INTERNATIONAL JOURNAL OF CONTROL
https://doi.org/10.1080/00207179.2020.1774077

Prescribing transient and asymptotic behaviour to deterministic systems with stochastic initial conditions

Martijn Desscher and Bayu Jayawardhana

Engineering and Technology Institute, Faculty of Science and Engineering Groningen, University of Groningen, Groningen, Netherlands

ABSTRACT

We study a containment control problem (CCP) and a shape control problem (SCP) for systems whose initial condition is a random variable with known distribution. The two control problems both require exponential convergence to a desired trajectory, which is complemented by either; (i) a required cumulative distribution over a prescribed containment set at a specific transient time for the CCP, or; (ii) a maximum distance between an attained and a desired probability density function of the state for the SCP. For the CCP, we obtain solutions for both linear and nonlinear systems by designing the closed-loop such that the initial pdf shrinks or contracts to a desired trajectory. For the SCP, we obtain solutions for linear systems and an admissible desired pdf, by designing the closed-loop such that the evolution of the pdf at the transient time is similar to the target pdf.

ARTICLE HISTORY
Received 19 July 2019
Accepted 18 May 2020

KEYWORDS
Control systems with prescribed transient behaviour; systems with stochastic initial conditions; contraction-based control method; robotic manipulator; process control

1. Introduction

One of the major focuses for many control engineering applications is the control of variations in the states of a process. Such variations are present in all aspects of all processes, but the magnitude of these variations either forces us to consider them explicitly, or allows us to ignore their effect. A stochastic noise, if present, can be a particular difficult type of perturbation to deal with. In this case, we generally consider a control system to be ‘good’ if the effect of this perturbation remains within acceptable bounds, which can be determined based on the system characteristics. A general underlying assumption is that the perturbation influences the vector field of the systems’ dynamics. Such problems are considered in the field of stochastic control (Aström, 1970; Bertsekas, 1976) and one of the ways to express knowledge on the state is through a probability density function (pdf) that changes over time (Kárný, 1996; Sun, 2006). This function then maps values of the state space to probabilities of occurrence. The Fokker–Planck equation or forward Kolmogorov equation (Gardiner, 1985; Risken, 1989) is a well-known result related to this approach, commonly specified for an Itô process (driven by a Wiener process), such as Brownian motion (Itô, 2004).

Efforts to improve control system design, combined with ongoing advances in sensor and actuator technologies, are increasingly successful in minimising the perturbation effects on the vector field. We can hence turn our attention to other sources of variation. A subject that has received little attention in the literature, is the variations in the initial conditions. This will be the focus of this paper. An example of an application where such variations are relevant is a high-precision and high-frequency manufacturing processes where natural and uncontrollable variation is present in the input materials. The presence of such undesirable *ab initio* conditions can decrease the end-products quality when it is not taken into account explicitly in the control design. In other words, although the use of high-precision systems can minimise product variations due to measurement noises and lack of accuracy in the actuator systems, it cannot pre-empt variability in the initial conditions.

In this paper, one of the main assumptions of the systems is that the initial conditions are random variables and that where we have apriori knowledge of the associated pdf. This accordingly assumes more information on the initial conditions than traditional approaches and allows for other analysis. Such formalism encompasses also the standard deterministic setting (with known initial condition) by taking a Dirac pdf. When both the initial time and initial state are random variables (which can represent time-varying initial conditions), our results are still applicable by considering the marginal distribution of the joint distribution on the initial state random variable, which leads to conservative results. For such systems with stochastic initial conditions, designing a control law with a particular asymptotic behaviour as the main criterion does not reveal extra information since all possible trajectories converge to the desired operating point or trajectory, independent of the initial state pdf. On the other hand, since the closed-loop systems’ transient behaviour is highly dependent on the initial state, we investigate in this paper control design methods where we take into account the evolution of the state pdf in the control design problems.

In our first main result, we consider a containment control problem (CCP) in Sections 3 and 4 where, in addition to achieving desired asymptotic behaviour, the controller needs to guarantee that the cumulative distribution of the trajectories over a prescribed set at a given transient time reaches a prescribed value.

CONTACT Martijn Dresscher martijndresscher@posteo.net TNO, Stieltjesweg 1, 2628CK Delft, Netherlands

Such a problem formulation can be related to the funnel control problem for deterministic systems (Ilchmann et al., 2002, 2007). In these papers, the control problem is to design control laws that guarantee the state trajectories from *all* initial conditions remain in a desired funnel, which contracts to the desired state. As the funnel must contain all initial conditions at the initial time, the results are conservative and applicable for linear systems with known relative degree, minimum phase with positive definite high-frequency gain matrix. In this case, our CCP case can be interpreted as a modification to the funnel control whose initial funnel need to cover only a set of initial state with the desired cumulative distribution at the initial time. As will be shown later in Sections 3 and 4, our setting is more general than the funnel one and it admits a general class of linear systems, namely all controllable linear systems.

In our second main result as presented in Section 5, we consider a different transient performance criterion where, in addition to the asymptotic behaviour requirement, the controller must ensure that the evolution of state pdf at a given transient time is close to a desired pdf. In particular, it must guarantee that the Hellinger distance between the two pdfs is less than a given prescribed level. A recent work related to the shaping control of pdf has also appeared in Buehler et al. (2016). In Buehler et al. (2016), a generic control framework using stochastic MPC is proposed for stochastic nonlinear systems, where the initial condition is a random variable and the disturbance is a stochastic process. In this work, a stochastic MPC problem is proposed where the distance of evolving pdf to a desired one must be minimised. In contrast to the result presented in Buehler et al. (2016) which does not yield an analytical solution, we restrict our problem only to the random initial condition case that has allowed us to construct simple control laws with a guaranteed level of performance and to provide rigorous analysis of the method. A practical implication of this limitation is that the obtained results are only valid for systems that allow for sufficiently accurate dynamical description through deterministic equations.

The preliminary works of our results have been presented in Dresscher and Jayawardhana (2017a, 2017b). In the present paper, we extend the preliminary results in Dresscher and Jayawardhana (2017a, 2017b) in several directions. Firstly, our Theorem 3.1 in Section 3 encompasses general pdf of initial state and we allow now for any well-posed reference trajectories instead of only equilibrium points as in Dresscher and Jayawardhana (2017a). Secondly, the results in Theorem 4.1 in Section 4 are extended to general nonlinear systems, instead of affine nonlinear systems as considered in Dresscher and Jayawardhana (2017b). Thirdly, the pdf matching approach as presented in Theorems 5.2 and 5.5 is novel when compared to Dresscher and Jayawardhana (2017a). We note that the results in Dresscher and Jayawardhana (2017a) rely on a coordinate transformation of the original system to another one such that the state pdf evolution and the target pdf belongs to the same class of pdf, which is not trivial to obtain.

The remainder of the paper is structured as follows. In Section 2 we will introduce the system dynamics, transient specifications to evaluate the performance of our controller, the two control problems that we consider and we will use a short example to show the non-triviality of our control problem. Sections 3–5 will present solutions to our control problems, where one control problem is solved for both the linear and nonlinear case, while the other is solved only for the linear case and under specific conditions. Sections 6 and 7 are used to present two non-trivial simulation results, where both control problems are considered. Lastly, we round up with the conclusions in Section 8.

2. Problem definition

We will use this section to formally present our control problems. For this purpose, we start by introducing our choice of dynamical system equations and transient specifications. After the problems have been introduced, we provide a simple example to show the non-triviality of the problems and some controller design considerations.

2.1 Dynamical system equations

Consider the general dynamical system given by

$$\dot{x} = f(x, u, t), \quad x(0) = x_0, \tag{1}$$

where $x \in X \subseteq \mathbb{R}^n$, $u \in U \subseteq R^m$ and $f : X \times U \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is a continuously differentiable vector field. Let us assume that x_0 is a known random variable defined on $X_0 \subset X$, satisfying a probability density function (pdf) $\phi_{x_0} : X_0 \to \mathbb{R}_{\geq 0}$. In this case, its forward solution $x(t)$ is random variable for all $t > 0$ and we denote the propagation of ϕ_{x_0} along (1) by $\phi_{x_0,t}$. Note that the usual setting where x_0 is deterministic is a particular class of this class of system where ϕ_{x_0} is simply given by a Dirac delta function.

2.2 Transient specifications

For defining transient behaviour specification corresponding to the evolution of $\phi_{x_0,t}$, there are two possibilities in defining the measure. For the first one, we can relate $\phi_{x_0,t}$ at a terminal time T or in an interval $[0, T]$ to a desired point (or desired set) or to a desired trajectory $x_d(t)$ defined on the time interval $[0, T]$, respectively. For the second one, we can relate $\phi_{x_0,t}$ to a (dynamic or stationary) target distribution.

With regards to the first possibility, we will discuss two measures that can be used to define transient specification based on $\phi_{x_0,t}$ and a given desired point or set. The first measure is given by the cumulative density of $\phi_{x_0,T}$ at the terminal time T over a prescribed set Ξ. More precisely, we can define the following measure

$$\Phi_{\Xi,T} := \int_{\Xi} \phi_{x_0,T}(\xi)\, \mathrm{d}\xi, \tag{2}$$

where $\Xi \subset X$ and T is the relevant terminal transient time. This transient specification is straightforward and it yields a scalar value. The second candidate measure is the second moment with respect to a point, which (for a single dimension) can be expressed as

$$\sigma(\phi_{x_0,T}, \mu) := \int_X (\xi - \mu)^2 \phi_{x_0,T}(\xi) \mathrm{d}\xi, \tag{3}$$

where μ is a desired point corresponding to the transient time of interest T. Notice that this expression is equal to the computation of the variance if $\phi_{x_0,T}$ is normally distributed with μ be the mean value. Both of the specifications above can be interpreted in a similar manner as the classical specifications of rise time or settling time and are thus highly relevant for any control problem. When comparing the two specifications, it is furthermore easy to see that the computation of the cumulative density is relatively simple. This transient specification has the advantage that we obtain a scalar-valued output which again simplifies interpretation and implementation in the control design.

For the second possibility, when considering (dis)similarity between two density functions, we can consider measures that are given by distances or divergences such as the Hellinger distance, Bhattacharyya distance, Kullback–Leibler divergence and Jeffrey's divergence (Ali & Silvey, 1966; Kailath, 1967; Kullback, 1997). A distance deserves preference over a divergence, since it produces the desired scalar-valued output. Note that the Bhattacharyya distance and the Hellinger distance are related to each other through the Bhattacharyya coefficient (Abou-Moustafa & Ferrie, 2012; Buehler et al., 2016). If ϕ_d denotes the target distribution at terminal time T, the Bhattacharyya coefficient is given by

$$BC(\phi_d, \phi_{x_0,T}) = \int_X \sqrt{\phi_d(\xi)\phi_{x_0,T}(\xi)}\mathrm{d}\xi. \tag{4}$$

If both distributions are equal then it will give Bhattacharyya coefficient of 1 and if they are dissimilar then the Bhattacharyya coefficient will be close to 0. Using this coefficient, the Hellinger distance is defined by

$$d_h(\phi_d, \phi_{x_0,T}) = \sqrt{1 - BC(\phi_d, \phi_{x_0,T})}, \tag{5}$$

and, correspondingly, the Bhattacharyya distance is given by

$$d_b(\phi_d, \phi_{x_0,T}) = -\ln(BC(\phi_d, \phi_{x_0,T})). \tag{6}$$

The two measures are relatively similar. However, the Hellinger distance is a proper metric, as discussed in Abou-Moustafa and Ferrie (2012), while the Bhattacharyya distance is only a semi-metric because it does not satisfy the triangle inequality. Our distance of choice is therefore the Hellinger distance.

