Programming Languages for Artificial Intelligence (PLAI) Research Group
THE UNIVERSITY OF BRITISH COLUMBIA

Sequential Core-Set Monte Carlo

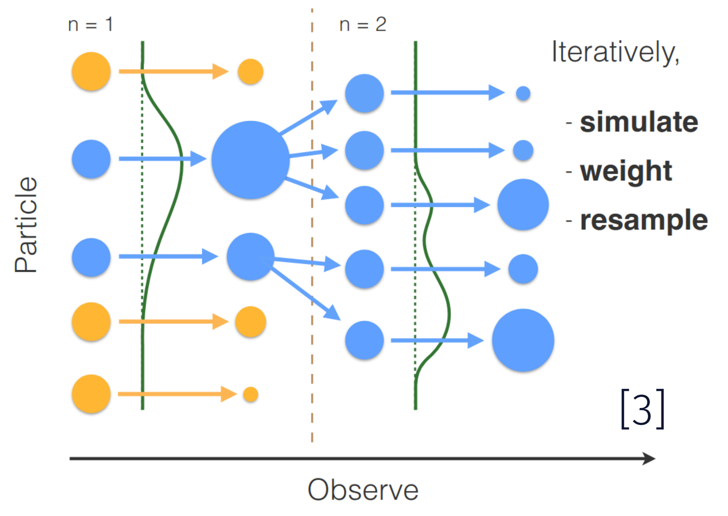
Boyan Beronov*, Christian Weilbach, Frank Wood
Department of Computer Science
Trevor Campbell
Department of Statistics
* beronov@cs.ubc.ca

NSERC CRSNG · DARPA · Mila · CIFAR

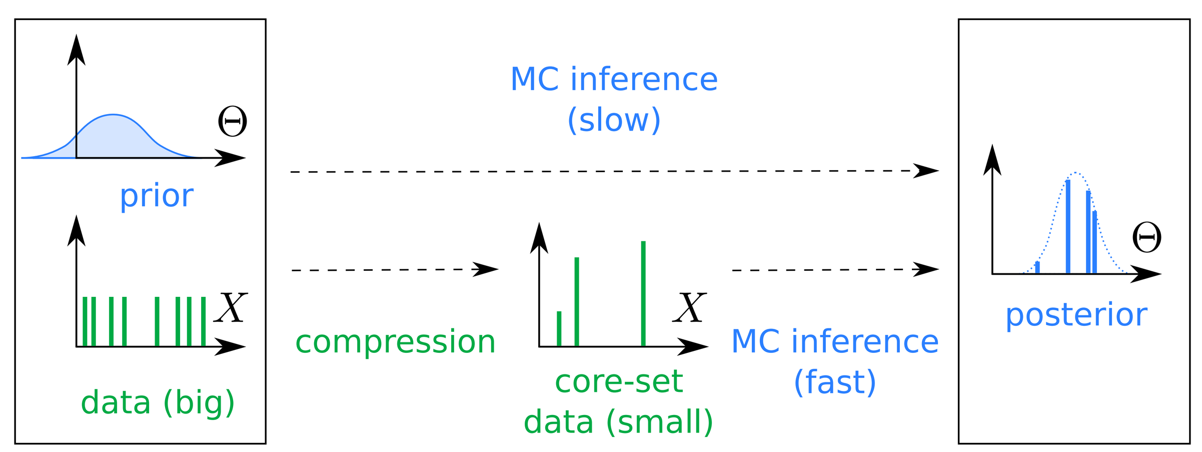
Background

Sequential Monte Carlo

- *generic* & *recursive* Bayesian inference
- *reweighting* & *resampling* updates to particle population
- *rejuvenation:* variance reduction via interleaved Markov chain Monte Carlo moves, requires access to *all past data* for asymptotic consistency

