applying the logarithm to the left side and defining L as the upper bound on $f'(\eta)$ defined in Proposition 1, one can estimate a lower bound k^*

$$k^* \geq \frac{\ln\left(\frac{\tau}{|\eta^* - \lambda_f|}\right)}{\ln(L)}. \tag{30}$$

In Fig. 5, we can observe the two convergence behaviors for η_k. Although $f(\eta)$ is increasing with η, $f(\eta_k)$ could be either increasing or decreasing with k. In Fig. 5, left, η_k decreases with k while in Fig. 5, right, $f(\eta_k)$ increases with k.

3.6 Implications of the Convergence

Consider the optimal control (20). As $\rho_{k+1}\eta_{k+1} = \frac{\eta_{k+1}}{\mu + (1+\sigma_u^2)\eta_{k+1}}$, one can deduce, that, when the convergence conditions described above hold, then if we consider large enough horizon values, we will have for the first planning steps

$$\mathbf{u}_k^* \approx -\frac{\eta^*}{\mu + (1+\sigma_u^2)\eta^*} \mathbf{x}_k. \tag{31}$$

This means that, in the case the convergence sufficient conditions are met then it is possible to have a good approximation of the function value and of the optimal controls for large enough horizon values.

Hence, there are two significant advantages in the convergence of the value function coefficients. First, we have numerically bounded costs when considering a bounded domain, and, second, we can estimate in advance the behavior of the optimal controls at the convergence horizon. Given these advantages, one can make decisions on the planning horizon and for example rule out the horizons $K > k^*$. Another natural question is which of these horizons is the most suitable in terms of the obtained final cost. This will be the object of Section 4.

3.7 Behavior of the Solutions in Extremal Cases

The system introduced in Section 2 includes different cost terms: one depending on the state values, one depending on the control values and one on the final state values. For the

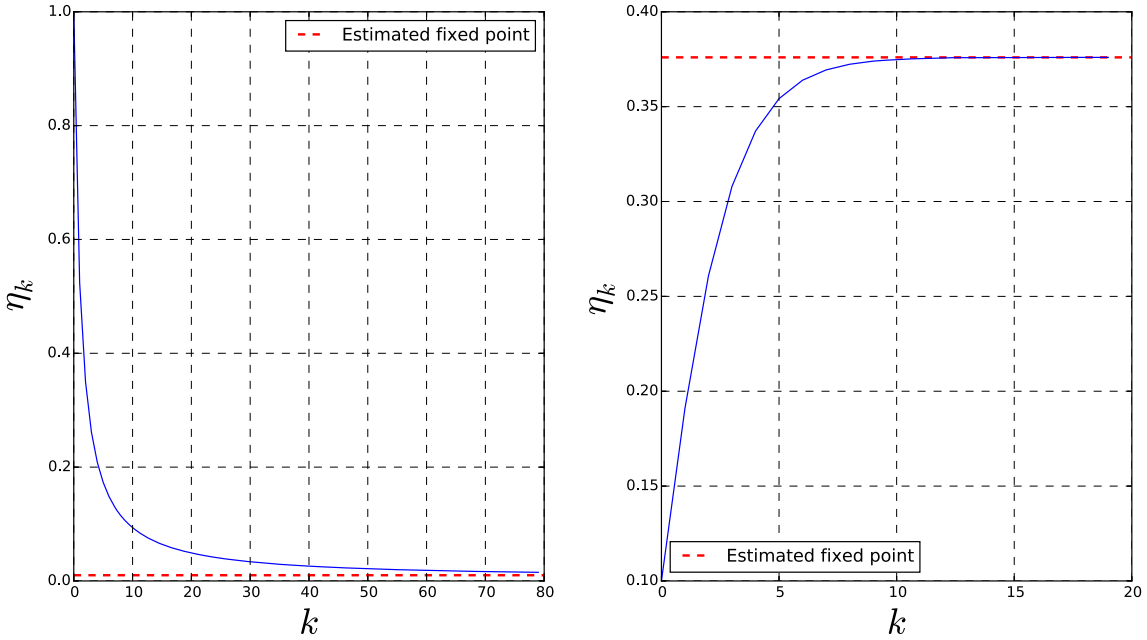

Fig. 5 Convergence cases: (Left) in interval $[\eta^*, \lambda_f]$, (Right) in interval $[\lambda_f, \eta^*]$

complete understanding of the system and of its solutions, we study here the behavior of the system for extremal values of its different parameters.

3.7.1 Predominant Instantaneous State Cost

In this section we consider the case where the value of the costs associated to the system states are high compared to the other terms. To quantify this notion properly, let us come back to Eq. 20. According to Section 3.3, when k grows, the term $w_k - 1$ will be smaller. This means that in the latest steps of our plan, $w_k - 1$ will have its lowest values. To ensure that the term associated with the state dominates the goal-related term throughout all the plan, we should have:

$$|w_{K-1} - 1| \gg 1.$$

A sufficient condition for this to occur can be derived from Eq. 22:

$$\frac{\lambda+\sigma_x^2\eta^-}{\lambda_f}\left(\frac{\mu}{\lambda}+\sigma_u^2+1\right)\left[\left(1+\frac{1}{\frac{\mu}{\lambda}+\sigma_u^2}\right)^{K-(K-1)}-1\right]\gg 1,$$

which leads to:

$$\lambda + \sigma_x^2\eta^- \gg \lambda_f\left(\frac{\mu+\lambda(\sigma_u^2+1)}{\mu+\lambda\sigma_u^2}\right).$$

To get this condition verified, we may set for example:

$$\lambda \gg \max\left(\frac{\mu}{\sigma_u^2}, \lambda_f\left(1+\frac{1}{\sigma_u^2}\right)\right), \tag{32}$$

which intuitively means that λ should be much larger than both μ and λ_f. If the above condition is satisfied, then, after some algebraic manipulations, it is equivalent to:

$$\mu \ll \lambda\sigma_u^2 < \eta_k\sigma_u^2,$$

then $\mu \ll \eta_k(1+\sigma_u^2)$. Note that this condition is independent of the selected planning horizon. From Eq. 20, the control in that case takes the simple form:

$$\mathbf{u}_k^* \approx -\frac{\eta_k}{\mu+\eta_k(1+\sigma_u^2)}\mathbf{x}_k.$$

As $\eta_k > \lambda$, then $\mu \ll \eta_k(1+\sigma_u^2)$ and the control can be approximated by:

$$\mathbf{u}_k^* \approx -\frac{1}{1+\sigma_u^2}\mathbf{x}_k, \tag{33}$$

which means that the state is driven to the origin, i.e. to the lowest values of the instantaneous state-dependent cost function. Note that the corresponding gain is $\frac{1}{1+\sigma_u^2}$, which only depends on the control-depending noise level and which takes lower values with higher noise coefficient, i.e., for large σ_u, smaller motions are preferred so as not to increment the cost function because of risky controls. Note that the optimal policy of Eq. 33 is the same as the one resulting from:

$$\min_{\mathbf{u}_{0:K-1}} E_{\xi_{\mathbf{u}_0},\xi_{\mathbf{x}_0},\ldots,\xi_{\mathbf{u}_{K-1}},\xi_{\mathbf{x}_{K-1}}}\left(\sum_{k=0}^{K-1}\mathbf{x}_k^T Q\mathbf{x}_k\right).$$

3.7.2 Predominant Goal Cost

Again, if one refers to Eq. 20 above, one can see that to have a significant part of the goal cost in the overall cost since the beginning of the planning horizon, then the proportion of the optimal control arising from the goal cost should dominate the state dependent cost term, starting from $k = 0$. This implies:

$$|w_0 - 1| \ll 1.$$

Because of the exponential nature of w_k with respect to $K - k$, and by taking into account Eq. 22, we have a sufficient condition for this to occur, which is:

$$\frac{\lambda(1+\sigma_x^2\eta^+)}{\lambda_f}(\sigma_u^2+1)\left[\left(1+\frac{1}{\sigma_u^2}\right)^K-1\right]\ll 1.$$

This may be satisfied by taking large enough values for λ_f:

$$\lambda_f \gg \lambda(1+\sigma_x^2\eta^+)(\sigma_u^2+1)\left(1+\frac{1}{\sigma_u^2}\right)^K.$$

The control under these conditions will be:

$$\mathbf{u}_k^* \approx -\rho_{k+1}\phi_{k+1}\lambda_f(\mathbf{x}_k-\mathbf{x}_f) = -\frac{\frac{1}{w_k}}{\frac{\mu}{\eta_{k+1}}+1+\sigma_u^2}(\mathbf{x}_k-\mathbf{x}_f).$$

When we have also $\lambda_f \gg \frac{\mu}{\sigma_u^2}$, then one can use Eq. 24 to show that:

$$\eta_k \approx \lambda_f\left(\frac{1}{1+\frac{1}{\sigma_u^2}}\right)^k,$$

so that the optimal control expression can be approximated by

$$\mathbf{u}_k^* \approx -\frac{1}{1+\sigma_u^2}(\mathbf{x}_k - \mathbf{x}_f). \tag{34}$$

Note that the solution for this case corresponds to leading the system states to the final configuration. Again, the weight associated with the optimal control depends on the noise coefficient σ_u^2. If it is null, the computed control leads the system to the final configuration, which remains there until the end of the planning horizon. Otherwise, the system gets closer to the final configuration at slower rates.