When we compare these two possibilities of defining measure, there is a significant difference between transient specification (2) and (5). The former does not impose a shape on the pdf of the states. The latter, on the other hand, does impose this shape and therefore places different (more strict) requirements on the controller design as we will show later in this paper. We will consider both the Hellinger distance (as in (5)) and the cumulative density (as in (2)) in the sequel.

Before we proceed with the control problem formulations, we would like to clarify some notation that we use for the defining distances in this paper. In order to maintain a clear distinction, we will always denote the Hellinger distance by $d_h(\phi_1, \phi_2)$, the Euclidean distance by $d_E(x_1, x_2)$ and the Finsler distance by $d_F(x_1, x_2)$. Here, ϕ_1 and ϕ_2 are two pdfs, while x_1 and x_2 are points in state space.

2.3 Control problems formulation

We are now ready to define our two control problems, based on two different transient specifications as given before.

Containment Control Problem (CCP): For a system as in (1), given a desired containment set Ξ, a desired containment level $p^* \in (0, 1)$, a transient time $T > 0$, a distance $d(\cdot, \cdot)$ and a target trajectory $x_d(t)$, design a control law $u(t) = k(x(t), t)$ such that

CCPa: $\Phi_{\Xi,T} \geq p^*$
CCPb: $\lim_{t\to\infty} d(x(t), x_d(t)) = 0$.

In the formulation as above, **CCPa** is the realisation of a minimum containment criterion during the transient. The control problem hence incorporates the cumulative density transient specification (as in (2)). This condition is complemented by **CCPb**, which requires convergence to a desired trajectory in the asymptote. The control objective for the **CCP** is illustrated in Figure 1.

Shape Control Problem (SCP): For the system in (1), given a desired pdf ϕ_d, a transient time T, a distance $d(\cdot, \cdot)$, a desired Hellinger distance $\ell \in [0, \infty)$ and a target trajectory x_d, design a control law $u(t) = k(x(t), t)$ such that

SCPa: $d_h(\phi_{x_0,T}, \phi_d) \leq \ell$
SCPb: $\lim_{t\to\infty} d(x(t), x_d(t)) = 0$.

Similar to the structure of the **CCP**, **SCPa** is the realisation of the transient performance criteria related to the Hellinger distance by requiring it to be smaller than ℓ. This condition is again complemented by the asymptotic convergence criterion as expressed in **SCPb**. The control objective for the **SCP** is illustrated in Figure 2.

2.4 Control problem example

Before moving on to our contributions, we will show the non-triviality of our control problem by considering the following simple example for the **CCP** with $n = 1$. The example furthermore highlights come of the considerations relevant for solving the **CCP** and **SCP**. We try to solve the **CCP** with a standard control law, namely a state feedback law, for a first-order linear system. For this first-order LTI system, we have $f(x, u) = ax + bu$, with $a \in \mathbb{R}$ and $b \in \mathbb{R}$. Furthermore, let $x_d(t) = x^*$. Applying the linear feedback

$$u = k(x - x^*) - \frac{a}{b}x^*, \tag{7}$$

with $k \in \mathbb{R}$, to (1) will lead us to the following simple expression of the closed-loop system

$$\dot{\tilde{x}} = (a + bk)\tilde{x}, \quad \tilde{x}(0) = x_0 - x^*, \tag{8}$$

where $\tilde{x} = x - x^*$ is the error state and the gain k can be chosen arbitrarily to ensure that $(a + bk) < 0$. To simplify things further, we will assume a normal distribution for the initial state, e.g. $\phi_{\tilde{x}_0} = \mathcal{N}(\mu - x^*, \sigma)$. Since we are interested in a non-trivial solution of control problem 1, we assume that $\mu \neq 0$. For defining the first transient behaviour specification of the closed-loop

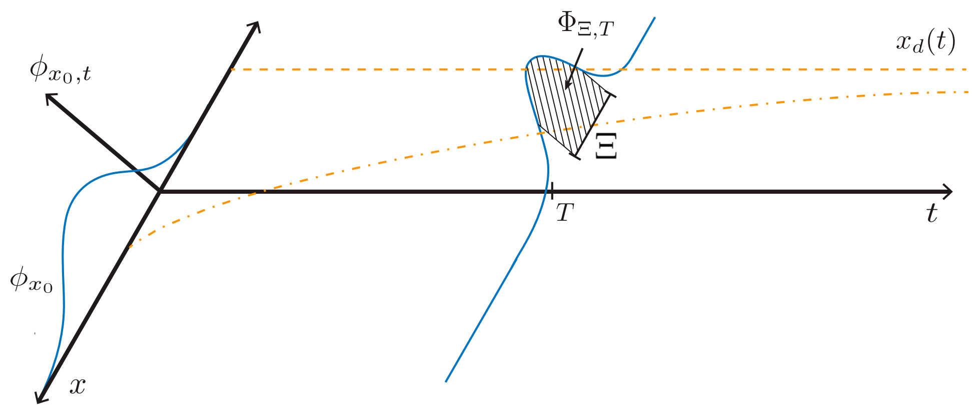

Figure 1. This figure illustrates the control objective of the **CCP**. The initial distribution ϕ_{x_0} changes with time such that it has a cumulative density over the set Ξ at time T, denoted as $\Phi_{\Xi,T}$. Furthermore, all possible initial values should converge to $x_d(t)$, which could result in the indicated trajectory for a specific initial state of ϕ_{x_0}.

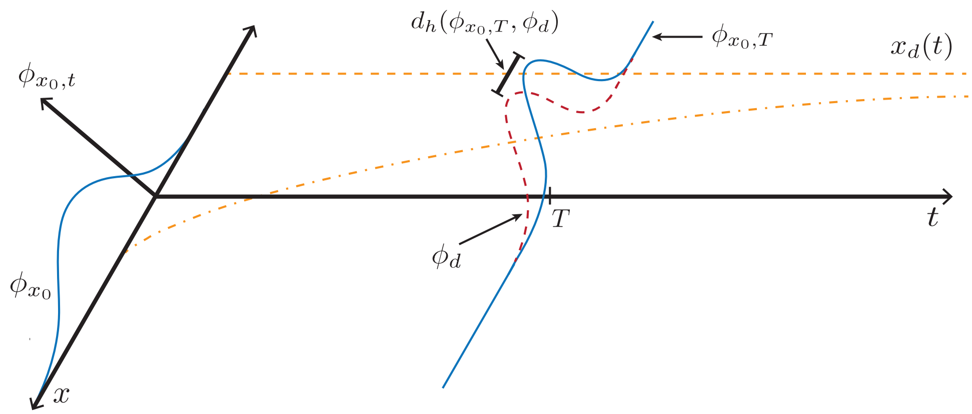

Figure 2. This figure illustrates the control objective of the **SCP**. The initial distribution ϕ_{x_0} changes with time, such that the distribution at time T, denoted as $\phi_{x_0,T}$, has a Hellinger distance $d_h(\phi_{x_0,T}, \phi_d)$ w.r.t. a desired distribution ϕ_d. Furthermore, all possible initial values should converge to $x_d(t)$, which can result in the indicated trajectory for a specific initial state of ϕ_{x_0}.

system (8), we take $\tilde{\Xi} = [x_{T,\text{low}}, x_{T,\text{up}}] - x^*$ where $x_{T,\text{low}}$ and $x_{T,\text{up}}$ are the lower and upper bound of the containment interval Ξ.

As we are dealing with a simple first-order linear system, we can use the bounds of $\tilde{\Xi}$ and the explicit solution of (8) to construct the image of this containment interval at time $t = 0$, which we denote as $\tilde{\Xi}_0$. In this way, the value $\Phi_{\tilde{\Xi},T}$ will be equivalent to cumulative distribution of $\tilde{x}_0$ on $\tilde{\Xi}_0$.

Based on the solution of (8), we have

$$\tilde{x}_{0,\text{low}} = e^{-(a+bk)T}\tilde{x}_{T,\text{low}} \tag{9}$$

and

$$\tilde{x}_{0,\text{up}} = e^{-(a+bk)T}\tilde{x}_{T,\text{up}}, \tag{10}$$

where, understandably, $\tilde{x}_{0,\text{low}}$ and $\tilde{x}_{0,\text{up}}$ are the lower and upper bound of $\tilde{\Xi}_0$.

Since $\phi_{\tilde{x}_0} = \mathcal{N}(\mu - x^*, \sigma)$, we can determine the maximum containment level $p_{\max}$ by solving

$$p_{\max} = \max_k \frac{1}{2}\left[\operatorname{erf}\left(\frac{e^{-(a+bk)T}\tilde{x}_{T,\text{up}} - \mu + x^*}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{e^{-(a+bk)T}\tilde{x}_{T,\text{low}} - \mu + x^*}{\sigma\sqrt{2}}\right)\right], \tag{11}$$

where erf is the error function. This quantity tells us that we will always have $\Phi_{\tilde{\Xi},T} \le p_{\max}$. This implies that if $p_{\max} < 1$, we cannot solve **CCP** for arbitrary containment level $p^* \in (0, 1)$.

In the following numerical example, we will demonstrate a case where a simple linear state-feedback control law without feedforward control cannot solve the **CCP** for an arbitrary containment level.

Example 2.1: Consider a system that satisfies

$$\dot{x} = u, \quad x(0) = x_0, \tag{12}$$

where we assume that $\phi_{x_0} = \mathcal{N}(10, 1)$. Furthermore, consider the desired containment set $\Xi = [4, 5]$ with a relevant transient time $T = 5$.

If we consider a non-zero desired equilibrium point of $x^* = 4$. Using the linear feedback controller as given before, we can obtain the gain $k < 0$ for any desired containment level $p^* \in (0, 1)$. For instance, by taking $k = -3.6776$, we get p^* very close to 1. Since $k < 0$, the closed-loop system is stable which implies that $x(t)$ converges to x^* as $t \to \infty$. Hence we achieve both **CCPa** and **CCPb**.

On the other hand, if we change the desired steady-state to $x^* = 0$ then the aforementioned feedback control will no longer solve **CP1** for arbitrary p^*. The main reason for this is that we can no longer design k such that **CCPa** is met for some desired containment level p^*. Indeed, solving (11) results in $p_{\max} = 0.7359 < 1$ which occurs for $k = -0.1617$. Hence, we can no longer find a feasible solution that satisfies both **CCPa** and **CCPb** for any $p^* > p_{\max}$.