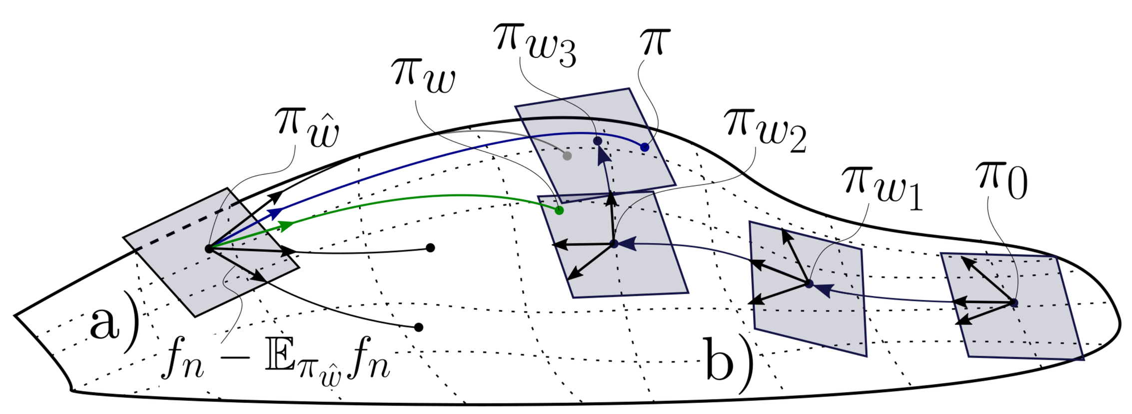


Bayesian core-sets



- *core-set compression:* data pre-processing for accelerating inference [1]
- generalization to *online belief representation* for variance reduction

Information geometric perspective



- centered log-likelihood functions as transportable *tangent vectors* to the statistical manifold of core-set posteriors
- *Fisher information metric* as similarity measure between observations for adaptive compression [4]
- online sparse approximation of density ratio (logarithmic map) *current belief / prior*

Motivation

Necessity of *streaming memory* mechanisms

- *online* inference & *continual* learning
- *large-scale* data (limited working memory)
- analysis by *synthesis* (dictionary, sketching)
- model interpretation by *examples*

***Problems* of previous approaches**

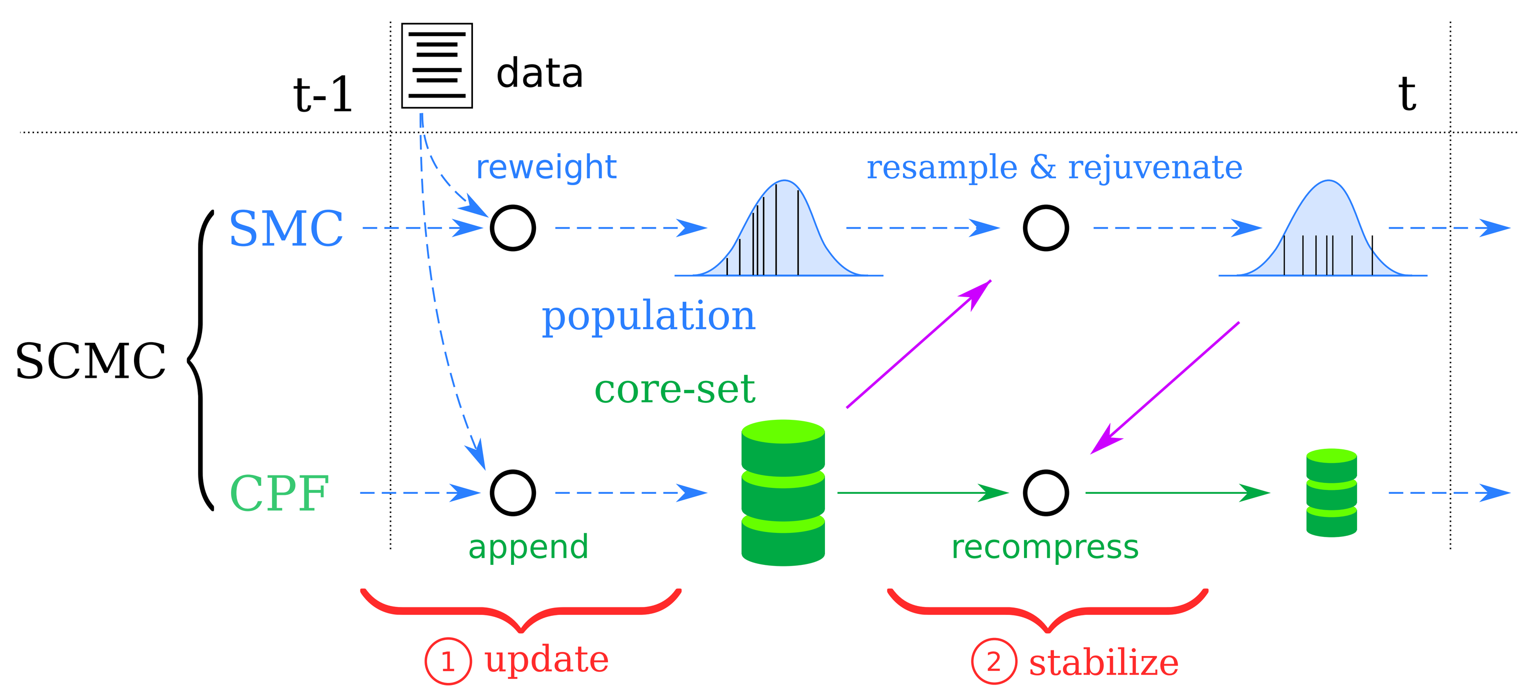
- repeated access to *all data*
- *heuristic* approximations
- *model-specific* methods
- *opaque* representations