These high values of λ_f correspond to the minimization of the cost function:

$$\min_{\mathbf{u}_{0:K-1}} E_{\xi_{\mathbf{u}_0},\xi_{\mathbf{x}_0},\ldots,\xi_{\mathbf{u}_{K-1}},\xi_{\mathbf{x}_{K-1}}}(l_K(\mathbf{x}_K)).$$

3.7.3 Predominant Control Penalty Cost, Without State-Dependent Noise

Again, the optimal control expression of Eq. 20 can be written as

$$\mathbf{u}_k^* = -\rho_{k+1}\phi_{k+1}\lambda_f w_{k+1}\left[\frac{1}{w_{k+1}}(\mathbf{x}_k - \mathbf{x}_f) - \frac{w_{k+1}-1}{w_{k+1}}\mathbf{x}_k\right].$$

The weights in the sum are between 0 and 1, and the term $\rho_{k+1}\phi_{k+1}\lambda_f w_{k+1}$ is a global factor over this control. Now note that when considering $\sigma_x^2 = 0$ (no state-dependent noise), we get from Eq. 22:

$$\rho_{k+1}\phi_{k+1}\lambda_f w_{k+1} < \frac{1}{\mu}\lambda_f\left(1 + \frac{\lambda}{\lambda_f}(\sigma_u^2+1) \times \left[\left(1+\frac{1}{\sigma_u^2}\right)^{K-k-1} - 1\right]\right).$$

Hence, with large enough values of μ:

$$\mu \gg \lambda_f\left(1 + \frac{\lambda}{\lambda_f}(\sigma_u^2+1)\left[\left(1+\frac{1}{\sigma_u^2}\right)^{K-1} - 1\right]\right),$$

the optimal control can be made as small as desired

$$\mathbf{u}_k^* \to 0. \tag{35}$$

This control keeps the system in its initial configuration and also corresponds to the solution to the problem:

$$\min_{\mathbf{u}_{0:K-1}} E_{\xi_{\mathbf{u}_0},\xi_{\mathbf{x}_0},\ldots,\xi_{\mathbf{u}_{K-1}},\xi_{\mathbf{x}_{K-1}}}\left(\sum_{k=0}^{K-1}\mathbf{u}_k^T R\mathbf{u}_k\right).$$

3.7.4 High Values of Control-Dependent Noise, Without State-Dependent Noise

As mentioned above, the control-dependent noise does have an impact on the obtained optimal controls. To follow this analysis, observe that the optimal controls behavior when σ_u^2 tends to infinity when, again, $\sigma_x^2 = 0$ (no state-dependent noise). First, note that from Eq. 21, ϕ_k tends to one; also, since η_k has λ as a lower bound,

$$\lim_{\sigma_u^2\to\infty}\left(\frac{1}{\mu + (1+\sigma_u^2)\eta_k}\right) = 0.$$

Hence,

$$\rho_{k+1}\phi_{k+1}\lambda_f w_{k+1} < \frac{1}{\sigma_u^2\lambda}\lambda_f\left(1 + \frac{\lambda}{\lambda_f}(\sigma_u^2+1) \times \left[\left(1+\frac{1}{\sigma_u^2}\right)^{K-k-1} - 1\right]\right).$$

Since the second part of this expression can be bounded when σ_u^2 tends to infinity, we obtain the optimal controls at the limit:

$$\mathbf{u}_k^* \to 0.$$

Interestingly this control is equivalent to the solution obtained for large values of μ. This result is intuitive considering that the expected state resulting from applying a noisy control may be highly expensive.

The following table summarizes the different cases described in this section and the optimal controls for each of them.

Dominant parameter	Sufficient condition	Optimal control	Interpretation
λ	$\lambda \gg \max\left(\frac{\mu}{\sigma_u^2}, \lambda_f\left(1+\frac{1}{\sigma_u^2}\right)\right)$,	$\mathbf{u}_k^* \approx -\frac{1}{1+\sigma_u^2}\mathbf{x}_k$	The control leads the system to the origin
μ	$\mu \gg, \lambda_f\left(1+\frac{\lambda}{\lambda_f}(\sigma_u^2+1)\left[\left(1+\frac{1}{\sigma_u^2}\right)^{K-1}-1\right]\right)$,	$\mathbf{u}_k^* \to 0$	The control keeps the system at its initial configuration
λ_f	$\lambda_f \gg \lambda\left(1+\sigma_x^2\eta^+\right)(\sigma_u^2+1)\left(1+\frac{1}{\sigma_u^2}\right)^K$	$\mathbf{u}_k^* \approx -\frac{1}{1+\sigma_u^2}(\mathbf{x}_k - \mathbf{x}_f)$	The control leads the system to the final configuration
σ_u^2	$\sigma_u^2 \to \infty, \quad \sigma_x^2 = 0$	$\mathbf{u}_k^* \to 0$	The control keeps the system at its initial configuration

4 Horizon Selection Without State-Depending Noise

Here, we are interested in describing the behavior of the total expected cost function $v_0(\mathbf{x}_0)$ with increasing horizon times K, and with fixed initial and goal configurations $\mathbf{x}_0$ and $\mathbf{x}_f$. We want to compare the benefits of selecting shorter or longer horizons for a specific planning problem. First, we study the behavior of the derivative of $v_0(\mathbf{x}_0)$ with respect to the horizon K and interpret its variations. We will show that in absence of state-dependent noise ($\sigma_x = 0$) the behavior of $v_0(\mathbf{x}_0)$ can be characterized in a geometric way.

In the next equations, we use a new notation, η_0^K and ϕ_0^K, to refer to the already defined quantities η_k and ϕ_k, evaluated at the time step $k = 0$ (index) and computed for horizon K (superscript). The optimal expected cumulative cost varies with the planning horizon, therefore we use a superscript K to indicate the planning horizon, and the index $k = 0$, as a subscript, to specify that we estimate the total expected cost, starting from the initial configuration. Besides, we use the notation Δ to indicate the difference between the expected total cost of two consecutive planning horizons, i.e. $\Delta\left[v_0^K(\mathbf{x}_0)\right] \stackrel{def}{=} v_0^K(\mathbf{x}_0) - v_0^{K-1}(\mathbf{x}_0)$.

By taking the difference between the total expected cost function evaluated at consecutive values $K-1$ and K, we get:

$$\Delta\left[v_0^K(\mathbf{x}_0)\right] = \frac{1}{2}(\eta_0^K - \eta_0^{K-1})\mathbf{x}_0^T\mathbf{x}_0 - (\phi_0^K - \phi_0^{K-1})\lambda_f\mathbf{x}_0^T\mathbf{x}_f - \frac{1}{2}\lambda_f^2\rho_0^K(\phi_0^K)^2\mathbf{x}_f^T\mathbf{x}_f. \quad (36)$$

By using the recurrence relations seen above in Eqs. 14 and 15, we can write the differences in the previous equation as functions of η_0^K as follows:

$$\begin{aligned}\Delta\eta_0^K &= \eta_0^K - \eta_0^{K-1} = \lambda - {\eta_0^K}^2\rho_0^K \\ \Delta\phi_0^K &= \phi_0^K - \phi_0^{K-1} = -\eta_0^K\rho_0^K\phi_0^K.\end{aligned} \quad (37)$$

Using the previous two equalities, one can rewrite the Eq. 36 as:

$$\Delta\left[v_0^K(\mathbf{x}_0)\right] = \frac{1}{2}\lambda\mathbf{x}_0^T\mathbf{x}_0 - \frac{1}{2}\rho_0^K((\eta_0^K)^2\mathbf{x}_0^T\mathbf{x}_0 - 2\phi_0^K\eta_0^K\lambda_f\mathbf{x}_0^T\mathbf{x}_f + \lambda_f^2(\phi_0^K)^2\mathbf{x}_f^T\mathbf{x}_f).$$

The second term in the previous equation can be written as the matrix product:

$$\left[\sqrt{\rho_0^K}\eta_0^K \;\; \sqrt{\rho_0^K}\phi_0^K\right]\begin{bmatrix}\mathbf{x}_0^T\mathbf{x}_0 & -\lambda_f\mathbf{x}_0^T\mathbf{x}_f \\ -\lambda_f\mathbf{x}_0^T\mathbf{x}_f & \lambda_f^2\mathbf{x}_f^T\mathbf{x}_f\end{bmatrix}\begin{bmatrix}\sqrt{\rho_0^K}\eta_0^K \\ \sqrt{\rho_0^K}\phi_0^K\end{bmatrix},$$

which allows to write:

$$\Delta\left[v_0^K(\mathbf{x}_0)\right] = \frac{1}{2}\lambda\mathbf{x}_0^T\mathbf{x}_0 - \frac{1}{2}\left[\sqrt{\rho_0^K}\eta_0^K \;\; \sqrt{\rho_0^K}\phi_0^K\right] \times \begin{bmatrix}\mathbf{x}_0^T\mathbf{x}_0 & -\lambda_f\mathbf{x}_0^T\mathbf{x}_f \\ -\lambda_f\mathbf{x}_0^T\mathbf{x}_f & \lambda_f^2\mathbf{x}_f^T\mathbf{x}_f\end{bmatrix}\begin{bmatrix}\sqrt{\rho_0^K}\eta_0^K \\ \sqrt{\rho_0^K}\phi_0^K\end{bmatrix}. \quad (38)$$

Note that the Eq. 38 is quadratic in the variables $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$, and that the other parameters do not depend on K. In particular, if there are critical points, they are located on the ellipse defined on the plane $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ by

$$\left[\sqrt{\rho_0^K}\eta_0^K \;\; \sqrt{\rho_0^K}\phi_0^K\right]\begin{bmatrix}\mathbf{x}_0^T\mathbf{x}_0 & -\lambda_f\mathbf{x}_0^T\mathbf{x}_f \\ -\lambda_f\mathbf{x}_0^T\mathbf{x}_f & \lambda_f^2\mathbf{x}_f^T\mathbf{x}_f\end{bmatrix} \times \begin{bmatrix}\sqrt{\rho_0^K}\eta_0^K \\ \sqrt{\rho_0^K}\phi_0^K\end{bmatrix} = \lambda\mathbf{x}_0^T\mathbf{x}_0. \quad (39)$$

From this geometrical object, it is possible to estimate the sign of $\Delta\left[v_0^K(\mathbf{x}_0)\right]$ and thus to understand the behavior of $v_0^K(\mathbf{x}_0)$. For example, if $\Delta\left[v_0^K(\mathbf{x}_0)\right] < 0$ then $v_0^K(\mathbf{x}_0)$ is decreasing. All the points that are outside the ellipse satisfy this inequality. Of course, there are two other possible

