BOUND/POSITIVITY PRESERVING AND ENERGY STABLE SAV SCHEMES FOR DISSIPATIVE SYSTEMS: APPLICATIONS TO KELLER-SEGEL AND POISSON-NERNST-PLANCK EQUATIONS

FUKENG HUANG AND JIE SHEN

ABSTRACT. We propose a new method to construct high-order, linear, positivity/bound preserving and unconditionally energy stable schemes for general dissipative systems whose solutions are positivity/bound preserving. The method is based on applying a new SAV approach to the transformed system with a suitable function transform. The resulting schemes enjoy remarkable properties such as positivity/bound preserving, unconditionally energy stable, can achieve high-order and with computational complexity similar to a semi-implicit scheme. We apply this approach to Keller-Segel and Poisson-Nernst-Planck (PNP) equations, and construct efficient numerical schemes which, in addition to positivity/bound preserving and energy dissipative, also conserve mass. Ample numerical results are presented to validate our theoretical claims.

1. INTRODUCTION

Many problems in sciences and engineering require their solutions to be positive or remain in a prescribed range, such as density, concentration, height, population, etc. Oftentimes, violation of the positivity or bound preserving in their numerical solutions renders the corresponding discrete problems ill posed, although the original problems are well posed. For these type of problems, it is of critical importance for the numerical schemes to be positivity or bound preserving. A particular class of such problems are the Wasserstein gradient flows which are gradient flows over spaces of probability distributions according to the topology defined by the Wasserstein metric [21, 28]. Important examples of Wasserstein gradient flows include the Poisson-Nernst-Planck (PNP) equations [26] and Keller-Segel equations [22, 17]. For these problems, in addition to positivity or bound preserving, it is also important for the numerical schemes to obey a discrete energy law. Many attempts have been made over the years in developing numerical schemes for the PNP and Keller-Segel equations.

For the PNP equations, a quite complicated entropy-based scheme with regularized free energy is constructed in [27] along with rigorous numerical analyses for a set of finite-element approximations; a mass-conservative finite difference scheme is constructed in [13]; an arbitrary-order energy dissipative schemes are constructed using a a discontinuous Galerkin (DG) method for 1-D PNP systems [23]; and most recently a fully discrete positivity-preserving and energy-dissipative finite difference scheme is developed in [19]. On the other hand, There exist a large number of numerical work for the PNP equations in the electric and medical engineering literature, see, for examples, [15, 25, 18] and the references therein.

For the Keller-Segel equations and related models, a finite volume scheme is developed with convergence proof in [12]; a second-order positivity preserving central-upwind scheme is constructed in [6] (see also [11, 10]); finite volume methods for a Keller-Segel system are considered with discrete energy dissipation and error estimates in [36]; and a positivity-preserving and asymptotic preserving method is constructed for a reformulated Keller-Segel system in [24] [36, 7]. We refer to the

²⁰⁰⁰ Mathematics Subject Classification. 65N35; 65N22; 65F05; 35J05.

Key words and phrases. positivity or bound preserving; SAV approach; energy stability; Keller-Segel; Poisson-Nernst-Planck; dissipative systems.

Department of Mathematics, Purdue University. This research is partially supported by NSF grants DMS-2012585 and AFOSR FA9550-20-1-0309. Emails: huang972@purdue.edu (F. Huang), shen7@purdue.edu (J. Shen).

F. HUANG AND J. SHEN

aforementioned papers and the references therein for more details on existing numerical schemes for Keller-Segel equations.

Some of these numerical schemes preserve positivity and/or some form of energy dissipation under certain conditions and specific spatial discretization. Oftentimes one needs to solve nonlinear systems at each time step. Very recently, an interesting approach is proposed to construct unconditionally energy stable and positivity/bound preserving for Keller-Segel equations in [30] and for PNP equations in [31]. However, these schemes require solving, at each time step, a nonlinear system which is a unique minimizer of a strictly convex functional. The question we would like to address in this paper is: for PDEs which preserve positivity or bound and satisfy an energy dissipation law, how to construct numerical schemes which are linear, positivity/bound preserving and unconditionally energy stable for any consistent spatial discretization?

The recently proposed scalar auxiliary variable (SAV) approach [32, 33] is a powerful tool to design unconditionally energy stable, linear schemes to a large class of gradient flows, and has been applied successfully to many challenging problems. However, it does not have mechanism to preserve bounds or positivity. On the other hand, a common strategy to enforce solutions to preserve bounds or positivity is to use a suitable function transform. A drawback of this approach is that the transformed equation becomes very complicated that it is very difficult to construct efficient and energy stable schemes for the transformed equation.

In this work, we propose a new class of bound/positivity preserving and energy stable schemes by combining the SAV approach and the function transform approach:

- make a suitable function transform to ensure positivity or bound preserving;
- use a recently proposed SAV approach [20] to design linear and unconditionally energy stable schemes for the transformed equation.

Our new schemes will enjoy the following remarkable properties:

- it can be used with high-order semi-implicit (i.e., IMEX) schemes;
- it is positivity or bound preserving;
- it is unconditionally energy dissipative;
- it only requires solving one set (instead of two in the original SAV approach) decoupled linear equations with constant coefficients at each time step, so the coding and computational complexity is similar to that of semi-implicit schemes;
- for problems with mass conservation as in PNP and KS equations, it also conserves mass.

The rest of the paper is organized as follows. In Section 2, we describe our approach for a general semi-linear or quasi-linear, dissipative system. In Section 3, we construct new schemes for the PNP equations, followed by the schemes for Keller-Segel equations in Section 4. In section 5, we present numerical examples to validate our schemes. Some concluding remarks are given in Section 6.

2. Bound/positivity preserving schemes through transform and SAV approaches

In order to clearly describe our idea, we consider a semi-linear or quasilinear parabolic system in the form

$$\frac{\partial u}{\partial t} - \Delta u + g(u) = 0, \qquad (2.1)$$

with either periodic or homogeneous Neumann boundary condition, where g(u) is a nonlinear function. The following discussions are still valid if we replace $-\Delta$ in (2.1) with more general or higher-order linear elliptic operators.

We assume that the above system satisfies a dissipation law in the form

$$\frac{dE(u)}{dt} = -(\mathcal{G}u, u), \qquad (2.2)$$

 $\mathbf{2}$

where E(u) is a typical energy functional given by

$$E[u] = \int_{\Omega} \left(\frac{1}{2}\mathcal{L}u \cdot u + F(u)\right) dx := E_0(u) + E_1(u),$$
(2.3)

 \mathcal{G} is a non-negative operator and \mathcal{L} is a self-adjoint, linear, non-negative operator. We also assume that E[u] is bounded from below, without loss of generality, we can assume $E[u] > 0, \forall u$.

Note that the above framework includes, as special cases, the L^2 gradient flows for which g(u) = F'(u) where F(u) is a given nonlinear function, $L = -\Delta$ and $(\mathcal{G}u, u) = (-\Delta u + g(u), -\Delta u + g(u))$.

Solutions of (2.1) is often bound/positivity preserving. It is desirable, and sometimes necessary such as in the case of PNP and Keller-Segel equations, for the numerical solutions to be also bound/positivity preserving. While it is possible to construct some fully discrete numerical methods which preserve the bounds/positivity using finite-differences or piecewise linear finite-elements for a class of (2.1) satisfying a maximum principle, it is in general very difficult to construct higher-order finite-elements or spectral methods which preserve bounds/positivity as well as energy dissipation.

While the SAV approach [33] provided a powerful approach to design numerical schemes which preserve energy dissipation, it does not have mechanism to preserve bounds or positivity. A common strategy to enforce solutions to preserve bounds or positivity is to use a suitable function transform. More precisely, given a prescribed range interval I which could be open, closed or half open, we can construct an invertible mapping $T : \mathbb{R} \to I$, and make the function transform u = T(v) in (2.1), leading to

$$\frac{\partial v}{\partial t} - \Delta v - \frac{T''(v)}{T'(v)} |\nabla v|^2 + \frac{1}{T'(v)} g(T(v)) = 0, \qquad (2.4)$$

with either periodic or homogeneous Neumann boundary condition, since $\frac{\partial u}{\partial n} = T'(v)\frac{\partial v}{\partial n}$. After we solve v from the above, we get u = T(v) whose range is included in I. Two typical cases are:

- I = (a, b): a suitable choice is $T(v) = \frac{b-a}{2} \tanh(v) + \frac{b+a}{2}$ so that the range of u = T(v) is still in I.
- $I = (0, \infty)$: a suitable choice is $T(v) = \exp(v/M)$, where M is a tunable parameter to prevent T(v) increases too fast, so that u = T(v) is always positive.

The main difficulty with this transformed approach is that the transformed equation (2.4) is much more complicated than (2.1), and it is difficult to design efficient and energy dissipative schemes. Fortunately, the recently proposed SAV approach [20] can provide a satisfactory solution as we show below.