In Example 2.1, we have shown that the previous simple linear feedback control law only allows us to solve the **CCP** for specific cases. Particularly, achieving a desired containment level p^* close to 1 may not be possible at all, even for the case of a simple integrator. This problem is exacerbated when we are interested in solving **SCP**.

3. Containment Control Problem for linear systems

We will start our exposition by considering the **CCP** in a linear time-invariant setting. The system (1) becomes

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \tag{13}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are the system matrices and x_0 is a random variable defined on $\mathbb{R}^n$. We are now ready to present our first result, which is an extension of Proposition 1 in Dresscher and Jayawardhana (2017a). The result below considers general distributions for the initial condition and state convergence to a desired state trajectory, while the previous result only considered normal distributions and convergence to a point. We remark that controllability in this context refers to the standard controllability property for LTI systems.

Theorem 3.1: *Consider the system as in* (13). *Let $T > 0$ be the given transient terminal time and x_d be the desired trajectory. Assume that the pair (A, B) is controllable and there exists a finite $\tau > T$ such that x_d is a solution to* (13) *(with an admissible input signal $u_d(t)$) for all $t \geq \tau$. Then the **CCP** is solvable for any p^* where $d(\cdot,\cdot) = d_E(\cdot,\cdot)$ is the Euclidean distance and the set $\Xi \subset \mathbb{R}^n$ is compact, connected and non-empty.*

Proof: The proof of the theorem follows a similar line as the proof of Proposition 1 in Dresscher and Jayawardhana (2017a). Consider the control law

$$u(t) = K(x(t) - x_r(t)) + u^*(t), \tag{14}$$

where x_r and u^* are the tracking reference signal and additional feedforward input signal to be designed.

Let us first define two closed balls. The first one is centred in ϵ_1 and has a radius κ_1 (which we will denote by $\mathbb{B}_{\kappa_1}(\epsilon_1)$). For this ball, ϵ_1 and κ_1 are such that

$$\int_{\mathbb{B}_{\kappa_1}(\epsilon_1)} \phi_{x_0}(\xi)\mathrm{d}\xi \geq p^*. \tag{15}$$

We will denote the second ball by $\mathbb{B}_{\kappa_2}(\epsilon_2)$. Since Ξ is compact, connected and non-empty, we can choose ϵ_2 and κ_2 such that $\mathbb{B}_{\kappa_2}(\epsilon_2) \subseteq \Xi$. Furthermore, we require $\kappa_1 > \kappa_2$. Define $x_r(t)$ and $u^*(t)$ with the following properties: (i) $x_r(t) = x_d(t)$, for all $t \geq \tau$; (ii) $x_r(0) = \epsilon_1$; (iii) $x_r(T) = \epsilon_2$; and (iv) $\dot{x}_r(t) = Ax_r(t) + Bu^*(t)$. Note that since the pair (A, B) is controllable, we can always find a control signal u^* that can bring the state from ϵ_1 at time 0 to ϵ_2 at time T, and subsequently, to $x_d(\tau)$ at τ. Furthermore, since $x_d(t)$ is a solution to (13) for $u_d(t)$ and $t \geq \tau$, we can let $u^*(t) = u_d(t)$ for $t \geq \tau$. Using such u^*, the tracking reference signal x_r is then given by the solution z of

$$\dot{z} = Az + Bu^*, \quad z(0) = \epsilon_1. \tag{16}$$

Define now $\zeta = x - x_r$ as the error signal between the state and the tracking reference signal. Note that with such coordinate transformation, if $\zeta(T) \in \mathbb{B}_{\kappa_2}(0)$ then, since $x_r(T) = \epsilon_2$, it implies that $x(T) \in \mathbb{B}_{\kappa_2}(\epsilon_2)$, which is a subset of Ξ. Also, it follows that $\zeta_0 \in \mathbb{B}_{\kappa_1}(0)$ implies that $x_0 \in \mathbb{B}_{\kappa_1}(\epsilon_1)$. Accordingly, the dynamics of the error signal where we have applied the control law (14) are given by

$$\dot{\zeta} = (A + BK)\zeta, \quad \zeta(0) = \zeta_0. \tag{17}$$

Let us now define a contraction exponential rate constant $\lambda = -\frac{1}{T}\ln(\kappa_2/\kappa_1)$. In the following, we will design K so that $\mathbb{B}_{\kappa_1}(0)$ under the closed-loop dynamics (17) will be contracted with an exponential rate of λ, to $\mathbb{B}_{\kappa_2}(0)$ at time T.

From the pair (A, B) being controllable, it follows that we can design K such that $A + BK$ has eigenvalues whose real part is less than $-\lambda$ (for example, by the well-known pole-placement method). This implies that

$$\|\zeta(t)\| \leq e^{-\lambda t}\|\zeta(0)\|$$

holds for all initial condition $\zeta(0)$. By our choice of λ as given before and by considering initial conditions along the boundary of $\mathbb{B}_{\kappa_1}(0)$ (having a Euclidean norm of κ_1),

$$\|\zeta(T)\| \leq \frac{\kappa_2}{\kappa_1}\|\zeta(0)\| = \kappa_2.$$

Hence,

$$\zeta(0) \in \mathbb{B}_{\kappa_1}(0) \Rightarrow \zeta(T) \in \mathbb{B}_{\kappa_2}(0) \Rightarrow x(T) \in \mathbb{B}_{\kappa_2}(\epsilon_2) \subseteq \Xi.$$

And, since (15) holds, we obtain

$$\Phi_{\Xi,T} \geq \int_{\mathbb{B}_{\kappa_1}(0)} \phi_{\zeta_0}(\xi)\,\mathrm{d}\xi \geq p^*.$$

In other words, **CCPa** is satisfied. Additionally, since we have $x_r(t) = x_d(t)$ for $t \geq \tau$, the following asymptotic property holds:

$$\lim_{t\to\infty} \zeta(t) = 0 \Rightarrow \lim_{t\to\infty} d_E(x(t), x_d(t)) = 0.$$

In other words, **CCPb** holds. This concludes the proof. ■

In the special case where $x_d(t)$ is a constant point, the result above becomes more simple. This is presented in the following corollary.

Corollary 3.2: *Consider the system as in* (13). *Assume that the pair (A, B) is controllable and that $x_d(t) = x^*$. Then, the **CCP** is solvable for any T and p^*, where $d(\cdot,\cdot) = d_E(\cdot,\cdot)$, the Euclidean distance, and Ξ is compact, connected and non-empty.*

Proof: The proof is almost identical to the proof of Theorem 3.1. Here, we let the reference signal satisfy the following: (i) $\lim_{t\to\infty} x_r(t) = x^*$; (ii) $x_r(0) = \epsilon_1$; (iii) $x_r(T) = \epsilon_2$; and (iv)

$\dot{x}_r(t) = Ax_r(t) + Bu^*(t)$. Notice that the new characteristic (i) is always achievable for a controllable system. The proof of **CCPa** remains unchanged. For **CCPb** we now obtain

$$\lim_{t\to\infty} \zeta(t) = 0 \Rightarrow \lim_{t\to\infty} d_E(x(t), x^*) = 0.$$

Hence, **CCPb** is satisfied. This concludes the proof. ■

4. Containment Control Problem for nonlinear systems

In this section, we will extend the result of Section 4 to the nonlinear case using recent results in contraction theory. For interested readers, we present relevant contraction results in Appendix. The overall main idea is that we use contraction results for quantifying the rate of decay among all trajectories which include the target trajectory. For further background reading on this subject, we refer the interested reader to Lohmiller and Slotine (1998), Jouffroy and Fossen (2010), Andrieu et al. (2016) andAndrieu et al. (2015).

Let us now proceed by presenting our solution to **CCP** for the nonlinear systems case. We can apply the contraction-based control design by implementing a control law for the system (1), such that a partial contraction with a desired contraction rate w.r.t. a desired reference trajectory $x_r(t)$ is achieved. The design procedure is accordingly an iterative process which should yield three components: (i) a reference trajectory x_r that starts close to the initial conditions and satisfies $x_r(T) \in \Xi$ and $\lim_{t\to\infty} d_F(x_r(t), x_d(t)) = 0$, with d_F a Finsler distance as in Definition A.2 in Appendix; (ii) a control law that yields a closed-loop system having x_r as a solution, and; (iii) a Finsler–Lyapunov function that yields a desired contraction rate λ for the closed-loop system.

We apply an adaptation of Lemma A.4 in Appendix as presented in Jouffroy and Fossen (2010) and Wang and Slotine (2005) to facilitate the control law design. Accordingly, for a given system (1), we assume the existence of a control law $u(t) = k(x, x_r, t)$ that causes the closed-loop system, given by

$$\dot{x} = f_c(x, x_r, t), \tag{18}$$

to be partially contracting with contraction region $C \subseteq X$. We require C to be such that $X_0 \subset C$. Furthermore, we require a minimum rate λ for all initial conditions belonging to a specific set that satisfies a condition similar to (15). For such a control law, we obtain λ, the choice of a set of initial conditions and the reference trajectory x_r as control design parameters for achieving the two control objectives in **CCP**.

Before presenting a particular design for λ and x_r in the following theorem, we will define an open ball induced by the Finsler distance d_F. For a given $\kappa > 0$ and a Finsler structure F satisfying (1) to (4) of Definition A.1 in Appendix, we define the ball with radius κ centred at x_1 by

$$\mathbb{D}_\kappa(x_1) = \{x_2 \in X \mid d_F(x_1, x_2) < \kappa\}. \tag{19}$$

The following theorem is an extension to the proposition in Desscher and Jayawardhana (2017b) (where we consider nonlinear systems that are affine w.r.t. the input) to general nonlinear systems.

Theorem 4.1: *Consider the system (1) with the control law $u(t) = k(x, x_r, t)$. Suppose that the closed-loop system defined by (18) is contracting w.r.t. x and a contraction region $C \supseteq X_0$, and that there exist points $\epsilon_1 \in C$, $\epsilon_2 \in \Xi$, a reference trajectory x_r and constants $\kappa_1, \kappa_2 > 0$ such that the following conditions hold.*

(1) *The reference signal x_r satisfies*

$$\dot{x}_r = f_c(x_r, x_r, t) \tag{20}$$

for all $t \geq 0$, $x_r(0) = \epsilon_1$, $x_r(T) = \epsilon_2$ and $\lim_{t\to\infty} d_F(x_r(t), x_d(t)) = 0$.

(2) *There are two sets $\mathbb{D}_{\kappa_1}(\epsilon_1)$ and $\mathbb{D}_{\kappa_2}(\epsilon_2)$ satisfying*

$$\int_{\mathbb{D}_{\kappa_1}(\epsilon_1)} \phi_{x_0}(\xi)\, \mathrm{d}\xi \geq p^* \tag{21}$$

and $\mathbb{D}_{\kappa_2}(\epsilon_2) \subseteq \Xi$.

(3) *The contraction rate λ satisfies*

$$\lambda \geq -\frac{\ln\left(\frac{\kappa_2}{\kappa_1}\right)}{T}, \tag{22}$$

for all $x_0 \in \mathbb{D}_{\kappa_1}(\epsilon_1)$.

