SCIENTIFIC REPORTS

OPEN

Free energy of a chemotactic model with nonlinear diffusion

Seung Ki Baek[1] & Beom Jun Kim[2]

The Patlak-Keller-Segel equation is a canonical model of chemotaxis to describe self-organized aggregation of organisms interacting with chemical signals. We investigate a variant of this model, assuming that the organisms exert effective pressure proportional to the number density. From the resulting set of partial differential equations, we derive a Lyapunov functional that can also be regarded as the free energy of this model, and minimize it with a Monte Carlo method to detect the condition for self-organized aggregation. Focusing on radially symmetric solutions on a two-dimensional disc, we find that the chemical interaction competes with diffusion so that aggregation occurs when the relative interaction strength exceeds a certain threshold. Based on the analysis of the free-energy landscape, we argue that the transition from a homogeneous state to aggregation is abrupt yet continuous.

Ants communicate with each other through the use of pheromones to adjust their collective behaviour[1–3]. This mechanism often leads to intriguing self-organized patterns. For example, their foraging path can be understood as solving a certain optimization problem in terms of time and energy costs[4–9], and the shape of the path is predictable by Fermat's principle of least time[10–12]. From a biological point of view, especially in the context of natural selection, it is highly plausible that an ant colony benefits from the ability of organizing a foraging path. It is also worth noting that the key ingredient is not an individual ant with little computational capacity, but the interaction in a group of such ants. It is thus regarded as an example of emergent phenomena[13] and the term 'swarm intelligence' has been coined to describe this idea. Various computational techniques can be categorized as based on swarm intelligence (see, e.g., refs 14 and 15). From a physical point of view, ants provide a good example of active matter[16], which can aggregate[17] or circulate[6] spontaneously and exhibit peculiar mechanical properties[18].

The Patlak-Keller-Segel equation is a canonical starting point to study organisms that interact by means of chemical attractants[19, 20]. This model treats the density of organisms $\rho(\mathbf{r}, t)$ and the concentration of chemical attractants $c(\mathbf{r}, t)$ as continuous variables, where $\mathbf{r}$ denotes spatial coordinates and t means time, and describes the interplay between them. The Patlak-Keller-Segel equation has been extensively studied by mathematicians and a variety of review papers are available (see, e.g., refs 21 and 22). One of characteristic features of this model is that the organisms can form a dense aggregate, developing a δ-function peak within a finite time, when the space has dimensionality $d > 1$. Although such a 'blow-up' phenomenon provides an approximate description for biological aggregation, it is not entirely realistic that the whole population collapses to a single point. Researchers have suggested various mechanisms to regularize this singularity: To name a few, there are density-dependent chemotactic sensitivity[23–26], nonlinear diffusion[27, 28], logistic damping[29], cross diffusion[30], and shear flows[31]. One may also refer to a review by Hillen and Painter[32] for many variations of the classical Patlak-Keller-Segel model. One may also refer to ref. 33 to see how it can be used to describe the organization of a foraging path.

This work adopts the idea of nonlinear diffusion[27, 28] to take into account the finite volume of the organisms, and analyse its consequences. Let us write down the following set of equations:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (-\chi_0 \rho \nabla c + D_0 \rho \nabla \rho) \tag{1}$$

$$\frac{\partial c}{\partial t} = f_0 \rho + \nu_0 \nabla^2 c - g_0 c, \tag{2}$$

where χ_0, D_0, f_0, ν_0, and g_0 are positive constants. The terms on the right-hand side of Eq. (1) represent chemotactic movement and nonlinear diffusion, respectively. On the other hand, the three terms on the right-hand side of Eq. (2) mean generation, diffusion, and degradation, respectively. According to the original derivation[27], the

[1]Department of Physics, Pukyong National University, Busan, 48513, Korea. [2]Department of Physics, Sungkyunkwan University, Suwon, 16419, Korea. Correspondence and requests for materials should be addressed to S.K.B. (email: seungki@pknu.ac.kr) or B.J.K. (email: beomjun@skku.edu)

nonlinear diffusion term derives from $\rho \nabla h(\rho)$ with a pressure function $h(\rho)$ due to crowding. If the pressure is expanded as a power series of density, as in the virial expansion, the choice of $h(\rho) \propto \rho$ corresponds to the lowest-order approximation, because the zeroth order clearly vanishes as $h(\rho = 0) = 0$. Some numerical observations have been reported in this case[28, 32]. Although $h(\rho)$ is effective pressure to describe collective motion phenomenologically, it is interesting to note that an ant aggregate has an elastic modulus, which has units of pressure, as a linear function of ρ, until the ants are so densely packed that their legs are compressed[18]. Note that the classical Patlak-Keller-Segel equation is interpreted as $h(\rho) \sim \ln \rho$ from this viewpoint.

In this work, we show that the system described by Eqs (1) and (2) has a Lyapunov functional whose time derivative is smaller than or equal to zero all the time. It will also be called the free energy on the analogy with statistical mechanics. In general, a Lyapunov functional is a powerful tool in analysing a dynamical system, and its existence can be utilized to study properties of a fixed point beyond the local stability analysis[34]. After examining two stationary states, of which one is homogeneous and the other is not, we investigate the Lyapunov functional in the normal-mode coordinates to examine the transition between the homogeneous and inhomogeneous states, restricting ourselves to radially symmetric solutions. We will minimize the Lyapunov functional with a Monte Carlo method because it is computationally efficient in studying long-time behaviour of the system. We then briefly check if the Monte Carlo results are consistent with those from the direct numerical integration of the partial differential equations. After characterizing the transition based on the free-energy landscape, we conclude this work.

Analysis

In this section, we begin with deriving the Lyapunov functional of Eqs (1) and (2). We are interested in homogeneous and inhomogeneous solutions and a transition between them. Of course, their stability can be studied in a standard way by adding small perturbation with the lowest nonzero mode, as will be demonstrated below. However, our main point is that the transition from the homogeneous distribution to aggregation can be analysed in detail by means of the Lyapunov functional, which contains the full spectrum of possible modes in this system.

Lyapunov functional. Before proceeding, we have to specify the boundary conditions of our model. In analysing Eqs (1) and (2), we consider a two-dimensional disc of radius l and choose the Neumann boundary conditions,

$$\frac{\partial \rho}{\partial r} = \frac{\partial c}{\partial r} = 0 \tag{3}$$

at $r = 0$ and $r = l$, where $r \equiv |\mathbf{r}|$ is the distance from the origin of the disc. This condition means that the organisms cannot enter or escape from the system across the boundary, which is the experimental situation under consideration. In other words, Eq. (1) is derived from a continuity equation with current $\mathbf{j} = -\chi_0 \rho \nabla c + D_0 \rho \nabla \rho$, which implies that it conserves the total mass of the organisms:

$$M = \int_0^{2\pi} \int_0^l \rho(r, \theta) r \, dr \, d\theta = \int \rho(r, \theta) dV, \tag{4}$$

where θ means the angle in the polar coordinates and dV is a volume element.

If we assume that the chemical attractant reaches a stationary state very quickly, so that the left-hand side of Eq. (2) can be taken to be approximately zero, we can solve the equation for c[35]. Let us consider the entire two-dimensional space for simplicity. The formal solution is then given as

$$c(\mathbf{x}) = -\frac{f_0}{\nu_0} \int d\mathbf{y} \mathcal{G}(\mathbf{x} - \mathbf{y}) \rho(\mathbf{y}), \tag{5}$$

where $\mathcal{G}$ is the Green function obtained in terms of K_0, the modified Bessel function of the second kind, as follows:

$$\mathcal{G}(\mathbf{x} - \mathbf{y}) = -\frac{1}{2\pi} K_0(\kappa |\mathbf{x} - \mathbf{y}|) \tag{6}$$

with $\kappa \equiv \sqrt{f_0/\nu_0}$. Plugging this into Eq. (1), we find that

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{E}}{\delta \rho} \right) \tag{7}$$

with

$$\mathcal{E} \equiv \frac{D_0}{2} \int \rho^2(\mathbf{x}) d\mathbf{x} + \frac{f_0 \chi_0}{2\nu_0} \iint \rho(\mathbf{x}) \mathcal{G}(\mathbf{x} - \mathbf{y}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \tag{8}$$

Note that the first term is equivalent to the participation ratio in the localization problem[36], and the second term can be interpreted as interaction energy between organisms at a distance. The participation ratio is minimized when ρ is distributed homogeneously, whereas the effective interaction potential, Eq. (6), make the organisms attract each other. If diffusion is dominant, i.e., $D_0 \nu_0 \gg f_0 \chi_0$, the interaction term becomes negligible and the

aggregation mediated by the chemical attractants will be suppressed. From Eqs ([7]) and ([8]), it is straightforward to see that

$$\frac{d\mathcal{E}}{dt} = -\int \left|\nabla\frac{\delta\mathcal{E}}{\delta\rho}\right|^2 \rho(\mathbf{x})d\mathbf{x},$$

(9)

which implies that $\mathcal{E}$ never increases as time goes by.