As in the usual SAV approach, we introduce a SAV to enforce energy dissipation (2.2). More precisely, we set $r(t) = \int_{\Omega} F(u) dx + C_0$ with $C_0 \ge E[u^0]$ and expand (2.4) with (2.2) as

$$\frac{\partial v}{\partial t} - \Delta v - \frac{T''(v)}{T'(v)} |\nabla v|^2 + \frac{1}{T'(v)} g(T(v)) = 0, \qquad (2.5a)$$

$$u = T(v), \tag{2.5b}$$

$$\frac{dE_0(u)}{dt} + \frac{dr}{dt} = -\frac{E_0(u) + r(t)}{E(u) + C_0} (\mathcal{G}u, u), \qquad (2.5c)$$

with $r(0) = \int_{\Omega} F(u(x, 0))dx + C_0$, it is clear that the above system is equivalent to (2.4) with (2.2). However, discretizing the above will allow us to easily construct schemes which is energy dissipative, in addition to bound/positivity preserving which is built into the system. We construct below k-th order BDF-Adams-Bashforth SAV schemes for (4.4b) in a uniform setting: treat the linear term Δv implicitly and use Adams-Bashforth extrapolation to deal with all nonlinear terms.

F. HUANG AND J. SHEN

More precisely, given r^n and (u^j, v^j) for $j = n, \dots, n-k+1$, we find $(v^{n+1}, u^{n+1}, r^{n+1}, \xi^{n+1})$ such that

$$\frac{\alpha_k v^{n+1} - A_k(v^n)}{\delta t} - \Delta v^{n+1} = \frac{T''(B_k(v^n))}{T'(B_k(v^n))} |\nabla B_k(v^n)|^2 - \frac{1}{T'(B_k(v^n))} g(B_k(u^n)), \quad (2.6)$$

$$\bar{u}^{n+1} = T(v^{n+1}), \tag{2.7}$$

$$\frac{1}{\delta t} \left(\frac{1}{2} \int_{\Omega} (\mathcal{L}\bar{u}^{n+1} \cdot \bar{u}^{n+1} - \mathcal{L}\bar{u}^n \cdot \bar{u}^n) dx + r^{n+1} - r^n \right) \\
= -\frac{\frac{1}{2} \int_{\Omega} \mathcal{L}\bar{u}^{n+1} \cdot \bar{u}^{n+1} dx + r^{n+1}}{\overline{\mathcal{L}}_{n+1} - \overline{\mathcal{L}}_{n+1}} (\mathcal{G}\bar{u}^{n+1}, \bar{u}^{n+1}),$$
(2.8)

$$\mathcal{E}[u^{n+1}] + C_0 \qquad (2.9)$$

$$\xi^{n+1} = \frac{\int_{\Omega} \frac{1}{2} \mathcal{L} \bar{u}^{n+1} \cdot \bar{u}^{n+1} dx + r^{n+1}}{r^{1-n+1} dx + r^{n+1}},$$

$$E[\bar{u}^{n+1}] + C_0 \qquad (1 c)$$

$$u^{n+1} = \eta_k^{n+1} \bar{u}^{n+1} \text{ with } \eta_k^{n+1} = 1 - (1 - \xi^{n+1})^{I_k}, \ I_k = \begin{cases} k+1, & \text{if } k \text{ is odd} \\ k, & \text{if } k \text{ is even} \end{cases},$$
(2.10)

where the constant α_k , operators A_k , B_k are defined by

first-order scheme:

$$\alpha_1 = 1, \ A_1(v^n) = v^n, \ B_1(h^n) = h^n;$$
(2.11)

2nd-order scheme:

$$\alpha_2 = \frac{3}{2}, \quad A_2(v^n) = 2v^n - \frac{1}{2}v^{n-1}, \quad B_2(h^n) = 2h^n - h^{n-1};$$
(2.12)

third-order scheme:

$$\alpha_3 = \frac{11}{6}, \quad A_3(v^n) = 3v^n - \frac{3}{2}v^{n-1} + \frac{1}{3}v^{n-2}, \quad B_3(h^n) = 3h^n - 3h^{n-1} + h^{n-2}; \tag{2.13}$$

fourth-order scheme:

$$\alpha_4 = \frac{25}{12}, \quad A_4(v^n) = 4v^n - 3v^{n-1} + \frac{4}{3}v^{n-2} - \frac{1}{4}v^{n-3}, \quad B_4(h^n) = 4h^n - 6h^{n-1} + 4h^{n-2} - h^{n-3}.$$
(2.14)

The formulae for k = 5 and k = 6 can be derived similarly.

Several remarks are in oder:

- Since we assume T is invertible, $T'(v) \neq 0$ so the above scheme is well defined. The range of the approximate solution $\bar{u}^{n+1} = T(v^{n+1})$ is obviously included in I.
- (2.6) is a k-th order approximation to (2.5a) with k-th order BDF for the linear terms and k-th order Adams-Bashforth extrapolation for the nonlinear terms. Hence, v^{n+1} is a k-th order approximation to $v(t_{n+1})$.
- (2.8) is a first-order approximation to (2.5c). Hence, r^{n+1} is a first order approximation to $E_1(u(\cdot, t_{n+1}))$ which implies that ξ^{n+1} is a first order approximation to 1. Hence, $\eta_k^{n+1} = 1 + O(\delta t)^{I_k}$ which implies that both \bar{u}^{n+1} and u^{n+1} are k-th order approximation of $u(t_{n+1})$.
- The above scheme can be efficiently implemented as follows:
 - determine v^{n+1} from (2.6);
 - set $\bar{u}^{n+1} = T(v^{n+1});$
 - with \bar{u}^{n+1} known, determine r^{n+1} explicitly from (2.8), and compute ξ^{n+1} from (2.9); - update u^{n+1} using (2.10), go to the next step.

The main cost is to solve v^{n+1} from (2.6) which is a linear equation with constant coefficients.

The above scheme looks similar to the scheme in [20], but there are some subtle differences, particularly in the choice of η_k^{n+1} . As we show below, this choice allows us to obtain a uniform bound on $(\mathcal{L}u^n, u^n)$, which in turn will play a crucial role in the error analysis as in [29].

Theorem 1. Without loss of generality, we assume $ab \leq 0$ if I = (a, b). Given u^i with range in I, $v^i = T^{-1}(u^i)$ and r^i for i = 0, 1, ..., k - 1. The scheme (2.6)-(2.10) admits a unique solution satisfying the following properties unconditionally:

4

- 1. Positivity or bound preserving: i.e., the range of \bar{u}^{n+1} and u^{n+1} is in I.
- 2. Unconditionally energy dissipation with a modified energy defined by $\bar{E}^n = \int_{\Omega} \frac{1}{2} \mathcal{L} \bar{u}^n \cdot \bar{u}^n dx + r^n$: More precisely, if $\bar{E}^n \ge 0$, we have $\bar{E}^{n+1} \ge 0$ and

$$\bar{E}^{n+1} - \bar{E}^n \le -\delta t \frac{E^{n+1}}{E[\bar{u}^{n+1}] + C_0} \left(\mathcal{G}\bar{u}^{n+1}, \bar{u}^{n+1} \right) \le 0.$$
(2.15)

3. Furthermore, if $E_1(u) = \int_{\Omega} F(u) dx$ is bounded from below, then for the k-th order schemes, there exists constant M_k , such that

$$(\mathcal{L}u^n, u^n)^{1/2} \le M_k, \,\forall n.$$

$$(2.16)$$

Proof. By construction, the scheme is obviously positivity or bound preserving for \bar{u}^{n+1} . We derive from (2.8) that

$$\bar{E}^{n+1} = \bar{E}^n / \left(1 + \frac{\delta t}{E[\bar{u}^{n+1}] + C_0} (\mathcal{G}\bar{u}^{n+1}, \bar{u}^{n+1}) \right).$$

Hence, if $\bar{E}^n \ge 0$, we have $\bar{E}^{n+1} \ge 0$, and (2.15) follows directly from (2.8). It follows from (2.9), (2.15) and $C_0 \ge E[u^0]$, $E[\bar{u}^{n+1}] > 0$ that

$$0 < \xi^{n+1} \le \frac{E[u^0] + C_0}{E[\bar{u}^{n+1}] + C_0} < 2, \tag{2.17}$$

which together with (2.10) imply

$$0 < (1 - \xi^{n+1})^{I_k} < 1, \ 0 < \eta_k^{n+1} < 1.$$
(2.18)

Hence, the range of u^{n+1} is also in I as $u^{n+1} = \eta_k^{n+1} \bar{u}^{n+1}$ for $I = (0, \infty)$ or I = (a, b) with $ab \le 0$.

If $E_1(u) = \int_{\Omega} F(u) dx$ is bounded from below, without loss of generality, we assume $E_1(u) > 1$. Denote $M := \bar{E}[u(\cdot, 0)]$, then (2.15) implies $\bar{E}^n \leq M, \forall n$. Now, it follows from (2.9) and the assumption of $E_1(u) > 1$ that

$$|\xi^{n+1}| = \frac{\bar{E}^{n+1}}{E[\bar{u}^{n+1}] + C_0} \le \frac{2M}{(\mathcal{L}\bar{u}^{n+1}, \bar{u}^{n+1}) + 2}.$$
(2.19)

Since $\eta_k^{n+1} = 1 - (1 - \xi^{n+1})^{I_k}$, there exists a polynomial P_k of degree $I_k - 1$ and a constant $M_k > 0$ such that

$$|\eta_k^{n+1}| = |\xi^{n+1} P_k(\xi^{n+1})| \le \frac{M_k}{(\mathcal{L}\bar{u}^{n+1}, \bar{u}^{n+1}) + 2}.$$
(2.20)

Therefore, by the fact $\sqrt{A} \leq A + 2$ for all $A \geq 0$, we have

$$(\mathcal{L}u^{n+1}, u^{n+1})^{1/2} = \eta_k^{n+1} (\mathcal{L}\bar{u}^{n+1}, \bar{u}^{n+1})^{1/2} \le M_k.$$
(2.21)

The above scheme can be directly applied to bound/positivity preserving L^2 gradient flows, including in particular the Allen-Cahn equation. In the following two sections, we shall extend the approach presented in this section to construct positivity preserving and energy stable schemes for Poisson-Nernst-Planck and Keller-Segel equations for which it is essential to preserve positivity.

