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
1 Background: JKO scheme for Wasserstein gradient flow
2 Variational time implicit scheme for dissipative systems
3 Optimization solver: the ALG2 Algorithm
4 High-order spatial discretization of variational time implicit schemes
5 Numerics
6 High-order computation of OT/MFP/MFG
7 Conclusion
The Wasserstein gradient flow

PDE:

\[ \partial_t \rho = \nabla \cdot \left( \rho \nabla \frac{\delta}{\delta \rho} \mathcal{E} \right), \quad (1) \]

with energy functional:

\[ \mathcal{E}(\rho) := \int_{\Omega} \left[ \alpha U_m(\rho(x)) + \rho(x) V(x) + \frac{1}{2} (W * \rho)(x) \rho(x) \right] \, dx, \quad (2) \]

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

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

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

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

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:

\[
\text{Dist}_{W_2}(\rho^0, \rho^1)^2 := \inf_{\nu, \rho} \int_0^1 \int_{\Omega} \|\nu(x, \tau)\|^2 \rho(x, \tau) \, dx \, d\tau,
\]

where the infimum is taken among \(\rho, \nu\) such that

\[
\partial_{\tau} \rho(x, \tau) + \nabla \cdot (\rho(x, \tau) \nu(x, \tau)) = 0, \quad \rho(x, 0) = \rho^0(x), \quad \rho(x, 1) = \rho^1(x).
\]

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

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

such that

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

\[ \rho^n = \arg \min_{\rho \in \mathcal{M}} \quad \frac{1}{2\Delta t} \int_{\Omega} \left( \frac{\|m(x)\|^2}{\rho(x)} \right) dx + \mathcal{E}(\rho), \]  

(6a)

where the minimization is over all functions \( m, \rho \) such that

\[ \rho(x) - \rho^{n-1}(x) + \nabla \cdot m(x) = 0. \]  

(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

1 Background: JKO scheme for Wasserstein gradient flow

2 Variational time implicit scheme for dissipative systems

3 Optimization solver: the ALG2 Algorithm

4 High-order spatial discretization of variational time implicit schemes

5 Numerics

6 High-order computation of OT/MFP/MFG

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

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

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 valid:

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

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

\[
\text{Dist}_{V_1, V_2}(\rho^0, \rho^1)^2 := \inf_{\nu_1, \nu_2, \rho} \int_0^1 \int_\Omega \left[ \|v_1(x, \tau)\|^2 V_1(\rho(x, \tau)) + |v_2(x, \tau)|^2 V_2(\rho(x, \tau)) \right] dxd\tau,
\]

(9a)

where the infimum is taken among \( \rho : \Omega \times [0, 1] \to \mathbb{R} \), \( \nu_1, \nu_2 : \Omega \times [0, 1] \to \mathbb{R}^d \), such that \( \rho \) satisfies a reaction-diffusion type equation with drift vector field \( \nu_1 \), drift mobility \( V_1 \), reaction ratio \( \nu_2 \), reaction mobility \( V_2 \), connecting initial and terminal densities \( \rho^0, \rho^1 \):

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

(9b)
One-step relaxation of variational time implicit schemes

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

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

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

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

**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

\[ \mathcal{E}(\rho^n) \leq \mathcal{E}(\rho^{n-1}), \quad \text{for } n \in \mathbb{N}_+. \]
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]:

\[ \frac{\partial \rho}{\partial t} - \nabla \cdot (\alpha \nabla \rho) = \rho(1 - \rho). \]
Strongly reversible reaction-diffusion system

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

$$\alpha_p^1 X_1 + \cdots + \alpha_p^M X_M \xrightleftharpoons[k^p_+]{k^p_-} \beta_p^1 X_1 + \cdots + \beta_p^M X_M,$$

(12)

where $p = 1, \ldots, R$ is the number of possible reactions, $\alpha^p = (\alpha_1^p, \ldots, \alpha_M^p), \beta^p = (\beta_1^p, \ldots, \beta_M^p) \in \mathcal{N}_0^M$ 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 $\rho_i$ of species $X_i$, we get the following reaction-diffusion system:

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

(13)

for $1 \leq i \leq M$, where $\rho = (\rho_1, \ldots, \rho_M)$ and the multi-index notation $\rho^{\alpha^p} := \prod_{i=1}^M \rho_i^{\alpha_i^p}$ 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]:

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

(14)

and denote the following mobility functions:

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

\[
V_{2,p}(\rho) = k^p \ell \left( \rho^{\alpha_p}, \rho^{\beta_p} \right), \quad \forall 1 \leq p \leq R. \tag{15b}
\]

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

\[
\partial_t \rho_i = \nabla \cdot \left( V_{1,i}(\rho_i) \nabla \frac{\delta}{\delta \rho} \mathcal{E}_i(\rho_i) \right) 
- \sum_{p=1}^{R} V_{2,p}(\rho) (\alpha_i^p - \beta_i^p) \sum_{j=1}^{M} (\alpha_j^p - \beta_j^p) \frac{\delta}{\delta \rho} \mathcal{E}_i(\rho_i).
\]

(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

\[
\frac{d}{dt} \sum_{i=1}^{M} E_i(\rho_i(\cdot, t)) = - \sum_{i=1}^{M} \int_{\Omega} \left| \nabla \frac{\delta}{\delta \rho} E_i(\rho_i)(x, t) \right|^2 V_{1,i}(\rho_i) \, dx 
- \sum_{p=1}^{R} \int_{\Omega} \left| \sum_{j=1}^{M} (\alpha_j^p - \beta_j^p) \frac{\delta}{\delta \rho} E_i(\rho_i) \right|^2 V_{2,p}(\rho) \, dx.
\]

(17)
One-step relaxation of variational time implicit schemes

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

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

(18a)

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

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

\[ \inf_u \sup_\Phi F(u) - G(\Phi) - (u, D\Phi)_\Omega, \]  

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

\[ u = (\rho, m, s), \]

with

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

and

\[ (u, D\Phi)_\Omega = \int_\Omega \left[ -\rho \Phi + (m, \nabla_x \Phi) + s \Phi \right] dx. \]
The augmented Lagrangian

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

$$\sup_u \inf_{\Phi,u^*} F^*(u^*) + G(\Phi) + (u, D\Phi - u^*) \Omega,$$

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

$$\sup_u \inf_{\Phi,u^*} L_r(\Phi,u,u^*),$$

where the augmented Lagrangian

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

in which $r$ is a positive parameter.
Algorithm 1  One iteration of ALG2 algorithm for variational implicit scheme (21).

- Step A: update $\Phi$. Minimize $L_r(\Phi, u, u^*)$ with respect to the first argument by solving the elliptic problem: Find $\Phi^\ell$ such that it solves
\[
\inf_{\Phi} L_r(\Phi, u^{\ell-1}, u^*, u^{*, \ell-1}).
\]
This step is a linear, constant-coefficient reaction-diffusion problem.

- Step B: update $u^*$. Minimize $L_r(\Phi, u, u^*)$ with respect to the last argument by solving the nonlinear problem: Find $u^{*, \ell}$ such that it solves
\[
\inf_{u^*} L_r(\Phi^\ell, u^{\ell-1}, u^*).
\]

- Step C: update $u$. This is a simple pointwise update for the Lagrange multiplier $u$:
\[
u^{\ell} = u^{\ell-1} + r(D\Phi^\ell - u^*, u^{*, \ell}). \tag{22}\]
Table of Contents

1. Background: JKO scheme for Wasserstein gradient flow
2. Variational time implicit scheme for dissipative systems
3. Optimization solver: the ALG2 Algorithm
4. High-order spatial discretization of variational time implicit schemes
5. Numerics
6. High-order computation of OT/MFP/MFG
7. Conclusion
High-order finite element spaces/nodal basis

Approximate $\Phi$ using high-order $H^1$-conforming finite element space

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

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

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

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 $\varphi_i(\xi_j) = \delta_{ij}, \ \forall 1 \leq j \leq N_W$, where $N_W$ is the dimension of the space $W_h^k$, $\delta_{ij}$ 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 $T_h$.

