Nested Sampling Approach to Set-membership Estimation

R. Paulen*, L. Gomoescu**, B. Chachuat**

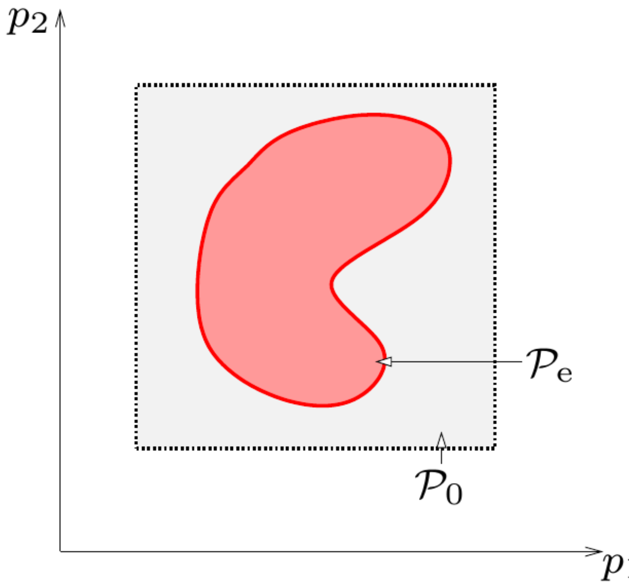
*Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Slovakia

** Centre for Process Systems Engineering, Department of Chemical Engineering

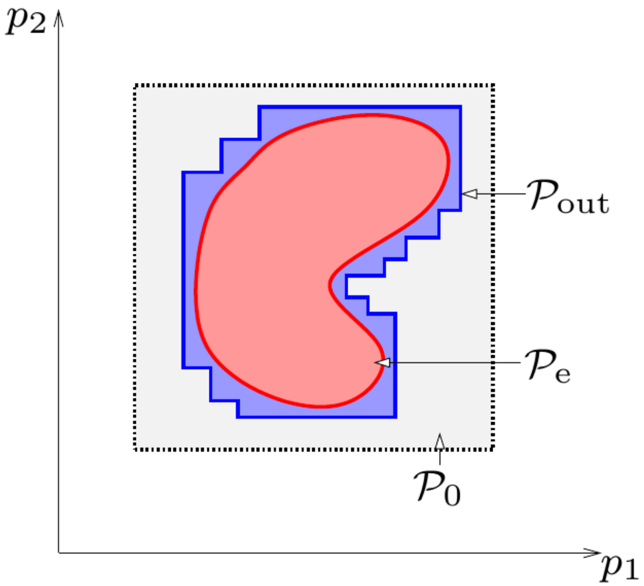
Imperial College London, United Kingdom

Motivation

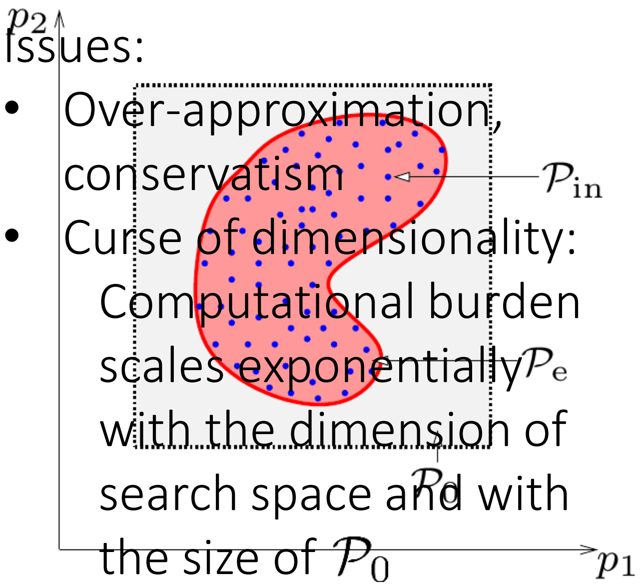
- Set-membership estimation is an alternative to statistical estimation
- It drops the need of knowing statistical distributions but requires set-based tools



$$\mathcal{P}_e := \{\boldsymbol{p} \in \mathcal{P}_0 \subset \mathbb{R}^n | \boldsymbol{h}(\dot{\boldsymbol{x}}, \boldsymbol{x}, \boldsymbol{p}) \leq 0\}$$



