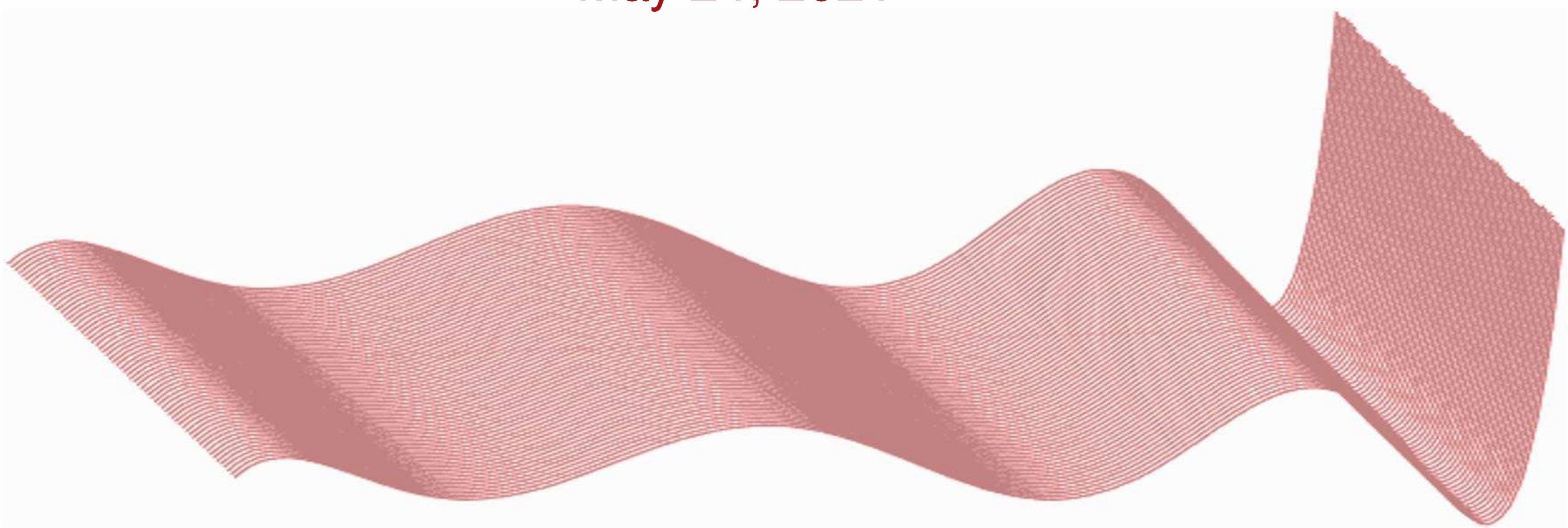
Sample-based Population Observers

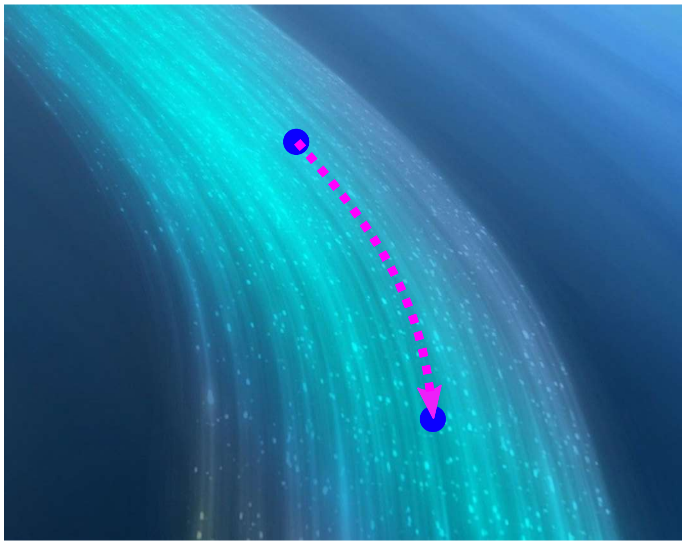
Shen Zeng
Washington University in St. Louis

2021 American Control Conference Workshop
Control of Distributions: Theory and Applications
May 24, 2021



Ensembles in control theory

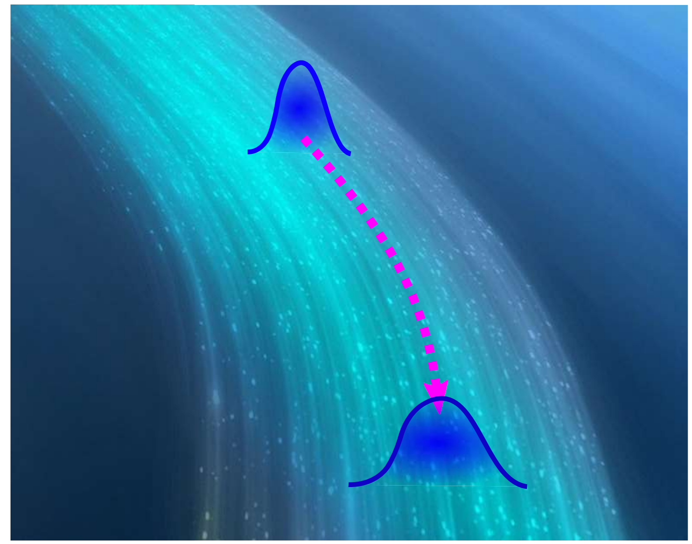
There is a recent trend in control theory to not just consider the dynamics of a *single point* but, more generally, *a distribution of points*.



A single system

$$\dot{x}(t) = f(x(t), u(t)),$$
$$y(t) = h(x(t))$$

with $x(0) = x_0 \in \mathbb{R}^n$.



An ensemble of systems

$$\dot{x}(t) = f(x(t), u(t)),$$
$$y(t) = h(x(t))$$

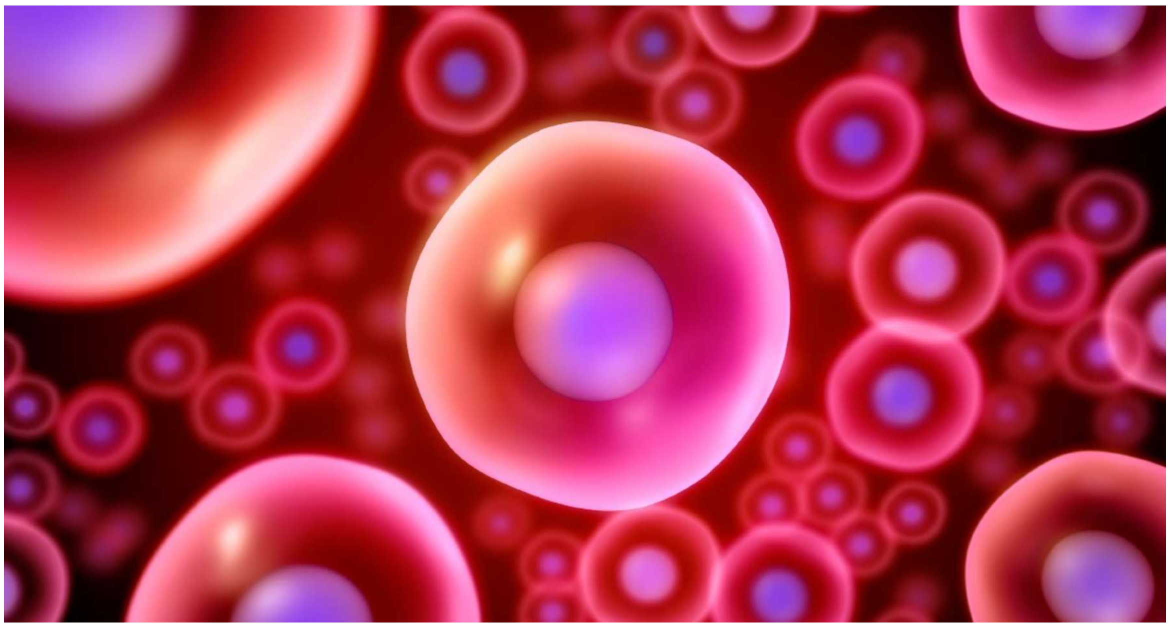
with $x(0) \sim \mathbb{P}_0$.

Interacting with populations of dynamical systems

Statistical mechanics: probability distribution as a model for ...

... *one uncertain system* or *an actual population of many systems*.

This model splits into two branches when considering inputs & outputs:



heterogeneous cell populations

Premise for ensembles: *interaction* on the population level only.

