EuroHPC Joint Undertaking

Data driven Computational Mechanics at EXascale

DCoMEX

Data driven Computational Mechanics at EXascale

Work program topic: EuroHPC-01-2019
Type of action: Research and Innovation Action (RIA)

COLLABORATION PLAN

DELIVERABLE D4.1

Version No 2

https://www.dcomex.eu
This project has received funding from the European High-Performance Computing Joint Undertaking Joint Undertaking ('the JU'), under Grant Agreement No 956201

DOCUMENTATION SUMMARY INFORMATION

Project Title	**Data driven Computational Mechanics at EXascale**
Project Acronym	DCoMEX
Project No	956201
Call Identifier	EuroHPC-01-2019
Project Start Date	01/04/2021
Related work package	WP 4
Related task(s)	Task 4.1
Lead Organisation	ETHZ
Submission date	06/03/2023
Re-submission date	
Dissemination Level	PU

Quality Control:

	Who	**Affiliation**	**Date**
Checked by internal reviewer	Petros Koumoutsakos		03/13/2023
Checked by WP Leader	Sergey Litvinov	ETHZ	06/03/2023
Checked by Project Coordinator	Vissarion Papadopoulos	NTUA	

Document Change History:

Version	**Date**	**Author(s)**	**Affiliation**	**Comment**
1.0	06/03/2023	Sergey Litvinov, Sebastian Kaltenbach	ETHZ	

Deliverable 4.1

A novel Transitional Markov Chain Monte Carlo Scheme (TMCMC) scheme was implemented in Korali and the results were published in the paper “The stress-free state of human erythrocytes: data driven inference of a transferable RBC model” [1]. In the following we present a summary of the findings which is closely based on the mentioned paper.

The developed TMCMC solver is based on the BASIS sampler [7] which is based on the algorithm introduced by Ching et al. [2]. The key idea of any algorithm based on the transitional Markov chain idea is to gradually transit from the prior to our posterior distribution. In more detail, for a prior $p(\theta)$ and a likelihood $p(D|\theta)$ we evaluate M intermediate target distributions:

$$p(\theta|D) \propto p(D|\theta)^{\gamma_i} p(\theta) \tag{1}$$

Here, $i = 1, ..., M$ and $\gamma_1 < \gamma_2 < ..... < \gamma_M$. The γ is increased slowly such that the evaluated distributions gradually changes from prior to the targeted posterior. The resulting set of samples can then be determined using weight factors as described in [2]. As for the DCoMEX project as well as for the application mentioned later, we are interested in hierarchical models we adapt the TMCMC solver for hierarchical stochastic models as suggested in [6]. This idea utilizes importance sampling to be able to deal with different levels/parts of the model individually. To obtain a general applicable solvers we do not make use of the gradient of the model with respect to the parameters, i.e. we do not require the model to be differentiable.

The mentioned sampling algorithm has been implemented into Korali [5]. In this report, we are presenting one example to underline the strength of our sampling method with regards to hierarchical models. We first introduce the model of the Red Blood Cell (RBC) to which we applied our sampling algorithm and show that it has the structure of a general hierarchical model. Subsequently we present the results obtained using the Korali implementation.

Computational model of the Red Blood Cell

RBC are modeled as a visco-elastic membrane enclosing a viscous fluid (cytoplasm). The membrane is discretized into a triangle mesh with N_v vertices, N_t triangles and N_e edges. Each vertex is a particle of mass m, position $\mathbf{r}_i$ and velocity $\mathbf{v}_i$, $i = 1, 2, \ldots, N_v$, that evolves in time according to Newton’s law of motion,

$$\begin{aligned} \dot{\mathbf{r}}_i &= \mathbf{v}_i, \\ \dot{\mathbf{v}}_i &= \mathbf{f}_i/m, \end{aligned}$$

where $\mathbf{f}_i$ is the force exerted on the i^{th} particle. The particles undergo forces coming from in-plane elasticity of the cytoskeleton, bending elasticity from the lipid bilayer, viscous contribution of the lipid bilayer and external forces such as those coming from the surrounding fluid. Additionally, area of the membrane and volume of the cell are constrained to model the incompressibility of the lipid bilayer and the cytosol, respectively, via energy penalization terms. All above contributions are written as

$$\mathbf{f}_i = -\nabla_i \left(U_{in-plane} + U_{bending} + U_{area} + U_{volume} \right) + \mathbf{f}_i^{visc} + \mathbf{f}_i^{fluct} + \mathbf{f}_i^{ext},$$

where ∇_i is the gradient with respect to the position $\mathbf{r}_i$.

