MathMet 2022

Nov 2-4, Paris, France

High-Dimensional Exponentiation with guaranteed Error Control for Bayesian Likelihood Approximation

N. Farchmin, P. Trunschke, M. Eigel, S. Heidenreich

November 3, 2022

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin

Weierstrass Institute for
Applied Analysis and Stochastics

Context and Motivation

Inference of hidden parameters: Bayesian Inverse Problem

Given a forward map $u\colon \mathbb{R}^M \to \mathbb{R}^J$ and noisy observations $\delta = u(y^*) + \eta \in \mathbb{R}^J$ with centered additive Gaussian noise $\eta \in \mathcal{N}(0, \Sigma)$, the Bayesian inverse problem reads

$$\pi_{y|\delta}(y) = Z^{-1} L(y;\delta)\pi_0(y)$$

with data likelihood

$$L(y;\delta) = \exp\left(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\right)$$

and normalization constant $Z = \mathbb{E}_{\pi_0}[L(\bullet;\delta)]$.

Context and Motivation

Inference of hidden parameters: Bayesian Inverse Problem

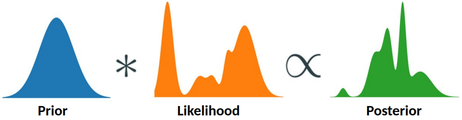
Given a forward map $u\colon \mathbb{R}^M \to \mathbb{R}^J$ and noisy observations $\delta = u(y^*) + \eta \in \mathbb{R}^J$ with centered additive Gaussian noise $\eta \in \mathcal{N}(0, \Sigma)$, the Bayesian inverse problem reads

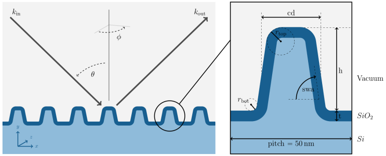
$$\pi_{y|\delta}(y) = Z^{-1} L(y;\delta)\pi_0(y)$$

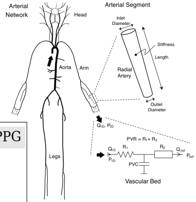
with data likelihood

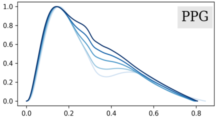
$$L(y;\delta) = \exp\left(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\right)$$

and normalization constant $Z = \mathbb{E}_{\pi_0}[L(\bullet;\delta)]$.









Context and Motivation

Inference of hidden parameters: Bayesian Inverse Problem

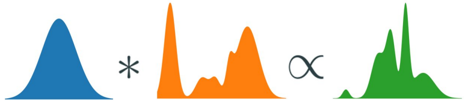
Given a forward map $u\colon \mathbb{R}^M \to \mathbb{R}^J$ and noisy observations $\delta = u(y^*) + \eta \in \mathbb{R}^J$ with centered additive Gaussian noise $\eta \in \mathcal{N}(0, \Sigma)$, the Bayesian inverse problem reads

$$\pi_{y|\delta}(y) = Z^{-1} L(y;\delta)\pi_0(y)$$

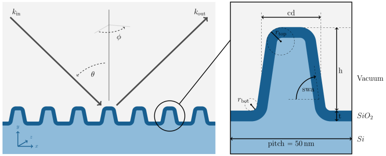
with data likelihood

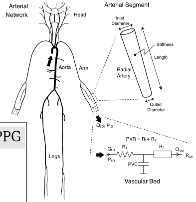
$$L(y;\delta) = \exp\left(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\right)$$

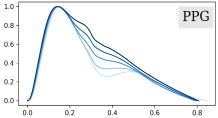
and normalization constant $Z = \mathbb{E}_{\pi_0}[L(\bullet;\delta)]$.



Prior Likelihood Posterior







Idea

If $h\colon \mathbb{R}^M \to \mathbb{R}$ is easy to approximate for $M \gg 1$ we can use this to construct $u = \exp(h)$

Approach and Setting for Exponentiation

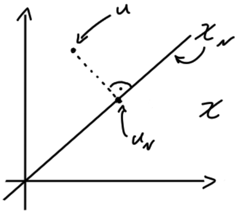
PDE and Galerkin projection

For $f(y) = \exp(h(y_0))\,\nabla h(y)$ and arbitrary $y_0 \in \mathbb{R}^M$, $u(y) = \exp(h(y)) - \exp(h(y_0))$ is the solution of

$$\nabla u - u\,\nabla h = f, \qquad \text{and} \qquad u(y_0) = 0.$$

The variational form for $B(w) = \nabla w - w\,\nabla h$ and $\mathcal{X} = \{w \in H^K(\mathbb{R}^M, \pi)\colon w(y_0) = 0\}$ reads:

$$\text{Find } u \in \mathcal{X} \text{ such that} \qquad \langle B(u)_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} = \langle f_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} \qquad \text{for all } m = 1, \dots, M \text{ and } v \in L^2(\mathbb{R}^M, \pi).$$



