Propagation of initial condition uncertainty for linear dynamical systems

Kurdyaeva, Tamara; Milias-Argeitis, Andreas

Published in:
2022 European Control Conference, ECC 2022

DOI:
10.23919/ECC55457.2022.9838074

2022 European Control Conference (ECC)
July 12-15, 2022. London, United Kingdom

Propagation of initial condition uncertainty for linear dynamical systems: beyond the Gaussian assumption

Tamara Kurdyaeva and Andreas Milias-Argeitis

Abstract— **Propagation of uncertainty in the initial conditions of a dynamical system is necessary in various control applications. While several generally applicable methods based on Monte Carlo simulation and surrogate modeling exist for this task, they can be computationally intensive or difficult to set up for complex initial distributions. Here, we propose an approach for studying the propagation of initial condition uncertainty, tailored to linear dynamical systems. Our approach uses a class of maximum entropy probability distributions to track the state probability density in time. For deterministic linear systems with initial state distributions belonging to this distribution class, our method results in set of ODEs that allow the exact calculation of the state distribution at any time point in time, generalizing the results known for Gaussian initial distributions. For systems perturbed by noise, we show that the state distribution can be efficiently approximated in a maximum entropy sense via the moment equations. Our results provide a powerful computational alternative to commonly used uncertainty propagation methods, and can be exploited in the construction of filtering and control methods for linear systems with uncertain initial conditions.**

I. Introduction

Studying the impact of initial condition uncertainty on model predictions is necessary for many practical applications of dynamical systems [1], [2]. Uncertainty in the initial conditions of a dynamical system can arise when reduced measurement accuracy prevents exact knowledge of the state [3], or when multiple copies of the same system are considered, each with its own set of initial conditions [4], [5]. This uncertainty can be described by a joint probability distribution over the initial system states, which produces a time-varying probability distribution over future states when propagated through the system dynamics. In the case of deterministic models based on Ordinary Differential Equations (ODEs), the time evolution of the initial probability distribution is governed by the Liouville partial differential equation [6]. When the dynamical system is perturbed by noise in the form of a (multidimensional) Wiener process, the evolution of the initial uncertainty is described by the Fokker-Planck partial differential equation [7].

Numerical solution of the Liouville and Fokker-Planck equations is typically computationally intensive, and only applicable for low-dimensional systems. Therefore, Monte Carlo simulation [2] is commonly used to propagate initial condition uncertainty, sometimes combined with the method of characteristics [8]. However, this method can converge slowly and also be computationally expensive. Surrogate modeling approaches such as Polynomial Chaos Expansions and Gaussian processes [9], [10] offer considerable speed-up, but the construction of accurate surrogate models for complex initial distributions can be challenging in practice. To fully circumvent the numerically intensive calculations of the state distribution, one can rather focus on tracking the time evolution of low-order moments of this distribution. However, the moment system is not closed when the underlying system is nonlinear, necessitating the use of more advanced approaches to accurately track these moments [11].

In this paper, we introduce a more efficient and simple approach for the propagation of initial condition uncertainty, which is tailored for linear dynamical systems affected by this type of uncertainty [4], [5], [12]. It has long been known that when the initial state of a linear system follows a (multivariate) Gaussian distribution, the state distribution remains Gaussian at all times, a fact that has been widely exploited, for example in the construction of the Kalman filter [13] and the propagation of initial condition uncertainty via Gaussian mixtures [14]. Here, we show that deterministic linear systems possess an even stronger property: when the initial condition follows a so-called Q-exponential distribution [15], then the state distribution remains within this family at all future time points, and its parameterization is easily obtainable via a set of linear ODEs. Since the Gaussian distribution is also a member of the Q-exponential family, our result provides a generalization of the Gaussian case. Although linear systems perturbed by Wiener processes do not enjoy the same invariance property, we show that Q-exponential distributions can still be fruitfully exploited to derive a maximum-entropy approximation of the state distribution, making use of the fact that the statistical moments of this distribution can be calculated exactly at any time point. Finally, we demonstrate the applicability of our approach in two examples with complex initial probability distributions, which are propagated through both deterministic and stochastic systems.

II. Preliminaries

A. Linear systems with initial condition uncertainty

We will consider systems described by a system of linear first-order differential equations (ODEs):

$$\frac{d\boldsymbol{x}_t}{dt} = A\boldsymbol{x}_t, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{1}$$

where $\boldsymbol{x}_t \in \mathbb{R}^N$ is the state vector and $A = (a_{ij}) \in \mathbb{R}^{N\times N}$ is a constant matrix. We further assume that $\boldsymbol{x}_0$ is subject

T. Kurdyaeva and A. Milias-Argeitis are with the Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands t.kurdyaeva@rug.nl, a.milias.argeitis@rug.nl

to uncertainty, which is described via a known probability density function $p_0(\boldsymbol{x})$; that is, $\boldsymbol{x}_0 \sim p_0(\boldsymbol{x})$.

