THE PARALLEL LIBRARY QUESO FOR UNCERTAINTY QUANTIFICATION

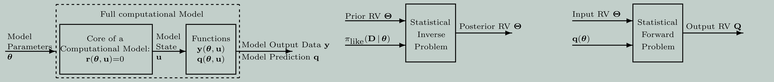
E. E. Prudencio, K. C. Estacio-Hiroms, K. W. Schulz

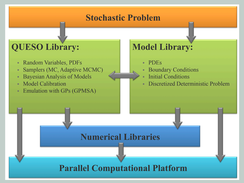
Acknowledgments

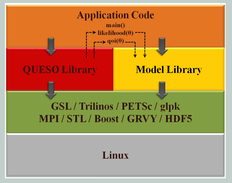
- Main collaborators:
  - In DOE laboratories: B. Adams (SNL), M. Eldred (SNL), J. Gattiker (LANL), D. Higdon (LANL), L. P. Swiler (SNL), B. Williams (LANL).
  - In academia: P. Bauman (UT), S. H. Cheung (NTU), F. Nobile (EPFL), G. Terejanu (USC), R. Tempone (KAUST).
- Sponsors:
  - 2008 - 2013: DOE NNSA, PSAAP Program
  - 2010 - 2011: DOE SNL-NM, Peridynamics Team
  - 2010 - 2012: KAUST, AEA2 Program
  - 2011 - 2013: AFOSR, RTC, DDDAS Program
  - 2011 - 2015: DOE SC, SciDAC3 Program
  - 2012 - 2013: DOE LANL and ORNL, CASL Project
  - 2012 - 2014: KAUST, AEA3 Program

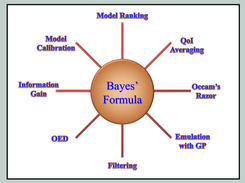
Objective

The QUESO library [PS12], from UT Austin, is a collection of parallel statistical algorithms and object-oriented programming constructs supporting research into the uncertainty quantification of mathematical models. Its development began in 2008, under the auspices of the Predictive Science Academic Alliance Program (PSAAP) of NNSA, DOE. Since 2011, QUESO has been part of the QUEST Institute, under the SciDAC3 program of SC, DOE. Some of its functionalities are also available to the DAKOTA toolkit [Eld13] from DOE.









Core Features: Object-Oriented and Parallel

Bayes' formula: posterior knowledge = data + model + prior knowledge

$$\underbrace{\pi_{\text{post}}(\boldsymbol{\theta}|\mathbf{d})}_{\text{posterior PDF}} = \frac{\overbrace{f(\mathbf{d}|\boldsymbol{\theta})}^{\text{likelihood}} \cdot \overbrace{\pi_{\text{prior}}(\boldsymbol{\theta})}^{\text{prior PDF}}}{\pi_{\text{data}}(\mathbf{d})}$$

The Likelihood:

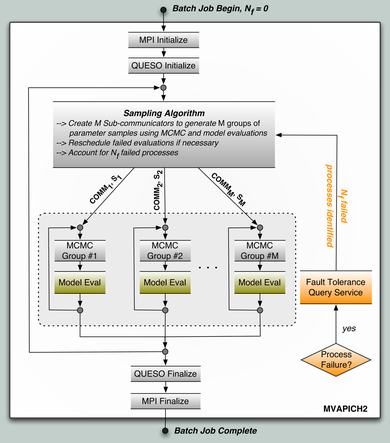
measurement + error = unknown reality = model output + model discrepancy

$$\mathbf{d}_{N\times 1} + \epsilon_{\text{data}} = \text{unknown reality} = [\mathbf{y}(\boldsymbol{\theta})]_{N\times 1} + \epsilon_{\text{model}}$$

$$\epsilon_{\text{data}} \sim \mathcal{N}(\mathbf{0}_{N\times 1}, \sigma^2_{\text{data}} \cdot \mathbf{I}_{N\times N});\ \epsilon_{\text{model}} \sim \mathcal{N}(\mathbf{0}_{N\times 1}, \sigma^2_{\text{model}} \cdot \mathbf{I}_{N\times N})$$

