Sample-based Population Observers ⋆

Shen Zeng [a]

[a] *Department of Electrical and Systems Engineering, Washington University in St. Louis*

Abstract

In this paper, a first sample-based formulation of the recently considered population observers, or ensemble observers, which estimate the state distribution of dynamic populations from measurements of the output distribution is established. The results presented in this paper yield readily applicable computational procedures that provide novel avenues to circumvent issues regarding the curse of dimensionality, which all previously developed techniques employing a kernel-based approach are inherently suffering from. The novel insights that eventually pave the way for all different kinds of sample-based considerations are in fact deeply rooted in the basic probabilistic framework underlying the problem, bridging optimal mass transport problems defined on the level of distributions with actual randomized strategies operating on the level of individual points. The conceptual insights established in this paper not only yield insight into the underlying mechanisms of sample-based ensemble observers but significantly advance our understanding of estimation and tracking problems for the class of ensembles of dynamical systems in general.

Key words: Observers, Large-scale systems, Nonlinear dynamical systems, Computed tomography

1 Introduction

The observability problem in systems theory systematically addresses a task fundamental to numerous scientific fields, particularly those close to physics, namely the extraction of information about the state of a dynamical process from knowledge of the underlying dynamics, and time series data of some less informative output measurement. The concept of observability together with the concept of controllability of a linear state-space model layed the basic foundation of a general theory of (control) systems (see [1, 2]), which has fundamentally reshaped the way we think about systems.

The same line of thoughts centered around the questions of controllability and observability are recently being investigated in relation to a new class of systems, consisting of populations of dynamical systems of the same structure with a given distribution in their states [3–9]. While a classical system can be thought of as a single point particle evolving in state-space (following the combined effect of a drift and a control vector field), for a population comprised of a large number of dynamical system, the point describing the state of the system would be replaced by a (probability) distribution of points, as suggested in Figure 1.

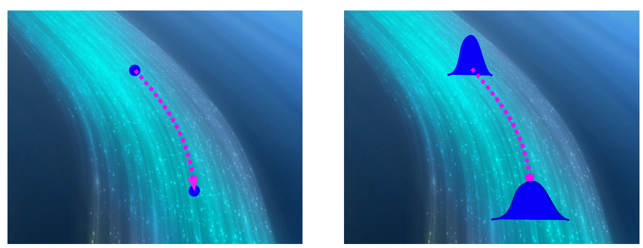

Fig. 1. The evolution of the state of a classical system is typically thought of as a point evolving in state space (left). In the same spirit, the dynamics of a population of systems is described by distributions of points (right).

Of course, the idea of considering probability distributions as a description of the state of a system is not new – in fact it traces back more than 100 years to the early beginnings of statistical mechanics, where the occurring probability distribution was already used both as a model for the state of one uncertain system or of an actual population of many systems, with a distribution in initial states. However, it has only recently become clear that once we look closer at the interface of really interacting with actual populations of systems, very distinct restrictions start to surface. This is where the probabilistic model splits into two branches, each with completely different interpretations with regard to what is being measured, and how we are able to exert control over the system.

⋆ This work was supported in part by the National Science Foundation under the award ECCS-1810202.

Corresponding author S. Zeng. Tel. +1-314-935-9357

Email address: `s.zeng@wustl.edu` (Shen Zeng).

A prime example that illustrates the fundamentally different interpration of the probabilistic setup for the situation of populations of dynamical systems is given by heterogeneous cell populations, such as cancer cell populations. For example, an important task for such heterogeneous cell populations is to estimate the specific distribution in states [1] or parameters, as such distribution can often be the key driver for heterogeneous responses to an external biochemical stimulus, like it is prominently observed with cancer, where we often see the survival of subpopulations during drug treatment. The given data for solving the estimation task are measurements of only a subset of molecule concentrations, which furthermore are increasingly being recorded via high-throughput devices called flow cytometers. By rapidly passing a stream of fluorescently labeled cells through a laser and fluorescence detectors, flow cytometers can easily gather concentration measurements of a vast number of cells. However, the ability to gather vast amounts of data comes at a cost. Namely, it is only possible to measure at the population level, which here, specifically, means that nothing can be said about an individual cell; it is only that a lot of measurements are being recorded and then stored in the form of histograms or other statistics. This circumstance may be described as a population-level observation, and is illustrated Figure 2.

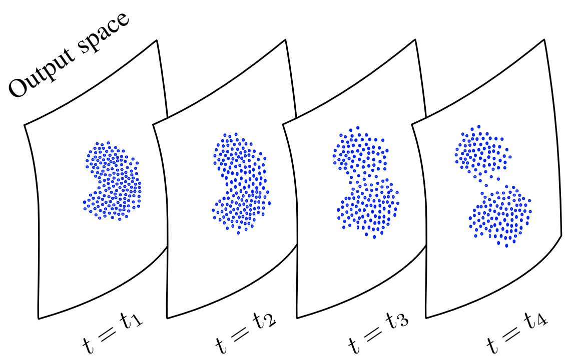

Fig. 2. An illustration of population snapshots. In each time step $t_1, \dots, t_4$ we have a snapshot of certain output values of a population. The crucial point is that in a snapshot, information relating an output value to the individual producing that output value is completely missing. Taken from [8].

While Figure 2 may give the impression that one is measuring many output trajectories of individual cells, but without recording the actual associations between measurements in different time points, the situation is in fact even more cumbersome for the example of cell populations. This is because we only get to measure each cell once, due to the simple reason that after it is measured, it is either destroyed or gone. Therefore, the measurements at different snapshots may stem from completely different individuals in the population; they do, however, all stem from the same population.

[1] The state of a single cell is typically described by the set of concentrations of different molecules or proteins, which are governed by regulatory networks that in turn can be described by ordinary differential equations.