The in-plane energy is

$$\begin{aligned} U_{in-plane} &= \frac{K_\alpha}{2} \oint \left(\alpha^2 + a_3\alpha^3 + a_4\alpha^4 \right) dA_0 + \mu \oint \left(\beta + b_1\alpha\beta + b_2\beta^2 \right) dA_0, \\ &\approx \frac{K_\alpha}{2} \sum_{k=1}^{N_t} \left(\alpha_k^2 + a_3\alpha_k^3 + a_4\alpha_k^4 \right) A_k + \mu \sum_{k=1}^{N_t} \left(\beta_k + b_1\alpha_k\beta_k + b_2\beta_k^2 \right) A_k, \end{aligned}$$

where α and β are the local dilation and shear strain invariants of the membrane, α_k and β_k are the area and shear strain invariants of the triangle k and A_k is the area of triangle k. The coefficients K_α and μ are the dilation and shear elastic moduli of the membrane, respectively. a_3, a_4, b_1 and b_2 are parameters that were set to $a_3 = -1$, $a_4 = 8$, $b_1 = 0.7$ and $b_2 = 0.75$.

The bending energy is chosen to [4]

$$U_{bending} = 2\kappa_b \oint (H - H_0)^2 \, dA \approx 2\kappa_b \sum_{k=1}^{N_v} A_k^v \left(H_k - H_0 \right)^2,$$

where κ_b is the bending energy coefficient, H_0 is the spontaneous mean curvature (set to 0 in the current work), H is the mean curvature and

$$H_k = \frac{1}{A_k^v} \sum_{\langle i,j \rangle_k} l_{ij}\phi_{ij}, \tag{2}$$

where the sum goes over all the neighboring dihedra of the vertex k, composed of triangles i and j. The length of the common edge between the triangles i and j is l_{ij} and ϕ_{ij} is the angle between the normal vectors of these two triangles, $\cos\phi_{ij} = \mathbf{n}_i \cdot \mathbf{n}_j$. The area A_k^v in is the area associated to the vertex k,

$$A_k^v = \frac{1}{3}\sum_{\langle i\rangle_k} A_i. \tag{3}$$

The penalization terms for the area and volume constrains of the cell are expressed

$$U_{area} = k_A \frac{(A - A_0)^2}{A_0}, \quad U_{volume} = k_V \frac{(V - V_0)^2}{V_0},$$

where A_0 and V_0 are the area and volume of the cell at rest. The coefficients k_A and k_v are chosen large enough to keep the membrane within a small deviation from the required area and volume.

The viscous force acting on each particles is given by

$$\mathbf{f}_i^{visc} = -\sum_j \gamma\left(\mathbf{v}_{ij}\cdot\mathbf{e}_{ij}\right)\mathbf{e}_{ij}, \tag{4}$$

where the sum goes over all connected vertices to i, γ is a friction coefficient, $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ and $\mathbf{e}_{ij}$ is the unit vector from $\mathbf{r}_j$ to $\mathbf{r}_i$. We define the viscosity of the membrane as $\eta_m = \gamma\sqrt{3}/4$. To satisfy the fluctuation-dissipation balance, the fluctuation forces are chosen as

$$\mathbf{f}_i^{fluct} = \sum_j \sigma\xi_{ij}\mathbf{e}_{ij}, \tag{5}$$

where $\sigma = \sqrt{2k_BT\gamma}$, k_BT is the temperature (in energy units) and ξ_{ij} is a random variable not correlated in time satisfying $\xi_{ij} = \xi_{ji}$, $\langle\xi_{ij}\rangle = 0$ and $\langle\xi_{ij}\xi_{lm}\rangle = \delta_{il}\delta_{jm} - \delta_{jl}\delta_{im}$. The shear and bending forces are computed by taking the negative gradient of the discretized energies with respect to the position of the mesh vertices.

