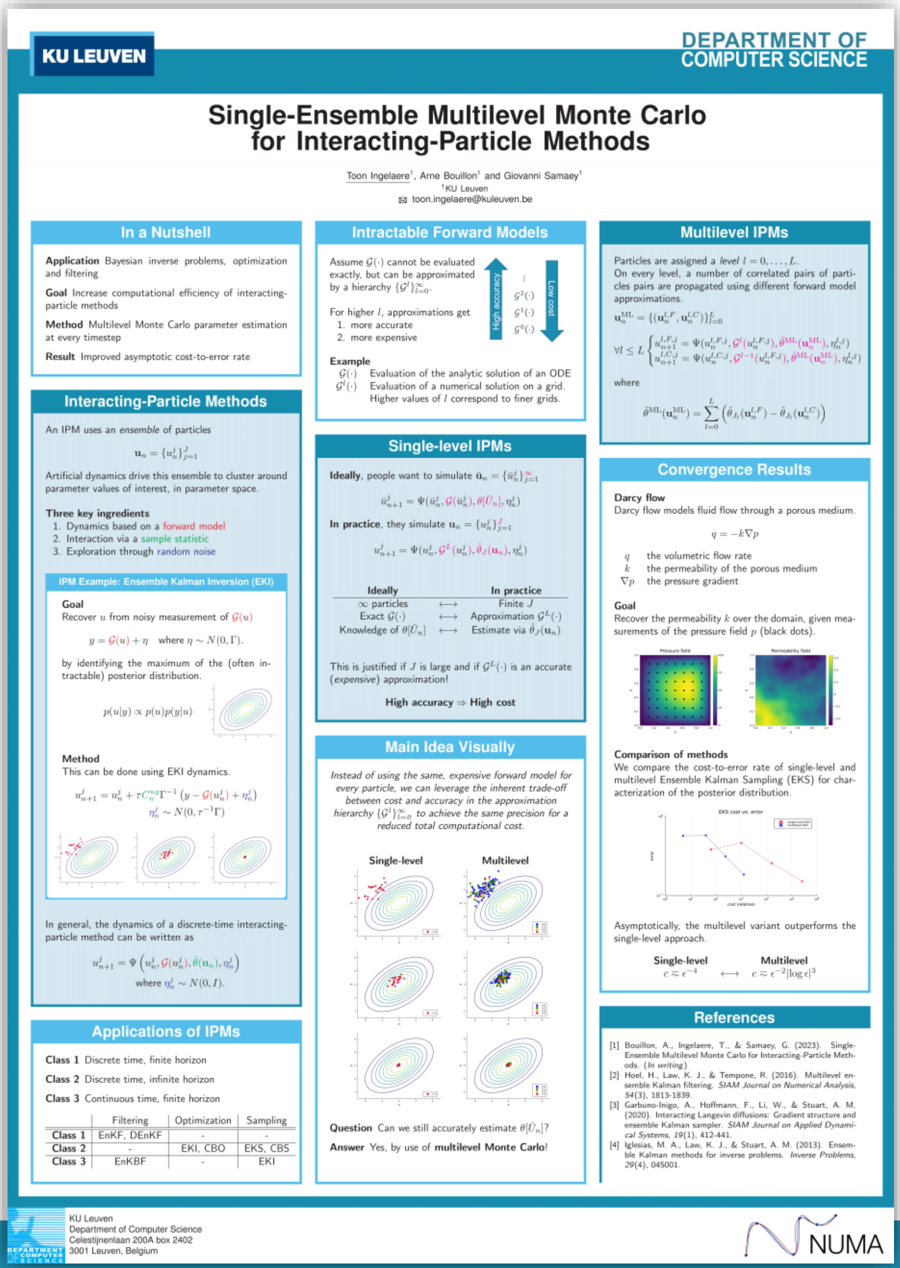
KU LEUVEN

Multilevel Interacting Particle Methods for Bayesian Inversion

Toon Ingelaere, A. Bouillon, G. Samaey

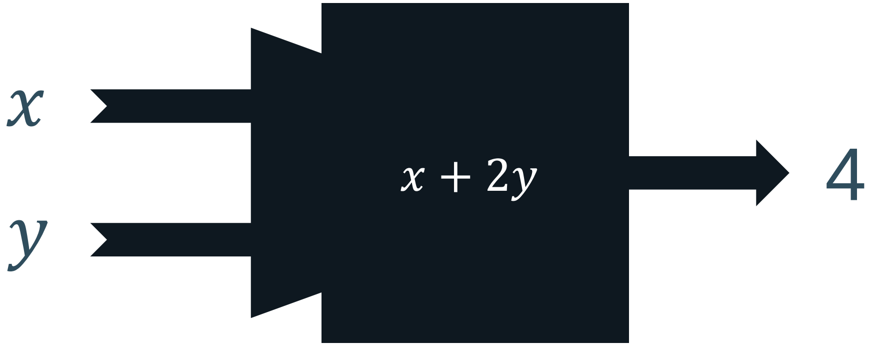

KU LEUVEN
DEPARTMENT OF COMPUTER SCIENCE
Single-Ensemble Multilevel Monte Carlo for Interacting-Particle Methods
Toon Ingelaere[1], Arne Bouillon[1] and Giovanni Samaey[1]
[1]KU Leuven
toon.ingelaere@kuleuven.be
In a Nutshell
Application Bayesian inverse problems, optimization and filtering
Goal Increase computational efficiency of interacting-particle methods
Method Multilevel Monte Carlo parameter estimation at every timestep
Result Improved asymptotic cost-to-error rate
Interacting-Particle Methods
An IPM uses an ensemble of particles
Artificial dynamics drive this ensemble to cluster around parameter values of interest, in parameter space.
Three key ingredients
1. Dynamics based on a forward model
2. Interaction via a sample statistic
3. Exploration through random noise
IPM Example: Ensemble Kalman Inversion (EKI)
Goal
Recover u from noisy measurement of G(u)
by identifying the maximum of the (often intractable) posterior distribution.
Method
This can be done using EKI dynamics.
In general, the dynamics of a discrete-time interacting-particle method can be written as
Applications of IPMs
Class 1 Discrete time, finite horizon
Class 2 Discrete time, infinite horizon
Class 3 Continuous time, finite horizon
Filtering
Optimization
Sampling
EnKF, DEnKF
EKI, CBO
EKS, CBS
EnKBF
EKI
Intractable Forward Models
High accuracy
Low cost
For higher l, approximations get
1. more accurate
2. more expensive
Example
Evaluation of the analytic solution of an ODE
Evaluation of a numerical solution on a grid.
Higher values of l correspond to finer grids.
Single-level IPMs
Ideally
In practice
High accuracy ⇒ High cost
Main Idea Visually
Single-level
Multilevel
Question Can we still accurately estimate
Answer Yes, by use of multilevel Monte Carlo!
Multilevel IPMs
Convergence Results
Darcy flow
Darcy flow models fluid flow through a porous medium.
the volumetric flow rate
the permeability of the porous medium
the pressure gradient
Goal
Comparison of methods
Asymptotically, the multilevel variant outperforms the single-level approach.
References
NUMA
KU Leuven
Department of Computer Science
Celestijnenlaan 200A box 2402
3001 Leuven, Belgium

