MULTISCALE MODEL. SIMUL.
Vol. 16, No. 4, pp. 1884–1915




# A HIGH ORDER STOCHASTIC ASYMPTOTIC PRESERVING SCHEME FOR CHEMOTAXIS KINETIC MODELS WITH RANDOM INPUTS*

SHI JIN†, HANQING LU‡, AND LORENZO PARESCHI§

**Abstract.** In this paper, we develop a stochastic Asymptotic-Preserving (sAP) scheme for the kinetic chemotaxis system with random inputs, which will converge to the modified Keller–Segel model with random inputs in the diffusive regime. Based on the generalized Polynomial Chaos (gPC) approach, we design a high order stochastic Galerkin method using implicit-explicit (IMEX) Runge–Kutta (RK) time discretization with a macroscopic penalty term. The new schemes improve the parabolic CFL condition to a hyperbolic type when the mean free path is small, which shows significant efficiency especially in uncertainty quantification (UQ) with multiscale problems. The sAP property will be shown asymptotically and verified numerically in several tests. Other numerical tests are conducted to explore the effect of the randomness in the kinetic system, with the goal of providing more intuition for the theoretic study of the chemotaxis models.

**Key words.** chemotaxis kinetic model, chemotaxis Keller–Segel model, diffusion limit, uncertainty quantification, asymptotic preserving, generalized polynomial chaos, stochastic Galerkin method, implicit-explicit Runge–Kutta methods

**AMS subject classifications.** 68Q25, 68R10, 68U05

**DOI.** 10.1137/17M1150840

## 1. Introduction.

Chemotaxis is the movement of an organism in response to a chemical stimulus (called a chemoattractant) approaching the regions of highest chemoattractant concentration. This process is critical to the early growth and subsequent development of the organism.

Mathematical study of this chemical system originates from the well-known (Patlak–)Keller–Segel model [34, 35, 36, 37, 44]. This model describes the drift-diffusion interactions between the cell density and chemoattractant concentration at a macroscopic level,

$$\partial_t \rho = \nabla \cdot (D\nabla\rho - \chi\rho\nabla s), \tag{1.1a}$$

$$\partial_t s = D_0 \Delta s + q(s,\rho), \tag{1.1b}$$

where $\rho(x,t) \geq 0$ is the cell density at position $x \in \mathbb{R}^n$ and time $t$; $s(x,t) \geq 0$ is the density of the chemoattractant; $D$ and $D_0$ are positive diffusive constants of the cells and the chemoattractant, respectively; and $\chi$ is the positive chemotactic sensitivity constant. In (1.1) the function $q(s,\rho)$ describes the interactions between the cell density and the chemoattractant, such as production and degradation. In

*Received by the editors October 16, 2017; accepted for publication (in revised form) July 23, 2018; published electronically November 20, 2018.

http://www.siam.org/journals/mms/16-4/M115084.html

**Funding:** This research was partially supported by NSF grants DMS-1522184 and DMS-1107291: RNMS KI-Net, by NSFC grant 91330203, and by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation.

†Department of Mathematics, University of Wisconsin, Madison, WI 53706, and School of Mathematical Science, Institute of Natural Sciences, MOE-LSEC and SHL-MAC, Shanghai Jiao Tong University, Shanghai 200240, China (sjin@wisc.edu).

‡Department of Mathematics, University of Wisconsin, Madison, WI 53706 (hlu57@wisc.edu).

§Department of Mathematics & Computer Science, University of Ferrara, Ferrara, 44121, Italy (lorenzo.pareschi@unife.it).

the literature, several modifications and studies of the Keller–Segel model have been conducted in recent years, e.g., [9, 13, 22, 23, 45, 46]. The one related to our study is the modified Keller–Segel model in [9],

$$\partial_t \rho = \nabla \cdot (D\nabla\rho - \chi\rho\nabla s), \tag{1.2a}$$

$$s = -\frac{1}{n\pi} \log|x| * \rho, \tag{1.2b}$$

where $n$ is the space dimension. Notice that in 2D, (1.1) and (1.2) are exactly the same if $q = 0$.

An important property of the Keller–Segel system is the blow up behavior, which depends on the dimension of the system and the initial mass [8, 20, 40, 49]. For the 2D Keller–Segel system (when (1.1) and (1.2) are equivalent), there exists a critical mass $M_c$ depending on the parameters of the system. When the initial mass $M < M_c$ (subcritical case), a global solution exists and presents a self-similar profile in long time; when the initial mass $M > M_c$ (supercritical case), the solution will blow up in finite time; and when the initial mass $M = M_c$ (critical case), the solution will blow up in infinite time. This property can be extended to 1D and 3D for the modified Keller–Segel system (1.2). The formula for the critical mass is given by

$$M_c = \frac{2n^2\pi D}{\chi}. \tag{1.3}$$

From another perspective, the chemotaxis can be described by a class of Boltzmann-type kinetic equations at a microscopic level. The kinetic description of the phase space cell density was first introduced by Alt [1, 2] via a stochastic interpretation of the "run" and "tumble" process of bacteria movements. Later, Othmer, Dunbar, and Alt formulated the following nondimensionalized chemotaxis kinetic system with parabolic scaling in [41]:

$$\varepsilon\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \frac{1}{\varepsilon}\int_V (T_\varepsilon f' - T_\varepsilon^* f)dv'. \tag{1.4}$$

Here $f(t, x, v)$ is the density function of cells at time $t \in \mathbb{R}^+$, position $x \in \mathbb{R}^n$ and moving with velocity $v \in V$, $V$ is a finite subset of $\mathbb{R}^n$. The small parameter $\varepsilon$ is the ratio of the mean running length between jumps to the typical observation length scale, and $f'$ is the abbreviation for $f(t, x, v')$. $T_\varepsilon = T_\varepsilon[s](t, x, v, v')$ with the property $T_\varepsilon^*[s](t, x, v, v') = T_\varepsilon[s](t, x, v', v)$ is the turning kernel operator depending on the density of chemoattractant $s(t, x)$, which also solves the Poisson equation (1.1b).

The relationship between the kinetic chemotaxis model (1.4) and the Keller–Segel model (1.1) was formally derived by Hillen and Othmer in [21, 42] using moment expansions. Then Chalub et al. gave a rigorous proof that the Keller–Segel system (1.2) (before blow up time in the supercritical case and for all time in the subcritical case) is the macroscopic limit (as $\varepsilon \to 0$) of the kinetic chemotaxis system (1.4) coupled with (1.2b) in three dimensions [11]. For a certain type of turning kernel $T_\varepsilon$ (the nonlocal model in section 2.1), [11] also proved the global existence of the solution to the kinetic system (1.4) for any initial conditions, which behaves completely differently from the Keller–Segel system. For other types of of turning kernel $T_\varepsilon$ (e.g., the local model in section 2.2), many questions are yet unsolved. Blow up may happen with supercritical initial mass, but the critical mass is different from the Keller–Segel equations [7]. The long time behavior of the subcritical case is yet unclear. Also, theoretic proof of the blow up in the 1D case is not available [47].

The microscopic kinetic model, with interesting properties and mysterious behaviors, makes it appealing to investigate the system numerically. Moreover, the global existence of the solution with nonlocal turning kernel could help us to understand the behavior of chemotaxis after Keller–Segel solutions blow up. One of the difficulties in solving the kinetic chemotaxis model, as with other multiscale kinetic equations, is the stiffness when $0 < \varepsilon \ll 1$. Classical algorithms require taking spatial and time steps of $O(\varepsilon)$, thus causing unaffordable computational cost. To overcome this difficulty, one has to design an *Asymptotic-Preserving* (AP) scheme, which discretizes the kinetic equations, with the mesh and the time step independent of $\varepsilon$, and preserves a consistent discretization of the limiting modified Keller–Segel equation as $\varepsilon \to 0$. The AP methods were first introduced in [25] and have been applied to a variety of multiscale kinetic equations. We refer the reader to [15, 16, 17, 26] for detailed reviews on AP schemes. In particular, AP schemes have been designed to solve 1D and 2D kinetic chemotaxis models in [10, 12], which are most relevant to our study.

The main issue we want to address in this paper concerns the uncertainties involved in the kinetic model due to modeling and experimental errors. For example, different turning kernels are proposed as operators that mimic the "run" and "tumble" process of cell movements and thus may contain uncertainties. Moreover, initial and boundary data, or other coefficients in the equations, could also be measured inaccurately. In such a system that behaves so sensitively to initial mass and turning kernel, only by quantifying the *intrinsic* uncertainties in the model could one gain a better understanding and a more reliable prediction of the chemotaxis from computational simulations, especially in the situation where many properties are not clarified by theoretic study.

The goal of this paper is to design a high order efficient numerical scheme such that uncertainty quantification (UQ) can be easily conducted. Only recently, studies in UQ have begun to develop for kinetic equations [24, 27, 28, 29, 32, 52, 14]. To deal with numerical difficulties for uncertainty and multiple scales simultaneously, the *stochastic Asymptotic-Preserving* (sAP) notion was first introduced in [32]. It requires that the discretization of a random kinetic equation, in the limit $\varepsilon \to 0$, becomes a good discretization of the limiting macroscopic random Keller–Segel equation. In the context of stochastic Galerkin (SG) methods, to be defined later, this requires that an SG approximation of the random kinetic equation, when $\varepsilon \to 0$, becomes an SG method for the limiting random Keller–Segel equation.

Since then, the *generalized Polynomial Chaos* (gPC)-based SG framework has been developed for a variety of kinetic equations [29, 32, 52, 14, 48]. In this paper, we conduct UQ under the same gPC-SG framework, which projects the uncertain kinetic equations into vectorized deterministic equations and thus allows us to extend the deterministic AP solver in [10]. We formally verify the sAP property by showing that the kinetic chemotaxis model with uncertainty after SG projection in a fully discrete setting, such as $\varepsilon \to 0$, automatically becomes a numerical discretization of the Keller–Segel equations with uncertainty after the SG projection. As realized in [32] and rigorously proved in [27, 39, 33], spectral accuracy is expected when using this gPC-SG method as long as the solution is sufficiently regular.

In addition, we improve the accuracy and efficiency of the numerical scheme by using the implicit-explicit (IMEX) Runge–Kutta (RK) methods (see [6, 5, 43] and the references therein) and the macroscopic penalization method. A similar approach was utilized in our previous work [30] for linear transport and radiative heat transfer equations with random inputs. In [30], we improved the parabolic CFL condition $\Delta t = O((\Delta x)^2)$ from [32] to a hyperbolic CFL condition $\Delta t = O(\Delta x)$, which allows significant savings in computational time.

The rest of the paper is organized as follows. In section 2, the kinetic models with random inputs of two different turning kernels are described, and the macroscopic limits of both models are formally derived. From section 3 to section 5, the numerical scheme for the kinetic chemotaxis equations is designed and the sAP properties are illustrated. In section 6, several numerical tests are presented to illustrate the accuracy and efficiency of our scheme. The sAP property is also verified numerically. Different properties, e.g., blow up, stationary solutions, etc., influenced by the introduced randomness of the local and nonlocal models, are explored for the chemotaxis system. The interactions between peaks involved with different sources of uncertainty are compared to show the dynamics. Finally, some conclusions are drawn in section 7.

**2. The kinetic description for chemotaxis.** The chemotaxis kinetic system with random inputs that we are going to study is (1.4) coupled with (1.2b) in 1D,

$$\varepsilon\frac{\partial f}{\partial t}+v\frac{\partial f}{\partial x}=\frac{1}{\varepsilon}\int_V (T_\varepsilon f'-T_\varepsilon^* f)dv', \tag{2.1a}$$

$$s=-\frac{1}{\pi}\log|x|*\rho,\quad \rho=\int_V f dv, \tag{2.1b}$$

where $x\in\Omega=[-x_{\max},x_{\max}]\subset\mathbb{R}$ and $v\in V=[-v_{\max},v_{\max}]\subset\mathbb{R}$.

The only difference from the classical deterministic case is that $f=f(t,x,v,z)$ and $s=s(t,x,z)$ are dependent on the random variable $z\in I_z\subset\mathbb{R}^d(d\geq 1)$, with compact support $I_z$, in order to account for random uncertainties.

Now we specify the turning kernel operator $T_\varepsilon$ in (2.1). Since the turning kernel $T_\varepsilon[s](t,x,z,v,v')$ measures the probability of a velocity jump of cells from $v$ to $v'$, it has the properties

$$\begin{aligned}&T_\varepsilon[s](t,x,z,v,v')\geq 0,\\ &T_\varepsilon[s](t,x,z,v,v')=F(z,v)+\varepsilon T_1+O(\varepsilon^2),\end{aligned} \tag{2.2}$$

where $F(z,v)$ is the equilibrium of velocity distribution and $T_1\geq 0$ characterizes the directional preference.

**2.1. The 1D nonlocal model.** Now we consider the nonlinear kernel introduced in [11] with uncertainty,

$$\begin{aligned}T_\varepsilon[s](t,x,z,v,v')=&\alpha_+(z)\psi(s(t,x,z),s(t,x+\varepsilon v,z))\\ &+\alpha_-(z)\psi(s(t,x,z),s(t,x-\varepsilon v',z)).\end{aligned} \tag{2.3}$$

The first term describes the cell movement to a new direction decided by the detection of the current environment and a probable new location, and the second term describes the influence of the past memory on the choice of the new moving direction. The values $\alpha_\pm(z)$ are experimental and characterize the intensity of these effects.

For simplicity, the past memory influence is neglected, namely $\alpha_-(z)=0$. Since $\alpha_+$ has an experimental nature, we introduce the randomness on $\alpha_+(z)>0$ with the probability density function $\lambda(z)$ for the random variable $z$ and take

$$\psi(s(t,x,z),s(t,x+\varepsilon v,z))=\bar{F}(v)+\delta^\varepsilon s(t,x,z,v), \tag{2.4}$$

where, omitting the dependence from $t$ and $z$ for notational simplicity, we define

$$\delta^\varepsilon s(x,v)=(s(x+\varepsilon v)-s(x))_+:=\begin{cases}s(x+\varepsilon v)-s(x) & \text{if } s(x+\varepsilon v)-s(x)>0,\\ 0 & \text{otherwise,}\end{cases} \tag{2.5}$$

and $\bar{F}(v)$ satisfies

$$\begin{cases} \int_V \bar{F}(v)dv = 1, \\ \bar{F}(v) = \bar{F}(|v|). \end{cases} \tag{2.6}$$

Notice that $\delta^\varepsilon s$ is an $O(\varepsilon)$ term which corresponds to $\varepsilon T_1$ in (2.2).

Then the kinetic system (2.1) becomes

$$\varepsilon\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} = \frac{\alpha_+(z)}{\varepsilon}\left[(\bar{F}(v) + \delta^\varepsilon s(v))\rho - \left(1 + \int_V \delta^\varepsilon s(v')dv'\right) f\right], \tag{2.7a}$$

$$s = -\frac{1}{\pi}\log|x| * \rho, \tag{2.7b}$$

together with positive initial conditions and reflecting boundary conditions for $f$. Reflecting boundary conditions for $s$ are imposed as follows:

$$f(0, x, z, v) = f^I(x, z, v) \geq 0, \tag{2.8a}$$

$$s(0, x, z) = s^I(x, z) \geq 0, \tag{2.8b}$$

$$f(t, \pm x_{\max}, z, v) = f(t, \pm x_{\max}, z, -v), \tag{2.8c}$$

$$\partial_x s|_{x=\pm x_{\max}} = 0. \tag{2.8d}$$

*Remark* 2.1. The global existence of the solution to (2.7) for fixed $z$ with any initial mass is proved in [11].