Parameters dependencies in the RBC model

Here we describe the dependencies (conditional probabilities) of each random variable in the RBC model. We assume that the parameters of the computational model are conditioned on the hyper-parameters with the following probability distributions:

$$\begin{aligned}
p(v\,|\,\boldsymbol{\psi}) &= \mathcal{TN}\left(v;\mu_v,\sigma_v,0,1\right),\\
p(\kappa_b\,|\,\boldsymbol{\psi}) &= \mathcal{TN}\left(\kappa_b;\mu_{\kappa_b},\sigma_{\kappa_b},a_{\kappa_b},b_{\kappa_b}\right),\\
p(\alpha\,|\,\boldsymbol{\psi}) &= \mathcal{N}\left(\alpha;\mu_\alpha,\sigma_\alpha\right), \quad \alpha\in\{\mu, b_2, \eta_m\},
\end{aligned}$$

where $\mathcal{N}\left(\cdot;\mu,\sigma\right)$ is the normal distribution with mean μ and standard deviation σ and $\mathcal{TN}\left(\cdot;\mu,\sigma,a,b\right)$ is the probability density of the truncated normal in the interval $[a,b]$. The bounds for the bending modulus are chosen as $a_{\kappa_b} = 1.074\times10^{-19}$ J and $b_{\kappa_b} = 1.074\times10^{-18}$ J. The hyper-parameter vector $\boldsymbol{\psi}$ comprises the mean and variance of all RBC parameters.

The first experiment consists in measuring the shape of a RBC in its equilibrium shape. The shape is quantified by the diameter D, minimum thickness h_{min} and maximum thickness h_{max} of the equilibrated cell. The experimental data are assumed to be a realization of the observable $\boldsymbol{y}_{eq,1} = (D, h_{min}, h_{max})$ following the forward statistical model

$$\boldsymbol{y}_{eq,1} = \mathbf{G}_{eq}(\boldsymbol{\vartheta}_1) + \sigma_{eq,1}\boldsymbol{\Sigma}\boldsymbol{\varepsilon},$$

where $\mathbf{G}_{eq} = (G_D, G_{h_{min}}, G_{h_{max}})$ with G_D, $G_{h_{min}}$ and $G_{h_{max}}$ representing D, h_{min} and h_{max} obtained from the computational model, respectively, $\sigma_{eq,1}$ is the relative error to be calibrated, $\boldsymbol{\Sigma}$ is a diagonal matrix with $\mathrm{diag}(\boldsymbol{\Sigma}) = \mathbf{G}_{eq}(\boldsymbol{\vartheta}_1)$ and $\boldsymbol{\varepsilon}\sim\mathcal{N}(0,\mathbf{1})$.

The second experiment consists in pulling the two ends of a RBC with opposite forces of magnitude F_{ext}. The experimental data report the diameters of the stretched cells along the two largest principal axes, D_{ax} and D_{tr}, for n different stretching force magnitudes. We assume that the axial and transverse diameters reported in the experiments are normally distributed around the output of the computational model,

$$\begin{aligned}
y_{ax,i,j} &= G_{D_{ax}}(F_{ext,j},\boldsymbol{\vartheta}) + \sigma_{ax,i}\varepsilon_{ax,i,j},\\
y_{tr,i,j} &= G_{D_{tr}}(F_{ext,j},\boldsymbol{\vartheta}) + \sigma_{tr,i}\varepsilon_{tr,i,j},
\end{aligned}$$

where $G_{D_{ax}}(F_{ext,j},\boldsymbol{\vartheta})$ and $G_{D_{tr}}(F_{ext,j},\boldsymbol{\vartheta})$, $j = 1,2,\ldots,n$, are the axial and transverse diameters estimated by the computational model at input stretching forces $F_{ext,j}$ and parameters $\boldsymbol{\vartheta}$, $\varepsilon_{ax,i,j}\sim\mathcal{N}(0,1)$ and $\varepsilon_{tr,i,j}\sim\mathcal{N}(0,1)$, with $i = 2,3$ and $j = 1,2,\ldots,n$. The standard deviations $\sigma_{ax,i}$ and $\sigma_{tr,i}$ are grouped into $\boldsymbol{\sigma}_{st,i} = (\sigma_{ax,i},\sigma_{tr,i})$ and the observable $y_{ax,i,j}$ and $y_{tr,i,j}$, $j = 1,2,\ldots,n$, are grouped into the vector $\boldsymbol{y}_{st,i}$ in the DAG.

