High-order spatial discretization for variational time implicit schemes: Wasserstein gradient flows and reaction-diffusion systems

Guosheng Fu, Stanley Osher, Wuchen Li

ACM Seminar @ University of South Carolina, Sep 22

** JCP 2023 (491) pp. 112375, JCP 2023 (491) pp. 112346

Table of Contents

The Wasserstein gradient flow

PDE:

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta}{\delta \rho} \mathcal{E} \right), \tag{1}$$

with energy functional:

$$\mathcal{E}(\rho) := \int_\Omega \left[\underbrace{\alpha U_m(\rho(x))}_{\textit{diffusion}} + \underbrace{\rho(x) V(x)}_{\textit{drift}} + \underbrace{\frac{1}{2}(W * \rho)(x)\rho(x)}_{\textit{aggregation}} \right] dx, \tag{2}$$

The PDE (1) is a gradient flow with an energy dissipation law

$$\frac{d}{dt}\mathcal{E}(\rho(\cdot, t)) = -\int_\Omega \|\nabla \frac{\delta}{\delta \rho}\mathcal{E}(\rho)(x,t)\|^2 \rho(x,t)\, dx \leq 0. \tag{3}$$

The JKO scheme [JKO98]

The Jordan-Kinderlehrer-Otto scheme, or minimizing movement scheme, proposed in [JKO98] is a variational time implicit scheme:
Definition: *Variational time implicit (JKO) scheme.* Denote $\Delta t > 0$ as a time step size. Consider the scheme below:

$$\rho^n = \arg\min_{\rho \in \mathcal{M}} \quad \frac{1}{2\Delta t} \mathrm{Dist}_{W_2}(\rho^{n-1}, \rho)^2 + \mathcal{E}(\rho). \tag{4}$$

where the distance functional $\mathrm{Dist}_{W_2}(\rho^{n-1}, \rho)^2$ is the Wasserstein-2 distance between current density ρ and previous step density ρ^{n-1}.

Positivity, mass conservation and energy dissipation are inbuilt in the JKO scheme, which are nontrivial to preserve.

The Wasserstein distance term involves solving a costly optimal transport problem at each step, which is a serious numerical difficulty for the practical implementation of the JKO scheme.

The dynamic JKO scheme [BCL16; Car+22]

The (dynamic) Benamou-Brenier formulation [BB00] of the Wasserstein-2 distance functional:

$$\mathrm{Dist}_{W_2}(\rho^0,\rho^1)^2 := \inf_{v,\rho} \quad \int_0^1 \int_\Omega \|v(x,\tau)\|^2 \rho(x,\tau) dx d\tau,$$

where the infimum is taken among ρ, v such that

$$\partial_\tau \rho(x,\tau) + \nabla \cdot (\rho(x,\tau) v(x,\tau)) = 0, \quad \rho(x,0) = \rho^0(x), \quad \rho(x,1) = \rho^1(x).$$

Using this definition, the JKO scheme can be converted to an (convex) control problem with linear constraints [BCL16; Car+22]:

$$\rho^n = \arg \min_{\rho_{\Delta t}, \rho(\cdot,\tau), m(\cdot,\tau)} \quad \frac{1}{2} \int_0^{\Delta t} \int_\Omega \frac{\|m(x,\tau)\|^2}{\rho(x,\tau)} dx d\tau + \mathcal{E}(\rho_{\Delta t}), \tag{5a}$$

such that

$$\partial_\tau \rho(x,\tau) + \nabla \cdot m(x,\tau) = 0, \quad \tau \in [0,\Delta t], \quad \rho(x,0) = \rho^{n-1}(x). \tag{5b}$$

One-step relaxation of JKO [LLW20; CGT20]

[LLW20; CGT20] introduced the following one-step relaxation of the JKO scheme to further drive down the computational cost.

Definition: *One-step relaxation of variational time implicit schemes.* Consider

$$\rho^n = \arg\min_{\rho\in\mathcal{M}} \underbrace{\frac{1}{2\Delta t}\int_\Omega \frac{\|m(x)\|^2}{\rho(x)}\,dx}_{\approx \frac{1}{2\Delta t}\mathrm{Dist}_{W_2}(\rho,\rho^{n-1})^2} + \mathcal{E}(\rho), \tag{6a}$$

where the minimization is over all functions m, ρ such that

$$\rho(x) - \rho^{n-1}(x) + \nabla\cdot m(x) = 0. \tag{6b}$$

This scheme forms a first-order implicit time scheme for Wasserstein gradient flows (1). With appropriate spatial discretization, the scheme (6) can then be efficiently solved using classical first-order proximal splitting methods [PPO14], e.g., ADMM/ALG2 [FG83], PDHG [CP11].

Table of Contents

Dissipative reaction-diffusion equation

We now extend the previous framework to the following dissipative reaction-diffusion equation:

$$\partial_t \rho = \nabla \cdot (V_1(\rho) \nabla \frac{\delta \mathcal{E}}{\delta \rho}(\rho)) - V_2(\rho) \frac{\delta \mathcal{E}}{\delta \rho}(\rho). \tag{7}$$

Here we require the two mobility functions $V_1(\rho)$ and $V_2(\rho)$ to be non-negative, so that the energy dissipation law is still value:

$$\frac{d}{dt}\mathcal{E}(\rho(\cdot, t)) = -\int_\Omega \left[\|\nabla \frac{\delta \mathcal{E}}{\delta \rho}(\rho)\|^2 V_1(\rho) + |\frac{\delta \mathcal{E}}{\delta \rho}(\rho)|^2 V_2(\rho) \right] dx \leq 0. \tag{8}$$

The metric distance operator

Definition: *Distance functional.* Define a distance functional $\mathrm{Dist}_{V_1,V_2}: \mathcal{M}\times\mathcal{M}\to\mathbb{R}_+$ as below. Consider the following optimal control problem:

$$\begin{aligned}&\mathrm{Dist}_{V_1,V_2}(\rho^0,\rho^1)^2\\ :=&\inf_{v_1,v_2,\rho}\int_0^1\int_\Omega\Big[\|v_1(x,\tau)\|^2V_1(\rho(x,\tau))+|v_2(x,\tau)|^2V_2(\rho(x,\tau))\Big]dxd\tau,\end{aligned}\tag{9a}$$

where the infimum is taken among $\rho\colon\Omega\times[0,1]\to\mathbb{R}$, $v_1,v_2\colon\Omega\times[0,1]\to\mathbb{R}^d$, such that ρ satisfies a reaction-diffusion type equation with drift vector field v_1, drift mobility V_1, reaction ratio v_2, reaction mobility V_2, connecting initial and terminal densities ρ^0, ρ^1:

$$\begin{cases}\partial_\tau\rho(x,\tau)+\nabla\cdot(V_1(\rho(x,\tau))v_1(x,\tau))=V_2(\rho(x,\tau))v_2(x,\tau),\quad\tau\in[0,1],\\ \rho(x,0)=\rho^0(x),\quad\rho(x,1)=\rho^1(x).\end{cases}\tag{9b}$$

One-step relaxation of variational time implicit schemes

Definition: *One-step relaxation of variational time implicit schemes.* Consider

$$\rho^n = \arg\min_{\rho\in\mathcal{M}} \underbrace{\frac{1}{2\Delta t}\int_\Omega \left[\frac{\|m(x)\|^2}{V_1(\rho(x))} + \frac{|s(x)|^2}{V_2(\rho(x))}\right] dx}_{\approx \frac{1}{2\Delta t}\mathrm{Dist}_{V_1,V_2}(\rho,\rho^{n-1})^2} + \mathcal{E}(\rho), \qquad (10a)$$

where the minimization is over all functions $m\colon \Omega \to \mathbb{R}^d$, $s\colon \Omega \to \mathbb{R}$, and $\rho\colon \Omega \to \mathbb{R}_+$, such that

$$\rho(x) - \rho^{n-1}(x) + \nabla \cdot m(x) = s(x). \qquad (10b)$$

Theorem 1 (Time implicit scheme entropy dissipation)

Denote the solution $\{\rho^n\}_{n\in\mathbb{N}}$ solving the variational implicit scheme (10). For any stepsize $\Delta t \geq 0$, we have

$$\mathcal{E}(\rho^n) \leq \mathcal{E}(\rho^{n-1}), \qquad \text{for } n \in \mathbb{N}_+.$$

Dissipative reaction-diffusion equation: examples

$V_1(\rho) = \rho$.

(i) $V_2(\rho) = c\,\rho^\gamma$ where $c \geq 0$ and $\gamma \in \mathbb{R}$, with a general $\mathcal{E}(\rho)$ given in (2). Here $\gamma = 1$ corresponds to the Wasserstein-Fisher-Rao metrics used in [Chi+18], and $\gamma = 0$ is related to unnormalized optimal transport [Lee+21].

(ii) $V_2(\rho) = c\,\frac{\rho-1}{\log(\rho)}$ where $c \geq 0$ with a general $\mathcal{E}(\rho)$ given in (2).

(iii) $V_2(\rho) = \frac{\rho(\rho-1)}{\alpha\log(\rho)}$, with linear diffusion $\mathcal{E}(\rho) := \int_\Omega \alpha\rho(x)(\log(\rho) - 1)dx$, where $\alpha > 0$. This model is the following Fisher–KPP equation; see [LLO22, Example 7]:

$$\frac{\partial \rho}{\partial t} - \nabla \cdot (\alpha \nabla \rho) = \rho(1-\rho). \tag{11}$$

Strongly reversible reaction-diffusion system

We consider M different chemical species $X_1, \dots, X_M$ reacting according to R mass-action laws:

$$\alpha_1^p X_1 + \cdots + \alpha_M^p X_M \underset{k_-^p}{\overset{k_+^p}{\rightleftharpoons}} \beta_1^p X_1 + \cdots + \beta_M^p X_M, \tag{12}$$

where $p = 1, \cdots, R$ is the number of possible reactions, $\boldsymbol{\alpha}^p = (\alpha_1^p, \cdots, \alpha_M^p), \boldsymbol{\beta}^p = (\beta_1^p, \cdots, \beta_M^p) \in \mathcal{N}_0^M$ are the vectors of the stoichimetric coefficients, and k_+^p, k_-^p the positive forward and backward reaction rates. We restrict ourselves to the *strongly reversible* case where $k_+^p = k_-^p = k^p > 0$.

Combining the mass-action laws (12) with (independent) isotropic linear diffusion with energy $\mathcal{E}_i(\rho_i) = \int_\Omega \rho_i(\log(\rho_i) - 1)\, dx$ for each density ρ_i of species X_i, we get the following reaction-diffusion system:

$$\partial_t \rho_i - \nabla \cdot \left(\gamma_i \rho_i \nabla \frac{\delta}{\delta \rho} \mathcal{E}_i(\rho_i) \right) = -\sum_{p=1}^{R} k^p (\alpha_i^p - \beta_i^p)(\boldsymbol{\rho}^{\boldsymbol{\alpha}^p} - \boldsymbol{\rho}^{\boldsymbol{\beta}^p}), \tag{13}$$

for $1 \le i \le M$, where $\boldsymbol{\rho} = (\rho_1, \cdots, \rho_M)$ and the multi-index notation $\boldsymbol{\rho}^{\boldsymbol{\alpha}^p} := \prod_{i=1}^M \rho_i^{\alpha_i^p}$ is used.