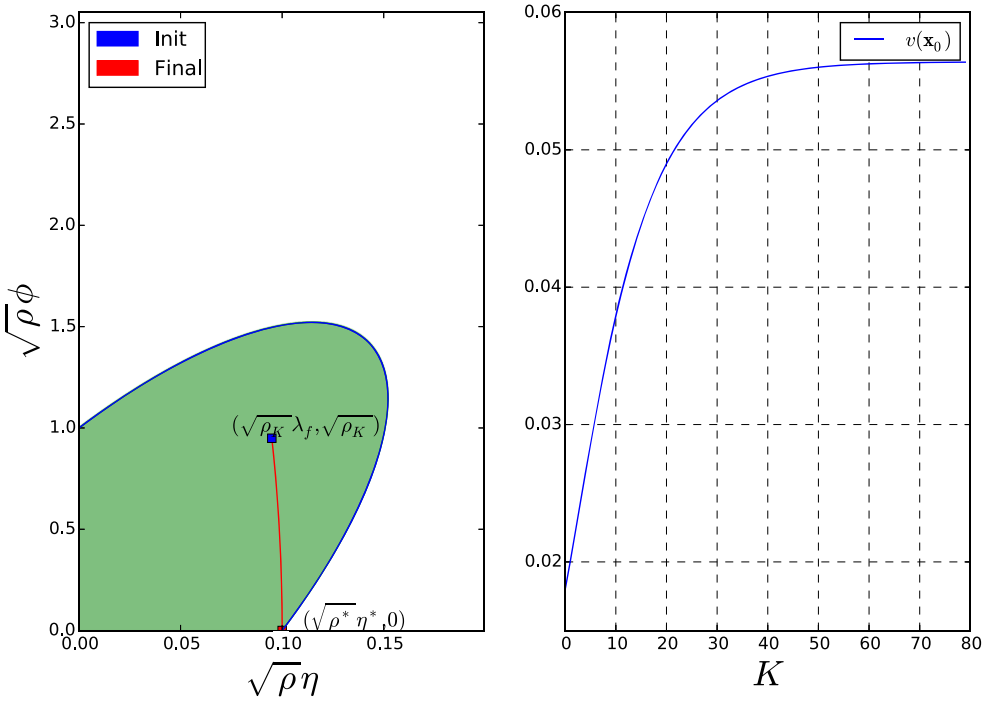

Fig. 6 Behavior of $v_0^K(\mathbf{x}_0)$ for differents horizons (case 1). Left: parameterization in the $(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K)$ space with parameters $\mathbf{x}_0 = (-0.3, 0.8)$, $\mathbf{x}_f = (0.3, 0.8)$, $\lambda_f = 0.1$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.01$ and $\mu = 1.0$. Right: function value at $\mathbf{x}_0$, varying with K

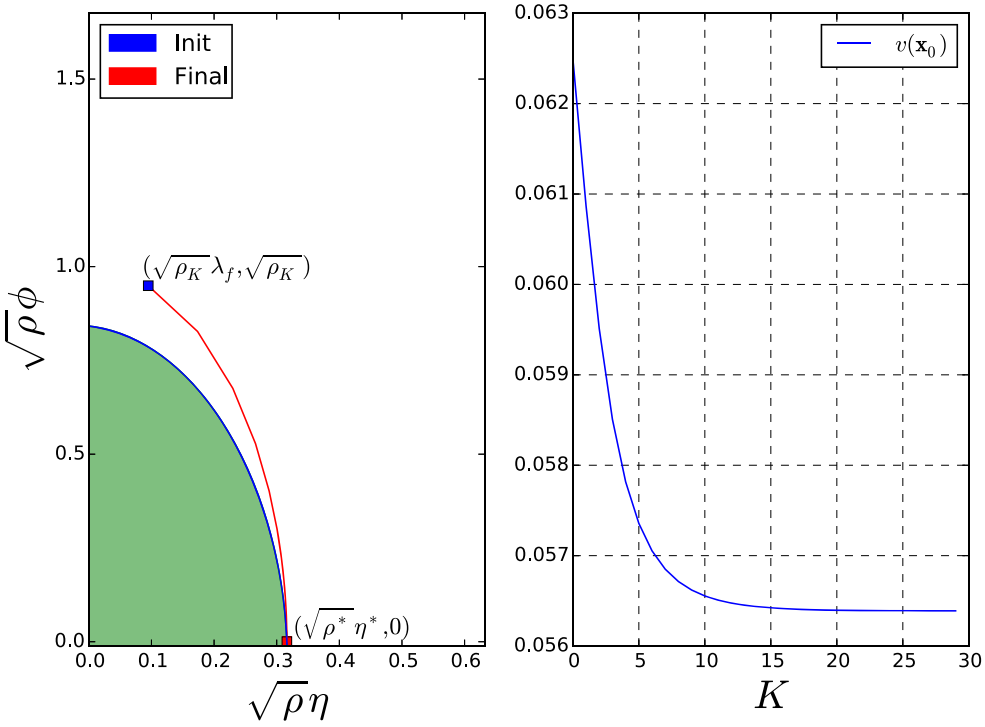

Fig. 7 Behavior of $v_0^K(\mathbf{x}_0)$ for differents horizons (case 2). Left: parameterization in the $(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K)$ space with parameters $\mathbf{x}_0 = (0.20, 0.20)$, $\mathbf{x}_f = (0.7, -0.8)$, $\lambda_f = 0.1$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.1$ and $\mu = 1.0$. Right: function value at $\mathbf{x}_0$, varying with K

cases. The second case corresponds to $\Delta\left[v_0^K(\mathbf{x}_0)\right] = 0$ and as we mentioned before, this case indicates the existence of a critical point in $v_0^K(\mathbf{x}_0)$. The last case occurs when $\Delta\left[v_0^K(\mathbf{x}_0)\right] > 0$. All the points inside the ellipse should satisfy this inequality. This case indicates that $v_0^K(\mathbf{x}_0)$ is increasing for horizon K.

Note that $\mathbf{x}_0$ and $\mathbf{x}_f$ are the initial and the final configuration of the system in the state space. The choice of the initial and final configuration, together with λ, λ_f, μ and σ_u, determines the geometry of an ellipse in the space $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ (as it was mentioned above). By varying the values of these parameters, we have found experimentally 4 basic behaviors, which we describe below.

In Fig. 6, we show the behavior of $v_0^K(\mathbf{x}_0)$ for the first case, where $v_0^K(\mathbf{x}_0)$ is a strictly increasing function of K. This occurs because applying controls with more planning steps gives higher overall costs. Hence, we go as quickly as

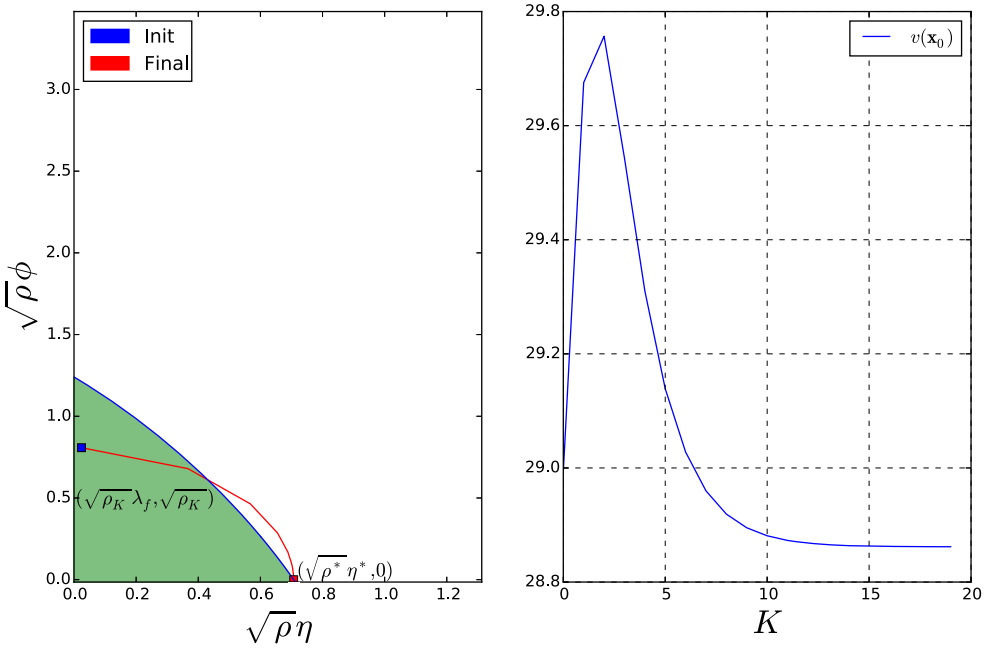

Fig. 8 Behavior of $v_0^K(\mathbf{x}_0)$ for differents horizons (case 3). Left: parameterization in the $(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K)$ space with parameters $\mathbf{x}_0 = (1.0, 2.0)$, $\mathbf{x}_f = (16.87, -39.0)$, $\lambda_f = 0.03$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.5$ and $\mu = 1.5$. Right: function value at $\mathbf{x}_0$, varying with K

possible to the final configuration and short term planning is more convenient. Each added planning step contributes through new added control-dependent terms (larger paths, as in Fig. 1), which explains the increasing trend.

In Fig. 7, in the second case, we have the opposite of case 1. In that case, the more planning steps, the lowest the final cost, because the increments in the path length (as in Fig. 1) are compensated by smaller state-dependent costs $\sum \mathbf{x_k^T x}_k$.

The third case is illustrated in Fig. 8. There is a K' such that, for horizons $K < K'$, the system is driven to the final configuration which makes the corresponding cost decreases, but the control cost is increasing even more, explaining the initial global increase. For larger K's, as in case 2, the state-dependent cost decrements dominate the control-dependent cost increments, explaining the global decrease.