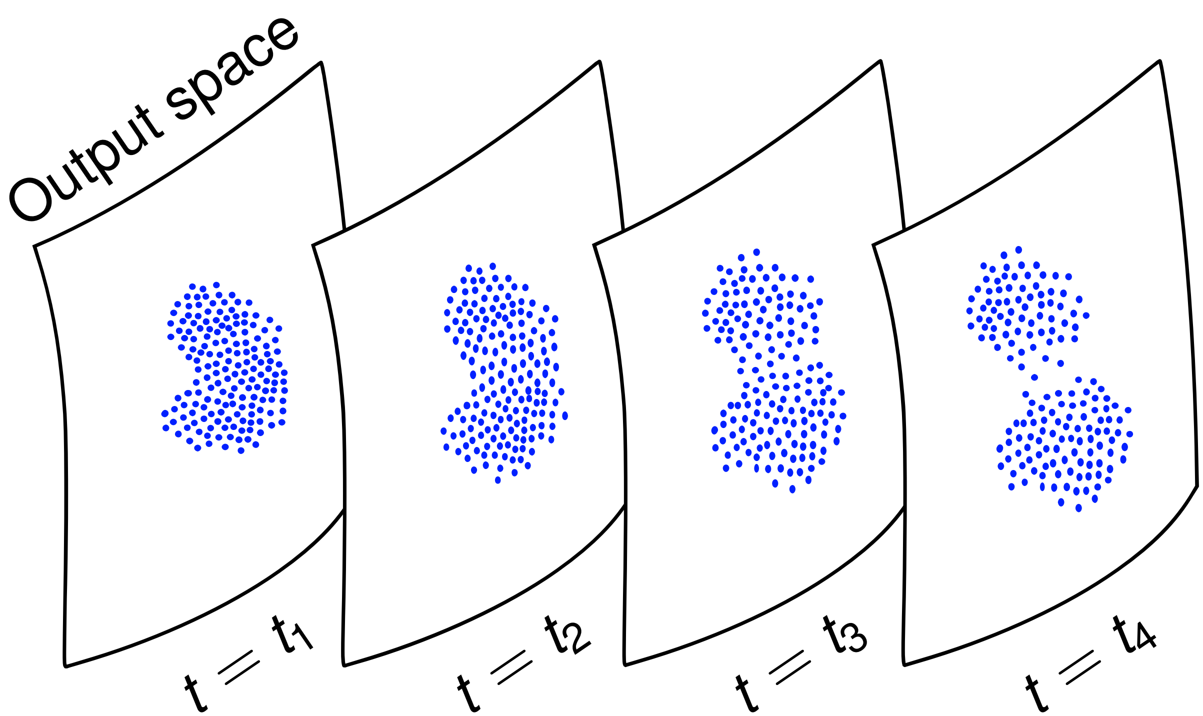
Table of Contents

Population level measurements

Reconstructing the distribution of states in cell populations ...

- J. Hasenauer et al. Identification of models of heterogeneous cell populations from population snapshot data. *BMC Bioinformatics*, 2011.
- H. T. Banks et al. A review of selected techniques in inverse problem nonparametric probability distribution estimation. *Journal of Inverse and Ill-Posed Problems*, 2012.

... which are often the cause of heterogeneous responses (cf. cancer).



Interpret aggregated output data as samples from output distribution!

Ensemble observability of dynamical systems

Reconstructing the distribution of states in cell populations ...

- J. Hasenauer et al. Identification of models of heterogeneous cell populations from population snapshot data. *BMC Bioinformatics*, 2011.
- H. T. Banks et al. A review of selected techniques in inverse problem nonparametric probability distribution estimation. *Journal of Inverse and Ill-Posed Problems*, 2012.

... *by just attempting to minimize the mismatch to the measured data.*

It was only recently that the *systems theoretic core* was formulated:

The ensemble observability problem

Consider a finite-dimensional system

$$\dot{x}(t) = f(x(t)),$$
$$y(t) = h(x(t)),$$

with initial state $x(0) \sim \mathbb{P}_0$. Under which conditions can $\mathbb{P}_0$ be reconstructed from the evolution of the distribution of $y(t)$?

S. Zeng. Ensemble Observability of Dynamical Systems. *Ph.D. dissertation*, 2016.

S. Zeng, and F. Allgöwer. On the Ensemble Observability of Dynamical Systems. *Proc. 22nd International Symposium on Mathematical Theory of Networks and Systems, 2016.*

Ensemble observability of linear systems

We consider a linear system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0,$$
$$y(t) = Cx(t),$$

and let $x_0 \sim \mathbb{P}_0$ be a random vector. Denote further $y(t) \sim \mathbb{P}_{y(t)}$.

Definition (Ensemble observability of linear systems)

A linear system (A, C) is said to be ensemble observable, if

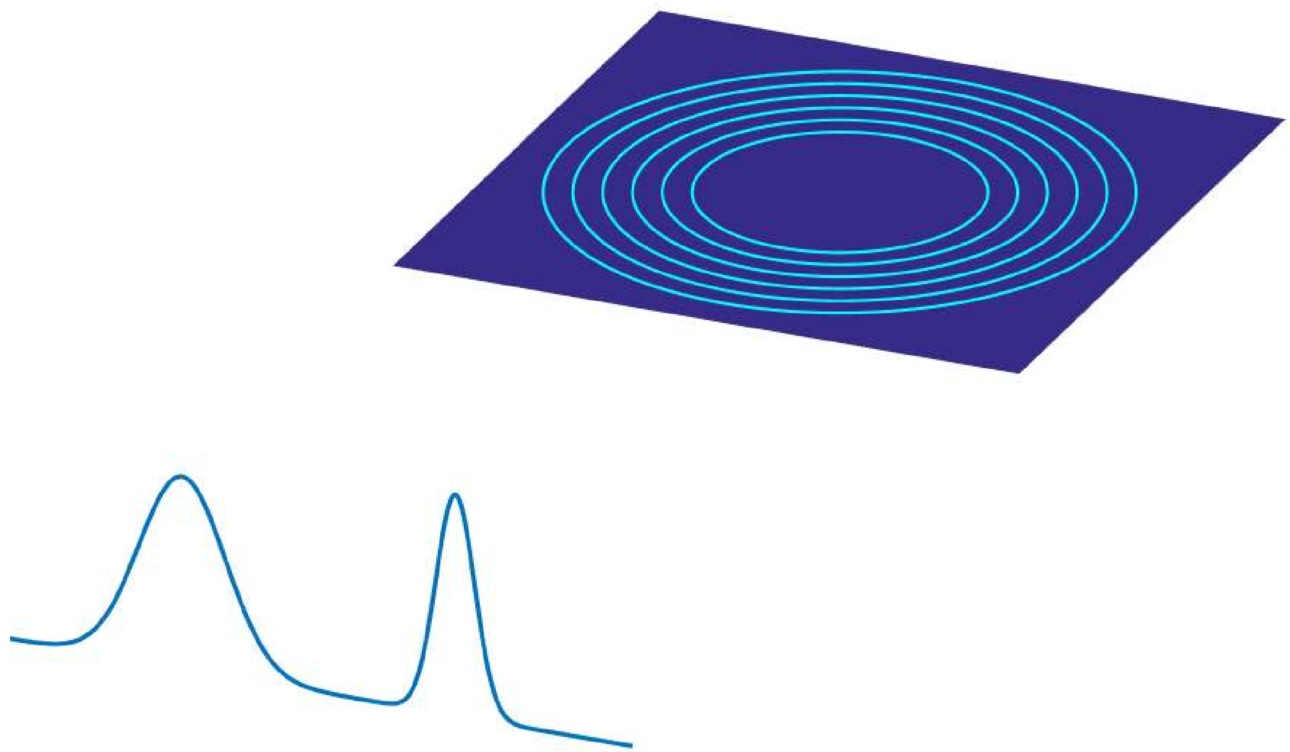
$$\left(\forall\, t \geq 0 \;:\; \mathbb{P}'_{y(t)} = \mathbb{P}''_{y(t)}\right) \;\Rightarrow\; \mathbb{P}'_0 = \mathbb{P}''_0,$$

where $\mathbb{P}'_0$ and $\mathbb{P}''_0$ are continuous probability distributions.

S. Zeng, S. Waldherr, C. Ebenbauer, and F. Allgöwer. Ensemble Observability of Linear Systems. *IEEE Transactions on Automatic Control*, vol. 61, 2016.

An illustration of the ensemble observability problem

Consider a two-dimensional harmonic oscillator with an output $y = x_1$.



Can we reconstruct $\mathbb{P}_0$ from *only observing* $\mathbb{P}_{y(t)}$ in this example?

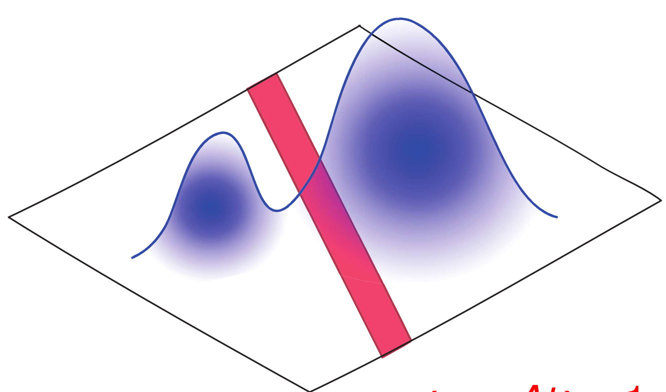
Table of Contents

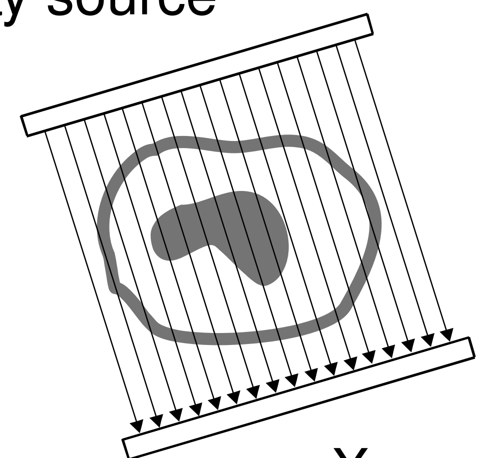
The reconstruction problem as an inverse problem

The output distribution $\mathbb{P}_{y(t)}$ is the pushforward of $\mathbb{P}_0$, i.e.