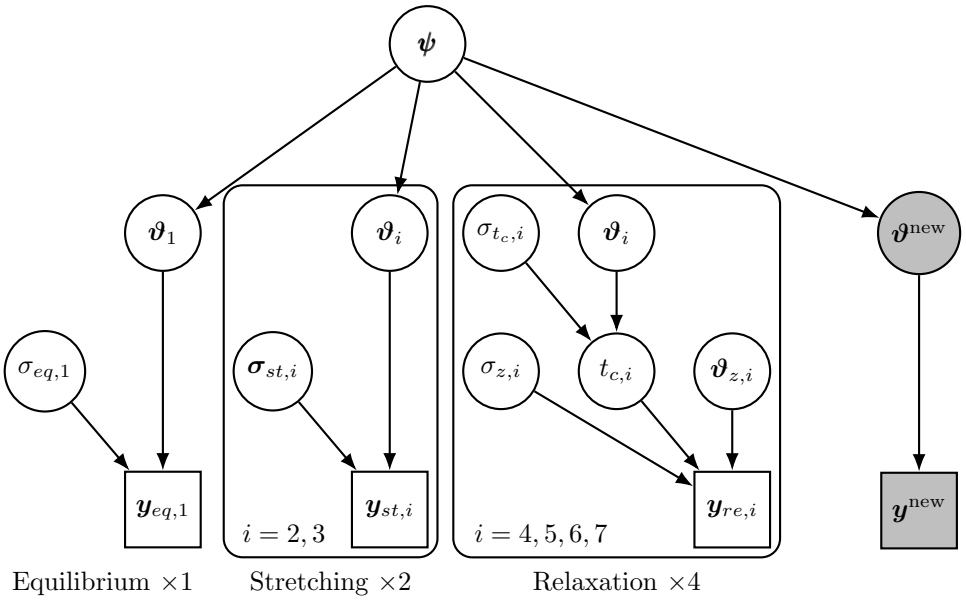

Figure 1: Hierarchical stocahstic model of the RBC

The last experiment corresponds to an initially stretched RBC relaxing to its resting shape. The ratio of the two main diameters, $z = D_{ax}/D_{tr}$, is recorded over time and has been measured experimentally in [3]. These measurements have been performed with arbitrary initial conditions, which would require to fit the corresponding stretching force to reach that shape for each parameter ϑ. This approach would require a prohibitive computational cost and we instead assume that the data follows an exponential decay as suggested in [3],

$$z(t, t_c) = z_\infty \frac{\Lambda + e^{-t/t_c}}{\Lambda - e^{-t/t_c}}, \tag{6}$$

where t_c is the characteristic relaxation time of the RBC, $\Lambda = (z_0 + z_\infty)/(z_0 - z_\infty)$, and z_0 and z_∞ are the initial and final ratio of the diameters, respectively, grouped in the parameter vector $\boldsymbol{\vartheta}_z = (z_0, z_\infty)$. The parameters of the RBC model do not influence the initial conditions, hence z_0 is not dependent on these parameters. Similarly, z_∞ can have a bias due to the experimental conditions (e.g. the orientation of the cell when measuring the diameters) that does not depend on the cell mechanical properties. Instead, we assume that only t_c depends on the RBC parameters.

The computational model consists in letting the cell relax from an arbitrary initial stretching position (chosen in the linear regime of the stretching curve, $z_0 \approx 2$) and computing the exponential decay $G_{t_c}(\boldsymbol{\vartheta})$ from a least square fit. The characteristic time t_c is assumed to follow a normal distribution with mean $G_{t_c}(\boldsymbol{\vartheta})$ and standard deviation σ_{t_c} (to be inferred),

$$t_{c,i} = G_{t_c}(\boldsymbol{\vartheta}_i) + \sigma_{t_c,i}\varepsilon_{t_c,i}, \quad i = 4, 5, 6, 7,$$

with $\varepsilon_{t_c,i} \sim \mathcal{N}(0, 1)$. The data is then modeled according to (6) with an additive Gaussian noise of standard deviation $\sigma_{z,i}$,