In the fourth case, the situation is the opposite. For small horizons, getting closer to the final configuration initially makes the corresponding cost decrease faster than the control cost increments, explaining the initial global decreasing tendency. For higher values of K, as in case 1, the control-dependent cost increments dominates the other variations, explaining the global increasing tendency.

In the following, we give explanations to the observed behaviors. First, the following lemmas enounce common characteristics to the $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ curves.

Lemma 1 *The sequence of values* $\sqrt{\rho_0^K}\eta_0^K$ *is monotonic with increasing* K.

Proof By using the intermediate value theorem, we can write:

$$\Delta\left[\rho_0^K\left(\eta_0^K\right)^2\right] = \frac{d}{dx}\left[\frac{x^2}{\mu + x\left(1+\sigma_u^2\right)}\right]_{x=\eta,\ \eta\in\left[\eta_0^{K-1},\eta_0^K\right]} \Delta\eta_0^K.$$

It can be verified that $\frac{d}{dx}\left[\frac{(x)^2}{\mu+x(1+\sigma_u^2)}\right] > 0$ for any $x > 0$, hence the sign of $\Delta\left[\rho_0^K\left(\eta_0^K\right)^2\right]$ is the one of $\Delta\eta_0^K$. As we have seen in Section 3, η_0^K is monotonic with increasing Ks. Hence, the sequence of $\frac{\partial}{\partial K}\rho_0^K(\eta_0^K)^2$'s and subsequently the sequence of $\sqrt{\rho_0^K}\eta_0^K$'s is monotonic (increasing or decreasing). This means that all our $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ curves either go the left or to the right. □

Lemma 2 *The sequence of values* $\sqrt{\rho_0^K}\phi_0^K$ *is decreasing with increasing* K.

Proof Consider $\Delta[\rho_0^K(\phi_0^K)^2]$; we can write:

$$\begin{aligned}
\Delta[\rho_0^K(\phi_0^K)^2] &= \rho_0^K\Delta\left[\phi_0^{K\,2}\right] + \phi_0^K\Delta\left[\rho_0^K\right] - \Delta\left[\phi_0^{K\,2}\right]\Delta\left[\rho_0^K\right] \\
&= \rho_0^K\Delta\left[\phi_0^{K\,2}\right] + \phi_0^{K-1}\Delta\left[\rho_0^K\right] \\
&= -\rho_0^K\left(\rho_0^{K-1}\phi_0^{K-1}\eta_0^{K-1}\right)\left(\phi_0^K + \phi_0^{K-1}\right) - \left(\phi_0^{K-1}\right)^2\rho_0^K\rho_0^{K-1}\Delta\left[\eta_0^K\right] \\
&= -\rho_0^K\rho_0^{K-1}\left(\left(\phi_0^{K-1}\right)^2\left(\frac{\mu\eta_0^{K-1} + \sigma_u^2\left(\eta_0^{K-1}\right)^2}{\mu + (1+\sigma_u^2)\eta_0^{K-1}} + \lambda(1+\sigma_u^2)\right) + \phi_0^{K-1}\phi_0^K\eta_0^{K-1}\right)
\end{aligned} \tag{40}$$

From Eq. 40, $\Delta[\rho_0^K(\phi_0^K)^2]$ is decreasing. Hence, the term $\sqrt{\rho_0^K}\phi_0^K$ is also decreasing. This means that when the horizon K increases, the trajectory of points $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ is going downwards. □

Lemma 3 *When* $K \to \infty$ *the point* $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ *converges to the intersection of the ellipse of* Eq. 39 *with the x axis (*$y = 0$*).*

Proof Under the convergence conditions, $\eta_0^\infty = \eta^*$ and $\phi_0^\infty = 0$. Recall that from Eq. 18, for high values of K, the value function at $\mathbf{x}_0$ tends to:

$$v_0^K(\mathbf{x}_0) \to \eta^*\mathbf{x}_0^T\mathbf{x}_0 + \frac{\lambda_f}{2}\mathbf{x}_f^T\mathbf{x}_f.$$

The expression above shows that the total expected cumulative cost function from a given starting point converges to a finite value, hence the differences $\Delta\left[v_0^K(\mathbf{x}_0)\right]$ tend to zero. This means that the point $\left(\sqrt{\frac{(\eta^*)^2}{1+\eta^*(1+\sigma_u^2)}}, 0\right)$ is a point located on the ellipse. It is the intersection of the ellipse with the x axis. To see this, we can evaluate the Eq. 39 at this point and we get:

$$\lambda - \frac{(\eta_0^K)^2}{\mu + \eta_0^K + \sigma_u^2\eta_0^K} = 0. \tag{41}$$

The value $\eta_0^K = \eta^*$ precisely satisfies this equation (see Eq. 28). □

The previous lemmas give us a few insights on the behavior of the curve $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ but it seems from the experiments described above that there is more: it appears that the curve does not cross more than once to the ellipse, i.e., that the value function has at most one critical value as a function of K. In the following, we provide a hint about how to show that this property holds.

First, we write the slope m_K for pair of points on the curve $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$:

$$m_K = \frac{\sqrt{\rho_0^K\left(\phi_0^K\right)^2} - \sqrt{\rho_0^{K-1}\left(\phi_0^{K-1}\right)^2}}{\sqrt{\rho_0^K\left(\eta_0^K\right)^2} - \sqrt{\rho_0^{K-1}\left(\eta_0^{K-1}\right)^2}}. \tag{42}$$

To ensure that the curve $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ crosses the ellipse at most once, we will show that m_K have a monotonic behavior (increasing or decreasing). Showing that m_K has a monotonic behavior using the definition above is complicated. However, if we assume that $\Delta\left[\sqrt{\rho_0^K}\phi_0^K\right]$ and $\Delta\left[\sqrt{\rho_0^K}\eta_0^K\right]$ are small enough,

$$m_K \approx \left(\frac{\frac{d}{dK}\left[\sqrt{\rho_0^K\left(\phi_0^K\right)^2}\right]}{\frac{d}{dK}\left[\sqrt{\rho_0^K\left(\eta_0^K\right)^2}\right]}\right) \text{ (in the continuous domain).} \tag{43}$$

From this approximation we get:

Lemma 4 *If* $\Delta\left[\sqrt{\rho_0^K}\phi_0^K\right]$ *and* $\Delta\left[\sqrt{\rho_0^K}\eta_0^K\right]$ *are small enough, and provided* $\mu > \frac{\lambda}{3}(1+\sigma_u^2)$*, then* m_K *has a monotonic behavior (increasing or decreasing) with increasing* K*.*

Proof Under the conditions above and after several algebraic developments using the properties of continuous derivatives in Eq. 43 we get:

$$\frac{\partial}{\partial K}[m_K] \approx -\frac{\phi_0^K}{\frac{\partial}{\partial}K}\left[\eta_0^K\right]\lambda\left(\frac{\mu v_K^2 + n_{\sigma_u} v_K \mu \eta_0^K + n_{\sigma_u}^2\left(\eta_0^K\right)^2 c_K - \lambda n_{\sigma_u}^2 c_K^2}{c_K^2 v_K^2}\right), \tag{44}$$

with:

$$\begin{aligned} c_K &\triangleq \mu + n_{\sigma_u}\eta_0^K \\ v_K &\triangleq 2\mu + n_{\sigma_u}\eta_0^K \\ n_{\sigma_u} &\triangleq \left(1+\sigma_u^2\right). \end{aligned} \tag{45}$$

On can show that whenever $3\mu > \lambda(1+\sigma_u^2)$, the numerator in the fraction above is positive, i.e., the slope m_K of the curve $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ is increasing or decreasing with increasing K. This shows that the concavity of the $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ curve keeps the same sign. □

Using these developments and with the conditions enounced above, one can write:

Theorem 1 *If* $\Delta\left[\sqrt{\rho_0^K}\phi_0^K\right]$ *and* $\Delta\left[\sqrt{\rho_0^K}\eta_0^K\right]$ *are smalll enough and provided* $\mu > \frac{\lambda}{3}(1+\sigma_u^2)$*, then the function of* K $\Delta\left[v_0^K(\mathbf{x}_0)\right]$ *has at most one critical point.*

Proof Considering the Lemmas 1, 2, 3 and 4 it is possible to deduce that the $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$ curve crosses the ellipse at most once. This means that $\Delta\left[v_0^K(\mathbf{x}_0)\right]$ has at most a critical point. □

Besides, because we know beforehand what the convergence point is, we have a taxonomy of behaviors of $v_0^K(\mathbf{x}_0)$, taking into account the position of the initial point with respect to the ellipse and the number of critical points.

First let us assume $3\mu > \lambda(1+\sigma_u^2)$. We define

$$\begin{aligned} \mathbf{I} &\triangleq \left(\sqrt{\rho_K^K}\eta_K^K, \sqrt{\rho_K^K}\phi_K^K\right) \\ &\triangleq \left(\sqrt{\frac{\lambda_f^2}{\mu + \lambda_f(1+\sigma_u^2)}}, \sqrt{\frac{1}{\mu + \lambda_f(1+\sigma_u^2)}}\right), \end{aligned}$$

i.e., $\mathbf{I}$ is the initial configuration of system in the space $\left(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K\right)$, meaning the point $\left(\sqrt{\rho_K^K}\eta_K^K, \sqrt{\rho_K^K}\phi_K^K\right)$ in that space, that is, the initial point in retro-time. Finally let us consider, the number of critical points of the function $v_0^K(\mathbf{x}_0)$.

Given these definitions and conditions, there are four possible behaviors of the function $v_0^K(\mathbf{x}_0)$, which are the following

1. If $\mathbf{I}$ is inside the ellipse and the function $v_0^K(\mathbf{x}_0)$ does not have critical points, then $v_0^0(\mathbf{x}_0)$ (resp. $v_0^\infty(\mathbf{x}_0)$) is

the global minimum (resp. maximum) over the possible K's and $\lambda_f(\mathbf{x}_0 - \mathbf{x}_f)^T(\mathbf{x}_0 - \mathbf{x}_f)$ is the global maximal value (see Fig. 6).

