hIPPYlib-MUQ: Scalable Markov Chain Monte Carlo Sampling Methods for Large-scale Bayesian Inverse Problems Governed by PDEs[1]

Noémi Petra[1]

joint work with:
Ki-Tae Kim[1], Umberto Villa[2], Matthew Parno[3],
Youssef Marzouk[4], and Omar Ghattas[5]

[1]Department of Applied Mathematics, University of California, Merced

[2]Department of Electrical & Systems Engineering, Washington University in St. Louis

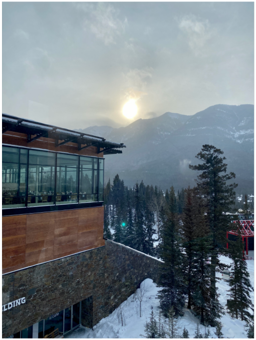
[3]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology

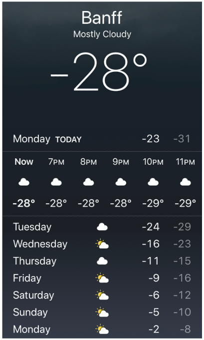
[4]Department of Mathematics, Dartmouth College

[5]Oden Institute, The University of Texas at Austin

"Women in Inverse Problems" Workshop
Banff International Research Station
December 9, 2021

[1]This work was supported by NSF ACI 1550593, 1550547, 1550487, 1429783 and DMS-CAREER-1654311.





"Mathematical Modelling in Glaciology" workshop at the Banff International Research Station (January 12-17, 2020).

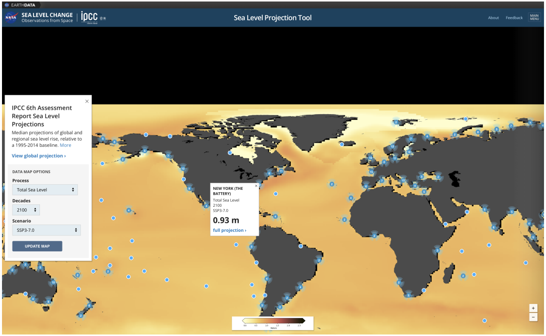
Once upon a time ...

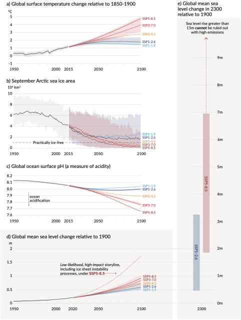
"Mathematical Modelling in Glaciology" workshop at the Banff International Research Station (January 12-17, 2020)[2].

[2]Photo credit BIRS website.

Need predictive models with quantified uncertainties to accurately anticipate future sea level rise.

Sea level projections from the IPCC 6th Assessment Report, 2021





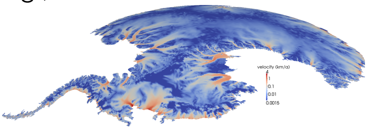
- Several port cities will be at risk from coastal flooding in the future.
- Ice flowing from ice sheets to ocean is primary contributor to sea level rise.

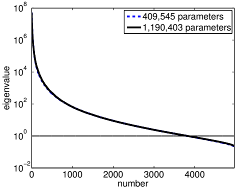
Details in: Masson-Delmotte, V. et al. "IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change", Cambridge University Press. In Press.

Challenges and the need to exploit problem structure and account for model error

Severe mathematical and computational challenges place significant barriers on improving predictability of ice sheet flow models, e.g.,

- complex and very high-aspect ratio (thin) geometry,
- highly nonlinear and anisotropic rheology,
- extremely ill-conditioned and large-scale linear and nonlinear algebraic systems that arise upon discretization,
- uncertain basal sliding parameter, basal topography, geothermal heat flux, and rheology,
- modeling error, etc.

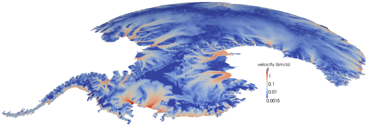


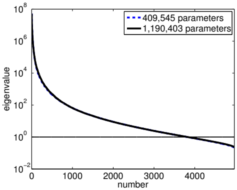


Details in: T. Isaac, N. Petra, G. Stadler, and O. Ghattas. "Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet", Journal of Computational Physics, 296, 348-368 (2015). *Selected as the 2019 SIAM Activity Group on Computational Science and Engineering Best Paper.*

How can we addresss (some of) these challenges?