Approach and Setting for Exponentiation

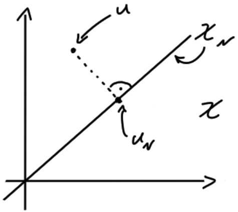
PDE and Galerkin projection

For $f(y) = \exp(h(y_0))\,\nabla h(y)$ and arbitrary $y_0 \in \mathbb{R}^M$, $u(y) = \exp(h(y)) - \exp(h(y_0))$ is the solution of

$$\nabla u - u\,\nabla h = f, \qquad \text{and} \qquad u(y_0) = 0.$$

The variational form for $B(w) = \nabla w - w\,\nabla h$ and $\mathcal{X} = \{w \in H^K(\mathbb{R}^M, \pi) \colon w(y_0) = 0\}$ reads:

$$\text{Find } u \in \mathcal{X} \text{ such that} \qquad \langle B(u)_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} = \langle f_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} \qquad \text{for all } m = 1, \ldots, M \text{ and } v \in L^2(\mathbb{R}^M, \pi).$$



original idea

$u(y) = \exp h(y)$ is the solution of the IVP

$$u'(y) = h'(y)u(y), \qquad \text{and} \qquad u(y_0) = \exp h(y_0)$$

Approach and Setting for Exponentiation

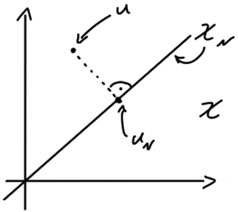
PDE and Galerkin projection

For $f(y) = \exp(h(y_0))\,\nabla h(y)$ and arbitrary $y_0 \in \mathbb{R}^M$, $u(y) = \exp(h(y)) - \exp(h(y_0))$ is the solution of

$$\nabla u - u\,\nabla h = f, \qquad \text{and} \qquad u(y_0) = 0.$$

The variational form for $B(w) = \nabla w - w\,\nabla h$ and $\mathcal{X} = \{w \in H^K(\mathbb{R}^M, \pi) \colon w(y_0) = 0\}$ reads:

$$\text{Find } u \in \mathcal{X} \text{ such that} \qquad \langle B(u)_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} = \langle f_m, v\rangle_{L^2(\mathbb{R}^M,\pi)} \qquad \text{for all } m = 1, \dots, M \text{ and } v \in L^2(\mathbb{R}^M, \pi).$$



original idea

$u(y) = \exp h(y)$ is the solution of the IVP

$$u'(y) = h'(y)u(y), \qquad \text{and} \qquad u(y_0) = \exp h(y_0)$$

different approach

We can also use a different ansatz: for $a, b \in \mathbb{R}$

$$u'' = [h'' + (h')^2]u, \qquad \text{and} \qquad u(a) = \exp h(a),\ \ u(b) = \exp h(b)$$

Energy Norm and Error Bounds

Lemma (energy norm)

$B\colon \mathcal{X} \to \mathcal{V}$ is injective (if $B(\mathcal{V}_a) \subseteq \mathcal{V}_t$) and $\|w\|_B := \|B(w)\|_{L^2(\pi)}$ is a norm on $\mathcal{X}$

Energy Norm and Error Bounds

Lemma (energy norm)

$B\colon \mathcal{X} \to \mathcal{V}$ is injective (if $B(\mathcal{V}_a) \subseteq \mathcal{V}_t$) and $\|w\|_B := \|B(w)\|_{L^2(\pi)}$ is a norm on $\mathcal{X}$

Relation to other norms

$$
\begin{aligned}
M = 1 \text{ and } h \in W^{1,\infty}(\mathbb{R}) \qquad &\rightsquigarrow \qquad \|u - w_N\|_B \leq C\|u - w_N\|_{H^1(\pi)} \\
(+) \quad \pi \text{ Gaussian and } |h'(y) - y/2| \geq \varepsilon \qquad &\rightsquigarrow \qquad \|u - w_N\|_{L^2(\mathbb{R},\pi)} \leq \varepsilon^{-1}\|u - w_N\|_B
\end{aligned}
$$

Energy Norm and Error Bounds

Lemma (energy norm)

$B\colon \mathcal{X} \to \mathcal{V}$ is injective (if $B(\mathcal{V}_a) \subseteq \mathcal{V}_t$) and $\|w\|_B := \|B(w)\|_{L^2(\pi)}$ is a norm on $\mathcal{X}$