We have derived Eq. ([8]) under the restriction that $\partial c/\partial t = 0$ only because $\mathcal{E}$ provides a simple physical interpretation in terms of ρ only. In fact, it is possible to construct a complete Lyapunov functional without such a restriction: Let us rescale the variables as $\tau = D_0 t$ and $c' = \frac{\chi_0}{D_0}c$. to rewrite Eqs ([1]) and ([2]) as

$$\frac{\partial\rho}{\partial\tau} = \nabla \cdot (-\rho\nabla c' + \rho\nabla\rho) = \nabla \cdot (\rho\nabla Z)$$

(10)

$$\frac{\chi_0}{\nu_0}\frac{\partial c'}{\partial\tau} = \nabla^2 c' - \frac{g_0}{\nu_0}c' + \frac{f_0\chi_0}{D_0\nu_0}\rho,$$

(11)

where $Z \equiv \rho - c'$. We can show that

$$\int dV\, Z\frac{\partial\rho}{\partial\tau} = \int dV\,\nabla \cdot (Z\rho\nabla Z) - \int dV\,\rho\,|\nabla Z|^2,$$

(12)

where the first term on the right-hand side vanishes due to the boundary conditions. By using Eq. ([12]), we can also show the following:

$$\frac{d}{d\tau}\int dV\,\rho Z = \int dV\,\rho\frac{\partial Z}{\partial\tau} + \int dV\,Z\frac{\partial\rho}{\partial\tau}$$

(13)

$$= \frac{d}{d\tau}\int dV\,\frac{\rho^2}{2} - \int dV\,\rho\frac{\partial c'}{\partial\tau} - \int dV\,\rho\,|\nabla Z|^2.$$

(14)

In addition, we have the following equality:

$$0 = \int dV\,\nabla \cdot \left(\frac{\partial c'}{\partial\tau}\nabla c'\right)$$

(15)

$$= \frac{\chi_0}{\nu_0}\int dV\left(\frac{\partial c'}{\partial\tau}\right)^2 + \frac{g_0}{\nu_0}\int dV\,c'\frac{\partial c'}{\partial\tau} - \frac{f_0\chi_0}{D_0\nu_0}\int dV\,\rho\frac{\partial c'}{\partial\tau} + \frac{d}{d\tau}\int dV\,\frac{|\nabla c'|^2}{2}.$$

(16)

Plugging Eq. ([14]) into Eq. ([16]), we get

$$-\frac{dW}{d\tau} = \frac{\chi_0}{\nu_0}\int dV\left(\frac{\partial c'}{\partial\tau}\right)^2 + \frac{f_0\chi_0}{D_0\nu_0}\int dV\,\rho\,|\nabla Z|^2,$$

(17)

where

$$W \equiv \frac{f_0\chi_0}{D_0\nu_0}\int dV\left(\frac{1}{2}\rho^2 - \rho c'\right) + \frac{g_0}{2\nu_0}\int dV\,|c'|^2 + \int dV\,\frac{|\nabla c'|^2}{2}.$$

(18)

It is clear from Eq. ([17]) that $dW/d\tau$ cannot be positive so that W does not increase when the system evolves according to Eqs ([1]) and ([2]). For this reason, this quantity is sometimes called the *free energy* of this system. The time derivative $dW/d\tau$ equals zero if $\partial c'/\partial\tau = 0$ and $\mathbf{j} \propto \nabla Z = 0$ everywhere that $\rho > 0$. The first integral of Eq. ([18]) consists of the participation ratio and the potential energy due to the coupling between ρ and c, whereas the other two integrals describe the chemical energy[37]. Likewise, one can argue that Eq. ([17]) contains the chemical production term $\propto(\partial c/\partial t)^2$ on its right-hand side, and that the last term corresponds to something referred to as entropy production in the classical Patlak-Keller-Segel model because it is related to the time derivative of the Shannon entropy[37]. In our nonlinear-diffusion model, the last term of Eq. ([17]) may be regarded as generalized entropy production in terms of the Tsallis entropy[38]. It is also worth noting that the integrands in Eq. ([18]) are all quadratic, which will turn out to be useful for our analysis.

Linear stability of a homogeneous stationary solution. Equations ([1]) and ([2]) admit a homogeneous stationary solution $\rho = \frac{g_0}{f_0}c = \rho_{const}$, where $\rho_{const} = M/(\pi l^2)$ from Eq. ([4]). In this state, Eq. ([18]) yields

$$W = \frac{M^2}{2\pi l^2}\frac{f_0\chi_0}{D_0\nu_0}\left(1 - \frac{f_0\chi_0}{D_0 g_0}\right).$$

(19)

The standard linear stability analysis assumes small perturbations ε_ρ and ε_c around this homogeneous solution to assume $\rho(\mathbf{r}, t) = \rho_{\text{const}} + \varepsilon_\rho(\mathbf{r}, t)$ and $c(\mathbf{r}, t) = \frac{f_0}{g_0}\rho_{\text{const}} + \varepsilon_c(\mathbf{r}, t)$. By collecting linear terms in ε_ρ and ε_c, we obtain

$$\frac{\partial}{\partial t}\begin{pmatrix}\varepsilon_\rho \\ \varepsilon_c\end{pmatrix} = \begin{pmatrix} 0 & 0 \\ f_0 & -g_0 \end{pmatrix}\begin{pmatrix}\varepsilon_\rho \\ \varepsilon_c\end{pmatrix} + \begin{pmatrix} D_0\rho_{\text{const}} & -\chi_0\rho_{\text{const}} \\ 0 & \nu_0 \end{pmatrix}\nabla^2\begin{pmatrix}\varepsilon_\rho \\ \varepsilon_c\end{pmatrix}. \tag{20}$$

Suppose that the perturbations are described as cylindrical harmonics, satisfying the following equation:

$$(\nabla^2 + k^2)\begin{pmatrix}\varepsilon_\rho \\ \varepsilon_c\end{pmatrix} = 0. \tag{21}$$

Each mode then takes the form of $J_n(kr)e^{\pm in\theta}e^{\eta t}$, where J_n means the Bessel function and η is its growth rate. The Neumann boundary conditions are expressed as $\frac{\partial}{\partial r}J_n(kl) = 0$. The lowest mode is thus found at $n = 0$, which means radially symmetric density fluctuations concentrated around the origin. The first zero of J_1 is located at $kl \approx 3.832\ldots$. If we solve the resulting eigenvalue problem:

$$\eta^2 + [k^2(D_0\rho_{\text{const}} + \nu_0) + g_0]\eta + \rho_{\text{const}}[k^2D_0(g_0 + k^2\nu_0) - k^2f_0\chi_0] = 0, \tag{22}$$

the stability condition is obtained as $k^2D_0(g_0 + k^2\nu_0) - k^2f_0\chi_0 > 0$. Note that it is independent of ρ_{const}, differently from the classical Patlak-Keller-Segel model[39], so that the system does not need critical mass for instability. This feature is, however, due to our particular choice of nonlinear diffusion. We find a necessary condition for the lowest mode to grow in time as follows:

$$k^2l^2 \approx 14.684 < \left(\frac{f_0\chi_0}{D_0\nu_0} - \frac{g_0}{\nu_0}\right)l^2 = K^2l^2, \tag{23}$$

where

$$K \equiv \sqrt{\frac{f_0\chi_0}{D_0\nu_0} - \frac{g_0}{\nu_0}}. \tag{24}$$

If we assume that $g_0 \ll 1$, the expression inside the square root of Eq. (24) is interpreted as a ratio between chemotactic strength and diffusivity. This small-g_0 limit is often plausible without altering the essential physics, because some ant pheromones last for days[40]. Equation (23) suggests that Kl will be an important dimensionless parameter that governs the aggregation phenomenon.