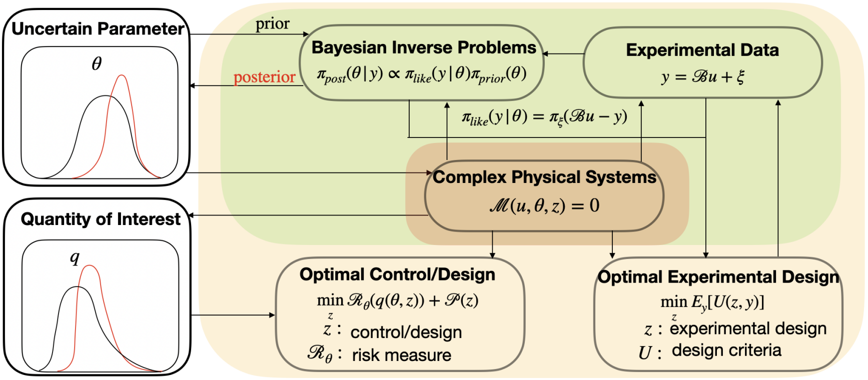
- Learn models from data, i.e., infer unknown/uncertain parameters from available data, e.g., satellite measurements of surface ice flow velocity (**statistical inverse problems governed by PDEs**)
- Apply/adapt/design fast, mesh-independent, structure exploiting, inner-product- and additional uncertainty-aware methods (**scalable, robust and efficient algorithms**)
- Collect data in an optimal way in order to minimize the uncertainty in the inferred parameters or in some predictive quantity of interest (**optimal experimental design**).





Details in: O. Ghattas and K. Willcox. "Learning physics-based models from data: perspectives from inverse problems and model reduction", Acta Numerica, Cambridge University Press (2021).

“outer-loop” problems[3],[4]



Outer-loop problems in forward uncertainty propagation, Bayesian inverse problems, optimal experimental design for optimal data acquisition, and optimization (control and/or design) under uncertainty.

[3] *outer loop:* repeatedly executes a forward model for differing parameters and variables.
[4] Schematic courtasy of Peng Chen (UT Austin) and Alen Alexanderian (NC State).

Motivation to develop an extensible software framework

- **Algorithmic/Computational challenges**:
  - Characterizing the posterior requires repeated evaluations of large-scale PDE models.
  - The posterior distribution often has a complex structure stemming from nonlinear parameter-to-observable maps and heterogeneous sources of data.
  - The parameters often are fields, which when discretized lead to very high-dimensional posteriors.
  - These difficulties render complex large-scale PDE-constrained Bayesian inverse problems intractable via standard MCMC methods.
- **Implementation challenges**:
  - Methods that facilitate the solution of Bayesian inverse problems governed by complex PDE models require a diverse and advanced background.
  - To be efficient these methods generally require first and second derivative information which can be cumbersome to derive.

Goal: Develop an extensible software framework aimed at overcoming these challenges and providing capabilities for additional algorithmic developments for large-scale deterministic and Bayesian inversion.

hIPPYlib: **I**nverse **P**roblems **PY**thon **lib**rary

- hIPPYlib contains the implementation of state-of-the-art scalable adjoint-based algorithms (and their extensions) for PDE-based deterministic and Bayesian inverse problems.
- hIPPYlib builds on FEniCS (a parallel finite element element library) for the discretization of the PDEs, and on PETSc (**P**ortable, **E**xtensible **T**oolkit for **S**cientific **C**omputation) for scalable and efficient linear algebra operations and solvers.
- hIPPYlib is implemented in a mixture of C++ and Python and has been released under the GNU General Public License version.
- One of the main features of this library is that it exposes specific aspects from the model setup to the inverse solution, which can be useful not only for research purposes but also for learning and teaching.
- hIPPYlib 3.0 can be downloaded from:

  http://hippylib.github.io

Details in: U. Villa, N. Petra, and O. Ghattas, *hIPPYlib: An Extensible Software Framework for Large-Scale Deterministic and Linearized Bayesian Inverse Problems*. ACM Transactions of Mathematical Software (TOMS), 47(2), Article 16 (2021)