$$\mathcal{P}_e \subseteq \mathcal{P}_{out}$$



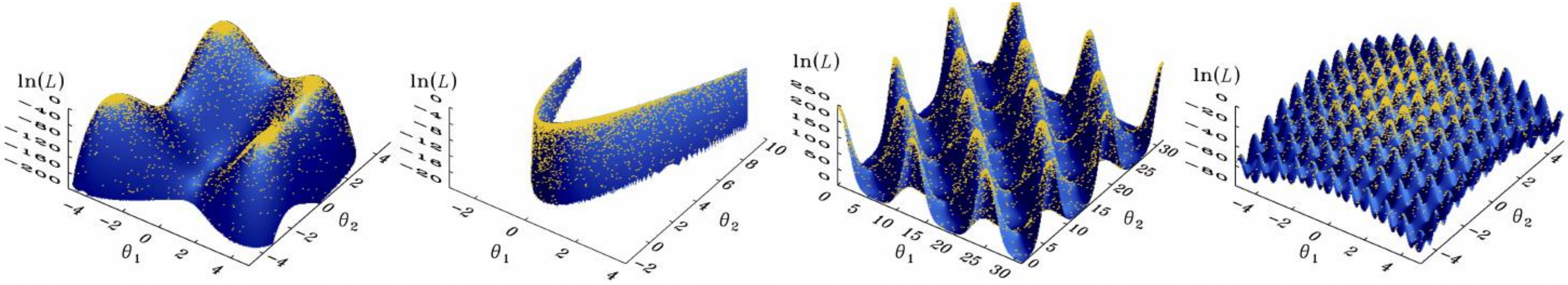
$$\mathcal{P}_{in} \subseteq \mathcal{P}_e$$

Issues:

- Over-approximation, conservatism
- Curse of dimensionality: Computational burden scales exponentially with the dimension of search space and with the size of $\mathcal{P}_0$

Motivation

- Sampling techniques well-developed for Bayesian statistics

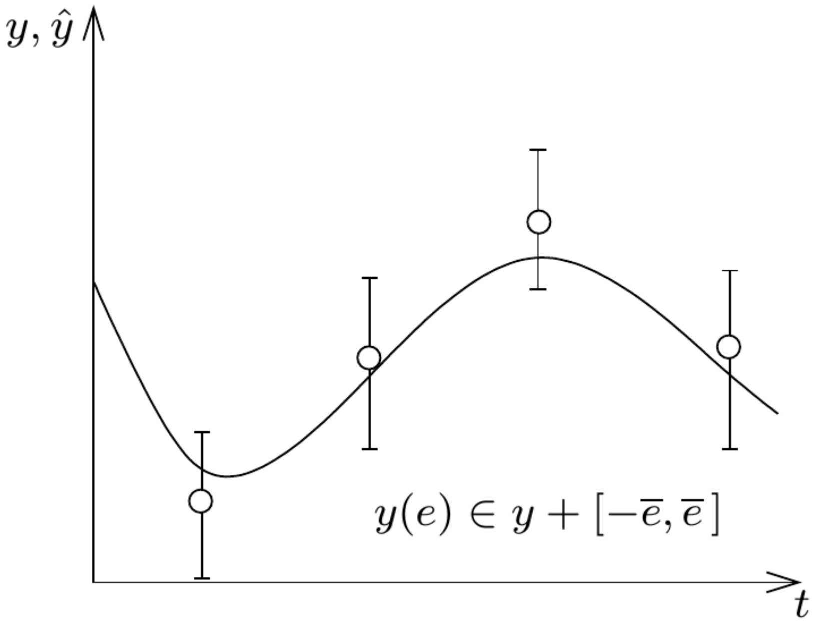


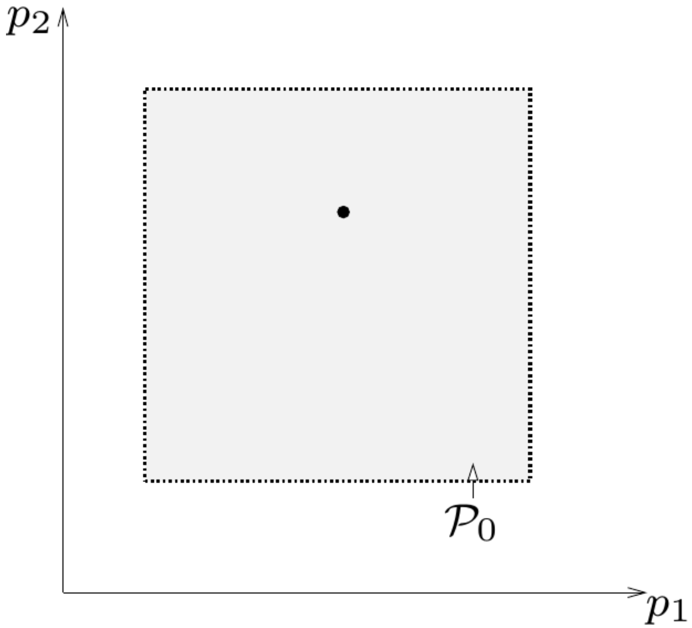
E. Corsaro and J. De Ridder (2014)