In addition, if the disc is so large that the boundary effects are negligible and there is a continuous spectrum of possible wavenumbers, the initial stage of instability from the homogeneous solution is governed by the most unstable mode with $k = k_u$ such that maximizes the positive η[26]. The wavenumber k_u can be expressed by the following formula:

$$k_u^2 = f_0\chi_0\sqrt{\frac{\rho_{\text{const}}}{D_0\nu_0}}\left(\frac{1}{\sqrt{D_0\rho_{\text{const}}} + \sqrt{\nu_0}}\right)^2, \tag{25}$$

where we take the limit of $g_0 \rightarrow 0$ to simplify the expression. Equation (25) will determine the typical length scale between aggregates, when the homogeneous initial state becomes unstable.

Inhomogeneous stationary solution. Let us now consider a radially symmetric stationary aggregate. The boundary conditions make the flux vanish everywhere, i.e., $\mathbf{j} = -\chi_0\rho\nabla c + D_0\rho\nabla\rho = 0$. It implies that

$$\rho = \frac{\chi_0}{D_0}(c - c_0) \tag{26}$$

with a constant of integration c_0. Substituting Eq. (26) into Eq. (2) with the stationarity condition, we obtain an inhomogeneous Helmholtz equation:

$$0 = \frac{f_0\chi_0}{D_0}(c - c_0) + \nu_0\nabla^2c - g_0c, \tag{27}$$

which has the following radially symmetric solution:

$$c(r) = AJ_0(Kr) + \frac{f_0\chi_0}{K^2D_0\nu_0}c_0, \tag{28}$$

where A is a constant describing the amplitude of aggregation, J_n is the Bessel function, and the wavenumber K has been defined in Eq. (24) above. Obviously, the solution is feasible only when the boundary condition is satisfied by $\frac{d}{dr}J_0(Kr)\big|_{r=l} = -KJ_1(Kl) = 0$, and let us suppose that this is the case. The constant A is bounded by a condition that both ρ and c must be non-negative everywhere. If we plug Eq. (28) into Eq. (26), we find that

$$\rho(r) = \frac{\chi_0}{D_0}\left[AJ_0(Kr) + \left(\frac{f_0\chi_0}{K^2 D_0 \nu_0} - 1\right)c_0\right].$$

(29)

The unknown constant c_0 can be explicitly determined from Eq. (4) because $\int_0^l J_0(Kr)r\,dr = 0$ as long as the boundary conditions are satisfied. After some algebra, we can write the results as

$$\rho = \frac{\chi_0}{D_0}AJ_0 + \rho_{\text{const}}$$

(30)

$$c = AJ_0 + c_{\text{const}},$$

(31)

where ρ_{const} and c_{const} define the homogeneous solution. We substitute these results into Eq. (18) to calculate the Lyapunov functional:

$$
\begin{aligned}
W &= \frac{f_0\chi_0}{D_0\nu_0}\int dV\left[\frac{1}{2}\left(\frac{\chi_0}{D_0}AJ_0 + m\right)^2 - \left(\frac{\chi_0}{D_0}AJ_0 + m\right)\frac{\chi_0}{D_0}\left(AJ_0 + \frac{f_0\chi_0}{K^2 D_0 \nu_0}c_0\right)\right] \\
&\quad + \frac{g_0}{2\nu_0}\int dV\left[\frac{\chi_0^2}{D_0^2}\left(AJ_0 + \frac{f_0\chi_0}{K^2 D_0 \nu_0}c_0\right)^2\right] + \frac{1}{2}\int dV\frac{\chi_0^2}{D_0^2}(KAJ_1)^2
\end{aligned}
$$

(32)

$$
\begin{aligned}
&= -\frac{f_0\chi_0^3 A^2}{2D_0^3\nu_0}\int dV J_0^2 + \left(\frac{1}{2}\frac{f_0\chi_0}{D_0\nu_0}m^2 - \frac{f_0^2\chi_0^3}{K^2 D_0^3 \nu_0^2}mc_0\right)\pi l^2 \\
&\quad + \frac{g_0\chi_0^2 A^2}{2D_0^2\nu_0}\int dV J_0^2 + \frac{f_0^2\chi_0^4 g_0}{2K^4 D_0^4 \nu_0^3}c_0^2\pi l^2 + \frac{\chi_0^2 K^2 A^2}{2D_0^2}\int dV J_1^2.
\end{aligned}
$$

(33)

We can see that the three integrals on the last line vanish altogether, if we note the definition of K [Eq. (24)] and the following identity:

$$
\begin{aligned}
K^2\int_0^l rJ_0^2(Kr)\,dr &= K\int_0^l \frac{d}{dr}\left[rJ_1(Kr)\right]J_0(Kr)\,dr \\
&= -K\int_0^l rJ_1(Kr)\frac{d}{dr}J_0(Kr)\,dr \\
&= K^2\int_0^l rJ_1(Kr)^2\,dr,
\end{aligned}
$$

(34)

which is valid under our assumption that $J_1(Kl) = 0$. As a result, we obtain

$$W = \frac{M^2}{2\pi l^2}\frac{f_0\chi_0}{D_0\nu_0}\left(1 - \frac{f_0\chi_0}{g_0 D_0}\right),$$

(35)

which is identical to the Lyapunov functional of the homogeneous solution [Eq. (19)]. It is consistent with the fact that the solution with K has neutral stability in the linear-stability analysis [see, e.g., Eq. (23)], according to which the radially symmetric mode $\propto J_0(kr)$ can survive only when k is smaller than K. Although we have assumed that the wavenumber K is compatible with the boundary condition, it is actually independent of l, which implies that the stationarity condition cannot be met exactly. If a perturbative mode with $k < K$ appears from the homogeneous state with satisfying the boundary conditions, therefore, it cannot be stationary: Its amplitude will grow exponentially at first, but cannot become arbitrarily large because of the non-negativity of ρ and c. The growth will stop when A reaches the largest value that does not violate the non-negativity. This scenario seems to suggest a jump in A as K crosses a threshold, and this scenario will be scrutinized below by considering a full spectrum of normal modes.

Normal-mode expansion. Let us decompose ρ and c into normal modes:

$$\rho(r, \theta, t) = \rho_{\text{const}} + \sum_{p=0}^{\infty}\sum_{m=1}^{\infty} J_p(j'_{p,m}r/l)\left[E_{pm}(t)\cos p\theta + F_{pm}(t)\sin p\theta\right]$$

(36)

$$c(r, \theta, t) = c_{\text{const}} + \sum_{p=0}^{\infty}\sum_{m=1}^{\infty} J_p(j'_{p,m}r/l)\left[G_{pm}(t)\cos p\theta + H_{pm}(t)\sin p\theta\right],$$

(37)

where j'_{pm} denotes the mth zero of $\frac{d}{dx}J_p(x)$. Note that

$$\int_0^{2\pi} \int_0^l r J_p(j'_{pm} r/l)\, e^{ip\theta}\, dr\, d\theta = 0, \tag{38}$$

so that Eq. (36) automatically conserves the total mass $M = \int_0^{2\pi} \int_0^l \rho(r, \theta) r\, dr\, d\theta = \rho_{\mathrm{const}} \pi l^2$. Likewise, the total amount of the chemical attractant is given as $c_{\mathrm{const}} \pi l^2$, which is, however, a function of time in general. It is straightforward to see the following orthogonality relation

$$\int_0^l r J_p(j'_{pu} r/l) J_p(j'_{pw} r/l)\, dr = -\frac{l^2}{2} J_p(j'_{pu}) \frac{d^2}{dx^2} J_p(j'_{pu}) \delta_{uw} = -\frac{l^2}{2} \phi_{pu} \delta_{uw}, \tag{39}$$

where δ_{uw} is the Kronecker delta and $\phi_{pu} \equiv J_p(j'_{pu}) \frac{d^2}{dx^2} J_p(j'_{pu})$.