These considerations led us to view these population snapshots as samples from an output distribution, and to further view the output distribution as the "total" output of the population. This idea was then formalized in terms of a novel systems theoretic setup in which a classical system

$$\begin{aligned} \dot{x} &= f(x) \\ y &= h(x), \end{aligned}$$

with $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^m$, is generalized by setting the initial state to be a random vector $x(0) \sim \mathbb{P}_0$, with a non-parametric probability distribution $\mathbb{P}_0$. This clearly leads to a probabilistic nature of the output as well, which we describe in terms of $y(t) \sim \mathbb{P}_{y(t)}$. The practical ensemble observability problem consists of reconstructing the initial state distribution $\mathbb{P}_0$ when given the evolution of the distribution of outputs $\mathbb{P}_{y(t)}$, which, again, is fundamentally different from classical filtering problems in which the measured data are *single realizations* of the output distribution associated to a *single uncertain* plant.

In [8], we first studied the ensemble observability problem in the linear case, where $f(x) = Ax$ and $h(x) = Cx$, both from a theoretical and practical perspective. The investigations of the underlying basic theoretical problem in particular also revealed a deep connection between such ensemble observability problems and mathematical tomography problems, providing crucial insights into the inner systems theoretic mechanisms and, from a practical perspective, also immediately rendered the problem amenable to computational solutions. The computational solutions, however, having been very much anchored in the tomography-based considerations, were inevitably affected by the curse of dimensionality. While problems in tomography most prominently take place in lower dimensions, specifically dimensions two or three, in the ensemble observability setup, such a restriction is naturally undesirable, as the dimension of the state space is in general typically higher. In [10], we already pointed out that in our quest to get better insights about the initial state distribution, we eventually want to circumvent the route over distributions, in which the output snapshots are first mapped into discretized distributions (histograms), from which a discretized initial state distribution is then to be reconstructed. Instead, a sample-based approach to the reconstruction problem was envisioned, which from a pragmatic standpoint seems very natural as well, as the measurement data is naturally given in terms of samples of the output in the first place, and not in terms of the distribution of the output, which is just a mathematical idealization introduced for the sake of studying the theoretical problem. In order to establish a sample-based framework, in this paper we derive a systematic practical procedure that takes the samples of the output distribution at different time points and eventually returns a set of points that mimic a set of real samples from the sought initial state distribution. This idea is realized by leveraging a novel connection to optimal mass transport problems [11, 12], which is in fact very fundamental, and leads to novel interesting theoretical insights, as well as challenging open problems.

The structure of the paper is as follows. In Section 2, we provide a very brief review of the ensemble observability problem, with its many different viewpoints and connections to other areas of mathematics. In particular, we will discuss an example of a nonlinear observability problem, which already provides some important hints towards establishing a sample-based approach. This sample-based approach is then fully established in Section 3, yielding both sample-based ensemble estimators and observers. All key steps in the introduction of the sample-based scheme are complemented by detailed illustrations and examples.

2 The ensemble observability problem

In this section we provide a rapid review of different aspects of the ensemble observability problem that are most relevant to the presentation of the novel insights and results put forth of this paper. In particular, we will put significant emphasis on the discussion of the relation between the ensemble observability problem and mathematical tomography problems, established in [8], by which the ensemble observability problem also first became amenable to comprehensible computational solutions.

We recall that in the general ensemble observability problem we ask under which conditions we can reconstruct the initial state distribution $\mathbb{P}_0$ when given the evolution of the distribution of outputs $\mathbb{P}_{y(t)}$, under a finite-dimensional (nonlinear) dynamical system. Furthermore, we are interested in practical reconstruction techniques for this problem. In [8], we first studied this problem in the linear case, both from a theoretical and practical perspective. To first build some intuition around the whole concept of ensemble observability, we consider an example with a two-dimensional harmonic oscillator

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x, \qquad y = \begin{pmatrix} 1 & 0 \end{pmatrix} x.$$

with a bimodal initial distribution as depicted in Figure 3. The measured output distribution corresponding to the output $y = x_1$ of the underlying linear system results from a marginalization of the state distribution over the second coordinate, i.e. from integration along the x_2-direction. Thus, when the system evolves, the state distribution is subject to both a transportation with the flow, and a marginalization over the second coordinate, resulting in an evolution of the output distribution, as suggested in Figure 3.

The question in the ensemble observability problem for the specific example is thus whether or not one can reconstruct the (initial) state distribution from only observing the evolution of the output distribution, shown in the lower left of Figure 3. Even though one might consider this a quite systems theoretic perspective on the problem, an answer to this problem is simply not immediate in this considered setting, which is a rather remarkable conclusion.

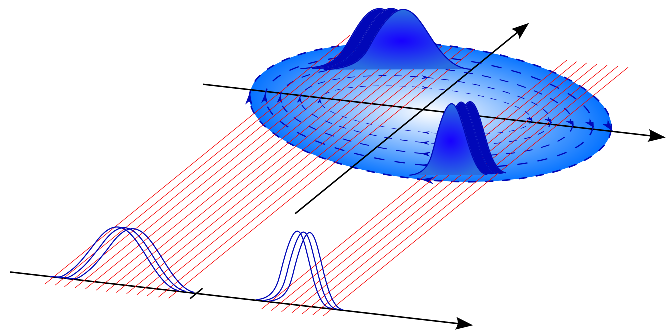

Fig. 3. Illustration of the ensemble observability problem for a two-dimensional harmonic oscillator with a bimodal initial distribution. The upper right shows the evolution of the state distribution. The lower left shows the evolution of the corresponding output distribution. Taken from [13].

Due to the aforementioned reasons, in [8] we took a different approach to the problem, which is to simply view and treat it as a (generic) inverse problem in a measure theoretic framework. In fact, the output distribution $\mathbb{P}_{y(t)}$ is related to the initial distribution $\mathbb{P}_0$ in a very basic way, namely through a pushforward relation

$$\mathbb{P}_{y(t)}(B_y) := \mathbb{P}_0((Ce^{At})^{-1}(B_y)) = \int_{(Ce^{At})^{-1}(B_y)} p_0(x)\,\mathrm{d}x.$$

The values of the output distribution are related to the initial density through these integrals over these preimages, which one can think of as a strips due linearity of $x \mapsto Ce^{At}x$, as well as the fact that the interesting cases occur only when C does not have full column rank. This basic perspective may be illustrated as in Figure 4. The remaining diffculty is then due to the fact that we only know the integrals over sets that stretch to infinity.

