

Replication / Ecology

# [Re] Reproductive pair correlations and the clustering of organisms

Coralie Picoche<sup>1,2, ID</sup>, William R. Young<sup>3</sup>, and Frédéric Barraquand<sup>1,2, ID</sup>

<sup>1</sup>Institute of Mathematics of Bordeaux, CNRS & University of Bordeaux, Talence, France – <sup>2</sup>Integrative and Theoretical Ecology, LabEx COTE, University of Bordeaux, Pessac, France – <sup>3</sup>Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0230, USA

Edited by  
(Editor)

Received  
X X X

Published  
X X X

DOI  
–

## Introduction

In the present work, we explicate and replicate the results of Young et al. 2001 “Reproductive pair correlations and the clustering of organisms” [1], an analysis of the formation of aggregates in an otherwise homogeneous environment mimicking marine small-scale hydrodynamics. Using an individual-based model of independent, random-walking particles (also called “brownian bugs”), they show that simple ecological processes such as birth and death in a turbulent and viscous flow leads to the formation of elongated clusters and therefore departure from the usual, homogeneous solution of the advection-diffusion-reaction equation for a large population. This could explain observed small-scale phytoplankton patchiness.

Aggregation of phytoplanktonic organisms, observed at macro- to micro-scales [2, 3], could explain part of their coexistence, as spatial clustering can help reduce interspecific interactions [4] and/or allow species to benefit from different organic substrates [5]. Due to their size, microphytoplankton organisms experience a mostly viscous environment in a laminar shear fields, with random, but homogeneous changes in directions due to turbulence [6]. Ecological phenomena as simple as growth and death, which occur at the phytoplankton scale, interact with these hydrodynamics processes and can lead to aggregates. In this context, a better understanding of the interactions between demographic stochasticity and environmental fluctuations at small scales could provide further explanation for the distribution and coexistence of these organisms in turbulent environments.

In addition to replicating the numerical and mathematical results of Young et al. 2001, we also wish to present the intermediate mathematical steps that were missing from the original paper, which should make this replication article more accessible to most readers, especially those without a fluid mechanics background.

---

Copyright © 2021 C. Picoche, W.R. Young and F. Barraquand, released under a Creative Commons Attribution 4.0 International license.  
Correspondence should be addressed to Coralie Picoche (coralie.picoche@u-bordeaux.fr)  
The authors have declared that no competing interests exist.  
Code is available at [https://github.com/CoraliePicoche/brownian\\_bug\\_fluid/code..](https://github.com/CoraliePicoche/brownian_bug_fluid/code..)

## Methods

### Brownian bug model

The brownian bug model is a discrete-time, individual-based model, here presented in its 2D formulation. Each particle is characterized by the vector of its Cartesian coordinates  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  and its original position on the y-axis at  $t = 0$  (a child particle inherits this attribute), this last characteristics being used only for representation purposes. Space is a  $L \times L$  square with periodic boundary conditions. Each timestep, of duration  $\tau$ , is divided into three substeps: (1) demographic processes, (2) diffusion, and (3) advection.

(1) Demographic processes take place during the first substep. The discrete-time Markov chain of the brownian bug model can be approximated as a continuous-time spatial birth-death process (described in the Supplementary Material), also referred to in Young et al. [1] as a Galton-Watson process. Each organism has a fixed probability ( $p$ ) of reproducing, dying ( $q$ ), or remaining unchanged ( $1 - p - q$ ). When an individual reproduces, a new organism appears on top of the parent. In the following,  $p = q = 0.5$ .

(2) Diffusion is modeled as a brownian motion, i.e.  $\mathbf{x}'(t) = \mathbf{x}(t) + \delta\mathbf{x}(t)$  where each component of  $\delta\mathbf{x}(t)$  follows a Gaussian distribution  $\mathcal{N}(0, \Delta)$  where  $D = \frac{\Delta^2}{2\tau}$  is the diffusivity.

(3) The turbulent flow governing advective stirring follows the Pierrehumbert random map [7].

$$x_1(t + \tau) = x'_1(t) + (U\tau/2) \cos[kx'_2(t) + \phi(t)] \quad (1)$$

$$x_2(t + \tau) = x'_2(t) + (U\tau/2) \cos[kx_1(t + \tau) + \theta(t)] \quad (2)$$

where  $\phi(t)$  and  $\theta(t)$  are random phases uniformly distributed between 0 and  $2\pi$ ,  $k = 2\pi/L$  (in the rest of the paper, we use  $k = 2\pi$ ) and  $U$  is the stretching parameter.

Unless otherwise specified, each simulation is initialized with  $N_0 = 20,000$  particles uniformly distributed in a  $1 \times 1$  square and run for 1000 timesteps.

### Pair correlation function $G(\mathbf{r}, t)$

The pair density function  $G(\mathbf{x}_i, \mathbf{x}_j, t)$  is defined so that  $G(\mathbf{x}_i, \mathbf{x}_j, t)dA_1dA_2$  is the probability of finding a pair of brownian bugs with one member in the area  $dA_1$  around  $\mathbf{x}_1$  and the other in the area  $dA_2$  around  $\mathbf{x}_2$ .  $G(\mathbf{r}, t)$  is actually called the pair correlation function in [1]. The radial density function  $g(r, t)$  is defined as  $G(\mathbf{x}_i, \mathbf{x}_j, t) = C^2 g(r, t)$  with  $r = |\mathbf{x}_i - \mathbf{x}_j|$ . As the pair correlation disappears when  $r \rightarrow \infty$ ,  $g \rightarrow 1$ .

