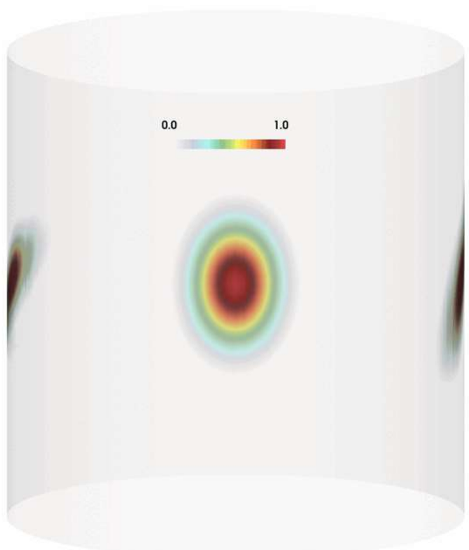
High Order Computation of MFC Barycenters with MFEM

Arjun Vijaywargiya University of Notre Dame October 22, 2024

MFC Barycenters



- Generalization of Wasserstein barycenters.
- Averages of probability densities.
- MFC builds on the Benamou-Brenier formulation.
- Comprehensive computational framework utilizing high order finite elements.

Reactive-diffusive Wasserstein Distance

$$\mathrm{Dist}_{\boldsymbol{V}_1,\boldsymbol{V}_2} : \mathcal{M}^N \times \mathcal{M}^N \to \mathbb{R}_+$$

Convex optimization with linear constraints!

Kinetic energy

Vectors of probability densities

Interaction energy

$$\mathrm{Dist}_{V_1,V_2}\left(\boldsymbol{\rho}^0,\boldsymbol{\rho}^1\right)^2 := \inf_{\boldsymbol{\rho},\underline{\boldsymbol{m}},\boldsymbol{s}} \int_0^T \int_\Omega \left(\sum_{i=1}^N \frac{||\boldsymbol{m}_i||^2}{V_{1,i}\left(\rho_i\right)} + \sum_{p=1}^R \frac{|s_p|^2}{V_{2,p}(\boldsymbol{\rho})} \right) dx dt + \sum_{i=1}^N \mathcal{E}_i(\rho_i)$$

Control variables

Positive and concave mobility functions.

such that

Set of reaction diffusion equations

$$\begin{cases} \partial_t \rho_i + \nabla \cdot \boldsymbol{m}_i = \sum_{p=1}^R \gamma_{i,p} s_p, & \forall 1 \le i \le N, \\ \boldsymbol{m}_i \cdot \boldsymbol{\nu}|_{\partial\Omega} = 0, \quad \boldsymbol{\rho}(0,x) = \boldsymbol{\rho}^0, & \boldsymbol{\rho}(T,x) = \boldsymbol{\rho}^1 \end{cases}$$

Barycenter problem:

$\boldsymbol{\rho}^{\mathbf{0}}$ Initial densities

$\boldsymbol{\rho}^{\mathbf{1}} = \varrho \mathbf{1}$

Unknown terminal density

Another control variable!

Unconstrained saddle point problem

Enforce constraints

$$\inf_{\boldsymbol{\rho},\underline{\boldsymbol{m}},\boldsymbol{s},\varrho}\sup_{\boldsymbol{\phi}}\int_0^T\left[\int_\Omega\left(\sum_{i=1}^N\frac{||\boldsymbol{m}_i||^2}{2V_{1,i}(\rho_i)}+\sum_{i=1}^N\frac{|s_i|^2}{2V_{2,i}(\rho_i,\rho_{i+1})}\right)dx-\boldsymbol{\mathcal{F}}(\boldsymbol{\rho})\right]dt$$

$$+\sum_{i=0}^N\int_0^T\int_\Omega(\partial_t\rho_i+\nabla_\Omega\cdot\boldsymbol{m}_i+s_{i-1}-s_i)\phi_i dxdt,$$