- This talk: A novel algorithm based on <u>nested sampling</u> that leverages efficient strategies from Bayesian estimation

Set-membership estimation (SME)

$$\text{model: } \hat{\boldsymbol{y}}(t_i) = \boldsymbol{F}(\boldsymbol{p}, t_i) \begin{cases} \text{dynamic system:} & \dot{\boldsymbol{x}}(t_i) = \boldsymbol{f}(\boldsymbol{x}(t_i), \boldsymbol{p}),\ \boldsymbol{x}(0) = \boldsymbol{x}_0(\boldsymbol{p}) \\ \text{output equation:} & \hat{\boldsymbol{y}}(t_i) = \boldsymbol{g}(\boldsymbol{x}(t_i), \boldsymbol{p}) \end{cases}$$





$$\mathcal{P}_\text{e} := \{\boldsymbol{p} \in \mathcal{P}_0 \mid -\bar{\boldsymbol{e}} \le \boldsymbol{y}(t_i) - \boldsymbol{F}(\boldsymbol{p}, t_i) \le \bar{\boldsymbol{e}}, \forall i \in \{1, \ldots, N\}\} := \{\boldsymbol{p} \in \mathcal{P}_0 | \boldsymbol{h}(\dot{\boldsymbol{x}}, \boldsymbol{x}, \boldsymbol{p}) \le 0\}$$

Nested sampling (NS) (Skilling, 2006)

$$\begin{array}{ccccccc} \Pr(\boldsymbol{y}|\boldsymbol{p}) & \times & \Pr(\boldsymbol{p}) & = & \Pr(\boldsymbol{y}) & \times & \Pr(\boldsymbol{p}|\boldsymbol{y}) \\ \text{Likelihood } \mathcal{L}(\boldsymbol{p}) & \times & \text{Prior } \pi(\boldsymbol{p}) & = & \text{Evidence } \mathcal{Z} & \times & \text{Posterior } \Pi(\boldsymbol{p}) \end{array}$$

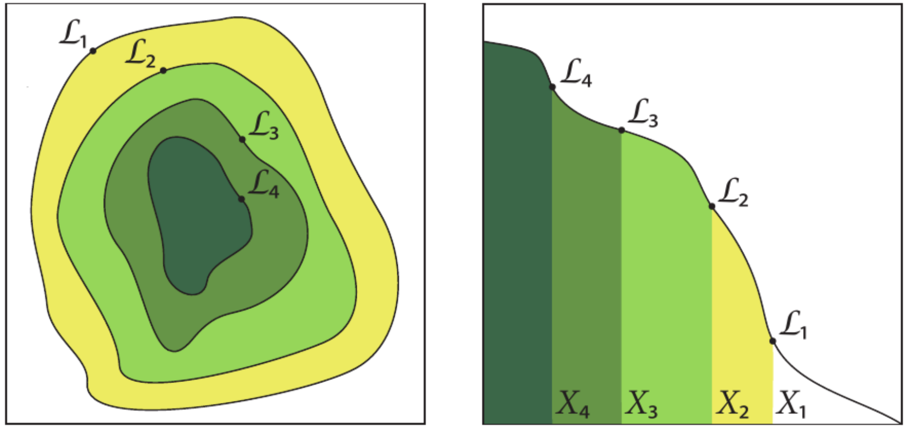
- $\mathbb{E}\,(\text{Posterior}) \propto \mathbb{E}\,(\text{Likelihood} \times \text{Prior}) = \int_{\boldsymbol{p}} \mathcal{L}(\boldsymbol{p})\pi(\boldsymbol{p})\mathrm{d}\boldsymbol{p}$ is hard to calculate.
- It is easier to get the probability mass contained within the level sets of $\mathcal{L}(\boldsymbol{p})$

$$X(\lambda) = \int_{\mathcal{L}(\boldsymbol{p})\geq\lambda} \pi(\boldsymbol{p})\mathrm{d}\boldsymbol{p}$$

- ...and calculate the evidence in 1D.