The FEniCS project

- FEniCS is an open-source computing platform for solving partial differential equations (PDEs).
- Enables users to quickly translate scientific models into efficient finite element code.
- It is easy to use due to the high-level Python and C++ interfaces.
- **FEniCS** is an acronym with FE representing **F**inite **E**lement, CS representing **C**omputational **S**oftware, and according to Anders Logg, a Senior Research Scientist with the FEniCS Project, "ni sits nicely in the middle". Andy Terrel also notes that the FEniCS software package was originally compiled at the University of Chicago, whose mascot is a phoenix, which likely inspired the name. (source: wiki)

http://fenicsproject.org/

hIPPYlib: Algorithms

Software contributions

1. A modular approach to define complex inverse problems governed by (possibly nonlinear or time-dependent) PDEs.
2. Implementation of adjoints and Hessian-actions needed to solve the deterministic inverse problem and to compute the maximum a posteriori (MAP) point of the Bayesian inverse problem.
3. A robust implementation of the inexact Newton-conjugate gradient (Newton-CG) algorithm together with line search algorithms to guarantee global convergence of the optimizer.
4. Implementation of randomized algorithms to compute the low-rank factorization of the misfit part of the Hessian.
5. Scalable algorithms to contruct and evaluate the Laplace approximation of the posterior.
6. Sampling capabilities to generate realizations of Gaussian random fields with a prescribed covariance operator.
7. An estimation of the pointwise variance of the prior distribution and Laplace approximation to the posterior.

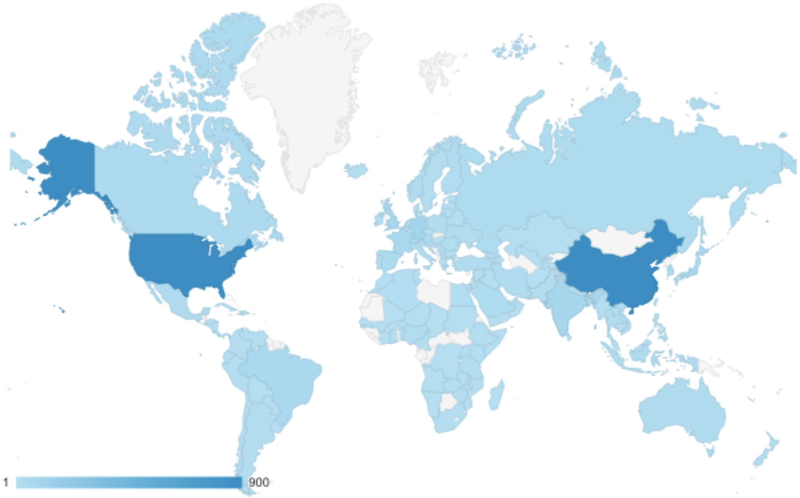
hIPPYlib as research tool

- Inversion for optical properties of biological tissues in optoacoustic tomography (WUSTL)
- Bayesian optimal experimental design for inverse problems in acoustic scattering (UT Austin)
- Inversion and control for CO2 sequestration with poroelastic models (UT Austin)
- Joint seismic-electromagnetic inversion (UT Austin)
- Inversion for coupled ice-ocean interaction (UT Austin)
- Inversion for material properties of cardiac tissue (UT Austin)
- Inference, prediction and optimization under uncertainty for turbulent combustion (UT Austin)
- Inference of constitutive laws in mechanics of nano-scale filaments (UC Merced)
- Bayesian approximation error (BAE) for inverse problems (UC Merced, U. of Auckland, WUSTL, RIT)
- Hierarchical off-diagonal low-rank approximation (HODLR) for Hessians in Bayesian inference applied to ice sheet inverse problems (UC Merced, NYU)
- Statistical treatment of inverse problems constrained by differential equations-based applied to power grid models with stochastic terms (UC Merced, LLNL, ANL)