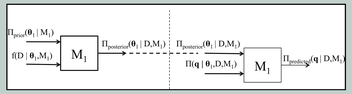
$$\sigma^2 \equiv \sigma^2_{\text{data}} + \sigma^2_{\text{model}} \Rightarrow \pi_{\text{like}}(\mathbf{d}|\boldsymbol{\theta}) \propto \frac{1}{\sqrt{\sigma^2 N}} \cdot \exp\left\{-\frac{1}{2}\frac{\|\mathbf{d}-\mathbf{y}(\boldsymbol{\theta})\|^2_2}{\sigma^2}\right\}$$

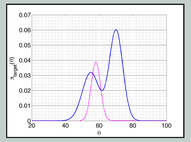
Problem Type	Main Entities
Deterministic	Vector and Matrix
Statistical	Vector Random Variable Joint Probability Density Realizer Covariance Matrix Random Field Covariance Function

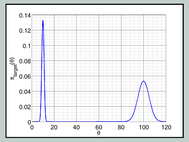


Sampling of (Multimodal) Distributions

Parallel multi-level algorithm [PC12]







Computation of Model Evidences

Bayes' Formula with Modeling Hypotheses

Intra Hypothesis:

$$\underbrace{\pi(\boldsymbol{\theta}_j|\mathbf{d},M_j)}_{\text{posterior PDF}} = \frac{\overbrace{f(\mathbf{d}|\boldsymbol{\theta}_j,M_j)}^{\text{likelihood}} \cdot \overbrace{\pi(\boldsymbol{\theta}_j|M_j)}^{\text{prior PDF}}}{\pi(\mathbf{d}|M_j)} = \frac{f(\mathbf{d}|\boldsymbol{\theta}_j,M_j)\,\pi(\boldsymbol{\theta}_j|M_j)}{\int f(\mathbf{d}|\boldsymbol{\theta}_j,M_j)\,\pi(\boldsymbol{\theta}_j|M_j)\,d\boldsymbol{\theta}_j}$$

Inter Hypotheses:

$$\underbrace{p(M_j|\mathbf{d},\mathcal{M})}_{\text{posterior plausibility}} = \frac{\overbrace{\pi(\mathbf{d}|M_j)}^{\text{evidence}} \cdot \overbrace{p(M_j|\mathcal{M})}^{\text{prior plausibility}}}{\pi(\mathbf{d}|\mathcal{M})} = \frac{\pi(\mathbf{d}|M_j)\,p(M_j|\mathcal{M})}{\sum_{j=1}^{m}\pi(\mathbf{d}|M_j)\,p(M_j|\mathcal{M})}$$

Extra Benefit: Information Theoretic Interpretation (Beck, Cheung, Ching, Muto, Yuen, 2005-2007):

$$\underbrace{\ln[\pi(\mathbf{d}|M_j)]}_{\text{log evidence}} = \underbrace{E\left[\ln[\pi(\mathbf{d}|\boldsymbol{\theta}_j,M_j)]\right]}_{\text{how well the hypothesis fits the data}} - \underbrace{E\left[\ln\frac{\pi(\boldsymbol{\theta}_j|\mathbf{d},M_j)}{\pi(\boldsymbol{\theta}_j|M_j)}\right]}_{\text{how much the hypothesis "learns" with the data}}$$

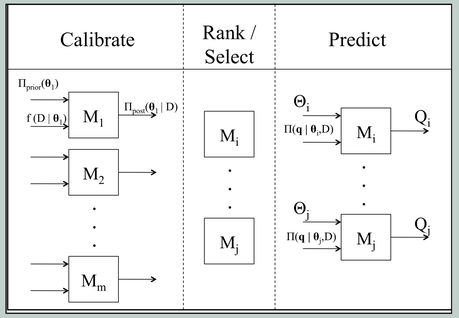
(expected information gain; EIG; model complexity; overfitting; Occam's razor)

Extra Benefit: Averaging of QoI Predictions:
With one hypothesis:

$$\pi(Q|\mathbf{d},M_j) = \int \pi(Q|\boldsymbol{\theta}_j,\mathbf{d},M_j)\,\overbrace{\pi(\boldsymbol{\theta}_j|\mathbf{d},M_j)}^{\text{"weight"}}\,d\boldsymbol{\theta}_j$$

With multiple hypotheses (averaging of QoI predictions):