**2.2. The 1D local model.** For the local model, we consider the turning kernel introduced in [7] with uncertainty,

$$T_\varepsilon = T_\varepsilon[s](t, x, z, v, v') = \alpha_+(z)\left[\bar{F}(v) + \varepsilon(v \cdot \nabla s(x))_+\right], \tag{2.9}$$

where $\bar{F}$ is the equilibrium function satisfying (2.6) and $\alpha(z) > 0$ describes the desire of the cell to change to a favorable direction, which could come with uncertainty. Similarly as in section 2.1, we introduce the randomness on $\alpha_+(z) > 0$. Then the kinetic equation (2.1) in 1D is

$$\varepsilon\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} = \frac{\alpha_+}{\varepsilon}\left[(\bar{F}(v) + \varepsilon(v \cdot \nabla s)_+)\rho - (1 + c_1\varepsilon|\nabla s|)f\right], \tag{2.10a}$$

$$s = -\frac{1}{\pi}\log|x| * \rho, \tag{2.10b}$$

with $c_1 = \int_V (v \cdot \nabla s/|\nabla s|)_+ dv = \frac{1}{2}\int_V |v|dv$. The same initial and boundary conditions in (2.8) are applied.

**2.3. The macroscopic limits.** The nonlocal kinetic model (2.7) and the local limit (2.10) give the same asymptotic limit when $\varepsilon \to 0$. Inserting the Hilbert expansion into (2.7a) and (2.10a) and collecting the same order terms, one can derive the classical modified Keller–Segel system for $\rho$ as $\varepsilon \to 0$,

$$\partial_t \rho = \partial_x\left(\frac{D}{\alpha_+}\partial_x\rho - \chi\rho\partial_x s\right), \tag{2.11a}$$

$$s = -\frac{1}{\pi}\log|x| * \rho, \tag{2.11b}$$

$$\partial_x \rho|_{x=\pm x_{\max}} = 0, \tag{2.11c}$$

$$\partial_x s|_{x=\pm x_{\max}} = 0, \tag{2.11d}$$

where

$$D = \int_V |v|^2 \bar{F}(v) dv, \quad \chi = \frac{1}{2} \int_V |v|^2 dv. \tag{2.12}$$

We refer the reader to [11] for the details.

**2.4. The critical mass with random inputs.** To derive the critical mass for system (2.11), we show, following [9], that the second momentum (with respect to $x$) of $\rho$ cannot remain positive for all time.

We use

$$\partial_x s = \partial_x \left( -\frac{1}{\pi} \log |x| * \rho \right) = -\frac{1}{\pi} \int_\Omega \frac{1}{x-y} \rho(t,y,z) dy = -\mathcal{H}\rho,$$

where $\mathcal{H}$ denotes the Hilbert transform [19], and the improper integral has to be understood in the principal value sense. Then

$$\begin{aligned}
\frac{d}{dt} \int_\Omega \frac{1}{2} |x|^2 \rho(t,x) dx &= \int_\Omega \frac{1}{2} |x|^2 \frac{\partial \rho}{\partial t}(t,x) dx \\
&= \int_\Omega \frac{1}{2} |x|^2 \partial_x \left( \frac{D}{\alpha_+(z)} \partial_x \rho - \chi \rho \partial_x s \right) dx \\
&= - \int_\Omega x \left( \frac{D}{\alpha_+(z)} \partial_x \rho - \chi \rho \partial_x s \right) dx \\
&= - \frac{D}{\alpha_+(z)} [x_{\max} \rho(x_{\max}) + x_{\max} \rho(-x_{\max})] + \frac{D}{\alpha_+(z)} M \\
&\quad - \frac{\chi}{\pi} \int_\Omega \rho(x) \lim_{\delta \to 0} \int_{|x-y|>\delta} \frac{x}{x-y} \rho(y) dy dx \\
&= - \frac{D}{\alpha_+(z)} x_{\max} [\rho(x_{\max}) + \rho(-x_{\max})] + \frac{D}{\alpha_+(z)} M \\
&\quad - \frac{\chi}{2\pi} \lim_{\delta \to 0} \int_\Omega \int_{|x-y|>\delta} \rho(x) \rho(y) dx dy \\
&= - \frac{D}{\alpha_+(z)} x_{\max} [\rho(x_{\max}) + \rho(-x_{\max})] \\
&\quad - \frac{\chi}{2\pi} M^2 \left( 1 - \frac{M_c(z)}{M} \right),
\end{aligned} \tag{2.13}$$

where

$$M_c(z) = \frac{2\pi D}{\chi \alpha_+(z)}, \tag{2.14}$$

and in (2.13), to simplify notation, we omit the dependence of $\rho$ from $t$ and $z$ unless it is necessary.

Here, to avoid unnecessary difficulties, we assume that the initial data are independent of $z$ and we use the conservation of mass; i.e., $M = \int_\Omega \rho dx$ is a constant independent of $z$. Therefore, uncertainties are due only to the value $\alpha_+(z)$ in the interaction kernel.

When $M > M_c(z)$, $\frac{d}{dt} \int_\Omega \frac{1}{2} |x|^2 \rho(t,x,z) dx \le -c < 0$, where $c$ is a positive constant. To preserve the positivity of this second moment (with respect to $x$), some singularity has to occur so that the above computation will not hold at a certain

time. The singularity is rigorously analyzed in [18, 4], and $\partial_x s$ is unbounded in this case. Thus blow up occurs.

When $M < M_c(z)$, the second moment (with respect to $x$) is locally controlled, and global existence of weak solution can be obtained [9].

*Remark* 2.2. When $n \geq 2$, the computation is similar, and the general formula for $M_c(z)$ is

$$M_c(z) = \frac{2n^2\pi D}{\chi\alpha_+(z)}.$$

In practice, one is more interested in the behavior of $\mathbb{E}[\rho(t,x,z)]$, the expected value of $\rho(t,x,z)$. We have the following theorem analyzing the influence of initial mass on $\mathbb{E}[\rho(t,x,z)]$.

THEOREM 2.1. *Suppose that the total mass $M$ is independent of $z$. Denote $\bar{M}_c$ as the critical mass for $\mathbb{E}[\rho(t,x,z)]$; i.e., when $M > \bar{M}_c$, $\mathbb{E}[\rho(t,x,z)]$ will blow up, and when $M < \bar{M}_c$, $\mathbb{E}[\rho(t,x,z)]$ will be bounded for all time. Then we have*

$$\bar{M}_c = \mathbb{E}[M_c(z)]. \tag{2.15}$$

*Proof.* Following the computations in (2.13) and omitting the dependence of $\rho$ from $t$ and $z$ unless it is necessary, we show that

$$\begin{aligned}
\frac{d}{dt}\int_\Omega \frac{1}{2}|x|^2\mathbb{E}[\rho(x,z)]dx &= \int_\Omega\int_{I_z}\frac{1}{2}|x|^2\frac{\partial\rho(x,z)}{\partial t}\lambda(z)dzdx \\
&= \int_\Omega\int_{I_z}\frac{1}{2}|x|^2\partial_x\left(\frac{D}{\alpha_+(z)}\partial_x\rho - \chi\rho\partial_x s\right)\lambda(z)dzdx \\
&= \int_{I_z}\left[\int_\Omega\frac{1}{2}|x|^2\partial_x\left(\frac{D}{\alpha_+(z)}\partial_x\rho - \chi\rho\partial_x s\right)dx\right]\lambda(z)dz \\
&= \int_{I_z}\left[-\frac{D}{\alpha_+(z)}[x_{\max}\rho(x_{\max}) + x_{\max}\rho(-x_{\max})]\right. \\
&\quad \left. -\frac{\chi}{2\pi}M^2\left(1-\frac{M_c(z)}{M}\right)\right]\lambda(z)dz \\
&= -\int_{I_z}\frac{D}{\alpha_+(z)}x_{\max}[\rho(x_{\max}) + \rho(-x_{\max})]\lambda(z)dz \\
&\quad -\frac{\chi}{2\pi}M^2\left(1-\frac{\mathbb{E}[M_c(z)]}{M}\right) \\
&\leq -\frac{\chi}{2\pi}M^2\left(1-\frac{\mathbb{E}[M_c(z)]}{M}\right).
\end{aligned} \tag{2.16}$$

Thus, $\bar{M}_c = \mathbb{E}[M_c(z)]$ is the critical mass for $\mathbb{E}[\rho(t,x,z)]$. ∎

*Remark* 2.3. The same conclusion holds for $n \geq 2$.

**3. The even-odd decomposition.** In this section, we apply the even-odd decomposition to reformulate the problem following the same procedure as in [10] for the deterministic kinetic model for chemotaxis.

**3.1. The 1D nonlocal model.** For $v > 0$, (2.7a) can be split into the two equations

$$\varepsilon\frac{\partial f^+}{\partial t} + v\frac{\partial f^+}{\partial x} = \frac{\alpha_+(z)}{\varepsilon}\left[(\bar{F}(v) + \delta^\varepsilon s(v))\rho - \left(1 + \int_V \delta^\varepsilon s(v')dv'\right) f^+\right], \tag{3.1a}$$

$$\varepsilon\frac{\partial f^-}{\partial t} - v\frac{\partial f^-}{\partial x} = \frac{\alpha_+(z)}{\varepsilon}\left[(\bar{F}(-v) + \delta^\varepsilon s(-v))\rho - \left(1 + \int_V \delta^\varepsilon s(v')dv'\right) f^-\right], \tag{3.1b}$$

where $f^+ = f^+(t,x,z,v) = f(t,x,z,v)$ and $f^- = f^-(t,x,z,v) = f(t,x,z,-v)$.

Now denote the even and odd parities

$$r(t,x,z,v) = \mathcal{R}[f] = \frac{1}{2}(f^+(t,x,z,v) + f^-(t,x,z,v)), \tag{3.2a}$$

$$j(t,x,z,v) = \mathcal{J}[f] = \frac{1}{2\varepsilon}(f^+(t,x,z,v) - f^-(t,x,z,v)). \tag{3.2b}$$

Then (3.1) becomes

$$\partial_t r + v\partial_x j = \frac{\alpha_+}{\varepsilon^2}[(\bar{F}(v) + \mathcal{R}[\delta^\varepsilon s])\rho - (1 + \langle\delta^\varepsilon s\rangle)r], \tag{3.3a}$$

$$\partial_t j + \frac{1}{\varepsilon^2}v\partial_x r = \frac{\alpha_+}{\varepsilon^2}(\mathcal{J}[\delta^\varepsilon s]\rho - (1 + \langle\delta^\varepsilon s\rangle)j), \tag{3.3b}$$

where

$$\langle\delta^\varepsilon s\rangle = \int_V \delta^\varepsilon s(x,v')dv', \tag{3.4a}$$

$$\rho = \int_V f dv = 2\int_{V^+} r dv,\ V^+ = \{v \in V | v \geq 0\}. \tag{3.4b}$$

Notice that, when $\varepsilon \to 0$, (3.3) yields

$$r = \frac{\bar{F}(v) + \mathcal{R}[\delta^\varepsilon s]}{1 + \langle\delta^\varepsilon s\rangle}\rho = \rho\bar{F}(v) + O(\varepsilon), \tag{3.5a}$$

$$j = \frac{\mathcal{J}[\delta^\varepsilon s]\rho - v\frac{\partial_x r}{\alpha_+}}{1 + \langle\delta^\varepsilon s\rangle} = v\left(\frac{1}{2}\partial_x s\rho - \frac{\partial_x r}{\alpha_+}\right) + O(\varepsilon). \tag{3.5b}$$

Substituting (3.5) into (3.3a) and integrating over $V^+$, one gets the same limiting Keller–Segel equations with random inputs as in (2.11).

**3.2. The 1D local model.** For the 1D local model, one can follow the same even-odd decomposition and obtain

$$\partial_t r + v\partial_x j = \frac{\alpha_+}{\varepsilon}\left[\left(\bar{F}(v) + \frac{\varepsilon}{2}|v\partial_x s|\right)\rho - (1 + c_1\varepsilon|\partial_x s|)r\right], \tag{3.6a}$$

$$\partial_t j + \frac{1}{\varepsilon^2}v\partial_x r = \frac{\alpha_+}{\varepsilon^2}\left[\frac{1}{2}v\partial_x s\rho - (1 + c_1\varepsilon|\partial_x s|)j\right]. \tag{3.6b}$$

The remaining work is the same as that in section 3.1.

**4. The gPC-SG formulation.** Now we deal with the random inputs using the gPC expansion via an orthogonal polynomial series to approximate the solution. That

is, for random variable $z \in \mathbb{R}^d$, one seeks

$$r(t,x,z,v) \approx r_N(t,x,z,v) = \sum_{k=1}^{K} \hat{r}_k(t,x,v)\Phi_k(z), \tag{4.1a}$$

$$j(t,x,z,v) \approx j_N(t,x,z,v) = \sum_{k=1}^{K} \hat{j}_k(t,x,v)\Phi_k(z), \tag{4.1b}$$

where

$$\left\{\Phi_k(z), 1 \le k \le K, K = \binom{d+N}{d}\right\}$$

are from $\mathbb{P}_N^d$, the $d$-variate orthogonal polynomials of degree up to $N \ge 1$, and are orthonormal:

$$\int_{I_z} \Phi_i(z)\Phi_j(z)\lambda(z)dz = \delta_{ij},\ 1 \le i,j \le K = \dim(\mathbb{P}_N^d). \tag{4.2}$$

Here $\delta_{i,j}$ is the Kronecker delta function (see [51]).

Now insert the approximation (4.1) into the governing equation (3.3) and enforce the residue to be orthogonal to the polynomial space spanned by $\{\Phi_1, \dots, \Phi_K\}$. Thus, we obtain a set of vector deterministic equations for $\hat{\mathbf{r}} = (\hat{r}_1, \dots, \hat{r}_K)^T$, $\hat{\mathbf{j}} = (\hat{j}_1, \dots, \hat{j}_K)^T$, and $\hat{\mathbf{s}} = (\hat{s}_1, \dots, \hat{s}_K)^T$:

$$\partial_t \hat{\mathbf{r}} + v\partial_x \hat{\mathbf{j}} = \frac{1}{\varepsilon^2}[\bar{F}(v)\mathbf{M}\hat{\boldsymbol{\rho}} + \mathbf{B}\hat{\boldsymbol{\rho}} - \mathbf{M}\hat{\mathbf{r}} - \mathbf{C}\hat{\mathbf{r}}], \tag{4.3a}$$

$$\partial_t \hat{\mathbf{j}} + \frac{1}{\varepsilon^2} v\partial_x \hat{\mathbf{r}} = \frac{1}{\varepsilon^2}(\mathbf{E}\hat{\boldsymbol{\rho}} - \mathbf{M}\hat{\mathbf{j}} - \mathbf{C}\hat{\mathbf{j}}), \tag{4.3b}$$

$$\hat{\mathbf{s}} = -\frac{1}{\pi}\log|x| * \hat{\boldsymbol{\rho}}, \tag{4.3c}$$

where

$$\hat{\boldsymbol{\rho}}(t,x) = \langle \hat{\mathbf{r}} \rangle = 2\int_{V^+} \hat{\mathbf{r}} dv, \tag{4.4}$$

and $\mathbf{M} = (m_{ij})_{1\le i,j\le K}$, $\mathbf{B}(\delta^\varepsilon s_N) = (b_{ij}(x,v))_{1\le i,j\le K}$, $\mathbf{C}(\langle \delta^\varepsilon s_N \rangle) = (c_{ij}(x))_{1\le i,j\le K}$, and $\mathbf{E}(\delta^\varepsilon s_N) = (e_{ij}(x,v))_{1\le i,j\le K}$ are $K \times K$ symmetric matrices with entries, respectively,

$$m_{ij} = \int_{I_z} \alpha_+(z)\Phi_i(z)\Phi_j(z)\lambda(z)dz, \tag{4.5a}$$

$$b_{ij}(x,v) = \int_{I_z} \alpha_+(z)\mathcal{R}[\delta^\varepsilon s_N]\Phi_i(z)\Phi_j(z)\lambda(z)dz, \tag{4.5b}$$

$$c_{ij}(x) = \int_{I_z} \alpha_+(z)\langle \delta^\varepsilon s_N \rangle \Phi_i(z)\Phi_j(z)\lambda(z)dz, \tag{4.5c}$$

$$e_{ij}(x,v) = \int_{I_z} \alpha_+(z)\mathcal{J}[\delta^\varepsilon s_N]\Phi_i(z)\Phi_j(z)\lambda(z)dz. \tag{4.5d}$$

As $\varepsilon \to 0^+$ in (4.3), since $\langle \delta^\varepsilon s_N \rangle = O(\varepsilon)$ and the matrices $\mathbf{M}$ and $\mathbf{C}$ are symmetric positive definite and thus invertible, we have

$$\hat{\mathbf{r}} = (\mathbf{M}+\mathbf{C})^{-1}(\bar{F}(v)\mathbf{M} + \mathbf{B})\hat{\boldsymbol{\rho}} = \bar{F}(v)\hat{\boldsymbol{\rho}} + O(\varepsilon), \tag{4.6a}$$

$$\hat{\mathbf{j}} = (\mathbf{M}+\mathbf{C})^{-1}(\mathbf{E}\hat{\boldsymbol{\rho}} - v\partial_x \hat{\mathbf{r}}) = \mathbf{M}^{-1}\mathbf{E}\hat{\boldsymbol{\rho}} - v\mathbf{M}^{-1}\partial_x \hat{\boldsymbol{r}} + O(\varepsilon). \tag{4.6b}$$

Plugging (4.6) into (4.3a) and integrating over $V^+$, one obtains

$$\partial_t \hat{\boldsymbol{\rho}} = \partial_x \left( D\mathbf{M}^{-1}\partial_x \hat{\boldsymbol{\rho}} - \chi \mathbf{G}\hat{\boldsymbol{\rho}} \right), \tag{4.7}$$

where $\mathbf{G} = \frac{1}{\chi}\mathbf{M}^{-1}\langle \mathbf{E} \rangle$.