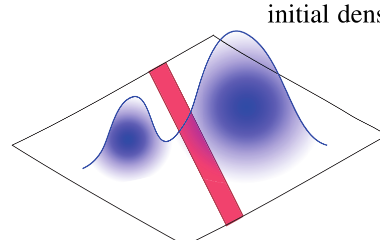

Fig. 4. Illustration of the relation between initial state distribution and output distribution at a given time. The value $\mathbb{P}_{y(t)}(B_y)$ is equal to the strip integrals $\int_{(Ce^{At})^{-1}(B_y)} p_0(x)\,\mathrm{d}x$, cf. [8].

Thus, we may only hope that as time changes, the directions of the strips, dictated by Ce^{At}, change and that the information for different directions can be combined to infer the integrand p_0. This is precisely the same problem as in tomography problems, where one wants to obtain a cross-section of an object by taking radiographs from different angles. Our study of the ensemble observability problem established a direct mathematical connection between (ensemble) observability and tomography problems [8].

Thus, in addition to the original, dynamic viewpoint, there is this second viewpoint associated to the ensemble observability problem in which we do not consider the evolution of the initial state distribution with the flow, but instead, the evolution of the "measurement directions", which are dictated by $\ker Ce^{At}$. For the example of the harmonic oscillator, the directions at which we take projections of the initial state distribution, rotate in a uniform counter-clockwise motion, which is in fact the canonical example of a tomography problem, by which the reconstructability of the ensemble observability problem for the harmonic oscillator becomes very clear. Figure 5 illustrates the duality between the two different viewpoints.

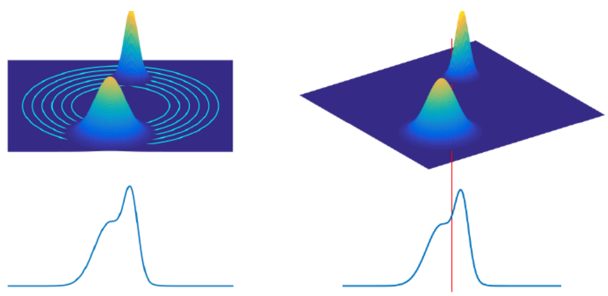

Fig. 5. Left: The distribution evolves with the flow, undergoing a rotation about the origin, and the measurement direction is fixed. Right: The distribution is held fix and we, as a (physical) observer, rotate around the object with our focus fixed on the center of the object. The observed densities are exactly the same in the two different setups.

The quite unexpected connection to tomography that was revealed in our investigation of the theoretical problem was effectively leveraged both for theoretical studies, as well as practical reconstruction schemes. In the former, the probabilistic analogue of the projection slice theorem, the Cramér-Wold theorem (see Section 3), yielded insightful algebraic geometric conditions for ensemble observability. In the latter, the Algebraic Reconstruction Technique from computed tomography provided a reconstruction method, which, unlike the previous approaches that treated the dynamic aspect more as a black box that is used only for forward simulation purposes, was anchored in a detailed systems theoretic analysis of the underlying problem; comparative studies in light of the new tomography-based viewpoint revealed significant weaknesses of the previous kernel-based reconstruction methods. The curse of dimensionality, however, was also not resolved in this new approach, so that it became apparent that a purely sample-based approach had to be derived.

To progress towards a sample-based viewpoint, the idea is to replace the formulations in terms of continuous probability distributions by sample-based descriptions. We adopt a description of the initial state distribution to be sought by finitely many particles $x^{(i)}$ with $i = 1, \dots, N$, and furthermore assume that a measurement at time $t \geq 0$ yields values $y^{(j)}$ with $j = 1, \dots, M$, which are samples from the output distribution at a given time. The pushforward relation for the continuous distributions then becomes

$$\frac{\#\{y^{(j)} \in B_y\}}{\#\{y^{(j)}\}} = \frac{\#\{x^{(i)} \in (h \circ \Phi_t)^{-1}(B_y)\ \}}{\#\{x^{(i)}\}},$$

which provides a sample-based description of the relation between the measured output histograms and the configuration of samples of the initial state distribution.

3 Formulating the ensemble state estimator

So far, we have articulated the need to consider a new type of approach in the computational ensemble observability problem, in which the sought state distribution is to be reconstructed by means of finitely many samples of it. The key problem in establishing this can be formalized as follows.

Problem 1 *Given a set of samples drawn from the output distribution, derive suitable update and correction rules for the particles of the sample-based ensemble state estimator so that these converge to a configuration that resembles a set of samples drawn from the unknown state distribution.*

It turns out that the presented sample-based description of the general pushforward equation in fact already contains all important ingredients to successfully establish a sample-based framework. It is noted however, that the exact implementation is still far from being obvious at this stage and requires further discussions. Essentially, the key idea that will enable our sample-based undertaking is in fact all along encoded in the Cramér-Wold theorem [14], which, in one of its different most prominent versions, states that if for two joint distributions *all* marginals distributions *in all directions* are the same, then the joint distributions are the same. Another way to put it is that a joint distribution is uniquely determined by its marginals in all different directions.

Theorem 1 (Cramér-Wold Theorem) *A distribution of a random vector X in $\mathbb{R}^n$ is uniquely determined by the family of distributions of $\langle v, X \rangle$, with $v \in \mathbb{S}^{n-1}$.*

The Cramér-Wold theorem can in fact be easily relaxed to cases in which marginal distributions are not available in all directions, but rather only in a smaller set of directions, which is closely related to the issue of limited angle tomography. In [8], we studied the underlying mathematical problem and were in particular able to provide complete insight into the connection between the required "minimal" set of directions and properties of (A, C), which would, analogous to the classical observability of a linear system, determine whether the underlying system is ensemble observable or not. As we will see, most examples of systems that are ensemble observable will not possess the property that $\ker Ce^{At}$ covers all possible "directions". A specific example illustrating this fact very clearly is a double integrator (see Section 4).

In light of this particular perspective on the Cramér-Wold theorem, the idea would thus be to produce samples in $\mathbb{R}^n$ so that the projections of the sample points in all available directions are as close as possible to the corresponding output histograms. The key to achieve this is to use an optimal transport approach to measure the closeness between the histograms of the projected samples and the output histograms and to devise a suitable correction strategy that will yield a matching of the two histograms.

