Bayesian statistics and Information Theory

Lecture 2: Probabilistic computations

… a.k.a. *how much do I know about the likelihood?*

Florent Leclercq

www.florent-leclercq.eu

Imperial Centre for Inference and Cosmology
Imperial College London

May 21st, 2019

ICIC
Imperial Centre
for Inference & Cosmology

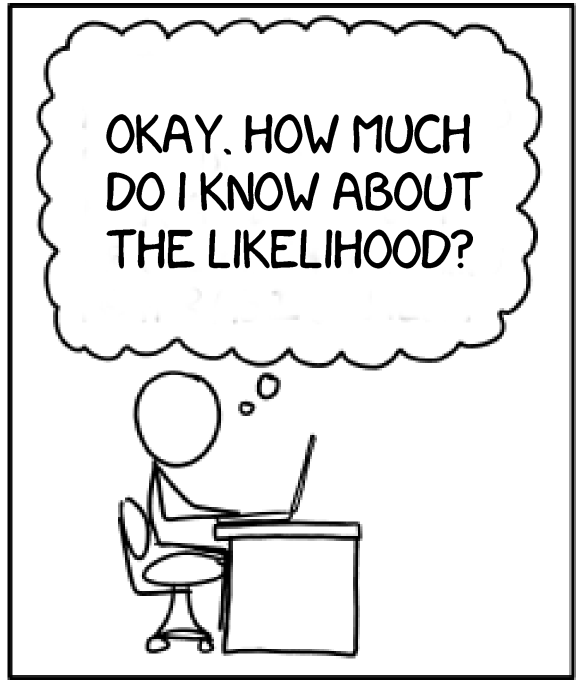
Imperial College
London

Outline: Lecture 2

- Which inference method to choose?
- Monte-Carlo integration, importance sampling, rejection sampling
- Markov Chain Monte Carlo: Metropolis-Hastings algorithm & Gelman-Rubin test
- The test pdf
- Slice sampling
- Gibbs sampling
- Hamiltonian sampling
- Approximate Bayesian Computation: Likelihood-free rejection sampling

Which inference method to choose?

Probabilistic computations: two approaches
(a very personal view)