Remark 1. We emphasize that both \bar{u}^{n+1} and u^{n+1} are k-th order approximation to $u(\cdot, t_{n+1})$.

We only considered the time discretization in this section. However, it is clear from the proof of the above theorem that, as long as the spatial approximations of \mathcal{G} and \mathcal{L} are still positive definite, the results of Theorem 1 also holds for the fully discrete schemes.

F. HUANG AND J. SHEN

3. Positivity preserving schemes for the Poisson-Nernst-Planck equation

We consider in this section the Poisson-Nernst-Planck (PNP) equation which describes the dynamics of N species of charged particles driven by Brownian motion and electric field (cf. [2, 14, 9] and the references therein). To simplify the presentation, we will focus on the two-component system (N = 2). The schemes can be easily extended to more general PNP system with N components.

3.1. **Poisson-Nernst-Planck equation.** We consider a two-component PNP system in the following form:

$$\frac{\partial c_1}{\partial t} = D_1 \nabla \cdot (\nabla c_1 + \chi_1 z_1 c_1 \nabla \phi), \qquad (3.1a)$$

$$\frac{\partial c_2}{\partial t} = D_2 \nabla \cdot (\nabla c_2 + \chi_1 z_2 c_2 \nabla \phi), \qquad (3.1b)$$

$$-\Delta\phi = \chi_2(z_1c_1 + z_2c_2), \tag{3.1c}$$

in an open bounded domain $\Omega \subset \mathbb{R}^d$ (d = 1, 2, 3) and supplemented with either periodic boundary condition, or no flux boundary conditions

$$\frac{\partial c_i}{\partial \vec{n}}|_{\partial\Omega} = 0, \ i = 1, 2; \frac{\partial \phi}{\partial \vec{n}}|_{\partial\Omega} = 0.$$
(3.2)

It is also possible to use the Dirichlet boundary condition $\phi|_{\partial\Omega} = 0$ or a Robin type boundary condition $(\alpha\phi + \beta \frac{\partial\phi}{\partial\pi})|_{\partial\Omega} = 0$. In the above, the unknown are c_i , the density of the *i*-th species, and ϕ , the internal electric

In the above, the unknown are c_i , the density of the *i*-th species, and ϕ , the internal electric potential, $D_i > 0$ is the diffusion constant of the *i*-th specie (i = 1, 2), z_i are the valence constant and χ_1, χ_2 are dimensionless parameters. To make the formulas below more concise, in the following we fix $z_1 = 1$, $z_2 = -1$ and $\chi_1 = \chi_2 = 1$.

Using the identity $\nabla \psi = \psi \nabla \log \psi$, we can rewrite (3.1) as a Wasserstein gradient flow

$$\frac{\partial c_1}{\partial t} = D_1 \nabla \cdot (c_1 \nabla \log c_1 + c_1 \nabla \phi), \qquad (3.3a)$$

$$\frac{\partial c_2}{\partial t} = D_2 \nabla \cdot (c_2 \nabla \log c_2 - c_2 \nabla \phi), \qquad (3.3b)$$

$$\Delta \phi = c_1 - c_2, \tag{3.3c}$$

with the free energy

$$E(c_1, c_2, \phi) = \int_{\Omega} c_1(\log c_1 - 1) + c_2(\log c_2 - 1) + \frac{1}{2} |\nabla \phi|^2 dx.$$
(3.4)

Indeed, taking the inner product of (3.3a) with $\log c_1 + \phi$ and of (3.3b) with $\log c_2 - \phi$, summing them up along with $(-\Delta \partial_t \phi = \partial_t (c_1 - c_2), \phi)$, we obtain the following energy law:

$$\frac{dE(c_1, c_2, \phi)}{dt} = -\int_{\Omega} \left(D_1 c_1 |\nabla(\log c_1 + \phi)|^2 + D_2 c_2 |\nabla(\log c_2 - \phi)|^2 \right) dx.$$
(3.5)

Note that the form of the free energy, as well as the well-posedness of (3.3), requires $c_1, c_2 > 0$. Therefore, it is of critical importance that numerical schemes for the PNP system preserve positivity.

On the other hand, we also derive from (3.3) and (3.2) that

$$\frac{d}{dt}\int_{\Omega}c_i dx = 0, \quad i = 1, 2, \tag{3.6}$$

i.e., the mass for each component is conserved.

3.2. Positivity preserving SAV schemes. As explained in Section 2, we can preserve the positivity using suitable function transforms. Since only c_1 , c_2 are positivity preserving, we only make function transform for c_1 , c_2 . More precisely, we introduce two new functions p_1 and p_2 through

$$c_i = T(p_i) := \exp(p_i), \ i = 1, 2,$$
(3.7)

which implies in particular $c_i > 0, i = 1, 2$.

Substituting (3.7) into (3.3a)-(3.3b), we obtain

$$\frac{\partial p_1}{\partial t} = D_1(\Delta p_1 + |\nabla p_1|^2 + \nabla p_1 \cdot \nabla \phi + \Delta \phi), \qquad (3.8a)$$

$$\frac{\partial p_2}{\partial t} = D_2(\Delta p_2 + |\nabla p_2|^2 - \nabla p_1 \cdot \nabla \phi - \Delta \phi).$$
(3.8b)

Note that for this transform, we have $T'(p_i) = T''(p_i) = T(p_i)$, so the transformed equations are not too complicated.

Next we split the free energy $E(c_1, c_2, \phi)$ into the sum of $E_0(\phi) := \frac{1}{2}(\nabla \phi, \nabla \phi)$ and $E_1(c_1, c_2) := \int_{\Omega} c_1(\log c_1 - 1) + c_2(\log c_2 - 1)dx$. It is clear that $E_1(c_1, c_2)$ is convex and bounded from below in the admissible set $\mathcal{D} := \{(c_1, c_2) : c_1, c_2 > 0\}$, so we assume that for some $C_0 > 0$,

$$E_1(c_1, c_2) \ge -C_0 + 1, \tag{3.9}$$

and define a SAV $r(t) = E_1(c_1, c_2) + C_0 > 1$. Then, the total free energy E and its time derivative can be rewritten as

$$E(c_1, c_2, \phi) = \frac{1}{2} (\nabla \phi, \nabla \phi) + r(t) = E_0(\phi) + r(t), \qquad (3.10a)$$

$$\frac{dE}{dt} = \frac{dE_0}{dt} + r_t. \tag{3.10b}$$

Denote $\mu_1 = \log c_1 + \phi$, $\mu_2 = \log c_2 - \phi$, we can reformulate (3.3) and (3.5) as

$$\frac{\partial p_1}{\partial t} = D_1(\Delta p_1 + |\nabla p_1|^2 + \nabla p_1 \cdot \nabla \phi + \Delta \phi), \qquad (3.11a)$$

$$\frac{\partial p_2}{\partial t} = D_2(\Delta p_2 + |\nabla p_2|^2 - \nabla p_2 \cdot \nabla \phi - \Delta \phi), \qquad (3.11b)$$

$$c_1 = \exp(p_1), \ c_2 = \exp(p_2),$$
(3.11c)

$$-\Delta\phi = c_1 - c_2, \tag{3.11d}$$

$$\frac{dE_0}{dt} + r_t = -\frac{E_0(\phi) + r(t)}{E(c_1, c_2, \phi) + C_0} \int_{\Omega} \left(D_1 c_1 |\nabla \mu_1|^2 + D_2 c_2 |\nabla \mu_2|^2 \right) dx.$$
(3.11e)

We remark that since the above system is equivalent to the original system (3.3), the masses of c_i are still conserved, but that of p_i are not.

We now construct k-th order SAV schemes $(1 \le k \le 6)$ for the above system in a uniform setting. Given $(c_i^j, p_i^j, \phi^j, r^j, \xi^j)$, $i = 1, 2, j = n, n - 1, \dots, n - k + 1$ such that

$$\int_{\Omega} c_i^j dx = \int_{\Omega} c_i^0 dx, \ i = 1, 2, \ j = n, n - 1, \cdots, n - k + 1,$$
(3.12)

we determine $(c_i^{n+1},p_i^{n+1},\lambda_i^{n+1}),\;i=1,2$ and $(\phi^{n+1},r^{n+1},\xi^{n+1})$ as follows:

$$\frac{\alpha_k p_i^{n+1} - A_k(p_i^n)}{\delta t} - D_i \Delta p_i^{n+1} = g_i(B_k(p_i^n), B_k(\phi^n)), \quad i = 1, 2,$$
(3.13)

$$\bar{c}_i^{n+1} = \exp(p_i^{n+1}), \ i = 1, 2,$$
(3.14)

$$\lambda_i^{n+1} \int_{\Omega} \alpha_k \bar{c}_i^{n+1} dx - \int_{\Omega} A_k(c_i^n) dx = 0, \ i = 1, 2,$$
(3.15)