Let us discuss this mathematically in the case that the states are n-dimensional and for one single output measurement of the form $y = \langle v, x \rangle$ with some $v \in \mathbb{R}^n$. Let the ensemble state estimator consist of N particles $\widehat{x}^{(i)}$, where N is (of course) taken to be sufficiently large and suppose that we have M scalar measurements $\langle v, x^{(j)} \rangle$, where the $x^{(j)}$ are samples from the joint distribution. We then produce a histogram for these measured samples and also produce a histogram for the projected estimator states $\langle v, \widehat{x}^{(i)} \rangle$ with the same bins $[y_\mu, y_{\mu+1}]$ with $\mu = 1, \ldots, \ell$. The situation is illustrated in Figure 6.

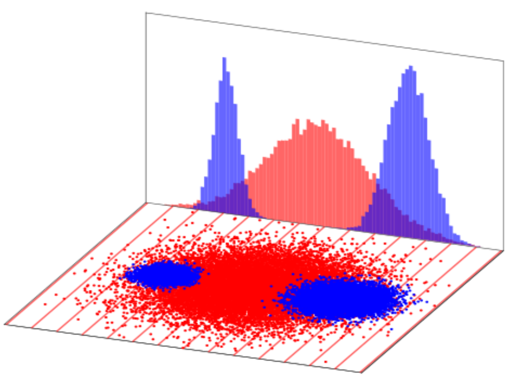

Fig. 6. This figure shows the sample points from the reference distribution (black) and the estimator's initial configuration of its sample points (red). The histograms of the marginalizations in one particular direction are illustrated in the back. By choosing the same bins for the two histograms, we can describe these as two vectors, whose entries are the (normalized) frequencies.

When the bins of the two histograms are identical, both histograms can be described by the vectors

$$q^v = \begin{pmatrix} q_1^v & \ldots & q_\ell^v \end{pmatrix}, \qquad \widehat{q}^v = \begin{pmatrix} \widehat{q}_1^v & \ldots & \widehat{q}_\ell^v \end{pmatrix}$$

containing the normalized frequency of projected samples in the respective ℓ bins. As such, they are probability vectors, i.e. $\|q^v\|_1 = \|\widehat{q}^v\|_1 = 1$. The aforementioned correction strategy is then given by "morphing" the probability vector $\widehat{q}^v$ into the probability vector q^v, i.e. to (optimally) redistribute the mass in the different bins of $\widehat{q}^v$ so as to obtain the mass distribution as specified in q^v. The problem of transforming one distribution into another by a suitable transport map is illustrated in Figure 7.

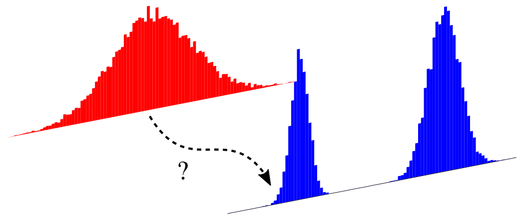

Fig. 7. This figure illustrates the idea of finding a way to transport one distribution into another, or, equivalently, transporting the associated probability vectors into another.

This is in fact the most basic instance of an optimal mass transport problem, namely one in a completely finite-dimensional setting. Here one is seeking for a so-called *transport plan*, which in the discrete setting is specified by a matrix $T \in \mathbb{R}^{\ell \times \ell}$ with non-negative entries so that

$$\sum_{i=1}^{\ell} T_{ij} = \widehat{q}_j^v, \qquad \sum_{j=1}^{\ell} T_{ij} = q_i^v.$$

The intepretation is that the entry T_{ij} would dictate how much of the (probability) "mass" $\widehat{q}_j^v$ in the jth bin of the histogram is to be transported to the ith bin, so that eventually $\widehat{q}^v$ will be completely transformed into q^v.

The aforementioned optimality is incorporated into this framework by additionally considering the cost functional

$$J = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} |i - j| \, T_{ij}.$$

From a physical perspective, this is a very reasonable choice as it favors transport plans that realize the transportation of one mass distribution into another in the most economical way. But this particular choice also leads to additional nice mathematical features, such as the fact that in this case the dual problem is a linear program involving only ℓ optimization variables instead of ℓ^2 variables. This is commonly referred to as the Kantorovich-Rubenstein duality. An even faster way to (approximately) solve this particular case of an optimal mass transport problem for large problem sizes (i.e. for large numbers of bins ℓ) is through the so-called method of Sinkhorn iterations [15].

Having solved the optimal transport problem, we obtain the transport plan T for mapping the two vectors containing the frequencies in the different bins, as illustrated in Figure 8.

In the last step, this transport plan obtained from solving the optimal transport problem on the level of population vectors needs to be converted into a correction scheme *on the level of the original particles*. This is described in the following:

(1) For all $\widehat{x}^{(i)}$ of the ensemble state estimator find the number μ of the bin in which $\langle v, \widehat{x}^{(i)} \rangle$ is contained in.

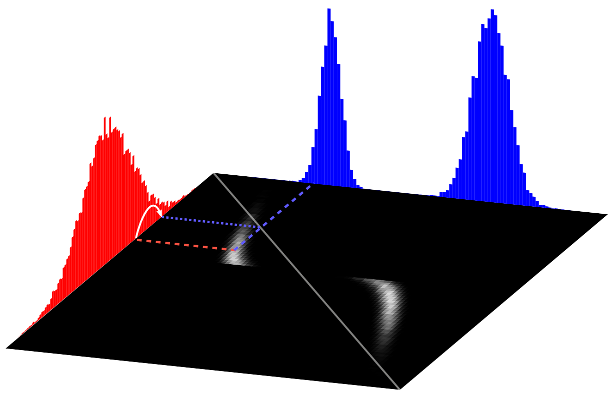

Fig. 8. This figure shows a visualization of the transport plan, with the intensity in a pixel corresponding to the magnitude of the corresponding entry in the transport plan matrix (gray scale), as well as the two marginal distributions (red and black). The red and black dashed lines indicate how the transport plan is related to the two corresponding marginal distributions. The dotted black line is the result of reflecting the black dashed line about the diagonal line, and highlights the position towards which the mass highlighted by the red dashed line is to be transported, as summarized by the white arrow between the two corresponding bins.