We denote the discrete $L^2(\Omega)$-inner product $(\cdot, \cdot)_h$ as

$$\langle u, v \rangle_h := \sum_{i=1}^{N_W} u(\xi_i)v(\xi_i)\omega_i, \quad (25)$$
Fully discrete scheme for reaction diffusion equation

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

$$\inf_{u \in [W_h^k]^4} \sup_{\Phi_h \in V_h^k, u_h^* \in [W_h^k]^4} L_{r,h}(\Phi_h, u_h, u_h^*),$$

(26a)

where $u_h = (\rho_h, m_0^h, m_1^h, s_h)$ is the collection of density/flux/source, $u_h^* = (\rho_h^*, m_0^h^*, m_1^h^*, s_h^*)$ is its dual, and

$$L_{r,h}(\Phi_h, u_h, u_h^*) := F_h^*(u_h^*) + G_h(\Phi_h) + (u_h, D\Phi_h - u_h^*)_h$$

$$+ \frac{r}{2}(D\Phi_h - u_h^*, D\Phi_h - u_h^*)_h,$$

(26b)

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

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

(26c)

$$F_h^*(u_h^*) := \sup_{u_h \in [W_h^k]^4} (u_h^*, u_h)_h - F_h(u_h),$$

(26d)

$$F_h(u_h) := \left(\frac{|m_0^h|^2 + |m_1^h|^2}{2 V_1(\rho_h)} + \frac{s_h^2}{2 V_2(\rho_h)}, 1\right)_h + \Delta t E_h(\rho_h).$$

(26e)
Practical ALG2 implementation

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

- Step A: update $\Phi^\ell_h$. Find $\Phi^\ell_h \in V^k_h$ such that, for all $\forall \psi_h \in V^k_h$,

$$
(\mathcal{D}\Phi^\ell_h, \mathcal{D}\psi_h)_h = (u^{*,\ell-1}_h - \frac{1}{r} u^{\ell-1}_h, \mathcal{D}\psi_h)_h - \frac{1}{r} (\rho^{old}_h, \psi_h)_h.
$$

(27)

- Step B: update $u^\ell_h$. Find $\rho^\ell_h$ such that it is the minimizer to the following functional of $\rho_h$

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

(28)

Then update $m^{0,\ell}_h, m^{1,\ell}_h, s^\ell_h$ according to

$$
m^{0,\ell}_h = \frac{rV_1(\rho^\ell_h m^0_h)}{r + V_1(\rho^\ell_h)}, \quad m^{1,\ell}_h = \frac{rV_1(\rho^\ell_h m^1_h)}{r + V_1(\rho^\ell_h)}, \quad s^\ell_h = \frac{rV_2(\rho^\ell_h s_h)}{r + V_2(\rho^\ell_h)},
$$

(29)

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

1. Background: JKO scheme for Wasserstein gradient flow
2. Variational time implicit scheme for dissipative systems
3. Optimization solver: the ALG2 Algorithm
4. High-order spatial discretization of variational time implicit schemes
5. Numerics
6. High-order computation of OT/MFP/MFG
7. Conclusion
We first consider the nonlinear Fokker-Plank equation
\[ \partial_t \rho - \Delta \rho^3 = \nabla \cdot (\rho x), \]
on the domain \( \Omega = [-1, 1] \times [-1, 1] \) with homogeneous Neumann boundary conditions. It is a Wasserestein gradient flow with energy
\[ E(\rho) := \int_\Omega \left( \frac{1}{2} \rho(x)^3 + \frac{1}{2} (x_0^2 + x_1^2) \rho(x) \right) dx, \]
where \( x = (x_0, x_1) \). This problem reaches a steady state solution
\[ \rho_{\text{steady}}(x_1, x_2) = \sqrt{(2C - (x_0^2 + x_1^2))_+}, \]
that satisfies either
\[ \frac{\delta E}{\delta \rho} = \frac{3}{2} \rho^2 + \frac{1}{2} (x_0^2 + x_1^2) = C, \]
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 \( \Omega \) is positive and smooth.
Ex1: Spatial convergence rates, Fokker-Plank equation

The $L^2$-convergence in the density $\rho$ is recorded in Table 1.