*Then, the control law $u(t) = k(x, x_r, t)$ solves the **CCP**.*

Proof: The proof of this theorem is similar to the proof in Dresscher and Jayawardhana (2017b). We will show that the initial ball $\mathbb{D}_{\kappa_1}(x_r(0))$, which has the desired minimum cumulative distribution, will contract to $\mathbb{D}_{\kappa_2}(x_r(T))$, which is contained in the desired containment set Ξ, at the transient time T.

By the hypothesis of the proposition, the closed-loop system is contracting w.r.t x. Furthermore, since x_r is an admissible solution to the system, by Lemma A.4, this implies that all trajectories starting in C converge to x_r. Notice that x_r is reachable since it is a solution to the closed-loop system which has a contraction property. We accordingly have that all trajectories starting in $\mathbb{D}_{\kappa_1}(\epsilon_1) \subset C$ converge with an exponential rate λ. Hence,

$$d_F(x_r(t), x(t)) \leq d_F(x_r(0), x_0)e^{-\lambda t}, \tag{23}$$

for all $x_0 \in \mathbb{D}_{\kappa_1}(\epsilon_1)$. For all initial conditions $x_0 \in \mathbb{D}_{\kappa_1}(\epsilon_1)$, as $x_r(0) = \epsilon_1$, we obtain that $d_F(x_r(0), x_0) \leq \kappa_1$. Hence

$$d_F(x_r(t), x(t)) \leq \kappa_1 e^{-\lambda t}. \tag{24}$$

Thus at time T, by the hypothesis on λ as in (22),

$$d_F(x_r(T), x(T)) \leq \frac{\kappa_2}{\kappa_1}\kappa_1 \Rightarrow d_F(x_r(T), x(T)) \leq \kappa_2. \tag{25}$$

Hence, for this λ we have

$$x_0 \in \mathbb{D}_{\kappa_1}(x_r(0)) \Rightarrow x(T) \in \mathbb{D}_{\kappa_2}(x_r(T)) \subseteq \Xi. \tag{26}$$

It follows that

$$\int_\Xi \phi_{x_0,T}(\xi)\, \mathrm{d}\xi \geq \int_{\mathbb{D}_{\kappa_2}(x_r(T))} \phi_{x_0,T}(\xi)\, \mathrm{d}\xi \geq p^*, \tag{27}$$

which implies that the first control objective **CCPa** is satisfied.

It remains to show that all trajectories converge to x_d as $t \to \infty$. Firstly, since C is such that it contains X_0, we have the contraction property for all initial conditions. Secondly, since the reference signal x_r is such that $\lim_{t\to\infty} d_F(x_r(t), x_d(t)) = 0$, it is sufficient to show that all contracting trajectories converge to x_r. From the partial contraction property, we have

$$\lim_{t\to\infty} d_F(x(t), x_r(t)) = 0 \Rightarrow \lim_{t\to\infty} d_F(x(t), x_d(t)) = 0, \quad (28)$$

for all $x_0 \in C$. Hence, **CCPb** is satisfied. This concludes the proof. ∎

5. Shape Control Problem for linear systems

We will now consider control design suitable for solving the **SCP**, where we want to obtain a desired closeness (which is defined using the Hellinger distance) to a prescribed distribution shape during the transient time. The problem is significantly more complicated than the **CCP** considered in previous sections, due to the well specified requirements on the shape of the pdf of the state during the transient. We will first consider the case when the initial pdf ϕ_{x_0} and the desired pdf ϕ_d are linearly matching, followed by an approach suitable for nonlinearly matching pdfs.

5.1 *Linearly matching initial and desired pdfs*

Before we continue, let us formally define our notion of (linearly) matching pdfs. This definition is based on matching procedures that are used in image processing applications (Inamdar et al., 2008; Shapiro & Stockman, 2001), which are typically performed through well-known operations of rotations, scaling, translation, shearing and/or reflections (Shapiro & Stockman, 2001).

Definition 5.1 (Matching probability density functions): For a given $Y \subset \mathbb{R}^n$, we call two pdfs $\phi : Y \to \mathbb{R}_{\geq 0}$ and $\varphi : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ (linearly) matching with respect to Y if there exist $\eta \in \mathbb{R}^{n\times 1}$, $\beta \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$ such that

$$\phi(x) = \lambda\varphi(\beta x + \eta) \quad (29)$$

holds for all $x \in Y$.

We note that the linearity refers to the application of a linear affine state transformation for matching both nonlinear maps ϕ and φ. For such linearly matching pdfs, we are now ready to present our controller design that can solve the **SCP**, with a bounded ℓ. This bound is then dependent on the matching of the pdfs (expressed through λ, β and η) and the system equations.

Theorem 5.2: *Assume that the hypothesis of Theorem* 3.1 *holds. Suppose that:* (i) *the target distribution at time $T > 0$ is given by ϕ_d and that ϕ_{x_0} and ϕ_d are matching with respect to X_0 for some $\eta \in \mathbb{R}^{n\times 1}$, invertible $\beta \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$; and* (ii) *there exists a finite $\tau > T$, such that $x_d(t)$ is a solution to the system* (13), *with an admissible input signal $u_d(t)$, for all $t \geq \tau$. Then, the* **SCP** *is solvable for ℓ and $K \in \mathbb{R}^{n\times m}$ satisfying*

$$\ell \geq \min_{\{K|\mathrm{spec}(A+BK)\in\mathbb{C}_-\}} \times \sqrt{1 - \int_X \sqrt{\frac{\phi_{x_0}(\tilde{\beta}^{-1}(\xi - \tilde{\eta}))\phi_{x_0}(\beta^{-1}(\xi - \eta))}{\tilde{\lambda}\lambda}}\, \mathrm{d}\xi}, \quad (30)$$

where

$$\tilde{\lambda} = \int_X \phi_{x_0}(\tilde{\beta}^{-1}(\xi - \tilde{\eta}))\, \mathrm{d}\xi, \quad (31)$$

$$\tilde{\beta} = e^{(A+BK)T}, \quad (32)$$

$$\tilde{\eta} = \mu_d - e^{(A+BK)T}\mu, \quad (33)$$

with μ the mean value of ϕ_{x_0} and μ_d the mean value of ϕ_d.

Proof: We will first prove the fulfilment of **SCPa**. We again consider the control law $u(t) = K(x(t) - x_r(t)) + u^*$. Similar to Theorem 3.1, we define $x_r(t)$ and $u^*(t)$ with the following properties: (i) $x_r(t) = x_d(t)$, for all $t \geq \tau$; (ii) $x_r(0) = \mu$; (iii) $x_r(T) = \mu_d$; and (iv) $\dot{x}_r(t) = Ax_r(t) + Bu^*(t)$. As before, since the pair (A, B) is controllable, we can always find a control signal u^* that can bring the state from μ at time 0 to μ_d at time T, and subsequently, to $x_d(\tau)$ at τ. Additionally, since $x_d(t)$ is a solution to (13) for $u_d(t)$ and $t \geq \tau$, we can let $u^*(t) = u_d(t)$ for $t \geq \tau$. For this control system, we define an error like signal as $\zeta(t) = x(t) - x_r(t)$. Also, due to ϕ_{x_0} and ϕ_d being matching, we have that $\lambda^{-1}\phi_{x_0}(x) = \phi_d(\beta x + \eta)$, for all $x \in X_0$. Let us then write a similar identity for $\phi_{x_0,T}$ as

$$\tilde{\lambda}^{-1}\phi_{x_0}(x(0)) = \phi_{x_0,T}(\tilde{\beta}x(0) + \tilde{\eta}).$$

Substituting our choices of $\tilde{\beta}$ and $\tilde{\eta}$ as given by (32) and (33) yields

$$\tilde{\lambda}^{-1}\phi_{x_0}(x(0)) = \phi_{x_0,T}\left(e^{(A+BK)T}\zeta(0) + \mu_d\right).$$

Notice that we furthermore have $x(T) = x_r(T) + \zeta(T)$ which by design satisfies the solution

$$x(T) = e^{(A+BK)T}\zeta(0) + \mu_d,$$

and hence we have

$$\tilde{\lambda}^{-1}\phi_{x_0}(x(0)) = \phi_{x_0,T}(x(T)) = \phi_{x_0,T}(\tilde{\beta}x(0) + \tilde{\eta}),$$

for all $x(0) \in X_0$. Notice that $\phi_{x_0,T}$ is accordingly the pdf of the state at time T for the closed-loop system. We have thus obtained that for all $x \in X_0$: (i). $\tilde{\lambda}^{-1}\phi_{x_0}(x) = \phi_{x_0,T}(\tilde{\beta}x + \tilde{\eta})$; and (ii). $\lambda^{-1}\phi_{x_0}(x) = \phi_d(\beta x + \eta)$ hold. Subsequently, we can define two coordinate transformations $y = \tilde{\beta}x + \tilde{\eta}$ and $z = \beta x + \eta$ whose inverses are given by $x = \tilde{\beta}^{-1}(y - \tilde{\eta})$ and $x =$

$\beta^{-1}(z-\eta)$, respectively. Accordingly, we obtain

$$\tilde{\lambda}^{-1}\phi_{x_0}(\tilde{\beta}^{-1}(y-\tilde{\eta})) = \phi_{x_0,T}(y), \tag{34}$$

$$\lambda^{-1}\phi_{x_0}(\beta^{-1}(z-\eta)) = \phi_d(z). \tag{35}$$

The inverse of $\tilde{\beta}$ always exists and we assume $z, y \in X$ for all $x \in X_0$. It then follows directly from (31) that $\phi_{x_0,T}$ satisfies

$$\int_X \phi_{x_0,T}(\xi)\,\mathrm{d}\xi = \frac{1}{\tilde{\lambda}}\int_X \phi_{x_0}(\tilde{\beta}^{-1}(\xi-\tilde{\eta}))\,\mathrm{d}\xi = 1. \tag{36}$$

Plugging (34) and (35) in (5) yields

$$d_h(\phi_{x_0,T},\phi_d) = \sqrt{1-\int_X\sqrt{\frac{\phi_{x_0}(\tilde{\beta}^{-1}(\xi-\tilde{\eta}))\phi_{x_0}(\beta^{-1}(\xi-\eta))}{\tilde{\lambda}\lambda}}\,\mathrm{d}\xi}.$$

Accordingly, for our simple choices of the control law, reference signal x_r and control signal u^*, we can always find a matrix K such that we satisfy **SCPa** for a maximum distance ℓ satisfying (30).

We are now left to prove **SCPb**. The proof for the asymptotic convergence to the reference signal follows directly from the design restriction on K that requires $(A+BK)$ to be Hurwitz. Additionally, since we have that $x_r(t) = x_d(t)$ for $t \geq \tau$, the asymptotic property holds:

$$\lim_{t\to\infty}\zeta(t) = 0 \Rightarrow \lim_{t\to\infty} d_E(x(t),x_d(t)) = 0.$$

In other words, **SCPb** holds. This concludes the proof. ■

As shown in Theorem 5.2, for given initial and target distributions ϕ_{x_0} and ϕ_d that are matching, there is a lower bound on achievable ℓ. However, there are cases when the lower bound is equal to zero for some specific combinations of A, B, ϕ_{x_0}, ϕ_d and T. In the following corollary, we show a particular example of such case.