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

An illustration of the situation:





A novel connection between observability (1960) and tomography (1960)

Analogy: inferring *internal structure* from *external measurements*.

S. Zeng, S. Waldherr, and F. Allgöwer. An inverse problem of tomographic type in population dynamics. *In Proc. 53rd Conference on Decision and Control*, 2014.

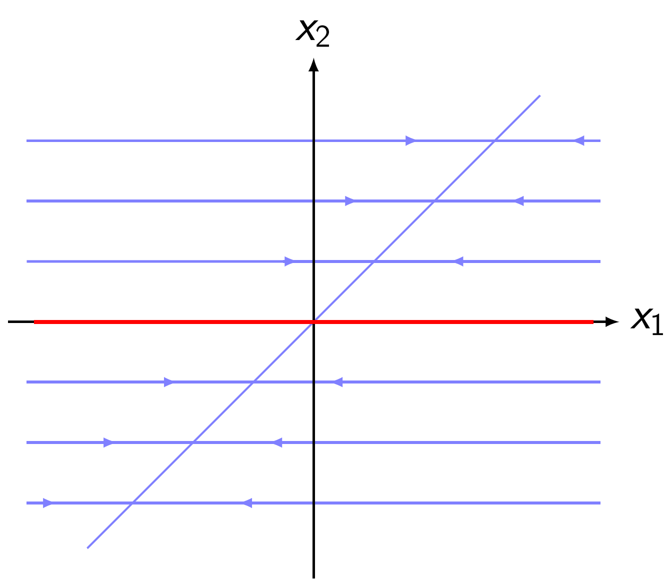
Connection to classical observability - Illustrative example

The "directions" given by $\ker Ce^{At} = e^{-At}(\ker C)$ depend on (A, C):

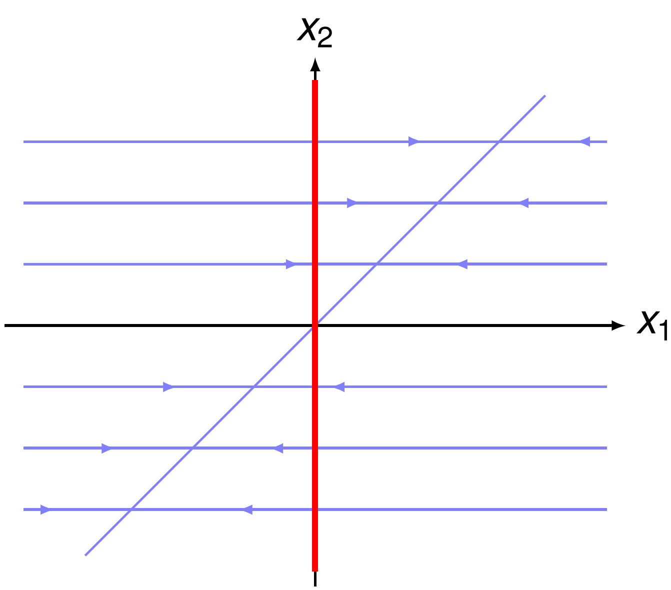
$$\dot{x} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} x \qquad\qquad \dot{x} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} x$$

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

unobservable system



observable system



Ensemble observability of nonlinear systems

S. Zeng, and F. Allgöwer. On the ensemble observability problem for nonlinear systems. *In Proc. 54th IEEE Conference on Decision and Control*, 2015.

Example

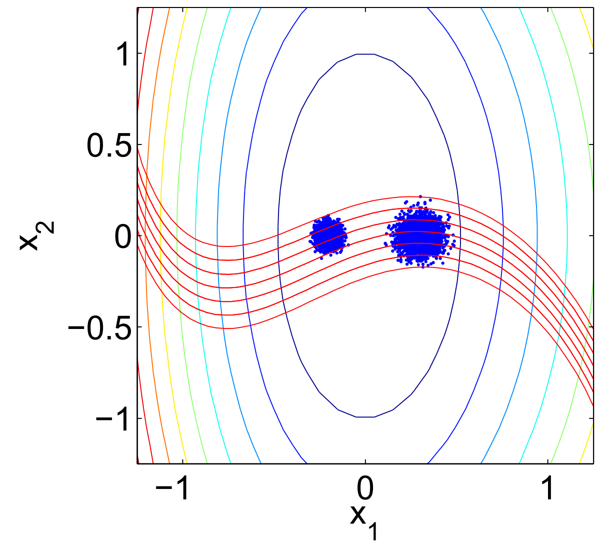
Consider the nonlinear oscillator

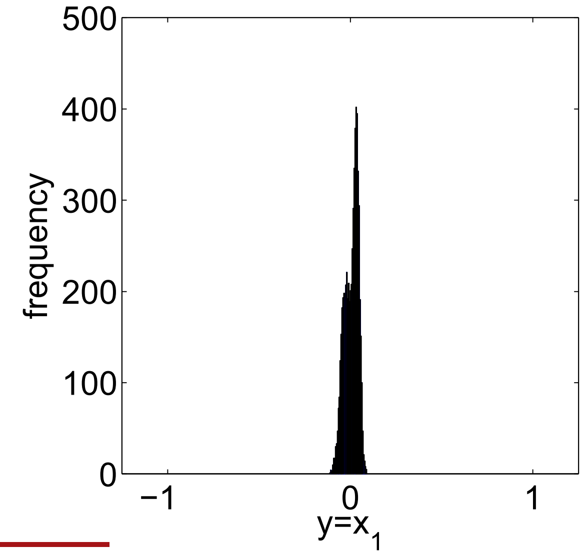
$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = -4x_1 + x_1^2,$$

with an output given by $y = x_1$.

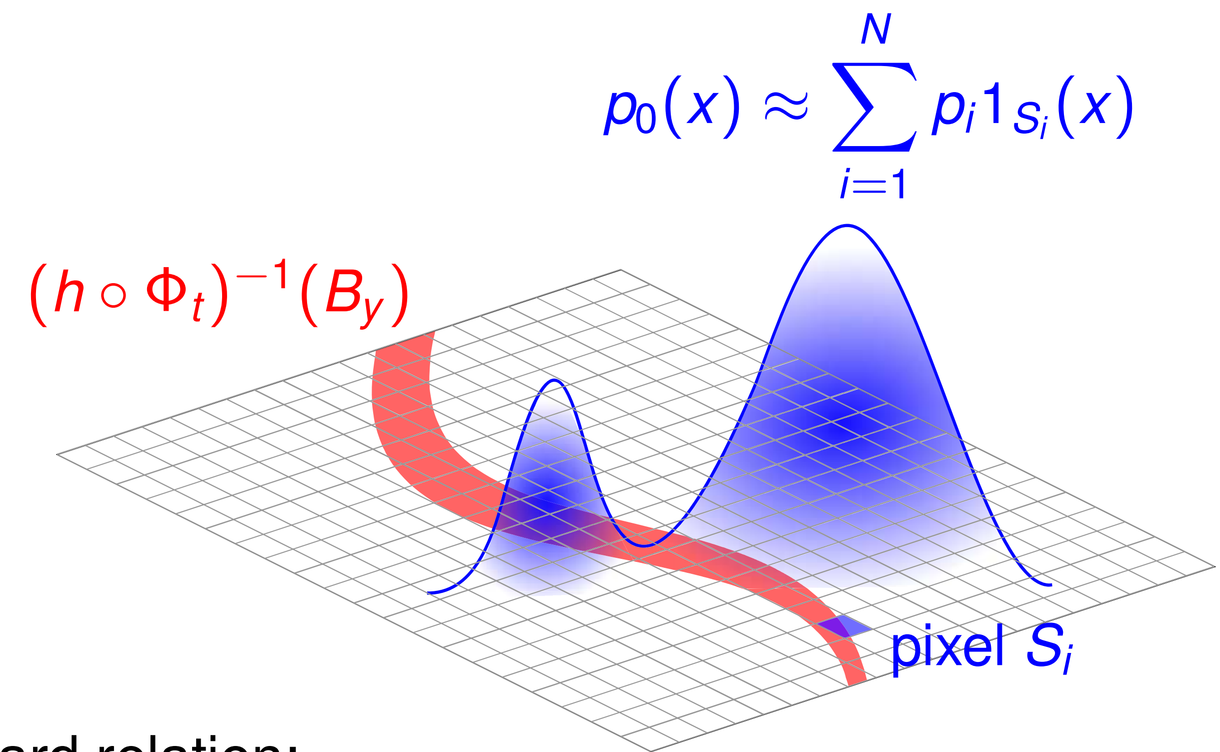
- The "scanning geometry" depends on the intertwining between the level sets of $h : \mathbb{R}^n \to \mathbb{R}^m$ and the flow $\{\Phi_t\}_{t\geq 0}$.
- We have $\int_{(h\circ\Phi_t)^{-1}(B_y)} p_0 \, \mathrm{d}x = \mathbb{P}_{y(t)}(B_y)$.
- By virtue of these insights, the nonlinear ensemble observability problem becomes amenable to computational solutions.