$$c_i^{n+1} = \lambda_i^{n+1} \bar{c}_i^{n+1}, \ i = 1, 2, \tag{3.16}$$

$$-\Delta\bar{\phi}^{n+1} = c_1^{n+1} - c_2^{n+1}, \tag{3.17}$$

$$\frac{1}{\delta t} \left(E_0(\bar{\phi}^{n+1}) - E_0(\bar{\phi}^n) + r^{n+1} - r^n \right)
= -\frac{E_0(\bar{\phi}^{n+1}) + r^{n+1}}{E(c_1^{n+1}, c_2^{n+1}, \bar{\phi}^{n+1}) + C_0} \int_{\Omega} \left(D_1 c_1^{n+1} |\nabla \mu_1^{n+1}|^2 + D_2 c_2^{n+1} |\nabla \mu_2^{n+1}|^2 \right) dx, \quad (3.18)$$

$$\xi^{n+1} = \frac{E_0(\phi^{n+1}) + r^{n+1}}{E(c_1^{n+1}, c_2^{n+1}, \bar{\phi}^{n+1}) + C_0},\tag{3.19}$$

$$\phi^{n+1} = \eta_k^{n+1} \bar{\phi}^{n+1} \text{ with } \eta_k^{n+1} = 1 - (1 - \xi^{n+1})^k, \tag{3.20}$$

together with homogeneous Neumann boundary conditions

$$\frac{\partial p_i^{n+1}}{\partial \vec{n}}|_{\partial\Omega} = 0, \ i = 1, 2; \frac{\partial \phi^{n+1}}{\partial \vec{n}}|_{\partial\Omega} = 0,$$
(3.21)

where $\mu_1^{n+1} = \log c_1^{n+1} + \bar{\phi}^{n+1}, \ \mu_2^{n+1} = \log c_2^{n+1} - \bar{\phi}^{n+1}, \ \alpha_k, \ A_k \text{ and } B_k \text{ are the same as in the last}$ section, and

$$g_1(p_1,\phi) = D_1(|\nabla p_1|^2 + \nabla p_1 \cdot \nabla \phi + \Delta \phi),$$

$$g_2(p_2,\phi) = D_1(|\nabla p_2|^2 - \nabla p_2 \cdot \nabla \phi - \Delta \phi).$$

Similar to the last section, we have the following remarks:

- Clearly, (3.13) is a k-th order semi-implicit scheme for (3.11a)-(3.11b). We then derive from (3.14)-(3.17) that λ_i^{n+1} is k-th order approximations to 1, c_i^{n+1} and $\bar{\phi}^{n+1}$ are k-th order approximations to $c_i(t_{n+1})$ and $\phi(t_{n+1})$.
- (3.18) is a first-order approximation to (3.11e), so r^{n+1} is a first-order approximation to $E_1(c_1^{n+1}, c_2^{n+1})$ and $\xi^{n+1} = 1 + O(\delta t)$ which implies that $\eta_k^{n+1} = 1 + O(\delta t^k)$. Therefore, ϕ^{n+1} is also a k-th order approximation of $\phi(t_{n+1})$.
- The scheme (3.13)- (3.20) can be efficiently implemented by the following steps:

 - 1. solve p_i^{n+1} from (3.13); 2. compute $\bar{c}_1^{n+1}, \bar{c}_2^{n+1}$ from (3.14) and compute λ_i^{n+1} explicitly from (3.15); 3. update c_1^{n+1}, c_2^{n+1} from (3.16) and solve $\bar{\phi}$ from (3.17); 4. compute r^{n+1} explicitly from (3.18) and then obtain ξ^{n+1} from (3.19);

 - 5. update ϕ^{n+1} from (3.20), goto next step.

The main computational cost is to solve the linear equations with constant coefficients in (3.13) and (3.17).

We have the following results:

Theorem 2. Given $c_i^j > 0$, $p_i^j = \log c_i^j$, ϕ^j , and r^j such that $\int_{\Omega} c_i^j dx = \int_{\Omega} c_i^0 dx$ for i = 1, 2 and $j = n, n-1, \ldots, n-k+1$. The scheme (3.13)-(3.20) admits a unique solution satisfying the following properties unconditionally:

- 1. Positivity preserving: c_1^{n+1} , $c_2^{n+1} > 0$. 2. Mass conserving: $\int_{\Omega} c_i^{n+1} dx = \int_{\Omega} c_i^0 dx$ for i = 1, 2.

3. Unconditionally energy dissipation with a modified energy defined by $\bar{E}^n = E_0(\bar{\phi}^n) + r^n$: More precisely, if $\bar{E}^n \ge 0$, we have $\bar{E}^{n+1} \ge 0$, $\xi^{n+1} \ge 0$ and

$$\bar{E}^{n+1} - \bar{E}^n = -\xi^{n+1} \int_{\Omega} \left(D_1 \, c_1^{n+1} |\nabla \mu_1^{n+1}|^2 + D_2 \, c_2^{n+1} |\nabla \mu_2^{n+1}|^2 \right) dx \le 0.$$
(3.22)

4. There exists constant M_k , such that

$$\sqrt{E_0[\phi^n]} \le M_k, \,\forall n. \tag{3.23}$$

Proof. From (3.14), we obviously have \bar{c}_1^{n+1} , $\bar{c}_2^{n+1} > 0$.

We derive from the assumption that $\int_{\Omega} c_i^j dx = \int_{\Omega} c_i^0 dx$ for i = 1, 2 and $j = n, n - 1, \dots, n - k + 1$, and the definition of coefficients α_k and A_k in section 2 that

$$\int_{\Omega} A_k(c_i^n) dx = \alpha_k \int_{\Omega} c_i^0 dx$$

It then follows from (3.12) and (3.15) that

$$\alpha_k \lambda_i^{n+1} \int_{\Omega} \bar{c}_i^{n+1} dx = \alpha_k \int_{\Omega} c_i^0 dx, \qquad (3.24)$$

which, along with $\bar{c}_i^{n+1} > 0$, implies that $\lambda_i^{n+1} > 0$. Hence, we have c_1^{n+1} , $c_2^{n+1} > 0$, and we derive from the above and (3.16) that $\int_{\Omega} c_i^{n+1} dx = \int_{\Omega} c_i^0 dx$ for i = 1, 2.

It follows from (3.18) that

$$E_{0}(\bar{\phi}^{n+1}) + r^{n+1} = \frac{E_{0}(\phi^{n}) + r^{n}}{1 + \delta t \frac{\int_{\Omega} \left(D_{1} c_{1}^{n+1} | \nabla \mu_{1}^{n+1} |^{2} + D_{2} c_{2}^{n+1} | \nabla \mu_{2}^{n+1} |^{2} \right) dx}{E(c_{1}^{n+1}, c_{2}^{n+1}, \bar{\phi}^{n+1}) + C_{0}} \ge 0.$$
(3.25)

Therefore, we derive from (3.19) that $\xi^{n+1} \ge 0$, which, together with (3.18), implies (3.22).

Denote $M := \overline{E}^0$, then (3.22) implies $\overline{E}^n \leq M, \forall n$. It follows from (3.18) and (3.9) that

$$|\xi^{n+1}| = \frac{\bar{E}^{n+1}}{E(c_1^{n+1}, c_2^{n+1}, \bar{\phi}^{n+1}) + C_0} \le \frac{M}{E_0(\bar{\phi}^{n+1}) + 1}.$$
(3.26)

Since $\eta_k^{n+1} = 1 - (1 - \xi^{n+1})^k$, there exists a polynomial P_{k-1} of degree k-1 and a constant $M_k > 0$ such that

$$|\eta_k^{n+1}| = |\xi^{n+1} P_{k-1}(\xi^{n+1})| \le \frac{M_k}{E_0(\bar{\phi}^{n+1}) + 1}.$$
(3.27)

Therefore, by the fact that $\sqrt{A} \leq A + 1$ for all $A \geq 0$, we derive

$$\sqrt{E_0[\phi^{n+1}]} = |\eta_k^{n+1}| \sqrt{E_0[\bar{\phi}^{n+1}]} \le M_k.$$

Remark 2. We emphasize that both \bar{c}_i^{n+1} (resp. $\bar{\phi}_i^{n+1}$) and c_i^{n+1} (resp. ϕ^{n+1}) are k-th order approximation to $c_i(\cdot, t_{n+1})$ (resp. $\phi(\cdot, t_{n+1})$), i = 1, 2.

Obviously, the positivity of c_i will be preserved with any spatial approximation of the schemes (3.13)-(3.20).

It is clear from the proof of the above theorem that the mass conservation and the energy dissipation (3.22) still hold for any fully discrete schemes.

4. Bound preserving schemes for Keller-Segel equations

We first introduce the Keller-Segel equations, followed by the construction of bound preserving schemes for one particular case of the Keller-Segel equations whose solution is bound preserving.

4.1. Keller-Segel equations. To fix the idea, we consider the following Keller-Segel system with only one organism and one chemoattractant in a bounded domain Ω :

$$\frac{\partial u}{\partial t} = D\big(\gamma \Delta u - \chi \nabla \cdot (\eta(u) \nabla \phi)\big),\tag{4.1a}$$

$$\tau \frac{\partial \phi}{\partial t} = \mu \Delta \phi - \alpha \phi + \chi u, \tag{4.1b}$$

with either periodic boundary conditions, or no-flux boundary conditions on u and the Neumann boundary conditions on ϕ ,

$$\gamma \frac{\partial u}{\partial \vec{n}} - \chi \eta(u) \frac{\partial \phi}{\partial \vec{n}} = 0, \ \frac{\partial \phi}{\partial \vec{n}} = 0 \ on \ \partial \Omega.$$
(4.2)

Here, the unknown are u, the concentration of the organism, and ϕ , the concentration of the chemoattractant. The parameters $D, \gamma, \chi, \tau, \mu, \alpha$ are all positive. The function $\eta(u) \ge 0$ describes the concentration-dependent mobility. It is a smooth function with $\eta(0) = 0$.