*Remark* 4.1. If one applies the gPC-SG formulation to the limiting Keller–Segel equation (2.11) directly, one gets

$$\partial_t \tilde{\boldsymbol{\rho}} = \partial_x \left( D\tilde{\mathbf{M}}\partial_x \tilde{\boldsymbol{\rho}} - \chi \tilde{\mathbf{G}}\tilde{\boldsymbol{\rho}} \right), \tag{4.8}$$

where $\tilde{\mathbf{M}} = (\tilde{m}_{ij})_{1\le i,j\le K}$ and $\tilde{\mathbf{G}} = (\tilde{g}_{ij})_{1\le i,j\le K}$ are $K \times K$ symmetric matrices with entries

$$\tilde{m}_{ij} = \int_{I_z} \frac{1}{\alpha_+(z)} \Phi_i(z)\Phi_j(z)\lambda(z)dz, \tag{4.9a}$$

$$\tilde{g}_{ij} = \int_{I_z} (\partial_x s_N)\Phi_i(z)\Phi_j(z)\lambda(z)dz. \tag{4.9b}$$

Although $\tilde{\mathbf{M}}$ is different from $\mathbf{M}^{-1}$, one can show that $\tilde{\mathbf{M}}\partial_x\tilde{\boldsymbol{\rho}}$ and $\mathbf{M}^{-1}\partial_x\hat{\boldsymbol{\rho}}$ are spectrally close to each other. The same property holds between $\tilde{\mathbf{G}}\tilde{\boldsymbol{\rho}}$ and $\mathbf{G}\hat{\boldsymbol{\rho}}$. To understand this, let us consider the simple identities

$$\begin{array}{c} u = a(z)v, \\ \Updownarrow \\ a(z)^{-1}u = v, \end{array} \tag{4.10}$$

where $u$ and $v$ are functions of $z$ and other independent variables, such as $x$, $v$, and $t$, in our setting. If we apply a gPC-SG projection to the above identities, using vector notation we get two different formulations:

$$\hat{\mathbf{U}} = A\hat{\mathbf{V}} \iff A^{-1}\hat{\mathbf{U}} = \hat{\mathbf{V}}, \tag{4.11a}$$

$$\tilde{A}\tilde{\mathbf{U}} = \tilde{\mathbf{V}}. \tag{4.11b}$$

Note that both (4.11a) and (4.11b) are spectrally close to (4.10), and therefore the solutions of (4.11a) and (4.11b) are spectrally close to each other. Thus $\hat{\mathbf{V}} - \tilde{\mathbf{V}}$ is spectrally small and so $A^{-1}\hat{\mathbf{U}} - \tilde{A}\tilde{\mathbf{U}}$ is spectrally small (although $A^{-1}$ and $\tilde{A}$ are not spectrally close).

**5. An efficient sAP scheme based on an IMEX-RK method.** One can apply the relaxation method as in [10] to the projected system (4.3), which falls into the sAP framework proposed in [32]. However, the method suffers from the parabolic CFL condition $\Delta t = O((\Delta x)^2)$.

Here we propose an efficient sAP scheme using the idea from [6] to get rid of the parabolic CFL condition. By adding and subtracting the term $\mu\bar{F}(v)\partial_x(D\tilde{\mathbf{M}}\partial_x\hat{\boldsymbol{\rho}} - \chi\tilde{\mathbf{G}}\hat{\boldsymbol{\rho}})$ in (4.3a) and the term $\phi v\partial_x\hat{\mathbf{r}}$ in (4.3b), we reformulate the problem into the

equivalent form

$$\partial_t \hat{\mathbf{r}} = - v\partial_x \hat{\mathbf{j}} - \mu \bar{F}(v)\partial_x (D\tilde{\mathbf{M}}\partial_x \hat{\boldsymbol{\rho}} - \chi \tilde{\mathbf{G}}\hat{\boldsymbol{\rho}}) + \frac{1}{\varepsilon^2}\left(\bar{F}(v)\mathbf{M}\hat{\boldsymbol{\rho}} + \mathbf{B}\hat{\boldsymbol{\rho}} - \mathbf{M}\hat{\mathbf{r}} - \mathbf{C}\hat{\mathbf{r}}\right) + \mu \bar{F}(v)\partial_x (D\tilde{\mathbf{M}}\partial_x \hat{\boldsymbol{\rho}} - \chi \tilde{\mathbf{G}}\hat{\boldsymbol{\rho}}) \tag{5.1a}$$

$$= f_1(\hat{\mathbf{r}}, \hat{\mathbf{j}}) + f_2(\hat{\mathbf{r}}, \hat{\mathbf{s}}), \tag{5.1b}$$

$$\partial_t \hat{\mathbf{j}} = - \phi v \partial_x \hat{\mathbf{r}} - \frac{1}{\varepsilon^2}\left[(1-\varepsilon^2\phi) v\partial_x \hat{\mathbf{r}} - \mathbf{E}\hat{\boldsymbol{\rho}} + \mathbf{M}\hat{\mathbf{j}} + \mathbf{C}\hat{\mathbf{j}}\right] = g_1(\hat{\mathbf{r}}) + g_2(\hat{\mathbf{r}}, \hat{\mathbf{j}}), \tag{5.1c}$$

$$\hat{\mathbf{s}} = - \frac{1}{\pi}\log|x| * \hat{\boldsymbol{\rho}} = h(\hat{\mathbf{r}}), \tag{5.1d}$$

where $\mathbf{M}, \tilde{\mathbf{M}}, \mathbf{B}, \mathbf{C}, \mathbf{E}$, and $\tilde{\mathbf{G}}$ are as defined in (4.5) and (4.9), and since $\hat{\boldsymbol{\rho}} = \langle \hat{\mathbf{r}} \rangle$, we have

$$f_1(\hat{\mathbf{r}}, \hat{\mathbf{j}}) = -v\partial_x \hat{\mathbf{j}} - \mu \bar{F}(v)\partial_x (D\tilde{\mathbf{M}}\partial_x \hat{\boldsymbol{\rho}} - \chi \tilde{\mathbf{G}}\hat{\boldsymbol{\rho}}), \tag{5.2a}$$

$$f_2(\hat{\mathbf{r}}, \hat{\mathbf{s}}) = \frac{1}{\varepsilon^2}\left(\bar{F}(v)\mathbf{M}\hat{\boldsymbol{\rho}} + \mathbf{B}\hat{\boldsymbol{\rho}} - \mathbf{M}\hat{\mathbf{r}} - \mathbf{C}\hat{\mathbf{r}}\right) + \mu \bar{F}(v)\partial_x (D\tilde{\mathbf{M}}\partial_x \hat{\boldsymbol{\rho}} - \chi \tilde{\mathbf{G}}\hat{\boldsymbol{\rho}}), \tag{5.2b}$$

$$g_1(\hat{\mathbf{r}}) = -\phi v \partial_x \hat{\mathbf{r}}, \tag{5.2c}$$

$$g_2(\hat{\mathbf{r}}, \hat{\mathbf{j}}) = -\frac{1}{\varepsilon^2}\left[(1-\varepsilon^2\phi) v\partial_x \hat{\mathbf{r}} - \mathbf{E}\hat{\boldsymbol{\rho}} + \mathbf{M}\hat{\mathbf{j}} + \mathbf{C}\hat{\mathbf{j}}\right]. \tag{5.2d}$$

Here we choose $\mu = \mu(\varepsilon)$ such that

$$\lim_{\varepsilon \to 0} \mu = 1, \qquad \mu = 0 \quad \text{if} \quad \varepsilon = O(1), \tag{5.3}$$

and choose $\phi = \phi(\varepsilon)$ such that

$$0 \le \phi \le \frac{1}{\varepsilon^2}. \tag{5.4}$$

The restriction on $\phi$ guarantees the positivity of $\phi(\varepsilon)$ and $(1-\varepsilon^2\phi(\varepsilon))$ so that the problem remains well-posed uniformly in $\varepsilon$. We make the same simple choice of $\phi$ as in [31]:

$$\phi(\varepsilon) = \min\left\{1, \frac{1}{\varepsilon^2}\right\}. \tag{5.5}$$

Now we apply an IMEX-RK scheme to system (5.1), where $(f_1, g_1)^T$ is evaluated explicitly and $(f_2, g_2)^T$ implicitly, and then we obtain

$$\hat{\mathbf{r}}^{n+1} = \hat{\mathbf{r}}^n + \Delta t \sum_{k=1}^{s} \tilde{b}_k f_1(\hat{\mathbf{R}}^k, \hat{\mathbf{J}}^k) + \Delta t \sum_{k=1}^{s} b_k f_2(\hat{\mathbf{R}}^k, \hat{\mathbf{S}}^k), \tag{5.6a}$$

$$\hat{\mathbf{j}}^{n+1} = \hat{\mathbf{j}}^n + \Delta t \sum_{k=1}^{s} \tilde{b}_k g_1(\hat{\mathbf{R}}^k) + \Delta t \sum_{k=1}^{s} b_k g_2(\hat{\mathbf{R}}^k, \hat{\mathbf{J}}^k), \tag{5.6b}$$

$$\hat{\mathbf{s}}^{n+1} = -\frac{1}{\pi}\log|x| * \hat{\boldsymbol{\rho}}^{n+1}, \tag{5.6c}$$

where the internal stages for $k = 1, \dots, s$ are defined as

$$\hat{\mathbf{R}}^k = \hat{\mathbf{r}}^n + \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} f_1(\hat{\mathbf{R}}^l, \hat{\mathbf{J}}^l) + \Delta t \sum_{l=1}^{k} a_{kl} f_2(\hat{\mathbf{R}}^l, \hat{\mathbf{S}}^l), \tag{5.7a}$$

$$\hat{\mathbf{J}}^k = \hat{\mathbf{j}}^n + \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} g_1(\hat{\mathbf{R}}^l) + \Delta t \sum_{l=1}^{k} a_{kl} g_2(\hat{\mathbf{R}}^l, \hat{\mathbf{J}}^l), \tag{5.7b}$$

$$\hat{\mathbf{S}}^k = -\frac{1}{\pi} \log|x| * \hat{\mathbf{P}}^k, \tag{5.7c}$$

and $\hat{\mathbf{P}}^k = \langle \hat{\mathbf{R}}^k \rangle$. Clearly, also the internal stages $\hat{\mathbf{R}}^k$ and $\hat{\mathbf{J}}^k$ depend on $(t, x, v)$, whereas $\hat{\mathbf{S}}^k$ depends on $(t, x)$ only.

It is obvious that the scheme is characterized by the $s \times s$ matrices

$$\tilde{A} = (\tilde{a}_{ij}), \quad A = (a_{ij}) \tag{5.8}$$

and the vectors $\tilde{b}, b \in \mathbb{R}^s$, which can be represented by a double tableau in the usual Butcher notation,

$$\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}^T \end{array}, \qquad \begin{array}{c|c} c & A \\ \hline & b^T \end{array}.$$

The coefficients $\tilde{c}$ and $c$ depend on the explicit part of the scheme,

$$\tilde{c}_i = \sum_{j=1}^{i-1} \tilde{a}_{ij}, \quad c_i = \sum_{j=1}^{i} a_{ij}. \tag{5.9}$$

In the literature, there are two main types of IMEX-RK schemes characterized by the structure of the matrix $A$. We are interested in the IMEX-RK method of type $A$ (see [6]), where the matrix $A$ is invertible.

As an example, we report the SSP(3,3,2) scheme, which is a second order IMEX scheme we use in section 6:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 1/2 & 1/2 & 0 \\ \hline & 1/3 & 1/3 & 1/3 \end{array}, \qquad \begin{array}{c|ccc} 1/4 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/4 & 0 \\ 1 & 1/3 & 1/3 & 1/3 \\ \hline & 1/3 & 1/3 & 1/3 \end{array}. \tag{5.10}$$

Note that the above IMEX scheme is *globally stiffly accurate* [6]; namely, the implicit tableau is stiffly accurate $a_{sj} = b_j$, $j = 1, \dots, s$, and the explicit tableau similarly satisfies $\tilde{a}_{sj} = \tilde{b}_j$, $j = 1, \dots, s-1$. As a consequence from (5.6), we have $\hat{\mathbf{r}}^{n+1} = \hat{\mathbf{R}}^s$, $\hat{\mathbf{j}}^{n+1} = \hat{\mathbf{J}}^s$, and $\hat{\mathbf{s}}^{n+1} = \hat{\mathbf{S}}^s$. As shown in [6], this property is essential in order to achieve the AP property in the presence of a penalization term. Therefore, in what follows, we will assume that our IMEX-RK scheme is globally stiffly accurate.

To obtain $\hat{\mathbf{R}}^k$ in each internal stage of (5.7), one needs $\hat{\mathbf{P}}^k$ and $\hat{\mathbf{S}}^k$ in the implicit part $f_2(\hat{\mathbf{R}}^k, \hat{\mathbf{S}}^k)$. These quantities can be obtained explicitly by the following procedure.