Relation to other norms

$$
\begin{array}{rlcl}
 & M = 1 \text{ and } h \in W^{1,\infty}(\mathbb{R}) & \rightsquigarrow & \|u - w_N\|_B \leq C\|u - w_N\|_{H^1(\pi)} \\
(+) & \pi \text{ Gaussian and } |h'(y) - y/2| \geq \varepsilon & \rightsquigarrow & \|u - w_N\|_{L^2(\mathbb{R},\pi)} \leq \varepsilon^{-1}\|u - w_N\|_B
\end{array}
$$

Theorem (equivalence of continuous residual and energy error)

$u \in \mathcal{X}$ solution of $\langle B(u), v\rangle = \langle f, v\rangle$ for $f(y) = \exp h(y_0)\,\nabla h(y) \in \mathcal{V}_t$

$$
\|u - w_N\|_B = \|\boldsymbol{f} - \boldsymbol{B}\boldsymbol{w_N}\|_2 \qquad \text{for all } w_N \in \mathcal{V}_a
$$

Energy Norm and Error Bounds

Lemma (energy norm)

$B\colon \mathcal{X} \to \mathcal{V}$ is injective (if $B(\mathcal{V}_a) \subseteq \mathcal{V}_t$) and $\|w\|_B := \|B(w)\|_{L^2(\pi)}$ is a norm on $\mathcal{X}$

Relation to other norms

$$\begin{aligned} M = 1 \text{ and } h \in W^{1,\infty}(\mathbb{R}) \qquad &\rightsquigarrow \qquad \|u - w_N\|_B \leq C\|u - w_N\|_{H^1(\pi)} \\ (+) \quad \pi \text{ Gaussian and } |h'(y) - y/2| \geq \varepsilon \qquad &\rightsquigarrow \qquad \|u - w_N\|_{L^2(\mathbb{R},\pi)} \leq \varepsilon^{-1}\|u - w_N\|_B \end{aligned}$$

Theorem (equivalence of continuous residual and energy error)

$u \in \mathcal{X}$ solution of $\langle B(u), v\rangle = \langle f, v\rangle$ for $f(y) = \exp h(y_0)\,\nabla h(y) \in \mathcal{V}_t$

$$\|u - w_N\|_B = \|\boldsymbol{f} - \boldsymbol{B}\boldsymbol{w_N}\|_2 \qquad \text{for all } w_N \in \mathcal{V}_a$$

- equivalence (efficiency and reliability) with constant 1
- operators can be represented efficiently in high dimensions (Tensor Trains)
- no computational overhead when using Galerkin approach with TTs

Energy Norm and Error Bounds

Lemma (energy norm)

$B\colon \mathcal{X} \to \mathcal{V}$ is injective (if $B(\mathcal{V}_a) \subseteq \mathcal{V}_t$) and $\|w\|_B := \|B(w)\|_{L^2(\pi)}$ is a norm on $\mathcal{X}$

Relation to other norms

$$\begin{aligned} M = 1 \text{ and } h \in W^{1,\infty}(\mathbb{R}) \quad &\rightsquigarrow \quad \|u - w_N\|_B \leq C\|u - w_N\|_{H^1(\pi)} \\ (+) \quad \pi \text{ Gaussian and } |h'(y) - y/2| \geq \varepsilon \quad &\rightsquigarrow \quad \|u - w_N\|_{L^2(\mathbb{R},\pi)} \leq \varepsilon^{-1}\|u - w_N\|_B \end{aligned}$$

Theorem (equivalence of continuous residual and energy error)

$u \in \mathcal{X}$ solution of $\langle B(u), v\rangle = \langle f, v\rangle$ for $f(y) = \exp h(y_0)\,\nabla h(y) \in \mathcal{V}_t$

$$\|u - w_N\|_B = \|\boldsymbol{f} - \boldsymbol{B}\boldsymbol{w_N}\|_2 \qquad \text{for all } w_N \in \mathcal{V}_a$$

- equivalence (efficiency and reliability) with constant 1
- operators can be represented efficiently in high dimensions (Tensor Trains)
- no computational overhead when using Galerkin approach with TTs
- bounds hold for arbitrary discrete function
- check error in iterative schemes
- use theory to check error in other approximation schemes
- **similar for other holonomic functions** (polynomials, rational functions, sin, cos, Bessel functions, erf, ...)