System reformation

We recast the above system (13) back to a system version of the general dissipative form (7) using appropriate mobility functions. We introduce the following function; see [Mie11]:

$$\ell(x,y) = \begin{cases} \frac{x-y}{\log(x)-\log(y)} & \text{for } x \neq y, \\ y & \text{for } x = y, \end{cases} \tag{14}$$

and denote the following mobility functions:

$$V_{1,i}(\rho_i) = \gamma_i \rho_i, \quad \forall 1 \leq i \leq M, \tag{15a}$$

$$V_{2,p}(\boldsymbol{\rho}) = k^p \, \ell\left(\boldsymbol{\rho}^{\boldsymbol{\alpha}^p}, \boldsymbol{\rho}^{\boldsymbol{\beta}^p}\right), \quad \forall 1 \leq p \leq R. \tag{15b}$$

Using these notations, it can be shown that (13) is equivalent to

$$\begin{aligned} \partial_t \rho_i = \nabla \cdot \left(V_{1,i}(\rho_i) \nabla \frac{\delta}{\delta \rho} \mathcal{E}_i(\rho_i) \right) \\ - \sum_{p=1}^{R} V_{2,p}(\boldsymbol{\rho})(\alpha_i^p - \beta_i^p) \sum_{j=1}^{M} (\alpha_j^p - \beta_j^p) \frac{\delta}{\delta \rho} \mathcal{E}_i(\rho_i). \end{aligned} \tag{16}$$

Energy dissipation for reaction-diffusion systems

It is now clear that the above system is purely dissipative as for the scalar case (7). That is, the first-time derivative of the energy functional is nonnegative and satisfies

$$\frac{d}{dt}\sum_{i=1}^{M}\mathcal{E}_i(\rho_i(\cdot,t)) = -\sum_{i=1}^{M}\int_\Omega \|\nabla\frac{\delta}{\delta\rho}\mathcal{E}_i(\rho)_i(x,t)\|^2 V_{1,i}(\rho_i)\,dx$$
$$-\sum_{p=1}^{R}\int_\Omega\left|\sum_{j=1}^{M}(\alpha_j^p-\beta_j^p)\frac{\delta}{\delta\rho}\mathcal{E}_i(\rho_i)\right|^2 V_{2,p}(\boldsymbol{\rho})\,dx. \tag{17}$$

One-step relaxation of variational time implicit schemes

Definition: *One-step relaxation of variational time implicit schemes for system* (16). Consider

$$\boldsymbol{\rho}^n = \arg\min_{\boldsymbol{\rho}\in[\mathcal{M}]^M} \quad \frac{1}{2\Delta t}\left(\sum_{i=1}^{M}\int_\Omega \frac{\|m_i\|^2}{V_{1,i}(\rho_i)}dx + \sum_{p=1}^{R}\int_\Omega \frac{\|s_p\|^2}{V_{2,p}(\boldsymbol{\rho})}dx\right) + \sum_{i=1}^{M}\mathcal{E}_i(\rho_i), \tag{18a}$$

where the minimization is over all functions $\boldsymbol{m}\colon \Omega \to [\mathbb{R}^d]^M$, $\boldsymbol{s}\colon \Omega \to [\mathbb{R}]^R$, and $\boldsymbol{\rho}\colon \Omega \to [\mathbb{R}_+]^M$, such that

$$\rho_i(x) - \rho_i^{n-1}(x) + \nabla\cdot m_i(x) = \sum_{p=1}^{R}(\alpha_i^p - \beta_i^p)s_p(x), \quad \forall 1 \le i \le M. \tag{18b}$$

Table of Contents

ALG2 [FG83]

We apply the augmented Lagrangian ALG2 algorithm [FG83] to solve the optimization problems (6), (10), (18). All three problems are of the form

$$\inf_{\boldsymbol{u}} \sup_{\Phi} F(\boldsymbol{u}) - G(\Phi) - (\boldsymbol{u}, \mathcal{D}\Phi)_\Omega, \tag{19}$$

where $\mathcal{D}(\Phi)$ is a *linear* differential operator for Φ, and $(\cdot,\cdot)_\Omega$ stands for the L^2-inner product on the domain Ω. For example, for scalar reaction-diffusion (10), we choose

$$\boldsymbol{u} = (\rho, m, s),$$

with

$$F(\boldsymbol{u}) = \frac{1}{2}\int_\Omega \Big[\frac{\|m\|^2}{V_1(\rho)} + \frac{|s|^2}{V_2(\rho)}\Big] dx + \Delta t \mathcal{E}(\rho), \quad G(\Phi) = \int_\Omega \rho^{n-1}\Phi \, dx,$$

and

$$(\boldsymbol{u}, \mathcal{D}\Phi)_\Omega = \int_\Omega \Big[-\rho\Phi + (m, \nabla_x \Phi) + s\Phi \Big] dx.$$

The augmented Lagrangian

The ALG2 algorithm starts with the dual formulation of the saddle-point problem (19):

$$\sup_{\boldsymbol{u}} \inf_{\Phi,\boldsymbol{u}^*} F^*(\boldsymbol{u}^*) + G(\Phi) + (\boldsymbol{u}, \mathcal{D}\Phi - \boldsymbol{u}^*)_\Omega, \tag{20}$$

where $F^*(\boldsymbol{u}^*) = \sup_{\boldsymbol{u}}(\boldsymbol{u}, \boldsymbol{u}^*)_\Omega - F(\boldsymbol{u})$ is the Legendre transform. The saddle point of the above system is equivalent to the saddle point of the following augmented Lagrangian form:

$$\sup_{\boldsymbol{u}} \inf_{\Phi,\boldsymbol{u}^*} L_r(\Phi, \boldsymbol{u}, \boldsymbol{u}^*), \tag{21}$$

where the augmented Lagrangian

$$L_r(\Phi, \boldsymbol{u}, \boldsymbol{u}^*) := F^*(\boldsymbol{u}^*) + G(\Phi) + (\boldsymbol{u}, \mathcal{D}\Phi - \boldsymbol{u}^*)_\Omega + \frac{r}{2}(\mathcal{D}\Phi - \boldsymbol{u}^*, \mathcal{D}\Phi - \boldsymbol{u}^*)_\Omega,$$

in which r is a positive parameter.

The ALG2 iteration

Algorithm 1 One iteration of ALG2 algorithm for variational implicit scheme (21).

- Step A: update Φ. Minimize $L_r(\Phi, \boldsymbol{u}, \boldsymbol{u}^*)$ with respect to the first argument by solving the elliptic problem: Find Φ^ℓ such that it solves

$$\inf_{\Phi} L_r(\Phi, \boldsymbol{u}^{\ell-1}, \boldsymbol{u}^{*,\ell-1}).$$

This step is a linear, constant-coefficient reaction-diffusion problem.

- Step B: update $\boldsymbol{u}^*$. Minimize $L_r(\Phi, \boldsymbol{u}, \boldsymbol{u}^*)$ with respect to the last argument by solving the nonlinear problem: Find $\boldsymbol{u}^{*,\ell}$ such that it solves

$$\inf_{\boldsymbol{u}^*} L_r(\Phi^\ell, \boldsymbol{u}^{\ell-1}, \boldsymbol{u}^*).$$

- Step C: update $\boldsymbol{u}$. This is a simple pointwise update for the Lagrange multiplier $\boldsymbol{u}$:

$$\boldsymbol{u}^\ell = \boldsymbol{u}^{\ell-1} + r(\mathcal{D}\Phi^\ell - \boldsymbol{u}^{*,\ell}). \tag{22}$$

Table of Contents

High-order finite element spaces/nodal basis

Approximate Φ using high-order H^1-conforming finite element space

$$V_h^k := \{v \in H^1(\Omega) : \quad v|_T \in \mathcal{Q}^k(T) \ \ \forall T \in \mathcal{T}_h\}, \tag{23}$$

Approximate density/flux/source using high-order L^2-conforming space

$$W_h^k := \{w \in L^2(\Omega) : \quad w|_T \in \mathcal{Q}^k(T) \ \ \forall T \in \mathcal{T}_h\}, \tag{24}$$

We equip the space W_h^k with a set of *nodal basis* $\{\varphi_i\}_{i=1}^{N_W} \subset W_h^k$ that satisfies $\varphi_i(\xi_j) = \delta_{ij}, \quad \forall 1 \le j \le N_W$, where N_W is the dimension of the space W_h^k, δ_{ij} is the Kronecker delta function, and $\{\xi_j\}_{i=1}^{N_W}$ is the collection of N_W Gauss-Legendre integration points with corresponding weights $\{\omega_j\}_{i=1}^{N_W}$ on the mesh $\mathcal{T}_h$.
We denote the discrete $L^2(\Omega)$-inner product $(\cdot,\cdot)_h$ as

$$(u, v)_h := \sum_{i=1}^{N_W} u(\xi_i) v(\xi_i) \omega_i, \tag{25}$$

Fully discrete scheme for reaction diffusion equation

Given density approximation ρ_h^{old} at the previous time step, find $\boldsymbol{u}_h, \boldsymbol{u}_h^* \in [W_h^k]^4$, and $\Phi_h \in V_h^k$ such that

$$\inf_{\boldsymbol{u}\in[W_h^k]^4} \quad \sup_{\Phi_h\in V_h^k, \boldsymbol{u}_h^*\in[W_h^k]^4} L_{r,h}(\Phi_h, \boldsymbol{u}_h, \boldsymbol{u}_h^*), \tag{26a}$$

where $\boldsymbol{u}_h = (\rho_h, m_h^0, m_h^1, s_h)$ is the collection of density/flux/source, $\boldsymbol{u}_h^* = (\rho_h^*, m_h^{0,*}, m_h^{1,*}, s_h^*)$ is its dual, and

$$\begin{aligned} L_{r,h}(\Phi_h, \boldsymbol{u}_h, \boldsymbol{u}_h^*) := F_h^*(\boldsymbol{u}_h^*) + G_h(\Phi_h) + (\boldsymbol{u}_h, \mathcal{D}\Phi_h - \boldsymbol{u}_h^*)_h \\ + \frac{r}{2}(\mathcal{D}\Phi_h - \boldsymbol{u}_h^*, \mathcal{D}\Phi_h - \boldsymbol{u}_h^*)_h, \end{aligned} \tag{26b}$$

in which $(\cdot,\cdot)_h$ is the volume integration rule given in (25), the operators

$$\mathcal{D}\Phi_h := (-\Phi_h, \partial_{x_0}\Phi_h, \partial_{x_1}\Phi_h, \Phi_h), \quad G_h(\Phi_h) := (\rho_h^{\text{old}}, \Phi_h)_h, \tag{26c}$$

$$F_h^*(\boldsymbol{u}_h^*) := \sup_{\boldsymbol{u}_h\in[W_h^k]^4} (\boldsymbol{u}_h^*, \boldsymbol{u}_h)_h - F_h(\boldsymbol{u}_h), \tag{26d}$$

$$F_h(\boldsymbol{u}_h) := \left(\frac{|m_h^0|^2 + |m_h^1|^2}{2V_1(\rho_h)} + \frac{s_h^2}{2V_2(\rho_h)}, 1\right)_h + \Delta t\, \mathcal{E}_h(\rho_h). \tag{26e}$$

Practical ALG2 implemenation

Algorithm 2 One iteration of ALG2 algorithm for (26a).