We will rewrite the Lyapunov functional [Eq. (18)] by using Eqs (36) and (37). The first term needs an integral of ρ^2 over the disc, which can be expressed as

$$\frac{1}{\pi l^2} \int \rho^2 dV = \rho_{\mathrm{const}}^2 - \sum_{m=1}^{\infty} \left[\phi_{0m} E_{0m}^2 + \frac{1}{2} \sum_{p=1}^{\infty} \phi_{pm} \left(E_{pm}^2 + F_{pm}^2 \right) \right] \tag{40}$$

by using the orthogonality relations. The integrals of ρc and c^2 can be done in a similar way. However, the last part of the Lyapunov functional [Eq. (18)] is more complicated: It is involved with an integral of $|\nabla c|^2$, which is decomposed into two terms:

$$\int |\nabla c|^2 dV = \int_0^{2\pi} \int_0^l \left| \frac{\partial c}{\partial r} \right|^2 r\, dr\, d\theta + \int_0^{2\pi} \int_0^l \frac{1}{r^2} \left| \frac{\partial c}{\partial \theta} \right|^2 r\, dr\, d\theta. \tag{41}$$

We again substitute Eqs (36) and (37) here to obtain

$$
\begin{aligned}
\int_0^{2\pi} \int_0^l \left| \frac{\partial c}{\partial r} \right|^2 r\, dr\, d\theta =\ & \pi \sum_{m=1}^{\infty} j_{1m}^2 J_0^2(j_{1m}) G_{0m}^2 \\
& + \frac{\pi}{l^2} \sum_{p=1}^{\infty} \sum_{mn} j'_{pm} j'_{pn} (G_{pm} G_{pn} + H_{pm} H_{pn}) \\
& \times \int_0^l r \left(\left. \frac{dJ_p(x)}{dx} \right|_{x=j'_{pm} r/l} \right) \left(\left. \frac{dJ_p(x)}{dx} \right|_{x=j'_{pn} r/l} \right) dr
\end{aligned}
\tag{42}
$$

and

$$
\begin{aligned}
\int_0^{2\pi} \int_0^l \frac{1}{r^2} \left| \frac{\partial c}{\partial \theta} \right|^2 r\, dr\, d\theta =\ & \pi \sum_{p=1}^{\infty} \sum_{mn} p^2 (G_{pm} G_{pn} + H_{pm} H_{pn}) \\
& \times \int_0^l \frac{1}{r} J_p(j'_{pm} r/l) J_p(j'_{pn} r/l)\, dr.
\end{aligned}
\tag{43}
$$

Note that the results still have the triple sums over p, m, and n, because we cannot enjoy the orthogonality between m and n when performing the integrals over r.

To circumvent the time-consuming evaluation of the triple sums, we focus on radially symmetric solutions by setting $p = 0$. If j_{pm} denotes the mth zero of $J_p(x)$, we can identify j'_{0m} with j_{1m} because $\frac{d}{dx} J_0(x) = -J_1(x)$. Therefore, Eq. (39) further simplifies to

$$\int_0^l r J_0(j_{1u} r/l) J_0(j_{1w} r/l)\, dr = \int_0^l r J_1(j_{1u} r/l) J_1(j_{1w} r/l)\, dr = \frac{l^2}{2} J_0^2(j_{1u}) \delta_{uw}, \tag{44}$$

where the first equality is derived in the same way as in Eq. (34), and the second one is the conventional orthogonality of the Bessel function[41]. Plugging Eqs (36) and (37) with $p = 0$ into the Lyapunov functional [Eq. (18)] and using the orthogonality, we find that
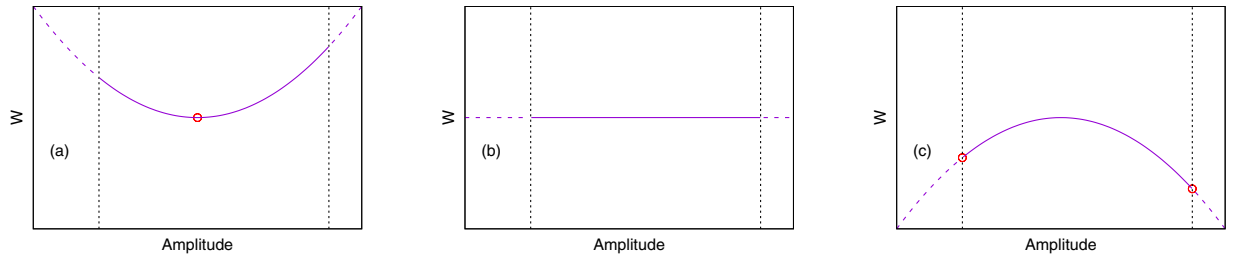
Figure 1. Sketches of the Lyapunov functional W along a principal axis, a combination of the amplitudes E_{01} and G_{01}, when (**a**) $Kl < j_{11}$, (**b**) $Kl = j_{11}$, and (**c**) $Kl > j_{11}$, respectively. The vertical dotted lines represent the physical constraint that both ρ and c should be non-negative, so that the system can explore only the landscapes of W drawn with solid lines. The small red circles show local minima of the given landscapes.

$$
\begin{aligned}
\frac{W}{\pi l^2} = & \; \frac{f_0 \chi_0}{D_0 \nu_0} \left[\frac{1}{2} \left(\rho_{\text{const}}^2 + \sum_{m=1}^{\infty} J_0^2(j_{1m}) E_{0m}^2 \right) \right. \\
& \left. - \frac{\chi_0}{D_0} \left(\rho_{\text{const}} c_{\text{const}} + \sum_{m=1}^{\infty} J_0^2(j_{1m}) E_{0m} G_{0m} \right) \right] \\
& + \frac{g_0 \chi_0^2}{2 D_0^2 \nu_0} \left(c_{\text{const}}^2 + \sum_{m=1}^{\infty} J_0^2(j_{1m}) G_{0m}^2 \right) \\
& + \frac{\chi_0^2}{2 D_0^2 l^2} \sum_{m=1}^{\infty} j_{1m}^2 J_0^2(j_{1m}) G_{0m}^2
\end{aligned}
\tag{45}
$$

$$
= \sum_{m=0}^{\infty} \frac{1}{2} J_0^2(j_{1m}) \left(E_{0m} \quad \frac{\chi_0}{D_0} G_{0m} \right)
\begin{pmatrix}
\dfrac{f_0 \chi_0}{D_0 \nu_0} & -\dfrac{f_0 \chi_0}{D_0 \nu_0} \\[2mm]
-\dfrac{f_0 \chi_0}{D_0 \nu_0} & \dfrac{g_0}{\nu_0} + \dfrac{j_{1m}^2}{l^2}
\end{pmatrix}
\begin{pmatrix}
E_{0m} \\[2mm]
\dfrac{\chi_0}{D_0} G_{0m}
\end{pmatrix},
\tag{46}
$$

where we have defined $E_{00} \equiv \rho_{\text{const}}$, $G_{00} \equiv c_{\text{const}}$, and $j_{10} \equiv 0$. We are interested in the minimum of Eq. (46), expecting that it captures the long-term behaviour of the system. The set of variables $\{c_{\text{const}}, E_{01}, E_{02}, \ldots, G_{01}, G_{02}, \ldots\}$ resulting from the minimization will be independent of the overall rescaling of W and thus determined by three dimensionless ratios, χ_0/D_0, g_0/f_0, and $\nu_0/(f_0 l^2)$. The first ratio measures the chemical sensitivity of the organism with respect to its nonlinear diffusivity. The next one measures the relative time scale between the generation and decay of the chemical attractant. Finally, the last one gives the typical time scale for the chemical attractant to diffuse into the whole system, measured with respect to the generation rate. Let us assume that each summand can be considered *separately* in this minimization problem. Then, for $m = 0$, only c_{const} varies, because ρ_{const} is fixed by the total mass M, and the optimal value for c_{const} equals $(f_0/g_0)\rho_{\text{const}}$ as we have already seen in the homogeneous stationary solution. For every other $m > 1$, we have a simple quadratic function of E_{0m} and G_{0m}. From an eigenvalue analysis, it is straightforward to see that the functional shape is elliptic when $j_{1m} > Kl$ and hyperbolic otherwise, where K is defined by Eq. (24). In the former case, the minimum is located at $E_{0m} = G_{0m} = 0$. In the latter case, the minima of Eq. (46) are found at $E_{0m} \propto G_{0m} = \pm\infty$, and the divergence must be regulated by the condition that both ρ and c are non-negative everywhere. The idea is sketched in Fig. 1 for $m = 1$. According to this argument, if Kl lies between j_{11} and j_{12}, for example, we will observe two local minima, one for $E_{01} \propto G_{01} > 0$ and the other for $E_{01} \propto G_{01} < 0$, while all the other E_{0m}'s and G_{0m}'s with $m > 1$ remain suppressed to zero. An interesting point in this picture is that the Lyapunov functional becomes independent of the amplitude of aggregation if Kl exactly equals j_{11}: An infinite number of states would have the same value of the Lyapunov functional. Therefore, even if the system converges to two different states as $Kl \to j_{11}^+$ and $Kl \to j_{11}^-$, respectively, there would be a continuous spectrum of states between them at $Kl = j_{11}$.