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

<table>
<thead>
<tr>
<th>dim($V_h^k$)</th>
<th>$k = 1$</th>
<th></th>
<th>$k = 2$</th>
<th></th>
<th>$k = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>2.362e-03 -</td>
<td></td>
<td>2.409e-04 -</td>
<td></td>
<td>2.628e-05 -</td>
</tr>
<tr>
<td>289</td>
<td>5.923e-04 2.00</td>
<td></td>
<td>3.298e-05 2.87</td>
<td></td>
<td>1.424e-06 4.21</td>
</tr>
<tr>
<td>1089</td>
<td>1.482e-04 2.00</td>
<td></td>
<td>4.232e-06 2.96</td>
<td></td>
<td>5.589e-08 4.67</td>
</tr>
<tr>
<td>4225</td>
<td>3.705e-05 2.00</td>
<td></td>
<td>5.326e-07 2.99</td>
<td></td>
<td>1.884e-09 4.89</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Case</th>
<th>$\alpha U_m(\rho)$</th>
<th>$V(x)$</th>
<th>$W(x)$</th>
<th>$L$</th>
<th>I.C.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\frac{</td>
<td>x</td>
<td>^4}{4} - \frac{</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>$\frac{</td>
<td>x</td>
<td>^2}{2} - \log(</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>$-\frac{1}{4} \log(</td>
<td>x</td>
<td>)$</td>
<td>$\frac{</td>
</tr>
<tr>
<td>4</td>
<td>$0.1 \rho^2$</td>
<td>$-\frac{1}{4} \log(</td>
<td>x</td>
<td>)$</td>
<td>$\frac{</td>
</tr>
<tr>
<td>5</td>
<td>$0.1 \rho^3$</td>
<td>0</td>
<td>$-\exp(-</td>
<td>x</td>
<td>^2)/\pi$</td>
</tr>
</tbody>
</table>
Ex2: Aggregation-drift-diffusion equations

For all cases, we take a computational domain with a $32 \times 32$ 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,

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

(30)

The same discretization setup as in the previous example is used, i.e., using polynomial degree $k = 4$ on a $32 \times 32$ 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 \times 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]:

\[
\begin{align*}
\partial_t \rho_1 - \frac{\gamma_1}{m} \Delta \rho_1^m &= -(k_+ \rho_1 \rho_2^2 - k_- \rho_2^3), \\
\partial_t \rho_2 - \gamma_2 \Delta \rho_2 &= (k_+ \rho_1 \rho_2^2 - k_- \rho_2^3).
\end{align*}
\]

It can be equivalently written as

\[
\begin{align*}
\partial_t \rho_1 &= \nabla \cdot \left( V_{1,1}(\rho_1) \nabla \frac{\delta E_1}{\delta \rho} (\rho_1) \right) - V_2(\rho_1, \rho_2) \left( \frac{\delta E_1}{\delta \rho}(\rho_1) - \frac{\delta E_2}{\delta \rho}(\rho_2) \right), \\
\partial_t \rho_2 &= \nabla \cdot \left( V_{1,2}(\rho_2) \nabla \frac{\delta E_2}{\delta \rho} (\rho_2) \right) + V_2(\rho_1, \rho_2) \left( \frac{\delta E_1}{\delta \rho}(\rho_1) - \frac{\delta E_2}{\delta \rho}(\rho_2) \right),
\end{align*}
\]

where

\[
\begin{align*}
V_{1,1}(\rho_1) &= \gamma_1(\rho_1)^m, \\
V_{1,2}(\rho_2) &= \gamma_2 \rho_2, \\
V_2(\rho_1, \rho_2) &= \ell(\kappa_1 \rho_1 \rho_2^2, \kappa_2 \rho_2^3),
\end{align*}
\]

with

\[
\begin{align*}
E_i(\rho_i) &= \rho_i (\log(\kappa_i \rho_i) - 1), \\
\ell(x, y) &= \frac{x - y}{\log(x) - \log(y)}.
\end{align*}
\]

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: $t = 0, 0.5, 1, 1.5, 2$
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]:

\[
\begin{align*}
\frac{\partial}{\partial t} \rho_1 &= \gamma_1 \Delta \rho_1 - (k_+^1 \rho_1 \rho_2^2 - k_+^3 \rho_2^3) - (k_+^3 \rho_1 - k_+^3 \rho_4), \\
\frac{\partial}{\partial t} \rho_2 &= \gamma_2 \Delta \rho_2 + (k_+^1 \rho_1 \rho_2^2 - k_+^3 \rho_2^3) - (k_+^2 \rho_2 - k_+^2 \rho_3), \\
\frac{\partial}{\partial t} \rho_3 &= (k_+^2 \rho_2 - k_-^2 \rho_3), \\
\frac{\partial}{\partial t} \rho_4 &= (k_+^3 \rho_1 - k_-^3 \rho_4) \text{.}
\end{align*}
\]

The physical parameters are chosen to be the following:

\[
\begin{align*}
\gamma_1 &= 1, \gamma_2 = 0.01, k_+^1 = 1, k_+^2 = 0.084, k_+^3 = 0.024, \quad k_-^i = 10^{-3} k_+^i \text{.}
\end{align*}
\]