Corollary 5.3: *Assume that the hypotheses in Theorem 5.2 holds. Suppose that there exists a $K \in \mathbb{R}^{n\times m}$ such that (32)and (33) satisfy $\tilde{\beta} = \beta$ and $\tilde{\eta} = \eta$. Then, the **SCP** is solvable for any $\ell \geq 0$.*

Proof: The result follows directly from (30) and (36), when $\tilde{\lambda} = \lambda$. Rewriting (36) for ϕ_d yields

$$\int_X \phi_d(\xi)\,\mathrm{d}\xi = \frac{1}{\lambda}\int_X \phi_{x_0}(\beta^{-1}(\xi-\eta))\,\mathrm{d}\xi = 1, \tag{37}$$

where $\beta = \tilde{\beta}$ and $\eta = \tilde{\eta}$ which also implies that $\lambda = \tilde{\lambda}$. Hence, (30) reduces to

$$\ell \geq \sqrt{1-\frac{1}{\lambda}\int_X \phi_{x_0}(\beta^{-1}(\xi-\eta))\,\mathrm{d}\xi} = 0.$$

Accordingly, **SCPa** holds for $\ell \geq 0$. The property **SCPb** follows from Theorem 5.2. ■

5.2 Nonlinearly matching initial and desired pdfs

In this section, we will propose a solution to the **SCP** for cases when the initial and the desired pdfs are nonlinearly matching. In this case, the results from the previous subsection can be extended to the situation when there exists nonlinear mappings that gives the relation between both pdfs.

Definition 5.4 (Nonlinearly matching probability density functions): For a given $Y, Z \subset \mathbb{R}^n$, we call two pdfs $\phi : Y \to \mathbb{R}_{\geq 0}$ and $\varphi : Z \to \mathbb{R}_{\geq 0}$ nonlinearly matching with respect to the tuple (Y, Z) if there exist a diffeomorphic mapping $\Psi : Y \to Z$, a function $\delta : Z \to \mathbb{R}_{>0}$ such that

$$\phi(x) = \delta\big(\Psi(x)\big)\varphi\big(\Psi(x)\big) \tag{38}$$

holds for all $x \in Y$.

Since we will later use Ψ in the coordinate transformation, the function δ becomes a normalising function that corrects for the elongation of the pdf in the transformed state space via the mapping Ψ. In this case, when we consider $Y = X$ (with X being the original state space domain) in the above definition, we have

$$\int_W \phi(\xi)\,\mathrm{d}\xi = \int_W \delta(\Psi(\xi))\varphi(\Psi(\xi))\,\mathrm{d}\xi, \tag{39}$$

holds for all $W \subseteq X$.

Theorem 5.5: *Assume the hypothesis of Theorem 5.2 holds. Suppose that; (i) the pdfs ϕ_{x_0} and ϕ_d are nonlinearly matching for a diffeomorphic map $\Psi : X \to X$, $\delta : X \to \mathbb{R}_{\geq 0}$; and (ii) there exists a finite $\tau > T$, such that $x_d(t)$ is a solution to the system (13) for all $t \geq \tau$ and for a corresponding admissible input signal $u_d(t)$. Then, the **SCP** is solvable for ℓ and $K \in \mathbb{R}^{n\times m}$ satisfying*

$$\ell \geq \min_{\{K|\mathrm{spec}(A+BK)\in\mathbb{C}_-\}} \times \sqrt{1-\int_X\sqrt{\frac{\phi_{x_0}(\tilde{\beta}^{-1}(\xi-\tilde{\eta}))\phi_{x_0}(\Psi^{-1}(\xi))}{\tilde{\lambda}\delta(\xi)}}\,\mathrm{d}\xi}, \tag{40}$$

where

$$\tilde{\lambda} = \int_X \phi_{x_0}(\tilde{\beta}^{-1}(\xi-\tilde{\eta}))\,\mathrm{d}\xi, \tag{41}$$

$\tilde{\beta} = e^{(A+BK)T}$, $\tilde{\eta} = \mu_d - e^{(A+BK)T}\mu$ with μ the mean value of ϕ_{x_0} and μ_d the mean value of ϕ_d.

Proof: The proof follows the same lines as the proof for Theorem 5.2. We have a controllable linear system with two nonlinearly matching pdfs ϕ_{x_0} and ϕ_d. Subsequently, we again consider $u(t) = K(x(t) - x_r(t)) + u^*$ and design $x_r(t)$ and u^* as in Theorem 5.2. Accordingly, (34) and (35) become

$$\tilde{\lambda}^{-1}\phi_{x_0}(\tilde{\beta}^{-1}(y-\tilde{\eta})) = \phi_{x_0,T}(y),$$

$$\delta(z)^{-1}\phi_{x_0}(\Psi^{-1}(z)) = \phi_d(z).$$

Using these substitutions, we obtain

$$d_h(\phi_{x_0,T},\phi_d) = \sqrt{1 - \int_X \sqrt{\frac{\phi_{x_0}(\tilde{\beta}^{-1}(\xi - \tilde{\eta}))\phi_{x_0}(\Psi^{-1}(\xi))}{\tilde{\lambda}\delta(\xi)}}\,\mathrm{d}\xi}. \tag{42}$$

And hence, we can always find a K to achieve a maximum distance ℓ as in (40), and thus satisfying **SCPa**. Furthermore, since K is such that $(A + BK)$ is Hurwitz, we obtain

$$\lim_{t\to\infty} \zeta(t) = 0 \Rightarrow \lim_{t\to\infty} d_E(x(t), x_r(t)) = 0.$$

Furthermore, since $x_d(t)$ is a solution to (13) for input $u_d(t)$, we can let $x_r(t) = x_d(t)$ for all $t \geq \tau$ and we hence also have $\lim_{t\to\infty} d_E(x(t), x_d(t)) = 0$. **SCPb** thus holds. This concludes the proof. ■

6. CCP controller simulation for a nonlinear robotic manipulator

In this section, we will evaluate a nonlinear contraction-based controller design for the **CCP**, applied a standard second-order mechanical system operating with 3-DOF, which is a SCARA robot as presented in Dresscher and Jayawardhana (2017b) and Reyes-Báez et al. (2017). The results are a short display of the results of our simulation, we refer the interested reader to Dresscher and Jayawardhana (2017b) for the full result. The controller design that we implement for this simulation is in accordance with Theorem 4.1.

6.1 Dynamics and controller design

The robot operates on the manifold $\mathcal{X} = \mathcal{Q} \times \mathbb{R}^3$, with states $q \in \mathcal{Q} = \mathcal{S}^1 \times \mathcal{S}^1 \times \mathbb{R}$, where $\mathcal{S}^1$ the unitary circumference, and $p \in \mathbb{R}^3$. Here, $q^\top = [\theta_1, \theta_2, z]$ is the generalised position, and $p^\top = [p_{\theta_1}, p_{\theta_2}, p_z] = M(q)\dot{q}$ is the generalised momentum, with $M(q) = M^\top(q)$ the inertia matrix, and $u^\top = [\tau_1, \tau_2, f]$ the generalised force. The system satisfies the port-Hamiltonian form

$$\begin{bmatrix}\dot{q}\\ \dot{p}\end{bmatrix} = \begin{bmatrix}0_3 & I_3\\ -I_3 & -D(q)\end{bmatrix}\begin{bmatrix}\frac{\partial H}{\partial q}(q,p)\\ \frac{\partial H}{\partial p}(q,p)\end{bmatrix} + \begin{bmatrix}0_3\\ G(q)\end{bmatrix}u, \tag{43}$$

where $H(q,p)$ is the Hamiltonian function, $D(q) = D^\top(q) : \mathcal{Q} \to \mathbb{R}^{3\times 3}_{\geq 0}$ is the damping matrix and $G(q) : \mathcal{Q} \to \mathbb{R}^{3\times 3}$ is the input matrix. For the Hamiltonian function we have the total energy as

$$H(q,p) = \frac{1}{2}p^\top M^{-1}(q)p + V(q) \tag{44}$$

with $V(q) = (m_1 + m_2 + m_3)gz$ the potential energy, where m_1, m_2 and m_3 are the masses of the robot manipulator links. For the mass matrix we have

$$\mathbf{M}(q) = \begin{bmatrix} M_{11} & M_{12} & 0\\ M_{12} & m_3 l_2^2 & 0\\ 0 & 0 & (m_1+m_2+m_3)g\end{bmatrix}, \tag{45}$$

where

$$M_{11} = (m_2 + m_3)l_1^2 + m_3 l_2^2 + 2m_3 l_1 l_2 \cos\theta_2,$$
$$M_{12} = m_3 l_2^2 + m_3 l_1 l_2 \cos\theta_2.$$

We assume stochastic initial conditions for the two rotational joints, satisfying a normal distribution. Hence, we have $q_0 \sim \mathcal{N}(\mu_q, \Sigma_q)$, where

$$\mu_q = \begin{bmatrix}\mu_{q,1}\\ \mu_{q,2}\\ \mu_{q,3}\end{bmatrix}, \quad \Sigma_q = \begin{bmatrix}\sigma_{q,1}^2 & 0 & 0\\ 0 & \sigma_{q,2}^2 & 0\\ 0 & 0 & 0\end{bmatrix}. \tag{46}$$

We take the system to be idle upon initialisation, e.g. $p_0^\top = [0,0,0]$. The initial conditions $x_0^\top = [q_0, p_0]^\top \sim \mathcal{N}(\mu, \Sigma)$ then satisfy

$$\mu = \begin{bmatrix}\mu_q\\ 0_{3\times 1}\end{bmatrix}, \quad \Sigma = \begin{bmatrix}\Sigma_q & 0_3\\ 0_3 & 0_3\end{bmatrix}. \tag{47}$$

In this example, we will use $D = \mathrm{diag}([0.2, 0.2, 0.2])$, $G = \mathbf{I}_3$, $[m_2, m_2, m_3] = [1.5, 1, 0.5]$, $[l_1, l_2] = [2, 1]$, $g = 9.81$, $\mu = [1,0,0]^\top$ and $[\sigma_{q,1}, \sigma_{q,2}] = [1,1]$. We consider for the desired trajectory $q_d = [\sin(t)+1, \sin(t), \sin(t)]^\top$ and $p_d(t) = M(q_d(t))\dot{q}_d(t)$. For the reference signal, we design q_r s.t. $q_r(0) = \mu_q$, $q_r(T) = \epsilon_q \in \Xi$ and $\lim_{t\to\infty} q_r(t) = q_d(t)$. We will discuss characteristics of Ξ shortly. Assume Ξ is such that we can take $\epsilon_q = [\sin(T)+1, \sin(T), \sin(T)]^\top$ and since we also have that $q_d(0) = \mu$, we can conveniently let $q_r(t) = q_d(t)$.