$$\mathcal{Z} = \int_0^1 \mathcal{L}(\boldsymbol{p})\mathrm{d}X$$

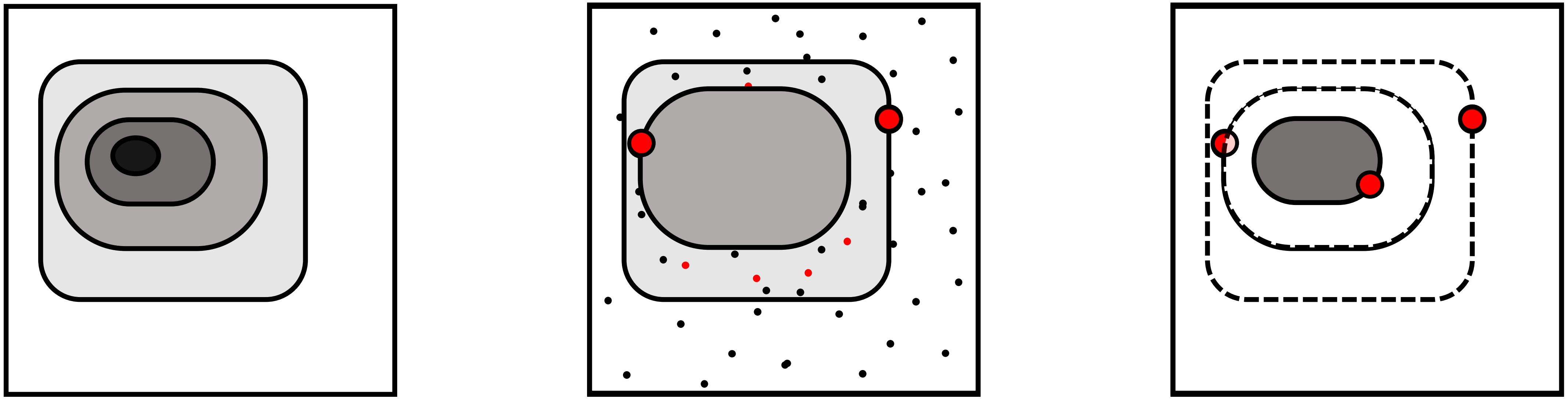
- The posterior is a free by-product.



Feroz et al. (2013)

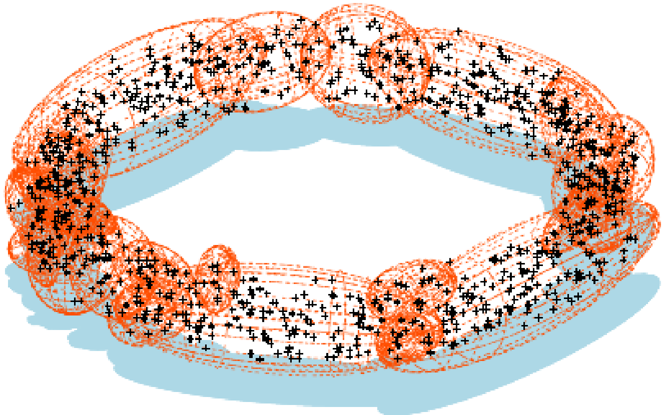
Nested sampling (NS) (Skilling, 2006)

- Initialization: Draw N_{live} samples (live points) from the prior in a list $\boldsymbol{L}$
- Iterations:
  1. Find the lowest value of $\mathcal{L}(\boldsymbol{p})$ within the list $\boldsymbol{L}$ and remove the point from the list
  2. Draw N_{prop} samples (proposals) from the prior within a region enclosing $\mathcal{L}(\boldsymbol{p})$
  3. Include them in the list $\boldsymbol{L}$ if their likelihood value is higher than the current lowest one



Nested sampling (NS) (Skilling, 2006)

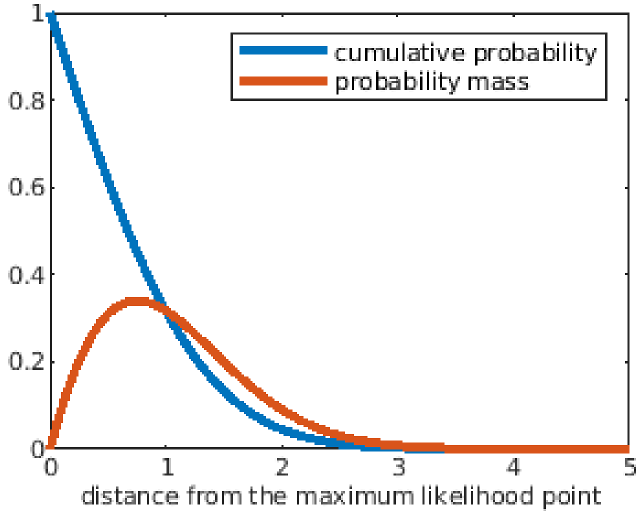
- Effective sampling from the prior is necessary for multimodal distributions (disconnected sets)
- E.g. X-means clustering algorithm (Pelleg and Moore, 2000)