2. If $\mathbf{I}$ is outside the ellipse and the function $v_0^K(\mathbf{x}_0)$ does not have critical points, then $v_0^0(\mathbf{x}_0)$ (resp. $v_0^\infty(\mathbf{x}_0)$) is the global maximum (resp. minimum) over the possible K's and $\lambda_f(\mathbf{x}_0 - \mathbf{x}_f)^T(\mathbf{x}_0 - \mathbf{x}_f)$ is the global minimal value (see Fig. 7).
3. If $\mathbf{I}$ is inside the ellipse and the function $v_0^K(\mathbf{x}_0)$ has a critical point at $K = K'$, then $v_0^{K'}(\mathbf{x}_0)$ is the global maximum of $v_0^K(\mathbf{x}_0)$ (see Fig. 8).
4. If $\mathbf{I}$ is outside the ellipse and the function $v_0^K(\mathbf{x}_0)$ has a critical point at $K = K'$ then $v_0^{K'}(\mathbf{x}_0)$ is the global minimum of $v_0^K(\mathbf{x}_0)$ (see Fig. 9).

Hence, we can determine, in finite time, the infinite horizon cumulative cost v_0^∞ associated to a set of system parameters (see Eq. 18). Given a small ϵ, it is always possible to determine a horizon K such that the associated cumulative cost is close enough to the its limit when K grows to infinity, i.e. such that $v_0^K(\mathbf{x}_0)$ satisfies $|v_0^\infty(\mathbf{x}_0) - v_0^K(\mathbf{x}_0)| < \epsilon$. By using the properties of the dynamic programming, when calculating the accumulated cost for K, we estimate the accumulated costs $v_0^1, v_0^2, \ldots, v_0^{K'}, \ldots, v_0^K$. This means that the cumulative cost associated with K' has been computed during the estimation of v_0^K. It is possible, then, to store the accumulated costs for the different values of K and select the one that has the lowest cost.

5 Simulation Results

The first experiments in this section are simulations related to Section 3.7, i.e. for extreme values of the different parameters involved in the cost function. Later on, we present simulations concerning the selection of the optimal horizon, as described in Section 4.

5.1 Description of our Simulated 3D System

We suppose we have a holonomic system that we want to operate wirelessly. If the distance between the system and the transmitting device (antenna) increases, the communication disturbances, i.e. the noise level in the communication, increases. A similar modelling is used in [10] to perform robot navigation based on the intensity of a signal transmitted by an antenna. Hence, we are interested in driving the system from an initial configuration to a final configuration with a finite number of controls, while minimizing the communication disturbances. These controls are noisy and, as it happens in real systems, as the control magnitude increases, so does the noise level in the motion. This problem can be adequately described by the model presented in Section 2.1. In this case, we consider the antenna as the origin of our frame of reference, and we estimate a plan that moves the robot from the initial configuration to the final configuration with the least possible noise on the transmission and with the goal of reaching the final configuration, without spending to much energy on the controls.

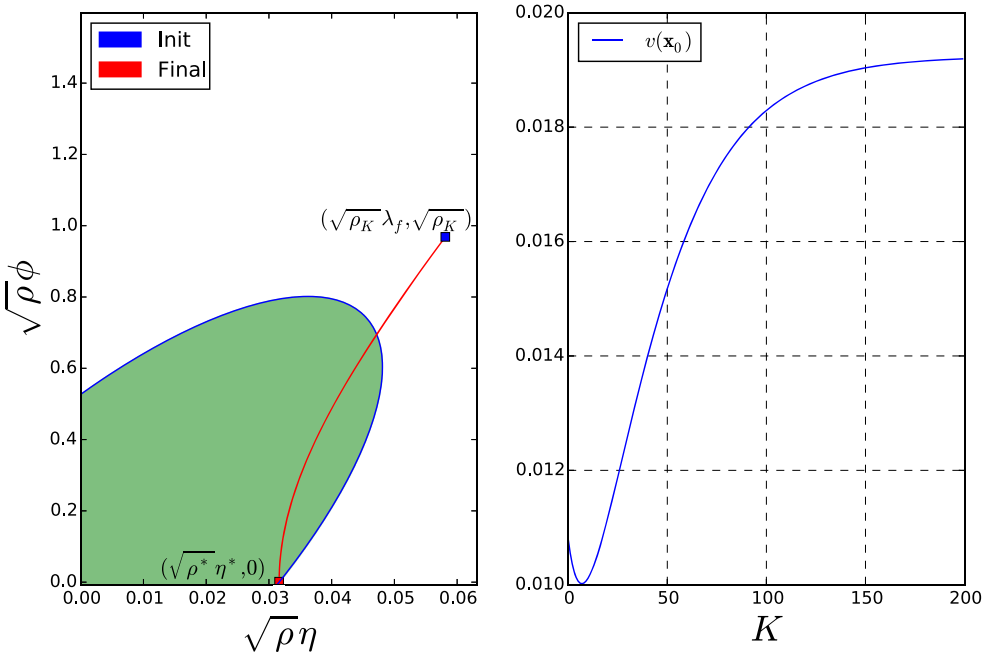

Fig. 9 Behavior of $v_0^K(\mathbf{x}_0)$ for differents horizons (case 4). Left: parameterization in the $(\sqrt{\rho_0^K}\eta_0^K, \sqrt{\rho_0^K}\phi_0^K)$ space with parameters $\mathbf{x}_0 = (-0.3, 0.8)$, $\mathbf{x}_f = (0.3, 0.8)$, $\lambda_f = 0.06$, $\sigma_x = 0.0$, $\sigma_u = 0.1$, $\lambda = 0.001$ and $\mu = 1.0$. Right: function value at $\mathbf{x}_0$, varying with K

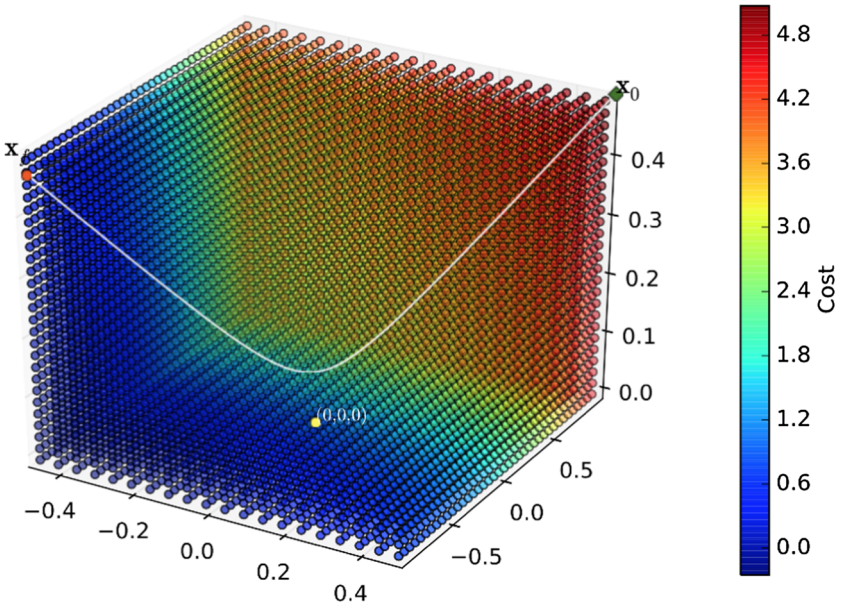

Fig. 10 Behavior of 3D system with parameters ($K = 50$, $\lambda = 1.0e-3$, $\mu = 1.0$, $\lambda_f = 1.0$, $\sigma_u^2 = 0.1$, $\sigma_x^2 = 0.1$). The color map indicates the cost: $\mathbf{x}^T Q\mathbf{x} + (\mathbf{x} - \mathbf{x}_f)^T Q_f(\mathbf{x} - \mathbf{x}_f)$

In the Fig. 10, we present an illustration for the simulation of such a three-dimensional system, with a planning horizon $K = 50$; we ensured that the differents terms of the cost function have similar scales (based on the analysis presented in the previous sections), in such a way that all terms of the cost function are considered. As expected, the initial controls tend to lead the states of the system close to the origin, while the latter controls lead the system towards the goal state.

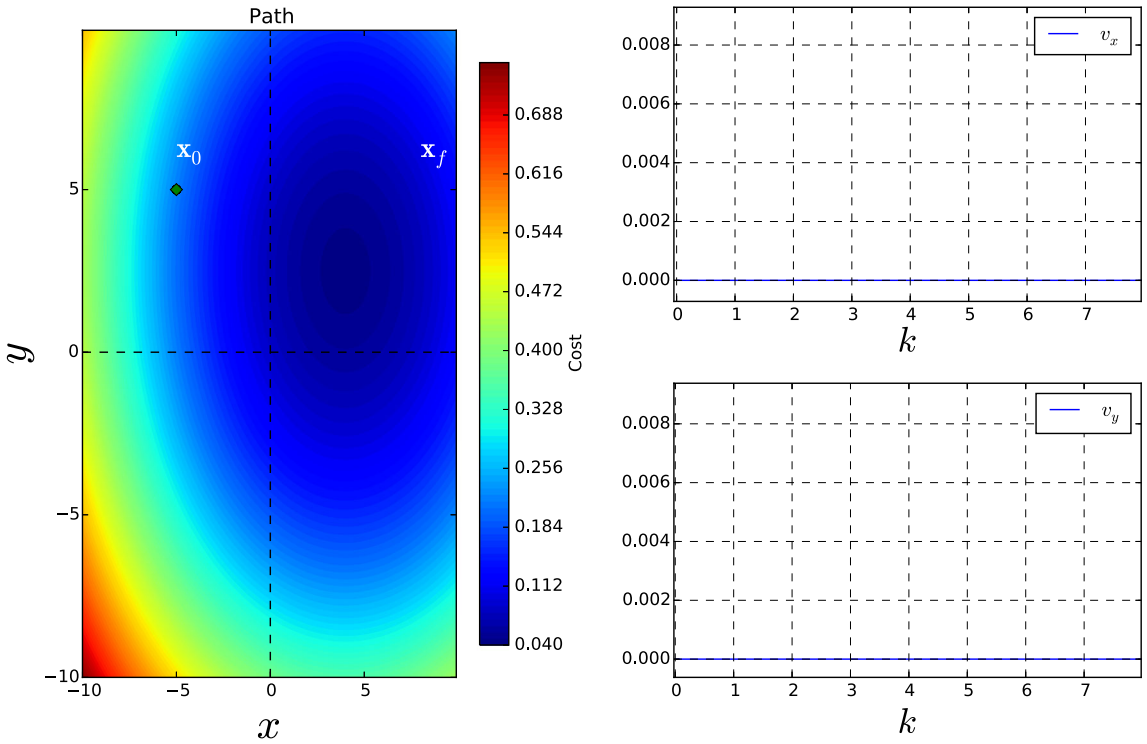

Fig. 11 Behavior of the system for large values of μ. Left: path and color map for the cost: $\mathbf{x}^T Q\mathbf{x} + (\mathbf{x} - \mathbf{x}_f)^T Q_f(\mathbf{x} - \mathbf{x}_f)$. Right: computed controls