We will also consider the stochastic version of (1), given by the linear stochastic differential equation (SDE)

$$d\boldsymbol{x}_t = A\boldsymbol{x}_t dt + \boldsymbol{\sigma} d\boldsymbol{W}_t \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{2}$$

where $\boldsymbol{W}(t)$ is a P-dimensional standard Wiener process, $\boldsymbol{\sigma} \in \mathbb{R}^{N\times P}$, and $\boldsymbol{x}_0 \sim p_0(\boldsymbol{x})$.

The above equations can be augmented with a deterministic input vector without affecting our derivations below. Likewise, A can also vary with time. To simplify the notation in the following, we will only consider the autonomous case.

B. Evolution of the probability density

The uncertainty introduced by $\boldsymbol{x}_0$ generates uncertainty in $\boldsymbol{x}_t, \forall t > 0$. The time evolution of the probability density function $p(t, \boldsymbol{x})$ associated with the state of (1) is governed by a first-order partial differential equation (PDE), known as the Liouville equation:

$$\frac{\partial p}{\partial t} + \sum_{i,j=1}^{N} a_{ij} x_j \frac{\partial p}{\partial x_i} + p \sum_{i=1}^{N} a_{ii} = 0, \tag{3}$$

with initial condition $p(0, \boldsymbol{x}) = p_0(\boldsymbol{x})$.

When $\boldsymbol{x}_t$ is described by (2), the probability density function $p(t, \boldsymbol{x})$ satisfies the Fokker-Planck PDE:

$$\frac{\partial p}{\partial t} + \sum_{i,j=1}^{N} a_{ij} x_j \frac{\partial p}{\partial x_i} + p \sum_{i=1}^{N} a_{ii} - \sum_{i,j=1}^{N} \frac{D_{ij}}{2} \frac{\partial^2 p}{\partial x_i x_j} = 0, \tag{4}$$

where $D = (D_{ij}) \in \mathbb{R}^{N\times N}$ is given by $D = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$.

C. Evolution of moments

Analytical solution of (3) is feasible via the method of characteristics, but becomes impractical for systems with more than three or four dimensions. On the other hand, analytical solution of (4) is only feasible in special cases, e.g. when $p_0(\boldsymbol{x})$ is the density of a multivariate normal distribution. Moreover, numerical solution of these PDEs can quickly become very costly as the state dimension N increases.

Instead of solving the PDEs, one can calculate the expectation of functions $f(\boldsymbol{x})$ with respect to $p(t, \boldsymbol{x})$ and use these quantities to study the effect of uncertainty in $\boldsymbol{x}_0$ on $\boldsymbol{x}_t$. In particular, if we consider $f(\boldsymbol{x}) = \boldsymbol{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_N^{\alpha_N}$ where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N) \in \mathbb{N}_0^N$ is a multi-index, the expected value of $f(\boldsymbol{x})$ with respect to $p(t, \boldsymbol{x})$ is an (uncentered) *moment* of $\boldsymbol{x}_t$:

$$\mu^{[\alpha]} := \mathbb{E}_{p(t,\boldsymbol{x})}\left[\boldsymbol{x}^\alpha\right] = \int_{\mathbb{R}^N} \boldsymbol{x}^\alpha p(t, \boldsymbol{x}) d\boldsymbol{x},$$

In the case of (1), one can easily write the time derivative of a given moment by applying the fundamental formula [6]

$$\frac{d}{dt}\mathbb{E}_{p(t,\boldsymbol{x})}\left[\boldsymbol{x}_t^\alpha\right] = \mathbb{E}_{p(t,\boldsymbol{x})}\left[\frac{d\boldsymbol{x}_t^\alpha}{dt}\right] = \\ = \mathbb{E}_{p(t,\boldsymbol{x})}\left[\sum_{i=1}^{N} \frac{\partial \boldsymbol{x}^\alpha}{\partial x_i} \sum_{j=1}^{N} a_{ij} x_j\right], \tag{5}$$

where $\mathbb{E}_{p(t,\boldsymbol{x})}$ denotes the expectation with respect to $p(t, \boldsymbol{x})$. Because (1) is linear, (5) results in a *closed* system of moment equations, as will be illustrated in Section III.

In the SDE case, the moment equations can be obtained by first applying Itô's lemma to $f(\boldsymbol{x}) = \boldsymbol{x}^\alpha$, which results in

$$d\boldsymbol{x}_t^\alpha = \sum_{i=1}^{N} \frac{\partial \boldsymbol{x}^\alpha}{\partial x_i} \sum_{j=1}^{N} a_{ij} dx_j + \frac{1}{2} \sum_{i,j=1}^{N} D_{ij} \frac{\partial^2 \boldsymbol{x}^\alpha}{\partial x_i \partial x_j} dt. \tag{6}$$

Taking the expectation of both sides of (6) with respect to $p(t, \boldsymbol{x})$, we arrive at the time derivative of the corresponding moment, as we will show in Section IV.