Experiment: Exponentiation of log-Likelihood

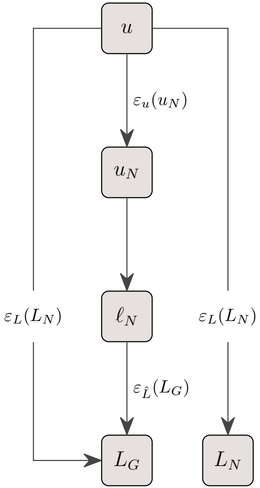
Forward map u solution to $-\nabla_x \cdot (\exp(\gamma(x,y))\,\nabla_x u(x,y)) = 1$, approximation u_N and observation δ, with $\Sigma = 10^{-6} I$, let

$$L(y) = \exp\Big(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\Big) \qquad \text{and} \qquad \hat{L}(y) = \exp\Big(\underbrace{-\frac{1}{2}\|\delta - u_N(y)\|^2_{\Sigma^{-1}}}_{\ell_N(y)}\Big).$$

For samples $\{y^{(i)}\}$ reconstruct u_N with $(y^{(i)}, u(y^{(i)}))$ and L_N with $(y^{(i)}, L(y^{(i)}))$.
Consider the errors

$$\mathrm{Res}(w_N) = \|\boldsymbol{f} - \boldsymbol{B}\boldsymbol{w_N}\|_2 \qquad \text{and} \qquad \varepsilon_w(w_N) = \frac{1}{N_{\mathrm{MC}}}\sum_{i=1}^{N_{\mathrm{MC}}} \frac{\|w(y^{(i)}) - w_N(y^{(i)})\|}{\|w(y^{(i)})\|}.$$

M	$\mathrm{Res}(L_N)$	$\mathrm{Res}(L_G)$	$\varepsilon_u(u_N)$	$\varepsilon_L(L_N)$	$\varepsilon_L(L_G)$	$\varepsilon_{\hat{L}}(L_G)$
5	$8.23 \cdot 10^{-6}$	$8.46 \cdot 10^{-7}$	$1.86 \cdot 10^{-2}$	$1.19 \cdot 10^{-4}$	$2.57 \cdot 10^{-4}$	$2.79 \cdot 10^{-6}$
10	$1.06 \cdot 10^{-5}$	$1.06 \cdot 10^{-6}$	$1.87 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$2.69 \cdot 10^{-4}$	$6.75 \cdot 10^{-6}$
20	$1.18 \cdot 10^{-5}$	$8.12 \cdot 10^{-7}$	$2.02 \cdot 10^{-2}$	$1.06 \cdot 10^{-4}$	$2.97 \cdot 10^{-4}$	$4.97 \cdot 10^{-6}$
30	$1.59 \cdot 10^{-5}$	$2.96 \cdot 10^{-6}$	$2.26 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$3.68 \cdot 10^{-4}$	$2.64 \cdot 10^{-6}$
40	$1.92 \cdot 10^{-2}$	$5.40 \cdot 10^{-3}$	$2.32 \cdot 10^{-2}$	$1.36 \cdot 10^{-4}$	$7.12 \cdot 10^{-4}$	$3.14 \cdot 10^{-5}$



Experiment: Exponentiation of log-Likelihood

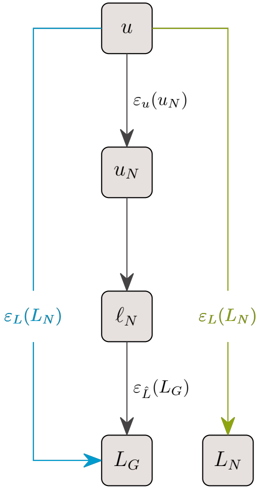
Forward map u solution to $-\nabla_x \cdot (\exp(\gamma(x,y))\, \nabla_x u(x,y)) = 1$, approximation u_N and observation δ, with $\Sigma = 10^{-6} I$, let

$$L(y) = \exp\left(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\right) \qquad \text{and} \qquad \hat{L}(y) = \exp\Big(\underbrace{-\frac{1}{2}\|\delta - u_N(y)\|^2_{\Sigma^{-1}}}_{\ell_N(y)}\Big).$$

For samples $\{y^{(i)}\}$ reconstruct u_N with $(y^{(i)}, u(y^{(i)}))$ and L_N with $(y^{(i)}, L(y^{(i)}))$.
Consider the errors

$$\mathrm{Res}(w_N) = \|\boldsymbol{f} - \boldsymbol{B} \boldsymbol{w_N}\|_2 \qquad \text{and} \qquad \varepsilon_w(w_N) = \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} \frac{\|w(y^{(i)}) - w_N(y^{(i)})\|}{\|w(y^{(i)})\|}.$$