5.2 Behavior of the Solutions in Extremal Cases

In this section, we show simulations for extreme values of the parameters involved in the cost function. For each example, the system performance is evaluated and we show that the results are consistent with the insights of Section 3.7.

The first example corresponds to $\mu \to \infty$. In Fig. 11, we can see the path and controls obtained for this simulation. The magnitude of the controls is close to zero, which causes the starting point $\mathbf{x}_0$ and the ending point $\mathbf{x}_K$ to be close. Although the initial configuration is relatively expensive, the applied controls maintain the system in this configuration due to the high cost of applying a control.

Figure 12 corresponds to $\lambda \to \infty$. As we mentioned it above, the controls drive the system to the origin as fast as possible. However, it is tempered by σ_u^2. Note that the color scale indicates that states near the origin have lower cost. Once the system states are near the origin, the control values remain close to zero.

In Fig. 13, we can observe the behavior of the system with $\lambda_f \to \infty$. The controls lead the system towards the final configuration. The scale color indicates that states far away from the target configuration $\mathbf{x}_f$ are those with a higher cost.

5.3 Optimal Planning Horizon

Taking as a starting point the results in Section 4, we present here two types of analysis for the 3D system described above: a quantitative analysis in terms of cost and a qualitative analysis based on the objectives achieved by the system.

We study the function value at the initial configutation $\mathbf{x}_0$, with different planning horizon; Note that a plan of horizon $K > K'$ might get smaller costs.

We had seen that there are four types of paths: The first is case is in which the minimal cost is obtained at the shortest horizon; The second corresponds to the case in which the best cost is obtained at horizons as largest as possible; Case 3 is a bit more complex, depending on the parameters it might be similar to case 1 or case 2. Finally, the case 4 corresponds to an intermediate horizon K' giving a globally minimal path cost.

For our first example, we consider that all the terms in the cost function have a comparable magnitude and that the initial and final configurations are near the origin. This configuration corresponds to the case 4 in Theorem 1. The simulation for this system configuration can be seen in Fig. 14. The optimal horizon in terms of cost $v_0^K(\mathbf{x}_0)$ is

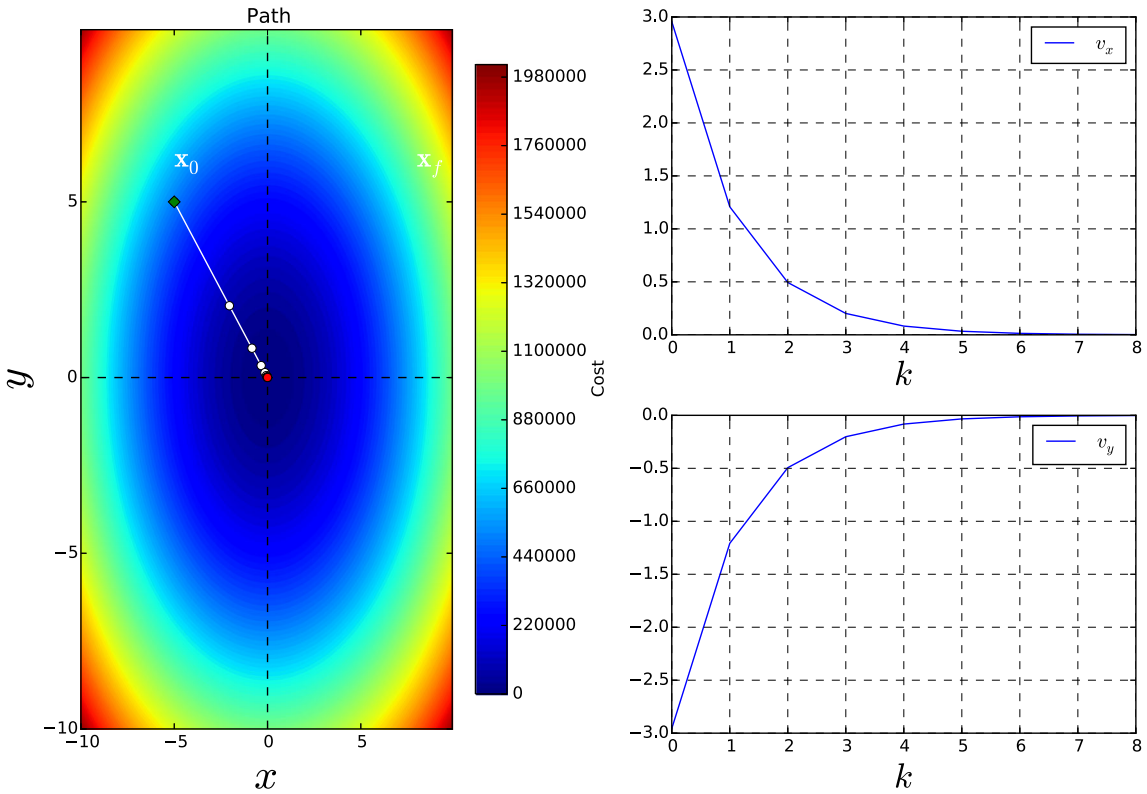

Fig. 12 Behavior of the system for large values of λ. Left: path and color map for the cost: $\mathbf{x}^T Q\mathbf{x}$. Right: computed controls

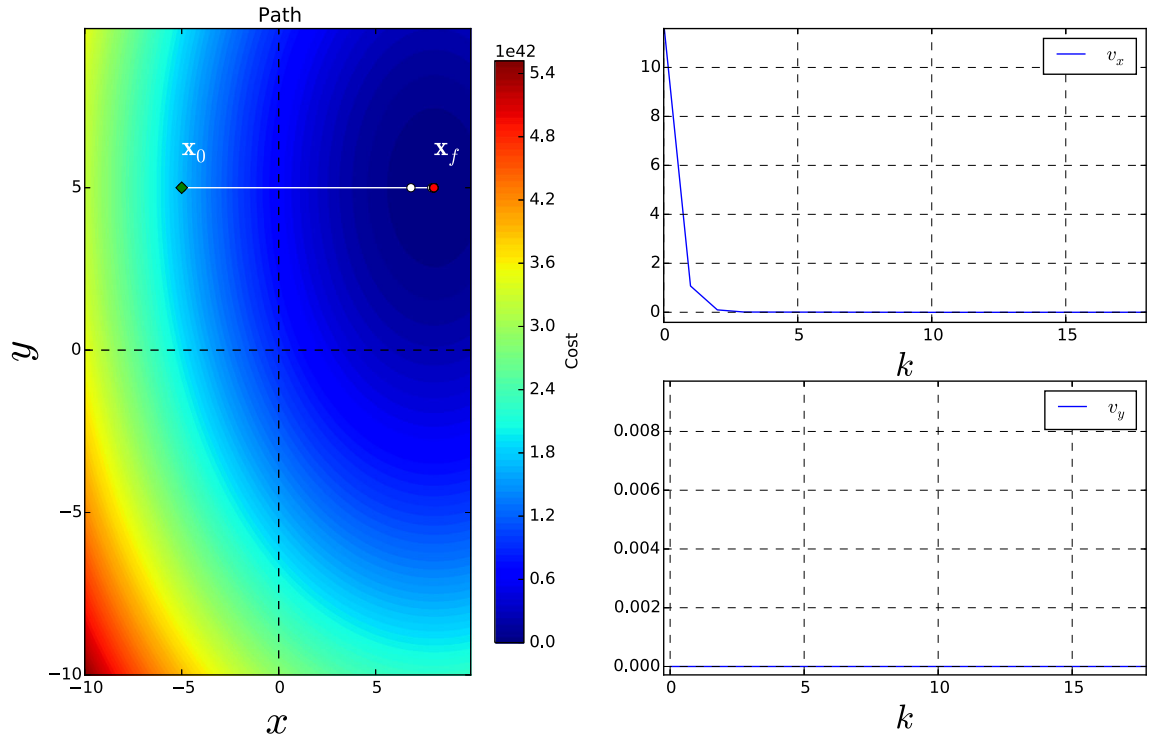

Fig. 13 Behavior of the system for large values of λ_f. Left: path and color map for the cost: $(\mathbf{x}-\mathbf{x}_f)^T Q_f (\mathbf{x}-\mathbf{x}_f)$. Right: computed controls