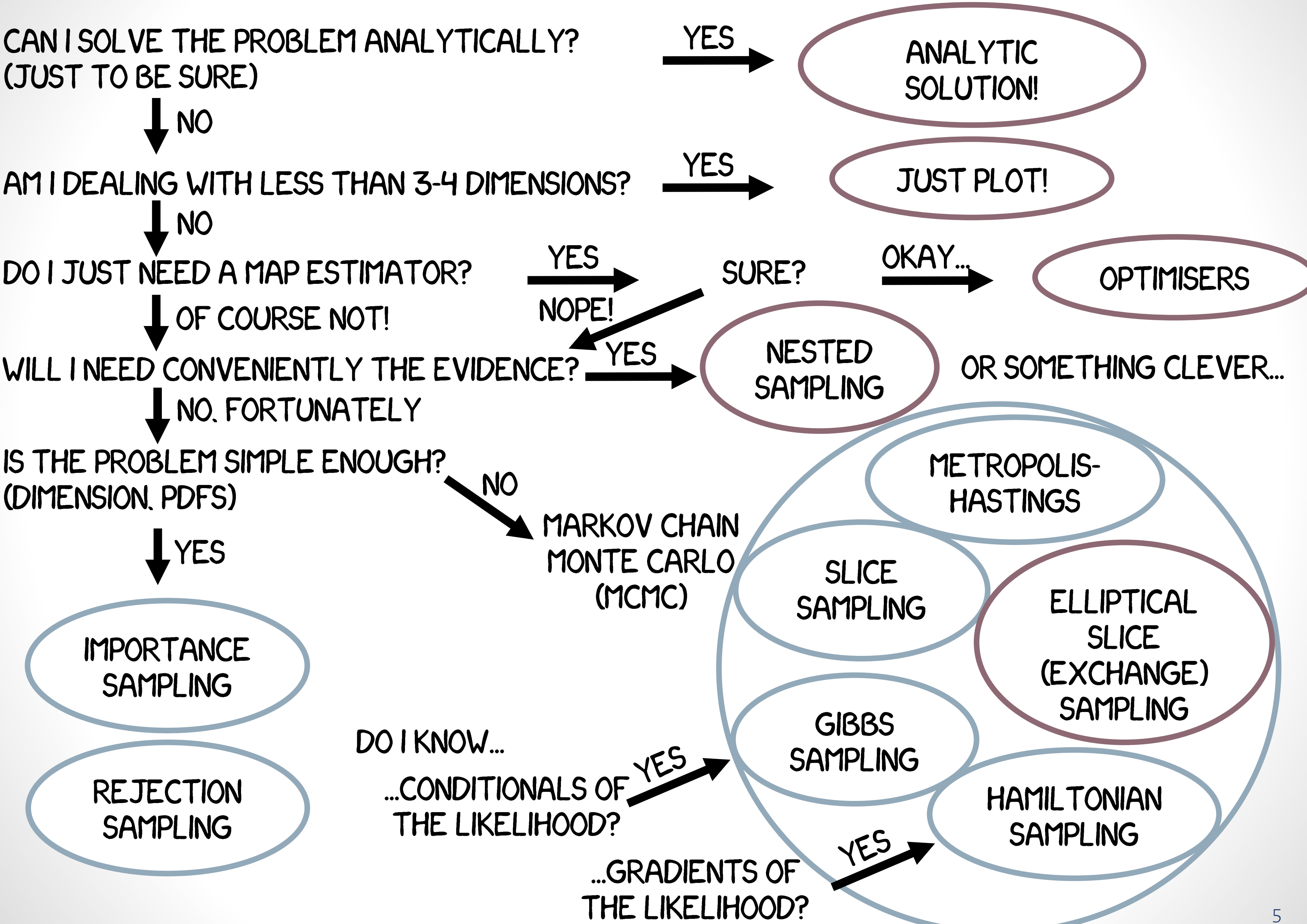
QUITE A BIT

ABSOLUTELY NOTHING

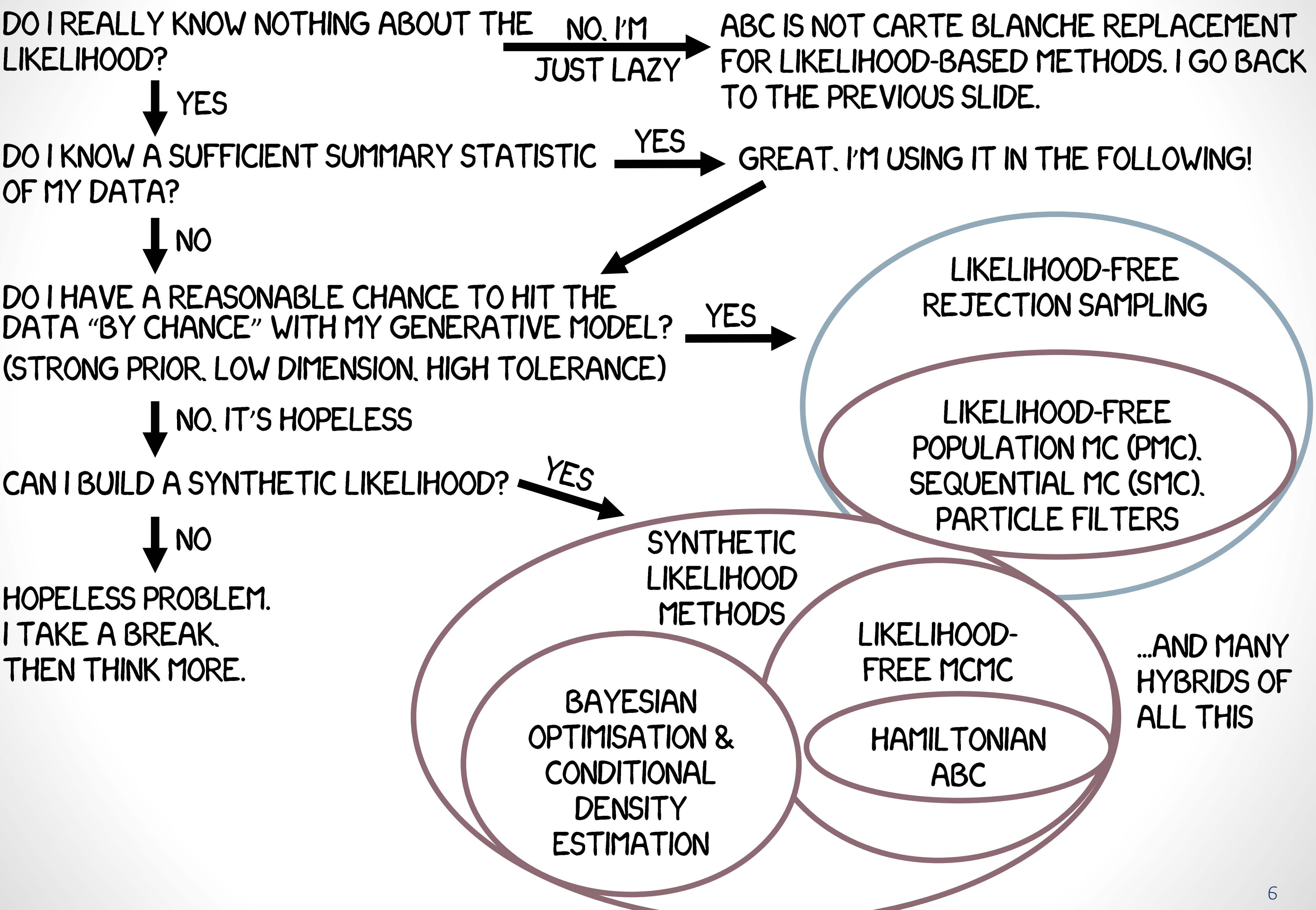
LIKELIHOOD-BASED METHODS:
EXACT BAYESIAN INFERENCE

LIKELIHOOD-FREE METHODS:
APPROXIMATE BAYESIAN COMPUTATION (ABC)

LIKELIHOOD-BASED METHODS: EXACT BAYESIAN INFERENCE



LIKELIHOOD-FREE METHODS: APPROXIMATE BAYESIAN COMPUTATION (ABC)



Monte-Carlo integration, importance sampling, rejection sampling

Notebook 7: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/Sampling_Importance_Rejection.ipynb

Markov Chain Monte Carlo: Metropolis-Hastings algorithm & Gelman-Rubin test

Notebook 8: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/MCMC_MH.ipynb

Slice sampling

Notebook 9: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/MCMC_Slice.ipynb