- Step A: update Φ_h^ℓ. Find $\Phi_h^\ell \in V_h^k$ such that, for all $\forall \psi_h \in V_h^k$,

$$(\mathcal{D}\Phi_h^\ell, \mathcal{D}\psi_h)_h = (\boldsymbol{u}_h^{*,\ell-1} - \frac{1}{r}\boldsymbol{u}_h^{\ell-1}, \mathcal{D}\psi_h)_h - \frac{1}{r}(\rho_h^{\text{old}}, \psi_h)_h. \tag{27}$$

- Step B: update $\boldsymbol{u}_h^\ell$. Find ρ_h^ℓ such that it is the minimizer to the following functional of ρ_h

$$\frac{1}{2r}\left(|\rho_h - r\overline{\rho}_h|^2, 1\right)_h + \left(\frac{r^2(|\overline{m}_h^0|^2 + |\overline{m}_h^1|^2)}{2(r + V_1(\rho_h))}, 1\right)_h + \left(\frac{r^2\,|\overline{s}_h|^2}{2(r + V_2(\rho_h))}, 1\right)_h + \Delta t\,\mathcal{E}_h(\rho_h). \tag{28}$$

Then update $m_h^{0,\ell}, m_h^{1,\ell}, s_h^\ell$ according to

$$m_h^{0,\ell} = \frac{rV_1(\rho_h^\ell \overline{m}_h^0)}{r + V_1(\rho_h^\ell)}, \quad m_h^{1,\ell} = \frac{rV_1(\rho_h^\ell)\overline{m}_h^1}{r + V_1(\rho_h^\ell)}, \quad s_h^\ell = \frac{rV_2(\rho_h^\ell \overline{s}_h)}{r + V_2(\rho_h^\ell)}, \tag{29}$$

- Step C. Finally, update $\boldsymbol{u}_h^{*,\ell}$ according to $\boldsymbol{u}_h^{*,\ell} = \overline{\boldsymbol{u}}_h - \boldsymbol{u}_h^\ell / r$..

Some remarks

- One iteration of ALG2 Algorithm 2 amounts to one linear reaction-diffusion equation solve (27), a point-wise update of the nonlinear equation (28) per quadrature point, and some vector updates. Hence one ALG iteration is of linear computational complexity.
- In practice we take the ALG parameter $r = 1$, and apply 200 ALG iterations before moving to the next time step.
- The framework can be generalized to the reversible reaction-diffusion system case, where we apply further splitting in Step A/B of the ALG2 algorithm so that only *linear scalar reaction-diffusion* equations and *point-wise scalar nonlinear minimization problems* needs to be solved. (still of linear complexity)
- More details can be found in the preprint [FOL23].

Table of Contents

Ex1: Spatial convergence rates, Fokker-Plank equation

We first consider the nonlinear Fokker-Plank equation

$$\partial_t \rho - \triangle \rho^3 = \nabla \cdot (\rho\, x),$$

on the domain $\Omega = [-1,1] \times [-1,1]$ with homogeneous Neumann boundary conditions. It is a Wasserestein gradient flow with energy

$$\mathcal{E}(\rho) := \int_\Omega \left(\frac{1}{2}\rho(x)^3 + \frac{1}{2}(x_0^2 + x_1^2)\rho(x) \right) dx,$$

where $x = (x_0, x_1)$. This problem reaches a steady state solution

$$\rho_{\text{steady}}(x_1, x_2) = \sqrt{\frac{(2C - (x_0^2 + x_1^2))_+}{3}},$$

that satisfies either

$$\frac{\delta \mathcal{E}}{\delta \rho} = \frac{3}{2}\rho^2 + \frac{1}{2}(x_0^2 + x_1^2) = C,$$

or $\rho = 0$. Here the constant C depends on the total mass of the initial condition, which we set to be $C = 2$ so that the solution on Ω is positive and smooth.

Ex1: Spatial convergence rates, Fokker-Plank equation

The L^2-convergence in the density ρ is recorded in Table 1.

Table 1: Convergence rates with different polynomial degree k applied to a 2D steady Fokker Plank equation.

$\dim(V_h^k)$	$k=1$		$k=2$		$k=4$	
81	2.362e-03	–	2.409e-04	–	2.628e-05	–
289	5.923e-04	2.00	3.298e-05	2.87	1.424e-06	4.21
1089	1.482e-04	2.00	4.232e-06	2.96	5.589e-08	4.67
4225	3.705e-05	2.00	5.326e-07	2.99	1.884e-09	4.89

Ex2: Aggregation-drift-diffusion equations

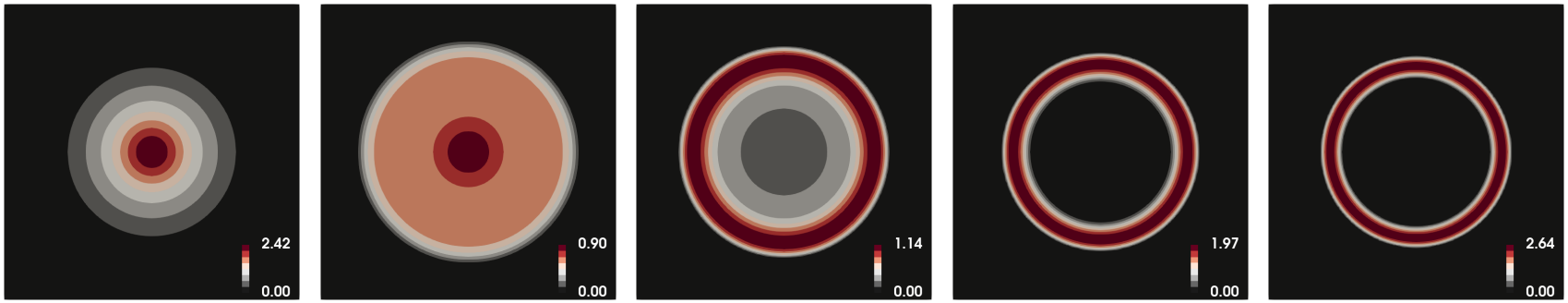
We consider Wasserstein gradient flow with five choices of energy (2) that including aggregation effects.

Table 2: Example 2. Five choices of energies, domain size, and initial condition.

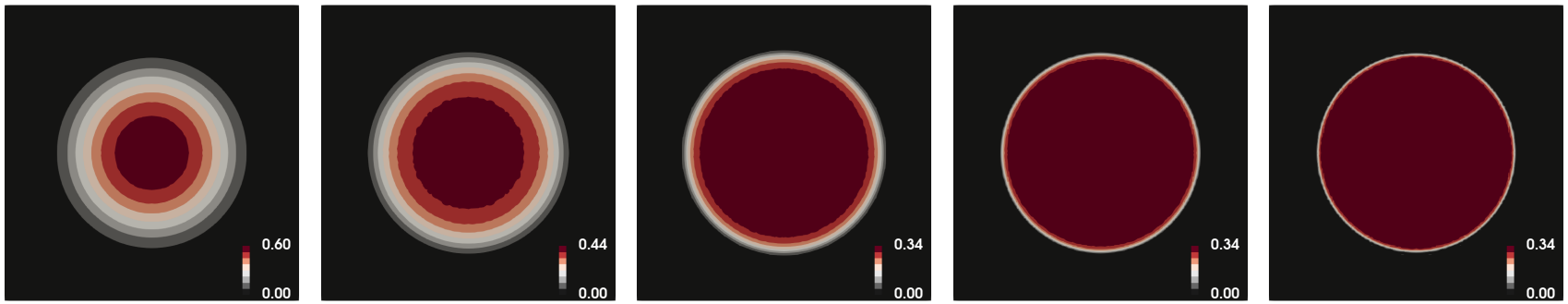
Case	$\alpha U_m(\rho)$	$V(x)$	$W(x)$	L	I.C.
1	0	0	$\frac{\lvert x\rvert^4}{4} - \frac{\lvert x\rvert^2}{2}$	1	$\frac{25}{2\pi}\exp(-\frac{25}{2}\lvert x\rvert^2)$
2	0	0	$\frac{\lvert x\rvert^2}{2} - \log(\lvert x\rvert)$	1.5	$\frac{25}{8\pi}\exp(-\frac{25}{8}\lvert x\rvert^2)$
3	0	$-\frac{1}{4}\log(\lvert x\rvert)$	$\frac{\lvert x\rvert^2}{2} - \log(\lvert x\rvert)$	1.5	$\frac{25}{8\pi}\exp(-\frac{25}{8}\lvert x\rvert^2)$
4	$0.1\rho^2$	$-\frac{1}{4}\log(\lvert x\rvert)$	$\frac{\lvert x\rvert^2}{2} - \log(\lvert x\rvert)$	1.5	$\frac{25}{8\pi}\exp(-\frac{25}{8}\lvert x\rvert^2)$
5	$0.1\rho^3$	0	$-\exp(-\lvert x\rvert^2)/\pi$	4	$0.25\chi_{[-3,3]\times[-3,3]}$

Ex2: Aggregation-drift-diffusion equations

For all cases, we take a computational domain with a 32×32 uniform square mesh, and use polynomial degree $k = 4$.