Suppose one has computed $\hat{\mathbf{R}}^l$ and $\hat{\mathbf{S}}^l$ for $l = 1, \dots, k-1$; then according to

(5.7a),

$$
\begin{aligned}
\hat{\mathbf{R}}^k =& \hat{\mathbf{r}}^n + \Delta t \sum_{l=1}^{k-1} \left( \tilde{a}_{kl} f_1(\hat{\mathbf{R}}^l, \hat{\mathbf{J}}^l) + a_{kl} f_2(\hat{\mathbf{R}}^l, \hat{\mathbf{S}}^l) \right) \\
&+ \Delta t a_{kk} \left[ \frac{1}{\varepsilon^2} (\bar{F}(v) \mathbf{M} \hat{\mathbf{P}}^k + \mathbf{B}^k \hat{\mathbf{P}}^k - \mathbf{M} \hat{\mathbf{R}}^k - \mathbf{C}^k \hat{\mathbf{R}}^k) \right. \\
&\left. + \mu \bar{F}(v) \partial_x (D \tilde{\mathbf{M}} \partial_x \hat{\mathbf{P}}^k - \chi \tilde{\mathbf{G}}^k \hat{\mathbf{P}}^k) \right] \\
=& \overline{\hat{\mathbf{R}}}^{k-1} + \Delta t a_{kk} \left[ \frac{1}{\varepsilon^2} (\bar{F}(v) \mathbf{M} \hat{\mathbf{P}}^k + \mathbf{B}^k \hat{\mathbf{P}}^k - \mathbf{M} \hat{\mathbf{R}}^k - \mathbf{C}^k \hat{\mathbf{R}}^k) \right. \\
&\left. + \mu \bar{F}(v) \partial_x (D \tilde{\mathbf{M}} \partial_x \hat{\mathbf{P}}^k - \chi \tilde{\mathbf{G}}^k \hat{\mathbf{P}}^k) \right].
\end{aligned} \tag{5.11}
$$

Here $\overline{\hat{\mathbf{R}}}^{k-1}$ represents all contributions in (5.11) from the first $k-1$ stages. Now one takes $\langle \cdot \rangle$ on both sides of (5.11) so that $[\bar{F}(v)\mathbf{M}\hat{\mathbf{P}}^k + \mathbf{B}^k\hat{\mathbf{P}}^k - \mathbf{M}\hat{\mathbf{R}}^k - \mathbf{C}^k\hat{\mathbf{R}}^k]$ is cancelled out. Now $\hat{\mathbf{P}}^k$ is obtained from the following diffusion equation in an implicit form:

$$
\hat{\mathbf{P}}^k - \Delta t a_{kk} \mu \partial_x \left( D \tilde{\mathbf{M}} \partial_x \hat{\mathbf{P}}^k - \chi \tilde{\mathbf{G}}^k \hat{\mathbf{P}}^k \right) = \left\langle \overline{\hat{\mathbf{R}}}^{k-1} \right\rangle. \tag{5.12}
$$

In order to avoid iterative solvers for nonlinear equations, in [6] an explicit scheme is obtained simply by computing the new value of $\hat{\mathbf{P}}^k$, explicitly setting $\mu = 0$ in (5.12), and then plugging the computed value back into (5.11) in order to compute $\hat{\mathbf{R}}^k$. Here, we use a different approach, based on a linearization of the implicit scheme obtained by replacing $\tilde{\mathbf{G}}^k$ by $\tilde{\mathbf{G}}^{k-1}$ in (5.12) to get

$$
\hat{\mathbf{P}}^k - \Delta t a_{kk} \mu \partial_x (D \tilde{\mathbf{M}} \partial_x \hat{\mathbf{P}}^k - \chi \tilde{\mathbf{G}}^{k-1} \hat{\mathbf{P}}^k) = \left\langle \overline{\hat{\mathbf{R}}}^{k-1} \right\rangle. \tag{5.13}
$$

This permits us to keep the implicit structure of (5.13) by avoiding iterative solvers. Although a careful stability analysis is missing, in our numerical tests this technique has shown better stability properties compared to the technique proposed in [6]. Clearly, both strategies do not affect the overall accuracy of the scheme and can be seen as predictor-corrector methods.

**5.1. The space discretization.** Second order accuracy in space is obtained using an upwind TVD scheme (with minmod slope limiter [38]) in the explicit transport part and central differences for other second order derivatives. For all space-dependent terms, we use notation $u_i \approx u(x_i)$ where $x_i = (i - 1/2)\Delta x$, $i \in \mathbb{Z}$.

During each internal stage of (5.7) the fully discrete scheme reads

$$
\begin{aligned}
\hat{\mathbf{R}}_i^k = \hat{\mathbf{r}}_i^n + \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} \Big\{ & -\frac{v}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l - \hat{\mathbf{J}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l - 2\hat{\mathbf{R}}_i^l + \hat{\mathbf{R}}_{i-1}^l) \\
& - \frac{v\phi^{1/2}}{4}(\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l + \boldsymbol{\beta}_{i+1}^l - \boldsymbol{\beta}_i^l) \\
& - \frac{\mu}{(\Delta x)^2}\bar{F}(v) D\tilde{\mathbf{M}} \left( \hat{\boldsymbol{\rho}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l \right) \\
& + \frac{\mu}{2\Delta x}\bar{F}(v)\chi \left( \tilde{\mathbf{G}}_{i+1}^l \hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l \hat{\mathbf{P}}_{i-1}^l \right) \Big\} \\
& + \Delta t \sum_{l=1}^{k} a_{kl} \Big\{ \frac{1}{\varepsilon^2} \left( \bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^l + \mathbf{B}_i^l \hat{\mathbf{P}}_i^l - \mathbf{M}\hat{\mathbf{R}}_i^l - \mathbf{C}_i^l \hat{\mathbf{R}}_i^l \right) \\
& + \frac{\mu}{(\Delta x)^2}\bar{F}(v) D\tilde{\mathbf{M}} \left( \hat{\boldsymbol{\rho}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l \right) \\
& - \frac{\mu}{2\Delta x}\bar{F}(v)\chi \left( \tilde{\mathbf{G}}_{i+1}^l \hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l \hat{\mathbf{P}}_{i-1}^l \right) \Big\},
\end{aligned} \tag{5.14a}
$$

$$
\begin{aligned}
\hat{\mathbf{J}}_i^k = \hat{\mathbf{j}}_i^n + \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} \Big\{ & -\frac{v\phi}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l - \hat{\mathbf{R}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l - 2\hat{\mathbf{J}}_i^l + \hat{\mathbf{J}}_{i-1}^l) \\
& - \frac{v\phi}{4}(\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l - \boldsymbol{\beta}_{i+1}^l + \boldsymbol{\beta}_i^l) \Big\} \\
& - \Delta t \sum_{l=1}^{k} a_{kl} \frac{1}{\varepsilon^2} \left\{ (1 - \varepsilon^2\phi) v \frac{\hat{\mathbf{R}}_{i+1}^l - \hat{\mathbf{R}}_{i-1}^l}{2\Delta x} - \mathbf{E}_i^l \hat{\mathbf{P}}_i^l + \mathbf{M}\hat{\mathbf{J}}_i^l + \mathbf{C}_i^l \hat{\mathbf{J}}_i^l \right\},
\end{aligned} \tag{5.14b}
$$

where

$$
\boldsymbol{\gamma}_i^l = \frac{1}{\Delta x}\text{minmod}\left( \hat{\mathbf{R}}_{i+1}^l + \phi^{-1/2}\hat{\mathbf{J}}_{i+1}^l - \hat{\mathbf{R}}_i^l - \phi^{-1/2}\hat{\mathbf{J}}_i^l, \right. \tag{5.15a}
$$

$$
\left. \hat{\mathbf{R}}_i^l + \phi^{-1/2}\hat{\mathbf{J}}_i^l - \hat{\mathbf{R}}_{i-1}^l - \phi^{-1/2}\hat{\mathbf{J}}_{i-1}^l \right), \tag{5.15b}
$$

$$
\boldsymbol{\beta}_i^l = \frac{1}{\Delta x}\text{minmod}\left( \hat{\mathbf{R}}_{i+1}^l - \phi^{-1/2}\hat{\mathbf{J}}_{i+1}^l - \hat{\mathbf{R}}_i^l + \phi^{-1/2}\hat{\mathbf{J}}_i^l, \right. \tag{5.15c}
$$

$$
\left. \hat{\mathbf{R}}_i^l - \phi^{-1/2}\hat{\mathbf{J}}_i^l - \hat{\mathbf{R}}_{i-1}^l + \phi^{-1/2}\hat{\mathbf{J}}_{i-1}^l \right). \tag{5.15d}
$$

Since $\hat{\mathbf{P}}^k$ can be obtained explicitly by (5.13), we can fully discretize $\hat{\mathbf{P}}_i^k$ as follows:

$$
\begin{aligned}
& \hat{\mathbf{P}}_i^k - \Delta t a_{kk} \frac{\mu}{(\Delta x)^2} \Big[ D\tilde{\mathbf{M}}(\hat{\mathbf{P}}_{i-1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i+1}^k) \\
& -\chi \left( \tilde{\mathbf{G}}_{i+\frac{1}{2}}^{k-1}(\hat{\mathbf{P}}_{i+1}^k - \hat{\mathbf{P}}_i^k) - \tilde{\mathbf{G}}_{i-\frac{1}{2}}^{k-1}(\hat{\mathbf{P}}_i^k - \hat{\mathbf{P}}_{i-1}^k) \right) \Big] = \left\langle \bar{\hat{\mathbf{R}}}_i^{k-1} \right\rangle,
\end{aligned} \tag{5.16}
$$

where $\tilde{\mathbf{G}}_{i\pm\frac{1}{2}}^{k-1}$ is an approximation of $\tilde{\mathbf{G}}^{k-1}$ at position $x_{i\pm\frac{1}{2}}$.

Then, using (5.16), the fully discretized $\hat{\mathbf{R}}_i^k$, and subsequently $\hat{\mathbf{J}}_i^k$, is obtained

from the following:

$$
\begin{aligned}
&\left(\mathbf{I}+\frac{a_{kk}\Delta t}{\varepsilon^2}(\mathbf{M}+\mathbf{C}_i^k)\right)\hat{\mathbf{R}}_i^k=\hat{\mathbf{r}}_i^n+\Delta t\sum_{l=1}^{k-1}\tilde{a}_{kl}\left\{-\frac{v}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l-\hat{\mathbf{J}}_{i-1}^l)\right.\\
&\quad+\frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l-2\hat{\mathbf{R}}_i^l+\hat{\mathbf{R}}_{i-1}^l)-\frac{v\phi^{1/2}}{4}(\boldsymbol{\gamma}_i^l-\boldsymbol{\gamma}_{i-1}^l+\boldsymbol{\beta}_{i+1}^l-\boldsymbol{\beta}_i^l)\\
&\quad-\frac{\mu}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^l-2\hat{\mathbf{P}}_i^l+\hat{\mathbf{P}}_{i-1}^l\right)\\
&\quad\left.+\frac{\mu}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^l\hat{\mathbf{P}}_{i+1}^l-\tilde{\mathbf{G}}_{i-1}^l\hat{\mathbf{P}}_{i-1}^l\right)\right\}\\
&\quad+\Delta t\sum_{l=1}^{k-1}a_{kl}\left\{\frac{1}{\varepsilon^2}\left[\bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^l+\mathbf{B}_i^l\hat{\mathbf{P}}_i^l-\mathbf{M}\hat{\mathbf{R}}_i^l-\mathbf{C}_i^l\hat{\mathbf{R}}_i^l\right]\right.\\
&\quad+\frac{\mu}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^l-2\hat{\mathbf{P}}_i^l+\hat{\mathbf{P}}_{i-1}^l\right)\\
&\quad\left.-\frac{\mu}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^l\hat{\mathbf{P}}_{i+1}^l-\tilde{\mathbf{G}}_{i-1}^l\hat{\mathbf{P}}_{i-1}^l\right)\right\}\\
&\quad+\Delta t a_{kk}\left\{\frac{1}{\varepsilon^2}\left[\bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^k+\mathbf{B}_i^k\hat{\mathbf{P}}_i^k\right]\right.\\
&\quad+\frac{\mu}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^k-2\hat{\mathbf{P}}_i^k+\hat{\mathbf{P}}_{i-1}^k\right)\\
&\quad\left.-\frac{\mu}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^k\hat{\mathbf{P}}_{i+1}^k-\tilde{\mathbf{G}}_{i-1}^k\hat{\mathbf{P}}_{i-1}^k\right)\right\},
\end{aligned}
\tag{5.17a}
$$

$$
\begin{aligned}
&\left(\mathbf{I}+\frac{a_{kk}\Delta t}{\varepsilon^2}(\mathbf{M}+\mathbf{C}_i^k)\right)\hat{\mathbf{J}}_i^k=\hat{\mathbf{j}}_i^n+\Delta t\sum_{l=1}^{k-1}\tilde{a}_{kl}\left\{-\frac{v\phi}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l-\hat{\mathbf{R}}_{i-1}^l)\right.\\
&\quad\left.+\frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l-2\hat{\mathbf{J}}_i^l+\hat{\mathbf{J}}_{i-1}^l)-\frac{v\phi}{4}(\boldsymbol{\gamma}_i^l-\boldsymbol{\gamma}_{i-1}^l+\boldsymbol{\beta}_{i+1}^l-\boldsymbol{\beta}_i^l)\right\}\\
&\quad-\Delta t\sum_{l=1}^{k-1}a_{kl}\frac{1}{\varepsilon^2}\left\{(1-\varepsilon^2\phi)v\frac{\hat{\mathbf{R}}_{i+1}^l-\hat{\mathbf{R}}_{i-1}^l}{2\Delta x}-\mathbf{E}_i^l\hat{\mathbf{P}}_i^l+\mathbf{M}\hat{\mathbf{J}}_i^l+\mathbf{C}_i^l\hat{\mathbf{J}}_i^l\right\}\\
&\quad-\Delta t a_{kk}\frac{1}{\varepsilon^2}\left\{(1-\varepsilon^2\phi)v\frac{\hat{\mathbf{R}}_{i+1}^k-\hat{\mathbf{R}}_{i-1}^k}{2\Delta x}-\mathbf{E}_i^k\hat{\mathbf{P}}_i^k\right\},
\end{aligned}
\tag{5.17b}
$$

In the above system $\left(1+\frac{a_{kk}\Delta t}{\varepsilon^2}(\mathbf{M}+\mathbf{C}_i^k)\right)$ is symmetric positive definite and thus invertible. After calculating all $\hat{\mathbf{R}}_i^k$ and $\hat{\mathbf{J}}_i^k$ for $k=1,\dots,s$, we can update $\hat{\mathbf{r}}_i^{n+1}$ and

$\hat{\mathbf{j}}_i^{n+1}$ in (5.6) as

$$
\begin{aligned}
\hat{\mathbf{r}}_i^{n+1} =& \hat{\mathbf{r}}_i^n + \Delta t \sum_{k=1}^{s} \tilde{b}_k \Big\{ -\frac{v}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^k - \hat{\mathbf{J}}_{i-1}^k) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^k - 2\hat{\mathbf{R}}_i^k + \hat{\mathbf{R}}_{i-1}^k) \\
& - \frac{v\phi^{1/2}}{4}(\boldsymbol{\gamma}_i^k - \boldsymbol{\gamma}_{i-1}^k + \boldsymbol{\beta}_{i+1}^k - \boldsymbol{\beta}_i^k) - \frac{\mu}{(\Delta x)^2}\bar{F}(v) D\tilde{\mathbf{M}} \left( \hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k \right) \\
& + \frac{\mu}{2\Delta x}\bar{F}(v)\chi \left( \tilde{\mathbf{G}}_{i+1}^k \hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k \hat{\mathbf{P}}_{i-1}^k \right) \Big\} \\
& + \Delta t \sum_{k=1}^{s} b_k \Big\{ \frac{1}{\varepsilon^2} \left[ \bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^k + \mathbf{B}_i^k \hat{\mathbf{P}}_i^k - \mathbf{M}\hat{\mathbf{R}}_i^k - \mathbf{C}_i^k \hat{\mathbf{R}}_i^k \right] \\
& + \frac{\mu}{(\Delta x)^2}\bar{F}(v) D\tilde{\mathbf{M}} \left( \hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k \right) \\
& - \frac{\mu}{2\Delta x}\bar{F}(v)\chi \left( \tilde{\mathbf{G}}_{i+1}^k \hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k \hat{\mathbf{P}}_{i-1}^k \right) \Big\},
\end{aligned} \tag{5.18a}
$$

$$
\begin{aligned}
\hat{\mathbf{j}}_i^{n+1} =& \hat{\mathbf{j}}_i^n + \Delta t \sum_{k=1}^{s} \tilde{b}_k \Big\{ -\frac{v\phi}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^k - 2\hat{\mathbf{J}}_i^k + \hat{\mathbf{J}}_{i-1}^k) \\
& - \frac{v\phi}{4}(\boldsymbol{\gamma}_i^k - \boldsymbol{\gamma}_{i-1}^k - \boldsymbol{\beta}_{i+1}^k + \boldsymbol{\beta}_i^k) \Big\} \\
& - \Delta t \sum_{k=1}^{s} b_k \frac{1}{\varepsilon^2} \left\{ (1 - \varepsilon^2\phi) v \frac{\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k}{2\Delta x} - \mathbf{E}_i^k \hat{\mathbf{P}}_i^k + \mathbf{M}\hat{\mathbf{J}}_i^k + \mathbf{C}_i^k \hat{\mathbf{J}}_i^k \right\},
\end{aligned} \tag{5.18b}
$$

where $\boldsymbol{\gamma}_i^k$ and $\boldsymbol{\beta}_i^k$ are defined as in (5.15).