(2) The entries of the μth row of the matrix T are treated as transition probabilities for $x^{(i)}$ to be moved from $\{x \in \mathbb{R}^n : y_\mu \leq \langle v, x\rangle \leq y_{\mu+1}\}$ to another bin.
(3) The actual transition of $\widehat{x}^{(i)}$ from the μth bin to another, say the νth, bin is realized by translating it in the direction of $v \in \mathbb{R}^n$, i.e.

$$\widehat{x}^{(i)} \leftarrow \widehat{x}^{(i)} + \lambda \frac{v}{\|v\|^2},$$

with a suitable displacement $\lambda \in \mathbb{R}$. This yields

$$\langle v, \widehat{x}^{(i)} + \lambda \frac{v}{\|v\|^2}\rangle = \langle v, \widehat{x}^{(i)}\rangle + \lambda.$$

To ensure a certain "regularity" of the resulting set of samples, the exact displacement is also randomized, allowing the corrected particle to lie anywhere in the νth bin with equal probability. More specifically, we choose $\lambda = -\langle v, \widehat{x}^{(i)}\rangle + \tilde{\lambda}$, where $\tilde{\lambda}$ is a random variable with a uniform distribution on $[y_\nu, y_{\nu+1}]$.

The algorithm for correcting the state distribution with respect to the marginal distributions in one direction given by $y = \langle v, x\rangle$ is summarized below.

Algorithm 1 Ensemble state estimator: One correction step

1: Compute the normalized frequencies of the samples $\{\langle v, x^{(i)}\rangle\}_{i=1,\ldots,M}$ and $\{\langle v, \widehat{x}^{(j)}\rangle\}_{j=1,\ldots,N}$ with respect to the bins $[y_\mu, y_{\mu+1}]$, where $\mu = 1, \ldots, \ell$, yielding the probability vectors q^v and $\widehat{q}^v$, respectively.
2: Compute the transport plan $T \in \mathbb{R}^{\ell \times \ell}$ for $\widehat{q}^v \rightsquigarrow q^v$.
3: Implement the transport plan on the level particles.

Figure 9 illustrates a situation, in which the estimator state has been corrected with respect to the highlighted direction, but admits a large deviation with respect to a different direction. Clearly, the above described plan will have to be repeated for sufficiently many directions $v \in \mathbb{R}^n$. Figure 10 illustrates the correction scheme for two consecutive correction steps. Note how in this particular case, with only two simple iterations, we are already able to achieve a quite acceptable reconstruction. The iteration over all different directions v itself can be iterated several times, similarly to the procedure in the Algebraic Reconstruction Technique in computed tomography. The intuitive idea is that by doing so, we expect to eventually end up with a configuration of particles $\widehat{x}^{(i)}$ whose projections along all given directions are at once in accordance with the actual data. By virtue of the Cramér-Wold theorem, in the idealized case that $N \to \infty$, as well as $\ell \to \infty$, and that all (a sufficient set of) directions are available, we would end up with a perfect approximation of the joint distribution by means of samples of the distribution. We stress, however, that this does not really qualify as a rigorous argument for the presented algorithm. One possible approach for a detailed convergence analysis is through the inherent connection to the classical Kaczmarz method [16], which can in fact be seen as a special case of the novel population-based scheme presented in this paper.

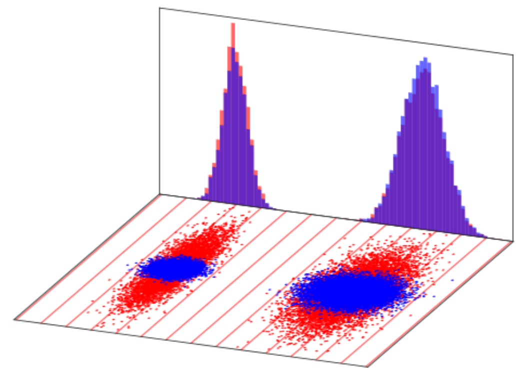

Fig. 9. This figure illustrates the situation in which the presented correction scheme has been carried out with respect to the highlighted direction. The illustrated marginal distribution of the estimator particles matches the marginal distribution of the particles from the actual initial state distribution. Note that the marginal distributions in other directions, e.g. that orthogonal to the highlighted one, are clearly not matched, which will eventually have to be addressed in further iteration steps.

3.1 The dynamic case

Up to now, we have described the correction scheme for the "static case" associated with the Cramér-Wold viewpoint. Regarding the dynamic state estimation problem for ensembles, the role of the linear and scalar $y = \langle v, x\rangle$ is taken by the nonlinear mapping $y = (h \circ \Phi_t)(x)$, which, in the very relevant linear case, specializes to $y = Ce^{At}x$. While the case of linear systems can be carried over from the static analysis in a rather straightforward fashion, the general case of nonlinear systems requires some further discussion.

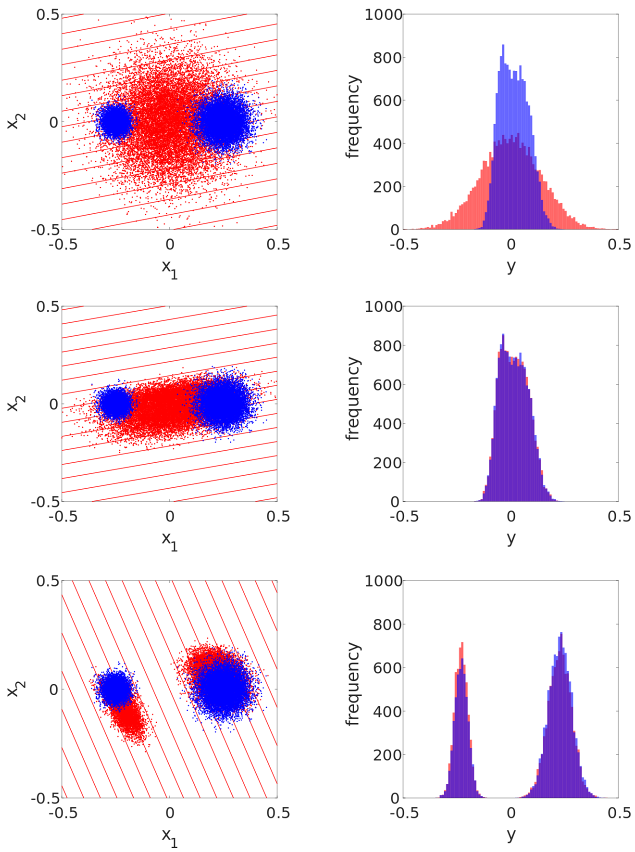

Fig. 10. First row: The actual initial distribution (black) and a prior estimate (red) are illustrated on the left. The right plot shows the histogram of the projections of the two distributions along the highlighted direction in the left plot. Middle row: The ensemble estimator's state is updated so that the marginals of the projections in the highlighted direction match. Last row: Illustration of a second update of the ensemble estimator's state associated to a different direction.