Practical reconstruction using the tomography framework

Introduce a grid on the state space via pixels S_i and approximate p_0:

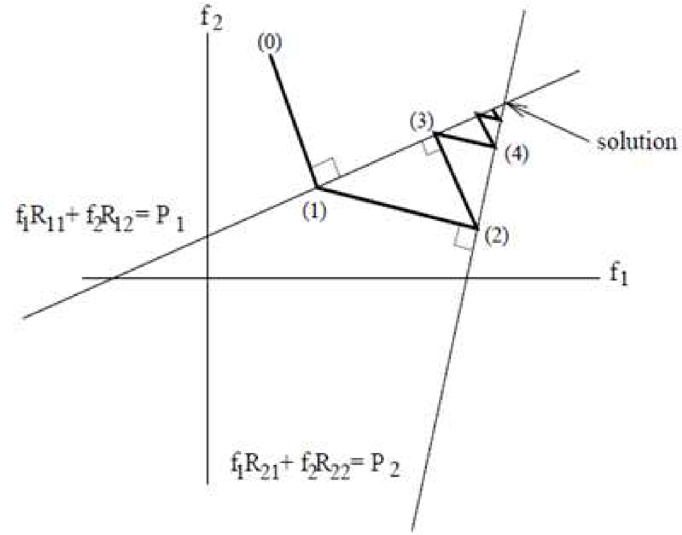


Pushforward relation:

$$\sum_{i=1}^{N} \left(\int_{(h \circ \Phi_t)^{-1}(B_y)} 1_{S_i} \, \mathrm{d}x \right) p_i = \mathbb{P}_{y(t)}(B_y).$$

Practical reconstruction using the tomography framework

The resulting huge, but sparse, system of linear equations is typically solved iteratively using projection-based methods.



Kaczmarz method

For $A \in \mathbb{R}^{N\times m}$ with $N \gg m$, and $b \in \mathbb{R}^N$, the iteration is given by

$$x_{k+1} = x_k + \frac{b_k - \langle a_k, x_k\rangle}{\|a_k\|^2} a_k.$$

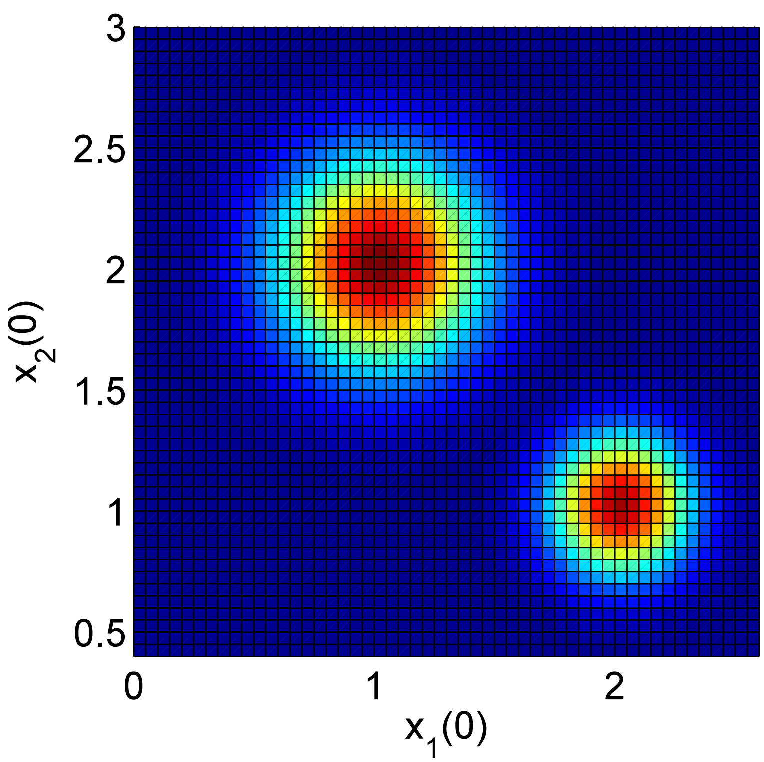
Once $k = N$, one may reinitialize $x_0 := x_N$ and iterate over k again.

A practical reconstruction example

Reconsider the observable system

$$\dot{x}(t) = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} x(t), \quad x(0) \sim \mathbb{P}_0$$
$$y(t) = \begin{pmatrix} 1 & 0 \end{pmatrix} x(t),$$

where $\mathbb{P}_0$ shall now be given as the following bimodal distribution.



A practical reconstruction example

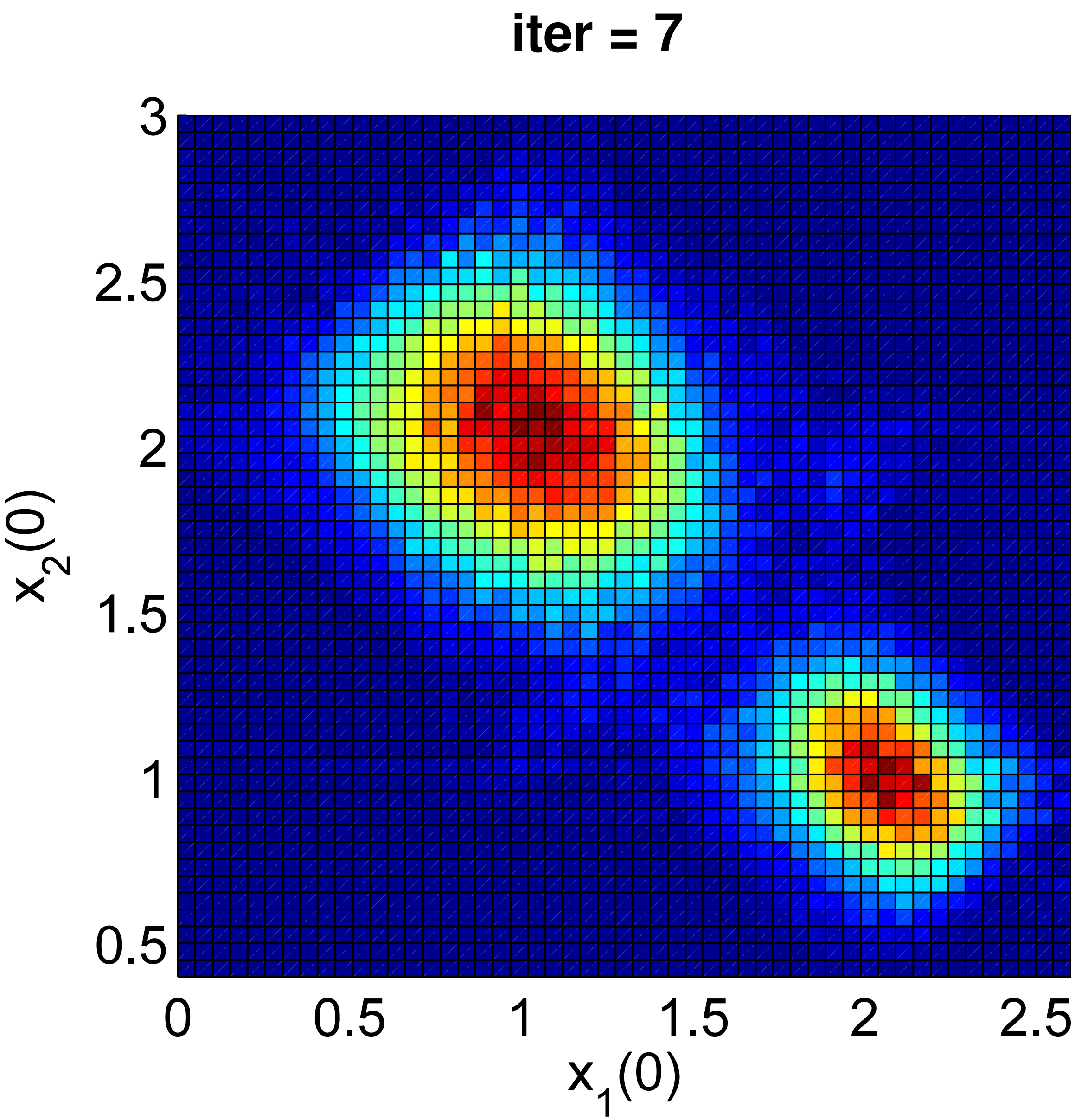


Table of Contents

Sample-based population observers

Recap: The actual data comes in the form of samples.