KU LEUVEN

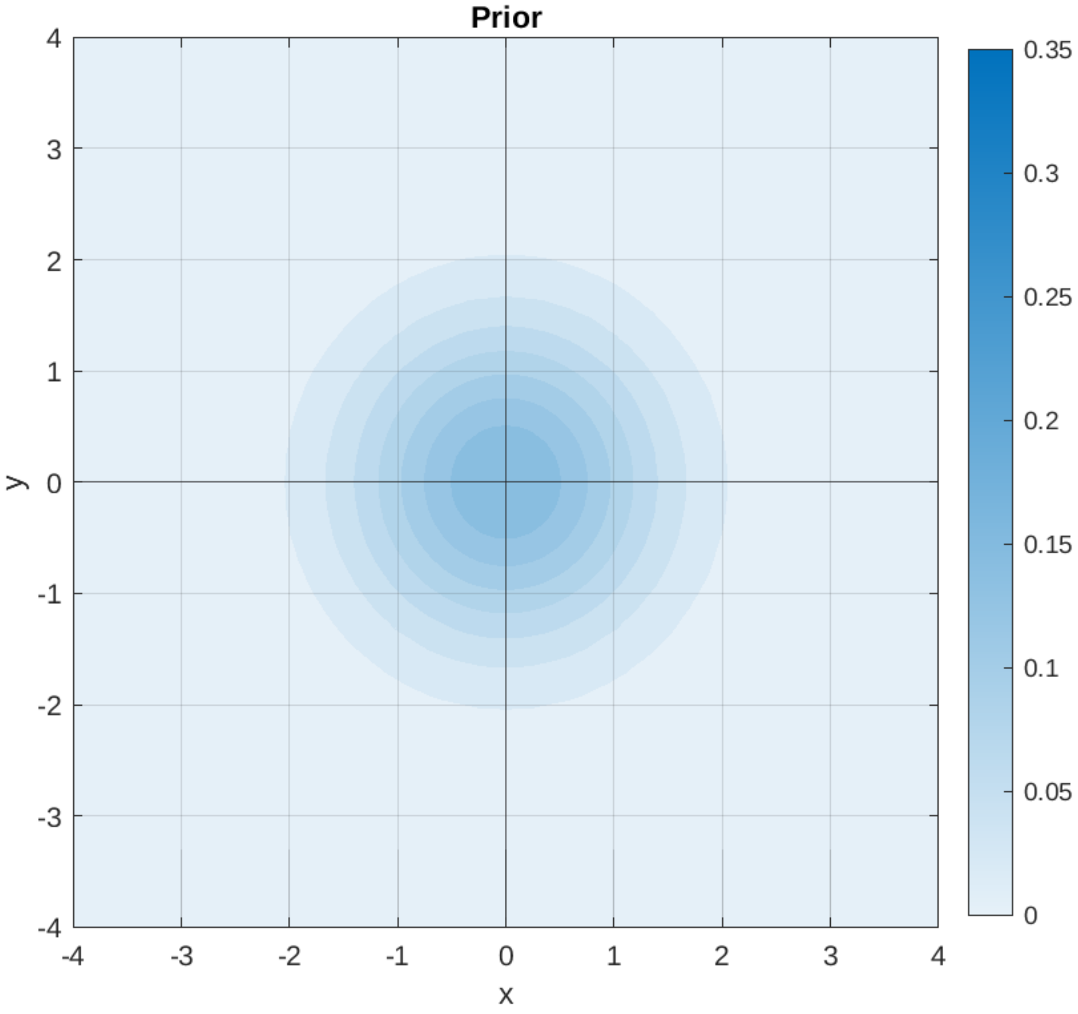
Chapter 1:
Inverse Problems

Let’s play a game

- I sample a point from a distribution.
- I feed it into *The Machine.*

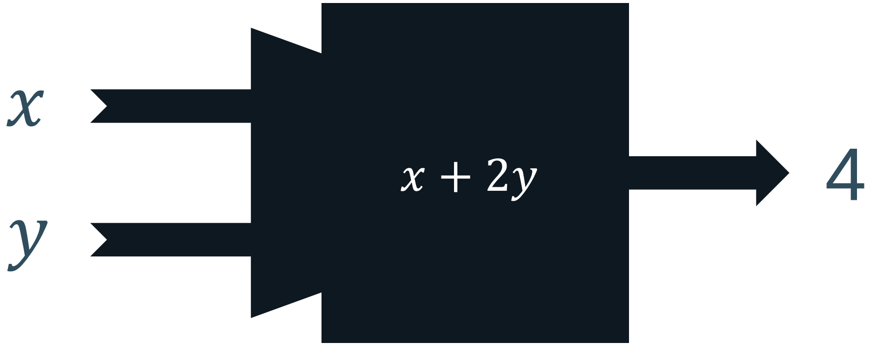


- The machine outputs* 4.
- Which point did I start with?

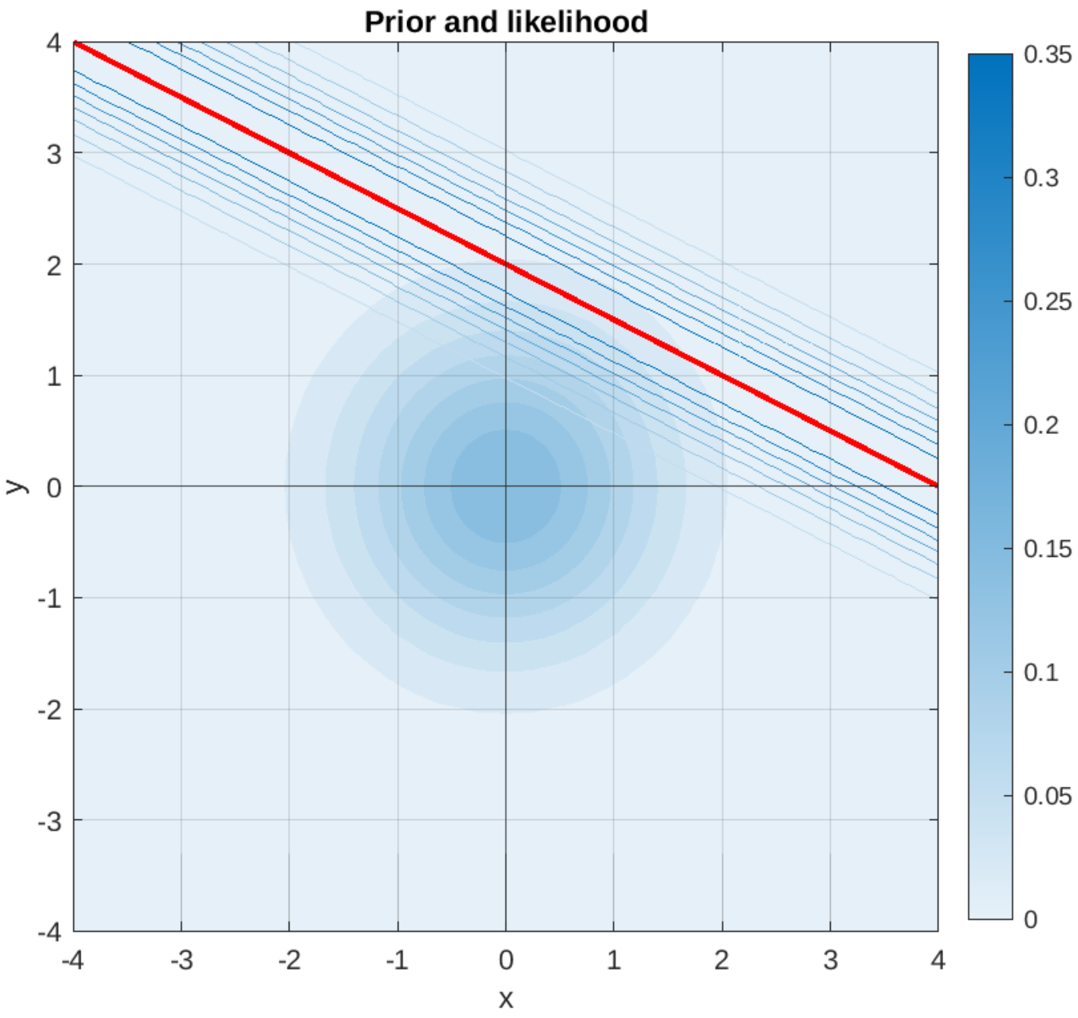


Let’s play a game

- I sample a point from a distribution.
- I feed it into *The Machine.*

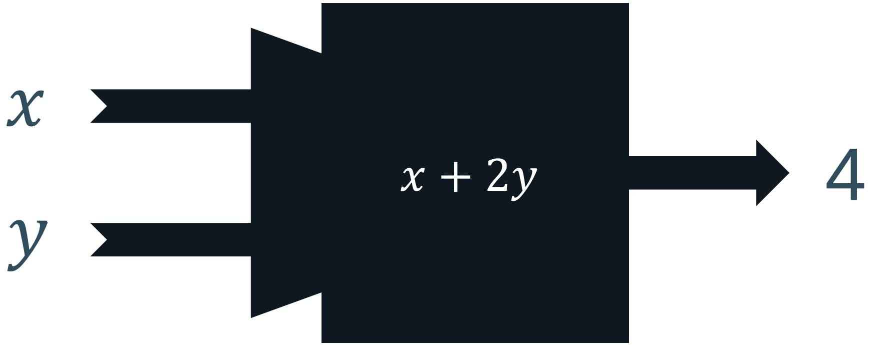


- The machine outputs* 4.
- Which point did I start with?