Numerical Results

Let us choose $\chi_0 = 4$ and set other parameters, ρ_{const}, D_0, ν_0, g_0, and l, to unity. With these parameters, the system reaches the threshold for aggregation, $Kl = j_{11}$, when $f_0 = f_0^* \approx 3.92$. We minimize the Lyapunov function for radially symmetric cases [Eq. (46)] with different values of f_0 by means of the Metropolis algorithm (see Method for details). In evaluating Eq. (46) numerically, we have to replace the infinite series by a partial sum, and the spatial resolution of the resulting expression will be enhanced as we include more and more modes in the summation. Here, let us use a partial sum up to $m = 19$ because it already captures the overall behaviour correctly. This choice implies that we have to work with 39 variables of c_{const}, E_{01}, …, G_{0m}. For the algorithm to search for the parameter space efficiently, we introduce a 'temperature' variable T, which helps the system escape from metastable local minima. We start with a sufficiently high temperature, say, $T = 10^1$, to explore a wide region of the
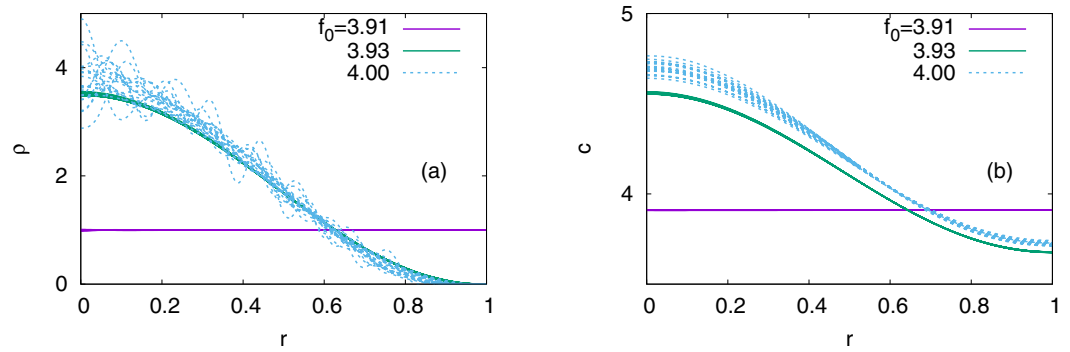
Figure 2. (**a**) Density of the organisms ρ and (**b**) the density of their chemical attractants c, obtained by minimizing a partial sum of Eq. (46) up to $m = 19$ with the Metropolis algorithm. We choose $\chi_0 = 4$, $\rho_{\mathrm{const}} = 1$, $D_0 = 1$, $\nu_0 = 1$, $g_0 = 1$, and $l = 1$. For each f_0, we run 20 independent samples, slowly lowering the 'temperature' from $T = 10$ to $T = 0$.
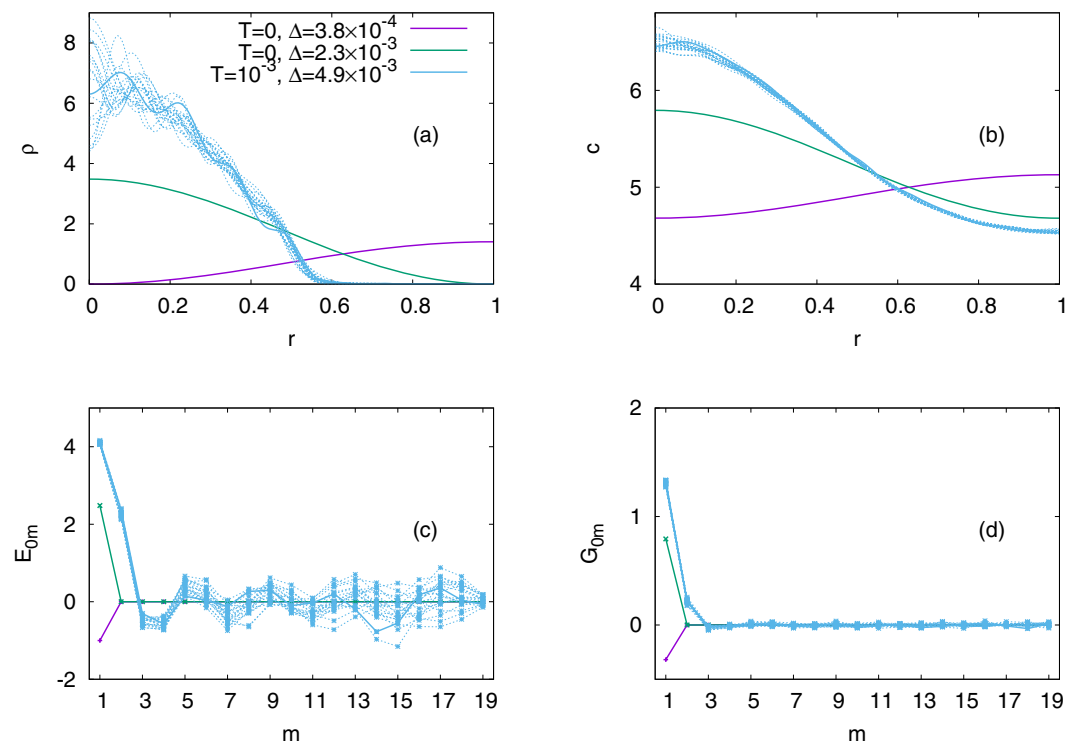

Figure 3. Minimization results of a partial sum of Eq. (46) up to $m = 19$, obtained by the Metropolis algorithm. In this plot, we show (**a**) the density of the organisms ρ, (**b**) that of the chemical attractants c, (**c**) the amplitudes E_{0m}'s for describing ρ, and (**d**) G_{0m}'s for c. We choose $f_0 = 5$, $\chi_0 = 4$, $\rho_{\mathrm{const}} = 1$, $D_0 = 1$, $\nu_0 = 1$, $g_0 = 1$, and $l = 1$. The initial condition is given by $c_{\mathrm{const}} = (f_0/g_0)\rho_{\mathrm{const}}$ and $E_{0m} = G_{0m} = 0$ in each case. For the zero-temperature case, i.e., $T = 0$, the system approaches either of two different local minima, represented by the purple and green lines, respectively. If we instead slowly cool down the system from $T = 10^1$ to $T \approx 10^{-3}$, we find high concentrations of ρ and c around $r = 0$ for all the 20 samples shown in this plot (the blue lines). Among the blue lines, the solid ones represent the sample with the best minimization result.

parameter space and then gradually lower the temperature down to $T = 0$. As argued above, we observe a sharp transition from a homogeneous solution to aggregation when f_0 exceeds $f_0^* \approx 3.92$, and the aggregation pattern is approximated to $J_0(j_{11}r/l)$ [Fig. 2]. From $f_0 = 3.93$ to $f_0 = 4.00$, on the other hand, the system remains qualitatively the same, although small variations exist from sample to sample. To sum up, the behaviour at $Kl \approx j_{11}$ is indeed explained by the assumption that the minimization of Eq. (46) can be carried out term by term.