Subsequently, the error system is given by

$$\zeta := \begin{bmatrix}\tilde{q}\\ \omega\end{bmatrix} = \begin{bmatrix}q - q_d\\ p - p_r\end{bmatrix}, \tag{48}$$

where p_r is a momentum reference signal, to be defined. The dynamics of $\tilde{q}$ are given by

$$\dot{\tilde{q}} = M^{-1}(\omega + q_d)p - M^{-1}(q_d)p_d. \tag{49}$$

We define $p_r = p_{d\omega} - \Lambda\tilde{q}$, with $p_{d\omega} = M(\tilde{q} + q_d)\dot{q}_d$ and $-\Lambda = \Lambda$ Hurwitz. Hence, we obtain the properties $\lim_{t\to\infty} q(t) = q_d(t)$ and $\lim_{t\to\infty} p_r(t) = p_d(t)$. We then take $\epsilon_p = p_r(T)$ and accordingly obtain, in reference to condition 1) in Theorem 4.1, that the reference signal satisfies $x_r(0) = (q_r(0)^\top, p_r(0)^\top)^\top = \mu$, $x_r(T) = (\epsilon_q^\top, \epsilon_p^\top)^\top$ and $\lim_{t\to\infty} d_F(x_r(t), x_d(t)) = 0$. The full error dynamics are given by

$$\begin{aligned}\dot{\tilde{q}} &= M^{-1}(\tilde{q} + q_d)(\omega - \Lambda\tilde{q})\\ \dot{\omega} &= -\left[\frac{\partial H}{\partial q}(q,p) + D\frac{\partial H}{\partial p}(q,p) - u + \dot{p}_r\right].\end{aligned} \tag{50}$$

Accordingly, we obtain the closed-loop error system by applying the control law

$$
\begin{aligned}
u &= u_{eq} + u_{at}, \\
u_{eq} &= \dot{p}_r + \frac{\partial H}{\partial q}(q, p_r) + D\frac{\partial H}{\partial p}(q, p_r), \\
u_{at} &= -K_d\frac{\partial H}{\partial p}(q, \omega) - M^{-1}(q)\Lambda\tilde{q} + \frac{\partial}{\partial q}(p_r^\top M^{-1}(q)\omega),
\end{aligned} \tag{51}
$$

where K_d is such that

$$
D + K_d + \frac{1}{2}I_3 - \frac{1}{4}(M^{-1} + M) > 0. \tag{52}
$$

Our system then specifies the contraction properties proven in, and for a virtual system as provided in Reyes-Báez et al. (2017). This virtual system admits x and $x_r = (q_d^\top, p_r^\top)^\top$ as solutions and we have therefore satisfied condition (1) in Theorem 4.1. The candidate *Finsler–Lyapunov function* for this virtual system, as in Definition A.1 in Appendix, is given by

$$
V_F(x_v, \delta x_v) = \frac{1}{2}\delta x_v^\top \Theta^\top P(\zeta)\Theta\delta x_v, \tag{53}
$$

with $\delta x_v = [\delta q_v, \delta p_v]^\top$ as in (A2),

$$
\Theta = \begin{bmatrix} I_3 & 0_3 \\ \Lambda & I_3 \end{bmatrix}, \tag{54}
$$

$$
P(\zeta) = \begin{bmatrix} \Lambda & 0_3 \\ 0_3 & M^{-1}(\tilde{q} + q_d) \end{bmatrix}. \tag{55}
$$

It follows that the distance as in Definition A.2 in Appendix satisfies

$$
d_F(x, x_r) = \inf_{\Gamma(x,x_r)} \int_I \sqrt{\left(V_F(\gamma(s))\frac{\partial \gamma(s)}{\partial s}\right)} \mathrm{d}s. \tag{56}
$$

Lastly, we obtain the property

$$
d_F(x(t), x_r(t)) < d_F(x(0), x_r(0))e^{-\lambda t}, \tag{57}
$$

related to a rate λ as

$$
\lambda(\zeta) = \min \mathtt{eig}(P^{1/2}(\zeta)\Upsilon(\zeta)P^{1/2}(\zeta)), \tag{58}
$$

with

$$
\Upsilon(\zeta) = \begin{bmatrix} 2M^{-1}(\tilde{q} + q_d) & (M^{-1}(\tilde{q} + q_d) - I_3) \\ (M^{-1}(\tilde{q} + q_d) - I_3) & 2(D + K_d) \end{bmatrix}. \tag{59}
$$

Notice that the infimum of (A4) for $d_F(x, x_r)$ is given by $\gamma(s) = [q_d^\top, p_r^\top]^\top + \zeta s$, hence $\frac{\partial \gamma(s)}{\partial s} = \zeta$. Subsequently, we consider $\Lambda = \text{diag}\{2, 2, 2\}$, $T = 10$, $p^* = 0.7$ and Ξ such that

$$
\Xi = \{x \mid d_F(x, x_r(T)) \le \kappa_2\} = \mathbb{D}_{\kappa_2}(x_r(T)), \tag{60}
$$

for $\kappa_2 = 6$. The first part of condition 2 of Theorem 4.1 is accordingly satisfied. We obtain the distance set for our initial conditions by taking the bivariate standard normal pdf around μ for θ_1 and θ_2, given by

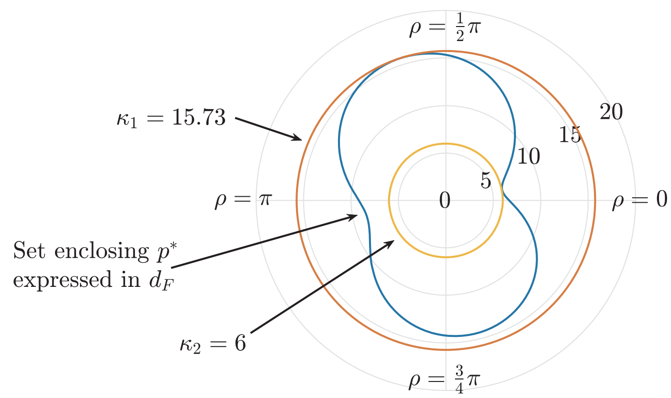

Figure 3. In this figure, we show the circumference of the two distance sets that are relevant for our simulation in Section 7; $\mathbb{D}_{\kappa_1}(x_r(0))$ and $\mathbb{D}_{\kappa_2}(x_r(T))$, and the contour of a set of initial conditions whose cumulative density is p^*. The angle (ρ) of this polar plot is interpretable with respect to θ_1 and θ_2, where $\rho = -\pi\mathbf{1}(-\theta_1 + \mu_1) + \tan^{-1}(\frac{\theta_1-\mu_1}{\theta_2-\mu_2})$, $\mathbf{1}(\cdot)$ being the step function.

$$
\phi_{x_0}(r, \theta) = \frac{r}{2\pi}e^{-0.5r^2}. \tag{61}
$$

Accordingly,

$$
\int_0^{1.5517} \frac{r}{2\pi}e^{-0.5r^2} \mathrm{d}r = 0.7. \tag{62}
$$

Subsequently, we map the contour of this radius through d_F to obtain the shape in Figure 3. These points can hence be captured by a distance set as in (19), $\mathbb{D}_{\kappa_1}(\mu)$, with $\kappa_1 = 15.73$. Accordingly, we have

$$
\int_{\mathbb{D}_{\kappa_1}(\mu)} \phi_{x_0}(\xi)\,\mathrm{d}\xi \ge p^*. \tag{63}
$$

We have accordingly satisfied condition (2) in Theorem 4.1 completely.

Let us now determine the minimal contraction rate λ and the corresponding K_d such that we achieve this rate. We find λ as

$$
\lambda \ge -\frac{\ln\left(\frac{\kappa_2}{\kappa_1}\right)}{T} = 0.0964. \tag{64}
$$

Accordingly, we can choose $K_d = \text{diag}\{3, 1, 25\}$ which is such that this minimal rate is always satisfied in accordance with (58), therefore satisfying condition (3) in Theorem 4.1. We are now ready to move on to the simulation results, as we have satisfied all criteria of Theorem 4.1.

6.2 Simulation results

The simulation results are shown in Figure 4, and 5. From Figure 4, it is easy to see that we have achieved

$$
\Phi_{\Xi,T} = 1 \ge p^* \tag{65}
$$

and thus satisfy **CCPa**. The performance of our system is strong, due to the following:

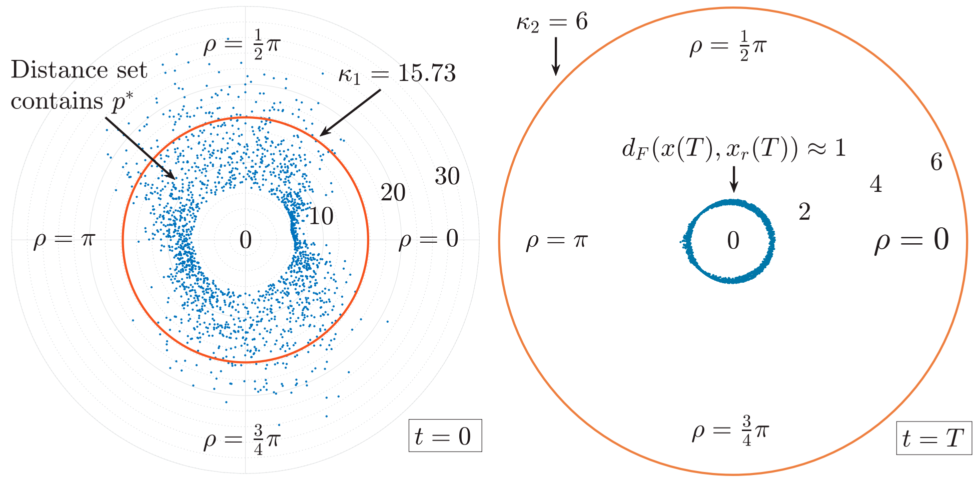

Figure 4. In this figure, we show the distances $d_F(x(t), x_r(t))$ of both the initial distribution at time $t = 0$ (left) and the distribution at time $t = T$ (right) for the simulation in Section 7. At time T, all trajectories are in the set $\Xi = \mathbb{D}_{\kappa_2}(x_r(T))$, hence we have achieved our desired performance $\Phi_{\Xi,T} \geq p^*$. The interpretation of the plot angle ρ is as in Figure 3.

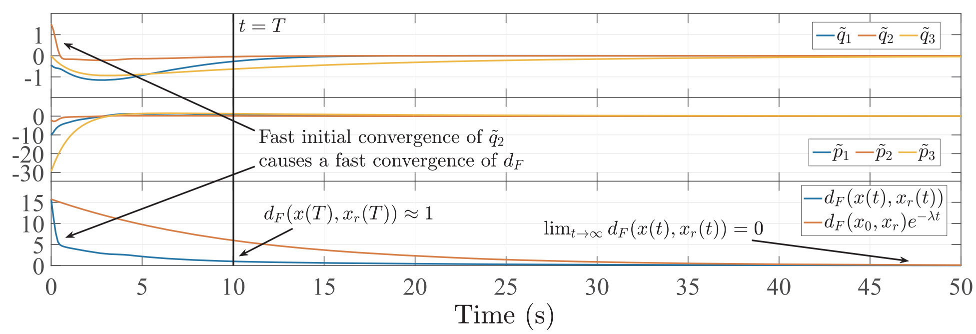

Figure 5. This figure depicts, for the simulation in Section 7, from top to bottom; (i) the time evolution of $\tilde{q}(t) = q(t) - q_d(t)$, (ii) $\tilde{p}(t) = p(t) - p_d(t)$ and (iii) the time evolution of the distance, for a trajectory satisfying $d_F(x(0), x_r(0)) = \kappa_1$. Furthermore, in the bottom plot we show the difference with the nominal decay, which is an effective upper bound. Lastly, the bottom plot shows convergence to a distance $d_F = 0$.