Feroz et al. (2009)

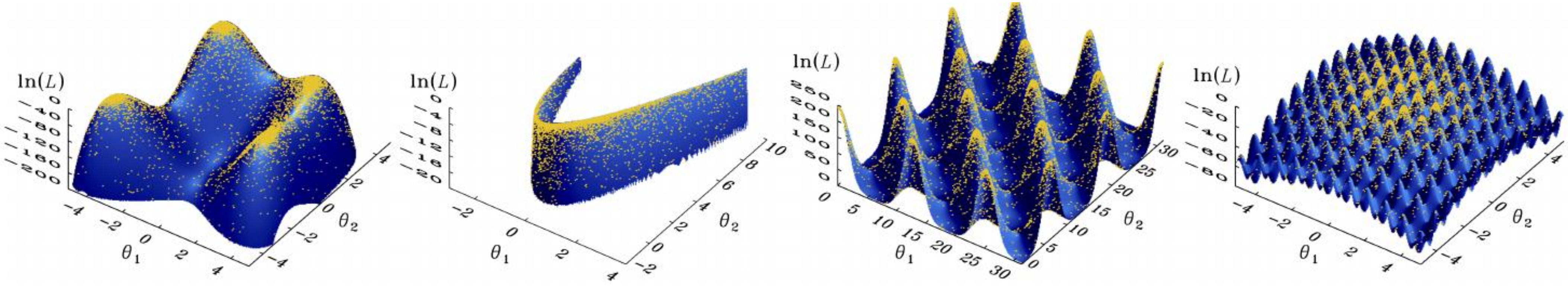
Stopping criterion:

- The algorithm estimates contribution of the enclosed probability mass
- Stop once the contribution to the mass from the live points is below threshold



Nested sampling (NS) – Features

- NS samples more sparsely from the prior in regions where the likelihood is low and more densely where the likelihood is high.
- The algorithm can be applied to multi-modal likelihood functions.
- The likelihood function evaluations can be easily parallelized.
- The procedure runs with an evolving collection of N_{live} points, where N_{live} can be chosen small for speed or large for accuracy.



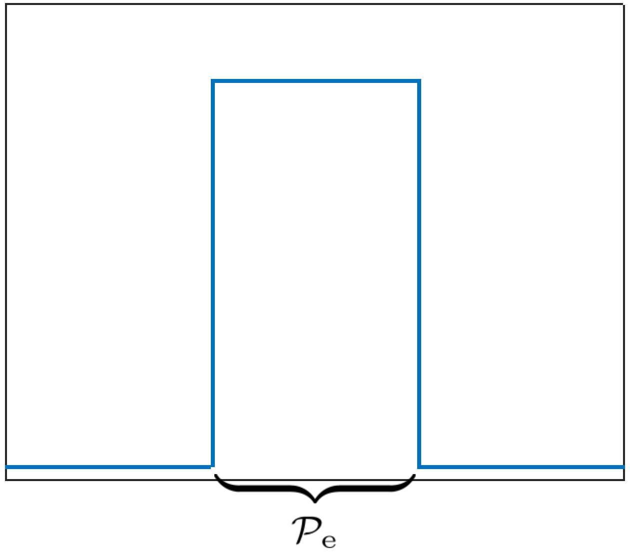
E. Corsaro and J. De Ridder (2014)

NS for Set-membership estimation (SME)

How to transform SME into Bayesian-like setup to apply NS?

Naïve way:

$$\pi(\boldsymbol{p}) := \begin{cases} 1, & \text{if } \boldsymbol{p} \in \mathcal{P}_0 \\ 0, & \text{otherwise} \end{cases} \qquad \mathcal{L}(\boldsymbol{p}) := \begin{cases} 1, & \text{if } \boldsymbol{p} \in \mathcal{P}_\mathrm{e} \\ 0, & \text{otherwise} \end{cases}$$