As f_0 increases, however, the assumption loses validity. In Fig. 3, we plot our numerical minimization results with $f_0 = 5$ while all the other parameters are the same as above. Then, the value of $Kl \approx 4.3589$ still falls between $j_{11} \approx 3.8317$ and $j_{12} \approx 7.0156$. If W_{ref} denotes the value of the Lyapunov functional of the homogeneous solution,
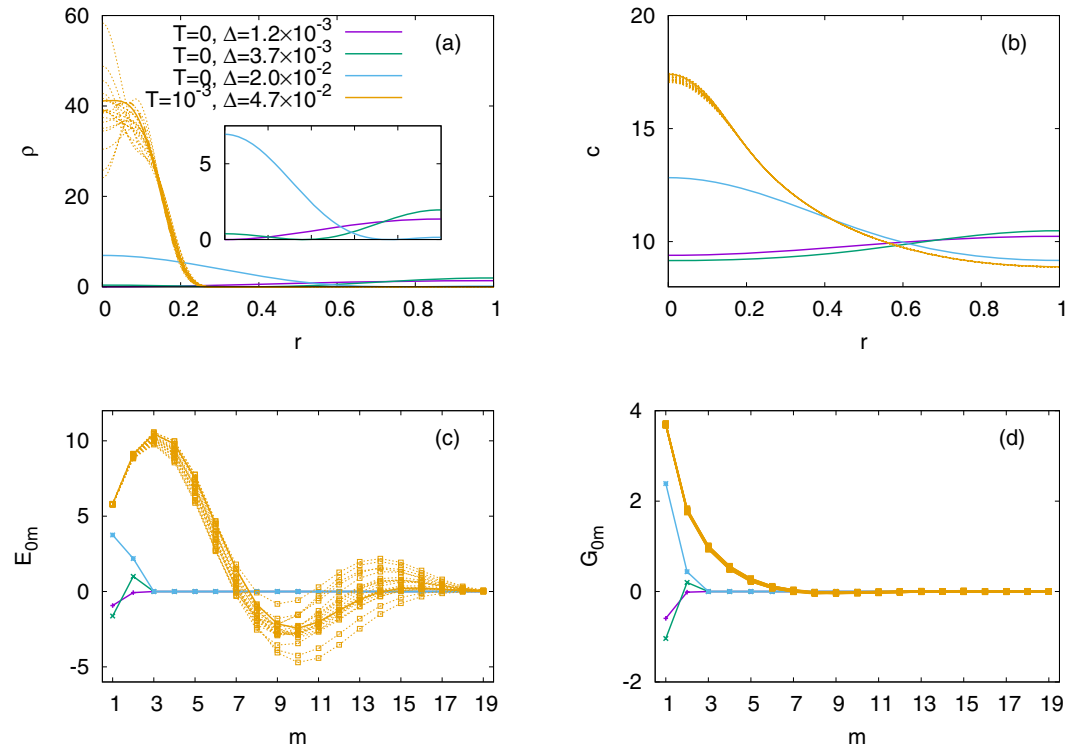
Figure 4. Minimization results of a partial sum of Eq. (46) up to $m = 19$, obtained by the Metropolis algorithm. We choose $f_0 = 10$ and $\chi_0 = 8$, and keep all the others the same as in Fig. 3. (**a**) The density of the organisms. (Inset) If we run the zero-temperature Metropolis algorithm starting with $E_{01} = E_{02} = \cdots = G_{01} = G_{02} = \cdots = 0$, the system approaches either of three different local minima, which are represented by the purple, green, and blue lines, respectively. We can also start from $T = 10$ and then cool down the system slowly. Performing this process with 20 independent samples, we plot their ρ at $T = 10^{-3}$ with the orange lines. Among the orange lines, the solid ones represent the sample with the best minimization result. The other panels show (**b**) the density of the chemical attractants, (**c**) the normal-mode amplitudes for ρ, and (**d**) those for c, respectively.

we see from Eq. (19) that $W_{\text{ref}}/\pi l^2 = -190$. To see the minimization performance, we check a relative difference from this value,

$$\Delta \equiv \frac{W_{\text{ref}} - W}{W_{\text{ref}}}. \tag{47}$$

We first run the Metropolis algorithm from $E_{0m} = G_{0m} = 0$ with fixing the temperature T to zero. We then find two different minima as expected: One describes a population concentration around $r = 0$, and the other shows an annular structure which is reminiscent of an ant mill[6]. These patterns nicely match with our picture in Fig. 1(c). Especially, the concentration around $r = 0$ is essentially the same pattern that we have shown in Fig. 2. However, if we start from $T = 10^1$ and gradually lower the temperature down to $T \approx 10^{-3}$, a better minimization result is achieved and it is characterized by systematic deviations of E_{0m} from zero for $m \leq 10$. The small yet finite temperature $T \approx 10^{-3}$ shows us how the modes are affected by environmental noises. Due to the excitation of high-m modes, we observe higher concentrations of ρ and c around the origin than expected from the zero-temperature case. Such coupling between modes would not be observed if Eq. (46) was minimized term by term. In Fig. 3, we see that E_{01} is considerably greater than that of the zero-temperature result. Higher modes with $m > 1$ should thus be excited to ensure the non-negativity of ρ, increasing W. Nevertheless, the reduction of W from $m = 1$ may well overtake the increment from $m > 1$, because each mode appears with a different weight in Eq. (46). The excitation of high-m modes becomes more pronounced as we go far above j_{11}: For example, let us choose $f_0 = 10.0$ and $\chi_0 = 8.0$, for which $Kl \approx 8.8882$ is greater than $j_{12} \approx 7.0156$ but lies below $j_{13} \approx 10.1735$. We observe that the zero-temperature Metropolis algorithm ends up with one of three different minima shown in Fig. 4(a). Once again, the annealing procedure from $T = 10^1$ to $T = 10^{-3}$ finds a much better result, concentrating the most of the population around $r = 0$. Note that the amplitudes E_{0m} exhibit a nontrivial structure in Fig. 4(c). It actually extends to even higher $m > 19$ if we take more modes into account in computing Eq. (46), but those higher modes hardly affect the radius of the aggregate in Fig. 4(a).

When the distribution $\rho(r)$ is given, the degree of aggregation can be estimated by the Shannon entropy:
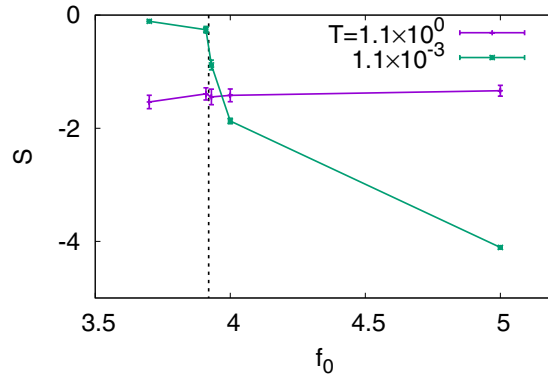
Figure 5. Shannon entropy [Eq. (48)] as a function of f_0. The other parameters are the same as in Figs 2 and 3. For each data point, we take an average over 20 independent samples. The vertical dotted line represents $f_0 = f_0^* \approx 3.92$ to make $Kl = j_{11}$.

$$S = -\int_0^{2\pi} \int_0^l \rho(r) \log \frac{\rho(r)}{\rho_{\text{const}}} r \, dr \, d\theta. \tag{48}$$

Figure 5 shows S as a function of f_0 at two different temperatures of the Monte Carlo calculation. The other parameters are set to the same as in Figs 2 and 3. When T is high, the system is insensitive to f_0, and S does not show any significant change. For low T, on the other hand, it becomes clear that a jump of S exists in the vicinity of $f_0 = f_0^*$. Recall that the separability assumption predicts that the system undergoes stepwise changes as f_0 increases, because Kl has to exceed j_{1m} to excite the mth mode ($m = 1, 2, \ldots$). That is, if the assumption was valid everywhere, all the higher modes would remain suppressed unless $Kl > j_{12}$, which requires $f_0 > 12.55$. However, our Monte Carlo results have shown that modes tend to be coupled to each other to reduce the free energy to a greater extent than predicted by the separability assumption. In other words, it implies that S jumps only once at f_0^* and then changes continuously for higher f_0, which is indeed the case in Fig. 5.