M	$\mathrm{Res}(L_N)$	$\mathrm{Res}(L_G)$	$\varepsilon_u(u_N)$	$\varepsilon_L(L_N)$	$\varepsilon_L(L_G)$	$\varepsilon_{\hat{L}}(L_G)$
5	$8.23 \cdot 10^{-6}$	$8.46 \cdot 10^{-7}$	$1.86 \cdot 10^{-2}$	$1.19 \cdot 10^{-4}$	$2.57 \cdot 10^{-4}$	$2.79 \cdot 10^{-6}$
10	$1.06 \cdot 10^{-5}$	$1.06 \cdot 10^{-6}$	$1.87 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$2.69 \cdot 10^{-4}$	$6.75 \cdot 10^{-6}$
20	$1.18 \cdot 10^{-5}$	$8.12 \cdot 10^{-7}$	$2.02 \cdot 10^{-2}$	$1.06 \cdot 10^{-4}$	$2.97 \cdot 10^{-4}$	$4.97 \cdot 10^{-6}$
30	$1.59 \cdot 10^{-5}$	$2.96 \cdot 10^{-6}$	$2.26 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$3.68 \cdot 10^{-4}$	$2.64 \cdot 10^{-6}$
40	$1.92 \cdot 10^{-2}$	$5.40 \cdot 10^{-3}$	$2.32 \cdot 10^{-2}$	$1.36 \cdot 10^{-4}$	$7.12 \cdot 10^{-4}$	$3.14 \cdot 10^{-5}$



Experiment: Exponentiation of log-Likelihood

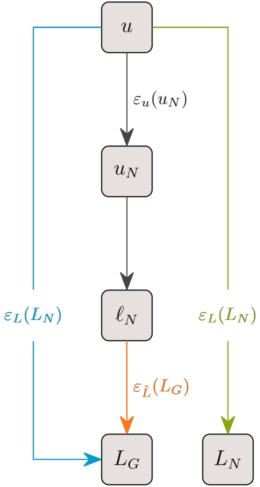
Forward map u solution to $-\nabla_x \cdot (\exp(\gamma(x,y))\,\nabla_x u(x,y)) = 1$, approximation u_N and observation δ, with $\Sigma = 10^{-6} I$, let

$$L(y) = \exp\Big(-\frac{1}{2}\|\delta - u(y)\|^2_{\Sigma^{-1}}\Big) \qquad \text{and} \qquad \hat{L}(y) = \exp\Big(\underbrace{-\frac{1}{2}\|\delta - u_N(y)\|^2_{\Sigma^{-1}}}_{\ell_N(y)}\Big).$$

For samples $\{y^{(i)}\}$ reconstruct u_N with $(y^{(i)}, u(y^{(i)}))$ and L_N with $(y^{(i)}, L(y^{(i)}))$.
Consider the errors

$$\mathrm{Res}(w_N) = \|\boldsymbol{f} - \boldsymbol{B}\boldsymbol{w}_N\|_2 \qquad \text{and} \qquad \varepsilon_w(w_N) = \frac{1}{N_{\mathrm{MC}}}\sum_{i=1}^{N_{\mathrm{MC}}} \frac{\|w(y^{(i)}) - w_N(y^{(i)})\|}{\|w(y^{(i)})\|}.$$

M	$\mathrm{Res}(L_N)$	$\mathrm{Res}(L_G)$	$\varepsilon_u(u_N)$	$\varepsilon_L(L_N)$	$\varepsilon_L(L_G)$	$\varepsilon_{\hat{L}}(L_G)$
5	$8.23 \cdot 10^{-6}$	$8.46 \cdot 10^{-7}$	$1.86 \cdot 10^{-2}$	$1.19 \cdot 10^{-4}$	$2.57 \cdot 10^{-4}$	$2.79 \cdot 10^{-6}$
10	$1.06 \cdot 10^{-5}$	$1.06 \cdot 10^{-6}$	$1.87 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$2.69 \cdot 10^{-4}$	$6.75 \cdot 10^{-6}$
20	$1.18 \cdot 10^{-5}$	$8.12 \cdot 10^{-7}$	$2.02 \cdot 10^{-2}$	$1.06 \cdot 10^{-4}$	$2.97 \cdot 10^{-4}$	$4.97 \cdot 10^{-6}$
30	$1.59 \cdot 10^{-5}$	$2.96 \cdot 10^{-6}$	$2.26 \cdot 10^{-2}$	$1.13 \cdot 10^{-4}$	$3.68 \cdot 10^{-4}$	$2.64 \cdot 10^{-6}$
40	$1.92 \cdot 10^{-2}$	$5.40 \cdot 10^{-3}$	$2.32 \cdot 10^{-2}$	$1.36 \cdot 10^{-4}$	$7.12 \cdot 10^{-4}$	$3.14 \cdot 10^{-5}$



Conclusion

Summary:

- non-intrusive (only requires TT represenation of h)
- freedom in **choice of PDE** (e.g. second-order problem) and thus of energy norm
- problem adaptable choice of initial condition / RHS of PDE
- free, **reliable and efficient** error estimator for any discrete function
- generalization to other **holonomic functions** (algebraic functions, sin & cos, sinh & cosh, $\log_b$, erf, generalized hypergeometric function, Bessel functions, ...)
- analysis easily extendible to other functions / PDEs