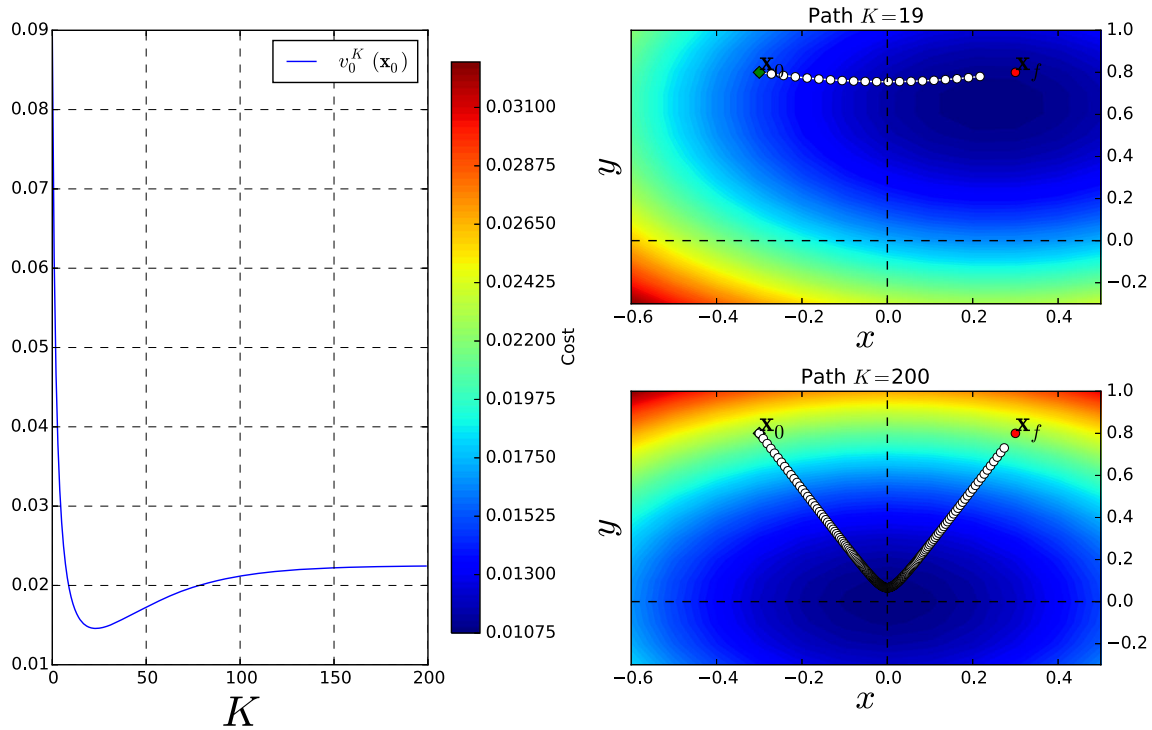

Fig. 14 Behavior of the system when all the terms in $v_0^K(\mathbf{x}_0)$ have a similar contribution. Left: $v_0^K(\mathbf{x}_0)$. Right: path for two differents values of K. The color map corresponds to $v_0^K(\mathbf{x})$

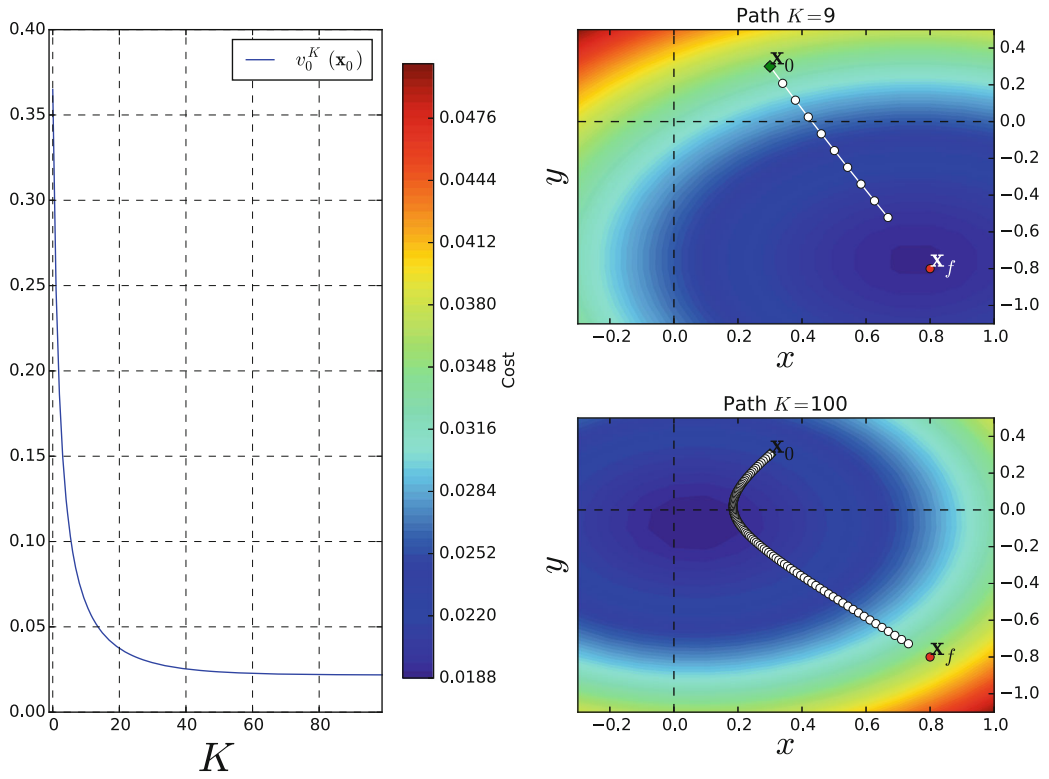

Fig. 15 Behavior of the system when all the terms in $v_0^K(\mathbf{x}_0)$ have a similar contribution. Right: path for two differents values of K. The color map corresponds to $v_0^K(\mathbf{x})$

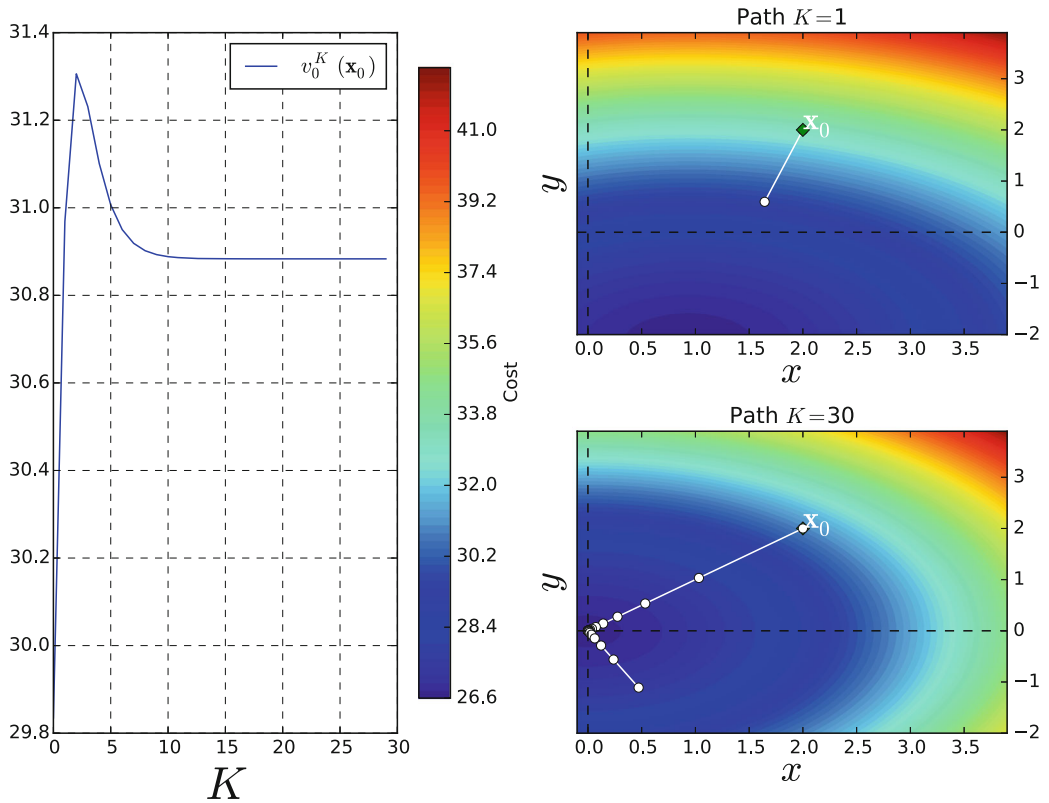

Fig. 16 Behavior of the system when $\mu \geq \lambda > \lambda_f$. Right: path for two differents values of K. The color map correspond to $v_0^K(\mathbf{x})$