D. Q-exponential distributions

For a random variable $\boldsymbol{x} \in \mathbb{R}^N$ with probability density $p(\boldsymbol{x})$, the differential entropy $H(\boldsymbol{x})$ is defined as

$$H(\boldsymbol{x}) = -\int_{\mathbb{R}^N} p(\boldsymbol{x}) \ln(p(\boldsymbol{x})) d\boldsymbol{x}. \tag{7}$$

Assume that the density $p(\boldsymbol{x})$ is unknown, but we know all moments of $\boldsymbol{x}$ up to a given order d. That is, given a set of real numbers c_α for all $\alpha \geq 0$ such that $|\alpha| \leq d$ (where $|\alpha| := \sum_{i=1}^{N} \alpha_i$), we assume that $\mu^{[\alpha]} = c_\alpha$. Then, among all probability densities supported on $\mathbb{R}^N$ that satisfy these moment constraints, the density that maximizes the differential entropy (7) has the form [16], [17]

$$p_Q(\boldsymbol{x}, \boldsymbol{\lambda}) = \exp\left(-\sum_{|\alpha|\leq d} \lambda_\alpha \boldsymbol{x}^\alpha\right), \tag{8}$$

where the real numbers λ_α are determined from the equations

$$\int_{\mathbb{R}^N} \boldsymbol{x}^\alpha \exp\left(-\sum_{|\alpha|\leq d} \lambda_\alpha \boldsymbol{x}^\alpha\right) d\boldsymbol{x} = c_\alpha, \tag{9}$$

for all α satisfying $|\alpha| \leq d$.

Distributions such as (8) are rarely studied outside the context of maximum entropy approximations, with a few notable exceptions [15], [18]. Following the nomenclature of [15], we will say that $\boldsymbol{x}$ follows a *Q-exponential distribution* if its density has the form

$$p_Q(\boldsymbol{x}, \boldsymbol{\lambda}) = \exp\left(-Q(\boldsymbol{x})\right), \tag{10}$$

where

$$Q(\boldsymbol{x}) = \sum_{|\alpha|\leq d} \lambda_\alpha x^\alpha. \tag{11}$$

Note that the constant of normalization in (10) is absorbed in the constant term of $Q(\boldsymbol{x})$. A necessary (but not sufficient) condition for (10) to be integrable over $\mathbb{R}^N$ is that the degrees of (11) with respect to each x_i are even.

Q-exponential distributions have a number of useful mathematical properties [15]. For what concerns our work, Q-exponential distributions can assume a wide range of shapes (e.g. multiple modes [18]) while requiring fewer parameters for their description than mixtures of normal distributions which are commonly used for this task. Moreover, they

have finite moments of all orders [15], so they can be used to approximate other commonly used distributions via the method of moments [15]. Finally, they can be viewed as generalizations of the multivariate normal distribution, which is also a member of the family for $d = 2$.

III. Propagation of initial condition uncertainty for linear ODEs

A. Propagation of moments

As discussed in Section II-C, instead of directly solving the Liouville PDE, we can track the time evolution of a set of moments of $p(t, \boldsymbol{x})$ with the help of (5), starting from the known moments of $p_0(\boldsymbol{x})$. To write the moment equations of (1) in compact form, we will define e_i as the multi-index with 1 in the i-th position and zeros in all other positions. Expanding the sum in (5) and applying the expectation operator, we obtain the following:

Lemma 3.1: For system (1) with uncertain initial conditions, any moment $\mu^{[\alpha]}$ with $\alpha = (\alpha_1, \ldots, \alpha_N)$, evolves according to the following equation:

$$\frac{d\mu^{[\alpha]}}{dt} = \sum_{i=1}^{N} \alpha_i \sum_{j=1}^{N} a_{ij} \mu^{[\alpha - e_i + e_j]} [\alpha - e_i + e_j \geq 0], \quad (12)$$

where the bracket is equal to 1 when the the inequality holds, and equal to 0 when it does not. Additionally, the equation for a moment of order $|\alpha| = d$ contains only moments of order d.

The assertion can be directly inferred from (5). It follows that if we are interested in all moments up to order d (i.e. $|\alpha| \leq d$, with at least one $\alpha_i > 0$ to exclude the constant function), (12) will provide a set of $\binom{N+d}{d} - 1$ ODEs which will form a closed system, i.e. the right-hand side will contain moments up to order d.