Conclusion

Summary:

- non-intrusive (only requires TT represenation of h)
- freedom in **choice of PDE** (e.g. second-order problem) and thus of energy norm
- problem adaptable choice of initial condition / RHS of PDE
- free, **reliable and efficient** error estimator for any discrete function
- generalization to other **holonomic functions** (algebraic functions, sin & cos, sinh & cosh, $\log_b$, erf, generalized hypergeometric function, Bessel functions, . . .)
- analysis easily extendible to other functions / PDEs

EFFICIENT APPROXIMATION OF HIGH-DIMENSIONAL EXPONENTIALS BY TENSOR NETWORKS

Volume 13, Issue 1, 2023, pp. 25-51
DOI: 10.1615/Int.J.UncertaintyQuantification.2022039164

Link to Article

Conclusion

Summary:

- non-intrusive (only requires TT represenation of h)
- freedom in **choice of PDE** (e.g. second-order problem) and thus of energy norm
- problem adaptable choice of initial condition / RHS of PDE
- free, **reliable and efficient** error estimator for any discrete function
- generalization to other **holonomic functions** (algebraic functions, sin & cos, sinh & cosh, $\log_b$, erf, generalized hypergeometric function, Bessel functions, ...)
- analysis easily extendible to other functions / PDEs

EFFICIENT APPROXIMATION OF HIGH-DIMENSIONAL EXPONENTIALS BY TENSOR NETWORKS

Volume 13, Issue 1, 2023, pp. 25-51
DOI: 10.1615/Int.J.UncertaintyQuantification.2022039164

Link to Article

PyThia

Git: gitlab1.ptb.de/pythia/pythia

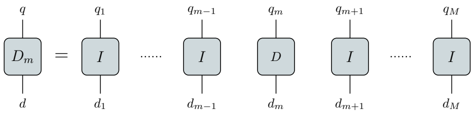
Doc: pythia-uq.rtfd.io

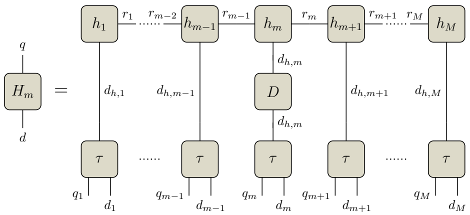
Efficient Implementation in Tensor Train Format

Use first-order optimality criterion:

$$u_N = \underset{w_N \in \mathcal{V}_a \cap \mathcal{X}}{\arg\min} \; \|\boldsymbol{f} - \boldsymbol{B} w_N\|_2^2 \quad \rightsquigarrow \quad \boldsymbol{W} u_N = \boldsymbol{b} \quad \text{with} \quad \boldsymbol{W} := \sum_{m=1}^{M} \boldsymbol{B}_m^{\mathsf{T}} \boldsymbol{B}_m \quad \text{and} \quad \boldsymbol{b} := \sum_{m=1}^{M} \boldsymbol{B}_m^{\mathsf{T}} \boldsymbol{f}_m$$

We have $\boldsymbol{B}_m = \boldsymbol{D}_m - \boldsymbol{H}_m$ and $\boldsymbol{f}_m \sim \boldsymbol{D}_m \boldsymbol{h}$ with:





Hence $\operatorname{rank}(\boldsymbol{B}_m) = \operatorname{rank}(h) + 1$ and $\operatorname{rank}(\boldsymbol{f}_m) = \operatorname{rank}(h)$

Laplace-like structure of $\boldsymbol{W}$ and $\boldsymbol{b}$ yields: $\operatorname{rank}(\boldsymbol{W}) = 2\,(\operatorname{rank}(h) + 1)^2 + 1$ and $\operatorname{rank}(\boldsymbol{b}) = 2\operatorname{rank}(h)\,(\operatorname{rank}(h) + 1)$

Assembly of Discrete System

Basic TT Operators:

i) differentiation: $\boldsymbol{D}_m = \boldsymbol{I}^{\otimes(m-1)} \otimes \boldsymbol{D} \otimes \boldsymbol{I}^{\otimes(M-m)}$ with $\boldsymbol{D}[i,j] := \langle p_i, p_j' \rangle_{L^2(\mathbb{R},\pi_m)}$

ii) multiplication by $\partial_m h$: $\boldsymbol{H}_m[\mu,\nu] = \sum_{k=1}^{r} \prod_{j=1}^{M} \boldsymbol{H}_{m,j}[k_j,\mu_j,\nu_j,k_{j+1}]$ with $\boldsymbol{\tau}_{i,j,k} = \langle p_i p_j, p_k \rangle_{L^2(\mathbb{R},\pi_m)}$