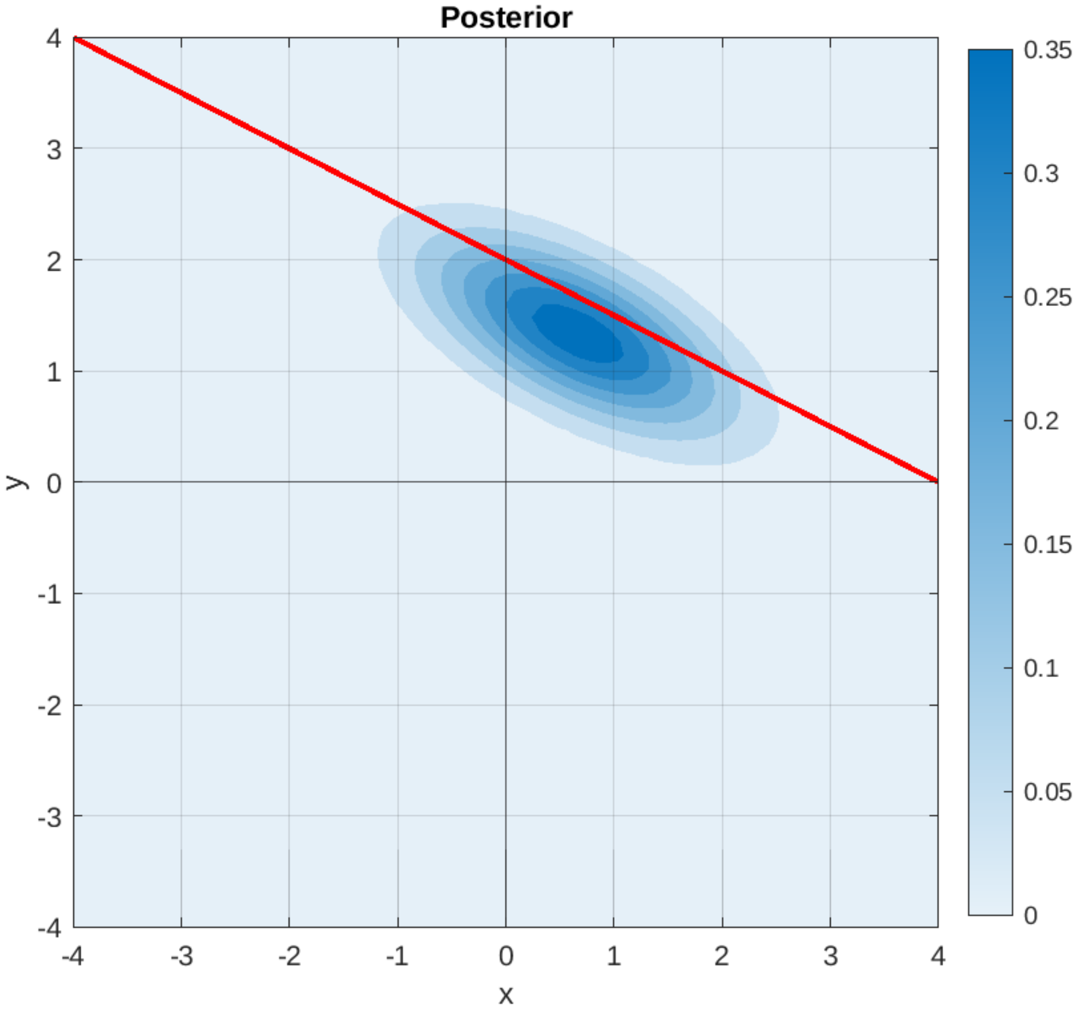


Let’s play a game

- I sample a point from a distribution.
- I feed it into *The Machine.*

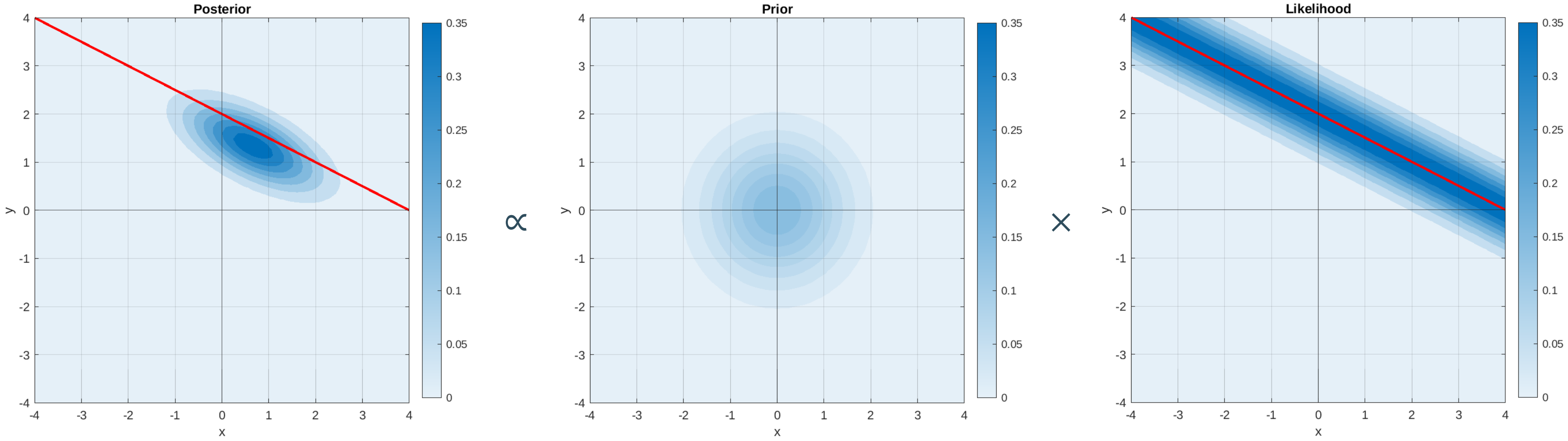


- The machine outputs* 4.
- Which point did I start with?



We can make this rigourous

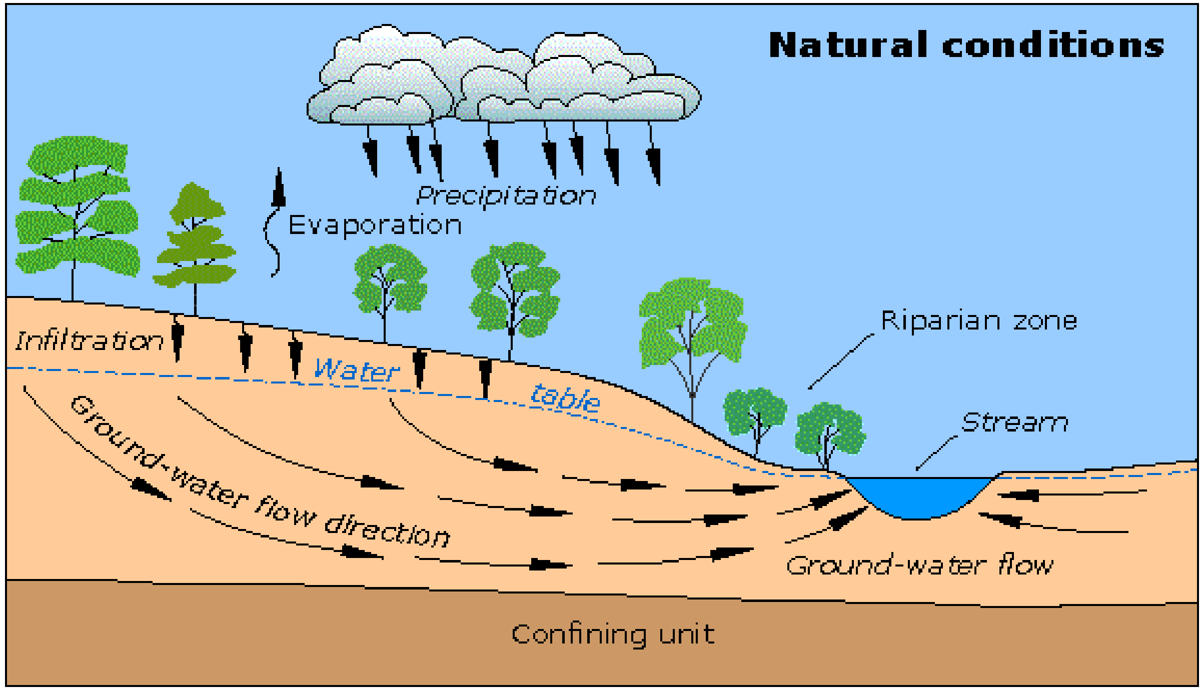
Bayes' rule: $\boldsymbol{p(\theta|d)} \propto p(\theta) \cdot p(d|\theta)$