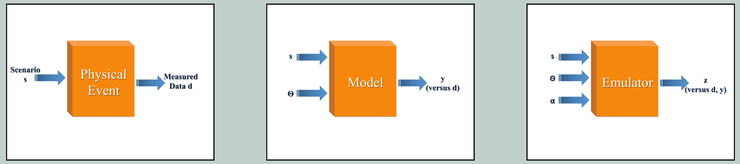
$$\pi(Q|\mathbf{d},\mathcal{M}) = \sum_{j=1}^{m}\pi(Q|\mathbf{d},M_j)\,\overbrace{p(M_j|\mathbf{d},\mathcal{M})}^{\text{"weight"}}$$



M_j	Log Evidence $\ln(\pi(\mathbf{d}\|M_j))$	Model Fitness $E[\ln(\pi_{\text{like}}(\cdot))]$	Model Complexity $E\left[\ln\frac{\pi_{\text{post}}(\cdot)}{\pi_{\text{prior}}(\cdot)}\right]$	Prior Plausibility $p(M_j\|\mathcal{M})$	Posterior Plausibility $p(M_j\|\mathbf{d},\mathcal{M})$
M_1	18.74	27.50	8.76	0.5	≈1
M_2	-6.41	7.48	13.89	0.5	≈0

Emulation with Gaussian Processes

Following GPMSA package from LANL [HGWR08]



Bayesian Filtering

Prediction (obtain a prior knowledge of the state):

$$\pi_{\text{post}}(\boldsymbol{\theta}^{(k)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)}) \longrightarrow \pi_{\text{prior}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)})$$

$$\pi_{\text{prior}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)}) = \int \pi_{\text{state}}(\boldsymbol{\theta}^{(k+1)}|\boldsymbol{\theta}^{(k)},\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)}) \cdot \pi_{\text{post}}(\boldsymbol{\theta}^{(k)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)})\,d\boldsymbol{\theta}^{(k)}$$

Correction (obtain a posterior knowledge of the state):

$$\pi_{\text{prior}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)}) \longrightarrow \pi_{\text{post}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k+1)},\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)})$$

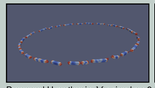
$$\pi_{\text{post}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k+1)},\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)}) = \frac{\pi_{\text{like}}(\mathbf{d}^{(k+1)}|\boldsymbol{\theta}^{(k+1)}) \cdot \pi_{\text{prior}}(\boldsymbol{\theta}^{(k+1)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)})}{\pi_{\text{data}}(\mathbf{d}^{(k+1)}|\mathbf{d}^{(k)},\ldots,\mathbf{d}^{(1)})}.$$

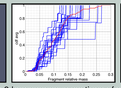
Other Ongoing and Planned Features

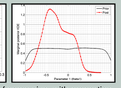
- Extreme-scale computing: load balancing and resilience (motivation for fault-tolerant versions of MPI).
- High dimensional parameter spaces: motivation to represent parallel vectors of parameters.
- Data mining and machine learning: motivation for nonparametric Bayesian.

Material Science

Peridynamics:







Proposed Hypothesis M_2: ring has 8 homogeneous sections of same size, with properties following a Gaussian distribution.

$$\begin{pmatrix} \log_{10}\left(\frac{s_{0,1}}{s_{0,\text{ref}}}\right) \\ \vdots \\ \log_{10}\left(\frac{s_{0,8}}{s_{0,\text{ref}}}\right) \\ \hline \log_{10}\left(\frac{Y_{\text{s}}}{Y_{\text{ref}}}\right) \\ \vdots \\ \log_{10}\left(\frac{Y_{\text{s}}}{Y_{\text{ref}}}\right) \end{pmatrix} \sim \begin{pmatrix} \mathcal{N}(\theta_1 \cdot \mathbf{1}_8, \mathbf{C}(a,\ell)) \\ \hline \mathcal{N}(\theta_2 \cdot \mathbf{1}_8, \mathbf{C}(a,\ell)) \end{pmatrix};\ s_{0,\text{ref}} = 0.18;\ Y_{\text{ref}} = 0.00176.$$

$$\mathbf{C}(a,\ell)_{ij} = a^2 \cdot \exp\left\{-\frac{1}{2}\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\ell}\right)^2\right\},\quad 1 \leqslant i,j \leqslant 8$$