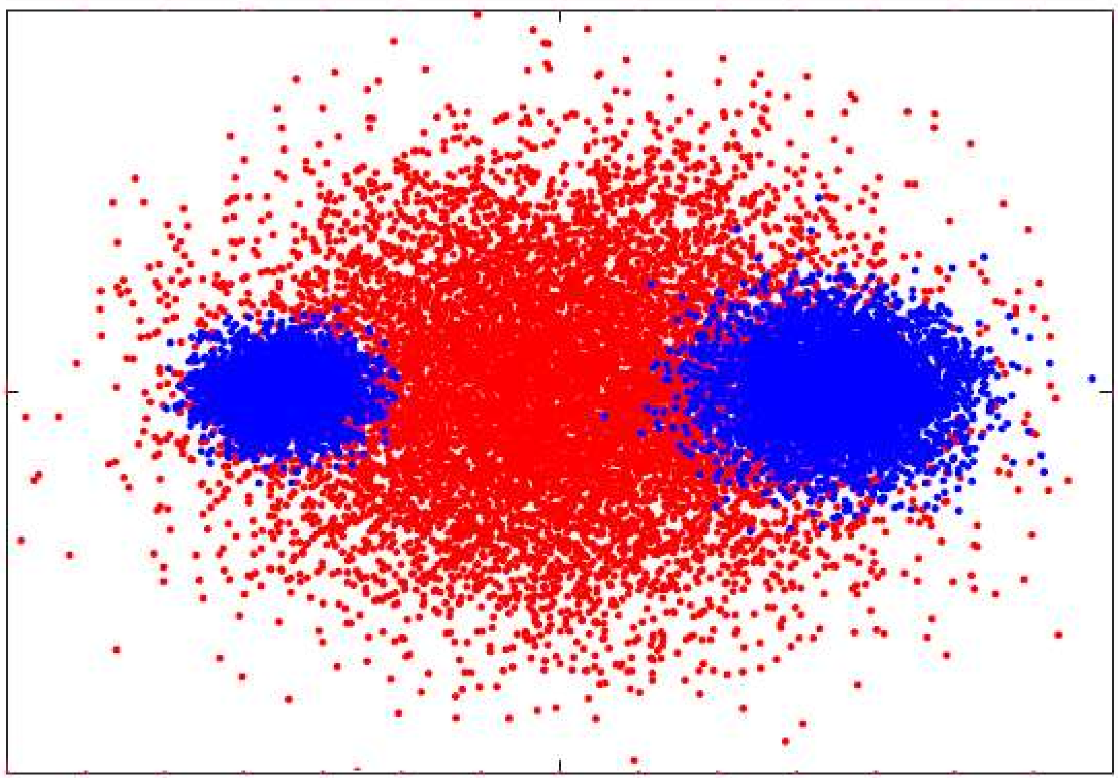
Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

Sample-based population observers

Recap: The actual data comes in the form of samples.

Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

Sample-based population observer in a picture:

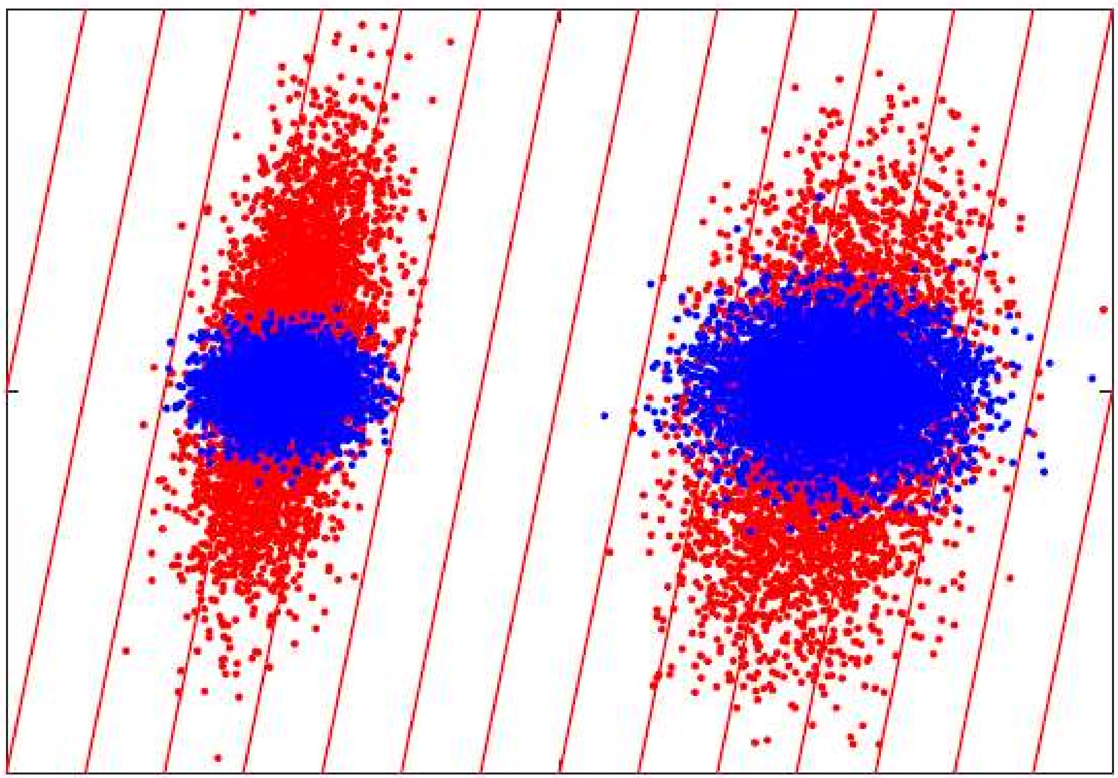


Sample-based population observers

Recap: The actual data comes in the form of samples.

Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

Sample-based population observer in a picture:

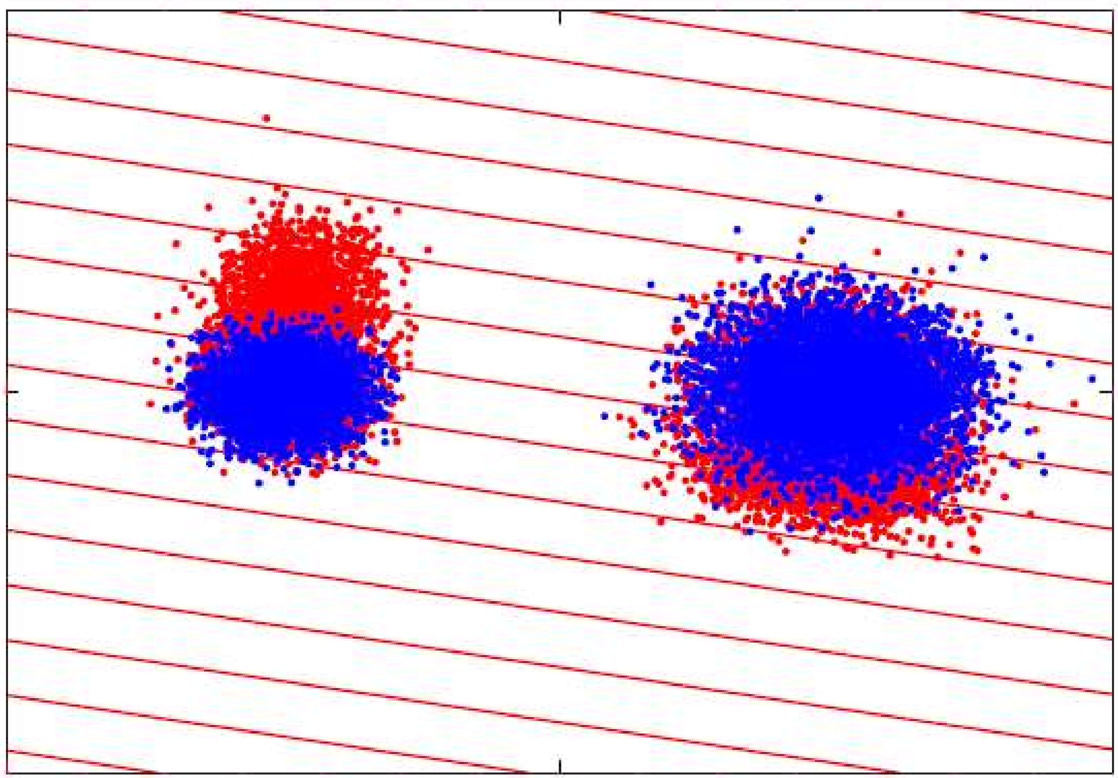


Sample-based population observers

Recap: The actual data comes in the form of samples.

Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

Sample-based population observer in a picture:

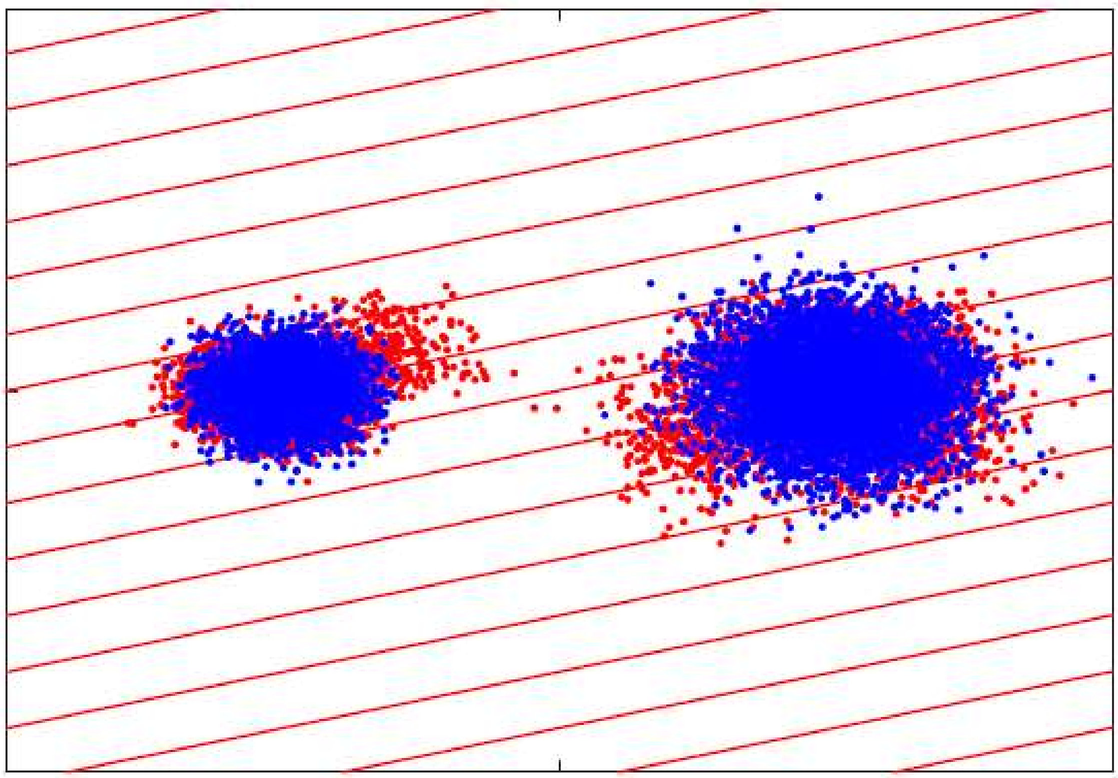


Sample-based population observers

Recap: The actual data comes in the form of samples.

Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

Sample-based population observer in a picture:

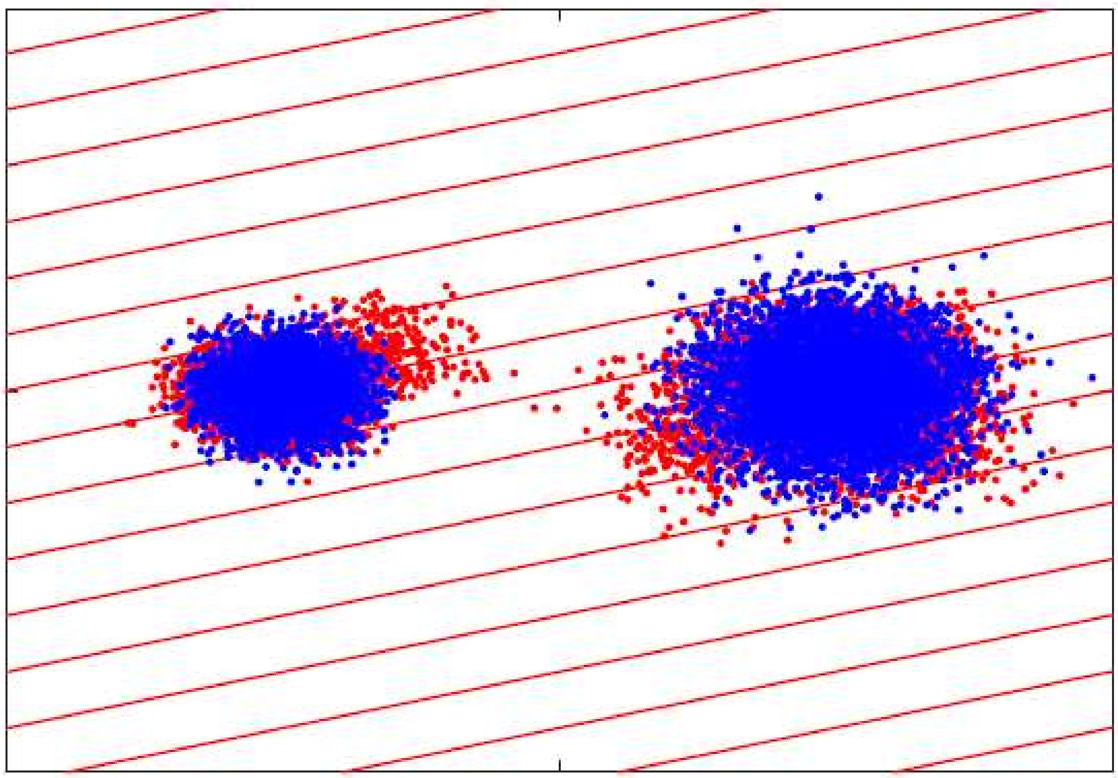


Sample-based population observers

Recap: The actual data comes in the form of samples.

Thought: It would be nice to also have a sample-based representation (vs. parametric representation) for the state distribution estimation!

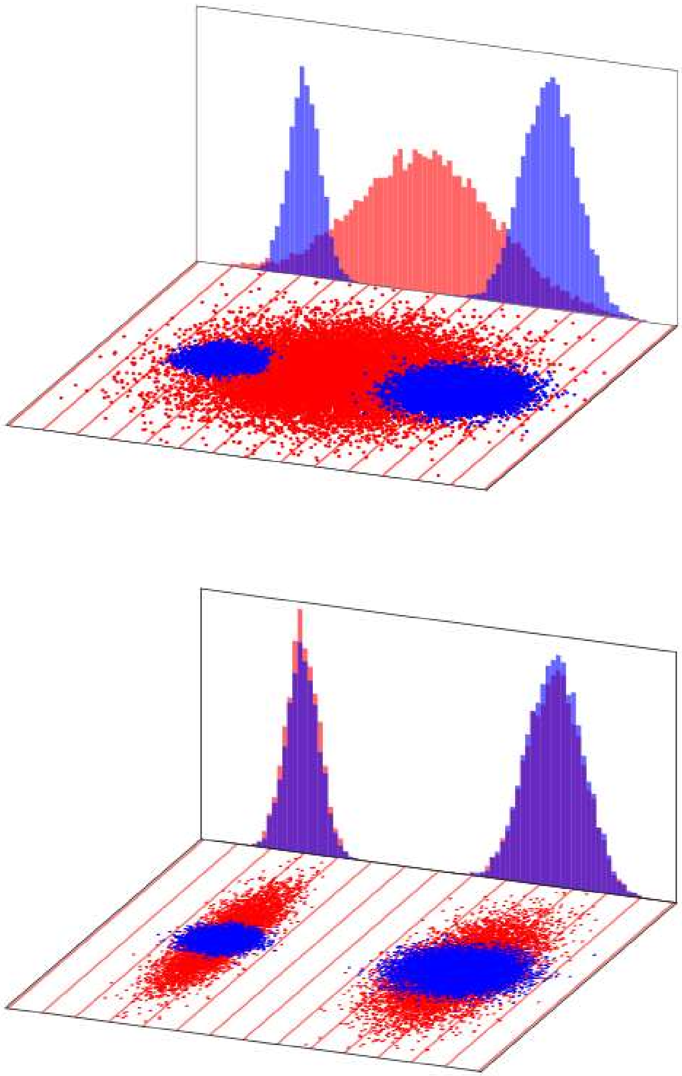
Sample-based population observer in a picture:



Good estimation in three iterations vs. thousands of iterations using ART!

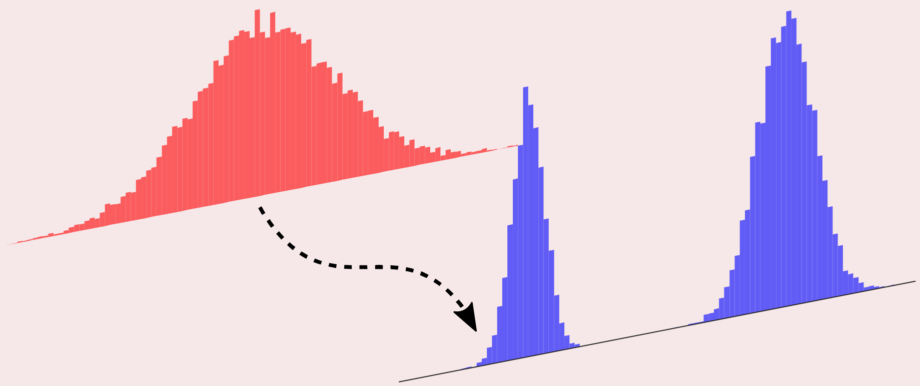
Sample-based population observers

S. Zeng. Sample-based Population Observers. *Automatica*, 2019.



Correction mechanism via Optimal Transport

On the level of the output, the transition from the red to the blue histogram



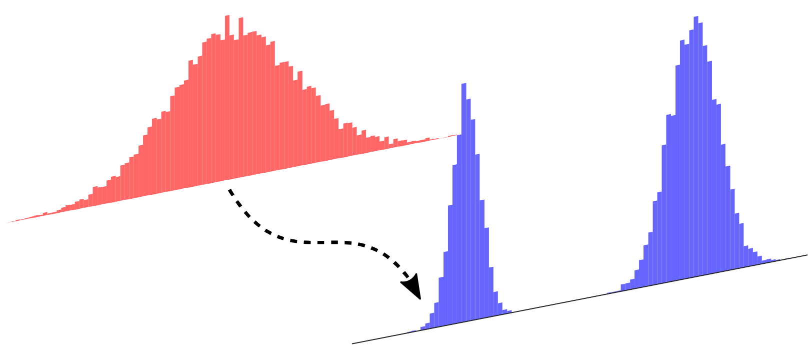
is achieved via a standard optimal transport formulation (described next).