B. Propagation of Q-exponential initial distributions

Besides calculating moments of (12), it is actually possible to calculate $p(t, \boldsymbol{x})$ if we assume that $p_0(\boldsymbol{x})$ is also a member of the Q-exponential family. In that case, as we will show below, $p(t, \boldsymbol{x})$ belongs to the Q-exponential family for all $t > 0$, and the solution of (3) reduces to the solution of a system of $\binom{N+d}{d}$ linear ODEs for tracking the time evolution of the parameters $\boldsymbol{\lambda}$ in (11). More concretely, we have:

Theorem 3.2: Consider the Liouville PDE (3) with initial condition $p_0(\boldsymbol{x}) = p_Q(\boldsymbol{x}, \boldsymbol{\lambda}_0) = \exp(-\sum_{|\alpha| \leq d} \lambda_{\alpha,0} x^\alpha)$. Then,

$$p(t, \boldsymbol{x}) = p_Q(\boldsymbol{x}, \boldsymbol{\lambda}(t)) = \exp(-\sum_{|\alpha| \leq d} \lambda_\alpha(t) x^\alpha) \quad (13)$$

for every α such that $|\alpha| \leq d$. Moreover, $\lambda_\alpha(t)$ for all $\alpha \neq 0$ satisfy the following linear ODE system:

$$\frac{d\lambda_\alpha}{dt} = -\sum_{i,j=1}^{N} a_{ij} (\alpha + e_i - e_j)_i \lambda_{\alpha + e_i - e_j} [\alpha + e_i - e_j \geq 0]. \quad (14)$$

with initial condition $\lambda_\alpha(0) = \lambda_{\alpha,0}$, where the bracket notation is interpreted as in Lemma 3.1. For $\alpha = 0$ (the constant term in $Q(\boldsymbol{x})$),

$$\frac{d\lambda_0}{dt} = \sum_{i=1}^{N} a_{ii}. \quad (15)$$

Proof: Substituting $p(t, \boldsymbol{x}) = \exp(-\sum_{|\alpha| \leq d} \lambda_\alpha(t) x^\alpha)$ in (3) gives

$$\Bigg[-\sum_{|\alpha| \leq d} \frac{d\lambda_\alpha}{dt} \boldsymbol{x}^\alpha + \sum_{i=1}^{N} \left(a_{ii} - \sum_{|\alpha| \leq d} \alpha_i \lambda_\alpha(t) \boldsymbol{x}^{\alpha - e_i} \left(\sum_{j=1}^{N} a_{ij} x_j \right) \right) \Bigg] p(t, \boldsymbol{x}) = 0$$

This implies that the term in square brackets needs to be identically zero. Rewriting this term and equating to zero, we obtain

$$-\sum_{|\alpha| \leq d} \frac{d\lambda_\alpha}{dt} \boldsymbol{x}^\alpha + \sum_{i=1}^{N} \left(a_{ii} - \sum_{j=1}^{N} a_{ij} \sum_{|\alpha| \leq d} \alpha_i \lambda_\alpha(t) \boldsymbol{x}^{\alpha - e_i + e_j} \right) = 0$$

In this equation, a polynomial of $\boldsymbol{x}$ is identically equal to zero. This will happen only if all polynomial coefficients vanish identically. Equating the coefficient of $\boldsymbol{x}^\alpha$ to zero for every $\alpha \neq 0$ results in (14). Finally, equating the coefficient of the constant term ($\alpha = 0$) to zero, we obtain (15). ∎

When $p_0(\boldsymbol{x})$ is unknown and only a set of moments up to order d are available, then the Q-exponential distribution (13) for $t \geq 0$ is equal to the maximum entropy distribution of the system given the available moment information.

Theorem 3.2 provides a direct generalization of the fact that when $x(0)$ in (1) follows a multivariate normal distribution, then the distribution of $x(t)$ is multivariate normal for all $t \geq 0$. In the case of Q-exponential distributions, however, the starting distribution can have a more complex shape than the normal.

In summary, we have shown that initial condition uncertainty of a linear ODE system can be propagated in two computationally efficient ways: via the moment equations, and via the Liouville PDE, provided $p_0(\boldsymbol{x})$ belongs to the Q-exponential family.

IV. Propagation of initial condition uncertainty for linear SDEs

A. Propagation of moments

Similarly to the case of linear ODEs (and for the same reasons), we can track the time evolution of moments for a system of linear SDEs with the help of (6). Making use of the multi-index notation of Lemma 3.1, we obtain

Lemma 4.1: For the stochastic system (2) with uncertain initial conditions, any moment $\mu^{[\alpha]}$ with $\alpha = (\alpha_1, \ldots, \alpha_N)$,

evolves according to the following equation:

$$\frac{d\mu^{[\alpha]}}{dt} = \sum_{i=1}^{N} \alpha_i \sum_{j=1}^{N} a_{ij} \mu^{[\alpha - e_i + e_j]} [\alpha - e_i + e_j \geq 0] + \\ + \frac{D_{ij}}{2} \sum_{i \neq j}^{N} \alpha_i \alpha_j \mu^{\alpha - e_i - e_j} [\alpha - e_i - e_j \geq 0] + \\ + \frac{D_{ii}}{2} \sum_{i=1}^{N} \alpha_i (\alpha_i - 1) x^{\alpha - 2e_i} [\alpha - 2e_i \geq 0] \quad (16)$$

where the brackets are equal to 1 when the the corresponding inequalities holds, and equal to 0 when they do not. Note that this moment system is again closed.