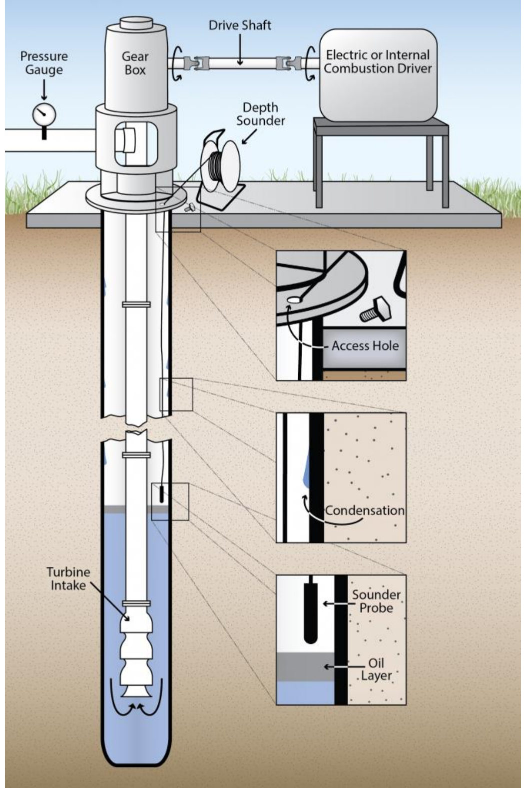
KU LEUVEN

Chapter 1.1:
A more interesting Inverse Problem

Let’s do something more interesting

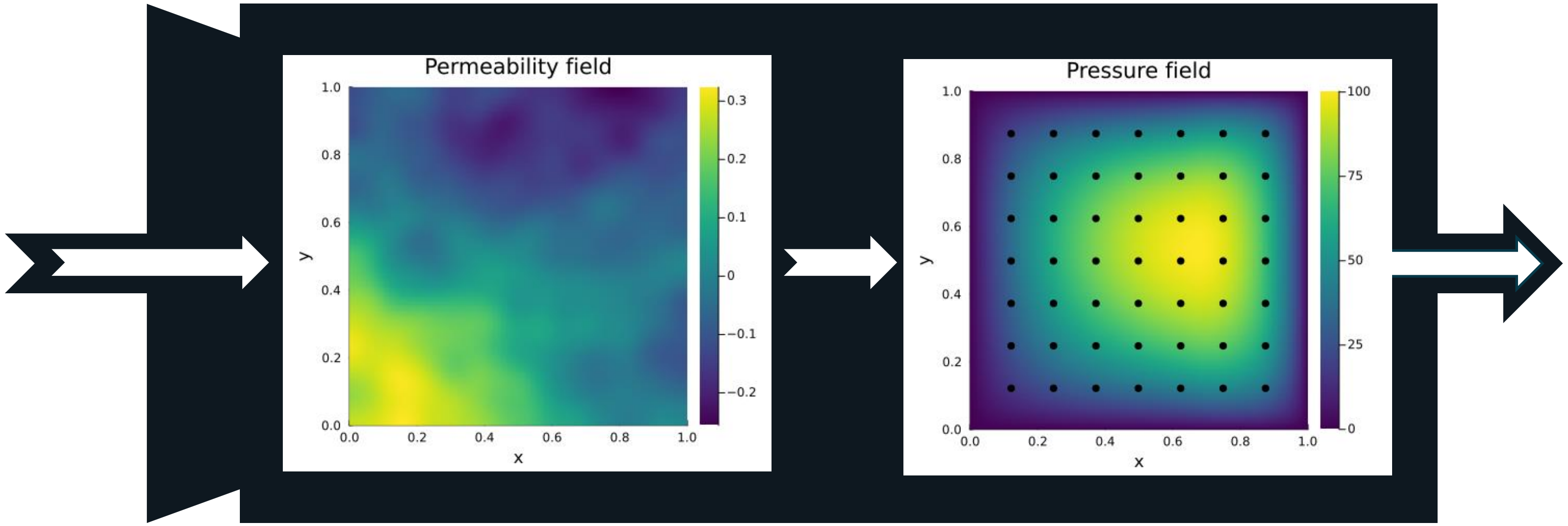


To simulate ground-water flow, researchers require knowledge of the permeability field of the subsurface.



Let’s do something more interesting

$$\begin{bmatrix}\theta_1\\ \theta_2\\ \vdots\\ \theta_{20}\end{bmatrix}$$

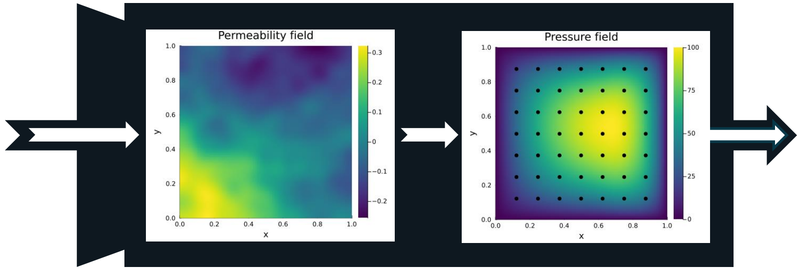


$$\begin{bmatrix}d_1\\ d_2\\ \vdots\\ d_{49}\end{bmatrix}$$

The game isn't always that simple…

- Take measurements $\boldsymbol{d}$.
- Assume a prior $p(\boldsymbol{\theta})$.

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{20} \end{bmatrix}$$



$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{49} \end{bmatrix}$$

- What is the posterior $p(\boldsymbol{\theta}|\boldsymbol{d})$?

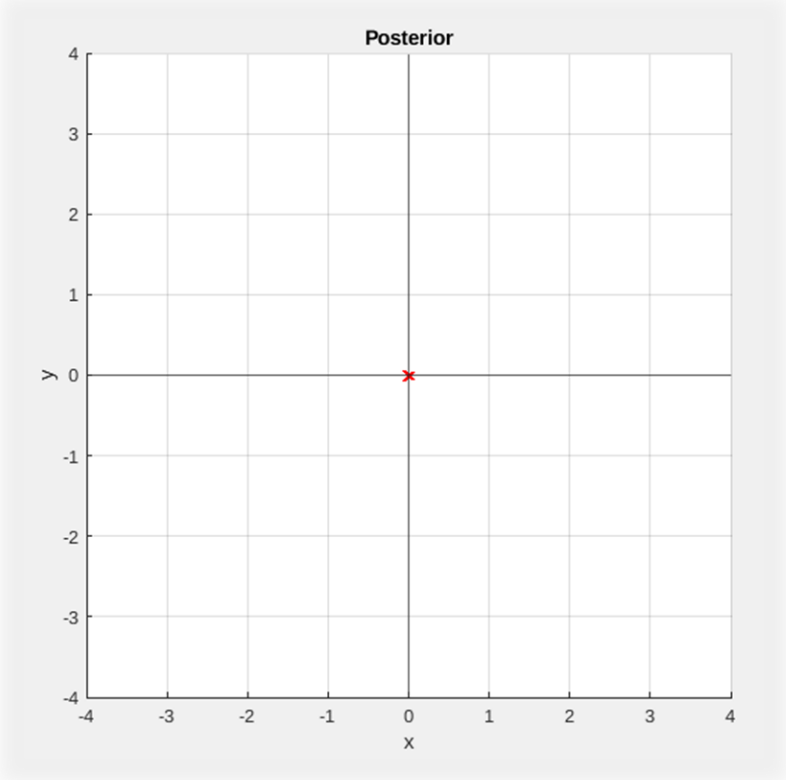
- We cannot simply 'evaluate' the posterior at every point in the **20-dimensional domain**

How can we characterize the posterior distribution?

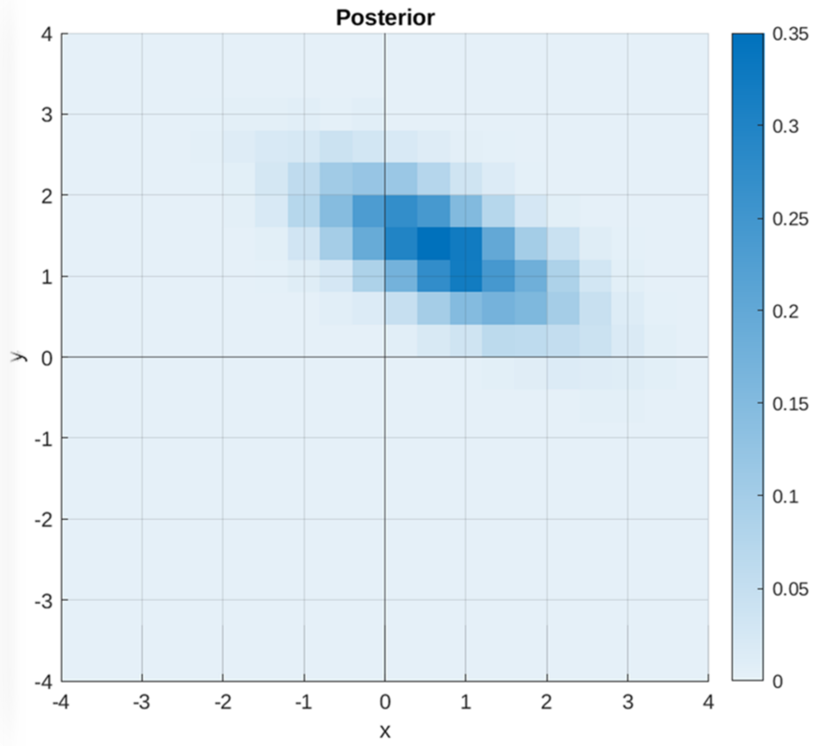
Chapter 2:
How can we characterize the posterior distribution?