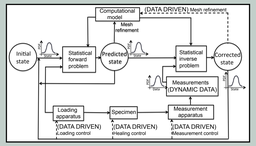
Likelihood for M_2:

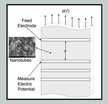
$$f(\mathbf{d}|\boldsymbol{\theta}_2,M_2) = \int_{\mathbb{R}^8}\int_{\mathbb{R}^8} f(\mathbf{d}|\mathbf{s}_0,\boldsymbol{Y},\boldsymbol{\theta}_2,M_2) \cdot \pi(\mathbf{s}_0,\boldsymbol{Y}|\boldsymbol{\theta}_2,M_2)\,d\mathbf{s}_0\,d\boldsymbol{Y}$$

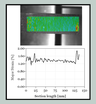
$$\approx \frac{1}{N_{\text{MC}}} \cdot \sum_{k=1}^{N_{\text{MC}}} f(\mathbf{d}|\mathbf{s}_0^{(k)},\boldsymbol{Y}^{(k)},\boldsymbol{\theta}_2,M_2)$$

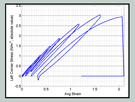
$$\propto \frac{1}{N_{\text{MC}}} \cdot \sum_{k=1}^{N_{\text{MC}}} \frac{1}{\sqrt{\sigma^2 N}} \cdot \exp\left\{-\frac{1}{2}\frac{\|\mathbf{d} - \mathbf{y}_2(\mathbf{s}_0^{(k)},\boldsymbol{Y}^{(k)},\boldsymbol{\theta}_2)\|^2_2}{\sigma^2}\right\}$$

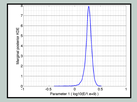
Stochastic Damage Monitoring:





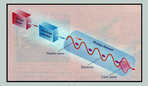




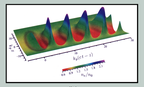


High Energy Physics

In collaboration with Tech-X, under the SciDAC3-HEP Community Petascale Project for Accelerator Science and Simulation (ComPASS).
Laser Plasma Acceleration (LPA); Direct Laser Acceleration (DLA).



(a)

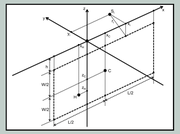


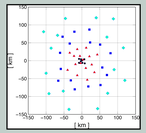
(b)

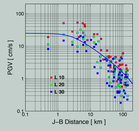
(a): W. Leemans and E. Esarey, "Laser-Driven Plasma-Wave Electron Accelerators", Physics Today, 62(3), pp. 44–29, 2009
(b): C. G. R. Geddes et al., "Large Fields for Smaller Facility Sources", SciDAC Review, Summer 2009, www.scidacreview.org

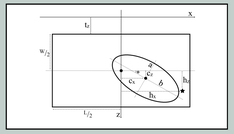
Accelerating fields: > 100 GV/m; Final energy: 100 MeV ⇔ 1 mm; 1 GeV ⇔ cm's; 10 GeV ⇔ 0.5 m.

Earthquake Seismology









Various Other Models

- Nuclear reactors, in collaboration with ORNL-LANL-SNL, under the Consortium for Advanced Simulation of Light Water Reactors (CASL).
- Tumor growth, turbulence, aerothermochemistry, chemical reactions, thermocoupling, ...

References

[Eld13] Michael S. Eldred et al.
DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis.
http://dakota.sandia.gov/, 1994-2013.

[HGWR08] D. Higdon, J. Gattiker, B. Williams, and M. Rightley.
Computer model calibration using high dimensional output.
Journal of the American Statistical Association, 103:570–583, 2008.

[PC12] E. E. Prudencio and S. H. Cheung.
Parallel adaptive multilevel sampling algorithms for the Bayesian analysis of mathematical models.
International Journal for Uncertainty Quantification, 2(3):215–237, 2012.

[PS12] E. E. Prudencio and K. W. Schulz.
The Parallel C++ Statistical Library QUESO: Quantification of Uncertainty for Estimation, Simulation and Optimization.
In M. Alexander et al., editor, *Euro-Par 2011 Workshops, Part I*, volume 7155 of *Lecture Notes in Computer Science*, pages 398–407. Springer-Verlag, Berlin Heidelberg, 2012.

prudenci@ices.utexas.edu
http://www.ices.utexas.edu
http://www.quest-scidac.org/