Contributions

***Online* sequential Monte Carlo (SMC) variant for *generic* exchangeable Bayesian models**

- *higher accuracy* under resource bounds via Bayesian core-set [1] rejuvenation
- first generic & statistically sound *streaming* Bayesian core-set compression
- *generalization* of projection filtering [2] to core-set posterior exponential family

Algorithm



***Rejuvenation* in sequential Monte Carlo (SMC) leverages core-set memory**

prior $p_0(\theta)$; exact MCMC target $p(\theta \mid x_1, \ldots, x_N) \propto p_0(\theta) \cdot \prod_{j=1}^{N} p(x_j \mid \theta)$

likelihood $p(x \mid \theta)$; *approximate* MCMC target $\approx p_0(\theta) \cdot \prod_{j=1}^{C} p(\bar{x}_j \mid \theta)^{\bar{w}_j}$

***Recompression* in core-set projection filter (CPF) leverages SMC particles**

- *sparse non-negative least squares* objectives in log-likelihood function space at *each filtering step*
- non-uniform (and hence recursively applicable) *generalization* of [5]

Objective for recompressed core-set weights $\hat{w}_j$ of centered potentials $f_j^{(t)}$ at belief state p_t under memory constraint M

$$\operatorname*{argmin}_{\hat{w} \in \mathbb{R}^{\bar{C}^{(t)}}} \mathbb{E}_{p_t}\left[\left(\sum_{j=1}^{\bar{C}^{(t)}} (\hat{w}_j - \bar{w}_j^{(t)}) \cdot f_j^{(t)}(\cdot\,; p_t)\right)^2\right]$$

$$\text{s.t.} \quad \hat{w} \geq 0\,, \quad \|\hat{w}\|_0 \leq M\,,$$

$$f_j^{(t)}(\theta; q) := \log p(\bar{u}_j^{(t)} \mid \theta) - \mathbb{E}_q\left[\log p(\bar{u}_j^{(t)} \mid \cdot)\right]$$