Characterizing the posterior

Optimization



Sampling



Sampling via Markov chain Monte Carlo

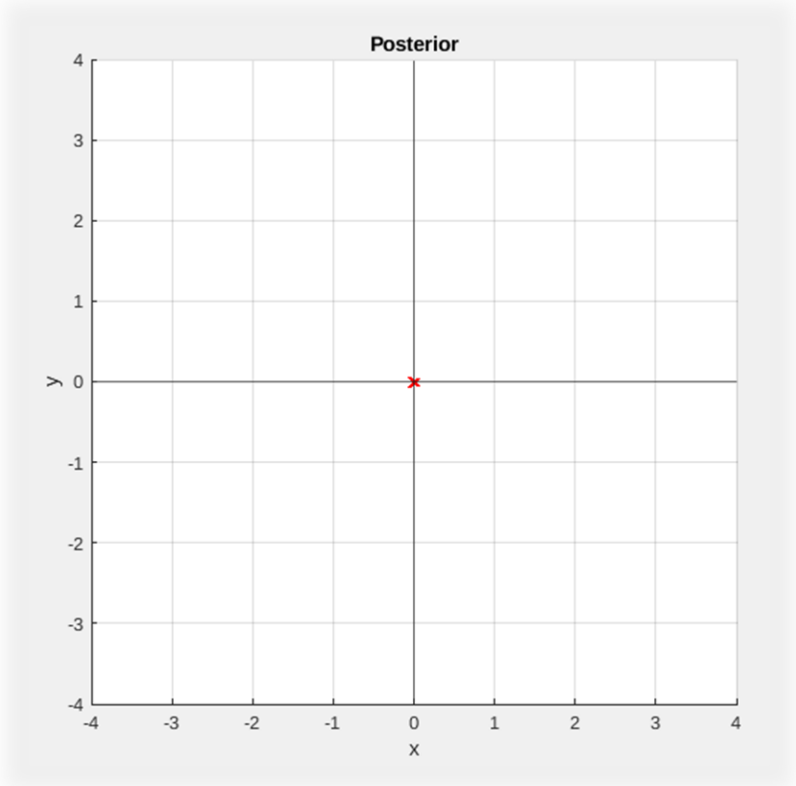
More samples = 

BUT

1. Inherently sequential
2. Evaluation of *The Machine* in every iteration

=> <u>Expensive</u> (if machine is expensive)

Sampling

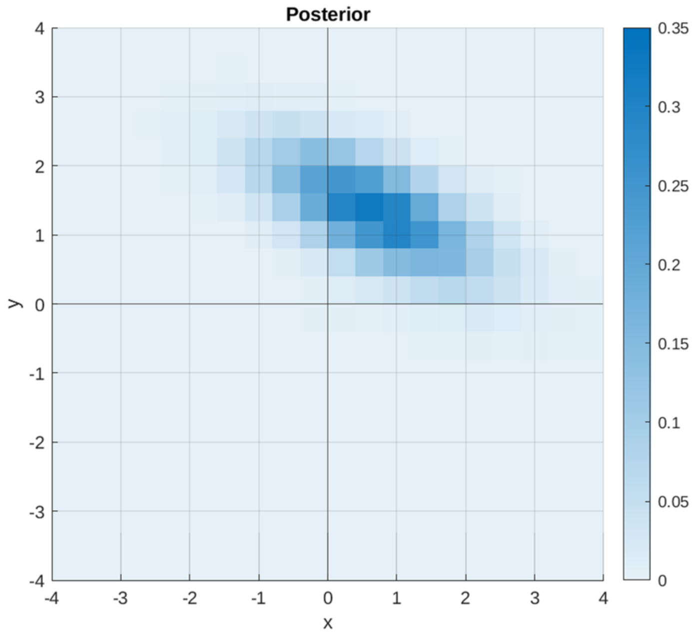


Chapter 3:
Interacting Particle Methods

Sampling via interacting-particle methods

Ensemble of **particles** in parameter space, evolve according to artificial dynamics

- Forward model (*Machine*)
- Interaction
- Randomness



Ensemble Kalman sampling [1]

$$u_{n+1}^{j} = u_n^j + \Delta t\, C_n^{ug} \Gamma^{-1} \left(y - G(u_n^j)\right) - \Delta t\, C_n^{uu} \Gamma_0^{-1} u_n^j + \sqrt{\Delta t}\, \eta_{n+1}^{j}, \quad j \in \{1, \dots, \infty\}$$

$$\eta_{n+1}^{j} \sim N(0, 2C_n^{uu})$$

$$C_n^{uu} = \mathbb{E}[(u_n - \mathbb{E}[u_n]) \otimes (u_n - \mathbb{E}[u_n])]$$

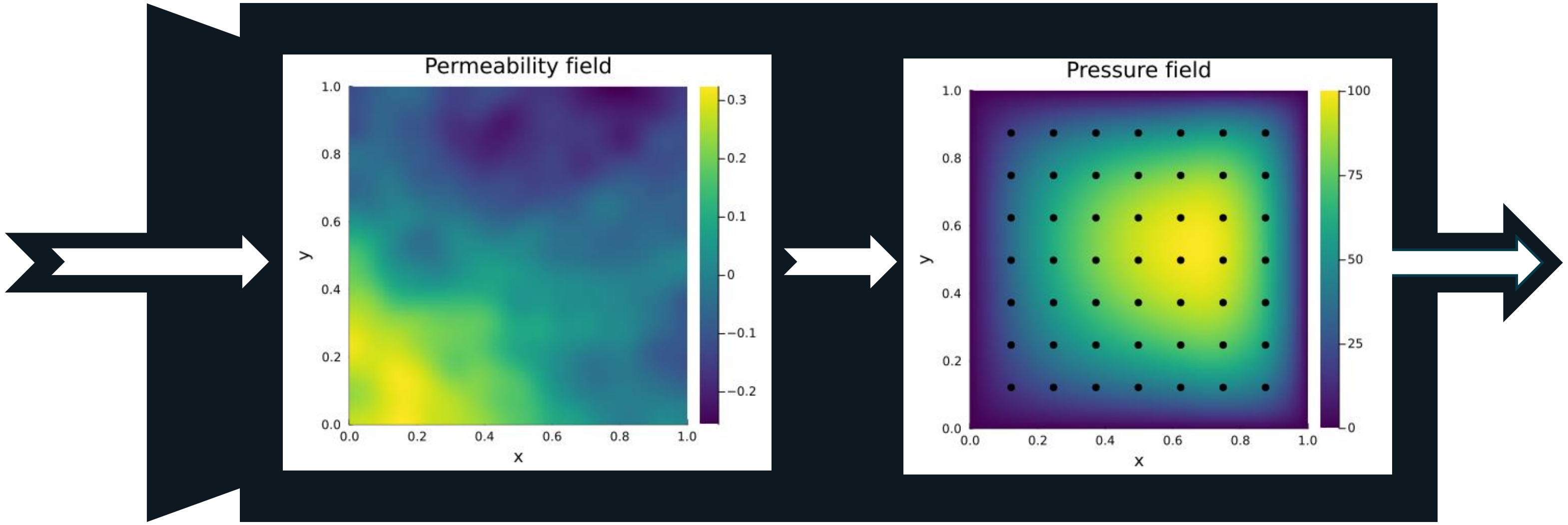
$$C_n^{ug} = \mathbb{E}[(u_n - \mathbb{E}[u_n]) \otimes (G(u_n) - \mathbb{E}[G(u_n)])]$$

[1] Garbuno-Inigo et al. *Interacting Langevin dynamics*. (2020)