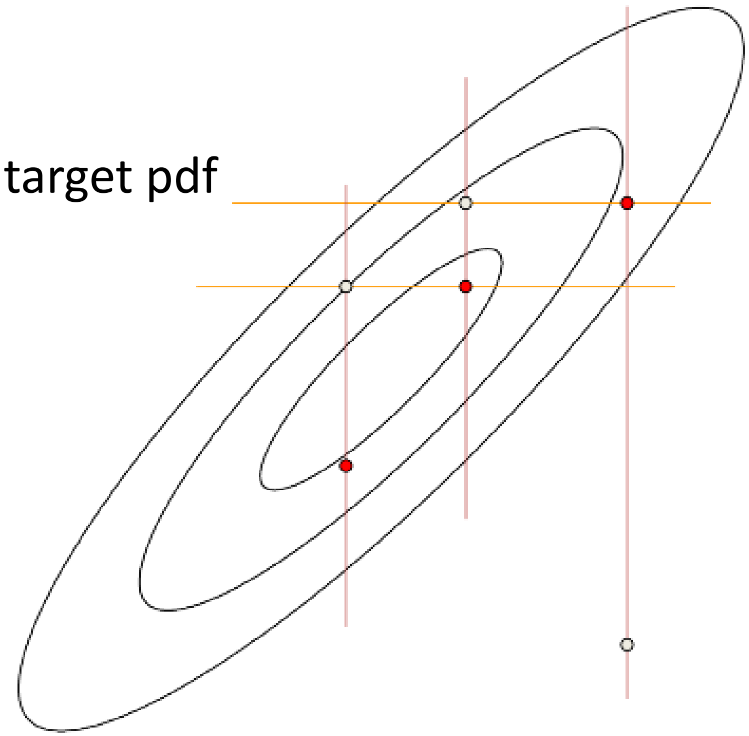
Gibbs sampling

Notebook 10: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/MCMC_Gibbs.ipynb

MCMC beyond Metropolis-Hastings

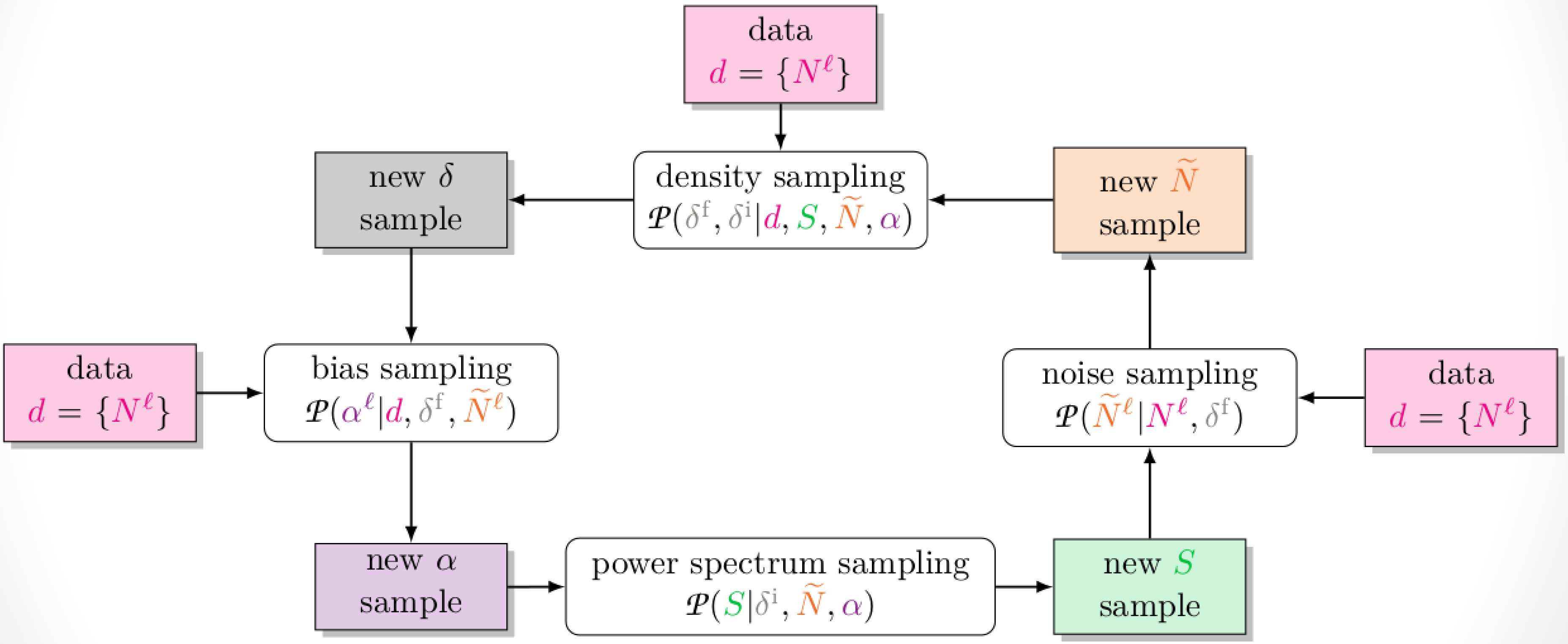
- Shortcomings of standard Metropolis-Hastings:
  - Tuning of proposal distributions
  - Curse of dimensionality