Integrate by parts

$$\inf_{\boldsymbol{\rho},\underline{\boldsymbol{m}},\boldsymbol{s},\varrho}\sup_{\boldsymbol{\phi}}\sum_{i=1}^N\left(\frac{||\boldsymbol{m}_i||^2}{2V_{1,i}(\rho_i)}+\frac{|s_i|^2}{2V_{2,i}(\rho_i,\rho_{i+1})},1\right)_{\Omega_t}-\int_0^T\boldsymbol{\mathcal{F}}(\boldsymbol{\rho})dt$$

$$-(\boldsymbol{\rho},\partial_t\boldsymbol{\phi})_{\Omega_t}-(\underline{\boldsymbol{m}},\underline{\nabla_\Omega\boldsymbol{\phi}})_{\Omega_t}-\left(\boldsymbol{s},\hat{\boldsymbol{\phi}}\right)_{\Omega_t}+\langle\varrho,\boldsymbol{\phi}\cdot\mathbf{1}\rangle_{t=T}-\langle\boldsymbol{\rho}^0,\boldsymbol{\phi}\rangle_{t=0}$$

$$(\boldsymbol{a},\boldsymbol{b})_{\Omega_t}=\int_0^T\int_\Omega\boldsymbol{a}\cdot\boldsymbol{b}dxdt$$

$$\langle\boldsymbol{a}\cdot\boldsymbol{b}\rangle_{t=s}=\int_\Omega\boldsymbol{a}(s,x)\cdot\boldsymbol{b}(s,x)dx$$

$$\underline{\nabla_\Omega\boldsymbol{\phi}}=(\nabla_\Omega\phi_1,\cdots,\nabla_\Omega\phi_N)$$

$$\mathcal{D}\phi_i=(\partial_t\phi_i,\nabla_\Omega\phi_i,\phi_i-\phi_{i+1})$$

Simplify notation

$$\inf_{\underline{\boldsymbol{u}},\varrho}\sup_{\boldsymbol{\phi}}\underbrace{H(\underline{\boldsymbol{u}})-\left(\underline{\boldsymbol{u}},\underline{\mathcal{D}\boldsymbol{\phi}}\right)_{\Omega_t}+\langle\varrho,\boldsymbol{\phi}\cdot\mathbf{1}\rangle_{t=T}-\left\langle\boldsymbol{\rho}^0,\boldsymbol{\phi}\right\rangle_{t=0}}_{:=\mathcal{L}(\underline{\boldsymbol{u}},\varrho,\boldsymbol{\phi})}$$

$$H(\underline{\boldsymbol{u}}):=\sum_{i=1}^N\int_0^T\int_\Omega\left[\frac{||\boldsymbol{m}_i||^2}{2V_{1,i}(\rho_i)}+\frac{|s_i|^2}{2V_{2,i}(\rho_i,\rho_{i+1})}\right]dxdt-\int_0^T\boldsymbol{\mathcal{F}}(\boldsymbol{\rho})dt$$

Discrete problem

Mesh

$\Omega_h = \{T_k\}_{k=1}^{N_S}$ — Space mesh

$\mathcal{I}_h = \{I_j\}_{j=1}^{N_T}$ — Time mesh

$\Omega_{t,h} := \mathcal{I}_h \otimes \Omega_h$ — Spacetime mesh

Finite element spaces

Lagrange multipliers → $V_h^k := \left\{ v \in H^1(\Omega_T) : v|_{I_j \times T_k} \in Q^k(I_j) \otimes Q^k(T_k), \ \ \forall j, k \right\}$

Physical variables → $W_h^{k-1} := \left\{ w \in L^2(\Omega_T) : w|_{I_j \times T_k} \in Q^{k-1}(I_j) \otimes Q^{k-1}(T_k), \ \ \forall j, k \right\}$

Terminal density → $M_h^{k-1} := \left\{ \mu \in L^2(\Omega) : \mu|_{T_k} \in Q^{k-1}(T_k), \ \ \forall k \right\},$

Discrete problem

$$\inf_{\underline{\boldsymbol{u}_h}, \varrho_h} \sup_{\boldsymbol{\phi}_h} F_h(\underline{\boldsymbol{u}_h}) - \left(\underline{\boldsymbol{u}_h}, \underline{\mathcal{D}\boldsymbol{\phi}_h} \right)_{\Omega_{t,h}} + \langle \varrho_h, \boldsymbol{\phi}_h \cdot \mathbf{1} \rangle_{t=T,h} - \left\langle \boldsymbol{\rho}_h^0, \boldsymbol{\phi}_h \right\rangle_{t=0,h}$$