The main obstruction in the nonlinear case is in the last step of implementing the correction strategy on the level of particles, in particular the displacement of particles from one (nonlinear) "bin" $\{x \in \mathbb{R}^n : y_m \leq (h \circ \Phi_t)(x) \leq y_{m+1}\}$ to another one. Now instead of directly operating on the particles used for approximating the initial state distribution, we can first propagate these using Φ_t and then implement the correction strategy for the resulting particles at time t. Here we still need one additional assumption in general, namely that the output mapping h is such that a displacement of particles from one partition $\{x \in \mathbb{R}^n : y_m \leq h(x) \leq y_{m+1}\}$ to another can be readily implemented. After the correction step on the level of forward propagated particles, we apply the reverse flow. Thus, a simple remedy by which the correction with respect to the output mapping $y = (h \circ \Phi_t)(x)$ is circumvented is given by splitting up what in the linear case can be naturally implemented in a single step into two steps. A detailed illustration of one correction step in the nonlinear case is shown in Figure 11, where the resulting action of the unfolded correction procedure is also clearly displayed

To summarize, in the above described unfolded correction procedure, we transport the state distribution of the estimator forward to a given measurement time, compare its output distribution with the measured output distribution at hand, and then implement the correction at that given measurement time. Then, after the transport plan has been implemented, the state distribution is transported backwards to the initial time. For linear systems, we can of course directly apply the correction in one simple step by means of the simpler relation $y = Ce^{At}x$.

3.2 On the architecture of the ensemble state estimator

At this point, we would like to draw some attention to the particular hierarchical architecture of this correction-based (particle) state estimator. The correction is essentially implemented by means of a two-layer feedback: First, the mismatch between the outputs of the estimator and of the actual system is evaluated on the *population level*, from which a correction on the population level is computed. In particular, at this stage, no attention is paid to individual systems but only the totality of individual systems. The correction in the next step on the other hand has to be actually realized by implementing it on the level of the individual particles. In particular, it cannot be fully implemented on the population level, i.e. by completely broadcasting an instruction to the systems in the ensemble. Rather, different individual systems in the population will be required to receive different instructions (in this case it is based on the bins they are located in). To summarize, though our presented scheme does not operate entirely on a population-level, it is also not a completely individual feedback. Rather it constitutes a quite simple to implement, yet very powerful *hybrid*, given by a two-layer structure, which we may refer to as a *population-level feedback*.

4 The ensemble observer

In the previous section, we presented a novel particle-based approach for estimating the initial state distribution of an ensemble from output samples. As for any such state estimation problem, we assumed to have all the measurements at different times stored and available to us at once. Another type of state reconstruction scheme is in a more dynamic spirit, in which the system's state is to be estimated online, i.e. at each time instant, the estimated state is updated based on the measurement received at that time point, or, more generally, from past measurements received up to that time point. From a more mathematical point of view, the problem considered in this section is the estimation of $p_{x(t)}$ from past output measurements $p_{y(\tau)}$, with $\tau \leq t$, which, when formulated in these more theoretical terms, we recognize to be analogous to a classical filtering problem. So far, approaches to implement such a filtering approach have not yielded any fruits.

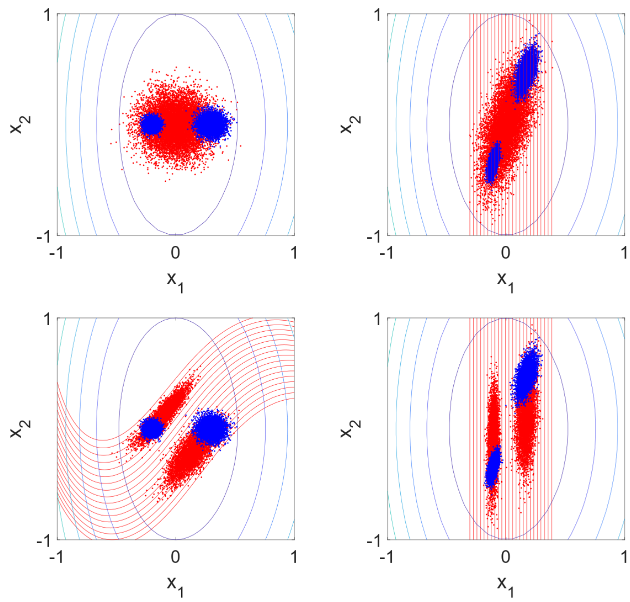

Fig. 11. Top left: The initial state distribution (black) and the estimated initial state distribution (red) before any correction step has been applied. Top right: The two distributions after being transported with the nonlinear oscillator to a given time point, as well as the level sets of the output measurement. Lower right: Correction step using optimal mass transport. Lower left: The transported corrected distribution, as well as the transported level sets.