Experiments

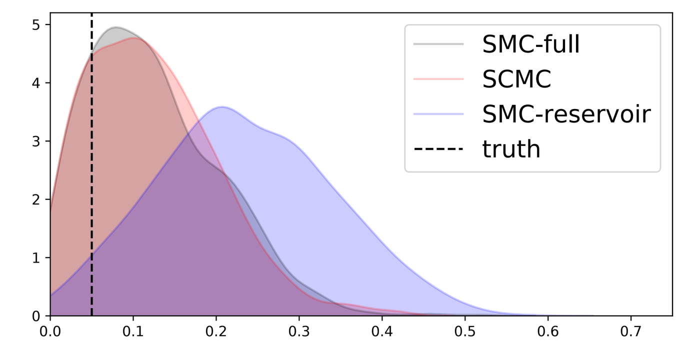
Method comparison

	Rejuvenation	Memory	Time	Error
SMC-full *(oracle)*	$O(T)$	$O(T^2)$	none	
SMC-reservoir *(baseline)*	$O(M)$	$O(MT)$	stoch.	
SCMC *(ours)*	$O(M)$	$O(MT)$	det.	

filtering steps T
memory size M

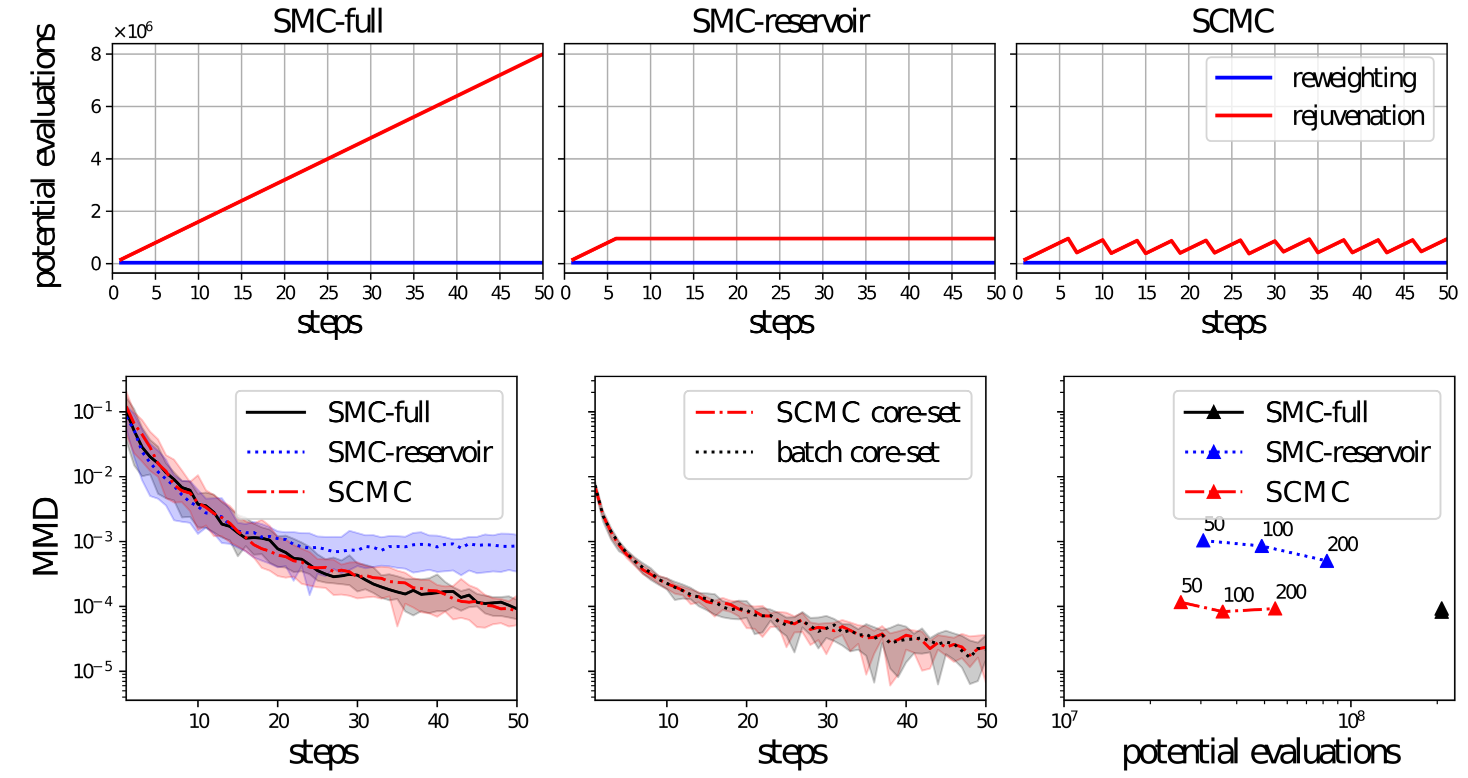
Autoregressive model (synthetic)