Chapter 4:
Multilevel algorithm

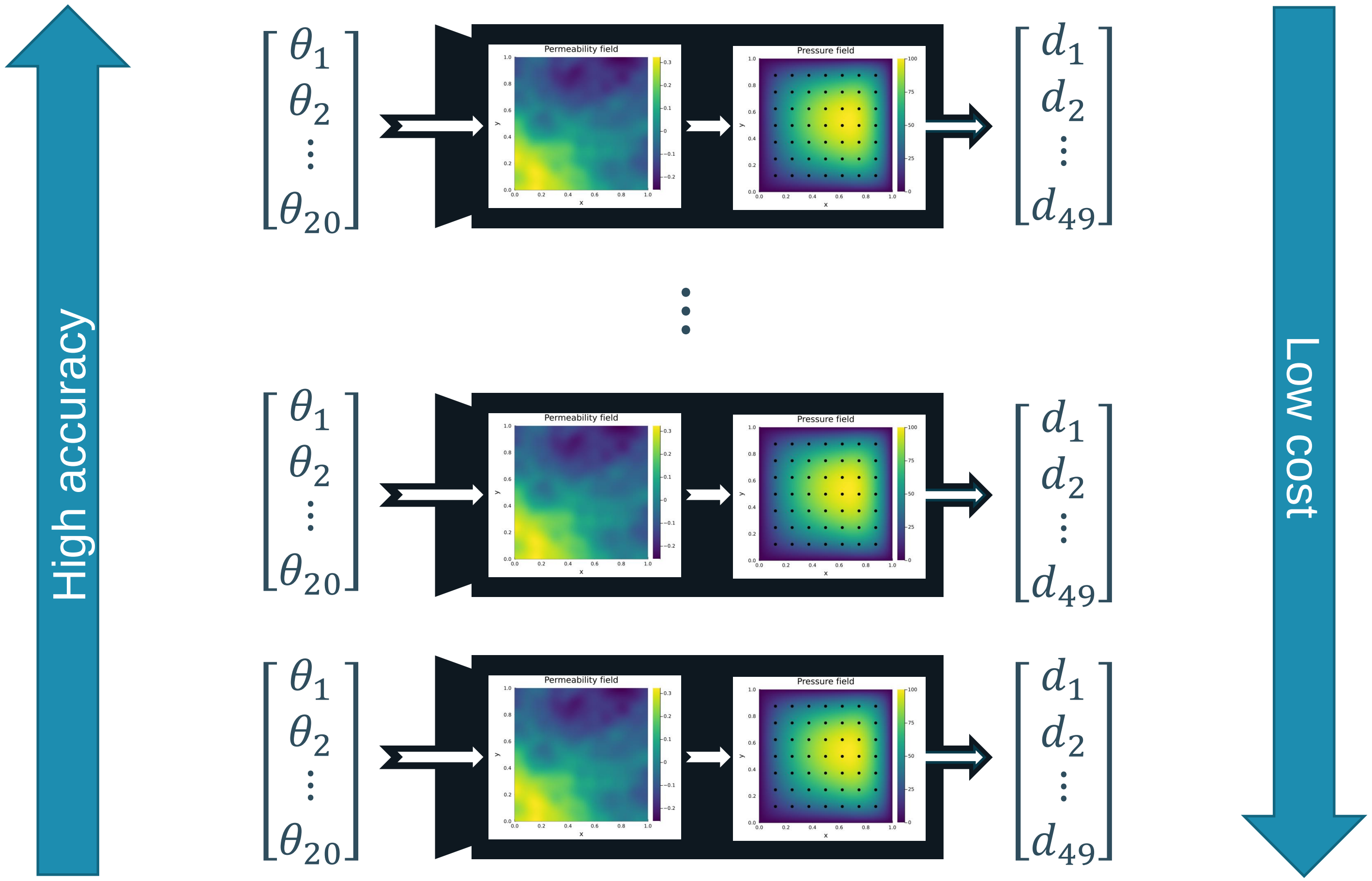
From a single *Machine*…

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{20} \end{bmatrix}$$



$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{49} \end{bmatrix}$$

… to a hierarchy of *Machines*



$$\begin{cases} u_{n+1}^{L,F,j} = u_n^{L,F,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^L(u_n^{L,F,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{L,j} \\ u_{n+1}^{L,C,j} = u_n^{L,C,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^{L-1}(u_n^{L,C,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{L,j} \end{cases} \quad j \in \{1, \dots, J_L\}$$

$$\vdots$$

$$\begin{cases} u_{n+1}^{2,F,j} = u_n^{2,F,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^2(u_n^{2,F,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{2,j} \\ u_{n+1}^{2,C,j} = u_n^{2,C,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^1(u_n^{2,C,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{2,j} \end{cases} \quad j \in \{1, \dots, J_2\}$$

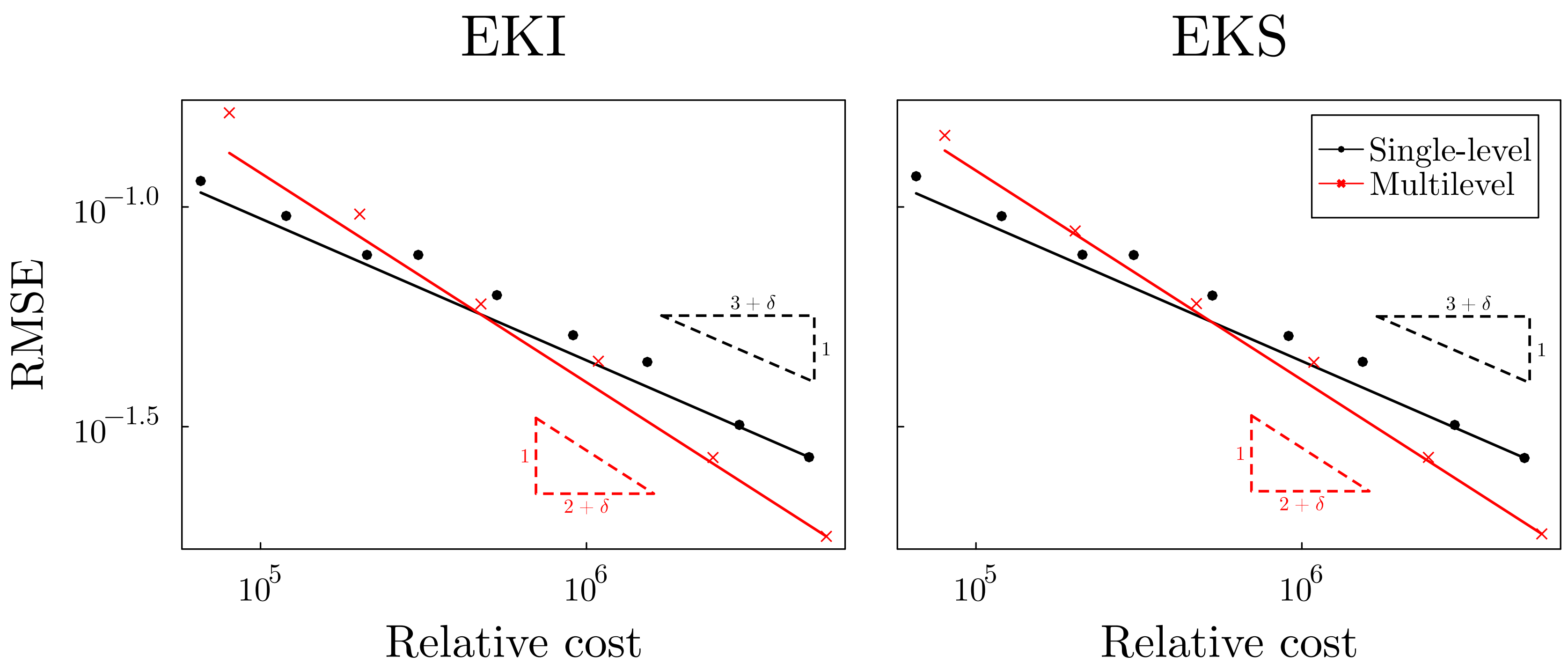
$$\begin{cases} u_{n+1}^{1,F,j} = u_n^{1,F,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^1(u_n^{1,F,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{1,j} \\ u_{n+1}^{1,C,j} = u_n^{1,C,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^0(u_n^{1,C,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{1,j} \end{cases} \quad j \in \{1, \dots, J_1\}$$

$$\left\{ u_{n+1}^{0,F,j} = u_n^{0,F,j} + \Delta t \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \left(y - G^0(u_n^{0,F,j}) \right) + \sqrt{\Delta t} \hat{C}_n^{ug,\mathrm{ML}} \Gamma^{-1} \eta_{n+1}^{0,j} \right. \quad j \in \{1, \dots, J_0\}$$

$$\hat{C}_n^{ug,\mathrm{ML}} = \hat{C}_n^{ug,0} + \sum_{l=1}^{L} \left[\hat{C}_n^{ug,l,F} - \hat{C}_n^{ug,l,C} \right]$$