4.1 Discussion of related approaches

To illustrate particular difficulties that were encountered in the aforementioned approaches, we shall highlight two approaches that one would rather naturally consider in this context. The first approach would consider a partial differential equation describing the evolution of the estimated state distribution. It is well-known that the original ensemble system can be described by a linear partial differential equation, the Liouville equation [7], given by

$$\frac{\partial}{\partial t} p(t,x) = -\mathrm{div}(p(t,x)f(x)),$$

where $p(t,\cdot)$ denotes the state density at time t. The output distribution results from the state distribution by a marginalization along $\ker C$, i.e.

$$p_{y(t)}(y) = \int_{Cx=y} p(t,x)\,\mathrm{d}S.$$

We denote the mapping $p(t,\cdot) \mapsto p_{y(t)}$ by $\mathcal{C}$. In the spirit of the classical Luenberger observer [17], having one part simulating the system and another part correcting based on the incoming output measurements as its basic design principle, it is indeed natural to consider an observer described by

$$\frac{\partial}{\partial t}\widehat{p}(t,x) = -\mathrm{div}(\widehat{p}(t,x)f(x)) + \mathcal{L}[\widehat{p}_{y(t)} - p_{y(t)}],$$

where $\widehat{p}_{y(t)} = \mathcal{C}\widehat{p}(t,\cdot)$. Defining $e(t,x) := \widehat{p}(t,x) - p(t,x)$ as the estimation error, in the approach based on partial differential equations, the problem boils down to designing the (linear) operator $\mathcal{L}$ so that the error dynamics

$$\frac{\partial}{\partial t} e(t,x) = -\mathrm{div}(e(t,x)f(x)) + (\mathcal{L}\mathcal{C}e)(t,x)$$

is asymptotically stable. However, due to the fact that the action of $\mathcal{C}$ is a rather unique one, not falling into any well-studied category of operators in the theory of infinite-dimensional systems theory [18], a general solution to this stabilization problem remains out of reach.

Another natural idea that circumvents the infinite-dimensional setting is to first discretize the state space, e.g. by approximating the considered probability density functions by piecewise constant functions, and then to reformulate the system dynamics for these finite-dimensional approximations. However, in trying to do so, we will at some point encounter a rather fundamental problem associated to this idea, which can be already seen for a simple linear oscillator. If the discretization of the state space is not tailored to the specific vector field at hand, say, we choose a simple discretization into pixels in $\mathbb{R}^2$, then the resulting discretized linear system will no longer admit the mass preserving property. This is because in implementing this discretization scheme, we inevitably have to truncate the discretization of the state space to some region of interest, of which the boundaries will suffer from leakage of mass, but will not provide mass from outside, the outside part being truncated. Thus, for an observer based on this idea of discretization, the part that simulates the system will not be able to reproduce the actual system behavior. In fact, the state generated by the simulation part will naturally converge to zero as the incoming flow inevitably has to be truncated, and the general trend will thus be that the whole mass will eventually leak out at the boundaries.

4.2 Towards sample-based population observers

Using the new insights from our first sample-based implementation of an ensemble state estimator, we are already able to formulate a new sample-based (online) ensemble observer. Recalling the filtering formulation introduced in the beginning of this section, where the problem is to estimate $p_{x(t)}$ from past output measurements $p_{y(\tau)}$, with $\tau \leq t$, the key to producing online estimates of the state distribution lies in the equation

$$(h \circ \Phi_{\tau - t})(x(t)) = y(\tau)$$

relating the state of a (single) nonlinear system at time t with its output measurement at an earlier time point $\tau \leq t$. This basic relation shows that the estimation of the current ensemble state distribution from past output distributions $\tau \leq t$ is inherently dual to the estimation of the ensemble initial state distribution from the generated output distributions (forward in time).

Combining this with our insights from the previous section, a sample-based population observer can be realized as follows.

Algorithm 2 Sample-based population observer

1: Given some time $t_k \geq 0$, and an earlier time t_{k-1}, we apply the flow $\Phi_{t_k - t_{k-1}}$ to the particles of the prior estimate $\mathbb{P}_{\widehat{x}(t_{k-1})}$ to obtain a first estimate of $\mathbb{P}_{\widehat{x}(t_k)}$.
2: For various different $\tau \leq t_k$, we apply the correction scheme described in the previous section with the sole modification of replacing Φ_t with $\Phi_{\tau - t_k}$.

4.2.1 Practical implementation for a double integrator

In this subsection, we illustrate a practical implementation of sample-based population observers in more detail. For this purpose, we consider the system

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x, \qquad y = \begin{pmatrix} 1 & 0 \end{pmatrix} x,$$

which is a simple double integrator. Here we can directly compute $Ce^{At} = \begin{pmatrix} 1 & t \end{pmatrix}$, allowing us for the specific example of a double integrator to write down the relation between angle α of $\ker Ce^{At}$ and time t explicitly as

$$\tan(\alpha) = \frac{x_2(t)}{x_1(t)} = t \quad \Leftrightarrow \quad \alpha = \arctan(t). \qquad (1)$$

This simple reading relation shows that unlike in the example of a harmonic oscillator, the maximal spread of achievable angles is inherently restricted to the range of $t \mapsto \arctan(t)$. Moreover, the explicit relation allows us to choose the time points of measurement in such a way that the corresponding set of angles is uniformly distributed, which in turn is expected to yield better results for the reconstruction.

Since unlike in the classical Kalman filter for linear systems, the availability of measurements at different time points $\tau \leq t$ is crucial for population observers, further discussions on the available measurements for the correction step with regard to computational effort and memory are required. It is a quite natural idea to first restrict the time points at which output data is available to $t - T_H \leq \tau \leq t$, where T_H denotes the horizon length and $[t - T_H, t]$ is called the moving horizon. Of course, when practically implementing such a moving horizon scheme, we also need to further assume that the measurement times are discrete. In this example, the specific horizon length $T_H = 3$ was chosen so as to guarantee a sufficiently large spread of available directions for each correction step. Note that by defining $\tau' := \tau - t$, which takes values in the interval $[-T_H, 0]$, we see that the directions are dictated by $\ker Ce^{A\tau'}$, where $\tau' \in [-T_H, 0]$, which result from transporting $\ker C$ forward in time in the interval $[0, T_H]$. Due to the simple relation between angles and times established in (1), the range of available angles would be $0°$ to $\arctan(3) \approx 71.56°$. In order to facilitate a wider spread, a longer horizon would need to be provided. For example, in order to have a spread of $85°$, it would already require a horizon length of $T_H = 11.43$.