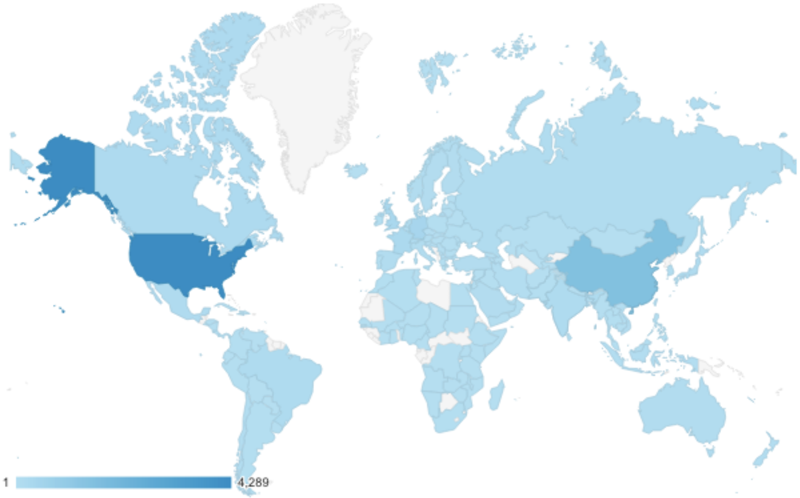
hIPPYlib as teaching tool

- Workshops, short-courses, summer schools:
  - The 2018 Gene Golub Summer School on “Inverse Problems: Systematic Integration of Data with Models Under Uncertainties”
  - The 2016 SAMSI Summer School on Optimization
  - The 2015 ICERM Idea-Lab on Inverse Problems and Uncertainty Quantification
  - Several pedagogical tutorials available at http://g2s3.com/labs
- Graduate level courses on inverse problems and imagining at
  - The University of Texas at Austin (O. Ghattas)
  - University of California, Merced (N. Petra)
  - Washington University in St. Louis (U. Villa)
  - New York University, Courant Institute (G. Stadler)
  - North Carolina State (A. Alexanderian)

hIPPYlib user community



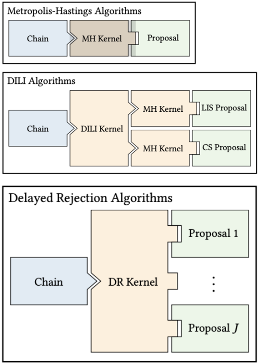
hIPPYlib user community



MUQ: MIT Uncertainty Quantification

- MUQ provide an easy-to-use framework for defining and solving UQ problems with complex models in C++ and Python. It allows:

- **sampling of non-Gaussian distributions** (e.g., Markov chain Monte Carlo and importance sampling),
- **approximating computationally intensive forward models** (e.g., polynomial chaos expansions and Gaussian process regression),
- **working with integral covariance operators** (e.g., Gaussian processes and Karhunen-Love decompositions), and
- **characterizing predictive uncertainties**.

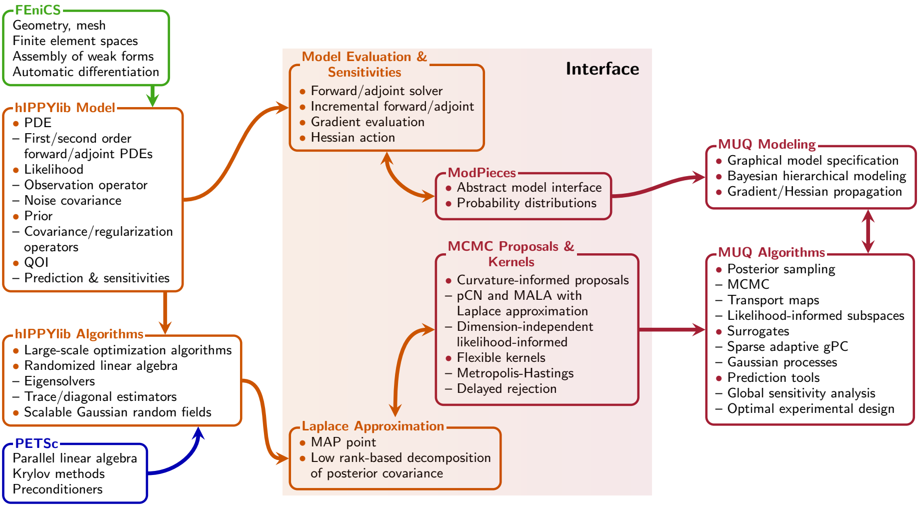


- MUQ can be downloaded from:

http://muq.mit.edu

Details in: Matthew Parno, Andrew Davis, and Linus Seelinger, *MUQ: The MIT Uncertainty Quantification Library*, JOSS, under review.