$$y_{re,i,j} = z(t_j, t_{c,i}) + \sigma_{z,i}\varepsilon_{z,i,j},$$

where t_j, $j = 1, 2, \ldots, n$ are the measurement times reported by the experiments and $\varepsilon_{z,i,j} \sim \mathcal{N}(0, 1)$. For a given data set $i \in \{4, 5, 6, 7\}$, the parameters $\boldsymbol{\vartheta}_i$, $\sigma_{t_c,i}$, $t_{c,i}$, $\boldsymbol{\vartheta}_{z,i} = (z_{0,i}, z_{\infty,i})$ and $\sigma_{z,i}$ are inferred all together to get the posterior of the RBC parameters given the relaxation data.

The model and its dependencies is summarized in Figure 1.

Sampling using Korali

The parameters of the RBC model are sampled using the Korali framework and our TMCMC implementation. Each sampling stage is performed using TMCMC with 50'000 samples. The resulting posterior distribution of the RBC parameters, $p(\boldsymbol{\vartheta}^{\text{new}} \mid \boldsymbol{d})$, is shown on the following figure. All distributions have a clear peak with relatively high uncertainties around the MAP as our data set has been quiet heterogen. The results obtained show that the Korali implementation can be used to obtain posterior distributions based on a hierarchical stochastic model.

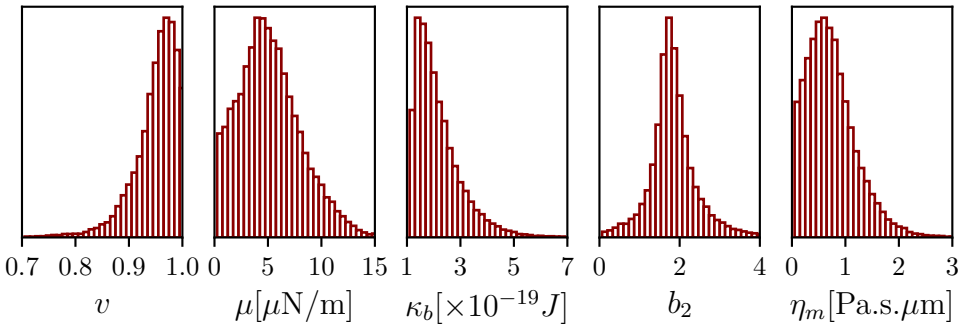

Figure 2: Posterior distribution for the parameters

References

[1] Lucas Amoudruz, Athena Economides, George Arampatzis, and Petros Koumoutsakos. The stress-free state of human erythrocytes: data driven inference of a transferable rbc model. *Biophysical journal*.

[2] Jianye Ching and Yi-Chu Chen. Transitional markov chain monte carlo method for bayesian model updating, model class selection, and model averaging. *Journal of engineering mechanics*, 133(7):816–832, 2007.

[3] R. M. Hochmuth, P. R. Worthy, and E. A. Evans. Red cell extensional recovery and the determination of membrane viscosity. *Biophysical Journal*, 26(1):101–114, 1979.

[4] Frank Jülicher. The morphology of vesicles of higher topological genus: conformal degeneracy and conformal modes. *Journal de Physique II*, 6(12):1797–1824, 1996.

[5] Sergio M. Martin, Daniel Wälchli, Georgios Arampatzis, Athena E. Economides, Petr Karnakov, and Petros Koumoutsakos. Korali: Efficient and scalable software framework for Bayesian uncertainty quantification and stochastic optimization. *Comput. Method. Appl. M.*, page 114264, nov 2021.

[6] Stephen Wu, Panagiotis Angelikopoulos, James L Beck, and Petros Koumoutsakos. Hierarchical stochastic model in bayesian inference for engineering applications: Theoretical implications and efficient approximation. *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, 5(1), 2019.

[7] Stephen Wu, Panagiotis Angelikopoulos, Costas Papadimitriou, and Petros Koumoutsakos. Bayesian annealed sequential importance sampling (BASIS): an unbiased version of transitional markov chain Monte Carlo. *ASCE-ASME J. Risk Uncertain. Eng. Sys. B*, aug 2017.