Primal Dual Hybrid Gradient

- Step 1: Proximal gradient ascent for $\boldsymbol{\phi}_h^{k+1}$:

$$\boldsymbol{\phi}_h^{k+1} = \underset{\boldsymbol{\phi}_h}{\text{argmax}} - \left(\underline{\boldsymbol{u}_h^k}, \underline{\mathcal{D}\boldsymbol{\phi}_h}\right)_{\Omega_{t,h}} + \left\langle \varrho_h^k, \boldsymbol{\phi}_h \cdot \mathbf{1} \right\rangle_{t=T,h} - \left\langle \boldsymbol{\rho}_h^0, \boldsymbol{\phi}_h \right\rangle_{t=0,h}$$
$$- \frac{1}{2\sigma_{\boldsymbol{\phi}_h}} \left(\underline{\mathcal{D}(\boldsymbol{\phi}_h - \boldsymbol{\phi}_h^k)}, \underline{\mathcal{D}(\boldsymbol{\phi}_h - \boldsymbol{\phi}_h^k)}\right)_{\Omega_{t,h}} - \frac{1}{2\sigma_{\boldsymbol{\phi}_h}} \left\langle (\boldsymbol{\phi}_h - \boldsymbol{\phi}_h^k) \cdot \mathbf{1}, (\boldsymbol{\phi}_h - \boldsymbol{\phi}_h^k) \cdot \mathbf{1} \right\rangle_{t=T,h}$$

Elliptical solves for the Lagrange multipliers

- Step 2: Extrapolation for $\tilde{\boldsymbol{\phi}}_h^{k+1}$:

$$\tilde{\boldsymbol{\phi}}_h^{k+1} = 2\boldsymbol{\phi}_h^{k+1} - \boldsymbol{\phi}_h^k$$

- Step 3: Proximal gradient descent for ($\underline{\boldsymbol{u}_h^{k+1}}$ and ϱ_h^{k+1}):

$$\varrho_h^{k+1} = \underset{\varrho_h}{\text{argmin}} \left\langle \varrho_h, \tilde{\boldsymbol{\phi}}_h^{k+1} \cdot \mathbf{1} \right\rangle_{t=T,h} + \frac{1}{2\sigma_{\underline{\boldsymbol{u}_h}}} \left(\varrho_h - \varrho_h^k, \varrho_h - \varrho_h^k\right)_{\Omega_{t,h}}$$

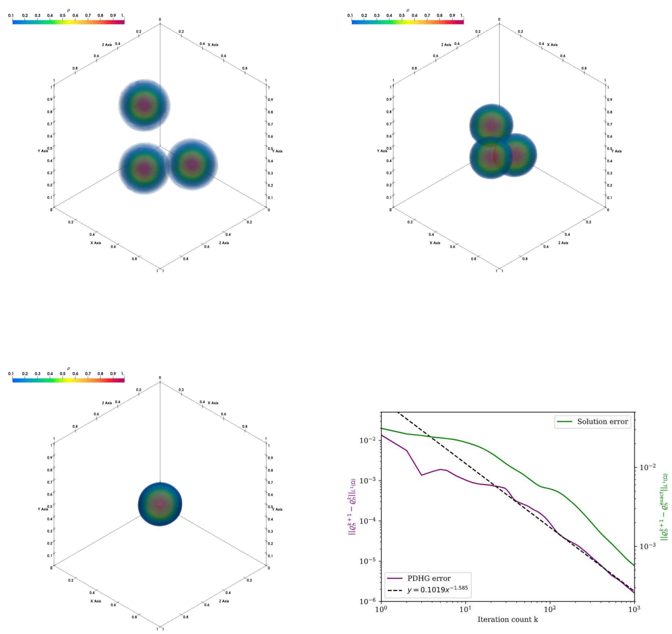
$$\underline{\boldsymbol{u}_h^{k+1}} = \underset{\underline{\boldsymbol{u}_h}}{\text{argmin}} \; F(\underline{\boldsymbol{u}_h}) - \left(\underline{\boldsymbol{u}_h}, \underline{\mathcal{D}\tilde{\boldsymbol{\phi}}_h^{k+1}}\right)_{\Omega_{t,h}} + \frac{1}{2\sigma_{\underline{\boldsymbol{u}_h}}} \left(\underline{\boldsymbol{u}_h} - \underline{\boldsymbol{u}_h^k}, \underline{\boldsymbol{u}_h} - \underline{\boldsymbol{u}_h^k}\right)_{\Omega_{t,h}}$$

Pointwise non-linear minimization at each degree of freedom.