B. Maximum entropy approximation of the state distribution

In the case of the Fokker-Planck PDE (4), $p(t, \boldsymbol{x})$ will not be a member of a Q-exponential family in general, even when $p_0(\boldsymbol{x})$ is Q-exponential. As one can verify by plugging (13) in (4), the second derivative terms will produce monomials of degree higher than d. The requirement that (4) holds, will in turn force all coefficients of monomials of order higher than 2 to be zero. In fact, only when $d = 2$, i.e. when $p_0(\boldsymbol{x})$ is a multivariate normal, does the solution of (4) remain within the same Q-exponential family (multivariate normal) for all $t \geq 0$.

Although (13) is not an exact solution of the Fokker-Planck PDE, it can still be viewed as a maximum entropy approximation of the actual solution at each time point, given a set of moments up to order d. To calculate this maximum entropy approximation, one could first solve the system of moments up to order d, and then solve the maximum entropy optimization problem at each time point of interest.

More concretely, we let $p_{FP}(t, \boldsymbol{x})$ denote the solution of (4) given an initial condition $p_0(\boldsymbol{x})$, and consider an ordering of the moments of $p_{FP}(t, \boldsymbol{x})$ up to order d: $\{\mu^{[\alpha(1)]}(t), \mu^{[\alpha(2)]}(t), \ldots, \mu^{[\alpha(M)]}(t)\}$, where $M = \binom{N+d}{d}$ is the total number of such moments (including the normalizing constant). Since these moments of $p_{FP}(t, \boldsymbol{x})$ can be computed from the closed system (16), they are known at each $t \geq 0$ (in practice, we can calculate these moments on arbitrarily chosen time grid $\{0, t_1, t_2, \ldots, t_K\}$).

Instead of solving a constrained maximization problem to calculate the maximum-entropy approximation to $p_{FP}(t, \boldsymbol{x})$ [16], one can obtain this approximation by solving an unconstrained minimization problem that corresponds to the dual of the original problem [16], [19], [20], [17]. At each time point $t_k, k = 1, \ldots, K$, this problem has the form

$$\min_{\boldsymbol{\lambda}(t_k)} \quad \ln Z(\boldsymbol{\lambda}(t_k)) + \sum_{i=1}^{M} \lambda_{\alpha(i)}(t_k) \mu^{[\alpha(i)]}(t_k), \quad (17)$$

where

$$Z(\boldsymbol{\lambda}(t_k)) = \int_{\mathbb{R}^N} \exp\left(- \sum_{1 \leq |\alpha| \leq d} \lambda_\alpha(t_k) \boldsymbol{x}^\alpha \right) d\boldsymbol{x},$$

and can be solved with the help of Newton's method [16], [20], [19]. Note that the gradient of the objective function in (17) contains moments of p_Q up to order d, while its Hessian contains moments up to order $2d$ [16], [19].

Although the objective function in (17) is strictly convex [16], [19], the numerical solution of this problem from an arbitrary starting point is practically very hard due to bad numerical conditioning [17], which necessitates the use of more advanced optimization methods [19]. In the case considered here, however, the solution of (17) can still be obtained with the standard Newton method. This is possible if the time grid used for the evaluation of the moments is dense enough, permitting the use of the solution at point t_{k-1}, $\boldsymbol{\lambda}(t_{k-1})$, to act as a good starting point for the optimization at t_k. In this way, starting from a known initial condition $\boldsymbol{\lambda}(0)$ (corresponding to the known initial moment vector $\boldsymbol{\mu}(0)$), one can sequentially evaluate the maximum entropy approximation of the Fokker-Planck solution at each time point. To summarize, for systems described by linear SDEs the state distribution does not belong to a Q-exponential family, unless the initial distribution is Gaussian. However, we can compute a maximum-entropy approximation $p_Q(\boldsymbol{x}, \boldsymbol{\lambda}(t))$ to $p_{FP}(t, \boldsymbol{x})$ using the numerical solutions of the moment equations as constraints.

V. Examples

We will present two examples to demonstrate the application of the proposed approach, and consider two-dimensional systems to facilitate visualization of the results. In both cases, initial distributions have shapes that are difficult to approximate with mixtures of multivariate Gaussian distributions.

A. A 2-dimensional linear system with a stable node

We first consider the following linear system:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -0.1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (18)$$

We assume that the initial value of the state vector $\boldsymbol{x}_0 = \begin{bmatrix} x_1(0) & x_2(0) \end{bmatrix}^T$ is uncertain and follows a "ring-shaped" distribution with the following probability density function:

$$p_0(\boldsymbol{x}) = \frac{1}{C_r} \exp\left(-\frac{1}{4}(x_1^2 + x_2^2 - 3)^2 (x_1^2 + x_2^2) \right) \quad (19)$$

where C_r is a normalizing constant which is necessary for $p_0(\boldsymbol{x})$ to integrate to 1 over the entire space. A top view of this bivariate distribution is displayed on Fig. 1.