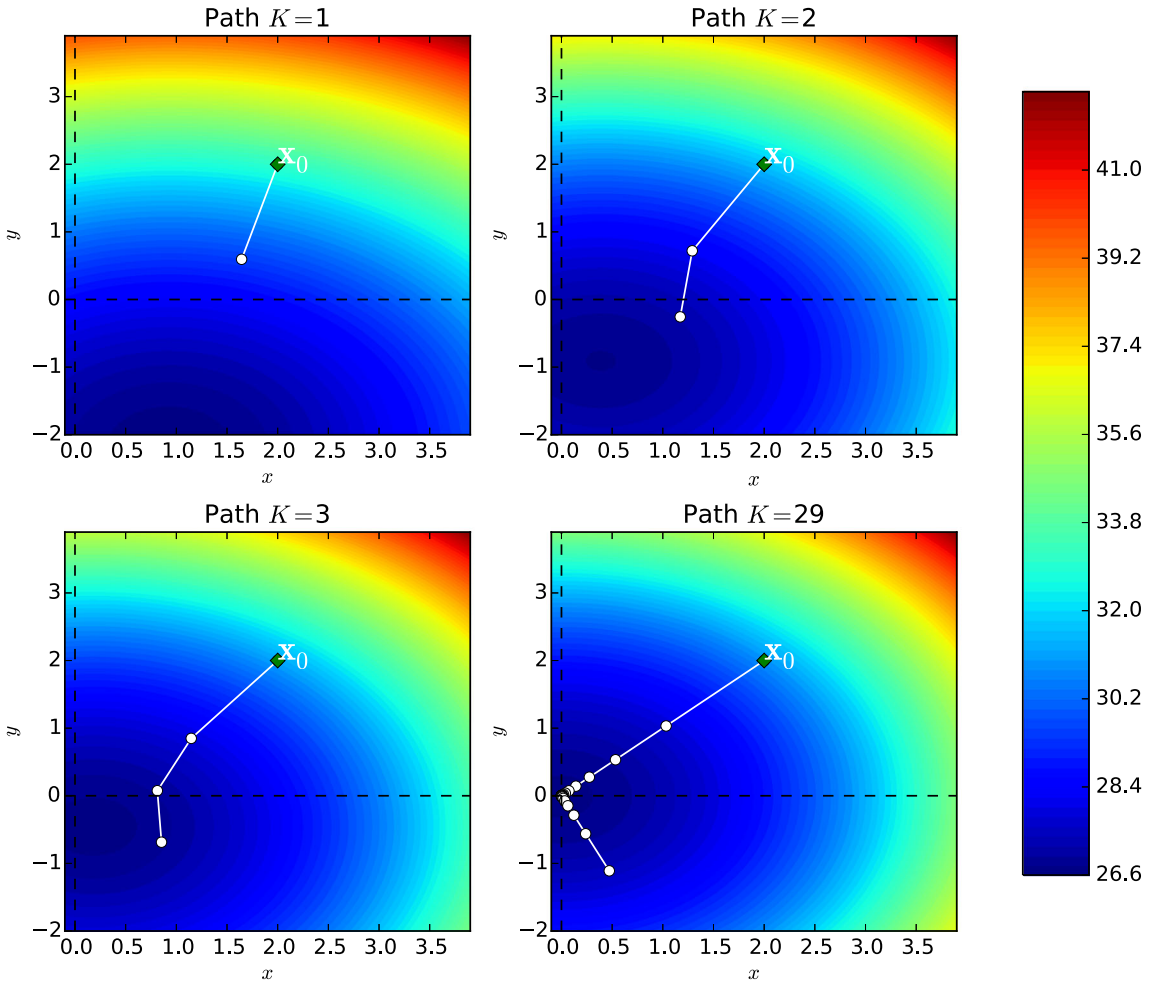

Fig. 17 Zoom over the path for the system shown in Fig. 16 for differents values of K. The color map correspond to $v_0^K(\mathbf{x})$

$K = 10$. The resulting policy leads the system near the goal configuration $\mathbf{x}_f$. On the other hand, the horizon $K = 50$ also leads the system close to $\mathbf{x}_f$ but $v_0^{50}(\mathbf{x}_0)$ is higher than $v_0^{10}(\mathbf{x}_0)$, because of the control penalties. The system has a global minumun at $K' = 10$.

For the system shown in Fig. 15, as the horizon is increased, the total cost is reduced, because state-dependent costs are reduced more than what control-dependent costs are increased. This example corresponds to case 2 in Theorem 5.

In Fig. 16, we can observe the behavior of a system corresponding to case 3. In this Figure, the minimal cost corresponds to shorthest horizons, which is similar to case 1. This system has a global maximum at $K' = 2$. This is because at K', the control-dependent cost is increased and the state-dependent cost is not reduced significantly. This can be seen in Fig. 17. In this Figure, shorter horizons $K \leq K'$ have higher costs on the states far away from the origin, which explains the increase in costs (i.e. the configuration $\mathbf{x}_0$ has a high cost in these horizons). As the K grows ($K > K'$) this cost is reduced, explaining the decreasing behavior.

Notice that even if the minimal cost is not related to convergence, when the convergence is reached, the robot is as close as possible of the goal configuration, paying a bounded cost. Planning more time steps does not provide any benefit.

6 Conclusions and Future Work

In this paper, we have studied solutions to the problem of generating motion policies for a navigation task defined as a multi-objective optimization problem, where the robot minimizes a cost function that combines the cost of being in a state, the cost of applying a control and the cost of not reaching the goal configuration under control and state uncertainty.

We have presented a quantitative and qualitative analysis of the expected behavior of the solutions to the Linear

Quadratic Gaussian (LQG) problem with control-dependent noise. We have related the behavior of the expected solutions to the LQG problem to the weights of the immediate costs, control costs and costs associated to the final configuration, and we have given useful bounds for understanding the effect of the objective function parameters.

We have also given conditions for the quadratic coefficient in the cost-to-go function to converge to a finite value when the horizon increases, which induces that the cost associated to the optimal motion policy is bounded by a finite value (for a bounded environment) and hence one can evaluate whether or not, the system can afford to pay that cost.

Finally, we have presented an analysis of the optimal planning horizon K for the overall cost or objective function used to assess the motion policy, under control dependent uncertainty. To the best of our knowledge this analysis is novel.

The analysis provided in this is work is directly applicable to the case of quadratic cost $\lambda \mathbf{x}^T \mathbf{x} + \mu \mathbf{u}^T \mathbf{u}$ and linear motion model $\dot{\mathbf{x}} = \mathbf{x} + \mathbf{u}$, in which $\mathbf{x}$ is the state, $\mathbf{u}$ is the control, but it can be extended to a larger class of problems. Qualitatively, it should hold for systems $\dot{\mathbf{x}} = \Lambda \mathbf{x} + \mathbf{u}$, being Λ a positive diagonal matrix. Then, it could also be extended to more general dynamic models $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$, when A is diagonalizable with positive eigenvalues. This is because, if $\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$, the change of variables $\mathbf{z} = \mathbf{Q}^{-1}\mathbf{x}$ leads to another system of the form $\dot{\mathbf{z}} = \Lambda \mathbf{z} + \mathbf{v}$, for which the discussed properties should hold.

References

1. Anderson, B.D.O., Moore, J.B.: Optimal control: linear quadratic methods. Prentice-Hall, Inc., Upper Saddle River (1990)
2. Bai, H., Hsu, D., Lee, W.S.: Integrated perception and planning in the continuous space: A pomdp approach. Int. J. Robot. Res. **33**(9), 1288–1302 (2014)
3. Barraquand, J., Ferbach, P.: Motion planning with uncertainty: the information space approach. In: 1995. Proceedings., 1995 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1341–1348 (1995)
4. Bertsekas, D.P.: Dynamic programming and optimal control. Volume I. Athena Scientific optimization and computation series. Mass. Athena Scientific, Belmont (2005)
5. Candido, S., Davidson, J.C., Hutchinson, S.: Exploiting domain knowledge in planning uncertain robot systems modeled as POMDPs. In: Proceedings of IEEE Int. Conf. on Robotics and Automation, ICRA 2010, Anchorage, pp. 3596–3603 (2010)
6. Kaelbling, L., Littman, M., Cassandra, A.: Planning and acting in partially observable stochastic domains. Artif. Intell. **101**(1-2), 99–134 (1998)
7. Kurniawati, H., Du, Y., Hsu, D., Lee, W.S.: Motion planning under uncertainty for robotic tasks with long time horizons. Int. J. Robot. Res. **30**(3), 308–323 (2011)
8. LaValle, S.M., Hutchinson, S.A.: An objective-based framework for motion planning under sensing and control uncertainties. Int. J. Robot. Res. **17**(1), 19–42 (1998)
9. Smith, C.R., Cheeseman, P.: On the representation and estimation of spatial uncertainty. Int. J. Robot. Res. **5**(4), 56–68 (1986)
10. Taylor, K., LaValle, S.M.: Intensity-based navigation with global guarantees. Auton. Robot. **36**(4), 349–364 (2014)
11. Thrun, S.: Monte Carlo POMDPs. In: Solla, S.A., Leen, T.K., Muller, K.-R. (eds.) Advances in Neural Information Processing Systems 12, pp. 1064–1070. MIT Press, Cambridge (2000)
12. Todorov, E.: A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In: Proceedings of the 2005, American Control Conference, pp. 300–306 (2005)
13. van den Berg, J., Abbeel, P., Goldberg, K.: LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information. Int. J. Robot. Res. **30**(7), 895–913 (2011)
14. van den Berg, J., Patil, S., Alterovitz, R.: Motion planning under uncertainty using iterative local optimization in belief space. Int. J. Robot. Res. **31**(11), 1263–1278 (2012)

Hugo Carlos graduated from Universidad Autonoma de Campeche (Mexico) in 2005. He got his master degree in 2011 at Centro de Investigacion en Matematicas and is now PhD candidate. His research interests are in computer vision and mobile robotics and optimal control.

Jean-Bernard Hayet graduated from Ecole Nationale Superieure de Techniques Avancees in 1999. He got his master degree from University Paris VI and his Ph.D. degree from University of Toulouse in 2003. His doctoral dissertation tackled the problem of landmark-based navigation with computer vision in indoor environments and was supervised by Dr. M. Devy and Dr. F. Lerasle. From 2003 to 2006 he did a post-doctoral stay in visual tracking at University of Liege with Pr. Justus Piater. Since 2007 he is researcher at the Centro de Investigacion en Matematicas, in Mexico, and is focusing his research on robot vision and visual tracking. He is member of the Mexican National System of Researchers at level 1.

Rafael Murrieta-Cid received the B.S degree in physics engineering from the Monterrey Institute of Technology and Higher Education, Monterrey, Mexico, in 1990 and the the Ph.D. degree from the Institut National Polytechnique, Toulouse, France, in 1998. His Ph.D. research was done with the Robotics Group (RIA) of the LAAS-CNRS. In 1998-1999, he was a Postdoctoral Researcher with the Department of Computer Science, Stanford University, Stanford CA, USA. During 2002-2004, he was a Postdoctoral Researcher with the Beckman Institute and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign (UIUC), Urbana, IL, USA. Since 2006, he has been with the Centro de Investigacion en Matematicas, CIMAT, Guanajuato, Mexico. In 2016, he was on a sabbatical leave at UIUC. His research interests include robotics, robot motion planning and control theory.