Following [6], we choose

$$
\mu = \mu(\varepsilon, \Delta x) = \exp(-\varepsilon^2/\Delta x). \tag{5.19}
$$

Thus, for large values of $\varepsilon$ (e.g., $\varepsilon = O(1)$), $\mu \approx 0$, and we could avoid the loss of accuracy caused by adding and subtracting the penalty term, whereas for very small values of $\varepsilon$ (e.g., $\varepsilon \to 0$), $\mu \to 1$, which guarantees that we have an implicit scheme for the limiting diffusion equation.

*Remark* 5.1. The full discrete scheme is obtained using the Gauss–Legendre quadrature nodes for the velocity discretization. Finally, to get the boundary conditions for $\hat{\mathbf{r}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{s}}$, we refer the reader to [31] for details.

**5.2. The sAP property.** Denote

$$
\begin{aligned}
f_1(\hat{\mathbf{R}}_i^l, \hat{\mathbf{J}}_i^l) =& -\frac{v}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l - \hat{\mathbf{J}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l - 2\hat{\mathbf{R}}_i^l + \hat{\mathbf{R}}_{i-1}^l) \\
& - \frac{v\phi^{1/2}}{4}(\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l + \boldsymbol{\beta}_{i+1}^l - \boldsymbol{\beta}_i^l) \\
& - \frac{\mu}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l\right) \\
& + \frac{\mu}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^l\hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l\hat{\mathbf{P}}_{i-1}^l\right),
\end{aligned} \tag{5.20a}
$$

$$
\begin{aligned}
f_2(\hat{\mathbf{R}}_i^l) =& \frac{1}{\varepsilon^2}\left[\bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^l + \mathbf{B}_i^l\hat{\mathbf{P}}_i^l - \mathbf{M}\hat{\mathbf{R}}_i^l - \mathbf{C}_i^l\hat{\mathbf{R}}_i^l\right] \\
& + \frac{\mu}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l\right) \\
& - \frac{\mu}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^l\hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l\hat{\mathbf{P}}_{i-1}^l\right),
\end{aligned} \tag{5.20b}
$$

$$
\begin{aligned}
g_1(\hat{\mathbf{R}}_i^l) =& -\frac{v\phi}{2\Delta x}(\hat{\mathbf{R}}_{i+1}^l - \hat{\mathbf{R}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x}(\hat{\mathbf{J}}_{i+1}^l - 2\hat{\mathbf{J}}_i^l + \hat{\mathbf{J}}_{i-1}^l) \\
& - \frac{v\phi}{4}(\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l - \boldsymbol{\beta}_{i+1}^l + \boldsymbol{\beta}_i^l),
\end{aligned} \tag{5.20c}
$$

$$
g_2(\hat{\mathbf{R}}_i^l, \hat{\mathbf{J}}_i^l) = \frac{1}{\varepsilon^2}\left[(1-\varepsilon^2\phi)v\frac{\hat{\mathbf{R}}_{i+1}^l - \hat{\mathbf{R}}_{i-1}^l}{2\Delta x} - \mathbf{E}_i^l\hat{\mathbf{P}}_i^l + \mathbf{M}\hat{\mathbf{J}}_i^l + \mathbf{C}_i^l\hat{\mathbf{J}}_i^l\right]. \tag{5.20d}
$$

From (5.17) we have

$$
\begin{pmatrix} \hat{\mathbf{R}}_i^1 \\ \hat{\mathbf{R}}_i^2 \\ \vdots \\ \hat{\mathbf{R}}_i^s \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{r}}_i^n \\ \hat{\mathbf{r}}_i^n \\ \vdots \\ \hat{\mathbf{r}}_i^n \end{pmatrix} + \Delta t \begin{pmatrix} 0 \\ \tilde{a}_{21} f_1(\hat{\mathbf{R}}_i^1, \hat{\mathbf{J}}_i^1) \\ \vdots \\ \sum_{l=1}^{s-1} \tilde{a}_{sl} f_1(\hat{\mathbf{R}}_i^l, \hat{\mathbf{J}}_i^l) \end{pmatrix} + \Delta t \mathbf{A} \begin{pmatrix} f_2(\hat{\mathbf{R}}_i^1) \\ f_2(\hat{\mathbf{R}}_i^2) \\ \vdots \\ f_2(\hat{\mathbf{R}}_i^s) \end{pmatrix}, \tag{5.21a}
$$

$$
\begin{pmatrix} \hat{\mathbf{J}}_i^1 \\ \hat{\mathbf{J}}_i^2 \\ \vdots \\ \hat{\mathbf{J}}_i^s \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{j}}_i^n \\ \hat{\mathbf{j}}_i^n \\ \vdots \\ \hat{\mathbf{j}}_i^n \end{pmatrix} + \Delta t \begin{pmatrix} 0 \\ \tilde{a}_{21} g_1(\hat{\mathbf{R}}_i^1) \\ \vdots \\ \sum_{l=1}^{s-1} \tilde{a}_{sl} g_1(\hat{\mathbf{R}}_i^l) \end{pmatrix} + \Delta t \mathbf{A} \begin{pmatrix} g_2(\hat{\mathbf{R}}_i^1, \hat{\mathbf{J}}_i^1) \\ g_2(\hat{\mathbf{R}}_i^2, \hat{\mathbf{J}}_i^2) \\ \vdots \\ g_2(\hat{\mathbf{R}}_i^s, \hat{\mathbf{J}}_i^s) \end{pmatrix}, \tag{5.21b}
$$

where

$$
\mathbf{A}_{K(i-1)+1:Ki,K(j-1)+1:Kj} = A_{i,j}\mathbf{I}_{K\times K}, \quad \mathbf{I}_{K\times K} \text{ is a } K\times K \text{ identity matrix}, \tag{5.22}
$$

and $A$ is defined as in (5.8).

Denote $\mathbf{W}$ as the inverse matrix of $\mathbf{A}$; then we obtain from (5.21)

$$\Delta t \begin{pmatrix} f_2(\hat{\mathbf{R}}_i^1) \\ f_2(\hat{\mathbf{R}}_i^2) \\ \vdots \\ f_2(\hat{\mathbf{R}}_i^s) \end{pmatrix} = \mathbf{W} \left[ \begin{pmatrix} \hat{\mathbf{R}}_i^1 \\ \hat{\mathbf{R}}_i^2 \\ \vdots \\ \hat{\mathbf{R}}_i^s \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{r}}_i^n \\ \hat{\mathbf{r}}_i^n \\ \vdots \\ \hat{\mathbf{r}}_i^n \end{pmatrix} - \Delta t \begin{pmatrix} 0 \\ \tilde{a}_{21} f_1(\hat{\mathbf{R}}_i^1, \hat{\mathbf{J}}_i^1) \\ \vdots \\ \sum_{l=1}^{s-1} \tilde{a}_{sl} f_1(\hat{\mathbf{R}}_i^l, \hat{\mathbf{J}}_i^l) \end{pmatrix} \right], \tag{5.23a}$$

$$\Delta t \begin{pmatrix} g_2(\hat{\mathbf{R}}_i^1, \hat{\mathbf{J}}_i^1) \\ g_2(\hat{\mathbf{R}}_i^2, \hat{\mathbf{J}}_i^2) \\ \vdots \\ g_2(\hat{\mathbf{R}}_i^s, \hat{\mathbf{J}}_i^s) \end{pmatrix} = \mathbf{W} \left[ \begin{pmatrix} \hat{\mathbf{J}}_i^1 \\ \hat{\mathbf{J}}_i^2 \\ \vdots \\ \hat{\mathbf{J}}_i^s \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{j}}_i^n \\ \hat{\mathbf{j}}_i^n \\ \vdots \\ \hat{\mathbf{j}}_i^n \end{pmatrix} - \Delta t \begin{pmatrix} 0 \\ \tilde{a}_{21} g_1(\hat{\mathbf{R}}_i^1) \\ \vdots \\ \sum_{l=1}^{s-1} \tilde{a}_{sl} g_1(\hat{\mathbf{R}}_i^l) \end{pmatrix} \right]. \tag{5.23b}$$

Since $\mathbf{W}$ has the same structure as $\mathbf{A}$, $\mathbf{W}$ should be a lower triangular matrix with entries

$$\mathbf{W}_{K(i-1)+1:Ki, K(j-1)+1:Kj} = \omega_{i,j} \mathbf{I}_{K\times K}, \tag{5.24}$$

where $W = (\omega_{i,j})$ is the inverse of the lower triangular matrix $A$ in (5.8).

Then one can rewrite (5.23) as

$$\Delta t f_2(\hat{\mathbf{R}}_i^k) = \sum_{l=1}^{k} \omega_{kl} \left[ \hat{\mathbf{R}}_i^l - \hat{\mathbf{r}}_i^n - \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} f_1(\hat{\mathbf{R}}_i^l, \hat{\mathbf{J}}_i^l) \right], \tag{5.25a}$$

$$\Delta t g_2(\hat{\mathbf{R}}_i^k, \hat{\mathbf{J}}_i^k) = \sum_{l=1}^{k} \omega_{kl} \left[ \hat{\mathbf{J}}_i^l - \hat{\mathbf{j}}_i^n - \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} g_1(\hat{\mathbf{R}}_i^l) \right]. \tag{5.25b}$$

More explicitly,

$$\begin{aligned}
&\Delta t \left\{ \frac{1}{\varepsilon^2} \left[ \bar{F}(v) \mathbf{M} \hat{\mathbf{P}}_i^k + \mathbf{B}_i^k \hat{\mathbf{P}}_i^k - \mathbf{M} \hat{\mathbf{R}}_i^k - \mathbf{C}_i^k \hat{\mathbf{R}}_i^k \right] \right. \\
&\left. + \frac{\mu}{(\Delta x)^2} \bar{F}(v) D \tilde{\mathbf{M}} \left( \hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k \right) - \frac{\mu}{2\Delta x} \bar{F}(v) \chi \left( \tilde{\mathbf{G}}_{i+1}^k \hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k \hat{\mathbf{P}}_{i-1}^k \right) \right\} \\
&= \sum_{l=1}^{k} \omega_{kl} \left\{ \hat{\mathbf{R}}_i^l - \hat{\mathbf{r}}_i^n - \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} \left[ -\frac{v}{2\Delta x} (\hat{\mathbf{J}}_{i+1}^l - \hat{\mathbf{J}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x} (\hat{\mathbf{R}}_{i+1}^l - 2\hat{\mathbf{R}}_i^l + \hat{\mathbf{R}}_{i-1}^l) \right. \right. \\
&\quad - \frac{v\phi^{1/2}}{4} (\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l + \boldsymbol{\beta}_{i+1}^l - \boldsymbol{\beta}_i^l) - \frac{\mu}{(\Delta x)^2} \bar{F}(v) D \tilde{\mathbf{M}} \left( \hat{\mathbf{P}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l \right) \\
&\quad \left. \left. + \frac{\mu}{2\Delta x} \bar{F}(v) \chi \left( \tilde{\mathbf{G}}_{i+1}^l \hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l \hat{\mathbf{P}}_{i-1}^l \right) \right] \right\},
\end{aligned} \tag{5.26a}$$

$$\begin{aligned}
&\Delta t \left\{ \frac{1}{\varepsilon^2} \left[ (1 - \varepsilon^2 \phi) v \frac{\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k}{2\Delta x} - \mathbf{E}_i^k \hat{\mathbf{P}}_i^k + \mathbf{M} \hat{\mathbf{J}}_i^k + \mathbf{C}_i^k \hat{\mathbf{J}}_i^k \right] \right\} \\
&= \sum_{l=1}^{k} \omega_{kl} \left\{ \hat{\mathbf{J}}_i^l - \hat{\mathbf{j}}_i^n - \Delta t \sum_{l=1}^{k-1} \tilde{a}_{kl} \left[ -\frac{v\phi}{2\Delta x} (\hat{\mathbf{R}}_{i+1}^l - \hat{\mathbf{R}}_{i-1}^l) + \frac{v\phi^{1/2}}{2\Delta x} (\hat{\mathbf{J}}_{i+1}^l - 2\hat{\mathbf{J}}_i^l + \hat{\mathbf{J}}_{i-1}^l) \right. \right. \\
&\quad \left. \left. - \frac{v\phi}{4} (\boldsymbol{\gamma}_i^l - \boldsymbol{\gamma}_{i-1}^l - \boldsymbol{\beta}_{i+1}^l + \boldsymbol{\beta}_i^l) \right] \right\}.
\end{aligned} \tag{5.26b}$$

Thus, multiplying both equations in (5.26) by $\varepsilon^2$ and letting $\varepsilon \to 0$, we have

$$\bar{F}(v)\mathbf{M}\hat{\mathbf{P}}_i^k + \mathbf{B}_i^k\hat{\mathbf{P}}_i^k - (\mathbf{M}+\mathbf{C}_i^k)\hat{\mathbf{R}}_i^k = 0, \tag{5.27a}$$

$$v\frac{\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k}{2\Delta x} - \mathbf{E}_i^k\hat{\mathbf{P}}_i^k + (\mathbf{M}+\mathbf{C}_i^k)\hat{\mathbf{J}}_i^k = 0. \tag{5.27b}$$

Now since $\mathbf{M}+\mathbf{C}_i^k$ is nonsingular, one obtains

$$\hat{\mathbf{R}}_i^k = (\mathbf{M}+\mathbf{C}_i^k)^{-1}(\bar{F}(v)\mathbf{M}+\mathbf{B}_i^k)\hat{\mathbf{P}}_i^k = \bar{F}(v)\hat{\mathbf{P}}_i^k + O(\varepsilon), \tag{5.28a}$$

$$\begin{aligned}\hat{\mathbf{J}}_i^k &= (\mathbf{M}+\mathbf{C}_i^k)^{-1}\left(\mathbf{E}_i^k\hat{\mathbf{P}}_i^k - v\frac{\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k}{2\Delta x}\right)\\ &= \mathbf{M}^{-1}\mathbf{E}_i^k\hat{\mathbf{P}}_i^k - v\mathbf{M}^{-1}\frac{\hat{\mathbf{R}}_{i+1}^k - \hat{\mathbf{R}}_{i-1}^k}{2\Delta x} + O(\varepsilon).\end{aligned} \tag{5.28b}$$

Inserting this back into (5.18a), thanks to the globally stiffly accurate property of the IMEX scheme, as $\varepsilon \to 0$ we get

$$\hat{\mathbf{r}}_i^{n+1} = \hat{\mathbf{r}}_i^n + \Delta t\sum_{k=1}^s \tilde{b}_k\hat{f}_1(\hat{\mathbf{R}}_i^k) + \Delta t\sum_{k=1}^s b_k\hat{f}_2(\hat{\mathbf{R}}_i^k), \tag{5.29}$$

where

$$\begin{aligned}\hat{f}_1(\hat{\mathbf{R}}_i^k) = {} & v^2\bar{F}(v)\frac{\mathbf{M}^{-1}}{4(\Delta x)^2}\left(\hat{\mathbf{R}}_{i+2}^k - 2\hat{\mathbf{R}}_i^k + \hat{\mathbf{R}}_{i-2}^k\right)\\ & - \bar{F}(v)D\frac{\tilde{\mathbf{M}}}{(\Delta x)^2}\left(\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k - \hat{\mathbf{P}}_{i-1}^k\right)\\ & - \frac{v^2}{4\Delta x}(\mathbf{M}^{-1}\mathbf{E}_{i+1}^k\hat{\mathbf{P}}_{i+1}^k - \mathbf{M}^{-1}\mathbf{E}_{i-1}^k\hat{\mathbf{P}}_{i-1}^k)\\ & + \frac{1}{2\Delta x}\bar{F}(v)\chi(\tilde{\mathbf{G}}_{i+1}^k\hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k\hat{\mathbf{P}}_{i-1}^k),\end{aligned} \tag{5.30a}$$

$$\begin{aligned}\hat{f}_2(\hat{\mathbf{R}}_i^k) = {} & \frac{1}{(\Delta x)^2}\bar{F}(v)D\tilde{\mathbf{M}}\left(\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k\right)\\ & - \frac{1}{2\Delta x}\bar{F}(v)\chi\left(\tilde{\mathbf{G}}_{i+1}^k\hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k\hat{\mathbf{P}}_{i-1}^k\right).\end{aligned} \tag{5.30b}$$