MFEM Implementation

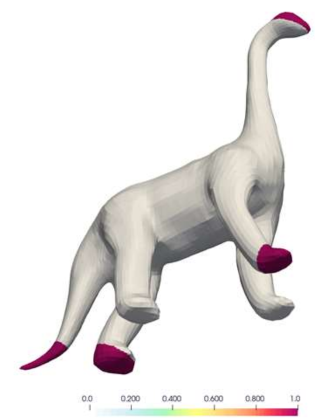
- Tensor product spacetime mesh.
- Define custom classes for the assembly of the linear systems and for quadrature functions.
- The elliptical problems for the Lagrange multipliers are solved using the Conjugate Gradient Solver and a Geometric Multigrid preconditioner.
- The non-linear minimization for the physical variables is done using a C++ implementation of the Brent method.
- Make full use of MFEM’s MPI-based parallelism and run the code on an HPC environment.

Numerical Results

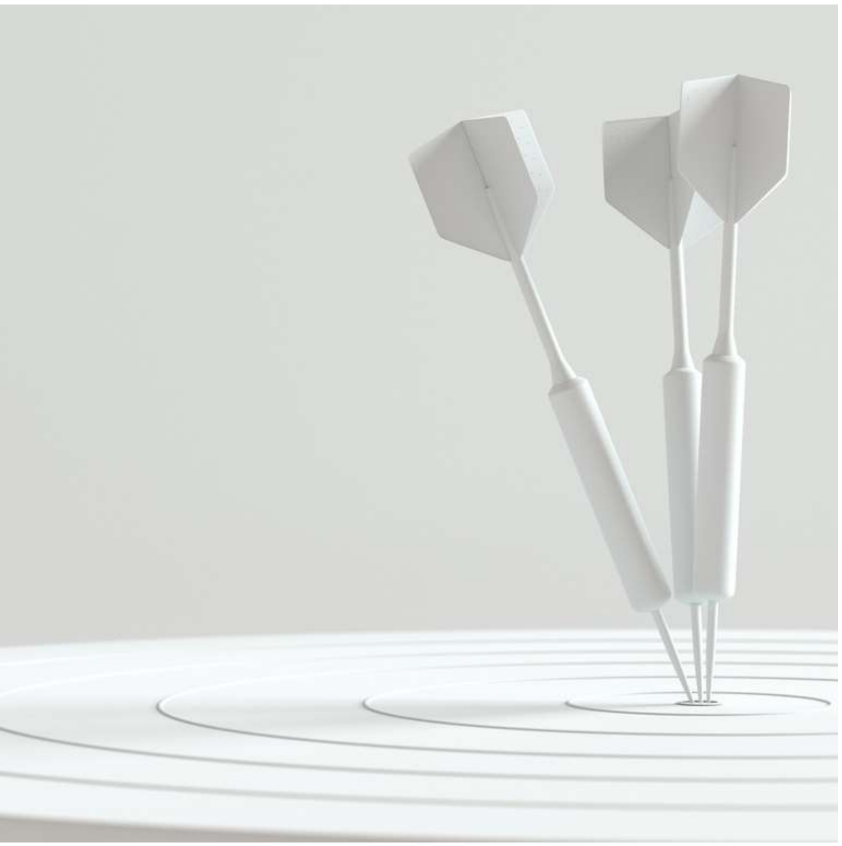


Barycenter of 3 gaussians

3D shape interpolation



Barycenter on a surface



Thank you

1. **A. Vijaywargiya**, S. McQuarrie, and A. Gruber, *Tensor Parametric Operator Inference with Hamiltonian Structure*, Computer Science Research Institute Summer Proceedings 2025, (to appear 2024).

2. **A. Vijaywargiya**, G. Fu, S. Osher, and W. Li, *Efficient Computation of Mean Field Control Based Barycenters from Reaction-Diffusion Systems*, Journal of Computational Physics, (in review 2024). Preprint available at: https://arxiv.org/abs/2404.01586.

3. **A. Vijaywargiya** and G. Fu, *Two Finite Element Approaches for the Porous Medium Equation That Are Positivity Preserving and Energy S table*, **J. Sci. Comput.**, **100**, 86 (2024). https://doi.org/10.1007/s10915-024-02642-x.