Proposed way:

$$\mathcal{P}_\mathrm{e} := \{\boldsymbol{p} \in \mathcal{P}_0 \,|\, -\overline{\boldsymbol{e}} \leq \boldsymbol{y}(t_i) - \boldsymbol{F}(\boldsymbol{p}, t_i) \leq \overline{\boldsymbol{e}}, \forall i \in \{1, \ldots, N\}\}$$

$$\mathcal{L}(\boldsymbol{p}) := \begin{cases} 1 & \text{if } \boldsymbol{p} \in \mathcal{P}_\mathrm{e} \\ \prod_{i=1}^{N} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{y}(t_i) - \boldsymbol{F}(\boldsymbol{p},t_i))^T \boldsymbol{Q} (\boldsymbol{y}(t_i) - \boldsymbol{F}(\boldsymbol{p},t_i))} & \text{otherwise} \end{cases}$$

$$Q := \mathrm{diag}^{-2}\left(\frac{1}{3}\overline{\boldsymbol{e}}\right)$$

DEUS 1.0.0 available

- **DE**sign under **U**ncertainty using **S**ampling methods
- A Python package that implements Nested sampling algorithm
  - Bayesian estimation
  - Set-membership estimation
  - Design space characterization
- Available at: https://github.com/omega-icl/DEUS
- Input files for all the presented case studies can be retrieved

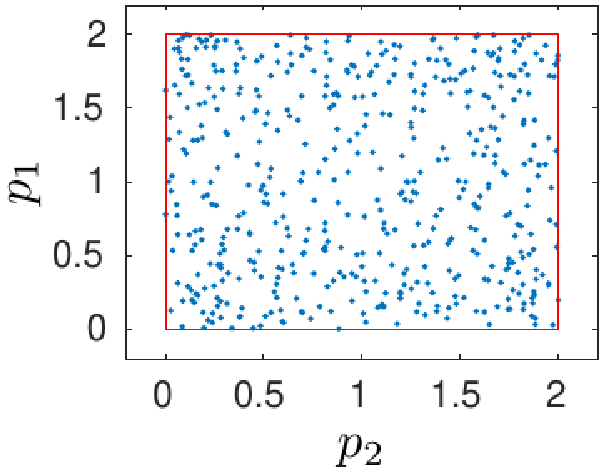
Sampling of an n-dimensional box

$$\mathcal{P}_\mathrm{e} := \{\boldsymbol{p} \in [-10^m, 10^m]^n \,|\, 0 \leq p_i \leq 2, \forall i \in \{0, \ldots, n\}\} \qquad n \in \{2, \ldots, 10\},\ m \in \{1, \ldots, 5\}$$

For $n = 2, m = 1$: Solution in less than 1s (naïve way of sampling takes 63s).

For $n = 10$:

- $m = 1$: 106s
- $m = 5$: 1100s

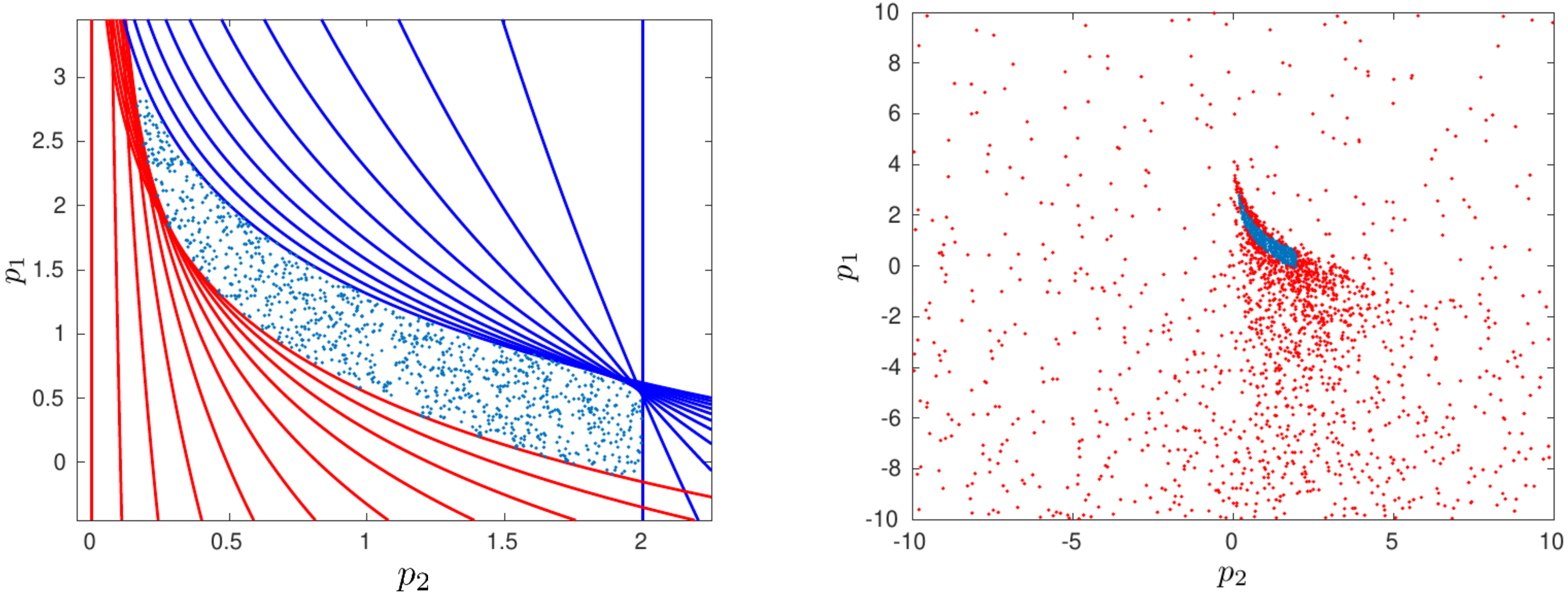


(a) $N_\mathrm{live} := 150$

The computational burden scales <u>quadratically</u> with the size of the search space n and <u>linearly</u> with the volume of initial search domain $\mathcal{P}_0$.

Simple nonlinear static problem

$$\mathcal{P}_\mathrm{e} := \left\{\boldsymbol{p} \in [-10,10]^2 \mid -1 \le \underbrace{p_1^* \mathrm{e}^{p_2^* t}}_{y} - \underbrace{p_1 \mathrm{e}^{p_2 t}}_{\hat{y}} \le 1, \forall t \in \{0, 0.1, \ldots, 1\}\right\} \qquad \boldsymbol{p}^* := (p_1^*, p_2^*)^T = (1,1)$$

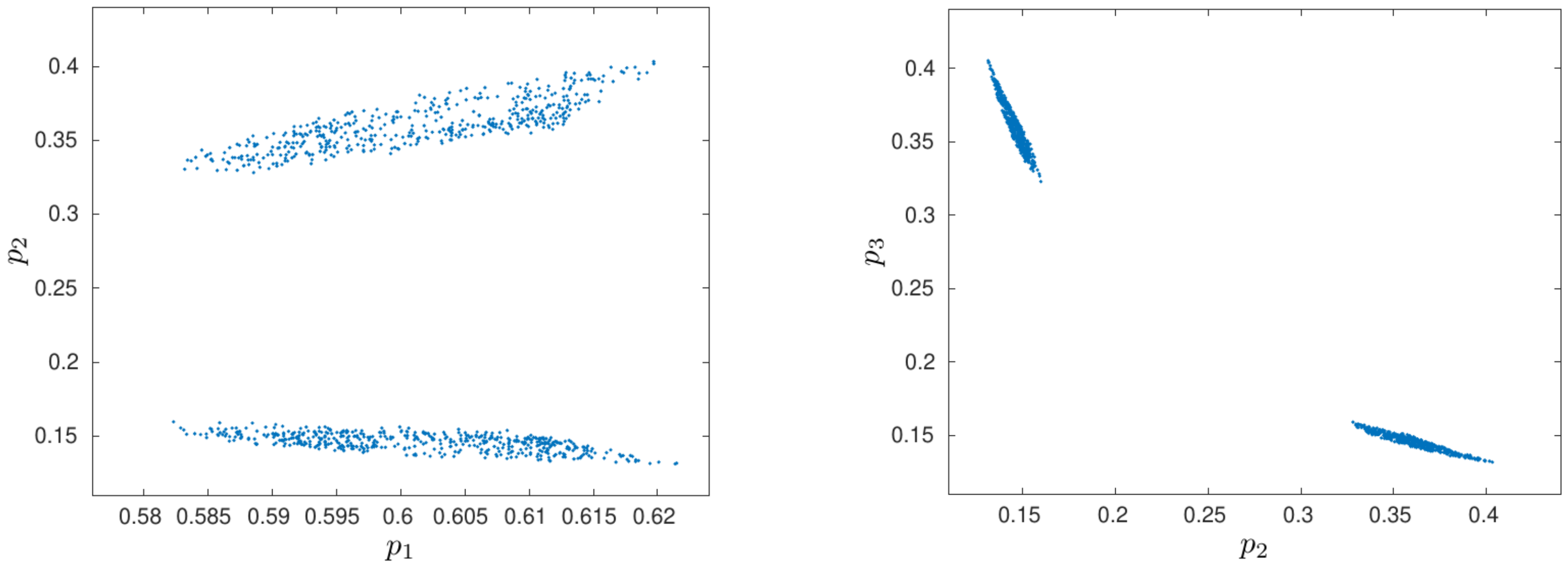


Solution in 3 seconds (no parallelization) after 7,200 function evaluations.

Jaulin, L. and Walter, E. (1993). Set inversion via interval analysis for nonlinear bounded-error estimation. Automatica, 29(4).