hIPPYlib-MUQ integration



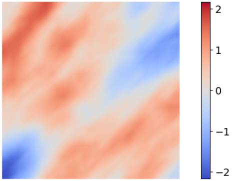
https://github.com/hippylib/hippylib2muq

$$-\nabla \cdot (\exp(m)\nabla u) = f \quad \text{in } \Omega \subset \mathbb{R}^d \quad \text{with bdry conditions}$$

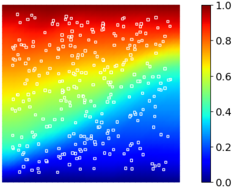
f : given source function

u : pressure (for subsurface flow)

m : unknown log-permeability



- the parameter-to-observable map $\mathcal{F}$ involves the solution of the PDE given m, followed by the application of an observation operator to extract the observations from the state.
- data/measurements: pressure u at points/parts of Ω
- parameter field/image: $m = m(x)$
- inverse problem: given (measurements of) u, find m



Applications: Estimating the log-permeability field for instance in subsurface flow.

Example: Coefficient field inversion in a PDE

Estimating the log-permeability field in subsurface flow

Parameter-to-observable map: often unique solution, well-posed

$$\boxed{m} \xrightarrow[\text{PDE/DAE}]{\text{forward}} \boxed{\mathcal{F}(m)} \xrightarrow{\text{Observations}} \boxed{\boldsymbol{d} = \mathcal{F}(m) + \boldsymbol{\eta}}$$

Inverse problem: ill-posed, under-determined

- Bayesian approach to ill-posedness: describe probability of all models that are consistent with the observations/data and any prior knowledge about the parameters:

$$d\mu_{\text{post}} \propto \exp\left\{ -\frac{1}{2}\|\mathcal{F}(m) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2}\|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}} \right\}.$$

  - The first term in the exponential is the negative log-likelihood.
  - The second term represent the negative log-prior. In `hIPPYlib` we work with a Gaussian prior, i.e., $m \sim \mathcal{N}(m_{\text{pr}}, \mathcal{C}_{\text{prior}})$.

Example: Coefficient field inversion in a PDE

Estimating the log-permeability field in subsurface flow

- The *maximum a posteriori* (MAP) point m_{MAP} is defined as the parameter field that maximizes the posterior distribution:

$$m_{\text{MAP}} := \underset{m \in \mathcal{M}}{\operatorname{argmin}}(-\log d\mu_{\text{post}}(m)) = \underset{m \in \mathcal{M}}{\operatorname{argmin}} \frac{1}{2}\|\mathcal{F}(m) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}}.$$

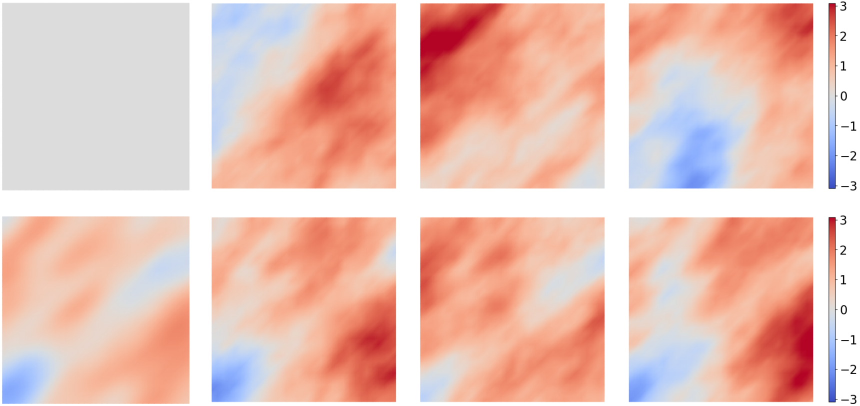
- When $\mathcal{F}$ is linear, due to the particular choice of prior and noise model, the posterior measure is Gaussian, $\mathcal{N}(m_{\text{MAP}}, \mathcal{C}_{\text{post}})$