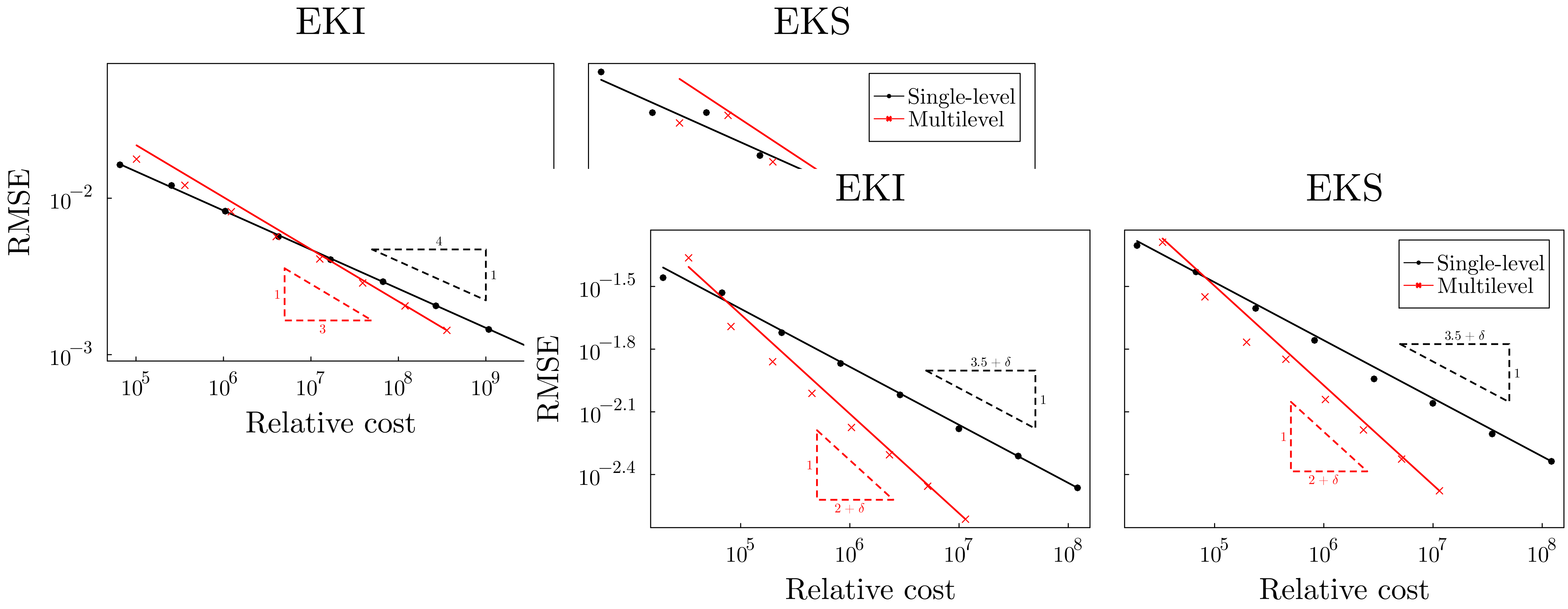
Multilevel interacting-particle methods

For a ‘clever’ choice of L and $\{J_l\}_0^L$… [2]



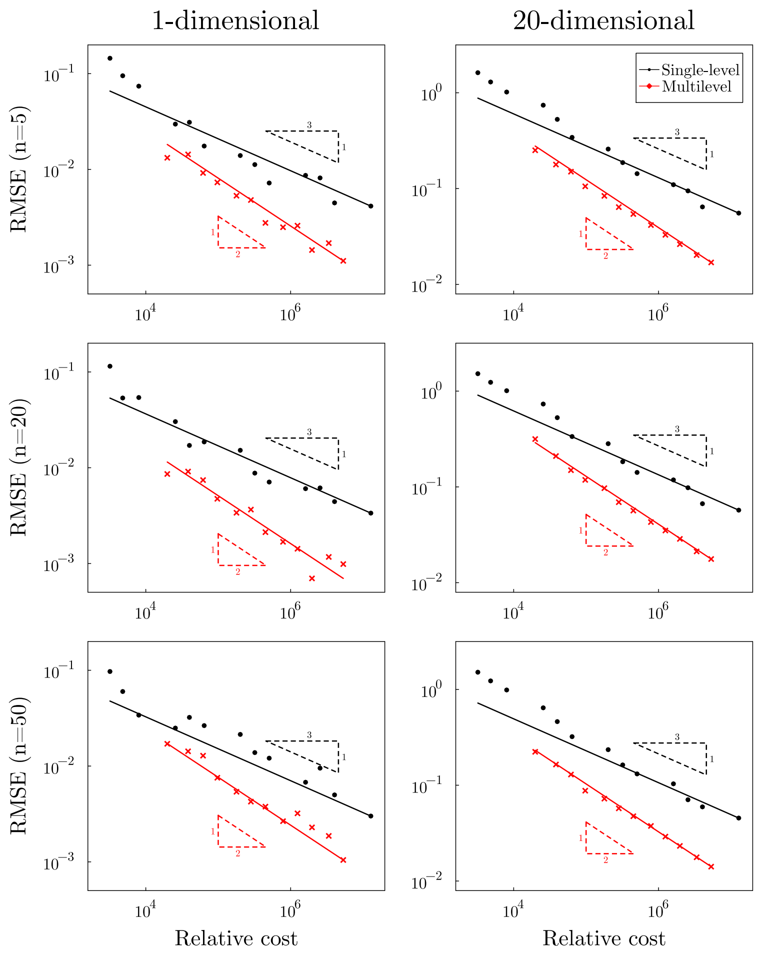
[2] A. Bouillon, T. Ingelaere, G. Samaey. *Single-ensemble multilevel Monte Carlo for discrete interacting-particle methods*. arXiv:2405.10146.

Different ‘machines’



[2] A. Bouillon, T. Ingelaere, G. Samaey. *Single-ensemble multilevel Monte Carlo for discrete interacting-particle methods*. arXiv:2405.10146.

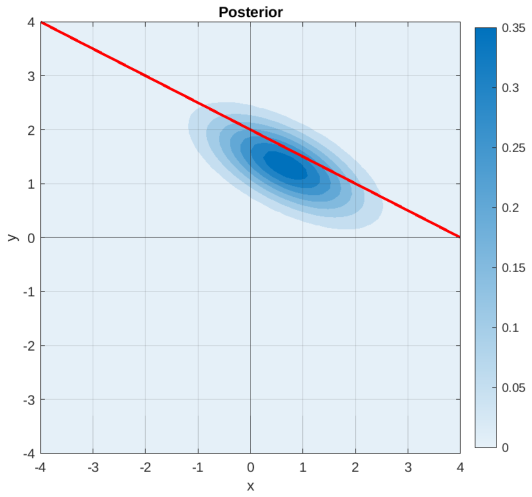
Different interacting-particle methods



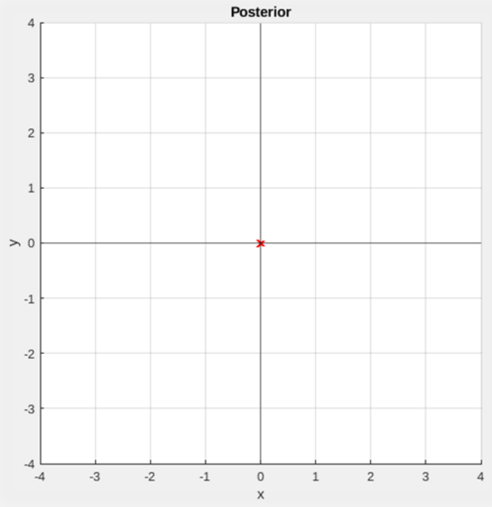
[2] A. Bouillon, T. Ingelaere, G. Samaey. *Single-ensemble multilevel Monte Carlo for discrete interacting-particle methods*. arXiv:2405.10146.