Three-parameter dynamic estimation

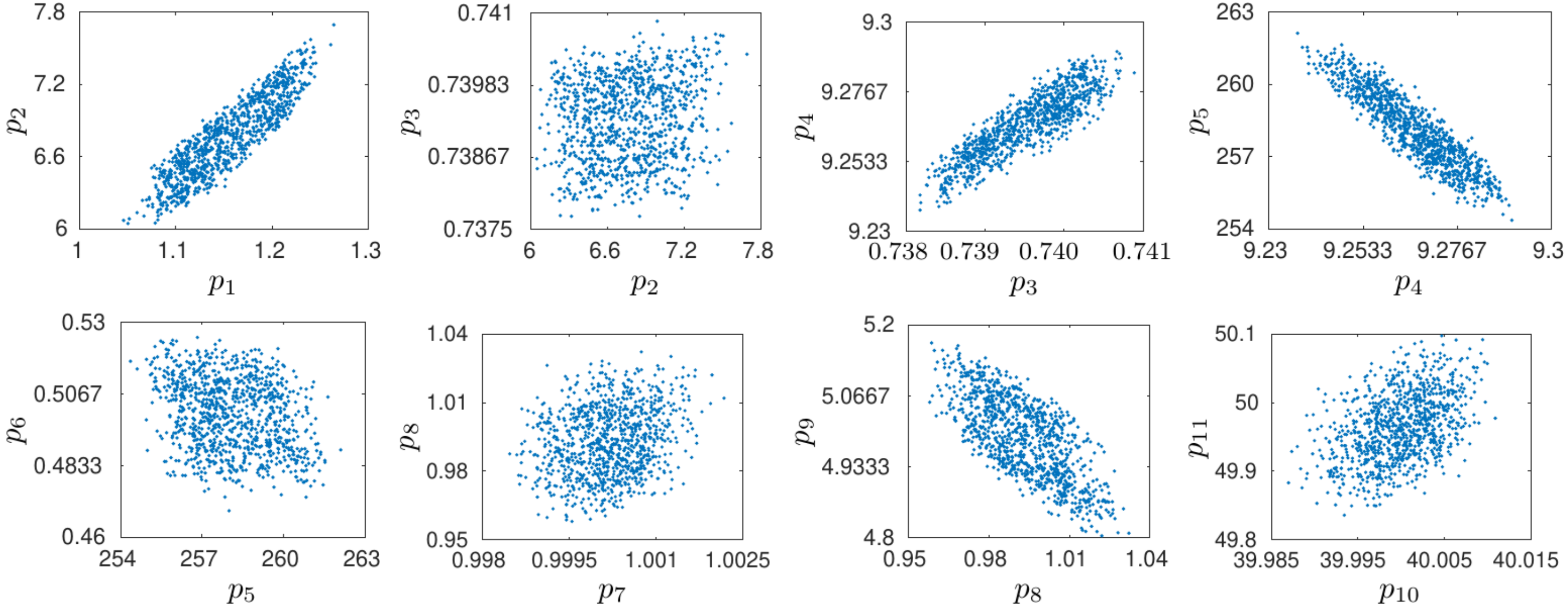
$$\text{model: } \hat{y}(t_i) = F(\boldsymbol{p}, t_i) \begin{cases} \text{dynamic system:} & \dot{x}_1(t) = -(p_1 + p_3)x_1(t) + p_2 x_2(t),\ x_1(0) = 1 \\ & \dot{x}_2(t) = p_1 x_1(t) - p_2 x_2(t),\ x_2(0) = 0 \\ \text{output equation:} & \hat{y}(t_i) = x_2(t_i),\ \forall t_i \in \{1, \ldots, 15\} \end{cases}$$



Kieffer, M. and Walter, E. (2011). International Journal of Adaptive Control & Signal Processing.

State/parameter dynamic estimation

- Anaerobic digester model with 6 states and 3 measured outputs
- 5 kinetic parameters and 6 initial conditions to estimate



- Inner approximation obtained in 1 hour of algorithm run
- Obtaining an outer-approximation is infeasible with state of the art tools

Conclusions

- A novel technique based on Bayesian nested sampling for inner-approximation of the solution set of set-membership estimation.
- Favourable computational performance obtained.
- Inner-approximation cannot replace a validated outer-approximation but it is helpful in certain specific tasks:
  - scenario-based optimization
  - identifiability analysis
  - a priori analysis of a set-membership problems

Questions:
radoslav.paulen@stuba.sk

- Acknowledgements: MSCA IF (Grant no. 790017), ITN (Grant no. 675585), ITMS 26240220084, APVV (Grant no. 15-0007)