(1) The distance set $\mathbb{D}_{\kappa_1}(\mu)$ is greater than a marginal set covering p^* fraction of initial conditions.
(2) The rate λ is a minimal rate, but the other eigenvalues from (58) generally cause the convergence to be faster than this minimum.

In Figure 5, we show the convergence of an initial condition that satisfies $d_F(x_0, x_r) = \kappa_1$. The asymptotic convergence, satisfying **CCPb**, to the reference signal can clearly be seen, as well as the fast decay of the initial distance $d_F(x_0, x_r)$.

7. Numerical evaluation of SCP for matching pdfs

In this section we will numerically evaluate the result of Theorem 5.2. The example will show that we can only solve the **SCP** for $\ell \geq 0$ when the initial pdf ϕ_{x_0} and the desired pdf ϕ_d have very specific (linear) matching properties.

Let us consider the classical second order mass-spring system with unitary parameters given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x(0) = x_0$$

where x_0 satisfies the pdf $\phi_{x_0} = \mathcal{N}(\mu_x, \Sigma_x)$, a normal distribution, with

$$\mu_x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \Sigma_x = \begin{bmatrix} 0.7 & -0.4 \\ -0.4 & 0.3 \end{bmatrix}$$

and $X = \mathbb{R}^2$. Let us furthermore define a relevant transient time $T = 5$ and a desired pdf $\phi_d = \mathcal{N}(\mu_d, \Sigma_d)$, with

$$\mu_d = \begin{bmatrix} 0.97 \\ 3.16 \end{bmatrix}, \quad \Sigma_d = \begin{bmatrix} 0.0004 & 0.0012 \\ 0.0012 & 0.0049 \end{bmatrix}.$$

Then, ϕ_{x_0} and ϕ_d satisfy the linear matching property $\phi_{x_0}(x) = \lambda\phi_d(\beta x + \eta)$, with

$$\eta = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \beta = \begin{bmatrix} 0.01 & -0.02 \\ 0.1 & 0.03 \end{bmatrix}, \quad \lambda = \sqrt{\frac{|\Sigma_d|}{|\Sigma_x|}}.$$

Indeed, let us consider the desired pdf as

$$\phi_d(z) = \frac{1}{2\pi\sqrt{|\Sigma_d|}} \exp\left(-\frac{1}{2}(z - \mu_d)^\top \Sigma_d^{-1}(z - \mu_d)\right),$$

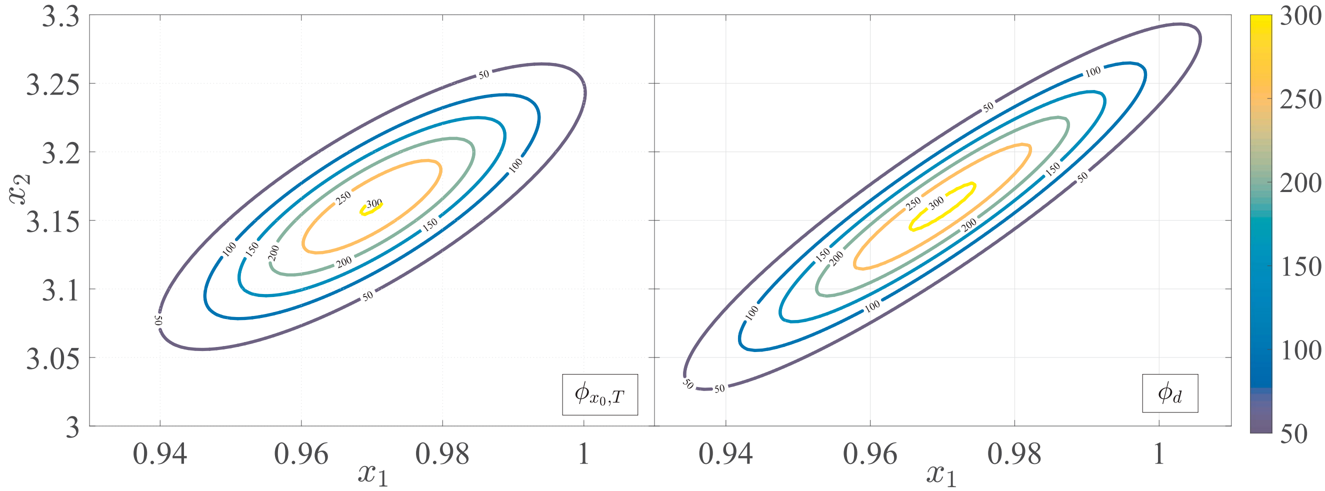

Figure 6. This figure shows level sets of the realised pdf $\phi_{x_0,T}$ and the desired pdf ϕ_d, for the simulation in Section 8. The pdf $\phi_{x_0,T}$ has a Hellinger distance $d_h(\phi_{x_0,T}, \phi_d) = 0.1799$ with ϕ_d.

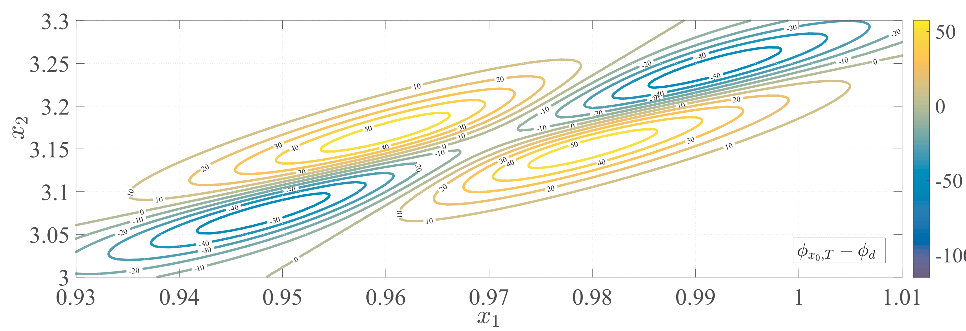

Figure 7. This figure shows the difference between $\phi_{x_0,T}$ and ϕ_d for the simulation in Section 8. This differences corresponds to Hellinger distance $d_h(\phi_{x_0,T}, \phi_d) = 0.1799$. The difference is computed by taking $\phi_{x_0,T}(x) - \phi_d(x)$.

with $z = \eta + \beta x$. Substituting z for x and letting $\mu_d = \eta + \beta \mu_x$ yields

$$\phi_d(\beta x + \eta) = \frac{1}{2\pi\sqrt{|\Sigma_d|}} \exp\left(-\frac{1}{2}(x-\mu_x)^\top \beta^\top \Sigma_d^{-1} \beta (x - \mu_x)\right)$$

and we thus have the relations $\mu_d = \eta + \beta\mu_x$ and $\Sigma_d = \beta\Sigma_x\beta^\top$. Substituting Σ_d and multiplying with λ then yields the matching property.

Let us furthermore assume that we have a reference signal $x_r(t)$ and a feedforward input u^*, satisfying (i). $x_r(t) = x_d(t)$ for all $t \geq \tau$, with $\tau > T$, (ii). $x_r(0) = \mu$, (iii). $x_r(T) = \mu_d$ and (iv). $\dot{x}_r = Ax_r(t) + Bu^*$. For this system, we then satisfy the conditions of Theorem 5.2 and we can therefore numerically evaluate the presented insights. Accordingly, we can express the minimal attainable distance ℓ through (30). As in Theorem 5.2, we take $\tilde{\beta} = e^{(A+BK)T}$ and $\tilde{\eta} = \mu_d - e^{(A+BK)T}\mu$. Subsequently, we can retrace the steps above to find a realisation $\phi_{x_0,T} = \mathcal{N}(\mu_T, \Sigma_T)$, where $\mu_T = \tilde{\eta} + \tilde{\beta}\mu_x$, $\Sigma_T = \tilde{\beta}\Sigma_x\tilde{\beta}^\top$ and $\tilde{\lambda} = \sqrt{\frac{|\Sigma_T|}{|\Sigma_x|}}$.

We are now ready to find a K that minimises $d_h(\phi_{x_0}, \phi_{x_0,T})$ through (30). For this realisation, this is $K = [\,-2.075 \;\; -0.21\,]$ and we find a corresponding minimum Hellinger distance between pdfs $\ell = 0.1799$. The obtained pdf is shown in Figure 6, the difference between the obtained and the desired pdf is shown in Figure 7. We remark that, in accordance with the above, the shown pdfs are not obtained through a time simulation. The pdfs ϕ_d and $\phi_{x_0,T}$ are instead generated directly, with analytically obtained mean and covariance values.

Let us now consider the interesting case where

$$\mu_d = \begin{bmatrix} 2 \\ 2.5 \end{bmatrix}, \quad \Sigma_d = 10^{-3}\begin{bmatrix} 0.6546 & 0.3921 \\ 0.3921 & 0.7042 \end{bmatrix},$$

corresponding to

$$\eta = \begin{bmatrix} 2.0691 \\ 2.3396 \end{bmatrix}, \quad \beta = \begin{bmatrix} 0.0065 & -0.0378 \\ 0.0567 & 0.0518 \end{bmatrix}.$$

We selected this example because, by choosing $K = [\,-0.5 \;\; -0.2\,]$, we obtain the equalities $\tilde{\beta} = \beta$ and $\tilde{\eta} = \eta$ and a minimum Hellinger distance of $\ell = 0$, in accordance with Corollary 5.3.

8. Conclusions

In this paper, we propose control design methods for solving a containment control problem and a shape control problem, applied to systems that have stochastic initial conditions. Both control problems prescribe a required performance during the transient, as well as the standard asymptotic convergence criterion. The containment control problem requires a minimum probability of the state belonging to a set during a specific transient time. We have provided solutions for both linear and nonlinear systems, where the latter relies on recent results in contraction-based control design. The shape control problem requires the attained probability density function at a specific transient time to have a maximum Hellinger distance with respect to a desired probability density function. We have provided solutions that are applicable for linear systems having initial and desired probability density functions that are either linearly or nonlinearly matching.

While our main results in Theorems 3.1–5.5 only provide sufficient conditions for the solvability of both control problems, they provide insightful knowledge on the structure and properties of the control laws. For example, the results related to the **CCP** show that recent results in contraction theory, differential passivity and incremental stability can be useful for future generalisation of these results. The results related to the **SCP** show that we can solve the problem without the need to solve Fokker–Planck equation that is in general not trivial.

An interesting direction for future research is to perform similar analysis when both the initial conditions and the vector field of the system are stochastic. It is expected that results presented here can be extended to such a situation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Abou-Moustafa, K. T., & Ferrie, F. P. (2012). A note on metric properties of some divergence measures: The Gaussian case. In *Proceedings of the 4th ACML*.