Sample-based population observers

When the bins of the histograms are identical, we can use vectors

$$q = (q_1 \quad \dots \quad q_\ell)$$
$$\widehat{q} = (\widehat{q}_1 \quad \dots \quad \widehat{q}_\ell)$$

containing the sample counts as useful representations.



Redistribute mass in the different bins of $\widehat{q}$ so as to obtain the mass distribution as specified in q.

Classic version of optimal mass transport in finite-dimensional setting

Find *transport plan* $T \in \mathbb{R}_+^{\ell \times \ell}$: entry T_{ij} dictates how much of the "mass" $\widehat{q}_j$ in the jth bin is to be moved into the ith bin:

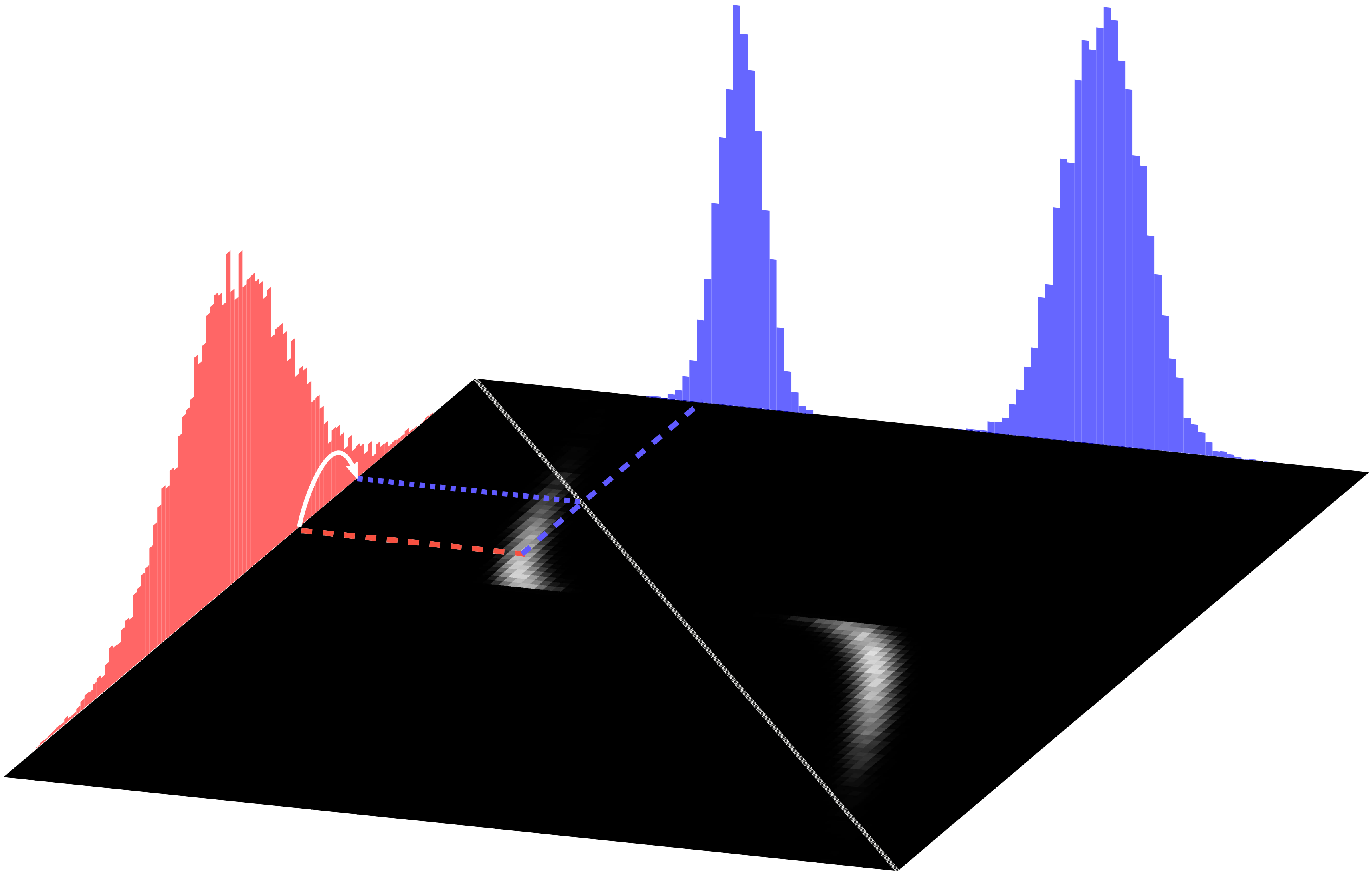
$$\text{minimize } J = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} |i - j| \, T_{ij}$$

$$\text{subject to } \sum_{i=1}^{\ell} T_{ij} = \widehat{q}_j, \; \sum_{j=1}^{\ell} T_{ij} = q_i.$$

The dual problem is a linear program involving only ℓ variables (Kantorovich-Rubenstein duality).

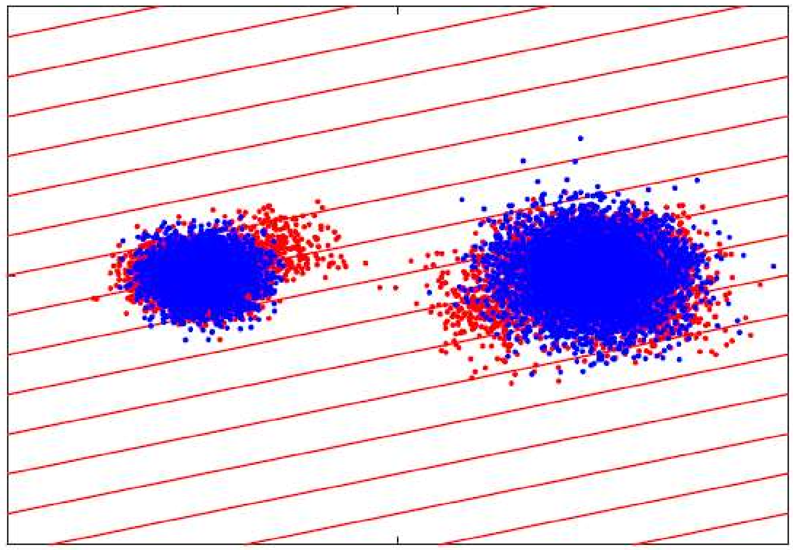
Sample-based population observers

Illustration of a transport plan:



Sample-based population observers

Sample-based population observer in a picture (cf. Kaczmarz method):



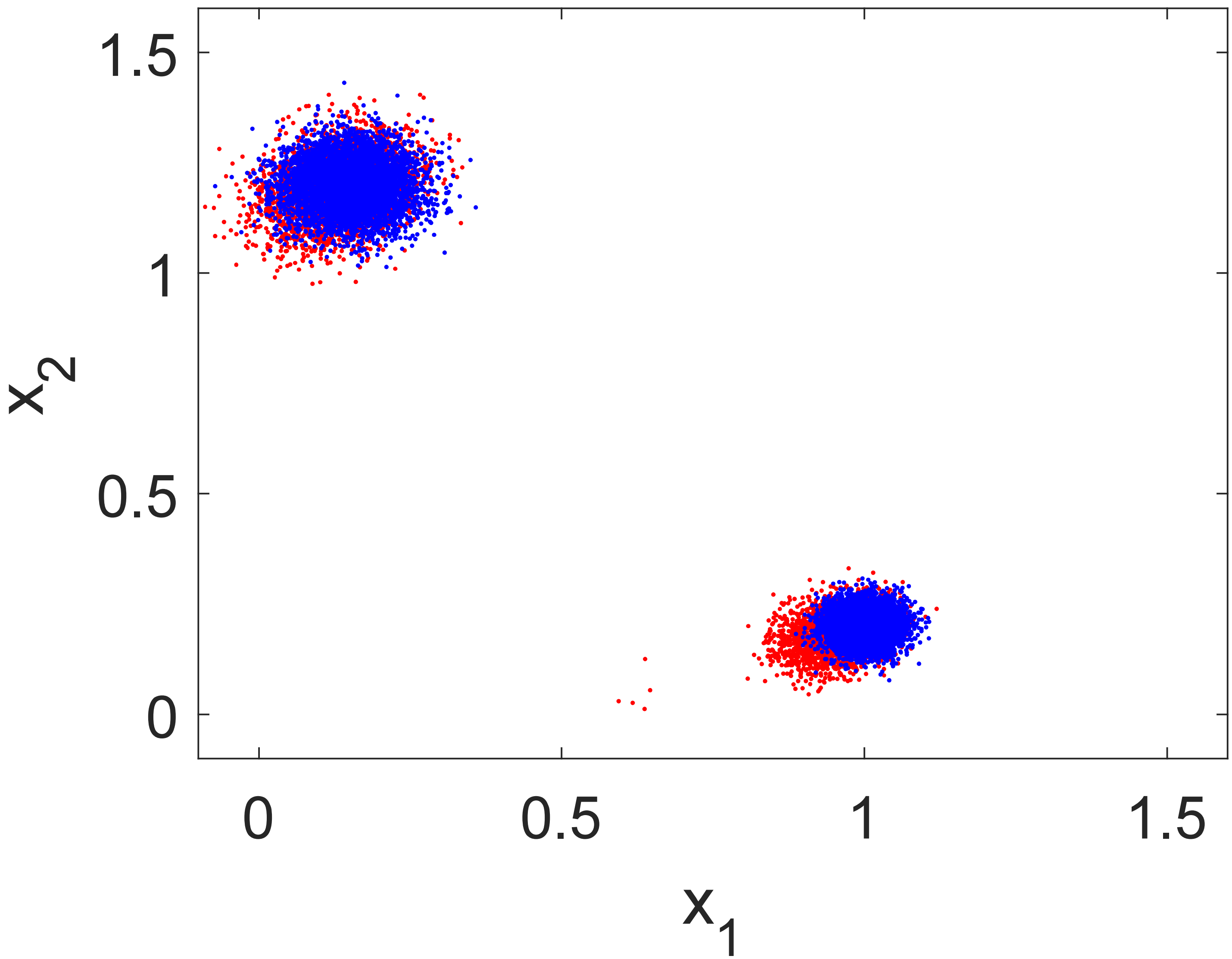
For $\tau \leq t$, we have $Ce^{A(\tau - t)}x(t) = Ce^{A\tau}e^{-At}x(t) = y(\tau)$.