$$\mathcal{C}_{\text{post}} = \mathcal{H}^{-1} = (\mathcal{F}^* \boldsymbol{\Gamma}^{-1}_{\text{noise}} \mathcal{F} + \mathcal{C}^{-1}_{\text{prior}})^{-1}, \qquad m_{\text{MAP}} = \mathcal{C}_{\text{post}}(\mathcal{F}^* \boldsymbol{\Gamma}^{-1}_{\text{noise}} \boldsymbol{d} + \mathcal{C}^{-1}_{\text{prior}} m_{\text{pr}}),$$

  where $\mathcal{F}^* : \mathbb{R}^q \rightarrow \mathcal{M}$ is the adjoint of $\mathcal{F}$.

- In the general case of nonlinear parameter-to-observable map $\mathcal{F}$ the posterior distribution is not Gaussian. In this case one needs to use sampling to characterize the posterior.

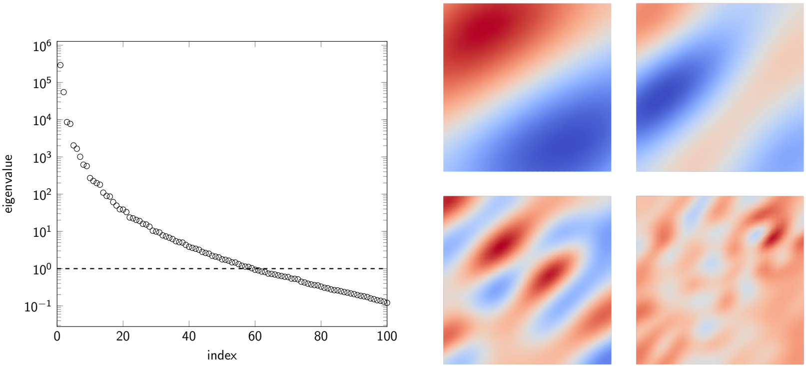
The prior and (Gaussian) posterior



Top: The prior mean (far left) and samples (right). Bottom: The (Gaussian) posterior mean (MAP) (far left) and samples (right).

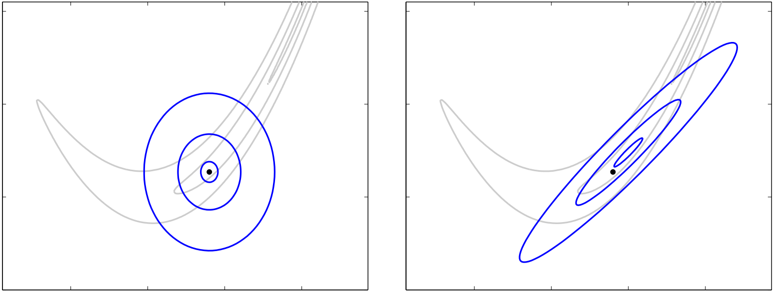
Low-rank-based posterior covariance

The spectrum (left) and eigenvectors (right) of the prior-preconditioned data misfit Hessian



Cost of low rank approximation (extracted via randomized SVD) measured in the number of matrix-free Hessian vector products is independent of the data and parameter dimensions.

Hessian informed MCMC



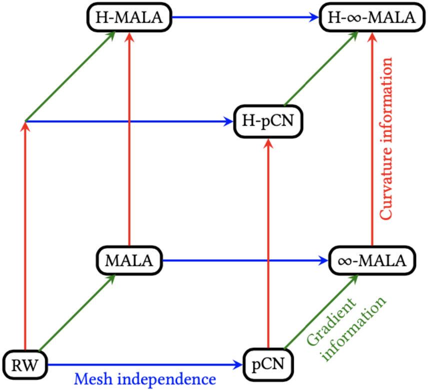
An isotropic (left) and curvature (Hessian) aware (right) MCMC proposal.

hIPPYlib-MUQ enables Hessian informed MCMC that

- exploits the curvature (Hessian) of the posterior $\rightarrow$ significant improvement of sampling performance
- uses low rank approximation of the Hessian $\rightarrow$ efficient Hessian operation

Exploring the posterior using MCMC methods

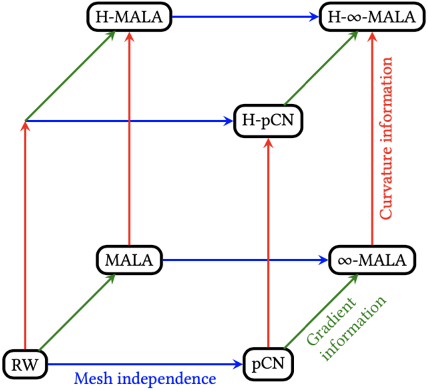
MCMC proposal distributions available in MUQ



The relationship of various MCMC proposal distributions with respect to mesh-refinement independence (blue arrow), gradient awareness (green arrow), and curvature awareness (red arrow).