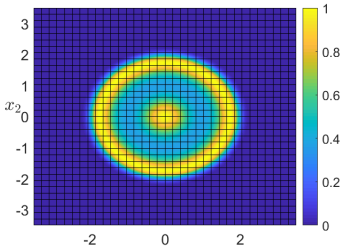

Fig. 1: *Top view of the bivariate "ring-shaped" probability density $p_0(\boldsymbol{x})$. Color intensity indicates the height of $p_0(\boldsymbol{x})$.*

This initial distribution is a member of the Q-exponential family, as can be easily seen from the form of probability

density function (19). Therefore, we can track the time evolution of the probability density of the state of (18) with initial condition $p_0(\boldsymbol{x})$ given by (19) making use of (14) and (15) to derive ODEs for the parameters $\boldsymbol{\lambda}(t)$ in (13), which in turn determine the solution of the Liouville PDE (3) according to Theorem 3.2. In this example, the polynomial $Q(\boldsymbol{x})$ in (19) has order 6 which leads to a system of 28 linear ODEs for the parameters $\lambda_\alpha, |\alpha| \leq 6$. Solving this system allows us to evaluate the Liouville solution $p_Q(\boldsymbol{x}, \boldsymbol{\lambda}(t))$ at any time point $t > 0$ using the corresponding $\boldsymbol{\lambda}(t)$ values. Examples of the resulting distributions are shown on Fig. 2.

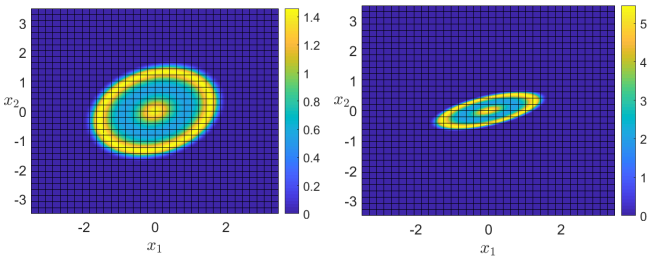

Fig. 2: *Snapshots of the state probability density of* (18)*, starting from the initial distribution* (19)*. Color intensity indicates the values of* $p_Q(t, \boldsymbol{x})$ *at the corresponding time points. Left:* $t = 0.13$*. Right:* $t = 0.57$*.*

We next assume that (18) is perturbed by a two-dimensional standard Wiener process and is modeled by a system of linear SDEs of the form (2), where $\boldsymbol{\sigma} = 2I_2$ with I_2 being the 2×2 unit matrix. According to the discussion of Section IV, we can approximate this target distribution at times $t > 0$ with a member of the Q-exponential family (with $Q(\boldsymbol{x})$ a 6-th degree polynomial) that maximizes the differential entropy subject to moment constraints obtained from state moments up to order 6, calculated from (16). We can then solve (17) over a time grid, as described in Section IV-B. Fig. 3 shows two snapshots of the resulting maximum entropy distributions.

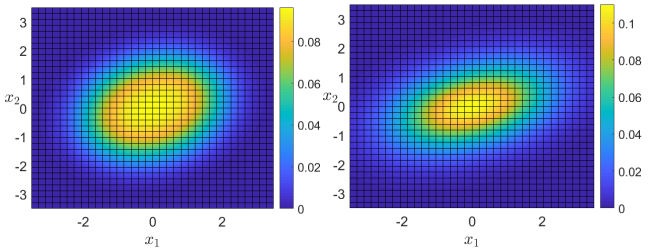

Fig. 3: *Two snapshots of the probability density of the state of the noise-perturbed version of* (18)*, using* $\sigma = 2I_2$*. Color intensity indicates the values of* $p_Q(t, \boldsymbol{x})$*. Left:*$t = 0.13$*. Right:* $t = 0.57$*.*

B. A 2-dimensional linear system with a stable focus

The next system we consider is governed by the following ODEs:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 13 \\ -5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{20}$$

We further assume that the initial value of the state vector $\boldsymbol{x}_0 = \begin{bmatrix} x_1(0) & x_2(0) \end{bmatrix}^T$ is uncertain and follows another non-Gaussian probability distribution $p_0(\boldsymbol{x})$, shown in Fig. 4. Its probability density function is obtained by changing the argument of the multivariate Gaussian density $\mathcal{N}(\boldsymbol{x}; 0, \Sigma)$ into $\phi_b(\boldsymbol{x}) = (x_1, x_2 + bx_1^2 - 100b)$; that is,

$$p_0(\boldsymbol{x}) = \frac{1}{C_b} \mathcal{N}(\phi_b(\boldsymbol{x}); 0, \Sigma), \tag{21}$$

where b controls the amount of distortion applied to the original Gaussian density and C_b is the normalization constant. For the probability function $p_0(\boldsymbol{x})$ considered here, we used $\Sigma = \mathrm{diag}(50, 1)$ and $b = 0.05$.