$$\boldsymbol{H}_{m,j}[k_j,\mu_j,\nu_j,k_{j+1}] = \sum_{i=1}^{d_{\mathrm{h}}} \boldsymbol{\tau}_{\mu_j,\nu_j,i} \boldsymbol{h}_j[k_j,i,k_{j+1}] \quad \text{and} \quad \boldsymbol{H}_{m,m}[k_m,\mu_m,\nu_m,k_{m+1}] = \sum_{i_1,i_2=1}^{d_{\mathrm{h}}} \boldsymbol{\tau}_{\mu_m,\nu_m,i_1} \boldsymbol{D}[i_1,i_2] \boldsymbol{h}_m[k_m,i_2,k_{m+1}]$$

Partial Derivative Operators: $\boldsymbol{B}_m := \boldsymbol{D}_m - \boldsymbol{H}_m$ and $\boldsymbol{f}_m := \exp\left(h(y_0)\right) \boldsymbol{D}_m \boldsymbol{h}$

Quadratic System: $\boldsymbol{W}\boldsymbol{u} = \boldsymbol{b}$ for $\boldsymbol{W} := \boldsymbol{P}\boldsymbol{P}^\intercal + \sum_{m=1}^{M} \boldsymbol{B}_m^\intercal \boldsymbol{B}_m$ and $\boldsymbol{b} := \sum_{m=1}^{M} \boldsymbol{B}_m^\intercal \boldsymbol{f}_m$

i) rank-1 basis evaluation tensor: $\boldsymbol{P} \in \mathbb{R}^{d_a^M}$ given by $\boldsymbol{P}[\mu] = P_\mu(y_0)$

ii) sum of partial derivative operators: $\boldsymbol{S} = \sum_{m=1}^{M} \boldsymbol{B}_m^\intercal \boldsymbol{B}_m$ $\quad$ ($\boldsymbol{C}_j := \boldsymbol{B}_{m,j}$ for any $m \neq j$)

$$\boldsymbol{S}_1 = \begin{bmatrix} \boldsymbol{B}_{1,1}^\intercal \boldsymbol{B}_{1,1} & \boldsymbol{C}_1^\intercal \boldsymbol{C}_1 \end{bmatrix}, \quad \boldsymbol{S}_j = \begin{bmatrix} \boldsymbol{C}_j^\intercal \boldsymbol{C}_j & 0 \\ \boldsymbol{B}_{j,j}^\intercal \boldsymbol{B}_{j,j} & \boldsymbol{C}_j^\intercal \boldsymbol{C}_j \end{bmatrix}, \quad \text{and} \quad \boldsymbol{S}_M = \begin{bmatrix} \boldsymbol{C}_M^\intercal \boldsymbol{C}_M \\ \boldsymbol{B}_{M,M}^\intercal \boldsymbol{B}_{M,M} \end{bmatrix}$$

iii) sum of partial derivative RHS:

$$\boldsymbol{b}_1 = \begin{bmatrix} \boldsymbol{B}_{1,1}^\intercal \boldsymbol{f}_{1,1} & \boldsymbol{C}_1^\intercal \boldsymbol{g}_1 \end{bmatrix}, \quad \boldsymbol{b}_j = \begin{bmatrix} \boldsymbol{C}_j^\intercal \boldsymbol{g}_j & 0 \\ \boldsymbol{B}_{j,j}^\intercal \boldsymbol{f}_{j,j} & \boldsymbol{C}_j^\intercal \boldsymbol{g}_j \end{bmatrix} \quad \text{and} \quad \boldsymbol{b}_M = \begin{bmatrix} \boldsymbol{C}_M^\intercal \boldsymbol{g}_M \\ \boldsymbol{B}_{M,M}^\intercal \boldsymbol{f}_{M,M} \end{bmatrix}$$

Second Order System and Energy Norm

For $y \sim \mathcal{U}([0,1])$ consider the second order system

$$w'' = (h'' + (h')^2)\, w \qquad \text{with} \qquad w(y_1) = \exp h(y_1) \text{ and } w(y_2) = \exp h(y_2).$$

Let $\tilde{h} = h'' + (h')^2$ and

$$c(y) = \exp\left(h(y_1)\right)\frac{y_2 - y}{y_1 - y_2} + \exp\left(h(y_2)\right)\frac{y - y_1}{y_1 - y_2}.$$