Posterior KDE, 1 parameter



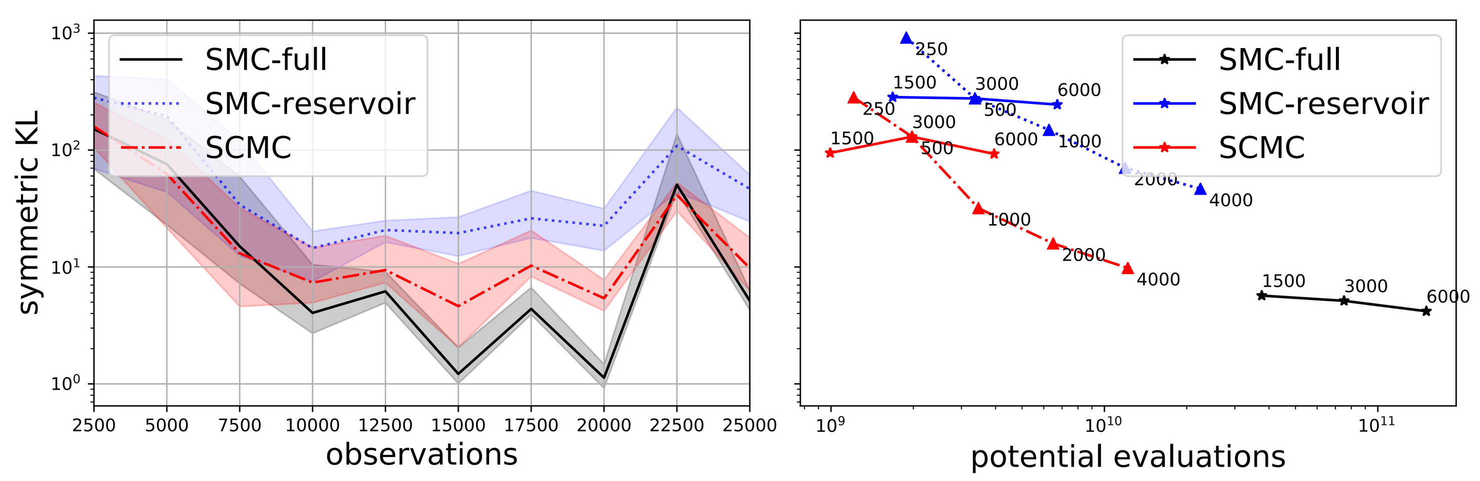
Normal-inverse-Wishart (synthetic)

Cost vs. posterior error, 27 parameters



Bayesian logistic regression (real data)

Cost vs. posterior error, 11 parameters



References

[1] Campbell, Broderick. 'Automated Scalable Bayesian Inference via Hilbert Coresets'. JMLR (2019).
[2] Brigo, Hanzon, LeGland. 'A Differential Geometric Approach to Nonlinear Filtering: The Projection Filter'. IEEE Transactions on Automatic Control (1998).
[3] http://www.robots.ox.ac.uk/~fwood/talks/2015/agi-keynote-2015-probabilistic-programming.pdf
[4] Campbell, Beronov. 'Sparse Variational Inference: Bayesian Coresets from Scratch'. NeurIPS (2019).
[5] Campbell, Broderick. 'Bayesian Coreset Construction via Greedy Iterative Geodesic Ascent'. ICML (2018).