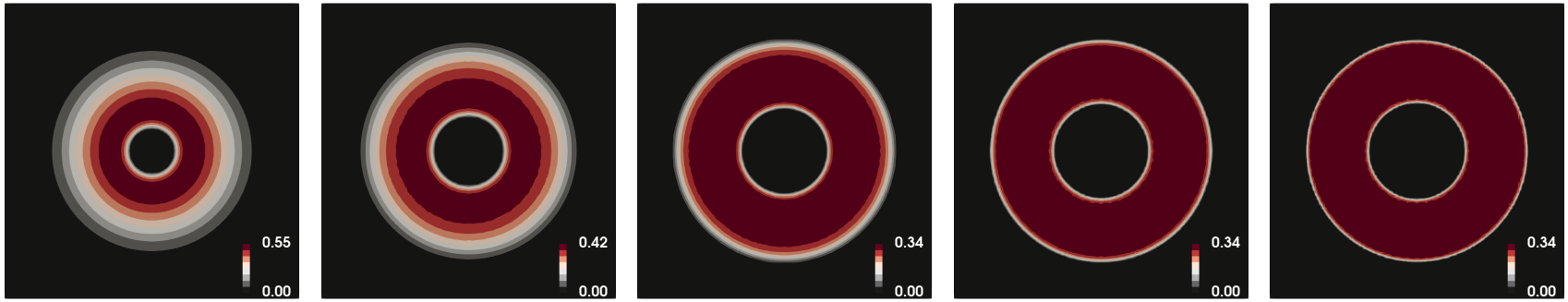


(a) Case 1. Left to right time: $t = 0.5, 1.5, 3.0, 6.0, 10$

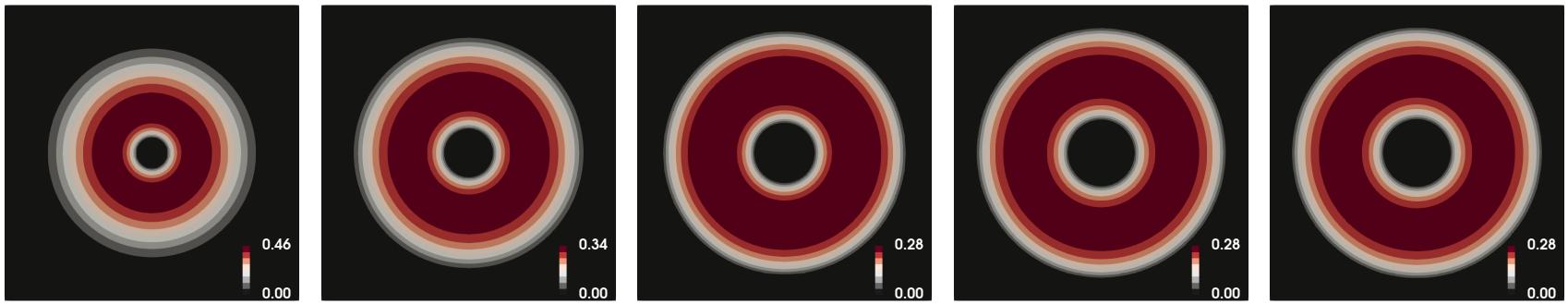


(b) Case 2. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$

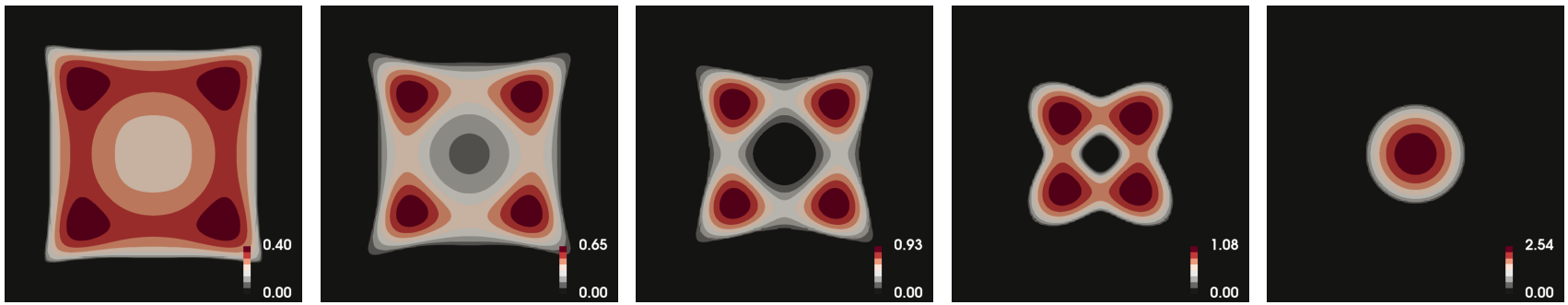
Figure 1: Snapshots of density contour at different times for different test cases.



(a) Case 3. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$



(b) Case 4. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$



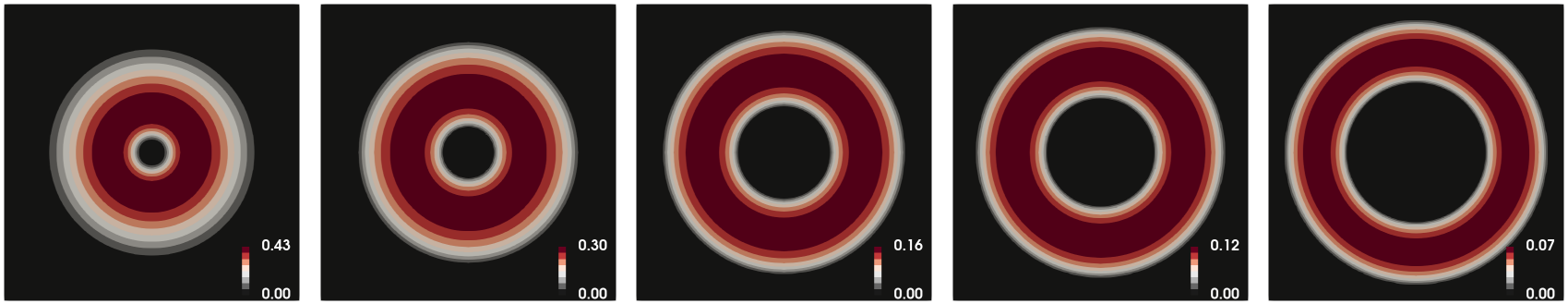
(c) Case 5. Left to right time: $t = 2, 4, 6, 10, 15$

Ex3: Scalar reaction-diffusion equation

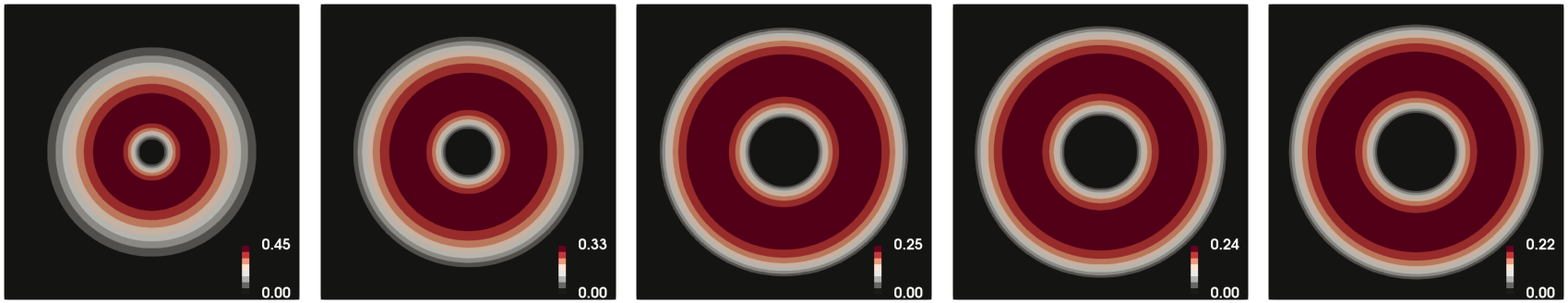
We take the Case 4 energy in Table 2, but consider the reaction-diffusion equation. Three choices of mobility coefficient $V_2(\rho)$ are used in this example, namely,

$$\begin{cases} \text{Type 1: } V_2(\rho) = 0.1, \\ \text{Type 2: } V_2(\rho) = 0.1\rho, \\ \text{Type 3: } V_2(\rho) = 0.1\frac{\rho-1}{\log(\rho)}. \end{cases} \tag{30}$$

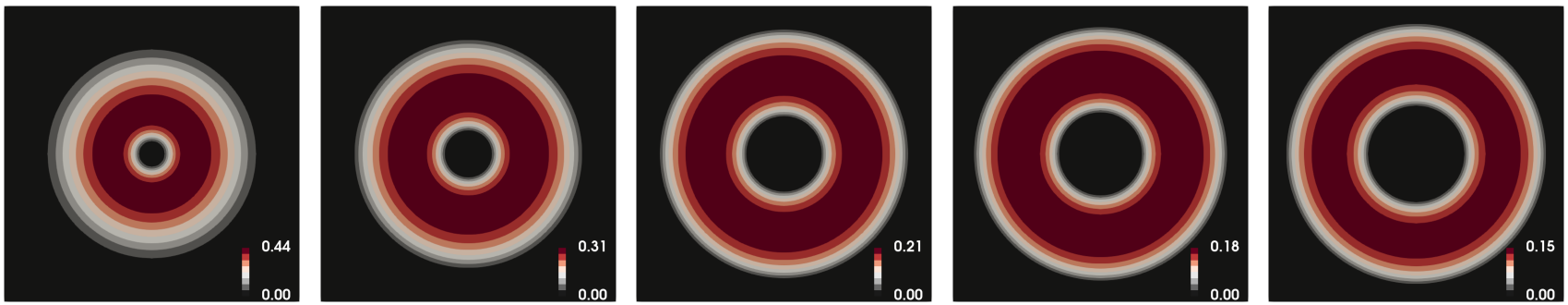
The same discretization setup as in the previous example is used, i.e., using polynomial degree $k = 4$ on a 32×32 uniform mesh with time step size $\Delta t = 0.05$, and final time $T = 3$.



(a) Case 4 energy, Type 1 reaction. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$



(b) Case 4 energy, Type 2 reaction. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$



(c) Case 4 energy, Type 3 reaction. Left to right time: $t = 0.2, 0.5, 1.5, 2.0, 3.0$

Ex4: Fisher-KPP equation

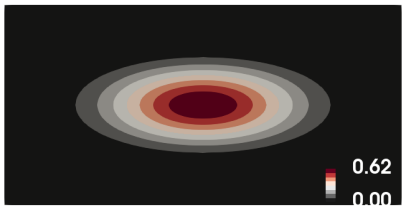
The PDE:

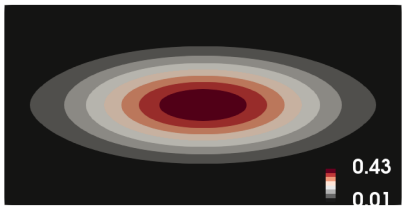
$$\partial_t \rho - \lambda_1 \partial_{x_0 x_0} \rho - \lambda_2 \partial_{x_1 x_1} \rho = \mu \rho (1 - \rho).$$

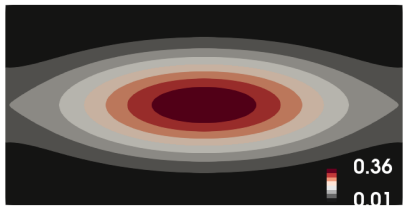
Diffusion parameters $\lambda_1 = 0.1, \lambda_2 = 0.01$, and $\mu > 0$ is the reaction coefficient to be specified. Initial condition is a flat top Gaussian:

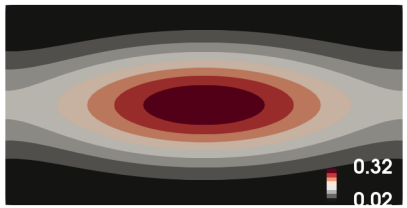
$$\rho_0(x_0, x_1) = \begin{cases} 1, & \text{if } x_0^2 + 4x_1^2 \leq 0.25 \\ \exp(-10(x_0^2 + 4x_1^2 - 0.25)), & \text{otherwise} \end{cases}$$

The computational domain is a rectangle $\Omega = [-2, 2] \times [-1, 1]$, which is discretized with a 32×16 square mesh. We use polynomial degree $k = 4$ We take time step size $\Delta t = 0.1$ and final time is $T = 4$.

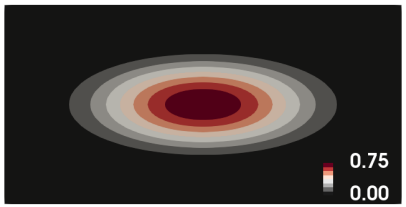


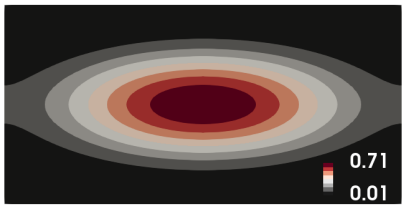


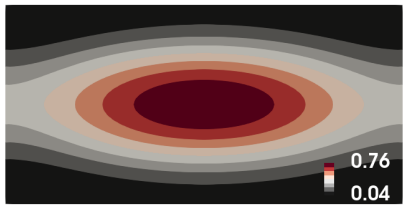


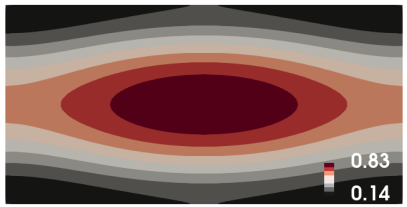


(a) Reaction coefficient $\nu = 0.1$. Left to right time: $t = 1, 2, 3, 4$

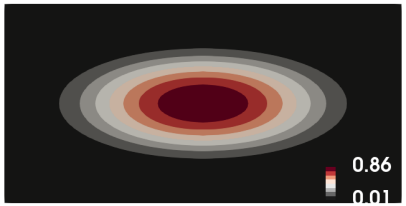


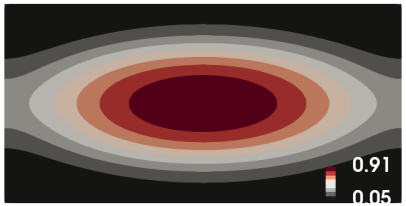


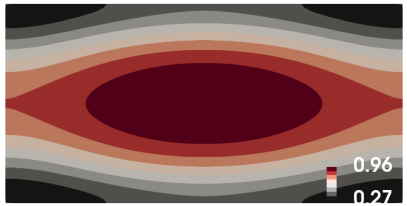


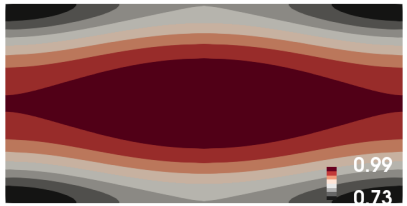


(b) Reaction coefficient $\mu = 0.5$. Left to right time: $t = 1, 2, 3, 4$









(c) Reaction coefficient $\mu = 1.0$. Left to right time: $t = 1, 2, 3, 4$

Figure 4: Snapshots of density contour at different times for different reaction coefficients.