It is also instructive to directly consider dynamics of Eqs (1) and (2) for the following reason: The idea behind our Monte Carlo calculation is that the result can describe long-time behaviour of the real dynamics. As mentioned in Method, the algorithm checks the non-negativity of ρ and c as well as the change of W, so that a Monte Carlo move will be rejected if it violates the non-negativity, even if it decreases W. On the other hand, the dynamics of Eqs (1) and (2) does not have such rejection but only continues with $dW/dt \leq 0$ [Eq. (17)]. Therefore, one may well ask if the dynamics always confines the system in a physical region where both ρ and c are non-negative. Fortunately, the answer is yes, as has been proved in ref. 27. We can thus safely move on to the next question, i.e., whether the long-time behaviour is consistent with the Monte Carlo result. Under radial symmetry, the equations are written as

$$\frac{\partial \rho}{\partial t} = -\chi_0 \left(\frac{\rho}{r} \frac{\partial c}{\partial r} + \frac{\partial \rho}{\partial r} \frac{\partial c}{\partial r} + \rho \frac{\partial^2 c}{\partial r^2} \right) + D_0 \left[\frac{\rho}{r} \frac{\partial \rho}{\partial r} + \left(\frac{\partial \rho}{\partial r} \right)^2 + \rho \frac{\partial^2 \rho}{\partial r^2} \right] \tag{49}$$

$$\frac{\partial c}{\partial t} = f_0 \rho + \nu_0 \left(\frac{1}{r} \frac{\partial c}{\partial r} + \frac{\partial^2 c}{\partial r^2} \right) - g_0 c. \tag{50}$$

We can integrate these equations numerically, e.g., with the Forward-Time Central-Space (FTCS) method[42], and the results are given in Fig. 6. We still use the same parameters as in Figs 2 and 3 to have a threshold at $f_0 = f_0^* \approx 3.92$. As expected, both ρ and c become flatter as time goes by when $f_0 < f_0^*$ [Fig. 6(a,b)]. On the other hand, when $f_0 > f_0^*$ [Fig. 6(c,d)], ρ and c instead converge to inhomogeneous distribution functions, respectively, which exactly match with the ones in Fig. 2. Moreover, the time evolution undergoes critical slowing down as we approach f_0^*. It is consistent with the linear-stability analysis in which the eigenvalue governing the mode growth (or decay) vanishes at the threshold. However, we also note that the naive FTCS scheme becomes unstable at large t, violating the non-negativity condition. This must be a numerical artefact because, as mentioned above, the dynamics itself preserves the non-negativity of ρ and c[27]. A better alternative could be to utilise the operator-splitting scheme[43], incorporating exact solutions of the porous-medium equation (see, e.g., ref. 44).

Summary

In summary, we have investigated a variant of the Patlak-Keller-Segel model in which pressure is assumed to increase linearly with the density of the organisms [Eqs (1) and (2)]. We have derived its Lyapunov functional W in Eq. (18), which may also be called the free energy of this system. The linear stability analysis of the homogeneous solution predicts a jump in the amplitude of aggregation as a parameter K, defined in Eq. (24), exceeds j_{11}/l. We have checked this transition by using the exact Lyapunov functional, simplified for radially symmetric
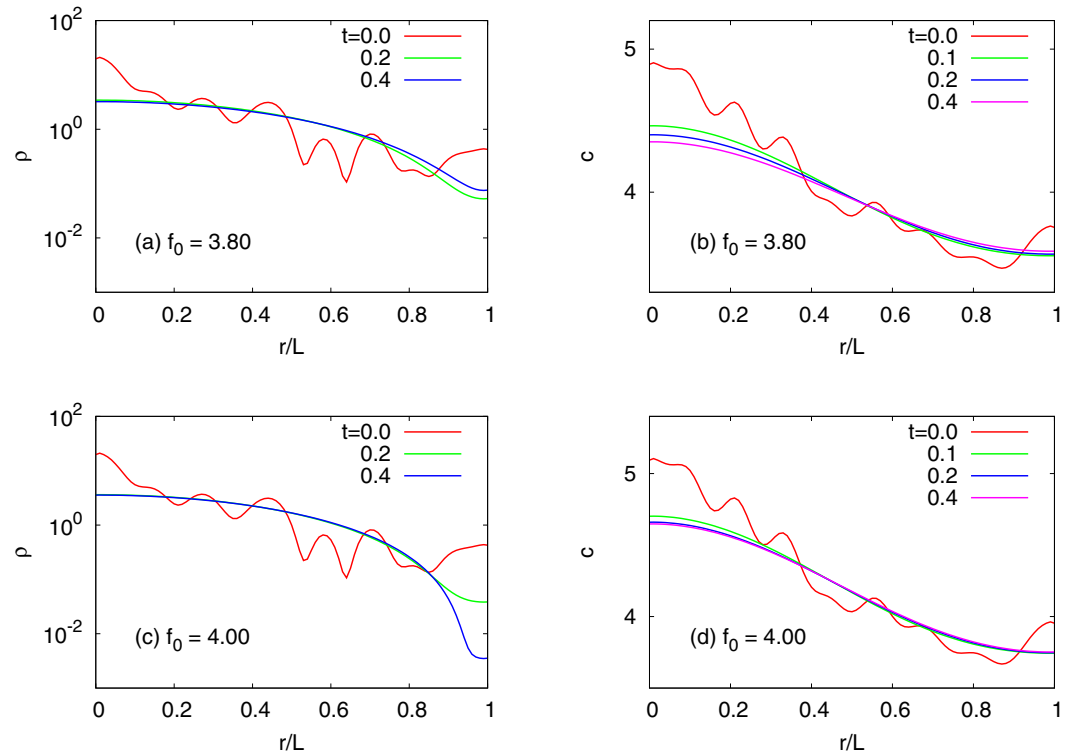
Figure 6. Direct numerical simulation of Eqs (1) and (2) under radial symmetry with the Forward-Time-Central-Space (FTCS) scheme. Panels (a,c) show $\rho(r, t)$ and the others do $c(r, t)$. We use the same parameter as in Figs 2 and 3, which means that the threshold corresponds to $f_0^* \approx 3.92$. Both for $f_0 = 3.80 < f_0^*$ (the upper panels) and $f_0 = 4.00 > f_0^*$ (the lower ones), the system starts from an identical configuration which is found by the Monte Carlo calculation at some high T. The only difference in the initial conditions is the total amount of the chemical attractants because we have set $c_{\text{const}} = (f_0/g_0)\rho_{\text{const}} \propto f_0$. The time step for integration is chosen to be $\Delta t = 10^{-7}$, and the horizontal axis is divided into 200 grid points. Note that the vertical axes are drawn on the log scale in panels (a,c) to see the behaviour of $\rho(r, t)$ near the boundary at $r = L$.

solutions [Eq. (46)]. The system converges to two different states depending on in which direction the transition point is approached. At the transition point, however, W is independent of the amplitude of aggregation and a continuous spectrum of infinitely many states exists between the two states with exactly the same value of W. The transition is thus continuous.

Our numerical calculation furthermore shows that Eq. (46) has multiple local minima (Figs 3 and 4). It is an open question if the existence of multiple local minima in W is due to the fact that we have restricted ourselves to radially symmetric solutions. That is, if we relaxed the symmetry requirement, some of the local minima could be connected to others via non-symmetric states. For example, the annular structure in Fig. 3(a) has relatively high W than other minima, and it is likely to collapse into another state in the presence of non-symmetric perturbations. At the same time, the extended parameter space could well introduce far more metastable states in the absence of the radial symmetry: Reference 28 shows us one of such states obtained with the finite-element method. To check those possibilities, we are currently working with the full normal-mode expression of W without the radial symmetry.

Method

In minimizing a partial sum of W from $m = 0$ to $m = \hat{m}$ [Eq. (46)] numerically, we treat the total mass M [Eq. (4)] and temperature T as input parameters. The initial state is defined by a set of variables, $G_{00} \equiv c_{\text{const}} = M/(\pi l^2)$ and $E_{01} = E_{02} = \cdots = G_{01} = G_{02} = \cdots = 0$, from which W is computed. Note that $E_{00} \equiv \rho_{\text{const}} = M/(\pi l^2)$ is a constant that will not be updated throughout the minimization procedure. We generate a neighbouring state in the following way: We first choose a mode $m \in [0, \ldots, \hat{m}]$. If $m > 0$, we add two independent random numbers r_E and r_G, each of which is taken from $[-0.1, 0.1]$, to the corresponding amplitudes E_{0m} and G_{0m}, respectively. If $m = 0$, on the other hand, only G_{00} will be updated by $r_G \in [-0.1, 0.1]$ because $E_{00} = \rho_{\text{const}}$ should remain constant. From this neighbouring state, we can calculate the Lyapunov functional, and let us denote its value W'. We basically employ the standard Metropolis algorithm to determine whether to accept the move to this neighbouring state: We first check if the move satisfies $W' \leq W$. Otherwise, we draw a random number from $[0, 1)$ and check if it is smaller than $\exp[(W - W')/T]$. If either of those two conditions is met, we proceed to check if the move leaves both ρ and c non-negative everywhere inside the disc by dividing the region into a sufficiently fine mesh compared to the variations of the highest mode with $\hat{m}$. In short, we carry out the move only if it is accepted by the Metropolis

algorithm without violating the non-negativity. One Monte Carlo step consists of $(\hat{m} + 1)$ such attempts to move to neighbouring states.