Now the difference between $\mathbf{M}^{-1}(\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k)$ and $\tilde{\mathbf{M}}(\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k)$ and the difference between $\frac{1}{\chi}\mathbf{M}^{-1}\mathbf{E}_i^k\hat{\mathbf{P}}_i^k$ and $\tilde{\mathbf{G}}_i^k\hat{\mathbf{P}}_i^k$ are both spectrally small (see Remark 4.1), and thus after integrating over $V^+$, $\hat{f}_1$ goes to 0 and one gets

$$\begin{aligned}\hat{\boldsymbol{\rho}}_i^{n+1} = \hat{\boldsymbol{\rho}}_i^n + \Delta t\sum_{k=1}^s b_k\Bigg[ & \bar{D}\tilde{\mathbf{M}}\frac{\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k}{(\Delta x)^2}\\ & -\bar{\chi}\frac{\tilde{\mathbf{G}}_{i+1}^k\hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^k\hat{\mathbf{P}}_{i-1}^k}{2\Delta x}\Bigg] + O((\Delta x)^2),\end{aligned} \tag{5.31}$$

where

$$\hat{\mathbf{P}}_i^k = \hat{\boldsymbol{\rho}}_i^n + \Delta t \sum_{l=1}^{k-1} a_{kl} \left[ \bar{D}\tilde{\mathbf{M}} \frac{\hat{\mathbf{P}}_{i+1}^l - 2\hat{\mathbf{P}}_i^l + \hat{\mathbf{P}}_{i-1}^l}{(\Delta x)^2} - \bar{\chi} \frac{\tilde{\mathbf{G}}_{i+1}^l \hat{\mathbf{P}}_{i+1}^l - \tilde{\mathbf{G}}_{i-1}^l \hat{\mathbf{P}}_{i-1}^l}{2\Delta x} \right] + \Delta t a_{kk} \left[ \bar{D}\tilde{\mathbf{M}} \frac{\hat{\mathbf{P}}_{i+1}^k - 2\hat{\mathbf{P}}_i^k + \hat{\mathbf{P}}_{i-1}^k}{(\Delta x)^2} - \bar{\chi} \frac{\tilde{\mathbf{G}}_{i+1}^{k-1} \hat{\mathbf{P}}_{i+1}^k - \tilde{\mathbf{G}}_{i-1}^{k-1} \hat{\mathbf{P}}_{i-1}^k}{2\Delta x} \right], \tag{5.32}$$

which is an implicit RK scheme for the projected limiting diffusion equation (4.8). Thus, the sAP property [32] of the proposed IMEX-RK scheme is formally justified.

## 6. Numerical tests.

**6.1. The 1D nonlocal deterministic model.** Our numerical tests are carried out with

$$x \in \Omega = [-1, 1],\ v \in V = [-1, 1],\ \alpha = 1,$$

$$\bar{F}(v) = \frac{1}{|V|}\mathbf{1}_V := \begin{cases} \dfrac{1}{|V|} & \text{if } v \in V, \\ 0 & \text{otherwise.} \end{cases}$$

The critical mass for the limiting Keller–Segel system given by formula (1.3) is

$$M_c = 2\pi.$$

The initial conditions are given by

$$\rho_I(x) = Ce^{-80x^2}, \quad f_I(x, v) = \rho_I(x)F(v),$$

where $C = C(M)$ is a constant determined by the total mass $M$.

For the deterministic case, we compare our results by the second order IMEX-RK method (5.10) (denoted by SSP2 in the figures) with the results by Carrillo and Yan [10] (denoted by CY in the figures). For both tests, we set $\Delta x = 0.005$. In their numerical tests, the CFL condition is

$$\Delta t = \max\left\{ \frac{\varepsilon \Delta x}{2}, \frac{\Delta x^2}{2} \right\}.$$

Obviously, when $\varepsilon$ is small, it suffers from the parabolic CFL condition for the diffusive nature of the Keller–Segel system.

For our IMEX-RK method, the choice of $\Delta t$ is given by

$$\Delta t = \lambda \Delta x,\ \lambda = 0.02,$$

which is much bigger than $\Delta x^2/2$.

**6.1.1. A supercritical mass.** It has been shown in [11] that the solution of the kinetic system can converge to the Keller–Segel system weakly in a finite time interval $[0, t^*]$, with $t^* < t_b$. Here $t_b$ is the blow up time of the corresponding Keller–Segel system.

For the supercritical case, we set

$$M = 4\pi > M_c = 2\pi,\ t = 0.003 < t_b \approx 0.0039. \tag{6.1}$$

Figure 1 shows that the solution to the kinetic equation $\rho$ converges to the solution of the Keller–Segel solution $\rho_0$ as $\varepsilon \to 0$ at time $t = 0.003 < t_b$. Our IMEX-RK results match very well with the results in [10].

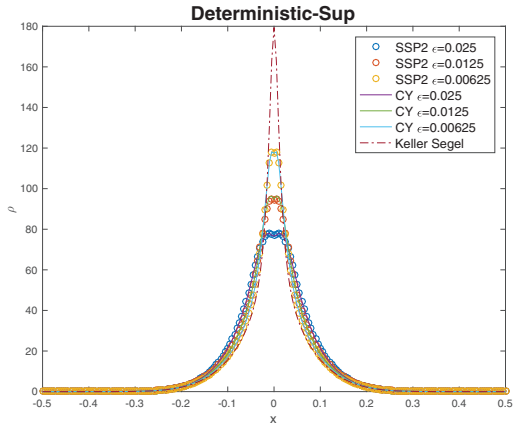


FIG. 1. *The 1D nonlocal deterministic model in the supercritical case. Solid lines are numerical results obtained in* [10] *and circles are numerical results obtained by the IMEX-RK method. The dashed line is the numerical solution of the Keller–Segel equations for reference.*

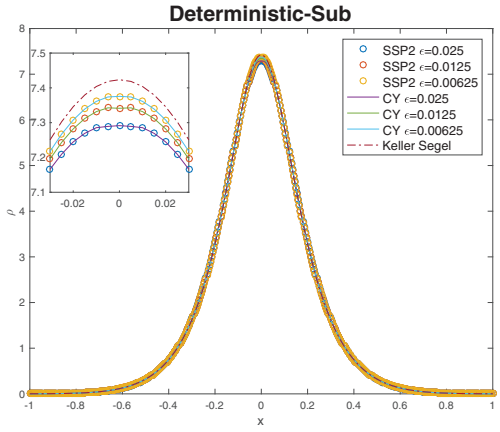


FIG. 2. *The 1D nonlocal deterministic model in the subcritical case. Solid lines are numerical results obtained in* [10] *and circles are numerical results obtained by the IMEX-RK method. The dashed line is the numerical solution of the Keller–Segel equations for reference.*

**6.1.2. A subcritical mass.** For the subcritical case, we set

$$M = \pi < M_c, \ t = 0.1. \tag{6.2}$$

Figure 2 shows convergence results similar to those of the supercritical case for a relatively long time $t = 0.1$. Also, good agreement between our new IMEX-RK solutions and the numerical results from [10] can be observed, even in the zoomed in area. Clearly, in both test cases we observe, as in [10], the first order convergence in $\varepsilon$ of the solution of the IMEX-RK method to the solution of the Keller–Segel model.

**6.2. The 1D nonlocal model with random inputs in the supercritical case.** Now we let

$$\alpha = 1 + 0.5z, \ z \sim U[-1, 1], \ M = 4\pi > \bar{M}_c \approx 2.197\pi.$$

Using the same mesh size as before, we also employ the stochastic collocation method (using 20 quadrature points) as reference solutions. In stochastic collocation, the deterministic solver can be applied directly to a set of selected sample points, and then the solution is approximated by interpolation of all sample solutions (see [50] for a review of stochastic collocation methods). Clearly, the collocation method is AP pointwise for each sample; however, if you use the solutions at the sample points to approximate the values at other points of $z$ by interpolation, or evaluate the expected value, variance, etc., using quadrature rules, then the error of interpolation or quadrature rule depends on the regularity of the solution, which may depend on $\varepsilon$. To avoid these aspects, the collocation solution we put as a reference solution is already the converged solution; namely, adding more sample points would not further change the solutions. The gPC expansion has been considered only up to fourth order in our numerical tests. The following are the comparisons of the two methods in mean and standard deviation for the supercritical case with the same initial mass and stopping time in (6.1). Given the gPC coefficients $\hat{\rho}_k$ of $\rho$, the mean and standard deviation are calculated as

$$\mathbb{E}[\rho] \approx \hat{\rho}_1, \quad \mathbb{S}[\rho] \approx \sqrt{\sum_{k=2}^{K} \hat{\rho}_k^2}.$$

**6.2.1. The sAP property.** Figure 3 shows that the IMEX-RK solution agrees well with results of [10] for all $\varepsilon$ regardless of whether it is combined with the gPC approach or the collocation approach to deal with the uncertainty. Small differences between the two methods, especially near the singularity for small $\varepsilon$, are observed due to different orders of accuracy, but the SG solution always matches the collocation solution accurately for the same deterministic solver. Figure 4 shows that both the mean and the standard deviation of the kinetic chemotaxis solutions tend to the quantities of the limiting Keller–Segel solution as $\varepsilon \to 0$ for fixed $\Delta t$ and $\Delta x$, which verifies the sAP property.

**6.2.2. Global existence and finite time blow up.** As proved in [11], the solution to the kinetic system (2.7) with the nonlocal turning kernel is bounded on $[0, T]$ for any time $T$. However, the Keller–Segel solution will blow up in finite time with a supercritical mass. We examine the mean value and standard deviation of $\|\rho\|_\infty$ for a relatively long time ($t \gg t_b$) in Figure 5. The uncertain systems show the same properties as the deterministic ones; e.g., the kinetic systems have global bound in the first and second moments for different $\varepsilon$, while the Keller–Segel solution will blow up in expected finite time.

**6.2.3. The stationary solution of the kinetic system.** The numerical tests in [10] suggest that the solution of the deterministic kinetic system with a supercritical initial mass stabilizes toward a stationary state after a long time. We also check whether the same property holds for the kinetic system with random inputs. We plot the mean and standard deviation of $\tilde{\rho}(x) = \varepsilon\rho(\varepsilon x)$ in Figure 6, which shows that both the mean and the standard deviation converge to some stationary state at a long time $t = 2$, while the mean agrees with the deterministic stationary solution.

**6.3. The interaction between peaks: The 1D nonlocal model with random initial data.** As shown in [3], the interactions between several peaks for the modified Keller–Segel system can be interpreted as optimal transportation. In the following numerical tests, we make some observations of the interaction changes in the kinetic system caused by different types of randomness in the initial data.

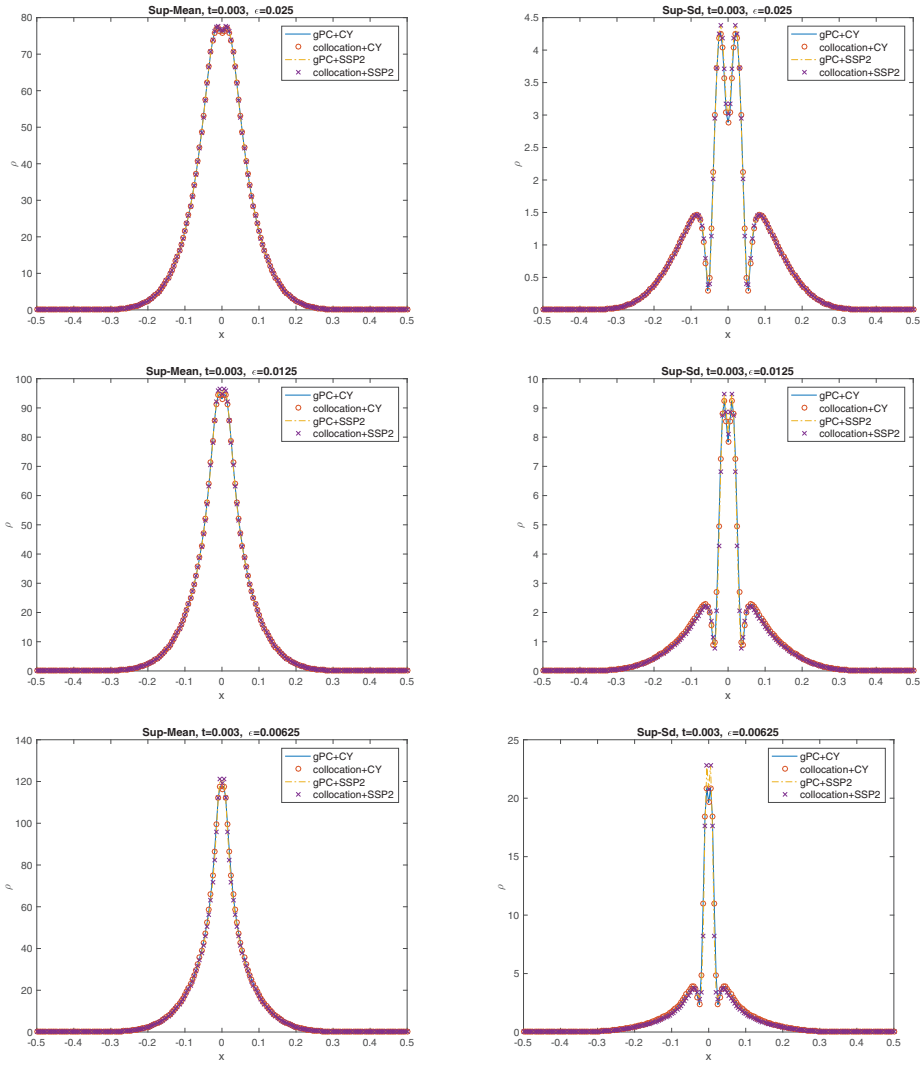


FIG. 3. *The 1D nonlocal random model in the supercritical case. The solid line is obtained by combining the deterministic solver* [10] *with the gPC method, and the circle is obtained by combining the deterministic solver* [10] *with the collocation method. The dashed line is obtained by the IMEX-RK using gPC, and the cross is obtained by the IMEX-RK using collocation. Different values of* $\varepsilon$ *are tested, and the two quantities of interest are mean value (left) and standard deviation (right).*

**6.3.1. Case 1: Two symmetric peaks, without enough mass in each peak.** In this case, we still have $M_c = 2\pi$ and $\bar{M}_c \approx 2.197\pi$. In Figure 7, we reproduced the deterministic attraction between two symmetric peaks with total mass $3\pi$. Then we input symmetric randomness in each peak; i.e., the total mass follows from uniform distribution from $2\pi$ to $4\pi$, keeping each peak without enough mass. Figure 8 shows that symmetric randomness keeps the attraction behavior exactly as in the deterministic case. Symmetric properties are preserved both in mean and standard deviation. However, in Figure 9 we input asymmetric randomness in each peak but keep the total mass fixed as $3\pi$. The two peaks will still be attracted in the center but present behavior different from that in the deterministic case. The

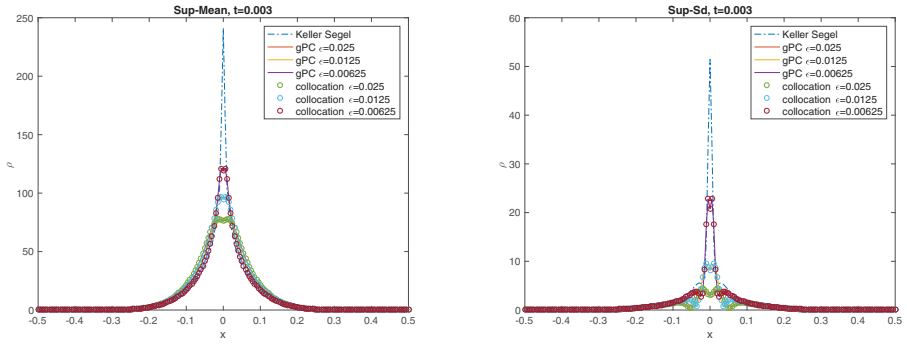


FIG. 4. *The 1D nonlocal random model in the supercritical case. The solid line is obtained by combining the deterministic solver IMEX-RK with the gPC-SG method, and the circle is obtained by combining the deterministic solver IMEX-RK with the collocation method. The dashed line is the gPC-SG solution of the limiting Keller–Segel equations with uncertainty. Different values of $\varepsilon$ are tested, and the two quantities of interest are mean value (left) and standard deviation (right).*