Moving horizon sample-based population observer:

- Carry a batch of previously measured output distributions.
- Estimation of $\mathbb{P}_{x(t)}$ via mass transport-based correction with select directions $Ce^{A(\tau_\ell - t)}$ and output distributions $\mathbb{P}_{y(\tau_\ell)}$.

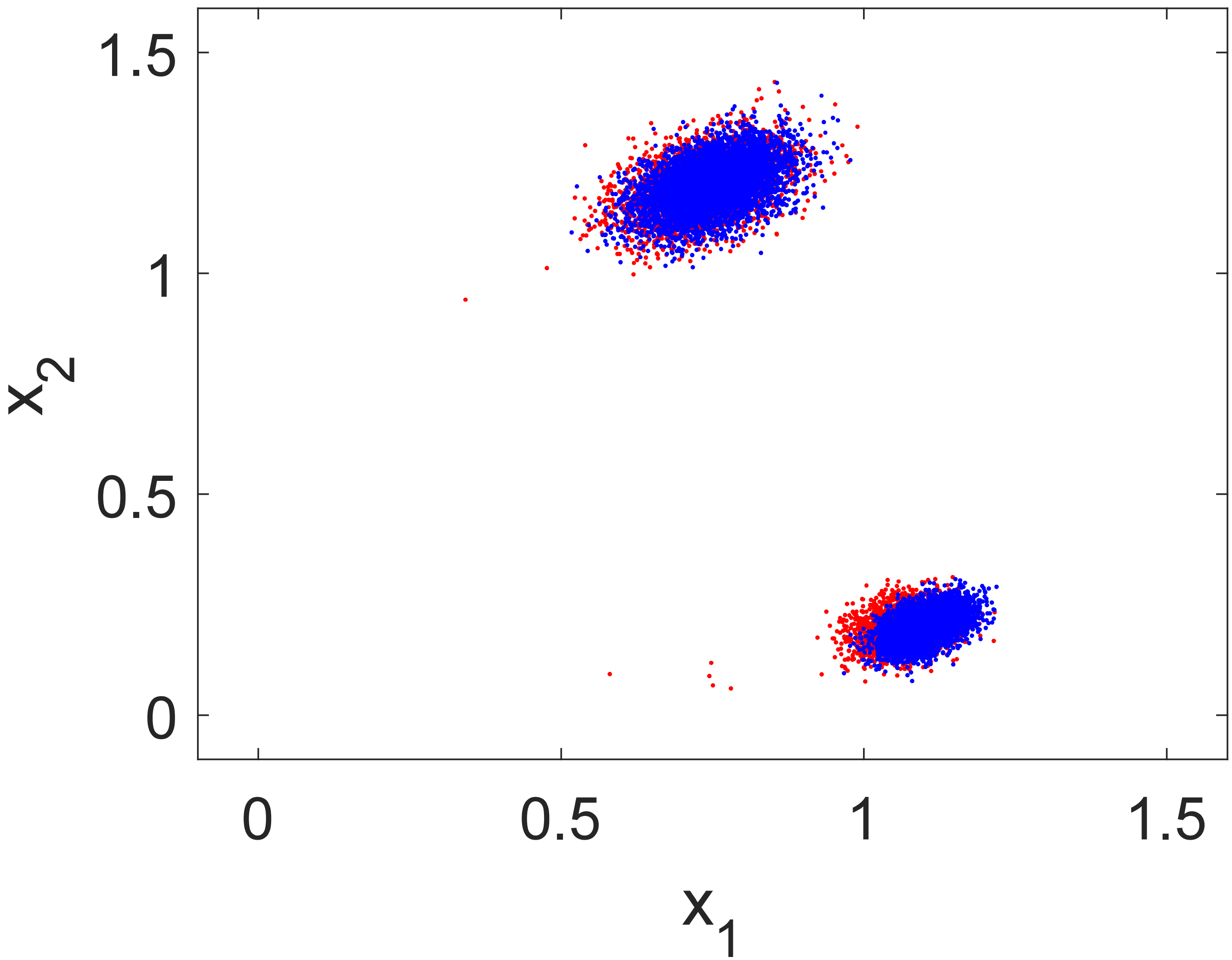
Sample-based population observers

Double integrator example:



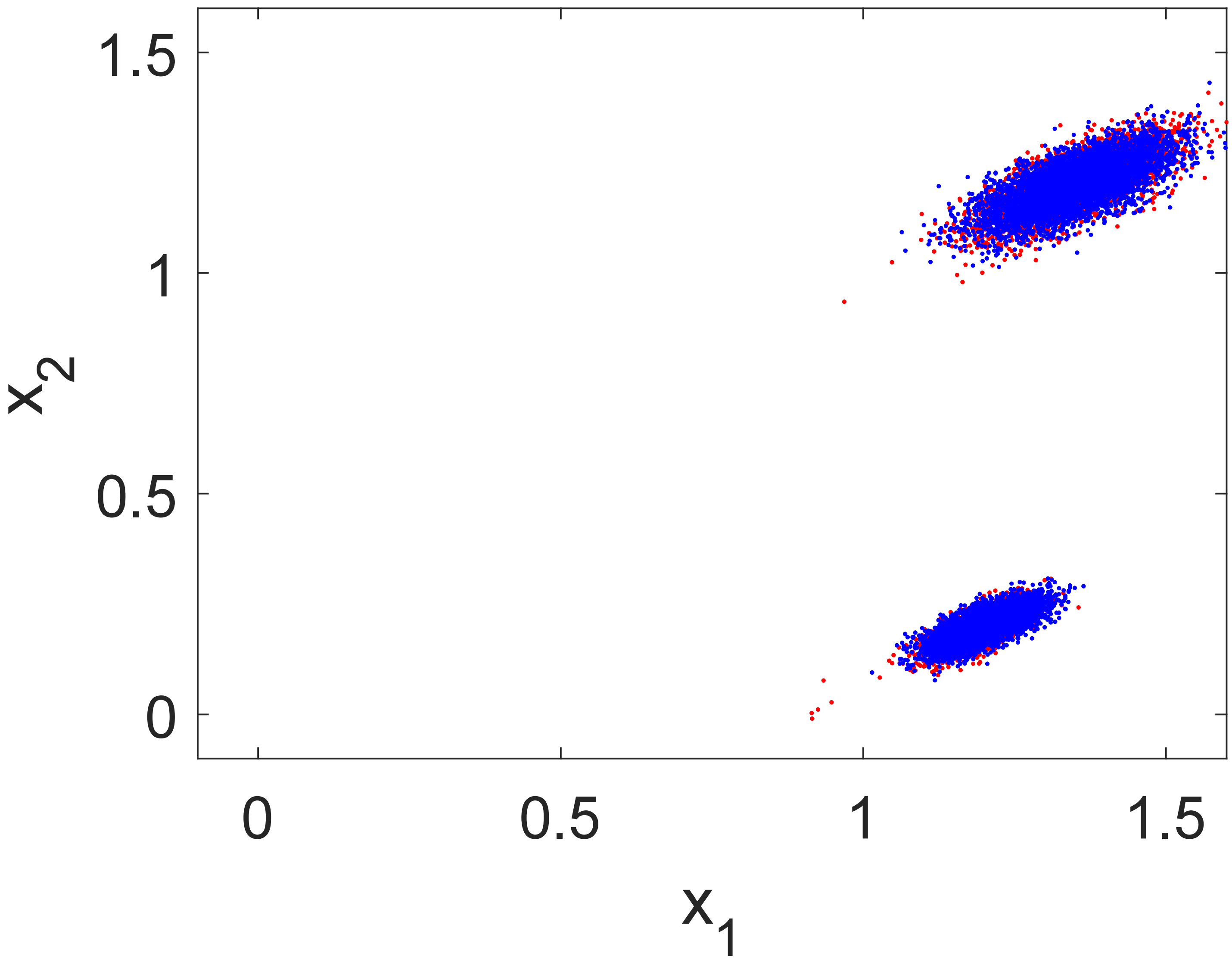
Sample-based population observers

Double integrator example:



Sample-based population observers

Double integrator example:



Summary

- Ensemble observability problem for dynamical systems:
  Inferring state distribution from corresponding output distributions.
- Basic systems theoretic concept related to population systems.
- First steps towards sample-based population observers made.