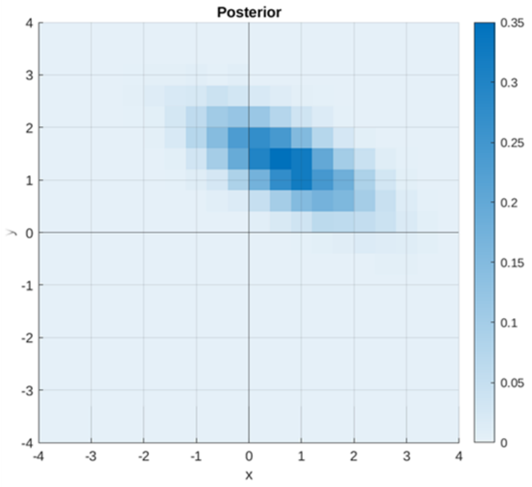
Conclusion

In summary

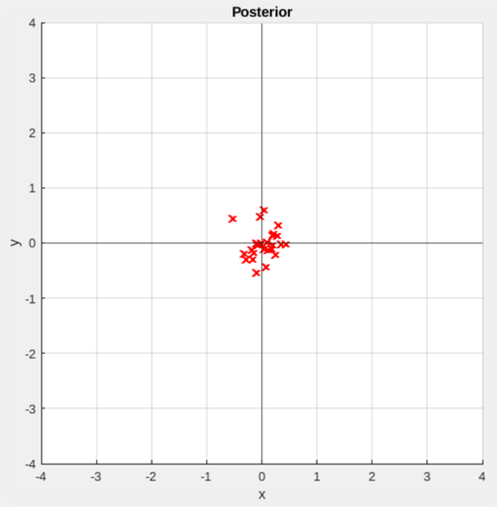


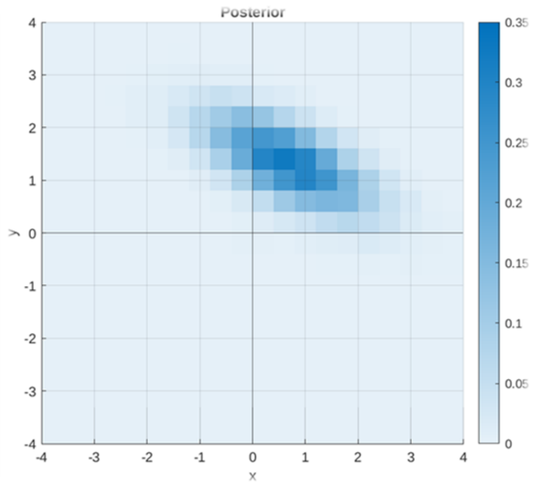
Intractible





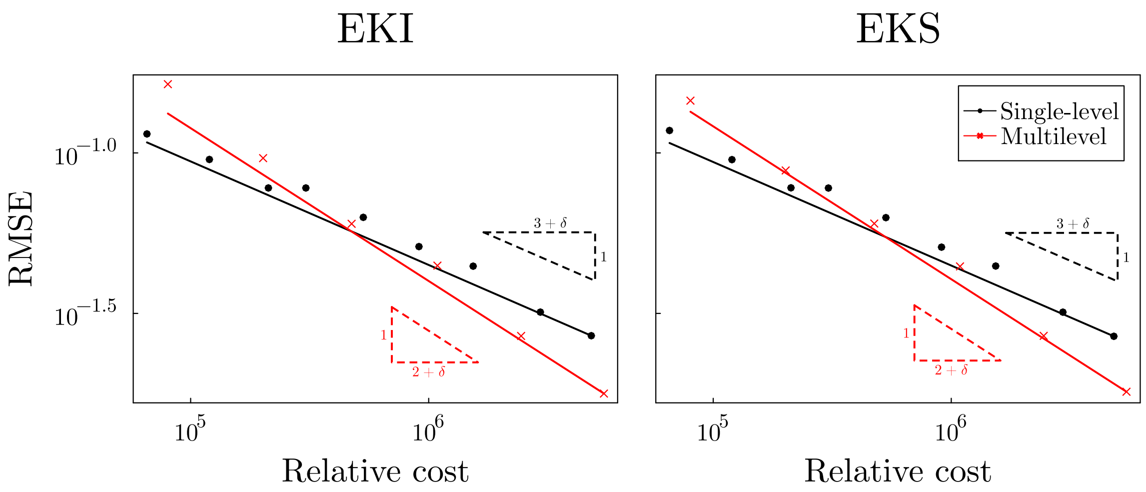
Sequential

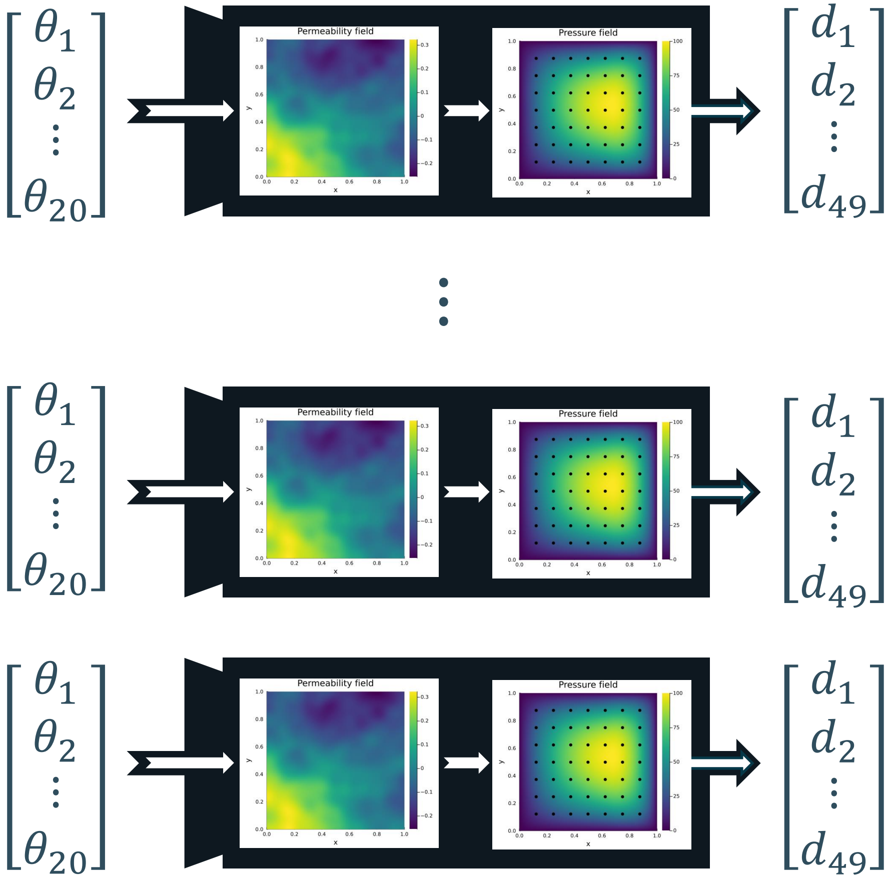




Our contributions

- Generalized multilevel approach from [3]
- Expressions for L and $\{J_l\}_0^L$, based on hierarchy
- Analytic, asymptotic convergence rates





[3] Hoel et al. *Multilevel ensemble Kalman filtering*. (2016)

Main sources

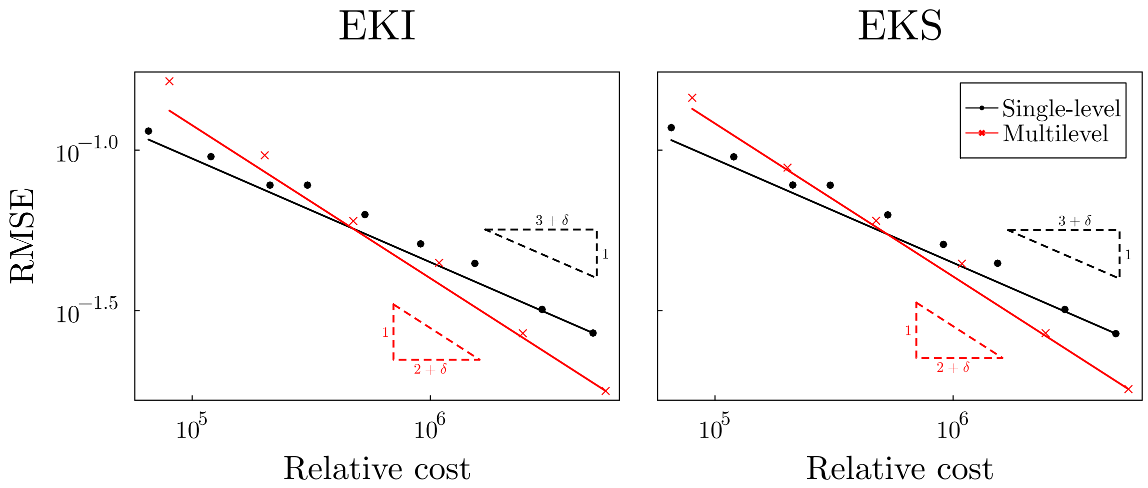
[1] Garbuno-Inigo, A., Hoffmann, F., Li, W., & Stuart, A. M. (2020). *Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler*. SIAM Journal on Applied Dynamical Systems, 19(1), 412-441.

[2] A. Bouillon, T. Ingelaere, G. Samaey. *Single-ensemble multilevel Monte Carlo for discrete interacting-particle methods,* arXiv:2405.10146.

[3] Hoel, H., Law, K. J., & Tempone, R. (2016). *Multilevel ensemble Kalman filtering. SIAM Journal on Numerical Analysis*, 54(3), 1813-1839.

Thanks! Questions?

A. Bouillon, T. Ingelaere, G. Samaey. *Single-ensemble multilevel Monte Carlo for discrete interacting-particle methods*. arXiv:2405.10146.



KU LEUVEN