Within the measurement horizon, the output distributions of the actual ensemble are of course not measured continuously, but at discrete time points. In the example, the times at which measurements of the output distributions are available are $t_k = t - 0.1k$, where $k = 1, \ldots, 30$. Out of these 30 measurement times only 10 are actually utilized by incorporating the measurement data for the correction steps at each prediction step. Of course, one could in fact increase the number of time points used for the reconstruction at each time step, and also increase the number of correction steps performed at each prediction step. This would, however, result in an increased computational load at each prediction step. We note that in the implementation of this example, the 10 time points are chosen randomly from the above measurement times in such a way that the distribution of corresponding angles would be as uniform as possible. More specifically, due to the nonlinear relation $\alpha = \tan(t)$, choosing random times t_k from a uniform distribution defined over $\{t_k\}$ would not result in a uniform distribution of the corresponding α_k. Instead, one has to sample with respect to a specific (discrete) distribution $t_k \sim P_t$, which guarantees that the distribution for $\alpha_k = \tan(t_k)$ is (close to) a uniform distribution. The detailed discussion of these issues, while of great practical importance, is beyond of the scope of this paper. In summary, with the above described procedure we obtain a quite satisfactory method for solving the continuous ensemble observability problem in an on-line fashion.

Figure 12 illustrates a resulting tracking process using the proposed method for an ensemble of double integrators.

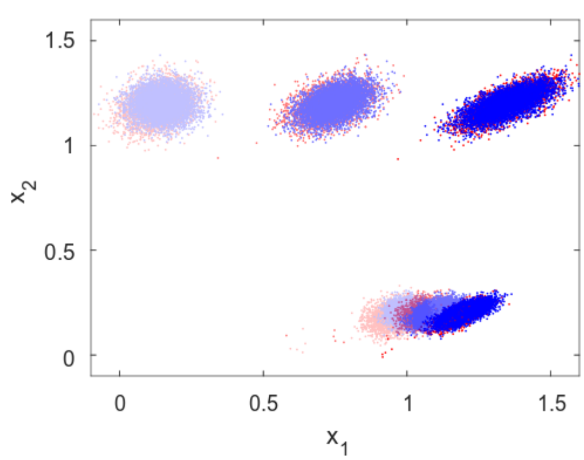

Fig. 12. Three successive predictions at $t = 0.5, 1.0, 1.5$ (distinguished by transparency), each of which is computed from 10 time points in the measurement horizon with $T_H = 3$.

5 Conclusions and Outlook

In the present paper, a first sample-based treatment of the estimation and observation problems associated to the recently emerging class of ensembles of dynamical systems was presented in an introductory manner. The sample-based approach completely circumvents the route over parameterizing the unknown nonparametric probability distribution, which is common to all previous approaches and a crucial aspect, as it inherently limits all previously considered algorithms to problem setups in which the state-space is low-dimensional.

The starting point for establishing a sample-based approach is the premise of strictly using a set of points in state space as a means to describe / track a distribution rather than to use other approximations such as histograms or more general kernel functions. The main challenge then was to devise an iterative strategy that operates by manipulations on the set of points which would eventually result in the convergence of the set of points to a configuration that could very well be a set of samples from the distribution of interest. From a conceptual point of view, a main result of this paper is the demonstration that optimal mass transport problems, as well as the classical Cramér-Wold device, when viewed through the lens of statitics, constitute crucial links in the endeavor to derive *sample-based* population observers.

A key feature of the correction scheme is the interesting two-layer structure that promotes a very basic and simple implementation: The corrective measures for the set of points is computed in a global fashion, based on population-level mismatches, but is eventually implemented on the level of individual particles by feeding population-level data to the individual particles, which compute their own correction by implementing a simple randomized strategy. As a prototype model for the more general scheme portrayed in a two-dimensional state-space, we may consider the system

$$\dot{x}(t) = \begin{pmatrix} \cos(\alpha(t)) \\ \sin(\alpha(t)) \end{pmatrix} (\cos(\alpha(t)) \;\; \sin(\alpha(t)))(x_{\text{ref}}(t) - x(t)),$$

where again the reference signal of the individual systems $x_{\text{ref}}(t)$ is obtained from population-level considerations and could differ for different systems in the population. Another interesting open problem in this regard is to derive optimal sequences of angles, possibly formulated in a stochastic framework, that yield a fast convergence for arbitrary configurations of sample points.

Lastly, while the problem formulation in this paper considers populations of dynamical systems, the particle-based tools in this paper might also be of use to particle filtering techniques for inferring the state of a single nonlinear system.

References

[1] R. E. Kalman, "On the general theory of control systems," in *Proc. First International Congress of Automatic Control*, 1959.

[2] ——, "Mathematical description of linear dynamical systems," *SIAM Journal on Applied Mathematics*, vol. 1, no. 2, pp. 152–192, 1963.

[3] R. Brockett and N. Khaneja, "On the stochastic control of quantum ensembles," in *System theory*. Springer, 2000, pp. 75–96.

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

[5] J.-S. Li and N. Khaneja, "Ensemble control of bloch equations," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 528–536, 2009.

[6] J.-S. Li, "Ensemble control of finite-dimensional time-varying linear systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 345–357, 2011.

[7] R. W. Brockett, "Notes on the control of the Liouville equation," in *Control of Partial Differential Equations*. Springer, 2012, pp. 101–129.

[8] S. Zeng, S. Waldherr, C. Ebenbauer, and F. Allgöwer, "Ensemble observability of linear systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1452–1465, 2015.

[9] S. Zeng, H. Ishii, and F. Allgöwer, "Sampled observability and state estimation of linear discrete ensembles," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2406–2418, 2017.

[10] S. Zeng and F. Allgöwer, "On the ensemble observability problem for nonlinear systems," in *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp. 6318–6323.

[11] C. Villani, *Optimal transport, old and New*. Springer, 2008.

[12] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport over a linear dynamical system," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2137–2152, 2017.

[13] S. Zeng, "Ensemble observability of dynamical systems," Ph.D. dissertation, University of Stuttgart, 2016.

[14] H. Cramér and H. Wold, "Some theorems on distribution functions," *Journal of the London Mathematical Society*, vol. 1, no. 4, pp. 290–294, 1936.

[15] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *Advances in neural information processing systems*, 2013, pp. 2292–2300.

[16] T. Strohmer and R. Vershynin, "A randomized Kaczmarz algorithm with exponential convergence," *Journal of Fourier Analysis and Applications*, vol. 15, no. 2, pp. 262–278, 2009.

[17] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.

[18] R. F. Curtain and H. Zwart, *An introduction to infinite-dimensional linear systems theory*. Springer, 1995, vol. 21.