We test the algorithm by running it at $T = 0$ to obtain the expected results such as in the inset of Fig. 4(a). To find a better minimum of W, we choose an annealing schedule as $T = 10 \times (1.2)^{-n}$ with $n = 0, 1, \ldots, 50$, and take 1.5×10^4 Monte Carlo steps at each T (Figs 3 and 4). We also note that we have added calculations with $T = 0$ at the end of this annealing schedule for clarity in Fig. 2.

References

1. Hölldobler, B. & Wilson, E. O. *The Ants* (The Belknap Press, Cambridge, 1990).
2. Vicsek, T. & Zafeiris, A. Collective motion. *Physics Reports* **517**, 71–140 (2012).
3. Vela-Pérez, M., Fontelos, M. A. & Garnier, S. From individual to collective dynamics in Argentine ants (*Linepithema humile*). *Math. Biosci.* **262**, 56–64 (2015).
4. Fewell, J. H. Energetic and time costs of foraging in harvester ants, *Pogonomyrmex occidentalis. Behav. Ecol. Sociobiol.* **22**, 401–408 (1988).
5. Denny, A. J., Wright, J. & Grief, B. Foraging efficiency in the wood ant, *Formica rufa*: is time of the essence in trail following? *Anim. Behav.* **61**, 139–146 (2001).
6. Couzin, I. D. & Franks, N. R. Self-organized lane formation and optimized traffic flow in army ants. *Proc. R. Soc. Lond. B* **270**, 139–146 (2003).
7. Dussutour, A., Fourcassié, V., Helbing, D. & Deneubourg, J.-L. Optimal traffic organization in ants under crowded conditions. *Nature* **428**, 70–73 (2004).
8. Tao, T., Nakagawa, H., Yamasaki, M. & Nishimori, H. Flexible foraging of ants under unsteadily varying environment. *J. Phys. Soc. Jpn.* **73**, 2333–2341 (2004).
9. Dussutour, A., Nicolis, S. C., Deneubourg, J.-L. & Fourcassié, V. Collective decisions in ants when foraging under crowded conditions. *Behav. Ecol. Sociobiol.* **61**, 17–30 (2006).
10. Goss, S., Aron, S., Deneubourg, J.-L. & Pasteels, J. M. Self-organized shortcuts in the Argentine ants. *Naturwissenschaften* **76**, 579–582 (1989).
11. Reid, C. R., Sumpter, D. J. T. & Beekman, M. Optimisation in a natural system: Argentine ants solve the Tower of Hanoi. *J. Exp. Biol.* **214**, 50–58 (2010).
12. Oettler, J. *et al.* Fermat's principle of least time predicts refraction of ant trails at substrate borders. *PLoS One* **8**, 1–7 (2013).
13. Vicsek, T. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229 (1995).
14. Dorigo, M. & Stützle, T. *Ant Colony Optimization* (A Bradford Book, London, 2004).
15. Mohamed Jafar, O. A. & Sivakumar, R. Ant-based clustering algorithms: A brief survey. *Int. J. Comput. Theor. Eng.* **2**, 1793–8201 (2010).
16. Marchetti, M. C. *et al.* Hydrodynamics of soft active matter. *Rev. Mod. Phys.* **85**, 1143 (2013).
17. Deneubourg, J. L., Lioni, A. & Detrain, C. Dynamics of aggregation and emergence of cooperation. *Biol. Bull.* **202**, 262–267 (2002).
18. Tennenbaum, M., Liu, Z., Hu, D. & Fernandez-Nieves, A. Mechanics of fire ant aggregations. *Nat. Mater.* **15**, 54–59 (2015).
19. Patlak, C. S. Random walk with persistence and external bias. *Bull. Math. Biophys.* **15**, 311–338 (1953).
20. Keller, E. F. & Segel, L. A. Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.* **26**, 399–415 (1970).
21. Horstmann, D. From 1970 until now: the Keller-Segel model in chemotaxis and its consequences I. *Jahresber. Dtsch. Math. Ver.* **105**, 103–165 (2003).
22. Horstmann, D. From 1970 until now: the Keller-Segel model in chemotaxis and its consequences II. *Jahresber. Dtsch. Math. Ver.* **106**, 51–69 (2004).
23. Biler, P. Global solutions to some parabolic-elliptic systems of chemotaxis. *Adv. Math. Sci. Appl.* **9**, 347–359 (1999).
24. Hillen, T. & Painter, K. Global existence for a parabolic chemotaxis model with prevention of overcrowding. *Adv. Appl. Math.* **26**, 280–301 (2001).
25. Painter, K. J. & Hillen, T. Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Quart.* **10**, 501–543 (2002).
26. Potapov, A. B. & Hillen, T. Metastability in chemotaxis models. *J. Dyn. Differ. Equ.* **17**, 293–329 (2005).
27. Kowalczyk, R. Preventing blow-up in a chemotaxis model. *J. Math. Anal. Appl.* **305**, 566–588 (2005).
28. Kowalczyk, R. & Szymańska, Z. On the global existence of solutions to an aggregation model. *J. Math. Anal. Appl.* **343**, 379–398 (2008).
29. Ignacio Tello, J. & Winkler, M. A chemotaxis system with logistic source. *Commun. Part. Diff. Eq.* **32**, 849–877 (2007).
30. Hittmeir, S. & Jüngel, A. Cross diffusion preventing blow-up in the two-dimensional Keller–Segel model. *SIAM J. Math. Anal.* **43**, 997–1022 (2011).
31. Bedrossian, J. & He, S. Suppression of blow-up in Patlak-Keller-Segel via shear flows. arXiv:1609.02866.
32. Hillen, T. & Painter, K. J. A user's guide to PDE models for chemotaxis. *J. Math. Biol.* **58**, 183–217 (2009).
33. Amorim, P. Modeling ant foraging: A chemotaxis approach with pheromones and trail formation. *J. Theor. Biol.* **385**, 160–173 (2015).
34. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Westview Press, Boulder, CO, 2001).
35. Fatkullin, I. A study of blow-ups in the Keller-Segel model of chemotaxis. *Nonlinearity* **26**, 81–94 (2013).
36. Edwards, J. T. & Thouless, D. J. Numerical studies of localization in disordered systems. *J. Phys. C* **5**, 807–820 (1972).
37. Calvez, V. & Corrias, L. The parabolic-parabolic Keller-Segel model in R^2. *Commun. Math. Sci.* **6**, 417–447 (2008).
38. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **52**, 479–487 (1988).
39. Childress, S. & Percus, J. K. Nonlinear aspects of chemotaxis. *Math. Biosci.* **56**, 217–237 (1981).
40. Regnier, F. E. & Law, J. H. Insect pheromones. *J. Lipid Res.* **9**, 541–551 (1968).
41. Boas, M. L. *Mathematical Methods in the Physical Sciences*, 3 edn. (Wiley, Hoboken, NJ, 2006).
42. Newman, M. E. J. *Computational Physics* (CreateSpace Independent, United States, 2013).
43. Lee, H. K., Kown, C. & Park, H. Equivalence of operator-splitting schemes for the integration of the Langevin equation. *J. Stat. Mech.: Theory Exp.* **2006**, P08021 (2006).
44. Pamuk, S. Solution of the porous media equation by Admomian's decomposition method. *Phys. Lett. A* **344**, 184–188 (2005).

Acknowledgements

Author Contributions

B.J.K. conceived the study. S.K.B. carried out the calculations. S.K.B. and B.J.K. wrote the paper.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.