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

\[
\begin{align*}
\frac{\partial}{\partial t} \rho_1 &= \gamma_1 \Delta \rho_1 - k_+^1 \rho_1 \rho_2^2 - k_+^3 (\rho_1 - 1), \\
\frac{\partial}{\partial t} \rho_2 &= \gamma_2 \Delta \rho_2 + k_+^1 \rho_1 \rho_2^2 - k_+^2 \rho_2 \text{,}
\end{align*}
\]

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 $\rho_2$ 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

1 Background: JKO scheme for Wasserstein gradient flow
2 Variational time implicit scheme for dissipative systems
3 Optimization solver: the ALG2 Algorithm
4 High-order spatial discretization of variational time implicit schemes
5 Numerics
6 High-order computation of OT/MFP/MFG
7 Conclusion
Dynamic Mean Field Planning

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

$$\inf_{(\rho, m)} \int_0^1 \int_{\Omega} L(\rho(t, x), m(t, x)) + A(\rho(t, x)) \, dx \, dt,$$

(31)

among all feasible density and flux \((\rho, m)\) 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 \(\rho_0\) and \(\rho_1\).

- 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 $\rho_1$ is not explicitly provided but it satisfies a given preference. The goal of MFG is to minimize the total cost

$$\inf_{(\rho,m)} \int_0^1 \int_{\Omega} \left[ L(\rho, m) + A(\rho) \right] \, dx dt + \int_{\Omega} \Gamma(\rho(1, x)) \, dx, \quad (32)$$

among all feasible $(\rho, m)$ 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:

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

(33a)

in which \( u = (\rho, m) \), and

\[
F(u) := \int_0^1 \int_\Omega \left[ L(\rho, m) + A(\rho) \right] \, dx\, dt,
\]

(33b)

\[
G(\phi) := \int_\Omega \left[ -\phi(1, x)\rho_1(x) + \phi(0, x)\rho_0(x) \right] \, dx,
\]

(33c)

\[
\mathcal{D}\Phi := (\partial_t \Phi, \nabla_x \Phi) \text{ is the space-time gradient}
\]

(33d)

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

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

with $\sigma = 0.1$. Note that we allow $\rho_T$ and $\rho_0$ to have different total masses here.

\[
\begin{align*}
\text{Case 1: } & A(\rho) = 0, & A^*(\rho^*) = & \begin{cases} 
0 & \text{if } \rho^* \leq 0, \\
+\infty & \text{if } \rho^* > 0,
\end{cases} \\
\text{Case 2: } & A(\rho) = c\rho^2, & A^*(\rho^*) = & \begin{cases} 
0 & \text{if } \rho^* \leq 0, \\
(\rho^*)^2/(4c) & \text{if } \rho^* > 0.
\end{cases} \\
\text{Case 3: } & A(\rho) = c\rho \log(\rho), & A^*(\rho^*) = & \exp(\rho^*/c - 1), \\
\text{Case 4: } & A(\rho) = c/\rho, & A^*(\rho^*) = & \begin{cases} 
-2\sqrt{-c\rho^*} & \text{if } \rho^* \leq 0, \\
+\infty & \text{if } \rho^* > 0.
\end{cases} \\
\text{Case 5: } & A(\rho) = & & \begin{cases} 
0 & \text{if } 0 \leq \rho \leq \rho_{\text{max}}, \\
+\infty & \text{else},
\end{cases} & A^*(\rho^*) = & \rho_{\text{max}}(\rho^*) +
\end{align*}
\]
Figure 4: Example 4.3. Snapshots of $\rho$ 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 4.1. Initial/final densities.
A MFP Example: ND $\rightarrow$ UCLA

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

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

(c) Case 3: $A(\rho) = 0.01/\rho$. ND $\rightarrow$ UCLA
A MFP Example: UCLA $\rightarrow$ USC

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

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

(c) Case 3: $A(\rho) = 0.01/\rho$. UCLA $\rightarrow$ USC
A MFP Example: USC → ND

(a) Case 1: $A(\rho) = 0$. USC → ND

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

(c) Case 3: $A(\rho) = 0.01 / \rho$. USC → ND
Table of Contents

1. Background: JKO scheme for Wasserstein gradient flow
2. Variational time implicit scheme for dissipative systems
3. Optimization solver: the ALG2 Algorithm
4. High-order spatial discretization of variational time implicit schemes
5. Numerics
6. High-order computation of OT/MFP/MFG
7. 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


References II


References III


References IV


References VI


References VII