Ex4: KPP, energy/mass evolution

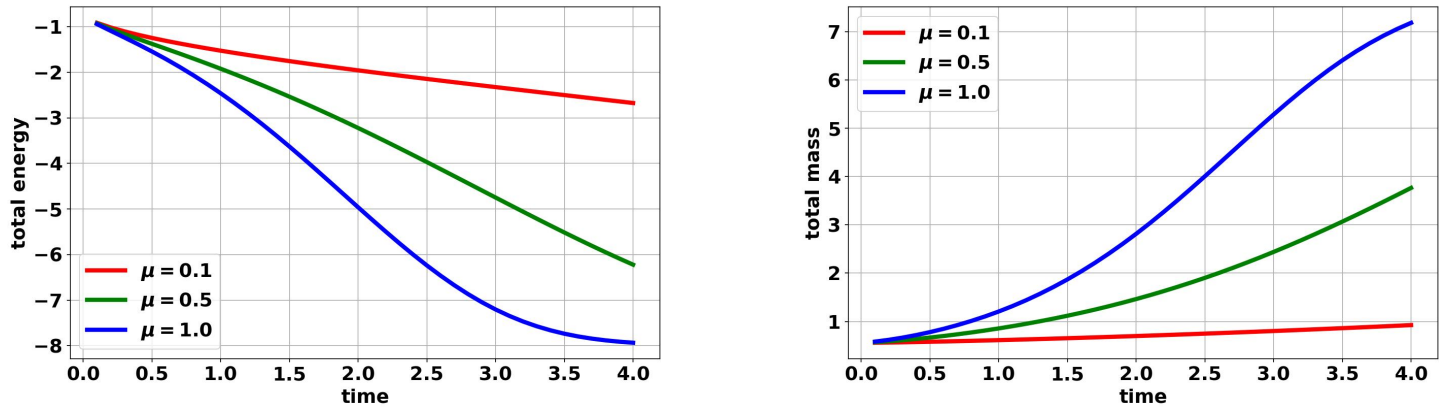


Figure 5: Evolution of total energy (left) and total mass (right) over time.

Ex5: Two-component reversible reaction-diffusion system

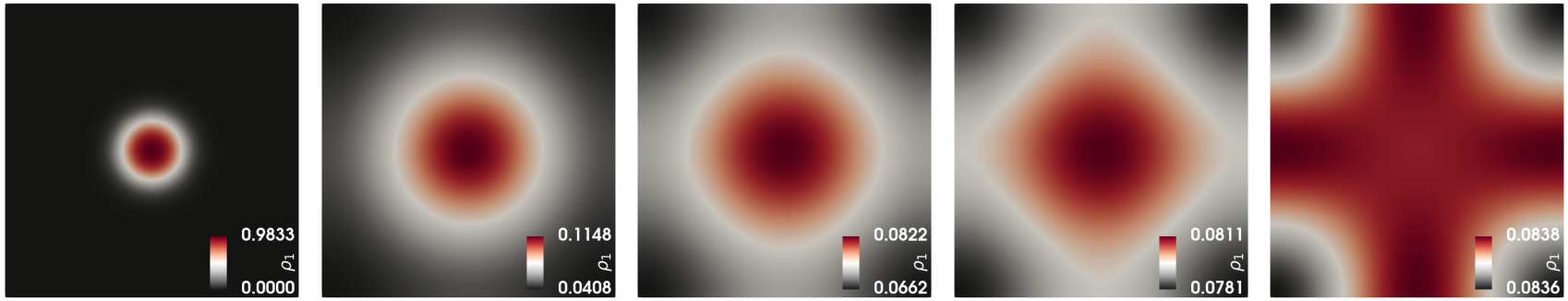
We consider the two-species model [LWW21]:

$$\partial_t \rho_1 - \frac{\gamma_1}{m}\Delta \rho_1^m = -(k_+ \rho_1 \rho_2^2 - k_- \rho_2^3),$$
$$\partial_t \rho_2 - \gamma_2 \Delta \rho_2 = (k_+ \rho_1 \rho_2^2 - k_- \rho_2^3).$$

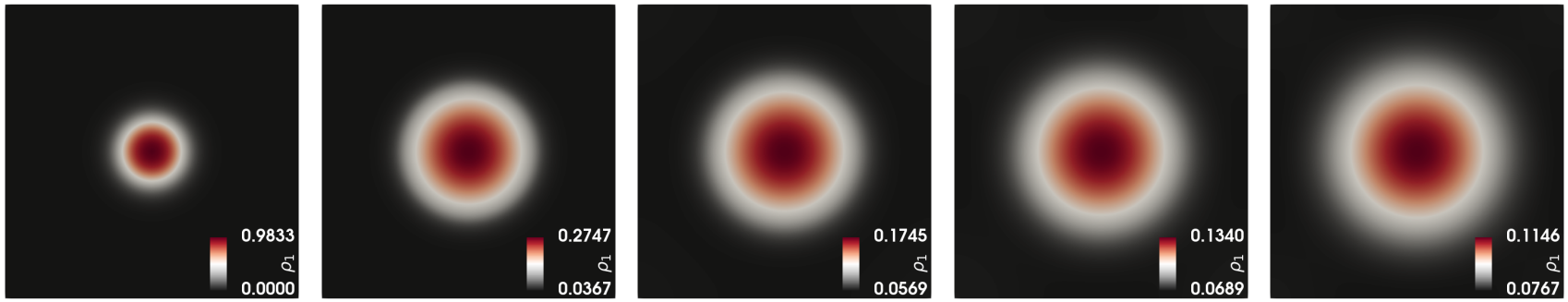
It can be equivalently written as

$$\partial_t \rho_1 = \nabla \cdot \left(V_{1,1}(\rho_1) \nabla \frac{\delta \mathcal{E}_1}{\delta \rho}(\rho_1) \right) - V_2(\rho_1, \rho_2) \left(\frac{\delta \mathcal{E}_1}{\delta \rho}(\rho_1) - \frac{\delta \mathcal{E}_2}{\delta \rho}(\rho_2) \right),$$
$$\partial_t \rho_2 = \nabla \cdot \left(V_{1,2}(\rho_2) \nabla \frac{\delta \mathcal{E}_2}{\delta \rho}(\rho_2) \right) + V_2(\rho_1, \rho_2) \left(\frac{\delta \mathcal{E}_1}{\delta \rho}(\rho_1) - \frac{\delta \mathcal{E}_2}{\delta \rho}(\rho_2) \right),$$

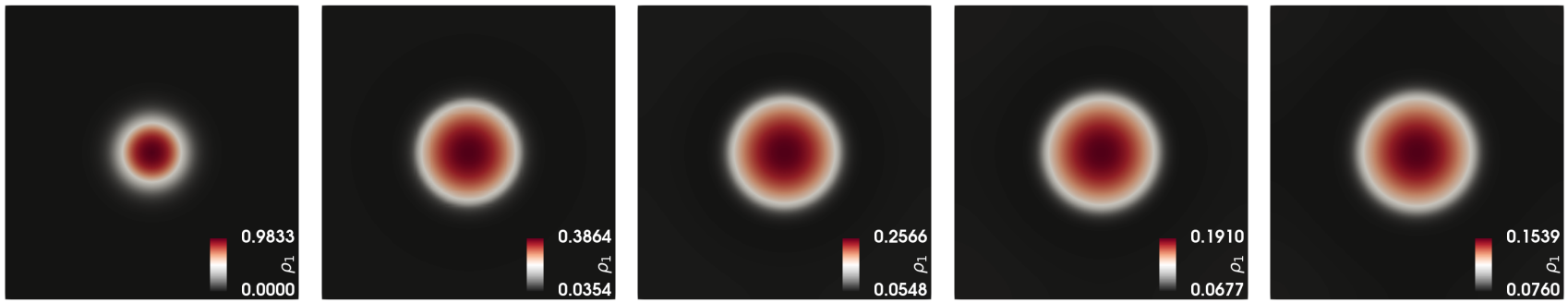
where $V_{1,1}(\rho_1) = \gamma_1 (\rho_1)^m$, $V_{1,2}(\rho_1) = \gamma_2 \rho_2$, $V_2(\rho_1, \rho_2) = \ell(\kappa_1 \rho_1 \rho_2^2, \kappa_2 \rho_2^3)$, with $\mathcal{E}_i(\rho_i) = \rho_i(\log(\kappa_i \rho_i) - 1)$, and $\ell(x, y) = \frac{x-y}{\log(x) - \log(y)}$.
We take parameters $k_+ = 1$ and $k_- = 0.1$, $\gamma_1 = 0.2$, $\gamma_2 = 0.1$.



(a) $V_{1,1}(\rho) = \gamma_1 \rho$. Left to right time: $t = 0, 0.5, 1, 1.5, 2$



(b) $V_{1,1}(\rho) = \gamma_1 \rho^2$. Left to right time: $t = 0, 0.5, 1, 1.5, 2$



(c) $V_{1,1}(\rho) = \gamma_1 \rho^3$. Left to right time: $t = 0, 0.5, 1, 1.5, 2$

Ex5: total energy/mass evolution

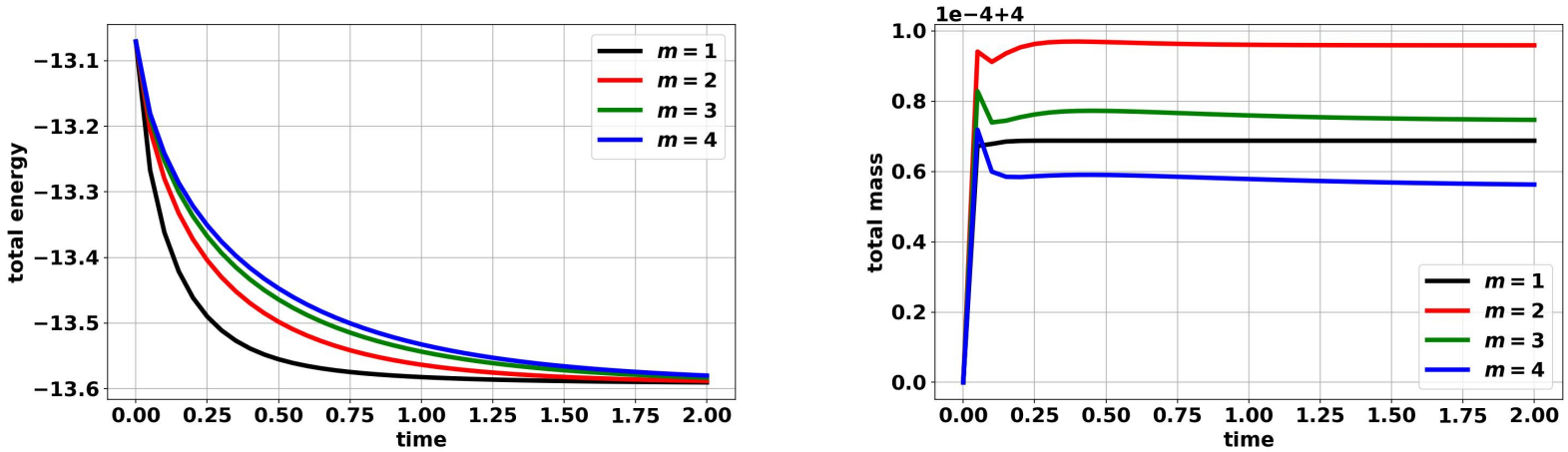


Figure 6: Example 5. Evolution of total energy (left) and total mass (right) over time with $V_{1,1}(\rho) = \gamma_1 \rho^m$.