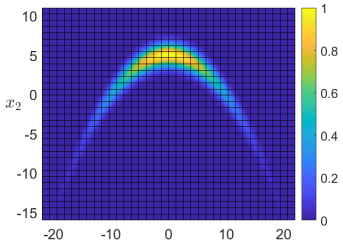

Fig. 4: *Top view of the bivariate "banana-shaped" probability density* $p_0(\boldsymbol{x})$*. Color intensity indicates the height of* $p_0(\boldsymbol{x})$*.*

It can be easily observed that (21) also belongs to the Q-exponential family with $Q(\boldsymbol{x})$ being a fourth degree polynomial. This means that we can apply (14) and (15) to study how the initial probability density $p_0(\boldsymbol{x})$ is transformed by (20). Since $Q(\boldsymbol{x})$ is a fourth-degree polynomial, the system of ODEs for $\lambda_\alpha, |\alpha| \leq 4$ contains 15 equations. From the snapshots of the Liouville PDE solutions shown on Fig. 5, we can observe that the oscillatory behaviour of the system causes the state distribution to rotate. At the same time, the initial uncertainty shrinks down to a very narrow distribution preserving the original shape, as expected.

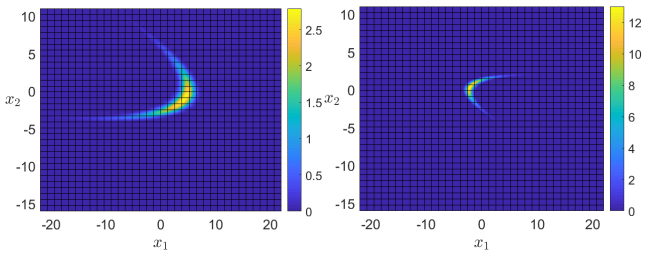

Fig. 5: *Two snapshots of the state probability density of* (20)*, starting from the initial distribution* (21)*. Color intensity indicates the values of* $p_Q(t, \boldsymbol{x})$ *at the corresponding time points. Left:* $t = 0.26$*. Right:* $t = 0.64$

Similarly to the previous example, we next examine how the time evolution of the initial uncertainty changes if (20) has an additive Wiener process component with the same $\boldsymbol{\sigma}$ as above. As discussed previously, we first track all moments of the system up to order 4 using (16), and then compute the maximum entropy approximation at every time step given the moment information. Fig. 6 summarises the results of this approach. In contrast to the deterministic system on Fig. 5, Fig. 6 shows wider distributions as the initial distribution is "smoothed out" by the input noise. The maximum entropy approximation is very closely approximating the actual state

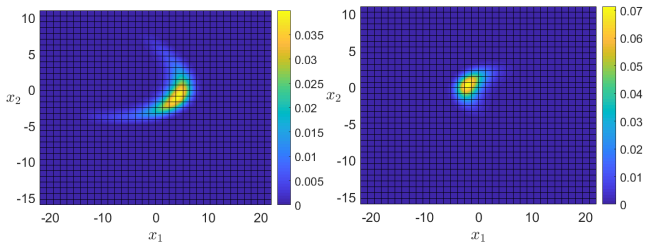

Fig. 6: *Two snapshots of the state probability density of the noise-perturbed version of* (20)*, using* $\sigma = 2I_2$*. Color intensity indicates the values of* $p_Q(t, \boldsymbol{x})$*. Left:* $t = 0.26$*. Right:* $t = 0.64$

distribution, as shown on Fig. 7, in which the Monte Carlo-based approximation of the Fokker-Planck solution is plotted.

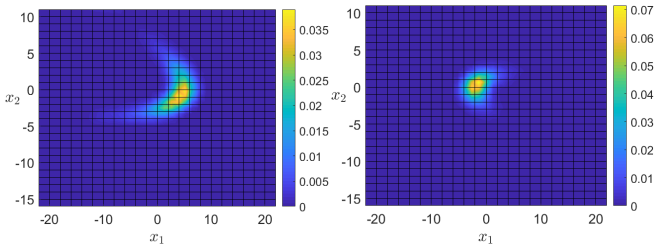

Fig. 7: *Empirical distributions approximating* $p_{FP}(t, \boldsymbol{x})$*, generated via Monte Carlo simulation [21] of the analytical SDE solution [22] using 50000 samples of the initial distribution. The initial samples were obtained with the help of a Metropolis–Hastings sampler. The panels show the Monte Carlo approximation of* $p_{FP}(t, \boldsymbol{x})$ *at* $t = 0.26$ *(left) and* $t = 0.64$ *(right)*