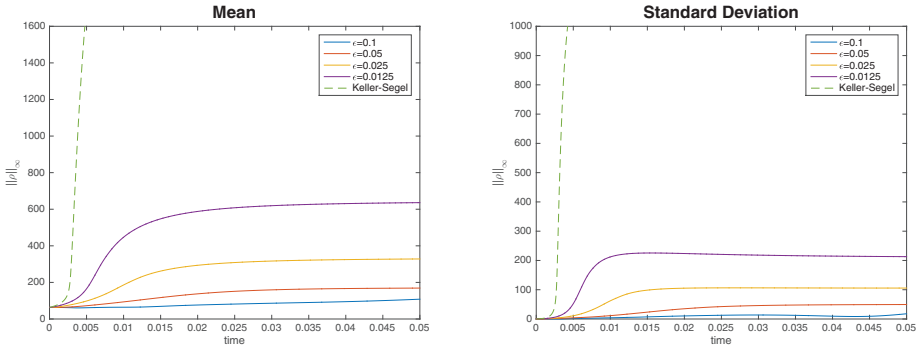


FIG. 5. *The 1D nonlocal random model in the supercritical case. The solid line is obtained by combining the deterministic solver IMEX-RK with the gPC-SG method, and the dashed line is the gPC-SG solution of the limiting Keller–Segel equations. $\rho$ in the infinity norm with different values of $\varepsilon$ is tested, and the two quantities of interest are mean value (left) and standard deviation (right).*

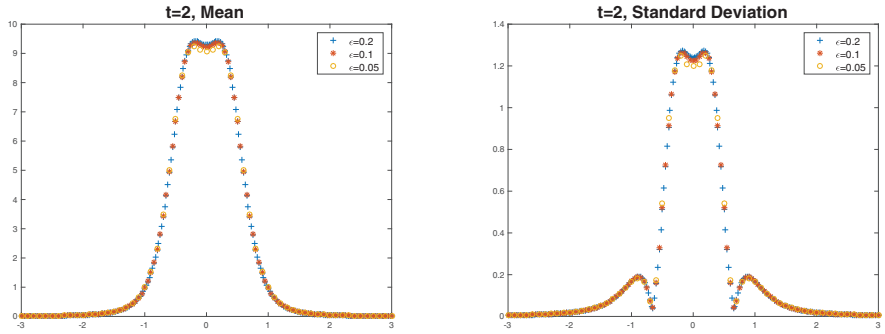


FIG. 6. *The 1D nonlocal random model in the supercritical case. The mean (left) and standard deviation (right) of the function $\varepsilon\rho(\varepsilon x)$ for different $\varepsilon$ are presented. $t = 2 \gg t_b$.*

asymmetric randomness in this case will widen the mean range of the center peak after concentration, in the sense that asymmetric initial data push the concentrated peak toward the direction with more initial mass.

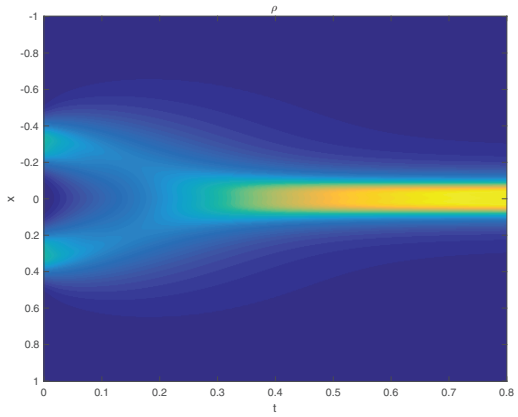


FIG. 7. *Deterministic solution of* $\rho(x,t)$ *with initial data* $f_0 = 4\sqrt{5\pi}\left(1.5e^{-80(x-0.3)^2} + 1.5e^{-80(x+0.3)^2}\right)$, $\varepsilon = 0.1$. *(Figure 8 in* [10]*.)*

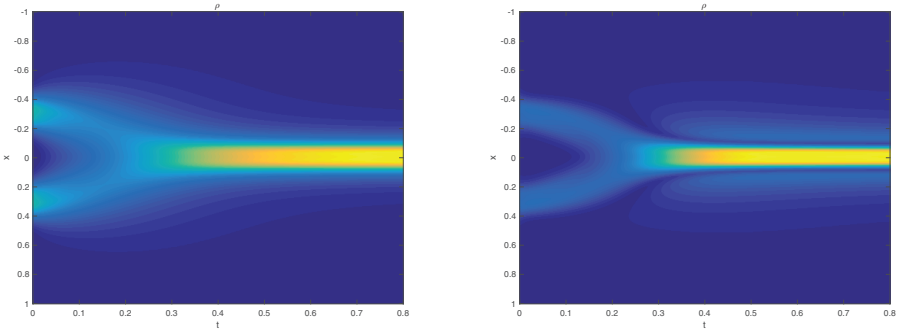


FIG. 8. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random initial condition* $f_0 = 4\sqrt{5\pi}\left((1.5+0.5z)e^{-80(x-0.3)^2} + (1.5+0.5z)e^{-80(x+0.3)^2}\right), z \sim \mathcal{U}[-1,1]$, $\varepsilon = 0.1$.

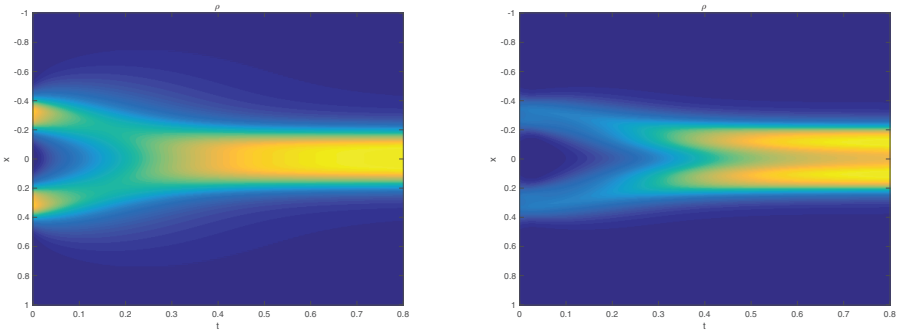


FIG. 9. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random initial condition* $f_0 = 4\sqrt{5\pi}\left((1.5+0.5z)e^{-80(x-0.3)^2} + (1.5-0.5z)e^{-80(x+0.3)^2}\right), z \sim \mathcal{U}[-1,1]$, $\varepsilon = 0.1$.

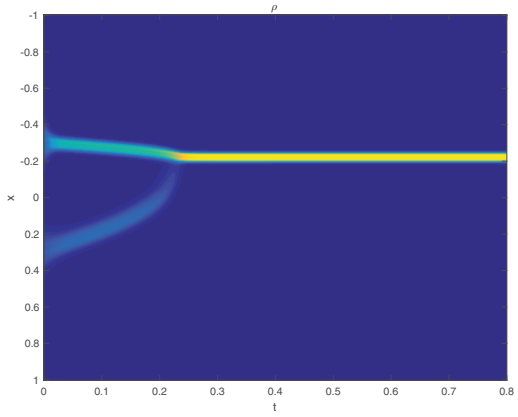


FIG. 10. *Deterministic solution of* $\rho(x,t)$ *with initial data* $f_0 = 4\sqrt{5\pi}\left(3e^{-80(x-0.3)^2} + 5e^{-80(x+0.3)^2}\right)$, $\varepsilon = 0.05$.

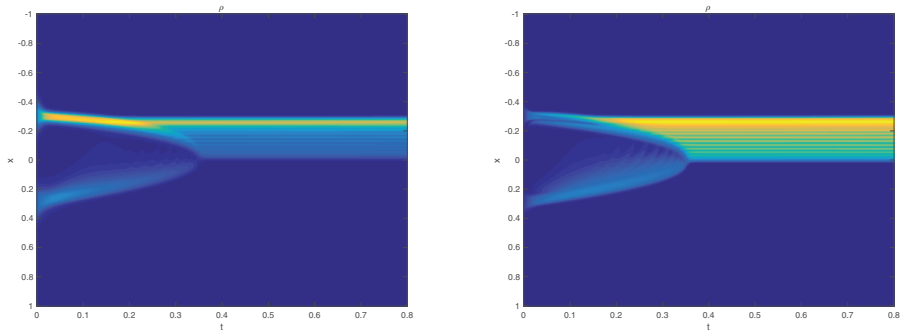


FIG. 11. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random initial condition* $f_0 = 4\sqrt{5\pi}\left((3+z)e^{-80(x-0.3)^2} + (5-z)e^{-80(x+0.3)^2}\right)$, $z \sim \mathcal{U}[-1,1]$, $\varepsilon = 0.05$.

**6.3.2. Case 2: Two asymmetric peaks with enough mass in each peak.** With $M_c = 2\pi$ and $\bar{M}_c \approx 2.197\pi$, we put asymmetric initial masses both larger than $2\pi$. Figure 10 shows results similar to those in Figure 10 of [10]. The mass in each peak is large enough to concentrate, but they will merge into a larger peak which is located closer to a larger initial peak due to asymmetry. Figure 11 shows the effect of the asymmetric randomness with total initial mass fixed. It can be observed in the mean and the standard deviation that the randomness affects the concentration time, location, and asymmetry, showing that the solution is sensitive to initial data.

**6.3.3. Case 3: Two asymmetric peaks (close), one below critical mass, one above critical mass.** From Figure 12 to Figure 15, we conduct a series of experiments with two asymmetric peaks, keeping one peak with enough mass and the other without enough mass. The deterministic case (Figure 12) shows that the peak with less mass will move toward the other in a short time, and then they continue to aggregate mass. In Figure 13, a small amount of randomness exchanged between two peaks will not change this tendency in the mean. The standard deviation in Figure 13 is asymmetric due to the asymmetric randomness in initial data. In Figure

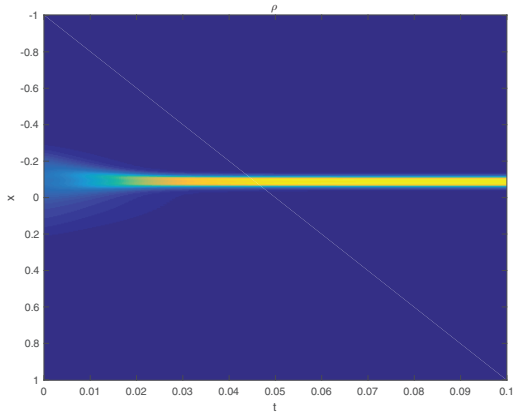


FIG. 12. *Deterministic solution of* $\rho(x,t)$ *with initial data* $f_0 = 4\sqrt{5\pi}\left(e^{-80(x-0.1)^2} + 5e^{-80(x+0.1)^2}\right)$, $\varepsilon = 0.05$.

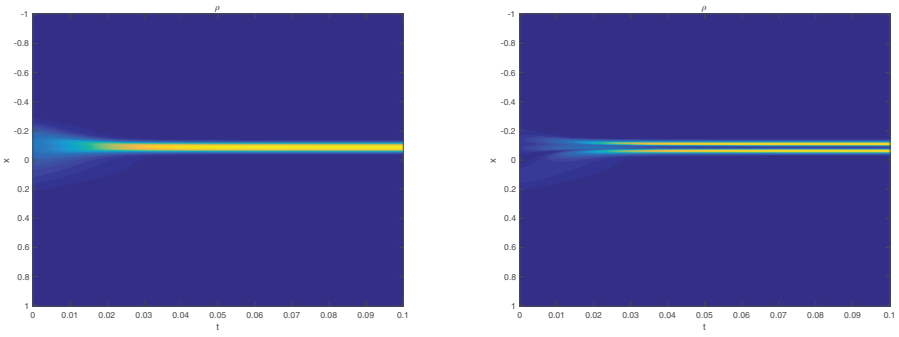


FIG. 13. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random initial condition* $f_0 = 4\sqrt{5\pi}\left((1+0.5z)e^{-80(x-0.1)^2} + (5-0.5z)e^{-80(x+0.1)^2}\right), z \sim \mathcal{U}[-1,1]$, $\varepsilon = 0.05$.

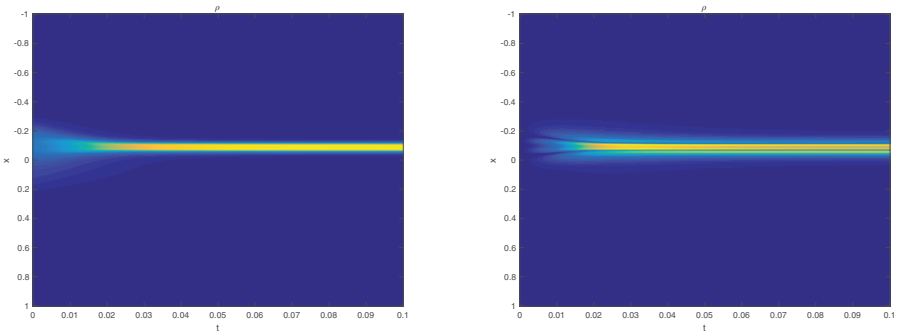


FIG. 14. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random* $\alpha = 1 + 0.5z, z \sim \mathcal{U}[-1,1]$, *deterministic initial data, and* $\varepsilon = 0.05$.

14, although mean values show no difference, the standard deviation is symmetric because the source of randomness comes from the diffusion coefficient $\alpha$. Figure 15 shows that the position of the two peaks has significant effects on the aggregation

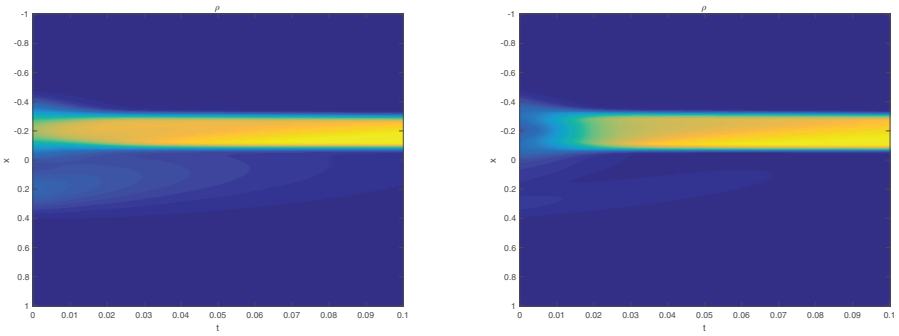


FIG. 15. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$*, respectively, with initial random position* $f_0 = 4\sqrt{5\pi}\left(e^{-80(x-(0.1+0.1z))^2} + 5e^{-80(x+(0.1+0.1z))^2}\right)$*,* $\varepsilon = 0.05$*.*

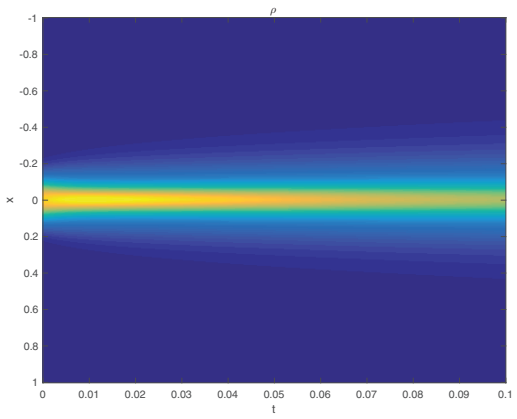


FIG. 16. *Deterministic solution of* $\rho(x,t)$ *with initial data* $f_0 = 1.5 \times 4\sqrt{5\pi}e^{-80x^2}$ *(subcritical),* $\varepsilon = 0.01$*.*

behavior in this case. From the mean and the standard deviation, one can observe that there exists a critical distance between the two peaks, beyond which the two peaks will not be able to merge. They will be separated to behave independently according to their initial mass.