Exploring the posterior using MCMC methods

MCMC proposal distributions available in MUQ



- RW: random walk
- MALA: Metropolis adjusted Langevin
- pCN: preconditioned Crank-Nicolson
- $\infty-$MALA: infinite-dimensional MALA
- H-·: Hessian-informed proposal

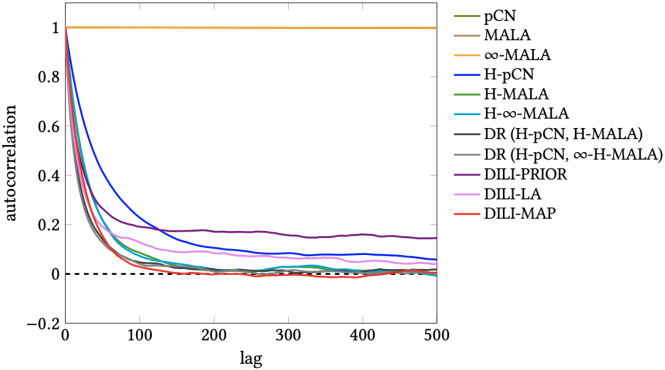
Exploring the posterior using MCMC methods

Convergence diagnostics and computational efficiency of the MCMC samples

Method	AR (%)	MPSRF	Min. ESS (index)	Max. ESS (index)	Avg. ESS	NPS/ES
pCN (5.0E-3)	24	2.629	25 (24)	225 (8)	84	5,952
MALA (6.0E-6)	48	2.642	26 (22)	874 (5)	148	10,135
∞-MALA (1.0E-5)	57	2.943	25 (23)	1,102 (5)	160	9,375
H-pCN (4.0E-1)	27	1.192	64 (1)	3,598 (15)	2,314	216
H-MALA (6.0E-2)	60	1.014	545 (1)	8,868 (19)	6,459	232
H-∞-MALA (1.0E-1)	71	1.016	582 (1)	8,417 (18)	5,905	254
DR (H-pCN (1.0E0), H-MALA (6.0E-2))	(4, 61)	1.013	641 (1)	12,522 (17)	9,222	215
DR (H-pCN (1.0E0), H-∞-MALA (2.0E-1))	(4, 48)	1.011	613 (1)	12,812 (17)	9,141	213
DILI-PRIOR (0.8, 0.1)	(60, 33)	1.064	314 (1)	4,667 (13)	3,216	548
DILI-LA (0.8, 0.1)	(83, 36)	1.017	562 (1)	10,882 (17)	7,192	245
DILI-MAP (0.8, 0.1)	(77, 22)	1.006	1,675 (1)	10,271 (20)	8,692	202

- Acceptance rate (AR), multivariate potential scale reduction factor (MPSRF), and effective sample sample size (ESS)
- NPS/ES: the number of forward and/or adjoint PDE solves required to draw a single independent sample
- NPS/ES can be used to measure the sampling efficiency and rank the methods in terms of computational efficiency.
- Under this metric, DILI-MAP is the most efficient method, DR and H-pCN are close seconds.
- 20 independent MCMC chains, each with 25,000 samples, total of 500,000 samples