Then $c'' = 0$, $c(y_1) = -\exp\left(h(y_1)\right)$ and $c(y_2) = -\exp\left(h(y_2)\right)$ and it follows, that $u = w + c$ solves

$$-u'' + \tilde{h}\, u = f \qquad \text{with} \qquad u(y_1) = u(y_2) = 0$$

for $f = \tilde{h}\, c$. Assuming $0 < \check{h} \leq \tilde{h}(y) \leq \hat{h} < \infty$ for a.a. $y \in [0,1]$, the weak formulation reads

$$B(u,v) = \langle f, v\rangle \qquad \text{for} \qquad B(u,v) = \langle u', v'\rangle + \langle \tilde{h}\, u, v\rangle.$$

For the energy norm induced by B it holds

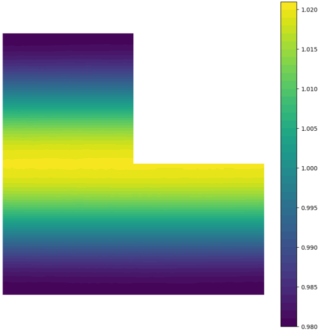
$$\min\{1, \check{h}\}\|w\|_{H^1} \leq \|w\|_B \leq \min\{1, \hat{h}\}\|w\|_{H^1}.$$

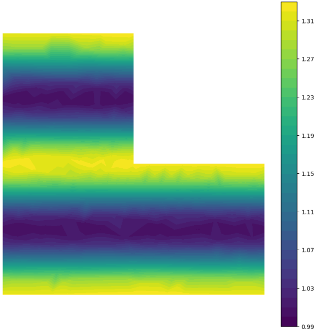
Diffusion Coefficient of Lognormal Darcy Equation

As a model problem we use $a(x, y) = \exp\left(\sum_{\ell=1}^{L} \gamma_\ell(x) y_\ell\right)$ with

$$\gamma_\ell(x) = \frac{9}{10\zeta(2)} \ell^{-2} \cos\big(2\pi\beta_1(\ell)x_1\big)\ \cos\big(2\pi\beta_2(\ell)x_2\big),$$

where $\beta_1(\ell) = \ell - k(\ell)\frac{k(\ell)+1}{2}$ and $\beta_2(\ell) = k(\ell) - \beta_1(\ell)$ for $k(\ell) = \lfloor -\frac{1}{2} + \sqrt{\frac{1}{4} + 2\ell} \rfloor$.

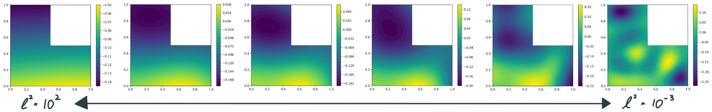




Lognormal Field with given Covariance Length

Let $a(x,\omega) = \exp(\gamma(x,\omega))$ with centered Gaussian random field γ with covariance

$$\mathrm{Cov}_\gamma(x,z) := \frac{1}{100}\exp\left(-\ell^{-2}\|x-z\|_2^2\right).$$



ℓ^2	$\varepsilon_{\gamma_{\hat M}}(\gamma_M)$	$r_{\max}(a_{\mathrm{PDE}})$	$\mathrm{res}(a_{\mathrm{PDE}})$	$\varepsilon_{a_M}(a_{\mathrm{PDE}})$	$r_{\max}(a_{\mathrm{VMC}})$	$\mathrm{res}(a_{\mathrm{VMC}})$	$\varepsilon_{a_M}(a_{\mathrm{VMC}})$
10	$1.59 \cdot 10^{-7}$	21	$1.62 \cdot 10^{-4}$	$4.10 \cdot 10^{-7}$	9	$4.97 \cdot 10^{-1}$	$6.59 \cdot 10^{-3}$
5	$9.30 \cdot 10^{-7}$	21	$1.99 \cdot 10^{-4}$	$1.68 \cdot 10^{-6}$	7	$5.51 \cdot 10^{-1}$	$3.01 \cdot 10^{-3}$
1	$5.79 \cdot 10^{-5}$	38	$1.31 \cdot 10^{-4}$	$8.04 \cdot 10^{-6}$	8	$4.87 \cdot 10^{-1}$	$1.05 \cdot 10^{-2}$
0.5	$3.27 \cdot 10^{-4}$	52	$2.10 \cdot 10^{-4}$	$2.69 \cdot 10^{-5}$	7	$3.09 \cdot 10^{-1}$	$1.35 \cdot 10^{-2}$
0.1	$8.57 \cdot 10^{-3}$	114	$5.28 \cdot 10^{-4}$	$3.21 \cdot 10^{-5}$	13	$1.30 \cdot 10^{0}$	$6.17 \cdot 10^{-2}$

Table 1: L-shaped domain with CG-1 FEM (3017 DoFs) for $M = 20$ and $d_{\mathrm{a}} = 10$. Also $\gamma_{\hat M}$ is computed for $\hat M = 100$.