**6.4. The 1D local model with random initial data.** Although theoretic study of the local model with supercritical mass is still not sufficient to understand the blow up behavior of the local kinetic chemotaxis system, numerical tests in [10] suggested blowing up density by using adapted grids. Instead of the blowing up property, we are more interested in studying the sensitivity effect brought about by the randomness around the critical mass. In Figure 16, the deterministic solution with subcritical initial data will stay bounded, as expected from theory. However, the solution continues to aggregate in Figure 17 if we introduce randomness into the initial mass ranging from subcritical mass to supercritical mass, with mean less than critical mass. It is more obvious in Figure 18 that the deterministic solution will remain bounded, while the mean of the random solution appears increasing in time. This indicates that the introduced randomness will influence the properties of the solution. If the range of the initial data contains supercritical regimes, the solution of the random system will behave quite differently from the deterministic one with average initial mass.

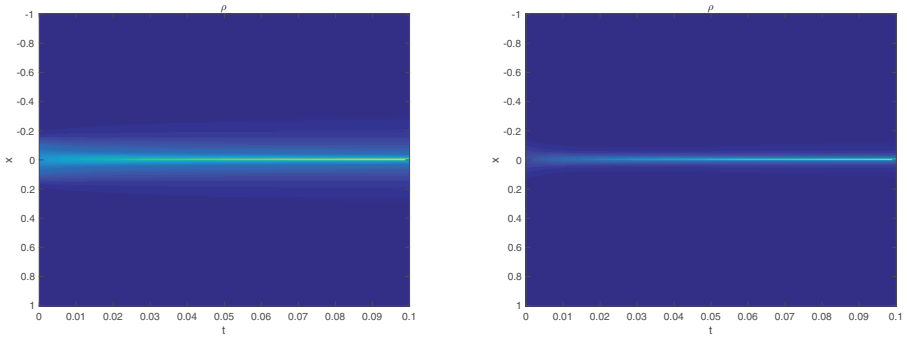


FIG. 17. *Left is the mean and right is the standard deviation of* $\rho(x,t,z)$, *respectively, with random initial data* $f_0 = (1.5+z)\times 4\sqrt{5\pi}e^{-80x^2}$, $\varepsilon = 0.01$.

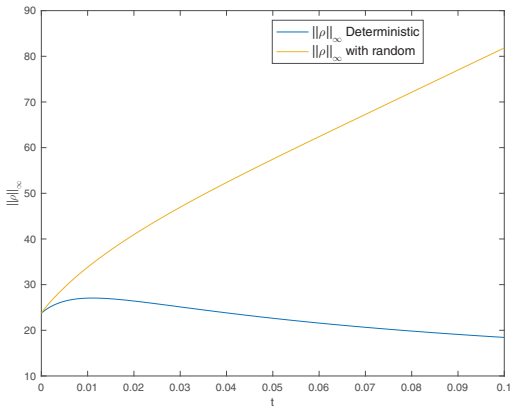


FIG. 18. *Comparison of* $\|\rho\|_\infty$ *in the deterministic solution and the mean solution.*

*Remark* 6.1. A stochastic collocation method is used in the test in subsection 6.4 to deal with $|\partial_x s|$ as follows: Once $\hat{\mathbf{s}} = (\hat{s}_1, \ldots, \hat{s}_K)^T$ is obtained at each time iteration, $\partial_x \hat{\mathbf{s}} = (\partial_x \hat{s}_1, \ldots, \partial_x \hat{s}_K)^T$ can be obtained using finite difference. Then $|\partial_x s(x,z)|$ can be approximated by $|\sum_{k=1}^K \partial_x \hat{s}_k(x)\Phi_k(z)|$. According to the probability density function of $z$, one can have a set of collocation points $\{z_j\}_{j=1}^M$ with corresponding weights $\{w_j\}_{j=1}^M$. ($M = 20$ points are used in our test.) Projecting $|\partial_x s(x,z)|$ onto the space $\{\Phi_1(z), \ldots, \Phi_K(z)\}$ in order to get the gPC coefficients $(\xi_1, \ldots, \xi_K)^T$ of $|\partial_x s|$ such that $|\partial_x s| \approx \sum_{k=1}^K \xi_k(x)\Phi_k(z)$, one can get

$$\begin{aligned}
\xi_k(x) &= \int_{I_z} |\partial_x s(x,z)|\Phi_k(z)\lambda(z)dz \\
&\approx \sum_{j=1}^M |\partial_x s(x,z_j)|\Phi_k(z_j)w_j \\
&\approx \sum_{j=1}^M \left|\sum_{i=1}^K \partial_x \hat{s}_i(x)\Phi_i(z_j)\right|\Phi_k(z_j)w_j, \qquad k = 1, \ldots, K.
\end{aligned}$$

Then $(\xi_1, \ldots, \xi_K)^T$ is used in the algorithm.

**7. Conclusion.** In this article, a high order efficient stochastic Asymptotic-Preserving (AP) scheme is designed for the kinetic chemotaxis system with random inputs. Compared with the previous work [10] for the deterministic kinetic chemotaxis equations, our new method, based on the generalized Polynomial Chaos (gPC) stochastic Galerkin (SC) approach to dealing with uncertainty, uses the implicit-explicit Runge–Kutta (IMEX-RK) method to gain high accuracy and utilize a macroscopic penalty to improve the CFL stability condition from parabolic type to hyperbolic type in the diffusive regime. The greater efficiency has been verified in comparison to existing schemes.

There is much remaining work for future study. Since the kinetic description of the chemotaxis system is more microscopic and consistent with the classical Keller–Segel equation, with more favorable properties (e.g., global existence for nonlocal turning kernel), it is important not only to complete the theory but also to conduct efficient numerical simulations, comparing them with experimental results. On one hand, many properties, which have been explored numerically in this paper and previous work [10, 12], remain to be verified by rigorous theory. On the other hand, the high order efficient method in this paper should be extended to 2D and 3D kinetic chemotaxis systems to support the theory in future work. Moreover, some general problems for uncertainty quantification (UQ), such as high dimensionality and rigorous sensitivity analysis, are to be further studied.

## REFERENCES


[1] W. Alt, *Orientation of cells migrating in a chemotactic gradient*, in Biological Growth and Spread, Springer, 1980, pp. 353–366.

[2] W. Alt, *Biased random walk models for chemotaxis and related diffusion approximations*, J. Math. Biol., 9 (1980), pp. 147–177.

[3] A. Blanchet, V. Calvez, and J. A. Carrillo, *Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model*, SIAM J. Numer. Anal., 46 (2008), pp. 691–721, https://doi.org/10.1137/070683337.

[4] A. Blanchet, J. Dolbeault, and B. Perthame, *Two-dimensional Keller-Segel model: Optimal critical mass and qualitative properties of the solutions*, Electron. J. Differential Equations (EJDE), 2006 (2006), 44.

[5] S. Boscarino, L. Pareschi, and G. Russo, *A unified IMEX Runge–Kutta approach for hyperbolic systems with multiscale relaxation*, SIAM J. Numer. Anal., 55 (2017), pp. 2085–2109, https://doi.org/10.1137/M1111449.

[6] S. Boscarino, L. Pareschi, and G. Russo, *Implicit-explicit Runge–Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit*, SIAM J. Sci. Comput., 35 (2013), pp. A22–A51, https://doi.org/10.1137/110842855.

[7] N. Bournaveas and V. Calvez, *Critical mass phenomenon for a chemotaxis kinetic model with spherically symmetric initial data*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 26 (2009), pp. 1871–1895.

[8] M. P. Brenner, P. Constantin, L. P. Kadanoff, A. Schenkel, and S. C. Venkataramani, *Diffusion, attraction and collapse*, Nonlinearity, 12 (1999), pp. 1071–1098.

[9] V. Calvez, B. Perthame, and M. Sharifi Tabar, *Modified Keller-Segel system and critical mass for the log interaction kernel*, in Stochastic Analysis and Partial Differential Equations, Contemp. Math. 429, Amer. Math. Soc., 2007, pp. 45–62.

[10] J. A. Carrillo and B. Yan, *An asymptotic preserving scheme for the diffusive limit of kinetic systems for chemotaxis*, Multiscale Model. Simul., 11 (2013), pp. 336–361, https://doi.org/10.1137/110851687.

[11] F. A. C. C. Chalub, P. A. Markowich, B. Perthame, and C. Schmeiser, *Kinetic models for chemotaxis and their drift-diffusion limits*, Monatsh. Math., 142 (2004), pp. 123–141.

[12] A. Chertock, A. Kurganov, M. Lukáčová-Medvid'ová, and S. Nur Özcan, *An asymptotic preserving scheme for kinetic chemotaxis models in two space dimensions*, Kinet. Relat. Models, 12 (2019), pp. 195–216.
[13] A. Chertock, A. Kurganov, X. Wang, and Y. Wu, *On a chemotaxis model with saturated chemotactic flux*, Kinet. Relat. Models, 5 (2012), pp. 51–95.
[14] N. Crouseilles, S. Jin, M. Lemou, and L. Liu, *Nonlinear Geometric Optics Based Multiscale Stochastic Galerkin Methods for Highly Oscillatory Transport Equations with Random Inputs*, preprint, https://arxiv.org/abs/1704.01019, 2017.
[15] P. Degond, *Asymptotic-preserving schemes for fluid models of plasmas*, in Numerical Models for Fusion, Panor. Synthèses 39/40, Soc. Math. France, 2013, pp. 1–90.
[16] P. Degond and F. Deluzet, *Asymptotic-preserving methods and multiscale models for plasma physics*, J. Comput. Phys., 336 (2017), pp. 429–457.
[17] G. Dimarco and L. Pareschi, *Numerical methods for kinetic equations*, Acta Numer., 23 (2014), pp. 369–520.
[18] J. Dolbeault and B. Perthame, *Optimal critical mass in the two dimensional Keller–Segel model in* $\mathbb{R}^2$, C. R. Math. Acad. Sci. Paris, 339 (2004), pp. 611–616.
[19] J. Duoandikoetxea Zuazo, *Fourier Analysis*, translated from the Spanish by D. Cruz-Uribe, Grad. Stud. Math. 29, Amer. Math. Soc., 2001.
[20] M. A. Herrero and J. J. L. Velázquez, *Chemotactic collapse for the Keller-Segel model*, J. Math. Biol., 35 (1996), pp. 177–194.
[21] T. Hillen and H. G. Othmer, *The diffusion limit of transport equations derived from velocity-jump processes*, SIAM J. Appl. Math., 61 (2000), pp. 751–775, https://doi.org/10.1137/S0036139999358167.
[22] T. Hillen and K. J. Painter, *A user's guide to PDE models for chemotaxis*, J. Math. Biol., 58 (2009), 183.
[23] D. Horstmann et al., *From* 1970 *until present: The Keller-Segel model in chemotaxis and its consequences*, I. Jahresber. Deutsch. Math.-Verein., 105 (2003), pp. 103–165.
[24] J. Hu and S. Jin, *A stochastic Galerkin method for the Boltzmann equation with uncertainty*, J. Comput. Phys., 315 (2016), pp. 150–168.
[25] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., 21 (1999), pp. 441–454, https://doi.org/10.1137/S1064827598334599.
[26] S. Jin, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: A review*, Riv. Math. Univ. Parma (N.S.), 3 (2012), pp. 177–216.
[27] S. Jin, J.-G. Liu, and Z. Ma, *Uniform spectral convergence of the stochastic Galerkin method for the linear transport equations with random inputs in diffusive regime and a micro–macro decomposition-based asymptotic-preserving method,* Res. Math. Sci., 4 (2017), 15, https://doi.org/10.1186/s40687-017-0105-1.
[28] S. Jin and L. Liu, *An asymptotic-preserving stochastic Galerkin method for the semiconductor Boltzmann equation with random inputs and diffusive scalings*, Multiscale Model. Simul., 15 (2017), pp. 157–183, https://doi.org/10.1137/15M1053463.
[29] S. Jin and H. Lu, *An asymptotic-preserving stochastic Galerkin method for the radiative heat transfer equations with random inputs and diffusive scalings*, J. Comput. Phys., 334 (2017), pp. 182–206.
[30] S. Jin, H. Lu, and L. Pareschi, *Efficient Stochastic Asymptotic-Preserving Imex Methods for Transport Equations with Diffusive Scalings and Random Inputs*, preprint, https://arxiv.org/abs/1703.03841, 2017.
[31] S. Jin, L. Pareschi, and G. Toscani, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numer. Anal., 38 (2000), pp. 913–936, https://doi.org/10.1137/S0036142998347978.
[32] S. Jin, D. Xiu, and X. Zhu, *Asymptotic-preserving methods for hyperbolic and transport equations with random inputs and diffusive scalings*, J. Comput. Phys., 289 (2015), pp. 35–52.
[33] S. Jin and Y. Zhu, *Hypocoercivity and Uniform Regularity for the Vlasov-Poisson-Fokker-Planck System with Uncertainty and Multiple Scales*, preprint, https://arxiv.org/abs/1704.00208, 2017.
[34] E. F. Keller and L. A. Segel, *Initiation of slime mold aggregation viewed as an instability*, J. Theoret. Biol., 26 (1970), pp. 399–415.
[35] E. F. Keller and L. A. Segel, *Model for chemotaxis*, J. Theoret. Biol., 30 (1971), pp. 225–234.
[36] E. F. Keller and L. A. Segel, *Traveling bands of chemotactic bacteria: A theoretical analysis*, J. Theoret. Biol., 30 (1971), pp. 235–248.

[37] E. F. Keller, *Assessing the Keller-Segel model: How has it fared?*, in Biological Growth and Spread, Springer, 1980, pp. 379–387.
[38] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Springer Science+Business Media, 1992.
[39] Q. Li and L. Wang, *Uniform regularity for linear kinetic equations with random input based on hypocoercivity*, SIAM/ASA J. Uncertainty Quantification, 5 (2017), pp. 1193–1219, https://doi.org/10.1137/16M1106675.
[40] T. Nagai, *Global existence of solutions to a parabolic system for chemotaxis in two space dimensions*, Nonlinear Anal., 30 (1997), pp. 5381–5388.
[41] H. G. Othmer, S. R. Dunbar, and W. Alt, *Models of dispersal in biological systems*, J. Math. Biol., 26 (1988), pp. 263–298.
[42] H. G. Othmer and T. Hillen, *The diffusion limit of transport equations* II: *Chemotaxis equations*, SIAM J. Appl. Math., 62 (2002), pp. 1222–1250, https://doi.org/10.1137/S0036139900382772.
[43] L. Pareschi and G. Russo, *Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation*, J. Sci. Comput., 25 (2005), pp. 129–155.
[44] C. S. Patlak, *Random walk with persistence and external bias*, Bull. Math. Biol., 15 (1953), pp. 311–338.
[45] B. Perthame, *PDE models for chemotactic movements: Parabolic, hyperbolic and kinetic*, Appl. Math., 49 (2004), pp. 539–564.
[46] B. Perthame, *Transport Equations in Biology*, Springer Science+Business Media, 2006.
[47] M. Sharifi tabar, *One-dimensional chemotaxis kinetic model*, NoDEA Nonlinear Differential Equations Appl., 18 (2011), pp. 139–172.
[48] R. Shu, J. Hu, and S. Jin, *A stochastic Galerkin method for the Boltzmann equation with multi-dimensional random inputs using sparse wavelet bases*, Numer. Math. Theory Methods Appl., 10 (2017), pp. 465–488.
[49] H. G. Othmer and A. Stevens, *Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks*, SIAM J. Appl. Math., 57 (1997), pp. 1044–1081, https://doi.org/10.1137/S0036139995288976.
[50] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
[51] D. Xiu and G. E. Karniadakis, *The Wiener–Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619–644, https://doi.org/10.1137/S1064827501387826.
[52] Y. Zhu and S. Jin, *The Vlasov–Poisson–Fokker–Planck system with uncertainty and a one-dimensional asymptotic preserving method*, Multiscale Model. Simul., 15 (2017), pp. 1502–1529, https://doi.org/10.1137/16M1090028.