VI. Discussion

We presented an approach for the propagation of initial condition uncertainty based on the Louville equation, applicable to linear dynamical systems when the distribution of initial conditions belongs to the Q-exponential family. For ODE-based linear models, we have shown that the exact solution of the Liouville equation remains within the Q-exponential family, and can be obtained by a set of linear ODEs for the distribution parameters. This fact makes our approach applicable to large linear systems with complex initial distributions. For linear SDEs, we have shown that the solution of the Fokker-Planck equation with a Q-exponential initial condition can be approximated at each $t > 0$ by a Q-exponential distribution obtained from a maximum entropy optimization problem constrained by the moments of the system. Instead of finding the maximum entropy approximation at discrete time points, one can alternatively use variational moment closure [23] to derive and solve a common set of differential equations for the coefficients of $p_Q(t, \boldsymbol{x})$ and its moments. However, numerical integration of these equations requires the evaluation of the moments of $p_Q(t, \boldsymbol{x})$ up to order $2d$ at each time step, which increases the computational burden of this method can be a limiting factor for larger systems, as the number of moments of a given order increases quickly with the system dimension.

Besides tracking the probability density of the system state over time, as done in our examples, we hope that our results will be further developed and used for the construction of computationally efficient filtering and control algorithms for linear systems with uncertain initial conditions.

References

[1] W. L. Oberkampf, S. M. DeLand, B. M. Rutherford, K. V. Diegert, and K. F. Alvin, "Error and uncertainty in modeling and simulation," *Reliability Engineering & System Safety*, vol. 75, no. 3, pp. 333–357, 2002.

[2] C. Yang and M. Kumar, "On the effectiveness of Monte Carlo for initial uncertainty forecasting in nonlinear dynamical systems," *Automatica*, vol. 87, pp. 301–309, 2018.

[3] E. S. Epstein, "The role of initial uncertainties in predicion," *Journal of Applied Meteorology and Climatology*.

[4] R. W. Brockett, "Optimal control of the Liouville equation," *AMS/IP Studies in Advanced Mathematics*, vol. 39, p. 23, 2007.

[5] J. Matis and T. Wehrly, "Stochastic models of compartmental systems," *Biometrics*, pp. 199–220, 1979.

[6] M. Ehrendorfer, *The Liouville equation and atmospheric predictability*. Cambridge University Press, 2006, p. 59–98.

[7] H. Risken and T. Frank, *The Fokker-Planck Equation: Methods of Solution and Applications*, ser. Springer Series in Synergetics. Springer Berlin Heidelberg, 1996.

[8] A. Halder and R. Bhattacharya, "Dispersion analysis in hypersonic flight during planetary entry using stochastic liouville equation," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 2, pp. 459–474, 2011.

[9] R. C. Smith, *Uncertainty quantification: theory, implementation, and applications*. SIAM, 2013, vol. 12.

[10] C. Schillings, M. Sunnåker, J. Stelling, and C. Schwab, "Efficient characterization of parametric uncertainty of complex (bio)chemical networks," *PLOS Computational Biology*, vol. 11, no. 8, pp. 1–16, 08 2015.

[11] T. Kurdyaeva and A. Milias-Argeitis, "Uncertainty propagation for deterministic models of biochemical networks using moment equations and the extended kalman filter," *Journal of the Royal Society Interface*, vol. 18, no. 181, p. 20210331, 2021.

[12] R. Brockett, "Notes on the control of the Liouville equation," in *Control of partial differential equations*, 2012, pp. 101–129.

[13] D. Simon, *Optimal state estimation: Kalman, H-infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[14] S. Roweis and Z. Ghahramani, "A unifying review of linear gaussian models," *Neural Computation*, vol. 11, no. 2, pp. 305–345, 1999.

[15] C. Urzúa, "A class of maximum-entropy multivariate distributions," *Communications in Statistics-Theory and Methods*, vol. 17, no. 12, pp. 4039–4057, 1988.

[16] L. R. Mead and N. Papanicolaou, "Maximum entropy in the problem of moments," *Journal of Mathematical Physics*, vol. 25, no. 8, pp. 2404–2417, 1984.

[17] R. V. Abramov, "The multidimensional maximum entropy moment problem: A review of numerical methods," *Communications in Mathematical Sciences*.

[18] L. Cobb, P. Koppstein, and N. H. Chen, "Estimation and moment recursion relations for multimodal distributions of the exponential family," *Journal of the American Statistical Association*, vol. 78, no. 381, pp. 124–130, 1983.

[19] Z. Wu, G. N. Phillips Jr, R. Tapia, and Y. Zhang, "A fast Newton algorithm for entropy maximization in phase determination," *SIAM review*, vol. 43, no. 4, pp. 623–642, 2001.

[20] X. Wu, "Calculation of maximum entropy densities with application to income distribution," *Journal of Econometrics*, vol. 115, no. 2, pp. 347–354, 2003.

[21] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM review*, vol. 43, no. 3, pp. 525–546, 2001.

[22] X. Mao, *Stochastic differential equations and applications*. Elsevier, 2007.

[23] L. Bronstein and H. Koeppl, "A variational approach to moment-closure approximations for the kinetics of biomolecular reaction networks," *The Journal of Chemical Physics*, vol. 148, no. 1, p. 014105, 2018.