- Gibbs sampling:
  - Uses conditionals of the target pdf



Modular probabilistic programming: example

ARES: *Algorithm for REconstruction and Sampling*



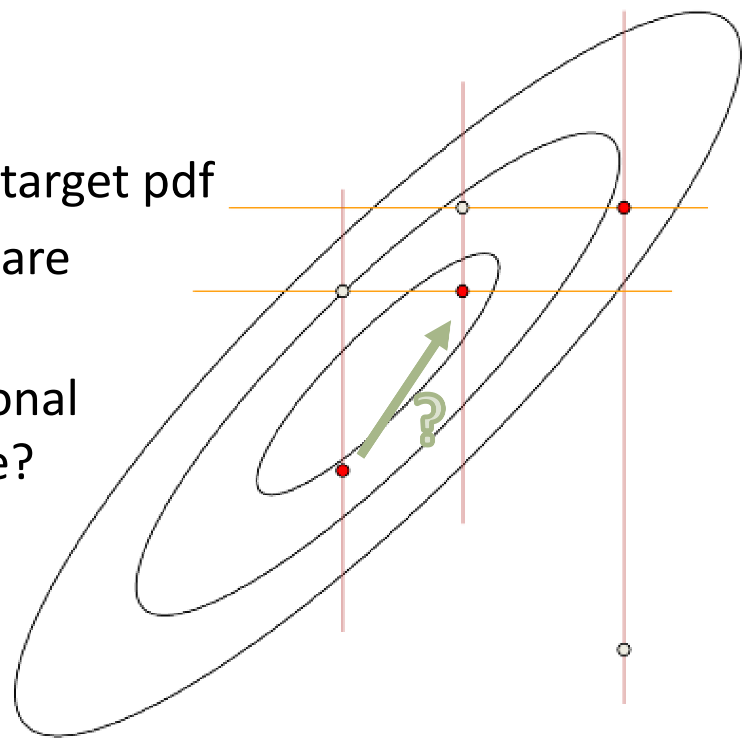
Jasche, Kitaura, Wandelt & Enßlin 2010, arXiv:0911.2493

Jasche & Wandelt 2013, arXiv:1306.1821

MCMC beyond Metropolis-Hastings

- Shortcomings of standard Metropolis-Hastings:
  - Tuning of proposal distributions
  - Curse of dimensionality

- Gibbs sampling:
  - Uses conditionals of the target pdf
  - Inefficient if parameters are strongly correlated
  - How does one take diagonal steps in parameter space?



Hamiltonian sampling

Notebook 11: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/MCMC_Hamiltonian.ipynb

Hamiltonian (Hybrid) Monte Carlo

- Use classical mechanics to solve statistical problems!
  - The potential: $\psi(\mathbf{x}) \equiv -\ln p(\mathbf{x})$
  - The Hamiltonian: $H(\mathbf{x},\mathbf{p}) \equiv \frac{1}{2}\mathbf{p}^\intercal \mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

$$(\mathbf{x},\mathbf{p}) \Rightarrow \left\{ \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}\mathbf{x}} \end{aligned} \right\} \Rightarrow (\mathbf{x}',\mathbf{p}')$$

gradients of the pdf

$$a(\mathbf{x}',\mathbf{x}) = \mathrm{e}^{-(H'-H)} = 1 \Leftarrow$$

acceptance ratio unity

- HMC **beats the curse of dimensionality** by:
  - Exploiting gradients
  - Using conservation of the Hamiltonian

Duane *et al.* 1987, Phys. Lett. B **195**, 2

Approximate Bayesian Computation: Likelihood-free rejection sampling

Notebook 12: https://github.com/florent-leclercq/Bayes_InfoTheory/blob/master/ABC_rejection.ipynb