Ex6: Reversible Gray-Scott model

We consider the 4-component reversible Gray-Scott model [Lia+22]:

$$\partial_t \rho_1 = \gamma_1 \Delta \rho_1 - (k_+^1 \rho_1 \rho_2^2 - k_-^1 \rho_2^3) - (k_+^3 \rho_1 - k_-^3 \rho_4),$$
$$\partial_t \rho_2 = \gamma_2 \Delta \rho_2 + (k_+^1 \rho_1 \rho_2^2 - k_-^1 \rho_2^3) - (k_+^2 \rho_2 - k_-^2 \rho_3),$$
$$\partial_t \rho_3 = (k_+^2 \rho_2 - k_-^2 \rho_3), \quad \partial_t \rho_4 = (k_+^3 \rho_1 - k_-^3 \rho_4).$$

The physical parameters are chosen to be the following:

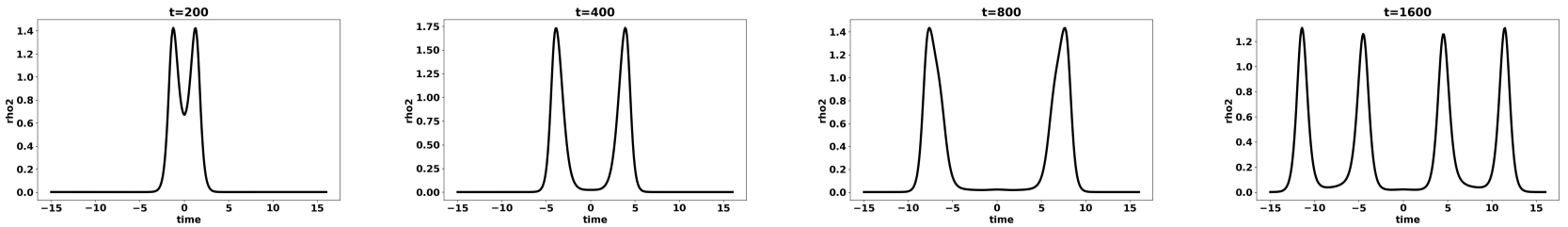
$$\gamma_1 = 1, \gamma_2 = 0.01, k_+^1 = 1, k_+^2 = 0.084, k_+^3 = 0.024, \quad k_-^i = 10^{-3} k_+^i.$$

This provides a good approximation to the irreversible Gray-Scott model:

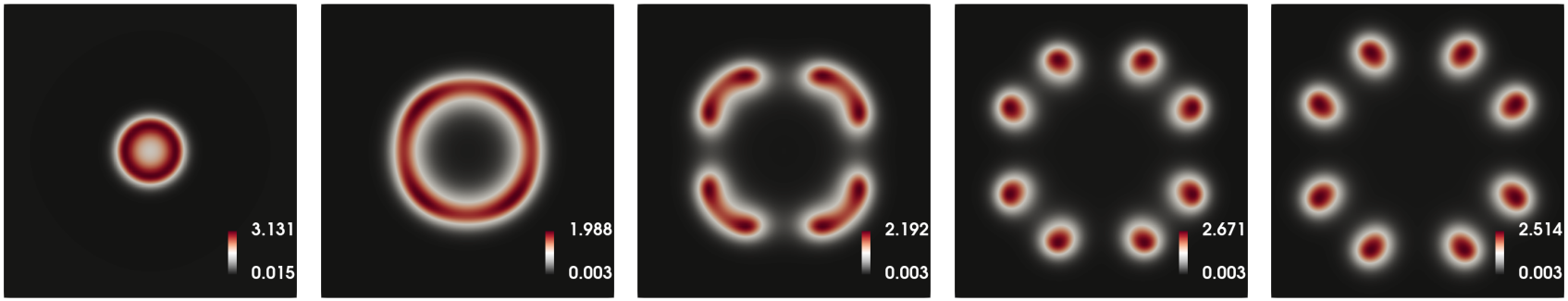
$$\partial_t \rho_1 = \gamma_1 \Delta \rho_1 - k_+^1 \rho_1 \rho_2^2 - k_+^3 (\rho_1 - 1),$$
$$\partial_t \rho_2 = \gamma_2 \Delta \rho_2 + k_+^1 \rho_1 \rho_2^2 - k_+^2 \rho_2,$$

which is widely used in pattern formations.

Ex6: density contour



(a) 1D results. Left to right time: $t = 200, 400, 800, 1600$.



(b) 2D results. Left to right time: $t = 100, 200, 300, 400, 500$

Figure 7: Snapshots of second-component density contour ρ_2 at different times for 1D (top) and 2D (bottom) simulations.

Ex6: total energy evolution

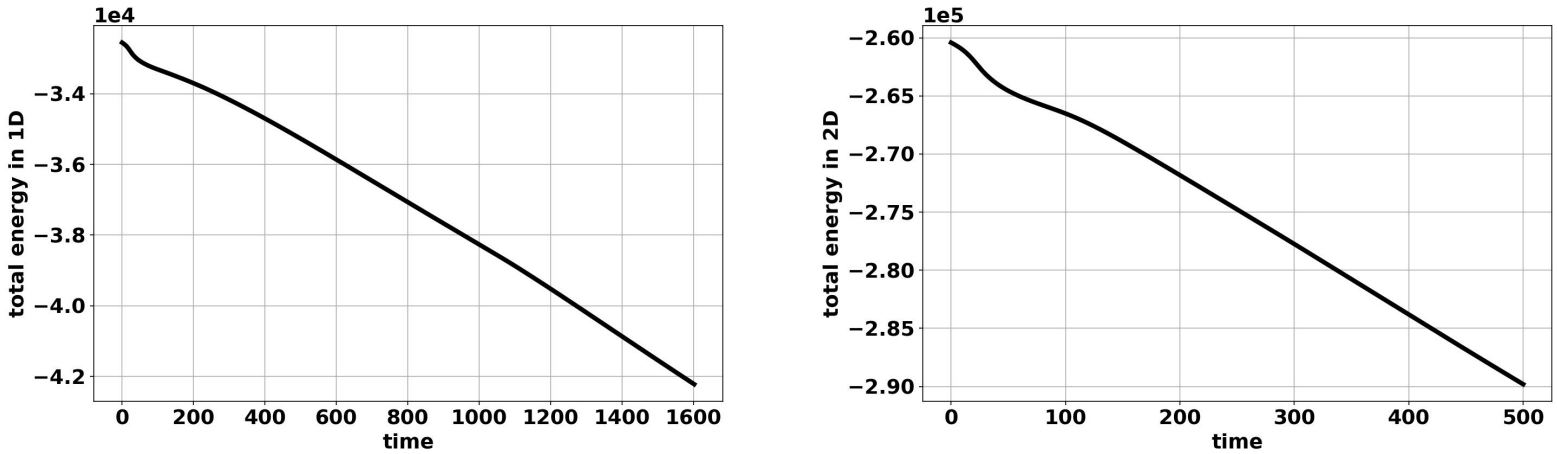


Figure 8: Evolution of total energy in 1D (left) and 2D (right).

*** show videos

Table of Contents

Dynamic Mean Field Planning

The goal of mean field planning (MFP) is to minimize a total cost

$$\inf_{(\rho,m)} \int_0^1 \int_\Omega L(\rho(t,x), m(t,x)) + A(\rho(t,x))\, dxdt. \tag{31}$$

among all feasible density and flux (ρ, m) such that

$$\partial_t \rho + \nabla_x \cdot m = 0 \ \text{ in } [0,1] \times \Omega, \quad \rho(0,\cdot) = \rho_0, \rho(1,\cdot) = \rho_1,$$

given initial and terminal densities ρ_0 and ρ_1.

- The dynamic cost function $L(\rho, m) = \frac{\|m\|^2}{2\rho}$ is related to optimal transport (and Wasserstein gradient flow).
- The interaction cost function $A(\rho)$ is usually taken to be convex.
- We are interested in the transport density $\rho(t,\cdot)$ for all $t \in (0,1)$.

Dynamic Mean Field Game (MFG)

For MFG, the terminal density ρ_1 is not explicitly provided but it satisfies a given preference. The goal of MFG is to minimize the total cost

$$\inf_{(\rho,m)} \int_0^1 \int_\Omega \left[L(\rho, m) + A(\rho) \right] dxdt + \underbrace{\int_\Omega \Gamma(\rho(1,x))dx}_{:=R(\rho(1,\cdot))}, \tag{32}$$

among all feasible (ρ, m) such that

$$\partial_t \rho + \nabla \cdot m = 0, \quad \text{in } [0,1] \times \Omega, \quad \rho(0,\cdot) = \rho_0.$$

- The MFG problem (32) is identical to a JKO step of Wasserstein gradient flow (4) if we take the terminal cost to be $\Delta t \mathcal{E}(\rho(1,\cdot))$ and remove the interaction cost $A(\rho) = 0$.
- But this time, we are more interested in the evolution of density along $t \in [0,1]$. Hence we typically do not use the one-step relaxation approach (6), which is of first-order accuracy in time.

The augmented Lagrangian formulation of MFG [BC15]

We reformulate the problem (32) into a saddle-point problem:

$$\inf_{\boldsymbol{u},\rho_1} \sup_{\Phi} F(\boldsymbol{u}) + R(\rho_1) - G(\phi) - (\boldsymbol{u}, \mathcal{D}\Phi)_{[0,T]\times\Omega}, \tag{33a}$$

in which $\boldsymbol{u} = (\rho, m)$, and

$$F(\boldsymbol{u}) := \int_0^1 \int_\Omega \left[L(\rho, m) + A(\rho)\right] dxdt, \tag{33b}$$

$$G(\phi) := \int_\Omega \left[-\phi(1,x)\rho_1(x) + \phi(0,x)\rho_0(x)\right] dx, \tag{33c}$$

$$\mathcal{D}\Phi := (\partial_t\Phi, \nabla_x\Phi) \text{ is the space-time gradient} \tag{33d}$$

This problem is of the form (19) and can be tackled via the ALG2 algorithm [BC15].