The model is a parabolic-parabolic system when $\tau > 0$, and a parabolic-elliptic system when $\tau = 0$.

The system (4.1) with (4.2) can be interpreted as a gradient flow about (u, ϕ) . To this end, we choose f(u) such that $f''(u) = 1/\eta(u)$, and define the free energy

$$E[u,\phi] = \int_{\Omega} (\gamma f(u) - \chi u\phi + \frac{\mu}{2} |\nabla \phi|^2 + \frac{\alpha}{2} \phi^2) dx.$$

$$(4.3)$$

Then writing $\Delta u = \nabla \cdot \left(\frac{1}{f''(u)} \nabla f'(u)\right)$, we can rewrite (4.1) as

$$\frac{\partial u}{\partial t} = D\nabla \cdot \left(\frac{1}{f''(u)}\nabla(\gamma f'(u) - \chi\phi)\right) = D\nabla \cdot \left(\frac{1}{f''(u)}\nabla\frac{\delta E}{\delta u}\right),\tag{4.4a}$$

$$\tau \frac{\partial \phi}{\partial t} = \mu \Delta \phi - \alpha \phi + \chi u = -\frac{\delta E}{\delta \phi}.$$
(4.4b)

Taking the inner products of (4.4a) with $\frac{\delta E}{\delta u}$, and of (4.4b) with $\frac{\partial \phi}{\partial t}$, and summing up the results, we obtain the energy dissipation law:

$$\frac{dE[u(t),\phi(t)]}{dt} = -\int_{\Omega} \left[D\frac{1}{f''(u)} \left(\nabla\frac{\delta E}{\delta u}\right)^2 + \tau \left(\frac{\partial\phi}{\partial t}\right)^2\right] dx.$$
(4.5)

We now consider several typical choices of $\eta(u)$ and the corresponding function f(u).

- (i) The classical Keller-Segel system: $\eta(u) = u$. We can choose $f(u) = u \log u u$ with the domain of definition $(0, +\infty)$. In this case, it is known that its solution can blow up in finite time if the initial mass is large enough [3, 4, 5].
- (ii) Keller-Segel system with a bounded mobility: a typical choice [34, 35] is $\eta(u) = \frac{u}{1+\kappa u}$ ($\kappa > 0$). In this case, we can choose $f(u) = u \log u - u + \kappa u^2/2$ with the domain of definition $(0, +\infty)$.
- (iii) Keller-Segel system with a saturation concentration: $\eta(u) = u(1 u/M)$, where M > 0 is the saturation concentration, and the mobility tends to zero when it is near saturation [8, 16]. In this case, we can choose $f(u) = u \log u + (M u) \log(1 u/M)$ with the domain of definition (0, M).

Hence, the solution of the Keller-Segel system is positivity preserving in cases (i) and (ii), and bound preserving in case (iii). Furthermore, we observe from (4.1) that

$$\frac{d}{dt} \int_{\Omega} u dx = 0. \tag{4.6}$$

To simplify the presentation, we shall only consider the third case where the solution is bound preserving. For the first and second order cases, the solution is positivity preserving, so one can construct positivity preserving schemes for these two cases similarly by replacing the mapping below with $Y(v) = \exp(v)$ as in the last section. 4.2. Bound preserving SAV schemes. We set $\eta(u) = u(1 - u/M)$ and $f(u) = u \log u + (M - u) \log(1 - u/M)$, and split $E[u, \phi]$ into two parts as follows

$$E[u,\phi] = \int_{\Omega} (\gamma f(u) - \chi u\phi + \frac{\alpha}{4}\phi^2) dx + \int_{\Omega} (\frac{\mu}{2}|\nabla\phi|^2 + \frac{\alpha}{4}\phi^2) dx = E_1[u,\phi] + E_0[\phi].$$
(4.7)

Note that $f(u) = u \log u + (M - u) \log(1 - u/M)$ implies that $u \in (0, M)$. Along with $\alpha > 0$ and f is strictly convex, it is easy to see that E_1 is bounded from below. Hence, there exists $C_0 > 0$ such that,

$$E_1[u,\phi] \ge -C_0 + 1. \tag{4.8}$$

Due to the form of f(u), it is necessary that the range of numerical solution is also in (0, M). To this end, we consider the transform

$$u = T(v) := \frac{M}{2} \tanh(v) + \frac{M}{2}.$$
(4.9)

As $tanh(x) \in (-1,1)$, $\forall x \in (-\infty, +\infty)$, then for $v \in (-\infty, +\infty)$, we have $u \in (0, M)$. Since ϕ is not bound preserving, we do not need to transform ϕ .

Substituting (4.9) into (4.1a), we obtain the equation for v

$$\frac{\partial v}{\partial t} = D\gamma \Delta v + D\gamma \frac{\tanh''(v)}{\tanh'(v)} |\nabla v|^2 - \frac{2D\chi}{M \tanh'(v)} \nabla \cdot (\eta(u) \nabla \phi).$$
(4.10)

Note that $\tanh'(x) = 1 - \tanh^2(x)$, we know $\tanh'(v) \neq 0$ and (4.10) is well-defined.

We introduce $r(t) = E_1(u, \phi) + C_0 \ge 1$. Then, we have

$$E[u,\phi] = \frac{\mu}{2} \left(\phi, -\Delta\phi\right)_{\Omega} + \frac{\alpha}{4} \left(\phi,\phi\right)_{\Omega} + r = E_0(\phi) + r, \qquad (4.11a)$$

$$\frac{d}{dt}E[u,\phi] = \mu(\phi_t, -\Delta\phi)_{\Omega} + \frac{\alpha}{2}(\phi_t, \phi)_{\Omega} + r_t = \frac{dE_0(\phi)}{dt} + r_t.$$
(4.11b)

We can reformulate (4.1) and (4.5) as

$$\frac{\partial v}{\partial t} = D\gamma \Delta v + \left(D\gamma \frac{\tanh''(v)}{\tanh'(v)} |\nabla v|^2 - \frac{2D\chi}{M\tanh'(v)} \nabla \cdot (\eta(u)\nabla\phi) \right), \tag{4.12a}$$

$$u = \frac{M}{2} \tanh(v) + \frac{M}{2},\tag{4.12b}$$

$$\tau \frac{\partial \phi}{\partial t} = \mu \Delta \phi - \alpha \phi + \chi u, \qquad (4.12c)$$

$$\frac{dE_0(\phi)}{dt} + r_t = -\frac{E_0(\phi) + r(t)}{E(u,\phi) + C_0} \int_{\Omega} \left[D\frac{1}{f''(u)} \left(\nabla\frac{\delta E}{\delta u}\right)^2 + \tau \left(\frac{\partial\phi}{\partial t}\right)^2\right] dx.$$
(4.12d)

We now construct k-th order schemes for (4.12) in a uniform setting.

Given (v^i, u^i, ϕ^i, r^i) , $i = n, n - 1, \dots, n - k + 1$, we find $(v^{n+1}, u^{n+1}, \phi^{n+1}, r^{n+1})$ as follows:

$$\frac{\alpha_k v^{n+1} - A_k(v^n)}{\delta t} - D\gamma \Delta v^{n+1} = g(B_k(v^n), B_k(u^n), B_k(\phi^n)),$$
(4.13)

$$\bar{u}^{n+1} = \frac{M}{2} \tanh(v^{n+1}) + \frac{M}{2}, \tag{4.14}$$

$$\lambda^{n+1} \int_{\Omega} \alpha_k \bar{u}^{n+1} dx - \int_{\Omega} A_k(u^n) dx = 0, \qquad (4.15)$$

$$u^{n+1} = \lambda^{n+1} \bar{u}^{n+1}, \tag{4.16}$$

$$\tau \frac{\alpha_k \bar{\phi}^{n+1} - A_k(\bar{\phi}^n)}{\delta t} = \mu \Delta \bar{\phi}^{n+1} - \alpha \bar{\phi}^{n+1} + \chi u^{n+1}, \qquad (4.17)$$

$$\frac{1}{\delta t} \Big(E_0(\bar{\phi}^{n+1}) - E_0(\bar{\phi}^n) + r^{n+1} - r^n \Big) \\
= -\frac{E_0(\bar{\phi}^{n+1}) + r^{n+1}}{E[\bar{u}^{n+1}, \bar{\phi}^{n+1}] + C_0} \int_{\Omega} \Big[\frac{D}{f''(\bar{u}^{n+1})} \Big(\nabla \frac{\delta E}{\delta u}(\bar{u}^{n+1}) \Big)^2 + \tau \Big(\frac{\bar{\phi}^{n+1} - \bar{\phi}^n}{\delta t} \Big)^2 \Big] dx,$$
(4.18)

$$\xi^{n+1} = \frac{E_0(\bar{\phi}^{n+1}) + r^{n+1}}{E[\bar{u}^{n+1}, \bar{\phi}^{n+1}] + C_0},\tag{4.19}$$

$$\phi^{n+1} = \eta_k^{n+1} \bar{\phi}^{n+1} \text{ with } \eta_k^{n+1} = 1 - (1 - \xi^{n+1})^k, \tag{4.20}$$

where the constant α_k , operators A_k , B_k are defined in Section 2, and

$$g(u, v, \phi) = D\gamma \frac{\tanh''(v)}{\tanh'(v)} |\nabla v|^2 - \frac{2D\chi}{M \tanh'(v)} \nabla \cdot (\eta(u) \nabla \phi).$$
(4.21)