Ali, S. M., & Silvey, S. D. (1966). A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society: Series B (Methodological)*, *28*(1), 131–142. https://doi.org/10.1111/j.2517-6161.1966.tb00626.x

Andrieu, V., Jayawardhana, B., & Praly, L. (2016). Transverse exponential stability and applications. *IEEE Transactions on Automatic Control*, *61*(11), 3396–3411. https://doi.org/10.1109/TAC.2016.2528050

Andrieu, V., Jayawardhana, B., & Tarbouriech, S. (2015). Necessary and sufficient condition for local exponential synchronization of nonlinear systems. In *54th IEEE Conference on Decision and Control (CDC)* (pp. 2981–2986).

Aström, K. (1970). *Introduction to stochastic control*. Academic Press.

Bertsekas, D. P. (1976). *Dynamic programming and stochastic control*. Academic Press.

Buehler, E. A., Paulson, J. A., & Mesbah, A. (2016). Lyapunov-based stochastic nonlinear model predictive control: Shaping the state probability distribution functions. In *American Control Conference (ACC)* (pp. 5389–5394).

Crouch, P. E., & Van der Schaft, A. J. (1987). *Variational and Hamiltonian Control Systems*. Springer.

Dresscher, M., & Jayawardhana, B. (2017a). Prescribing transient and asymptotic behaviour of LTI systems with stochastic initial conditions. *IFAC-PapersOnLine*, *50*(1), 1822–1827. https://doi.org/10.1016/j.ifacol.2017.08.182

Dresscher, M., & Jayawardhana, B. (2017b). Prescribing transient and asymptotic behaviour of non-linear systems with stochastic initial conditions. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)* (pp. 1957–1962).

Forni, F., & Sepulchre, R. (2014). A differential Lyapunov framework for contraction analysis. *IEEE Transactions on Automatic Control*, *59*(3), 614–628. https://doi.org/10.1109/TAC.2013.2285771

Gardiner, C. W. (1985). *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Springer.

Ilchmann, A., Ryan, E. P., & Sangwin, C. J. (2002). Tracking with prescribed transient behaviour. *ESAIM: Control, Optimisation and Calculus of Variations*, *7*, 471–493. https://doi.org/10.1051/cocv:2002064

Ilchmann, A., Ryan, E. P., & Townsend, P. (2007). Tracking with prescribed transient behavior for nonlinear systems of known relative degree. *SIAM Journal on Control and Optimization*, *46*(1), 210–230. https://doi.org/10.1137/050641946

Inamdar, S., Bovolo, F., Bruzzone, L., & Chaudhuri, S. (2008). Multidimensional probability density function matching for preprocessing of multitemporal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(4), 1243–1252. https://doi.org/10.1109/TGRS.2007.912445

Itô, K. (2004). *Diffusion processes*. Wiley.

Jouffroy, J., & Fossen, T. I. (2010). A tutorial on incremental stability analysis using contraction theory. *Modeling, Identification and Control: A Norwegian Research Bulletin*, *31*(3).93–106. https://doi.org/10.4173/mic.2010.3.2

Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communications*, *15*(1), 52–60. https://doi.org/10.1109/TCOM.1967.1089532

Kárný, M. (1996). Towards fully probabilistic control design. *Automatica*, *32*(12), 1719–1722. https://doi.org/10.1016/S0005-1098(96)80009-4

Kullback, S. (1997). *Information theory and statistics*. Courier Corporation.

Lohmiller, W., & Slotine, J.-J. E. (1998). On contraction analysis for nonlinear systems. *Automatica*, *34*(6), 683–696. https://doi.org/10.1016/S0005-1098(98)00019-3

Reyes-Báez, R., van der Schaft, A., & Jayawardhana, B. (2017). Tracking control of fully-actuated mechanical port-Hamiltonian systems using sliding manifold and contraction. *IFAC-PapersOnLine*, *50*(1), 8256–8261.

Risken, H. (1989). *Fokker-Planck equation: Methods of solution and applications*. Springer.

Shapiro, L., & Stockman, G. C. (2001). *Computer vision*. Prentice Hall.

Slotine, J. E., & Wang, W. (2005). A study of synchronization and group cooperation using partial contraction theory. In V. Kumar, N. Leonard & A.S. Morse (Eds.), *Cooperative control* (pp. 207–228). Springer. https://doi.org/10.1007/978-3-540-31595-7_12

Sun, J. (2006). *Stochastic dynamics and control*. Elsevier.

Wang, W., & Slotine, J. E. (2005). On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics*, *92*(1), 38–53. https://doi.org/10.1007/s00422-004-0527-x

Appendix. Contraction preliminaries

Consider the system (1). The contraction-based control method is applied through a control law $u(t) = k(\bar{x}(t), t)$, such that the resulting closed-loop system is contracting. Here, $\bar{x}$ can be a dependence on the state, a desired state, or both. Accordingly, we obtain a closed-loop system

$$\dot{x} = f(x, \bar{x}, t), \tag{A1}$$

where f is still a continuously differentiable, e.g. $\mathbb{C}^1$, vector field. For each point $x \in X$ we denote its tangent space as T_xX. Furthermore, let $TX = \bigcup_{x\in X}\{x\} \times T_xX$ be the tangent bundle of X.

The contraction analysis is performed on the prolonged system (Crouch & Van der Schaft, 1987), which is obtained by combining the system (A1) with its variational system. The prolonged system is then given by

$$\begin{aligned} \dot{x} &= f(x, \bar{x}, t), \\ \dot{\delta x} &= \frac{\partial f}{\partial x}(x, \bar{x}, t)\delta x, \end{aligned} \tag{A2}$$

where δx is tangent vector and $(x, \delta x, t) \in T\mathcal{X} \times \mathbb{R}_{\geq 0}$. Accordingly, we can consider a system to be contracting if relevant vector lengths δx (defined by a distance) are uniformly decreasing for all trajectories that start in a certain set. The natural choice for a distance is the Finsler distance, which is related to a Finsler–Lyapunov function. We adopt the corresponding definitions from Forni and Sepulchre (2014).

Definition A.1 (Finsler–Lyapunov function): A $\mathbb{C}^1$ function $V_F : TX \to \mathbb{R}_{\geq 0}$, that maps every $(x, \delta x) \in TX$ to $V_F(x, \delta x) \in \mathbb{R}_{\geq 0}$, is a candidate Finsler–Lyapunov function for (A2), if there exist $c_1, c_2 \in \mathbb{R}_{\geq 0}$, $p \in \mathbb{R}_{\geq 1}$, and a Finsler structure $F : TX \to \mathbb{R}_{\geq 0}$ such that, $\forall (x, \delta x) \in TX$,

$$c_1 F(x, \delta x)^p \leq V_F(x, \delta x, t) \leq c_2 F(x, \delta x)^p, \tag{A3}$$

where the Finsler structure F satisfies the following conditions:

(1) F is a $\mathbb{C}^1$ function for each $(x, \delta x) \in TX$ such that $\delta x \neq 0$;
(2) $F(x, \delta x) > 0$ for each $(x, \delta x) \in TX$ such that $\delta x \neq 0$;
(3) $F(x, \lambda \delta x) = \lambda F(x, \delta x)$ for each $\lambda \geq 0$ and each $(x, \delta x) \in TX$;
(4) $F(x, \delta x_1 + \delta x_2) < F(x, \delta x_1) + F(x, \delta x_2)$ for each $(x, \delta x_1), (x, \delta x_2) \in TX$ such that $\delta x_1 \neq \lambda \delta x_2$ for any given $\lambda \in \mathbb{R}$.

It follows that the Finsler–Lyapunov function is a measure of length of the tangent vector, the corresponding Finsler distance is then obtained through integration.

Definition A.2 (Finsler distance): Consider a candidate Finsler–Lyapunov function V_F on X and the associated Finsler structure F as in Definition A.1. For any two points $(x_1, x_2) \in X \times X$, let $\Gamma(x_1, x_2)$ be the collection of piecewise $\mathbb{C}^1$ curves $\gamma : \mathcal{I} \to X$, $\mathcal{I} := \{s \in \mathbb{R} \mid 0 \leq s \leq 1\}$, $\gamma(0) = x_1$ and $\gamma(1) = x_2$. The distance $d_F : X \times X \to \mathbb{R}_{\geq 0}$ induced by F satisfies

$$d_F(x_1, x_2) := \inf_{\Gamma(x_1, x_2)} \int_{\mathcal{I}} F(\gamma(s), \dot{\gamma}(s)) \mathrm{d}s. \tag{A4}$$

We are now ready to present the existing results on contraction for nonlinear systems.

Lemma A.3 (Contraction): *Consider the system* (A2) *on the smooth manifold X with F a $\mathbb{C}^2$ function, a connected and forward invariant set $C \subseteq X$ and a function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Let V_F be a candidate Finsler–Lyapunov function such that,*

$$\frac{\partial V_F(x, \delta x)}{\partial x} f(x, \bar{x}, t) + \frac{\partial V_F(x, \delta x)}{\partial \delta x} \frac{\partial f(x, \bar{x}, t)}{\partial x} \delta x \leq \alpha(V_F(x, \delta x)) \tag{A5}$$

for each $t \in \mathbb{R}_{\geq 0}$, $x \in C \subseteq X$, and $\delta x \in T_x X$. Then, (A2) *is*

- *incrementally stable (IS) on C if $\alpha(s) = 0$, for each $s \geq 0$;*
- *incrementally asymptotically stable (IAS) on C if α is a class $\mathcal{K}$ function;*
- *incrementally exponentially stable (IES) on C if $\alpha(s) = \lambda s$.*

We refer to Forni and Sepulchre (2014) for the proof of this lemma.

The above can then be interpreted as follows. A system (A2) is *contracting* if for a Finsler distance d_F, there exists a Finsler–Lyapunov function V_F as in Lemma A.3 and $\alpha \in \mathcal{K}$ such that (A5) holds. The system is said to be *exponentially contracting* in the case $\alpha(s) = \lambda s$. Here, C is the contraction region and V_F the contraction measure.

For a system that is contracting, all trajectory starting in the contraction region will converge to a single trajectory. However, which trajectory is not specified by the contraction property. In order to specify this, we can use the result on partial contraction as in Reyes-Báez et al. (2017); Slotine and Wang (2005).

Lemma A.4 (Partial contraction): *Consider the nonlinear system as in* (A1) *with an admissible target trajectory $x_r(t)$, i.e. x_r satisfies $\dot{x}_r(t) = f(x_r(t), x_r(t), t)$ for all $t \geq 0$. Consider a virtual system*

$$\dot{x} = f(x, x_r, t). \tag{A6}$$

If the virtual system (A6) *is contracting w.r.t. x, then x converges to $x_r(t)$.*

The proof of the lemma follows the result from Slotine and Wang (2005) and Forni and Sepulchre (2014).

The system (A1) is then called *partially contracting* if it satisfies the hypothesis in Lemma 8.4 for a given admissible target trajectory x_r.