High-order computation for MFG

- Our contribution in [Fu+23] is to discretize the saddle-point problem in (33) using high-order **space-time** finite element spaces on $[0,1] \times \Omega$.
- In particular, we discretize Φ using a high-order space-time continuous finite element space V_h^k, and the other physical variables using a high-order **nodal** discontinuous integration rule space W_h^k.
- It achieves high-order accuracy in both space and time. (first in the literature)

A MFG Example

4.3. MFG with obstacles

We consider a similar setting as in Example 4.2, where we consider a MFG problem with terminal cost

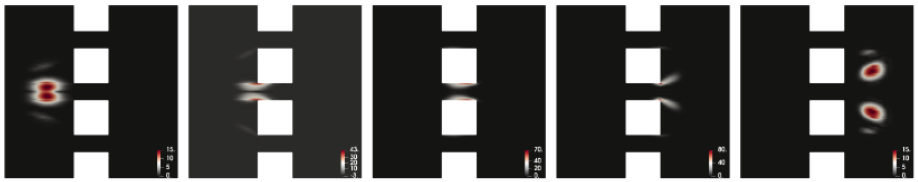
$$\Gamma(\rho) := \begin{cases} \frac{1}{2}(\rho - \rho_T)^2 & \text{if } \rho \geq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

where the target density

$$\rho_T := \frac{1}{2\pi\sigma^2}\left(\exp(-\frac{1}{2\sigma^2}|\boldsymbol{x} - (0.65, 0.3)|^2) + \exp(-\frac{1}{2\sigma^2}|\boldsymbol{x} - (0.65, -0.3)|^2)\right)$$

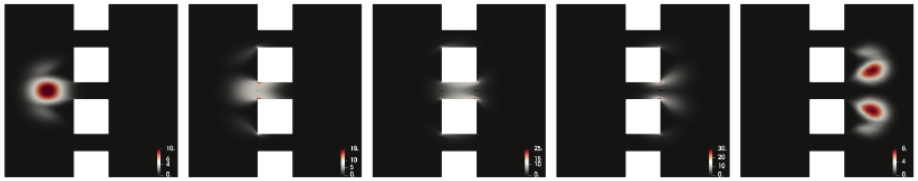
with $\sigma = 0.1$. Note that we allow ρ_T and ρ_0 to have different total masses here.

$$\begin{cases} \text{Case 1}: A(\rho) = 0, \quad A^*(\rho^*) = \begin{cases} 0 & \text{if } \rho^* \leq 0, \\ +\infty & \text{if } \rho^* > 0. \end{cases}, \\ \text{Case 2}: A(\rho) = c\rho^2, \quad A^*(\rho^*) = \begin{cases} 0 & \text{if } \rho^* \leq 0, \\ (\rho^*)^2/(4c) & \text{if } \rho^* > 0. \end{cases}, \\ \text{Case 3}: A(\rho) = c\rho\log(\rho), \quad A^*(\rho^*) = \exp(\rho^*/c - 1), \\ \text{Case 4}: A(\rho) = c/\rho, \quad A^*(\rho^*) = \begin{cases} -2\sqrt{-c\rho^*} & \text{if } \rho^* \leq 0, \\ +\infty & \text{if } \rho^* > 0. \end{cases} \\ \text{Case 5}: A(\rho) = \begin{cases} 0 & \text{if } 0 \leq \rho \leq \rho_{\max}, \\ +\infty & \text{else.} \end{cases} \quad A^*(\rho^*) = \rho_{\max}(\rho^*)_+ \end{cases}$$



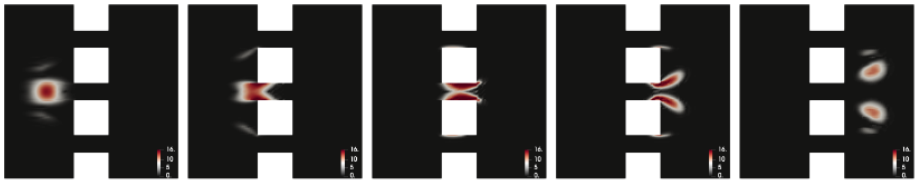
(a) Case 1

(b) Case 2



(c) Case 3

(d) Case 4



(e) Case 5

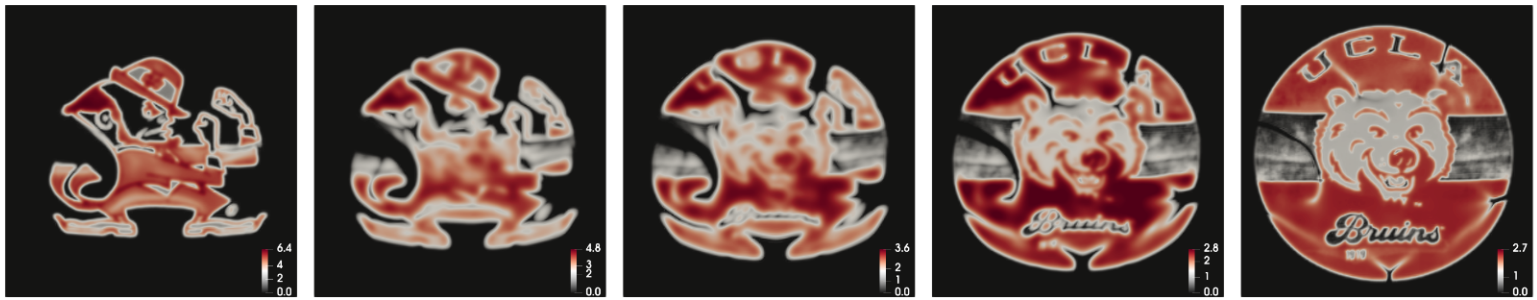
Figure 4: Example 4.3. Snapshots of ρ at $t = 0.1, 0.3, 0.5, 0.7, 0.9$ (left to right).

A MFP Example between mascot images

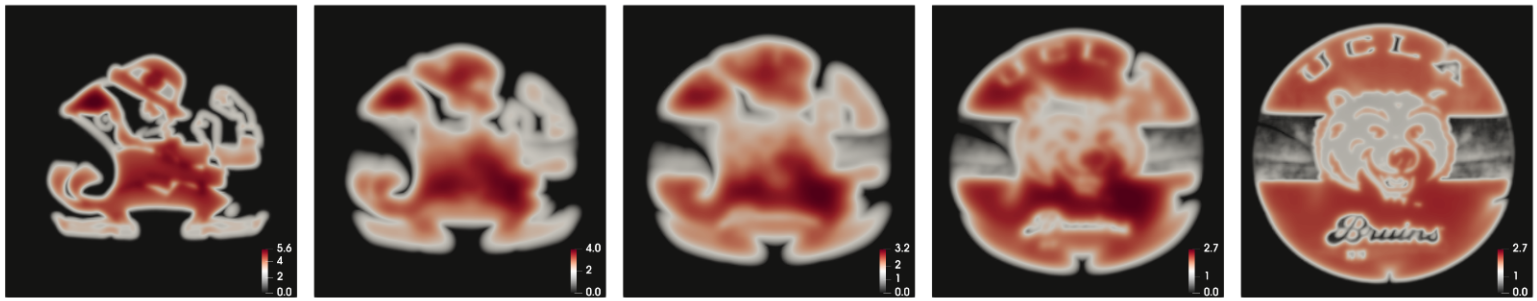
(a) ND (Leprechaun) (b) UCLA (Bruins) (c) USC (Gamecocks)

Figure 5: Example 4.4. Initial/final densities.