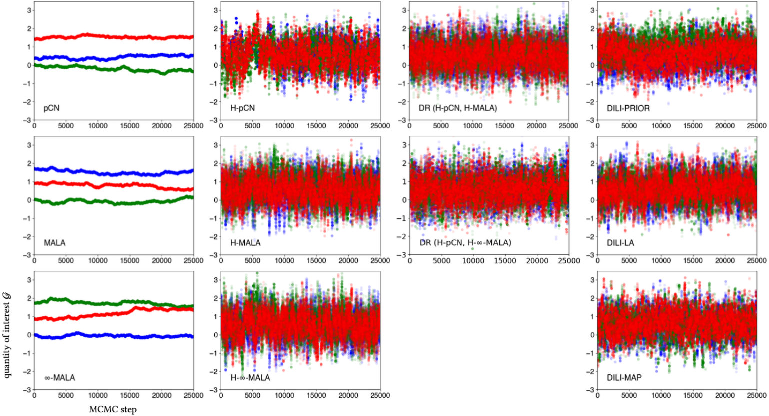
Convergence of MCMC samples
of a QoI (log effective permeability)



Autocorrelation function estimate of the log effective permeability for several MCMC methods. This result can be used to estimate the effective sample size.

Performance comparison

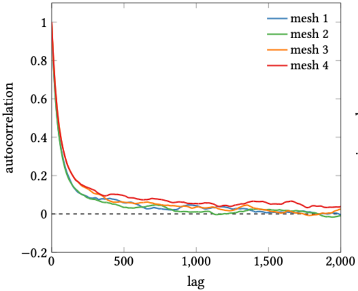
Trace plots of the quantity of interest (log of normal flux through the bottom boundary)

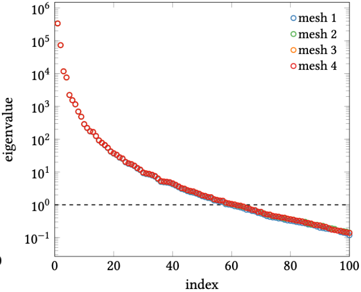


- MALA: Metropolis-adjusted Langevin algorithm
- DR: Delayed rejection (Compose two proposals to increase AR)
- DILI: Dimension independent likelihood informed (Requires off-line sampling of Hessian according to prior, Laplace approximation, or posterior distribution)

Mesh independence of H-pCN

Dimension (state, parameter)	AR (%)	MPSRF	Min. ESS (index)	Max. ESS (index)	Avg. ESS
(4,225, 1,089)	27	1.192	64 (1)	3,598 (15)	2,314
(16,641, 4,225)	24	1.333	63 (1)	3,221 (18)	1,830
(66,049, 16,641)	23	1.075	209 (1)	3,073 (11)	1,940
(263,169, 66,049)	22	1.117	102 (2)	3,276 (15)	1,767





The convergence of samples is almost independent with respect to the MPSRF and the autocorrelation function.

Summary

- We have presented a robust and scalable software framework for the solution of large-scale Bayesian inverse problems governed by PDEs.
- The software integrates two complementary open-source software libraries, hIPPYlib and MUQ, resulting in a unique software framework that addresses the prohibitive nature of Bayesian solution of inverse problems governed by PDEs.
- The main objectives of hIPPYlib-MUQ are to
  1. provide to domain scientists a suite of sophisticated and computationally efficient MCMC methods that exploit Bayesian inverse problem structure; and
  2. allow researchers to easily implement new methods and compare against the state of the art.

Details in: K.T. Kim, N. Petra, U. Villa, M. Parno, Y. Marzouk, and O. Ghattas. *hIPPYlib-MUQ: Scalable Markov Chain Monte Carlo Sampling Methods for Large-scale Bayesian Inverse Problems Governed by PDEs* (Submitted to TOMS) https://arxiv.org/abs/2112.00713

https://github.com/hippylib/hippylib2muq

Summary

- Despite the fast and dimension-independent convergence of these advanced structure-exploiting MCMC methods, many Bayesian inverse problems governed by expensive-to-solve PDEs remain out of reach.
- In such cases, hIPPYlib-MUQ can be used as a prototyping environment to study new methods that further exploit problem structure, for example through
  1. the use of various reduced models/surrogates, or
  2. via advanced Hessian approximations that go beyond low rank.

Details in: K.T. Kim, N. Petra, U. Villa, M. Parno, Y. Marzouk, and O. Ghattas. *hIPPYlib-MUQ: Scalable Markov Chain Monte Carlo Sampling Methods for Large-scale Bayesian Inverse Problems Governed by PDEs* (Submitted to TOMS) https://arxiv.org/abs/2112.00713

https://github.com/hippylib/hippylib2muq