Essential properties of the above schemes are as follows:

- (4.13) and (4.17) are k-th order semi-implicit schemes for (4.12a) and (4.12c), (4.15) is kth order approximation to (4.6), which imply that v^{n+1} , λ^{n+1} , u^{n+1} , ϕ^{n+1} are k-th order approximations to $v(t_{n+1}), 1, u(t_{n+1}), \phi(t_{n+1})$.
- (4.18) is a first-order approximation to (4.12d), which implies that r^{n+1} is a first-order approximation to $r(t_{n+1})$. Then, (4.19) implies that $\xi^{n+1} = 1 + O(\delta t)$, which in turn implies $\eta_k^{n+1} = 1 + O(\delta t)^k$ and ϕ^{n+1} is a k-th order approximations to $\phi(t_{n+1})$.
- The above scheme can be efficiently implemented as follows:
 - 1. solve v^{n+1} from (4.13);
 - 2. compute \bar{u}^{n+1} from (4.14) and compute λ^{n+1} explicitly from (4.15);

 - 3. update u^{n+1} from (4.16); 4. with u^{n+1} known, solve $\overline{\phi}^{n+1}$ from (4.17);
 - 5. with \bar{u}^{n+1} , $\bar{\phi}^{n+1}$ known, determine r^{n+1} explicitly from (4.18);
 - 6. compute ξ^{n+1} from (4.19) and update ϕ^{n+1} from (4.20), go to the next step.

We have the following results:

Theorem 3. Given u^i , ϕ^i , v^i and r^i such that

$$\int_{\Omega} u^{i} dx = \int_{\Omega} u^{0} dx, \quad i = n, n - 1, \dots, n - k + 1.$$
(4.22)

Then, the scheme (4.13)-(4.20) admits a unique solution satisfying the following properties unconditionally:

- 1. Bound preserving for \bar{u}^{n+1} : i.e., the range of \bar{u}^{n+1} is in (0, M).
- 2. Mass conservation: i.e., $\int_{\Omega} u^{n+1} dx = \int_{\Omega} u^0 dx$.
- 3. Unconditionally energy dissipation with a modified energy defined by $\bar{E}^n = E_0(\bar{\phi}^{n+1}) + r^n$: More precisely, if $\overline{E}^n \ge 0$, we have $\overline{E}^{n+1} \ge 0$, $\xi^{n+1} \ge 0$ and

$$\bar{E}^{n+1} - \bar{E}^n = -\xi^{n+1} \int_{\Omega} \left[\frac{1}{f''(\bar{u}^{n+1})} \left(\nabla \frac{\delta E}{\delta u}(\bar{u}^{n+1}) \right)^2 + \tau \left(\frac{\bar{\phi}^{n+1} - \bar{\phi}^n}{\delta t} \right)^2 \right] dx \le 0.$$
(4.23)

4. There exists constant M_k , such that

$$\sqrt{E_0[\phi^n]} = \sqrt{\int_{\Omega} (\frac{\mu}{2} |\nabla \phi^n|^2 + \frac{\alpha}{4} (\phi^n)^2) dx} \le M_k, \,\forall n.$$

$$(4.24)$$

Proof. The proof is essentially the same as that of Theorem 2. For the readers' convenience, we still carry it out below.

We derive from (4.14) that the range of \bar{u}^{n+1} is in (0, M).

Noting the definition of coefficients α_k and A_k in Section 2, it follows from (4.22) and (4.15) that

$$\alpha_k \lambda^{n+1} \int_{\Omega} \bar{u}^{n+1} dx = \alpha_k \int_{\Omega} u^0 dx, \qquad (4.25)$$

which implies $\lambda^{n+1} > 0$, and consequently $u^{n+1} > 0$. Furthermore, along with (4.16), it also implies $\int_{\Omega} u^{n+1} dx = \int_{\Omega} u^0 dx$.

It follows from (4.18) that

$$E_0(\bar{\phi}^{n+1}) + r^{n+1} = \frac{E_0(\phi^n) + r^n}{1 + \frac{\delta t \int_{\Omega} \left[\frac{D}{f''(\bar{u}^{n+1})} \left(\nabla \frac{\delta E}{\delta u}(\bar{u}^{n+1})\right)^2 + \tau \left(\frac{\bar{\phi}^{n+1} - \bar{\phi}^n}{\delta t}\right)^2\right] dx}{E(\bar{u}^{n+1}, \bar{\phi}^{n+1}) + C_0} \ge 0.$$

Therefore, we derive from (4.19) that $\xi^{n+1} \ge 0$, which, together with (4.18), implies the energy dissipation.

Denote $M := \overline{E}^0$, then (4.23) implies $\overline{E}^n \leq M, \forall n$. Now, it follows from (4.19) and (4.8) that

$$|\xi^{n+1}| = \frac{\bar{E}^{n+1}}{E(\bar{u}^{n+1}, \bar{\phi}^{n+1}) + C_0} \le \frac{M}{E_0(\bar{\phi}^{n+1}) + 1}.$$
(4.26)

Since $\eta_k^{n+1} = 1 - (1 - \xi^{n+1})^k$, there exists a polynomial P_{k-1} of degree k - 1 and a constant $M_k > 0$ such that

$$|\eta_k^{n+1}| = |\xi^{n+1} P_{k-1}(\xi^{n+1})| \le \frac{M_k}{E_0(\bar{\phi}^{n+1}) + 1}.$$
(4.27)

Therefore, by the fact that $\sqrt{A} \leq A + 1$ for all $A \geq 0$, we obtain

$$\sqrt{E_0[\phi^{n+1}]} = |\eta_k^{n+1}| \sqrt{E_0[\bar{\phi}^{n+1}]} \le M_k.$$

We only consider the semi-discretization in time in this paper. As for fully discretizations, we have the following remarks:

Remark 3. We emphasize that both \bar{u}^{n+1} (resp. $\bar{\phi}_i^{n+1}$) and u^{n+1} (resp. ϕ^{n+1}) are k-th order approximation to $u(\cdot, t_{n+1})$ (resp. $\phi(\cdot, t_{n+1})$). While only the range of \bar{u}^{n+1} is guaranteed in (0, M), the range of u^{n+1} is in $(0, M + O(\delta t^k))$.

The positivity of \bar{u}^{n+1} and u^{n+1} will be preserved with any spatial approximation of the schemes (4.13)-(4.20).

It is also clear from the proof of the above theorem that the mass conservation and the energy dissipation (4.23) still hold for any fully discrete schemes.

One can easily extend these schemes to deal with Keller-Segel equations with multiple organisms. We leave the detail to the interested readers.

5. Numerical examples

In this section, we provide some numerical examples to validate our numerical schemes.

5.1. Allen-Cahn equation with a singular potential. We first use the schemes presented in Section 2 to solve the Allen-Cahn equation with a singular potential. In all examples for the Allen-Cahn equation, we consider problems with periodic boundary conditions and use a Fourier-spectral method to discretize in space.

Example 1. We consider the Allen-Cahn equation [1]

$$\partial_t u = -\frac{\delta E}{\delta u} = \varepsilon^2 \Delta u + \lambda u - \ln(1+u) + \ln(1-u), \tag{5.1}$$

where $\varepsilon > 0, \lambda > 0$ and

$$E(u) = \int_{\Omega} \left(\frac{\varepsilon^2}{2} |\nabla\phi|^2 - \frac{\lambda}{2}u^2 + (1+\phi)\ln(1+u) + (1-u)\ln(1-u)\right) dx,$$
(5.2)

is the free energy with a singular potential. The well posedness of the above equation requires that $u \in (-1, 1)$.

We use the transformation $u = \tanh(v)$, in the scheme (2.6)-(2.10).

We first test the accuracy with the following exact solution and the corresponding external forcing \boldsymbol{f}

$$u(x, y, t) = \left(\exp(-\sin^2(\pi x)) - \exp(-\sin^2(\pi y))\right)\sin(t),$$
$$f = \partial_t u + \frac{\delta E}{\delta u}.$$

The parameters are chosen as $\varepsilon = 0.1$, $\lambda = 3$ and the computational domain is $(0, 2) \times (0, 2)$. Fourier spectral method with 96 × 96 modes is used for special discretization. we plot in Figure 1 (a) the errors of the first- and second-order schemes at $t_n = 1$, and in Figure 1 (b), the errors of the thirdand fourth-order schemes at $t_n = 1$. Expected convergence rates are observed for all cases.



FIGURE 1. (*Example 1.*) Accuracy test for the Allen-Cahn equation using the new SAV/BDFk schemes (k = 1, 2, 3, 4).

Next, we consider the spinodal decomposition of a homogeneous mixture into two coexisting phases governed by the Allen-Cahn equation. The parameters are chosen as $\varepsilon = 0.005$, $\lambda = 3$ and the computational domain is $(0, 1) \times (0, 1)$. The time step is set to $\delta t = 0.001$. Fourier spectral method with 256 × 256 modes is used for space discretization. The initial condition is chosen as a random variable with uniform distribution in [-0.05, 0.05]. We plot the evolution of energy, the evolution of max u, min u and four snapshots in Figure 2.