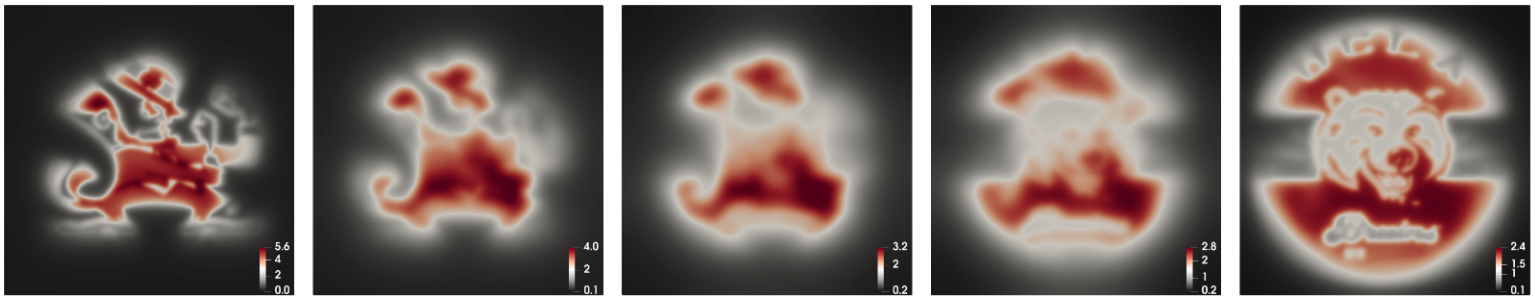
A MFP Example: ND→ UCLA



(a) Case 1: $A(\rho) = 0$. ND $\to$ UCLA

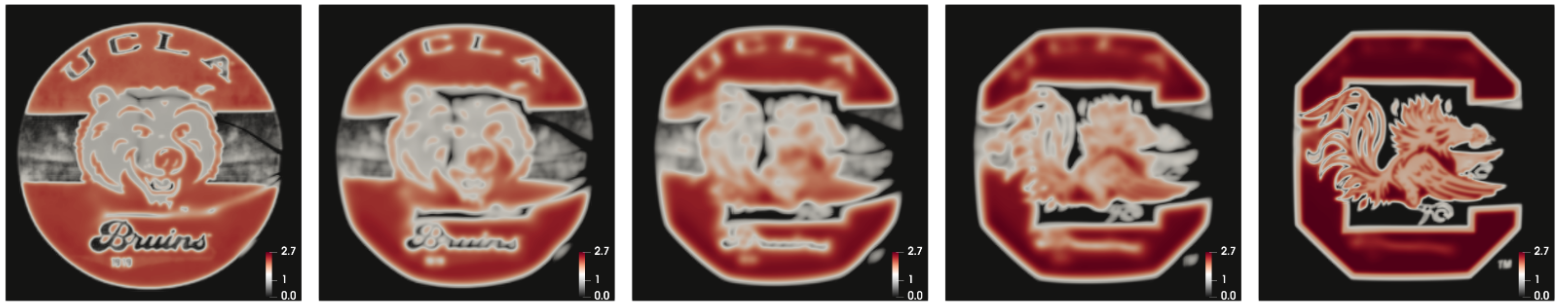


(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. ND $\to$ UCLA

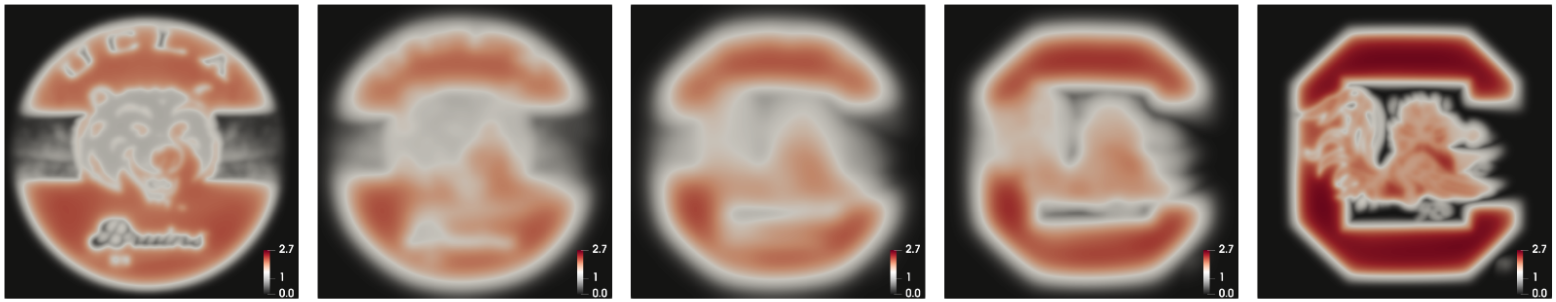


(c) Case 3: $A(\rho) = 0.01/\rho$. ND $\to$ UCLA

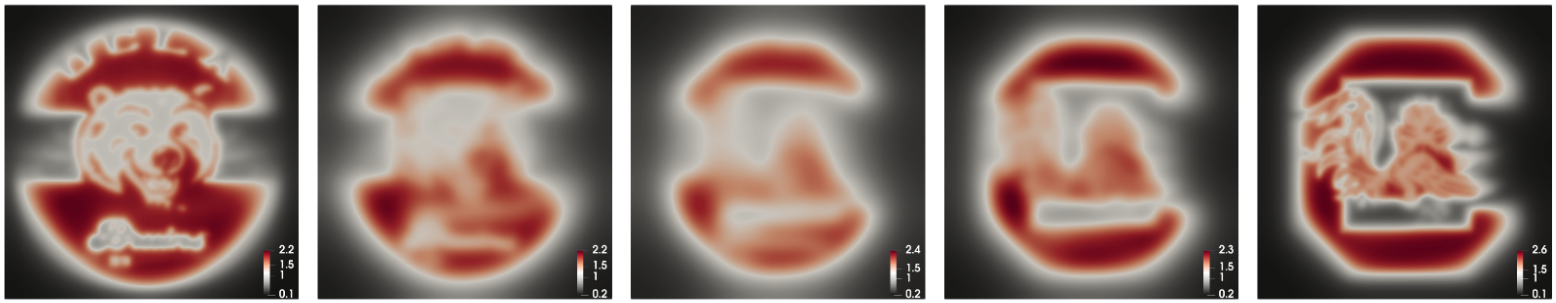
A MFP Example: UCLA→ USC



(a) Case 1: $A(\rho) = 0$. UCLA $\to$ USC

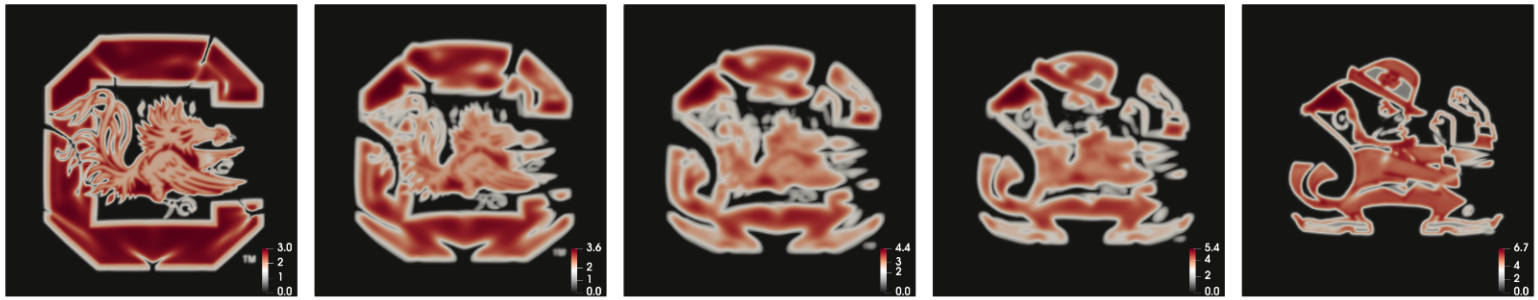


(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. UCLA $\to$ USC

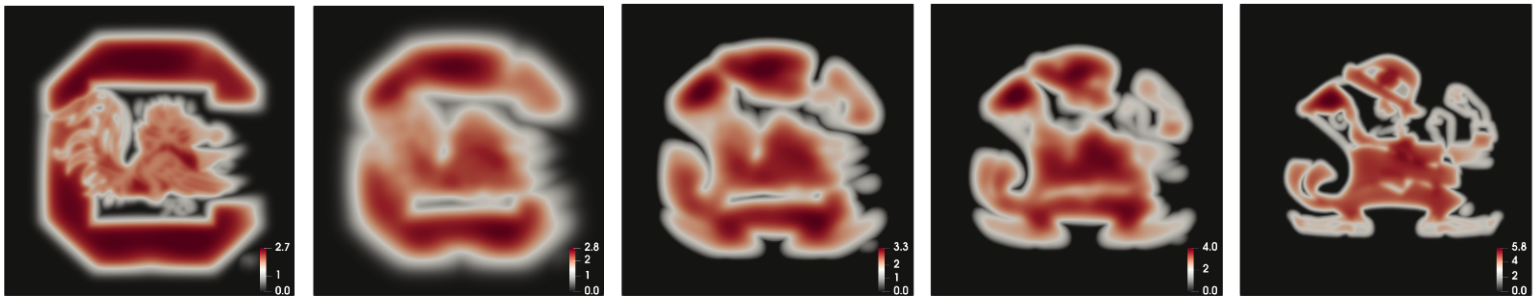


(c) Case 3: $A(\rho) = 0.01/\rho$. UCLA $\to$ USC

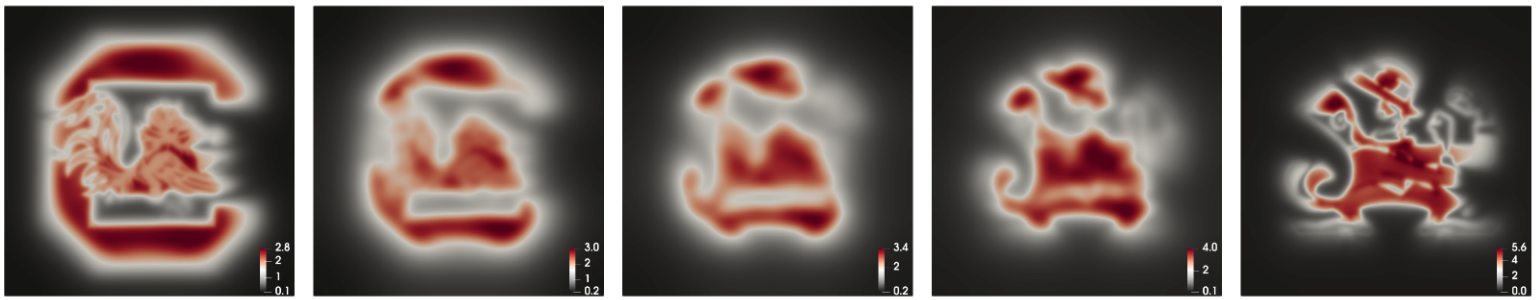
A MFP Example: USC→ND *** show videos



(a) Case 1: $A(\rho) = 0$. USC $\to$ ND



(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. USC $\to$ ND



(c) Case 3: $A(\rho) = 0.01/\rho$. USC $\to$ ND

Table of Contents

Conclusion

- A high-order spatial FEM discretization of variational time implicit schemes for dissipative reaction-diffusion systems.
- A high-order space-time FEM discretization of OT/MFP/MFG.
- Proximal splitting optimization solver (ALG2) with linear complexity for each ALG iteration.

Thanks for your attention!

References I

[BB00] Jean-David Benamou and Yann Brenier. "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem". In: *Numerische Mathematik* 84.3 (2000), pp. 375–393.

[BC15] Jean-David Benamou and Guillaume Carlier. "Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations". In: *Journal of Optimization Theory and Applications* 167.1 (2015), pp. 1–26.

[BCL16] Jean-David Benamou, Guillaume Carlier, and Maxime Laborde. "An augmented Lagrangian approach to Wasserstein gradient flows and applications". In: *ESAIM: Proceedings and Surveys* 54 (2016), pp. 1–17.

References II

[Car+22] José A. Carrillo et al. "Primal dual methods for Wasserstein gradient flows". In: *Found. Comput. Math.* 22.2 (2022), pp. 389–443. ISSN: 1615-3375. DOI: 10.1007/s10208-021-09503-1. URL: https://doi-org.proxy.library.nd.edu/10.1007/s10208-021-09503-1.

[CGT20] Clément Cancès, Thomas O. Gallouët, and Gabriele Todeschi. "A Variational Finite Volume Scheme for Wasserstein Gradient Flows". In: *Numerische Mathematik* 146.3 (2020), pp. 437–480.

[Chi+18] Lénaıc Chizat et al. "Unbalanced optimal transport: dynamic and Kantorovich formulations". In: *J. Funct. Anal.* 274.11 (2018), pp. 3090–3123. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2018.03.008. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jfa.2018.03.008.

References III

[CP11] Antonin Chambolle and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *J. Math. Imaging Vision* 40.1 (2011), pp. 120–145. ISSN: 0924-9907. DOI: 10.1007/s10851-010-0251-1. URL: https://doi-org.proxy.library.nd.edu/10.1007/s10851-010-0251-1.

[FG83] Michel Fortin and Roland Glowinski. *Augmented Lagrangian methods*. Vol. 15. Studies in Mathematics and its Applications. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer. North-Holland Publishing Co., Amsterdam, 1983, pp. xix+340. ISBN: 0-444-86680-9.

References IV

[FOL23] Guosheng Fu, Stanley Osher, and Wuchen Li. "High order spatial discretization for variational time implicit schemes: Wasserstein gradient flows and reaction-diffusion systems". In: *arXiv:2303.08950 [math.NA]* (2023).

[Fu+23] Guosheng Fu et al. "High order computation of optimal transport, mean field planning, and mean field games". In: *arXiv:2302.02308 [math.NA]* (2023).

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. "The Variational Formulation of the Fokker–Planck Equation". In: *SIAM Journal on Mathematical Analysis* 29.1 (1998), pp. 1–17. DOI: 10.1137/S0036141096303359.

[Lee+21] Wonjun Lee et al. "Generalized unnormalized optimal transport and its fast algorithms". In: *J. Comput. Phys.* 436 (2021), Paper No. 110041, 24. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2020.110041. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jcp.2020.110041.

References V

[Lia+22] Jiangyan Liang et al. "On a reversible Gray-Scott type system from energetic variational approach and its irreversible limit". In: *J. Differential Equations* 309 (2022), pp. 427–454. ISSN: 0022-0396. DOI: 10.1016/j.jde.2021.11.032. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jde.2021.11.032.

[LLO22] Wuchen Li, Wonjun Lee, and Stanley Osher. "Computational mean-field information dynamics associated with reaction-diffusion equations". In: *J. Comput. Phys.* 466 (2022), Paper No. 111409, 30. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2022.111409. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jcp.2022.111409.

References VI

[LLW20] Wuchen Li, Jianfeng Lu, and Li Wang. "Fisher information regularization schemes for Wasserstein gradient flows". In: *J. Comput. Phys.* 416 (2020), pp. 109449, 24. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2020.109449. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jcp.2020.109449.

[LWW21] Chun Liu, Cheng Wang, and Yiwei Wang. "A structure-preserving, operator splitting scheme for reaction-diffusion equations with detailed balance". In: *J. Comput. Phys.* 436 (2021), Paper No. 110253, 22. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2021.110253. URL: https://doi-org.proxy.library.nd.edu/10.1016/j.jcp.2021.110253.

References VII

[Mie11] Alexander Mielke. "A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems". In: *Nonlinearity* 24.4 (2011), pp. 1329–1346. ISSN: 0951-7715. DOI: 10.1088/0951-7715/24/4/016. URL: https://doi-org.proxy.library.nd.edu/10.1088/0951-7715/24/4/016.

[PPO14] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. "Optimal transport with proximal splitting". In: *SIAM Journal on Imaging Sciences* 7.1 (2014), pp. 212–238.