FIGURE 2. Example 1. Spinodal decomposition by the Allen-Cahn equation. The simulation is obtained with $\delta t = 0.001$ using the scheme (2.6)-(2.10).

5.2. **Two-component PNP system.** We present here numerical results of using the scheme (3.13)-(3.20) to solve the two-component PNP system (3.1).

Example 2. We test accuracy by considering the two-component PNP system (3.3), i.e. we fix $z_1 = 1$, $z_2 = -1$ and $\chi_1 = \chi_2 = 1$ in (3.1). We first consider the following manufactured exact solutions in $\Omega = (-0.5, 0.5) \times (-0.5, 0.5)$ with suitable external forcing:

$$c_1(x, y, t) = 1.1 + \sin(\pi x)\sin(\pi y)\sin(t), \qquad (5.3a)$$

$$c_2(x, y, t) = 1.1 - \sin(\pi x) \sin(\pi y) \sin(t),$$
 (5.3b)

$$\phi(x, y, t) = \frac{1}{\pi^2} \sin(\pi x) \sin(\pi y) \sin(t).$$
(5.3c)

In this example, we use Legendre spectral-Galerkin method and $(N_x, N_y) = (40, 40)$. Other parameters are $D_1 = D_2 = 1$. Define the L^2 -error at t_n as $\sqrt{\|c_1^n - c_1(t_n)\|^2 + \|c_2^n - c_2(t_n)\|^2}$, we plot in Figure 3 (a) the errors of the first- and second-order schemes at $t_n = 1$, and in Figure 3 (b), the errors of the third- and fourth-order schemes at $t_n = 10$. Expected convergence rates are observed for all cases.

Next, we test the accuracy in the computational domain $\Omega = (0, 2\pi) \times (0, 2\pi)$ with periodic boundary condition and the initial conditions are given by

$$c_1(x, y, 0) = 1.1 + \sin(x)\cos(y), \tag{5.4a}$$

$$c_2(x, y, 0) = 1.1 - \sin(x)\cos(y).$$
 (5.4b)

In this example, we use Fourier-spectral method to discretize in space and $(N_x, N_y) = (128, 128)$. Other parameters are $D_1 = D_2 = 1$. We generate the reference solution by the fourth-order scheme with $\delta t = 0.0001$. Define the L^2 -error at t_n as above, we plot in Figure 3 (c) the errors of the firstand second-order schemes at $t_n = 0.1$, and in Figure 3 (d), the errors of the third- and fourth-order schemes at $t_n = 0.1$. Expected convergence rates are observed for all cases.



(a) BDF1 and BDF2 for PNP with known exact solution (b) BDF3 and BDF4 for PNP with known exact solution



(c) BDF1 and BDF2 for PNP with unknown exact solution (d) BDF3 and BDF4 for PNP with unknown exact solution

FIGURE 3. *Example* 2. Accuracy test for PNP equation using the SAV/BDFk schemes (k = 1, 2, 3, 4).

Example 3. In this example, we test the so-called Gouy-Chapman model [13] which is used to describe the evolution of the distributions of the ions.

We consider the PNP system (3.1) in (-1, 1) with the following parameters: $D_1 = D_2 = 1$, $z_1 = 1$, $z_2 = -1$, and $\chi_1 = 3.1$, $\chi_2 = 125.4$. The boundary conditions for c_i and ϕ are given as

$$\partial_x c_i + z_i \chi_1 c_i \partial_x \phi = 0, \ i = 1, 2, \tag{5.5}$$

$$\alpha\phi(t,-1) - \beta\phi_x(t,-1) = f_{-1}, \ \alpha\phi(t,1) + \beta\phi_x(t,1) = f_1, \ t \ge 0,$$
(5.6)

with $\alpha = 1$, $\beta = 4.63 \times 10^{-5}$, $f_{-1} = 1$ and $f_1 = -1$. For space discretization, we use Legendre spectral-Galerkin method. We set $\delta t = 0.001$ and used 80 nodes in space. The initial condition on c_i are $c_i(x, 0) = 1$, i = 1, 2 for all $-1 \le x \le 1$. The profiles of c_1 , c_2 and ϕ at different times are plotted in Fig 4, which are consistent with the results in [13]. In Fig 4(d), we also plot the mass evolution of c_i and p_i with $c_i = \exp(p_i)$, i = 1, 2. We can see the masses of c_i are well conserved, but that of p_i are not.

5.3. Keller-Segel equations. In this subsection, we present numerical results of using scheme (4.13)-(4.20) to solve the Keller-Segel equations (4.1).



FIGURE 4. Example 3. Gouy-Chapman model: Profiles of c_1, c_2 and ϕ

Example 4. We test the accuracy of the scheme. First consider the one-specie parabolic-elliptic $(\tau = 0)$ Keller-Segel equations (4.1) in $\Omega = (-0.5, 0.5) \times (-0.5, 0.5)$ with external forcing such that the exact solutions are given by

$$u(x, y, t) = \sin(\pi x)\sin(\pi y)\sin(t) + 1.1,$$
(5.7a)

$$\phi(x, y, t) = \frac{1}{2\pi^2} \sin(\pi x) \sin(\pi y) \sin(t) + 1.1.$$
(5.7b)

Other parameters are $D = \gamma = \mu = \alpha = \chi = 1$, M = 5. We use Legendre spectral-Galerkin method and $(N_x, N_y) = (40, 40)$ in space. Define the L^2 -error as $\sqrt{\|u^n - u(t_n)\|^2 + \|\phi^n - \phi(t_n)\|^2}$, we plot in Figure 5 (a) the errors at $t_n = 1$ for the first- and second-order schemes, and in Figure 5 (b) the errors at $t_n = 10$ for the third- and fourth-order schemes.

Next, we test the accuracy in $\Omega = (0, 2\pi) \times (0, 2\pi)$ with periodic boundary condition and the initial conditions are given by

$$u(x, y, 0) = \sin(x)\sin(y) + 1.1.$$
(5.8)

In this example, we use Fourier-spectral method to discretize in space and $(N_x, N_y) = (128, 128)$. Other parameters are $D = \gamma = \mu = \alpha = \chi = 1$, M = 3. We generate the reference solution by the fourth-order scheme with $\delta t = 0.0001$. Define the L^2 -error as above, we plot in Figure 5 (c) the errors at $t_n = 0.1$ for the first- and second-order schemes, and in Figure 5 (d) the errors at $t_n = 0.1$ for the third- and fourth-order schemes. As in the Allen-Cahn case and the PNP case, the expected convergence rates are observed for all cases.



(a) BDF1 and BDF2 for Keller-Segel with known exact (b) BDF3 and BDF4 for Keller-Segel with known exact solution



(c) BDF1 and BDF2 for Keller-Segel with unknown exact (d) BDF3 and BDF4 for Keller-Segel with unknown exact solution

FIGURE 5. Example 4. Accuracy test for Keller-Segel equations using the SAV/BDFk (4.13)-(4.20) (k = 1, 2, 3, 4).

Example 5. In this example, we consider the one-specie parabolic-elliptic ($\tau = 0$) Keller-Segel equations with the following initial condition

$$u(x, y, 0) = 4 \exp\left(-\frac{(x - L/2)^2 + (y - L/2)^2}{4}\right)$$
(5.9)

such that the total mass is large enough that chemotaxis happens, in $(0, 2\pi) \times (0, 2\pi)$ with the homogeneous Neumann boundary conditions. We use Legendre spectral-Galerkin method method with $(N_x, N_y) = (64, 64)$ nodes to discretize in space, and the second-order scheme with time step $\delta t = 0.001$. The parameters are chosen as $D = \gamma = \mu = 1$, $\chi = 1$, M = 100, $\alpha = 0.1$ and $L = 2\pi$.

We carry out simulation until the system reaches steady state at t = 8. Several snapshots of concentration at different times are shown in Figure 6, where we plot the snapshots by using smaller time steps and more nodes in the right hand side , and evolutions of max u, mass of u, mass of v and

energy are shown in Figure 7. These results agree well with those in [30] computed with a nonlinear scheme. In particular, the energy is dissipative at all time, and the mass of u is conserved up to machine accuracy.

Example 6. We consider the one specie parabolic-elliptic system with an initial condition with two bulges, given by

$$u(x,y,0) = 2\exp(-\frac{(x-3L/8)^2 + (y-3L/8)^2}{4}) + 2\exp(-\frac{(x-5L/8)^2 + (y-5L/8)^2}{4})$$
(5.10)

with $L = 4\pi$. We take M = 50 while all other settings are the same as Example 5. We use the third-order scheme, and plot the evolution of energy, maximum concentration and four snapshots of u in Figure 8. We observe that the energy is dissipative at all times, the maximum of u increases while the support of u shrinks to maintain the mass conservation.

Example 7. In this example, we consider the parabolic-elliptic Keller-Segel system with two species:

$$\frac{\partial u_1}{\partial t} = D_1 \big(\gamma_1 \Delta u_1 - \chi_1 \nabla \cdot (\eta_1(u_1) \nabla \phi) \big), \tag{5.11a}$$

$$\frac{\partial u_2}{\partial t} = D_2 \big(\gamma_2 \Delta u_2 - \chi_2 \nabla \cdot (\eta_2(u_2) \nabla \phi) \big), \tag{5.11b}$$

$$0 = \mu \Delta \phi - \alpha \phi + \chi_1 u_1 + \chi_2 u_2, \qquad (5.11c)$$

with the initial conditions

$$u_1(x,y,0) = u_2(x,y,0) = \phi(x,y,0) = 4\exp\left(-\frac{(x-L/2)^2 + (y-L/2)^2}{4}\right).$$
 (5.12)

The parameters are chosen as $D_1 = D_2 = \gamma_1 = \gamma_2 = \mu = \chi_1 = 1$, $\alpha = 0.1$ with all other settings are the same as Example 5. We use the first order scheme for this example. The results with two different chemotactic sensitivities with $\chi_2 = 0.1$ and $\chi_2 = 0.01$ are plotted in Figure 9 and 10 respectively.

In both cases, we observe accumulation for u_1 , while for u_2 , it diffuses first and then accumulates in the case $\chi_2 = 0.1$, and it keeps diffusing in the case $\chi_2 = 0.01$. These results are consistent with the results in [30].

6. Concluding remarks

For PDEs whose solutions are required to be positive or in a prescribed range, it is of critical importance to construct numerical schemes which are positivity or bound preserving. If the PDEs are also energy dissipative and/or mass conservative, it is important that the numerical schemes would be energy dissipative and/or mass conservative at the discrete level.

In this paper, we proposed a new approach to construct linear, positivity/bound preserving and unconditionally energy stable schemes for general dissipative systems whose solutions are positivity/bound preserving. The essential ideas of this new approach are (i) to first make a function transform so that the solution will always be positivity/bound preserving, and (ii) apply a new SAV approach presented in [20] to the transformed system and the original energy dissipation law to construct efficient and accurate time discretization schemes.

The resulting schemes enjoy remarkable properties such as positivity/bound preserving, unconditionally energy stable, can achieve high-order and with computational complexity similar to a semiimplicit scheme. We applied this approach to Allen-Cahn equation with a singular potential, and to Keller-Segel and Poisson-Nernst-Planck (PNP) equations which can be classified as Wasserstein gradient flows with an additional property of mass conservation.

While we only discussed semi-discretization in time in this paper, we pointed out that the energy dissipation, positivity or bound preserving and mass conservation can all be naturally carried over to consistent fully discretizations.



FIGURE 6. Example 5. Simulation of Keller-Segel equations with chemotaxis.



FIGURE 7. Example 5. Simulation of Keller-Segel equations with chemotaxis.

References

- S. M. Allen and J. W. Cahn. A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. Acta Metall. Mater., 27:1085–1095, 1979.
- [2] Martin Z. Bazant, Katsuyo Thornton, and Armand Ajdari. Diffuse-charge dynamics in electrochemical systems. *Physical review E*, 70(2):021506, 2004.
- [3] Piotr Biler and Tadeusz Nadzieja. Existence and nonexistence of solutions for a model of gravitational interaction of particles, i. In *Colloquium Mathematicae*, volume 66, pages 319–334, 1993.
- [4] Adrien Blanchet, Jean Dolbeault, and Benoît Perthame. Two-dimensional Keller-Segel model: Optimal critical mass and qualitative properties of the solutions. *Electronic Journal of Differential Equations (EJDE)*[*electronic only*], 2006:Paper–No, 2006.
- [5] Vincent Calvez, Lucilla Corrias, et al. The parabolic-parabolic Keller-Segel model in R2. Communications in Mathematical Sciences, 6(2):417–447, 2008.
- [6] Alina Chertock and Alexander Kurganov. A second-order positivity preserving central-upwind scheme for chemotaxis and haptotaxis models. *Numerische Mathematik*, 111(2):169, 2008.
- [7] Luis Neves De Almeida, Federica Bubba, Benoît Perthame, and Camille Pouchol. Energy and implicit discretization of the Fokker-Planck and Keller-Segel type equations. arXiv preprint arXiv:1803.10629, 2018.
- [8] Yasmin Dolak and Christian Schmeiser. The Keller-Segel model with logistic sensitivity function and small diffusivity. SIAM Journal on Applied Mathematics, 66(1):286–308, 2005.
- Bob Eisenberg. Ionic channels in biological membranes-electrostatic analysis of a natural nanotube. Contemporary Physics, 39(6):447–466, 1998.
- [10] Yekaterina Epshteyn. Upwind-difference potentials method for Patlak-Keller-Segel chemotaxis model. Journal of Scientific Computing, 53(3):689–713, 2012.
- [11] Yekaterina Epshteyn and Alexander Kurganov. New interior penalty discontinuous Galerkin methods for the Keller–Segel chemotaxis model. SIAM Journal on Numerical Analysis, 47(1):386–408, 2009.
- [12] Francis Filbet. A finite volume scheme for the Patlak-Keller-Segel chemotaxis model. Numerische Mathematik, 104(4):457-488, 2006.
- [13] Allen Flavell, Michael Machen, Bob Eisenberg, Julienne Kabre, Chun Liu, and Xiaofan Li. A conservative finite difference scheme for Poisson-Nernst-Planck equations. *Journal of Computational Electronics*, 13(1):235–249, 2014.
- [14] Herbert Gajewski and Konrad Gröger. On the basic equations for carrier transport in semiconductors. Journal of mathematical analysis and applications, 113(1):12–35, 1986.
- [15] Carl L Gardner, Wolfgang Nonner, and Robert S Eisenberg. Electrodiffusion model simulation of ionic channels: 1D simulations. Journal of Computational Electronics, 3(1):25–31, 2004.
- [16] Thomas Hillen and Kevin Painter. Global existence for a parabolic chemotaxis model with prevention of overcrowding. Advances in Applied Mathematics, 26(4):280–301, 2001.
- [17] Thomas Hillen and Kevin J Painter. A user's guide to PDE models for chemotaxis. Journal of mathematical biology, 58(1-2):183, 2009.
- [18] Tzyy-Leng Horng, Tai-Chia Lin, Chun Liu, and Bob Eisenberg. PNP equations with steric effects: a model of ion flow through channels. *The Journal of Physical Chemistry B*, 116(37):11422–11441, 2012.
- [19] Jingwei Hu and Xiaodong Huang. A fully discrete positivity-preserving and energy-dissipative finite difference scheme for Poisson-Nernst-Planck equations. Numer. Math., 145(1):77–115, 2020.
- [20] Fukeng Huang, Jie Shen, and Zhiguo Yang. A highly efficient and accurate new scalar auxiliary variable approach for gradient flows. SIAM Journal on Scientific Computing, 42(4):A2514–A2536, 2020.
- [21] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal., 29(1):1–17, 1998.



FIGURE 8. Example 6. Simulation of Keller-Segel equations with initial condition (5.10)

- [22] Evelyn F. Keller and Lee A. Segel. Initiation of slime mold aggregation viewed as an instability. Journal of Theoretical Biology, 26(3):399 – 415, 1970.
- [23] Hailiang Liu and Zhongming Wang. A free energy satisfying discontinuous Galerkin method for one-dimensional Poisson-Nernst-Planck systems. J. Comput. Phys., 328:413–437, 2017.
- [24] Jian-Guo Liu, Li Wang, and Zhennan Zhou. Positivity-preserving and asymptotic preserving method for 2D Keller-Segal equations. Math. Comp., 87(311):1165–1189, 2018.



FIGURE 9. Example 7. Simulation with $\chi_2 = 0.1$

- [25] Courtney L Lopreore, Thomas M Bartol, Jay S Coggan, Daniel X Keller, Gina E Sosinsky, Mark H Ellisman, and Terrence J Sejnowski. Computational modeling of three-dimensional electrodiffusion in biological systems: application to the node of ranvier. *Biophysical journal*, 95(6):2624–2635, 2008.
- [26] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser. Semiconductor equations. Springer-Verlag, Vienna, 1990.
- [27] Andreas Prohl and Markus Schmuck. Convergent discretizations for the Nernst-Planck-Poisson system. Numerische Mathematik, 111(4):591-630, 2009.
- [28] Filippo Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. Bulletin of Mathematical Sciences, 7(1):87–154, 2017.



FIGURE 10. Example 7. Simulation with $\chi_2 = 0.01$

- [29] Jie Shen and Jie Xu. Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows. SIAM Journal on Numerical Analysis, 56(5):2895–2912, 2018.
- [30] Jie Shen and Jie Xu. Unconditionally bound preserving and energy dissipative schemes for a class of Keller–Segel Equations. SIAM Journal on Numerical Analysis, 58(3):1674–1695, 2020.
- [31] Jie Shen and Jie Xu. Unconditionally positivity preserving and energy dissipative schemes for Poisson-Nernst-Planck equations. arXiv preprint arXiv:2007.06132, 2020.
- [32] Jie Shen, Jie Xu, and Jiang Yang. The scalar auxiliary variable (SAV) approach for gradient flows. J. Comput. Phys., 353:407–416, 2018.

- [33] Jie Shen, Jie Xu, and Jiang Yang. A new class of efficient and robust energy stable schemes for gradient flows. SIAM Review, 61(3):474–506, 2019.
- [34] Juan JL Velázquez. Point dynamics in a singular limit of the Keller–Segel model 1: Motion of the concentration regions. SIAM Journal on Applied Mathematics, 64(4):1198–1223, 2004.
- [35] Juan JL Velázquez. Point dynamics in a singular limit of the Keller–Segel model 2: Formation of the concentration regions. SIAM Journal on Applied Mathematics, 64(4):1224–1248, 2004.
- [36] Guanyu Zhou and Norikazu Saito. Finite volume methods for a Keller–Segel system: discrete energy, error estimates and numerical blow-up analysis. *Numerische Mathematik*, 135(1):265–311, 2017.