**Derivation of  $G(\mathbf{r}, t)$**  – All details of the derivation of  $G(r, t)$  are to be found in the Supplementary Material. We finally obtain:

$$\frac{\partial G}{\partial t} = 2Dr^{1-d} \frac{\partial}{\partial r} \left( r^{d-1} \frac{\partial G}{\partial r} \right) + 2(\lambda - \mu)G + \gamma r^{1-d} \frac{\partial}{\partial r} \left( r^{d+1} \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\mathbf{x}) \quad (3)$$

where  $\mathbf{x}$  is the position of the particle.

We focus on the case  $d = 2$  and  $\lambda = \mu$ , which means Eq. 3 can be reduced to

$$\frac{\partial G}{\partial t} = \frac{2D}{r} \frac{\partial}{\partial r} \left( r \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r} \frac{\partial}{\partial r} \left( r^3 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\mathbf{x}) \quad (4)$$

The value of  $\gamma$  is computed from simulations (see Supplementary Material).

**Analytical solution with advection** – The analytical solutions of  $G(r, t)$  with and without advection could only be found with the indications of WY. In the presence of advection ( $\gamma \neq 0$ ), a steady-state solution can be found; without advection, there is no steady-state.

$$\begin{aligned}
 \frac{2D}{r} \frac{\partial}{\partial r} \left( r \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r} \frac{\partial}{\partial r} \left( r^3 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\mathbf{x}) &= 0 \\
 \Leftrightarrow 2\pi r \left( \frac{2D}{r} \frac{\partial}{\partial r} \left( r \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r} \frac{\partial}{\partial r} \left( r^3 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\mathbf{x}) \right) &= 0 \\
 \Leftrightarrow 2\pi \left( 2D \frac{\partial}{\partial r} \left( r \frac{\partial G}{\partial r} \right) + \gamma \frac{\partial}{\partial r} \left( r^3 \frac{\partial G}{\partial r} \right) \right) + 2\pi r 2\lambda C \delta(\mathbf{x}) &= 0 \quad (5)
 \end{aligned}$$

We can then integrate Eq. 5 over a small area centered on a particle, with radius  $\rho$ . Let us first note that

$$\begin{aligned}
 \int_{\mathbb{R}^2} \delta(\mathbf{x}) d^2 \mathbf{x} &= 1 \\
 \Leftrightarrow \int_0^{2\pi} \int_0^\rho \delta(r') \delta(\theta) r' dr' d\theta &= 1 \\
 \Leftrightarrow 2\pi \int_0^\rho \delta(\mathbf{x}') r' dr' &= 1 \quad (6)
 \end{aligned}$$

where  $\mathbf{x}'$  is the equivalent of  $\mathbf{x}$  in polar coordinates.

Using Eq. 5 and 6, we can integrate between 0 and  $\rho$ ,

$$\begin{aligned}
 0 &= 2\pi \left( 2D\rho \frac{\partial G}{\partial r} + \gamma\rho^3 \frac{\partial G}{\partial r} \right) + 2\lambda C \\
 \Leftrightarrow \frac{\partial G}{\partial r} &= -\frac{1}{2\pi} \frac{2\lambda C}{2D\rho + \gamma\rho^3} \quad (7)
 \end{aligned}$$

Eq. 7 can now be integrated between  $\rho$  and  $\infty$ , knowing that  $G(\infty) = C^2$ .

$$C^2 - G(\rho) = -\frac{1}{2\pi} \int_\rho^\infty \frac{2\lambda C}{2Dr + \gamma r^3} dr \quad (8)$$

Using the variable change  $u = 2Dr + \gamma r^3$ , the integral is equivalent to  $\int \frac{u'}{u} du$ .

$$C^2 - G(\rho) = -\frac{\lambda C}{2\pi} \frac{1}{4D} [\log(\gamma) - \log(\frac{2D}{r^2} + \gamma)] \quad (9)$$

$$\Leftrightarrow G(\rho) = \frac{\lambda C}{8\pi D} \log \left( \frac{2D + \gamma r^2}{\gamma r^2} \right) + C^2 \quad (10)$$

Finally, the pair correlation function  $g = G/C^2$  is defined as

$$g = \frac{\lambda}{8\pi DC} \log \left( \frac{2D + \gamma r^2}{\gamma r^2} \right) + 1 \quad (11)$$

**Analytical solution without advection** – When  $U = 0$ ,  $\gamma = 0$  and there is no steady solution. We can get back to Eq. 4.

$$\frac{\partial G}{\partial t} = \frac{2D}{r} \frac{\partial}{\partial r} \left( r \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\mathbf{x}) \quad (12)$$

Assuming an isotropic environment (and switching to the polar coordinate system), this means

$$\frac{\partial G}{\partial t} - 2D\Delta G = 2\lambda C\delta(\mathbf{x}) \quad (13)$$

where  $\Delta = \nabla^2$  is the Laplacian operator.

We therefore have

$$\mathcal{L}G(\mathbf{x}, t) = 2\lambda C\delta(\mathbf{x}) \quad (14)$$

where  $\mathcal{L}$  is the linear differential operator  $\partial_t - 2D\Delta$ .

We can use the Green's function  $H$ , defined with  $\mathcal{L}H = \delta(\mathbf{x}, t) = \delta(\mathbf{x})\delta(t)$ .

By definition, we know that  $G(y) = \int H(y, s)2\lambda C\delta(s)ds$  (where  $y = (\mathbf{x}, t)$ ) is a solution to Eq. 14.

$$\begin{aligned} G(\mathbf{x}, t) &= 2\lambda C \int_{\mathbb{R}^2} \int_0^t H(\mathbf{x} - \mathbf{x}', t') \delta(\mathbf{x}') d^2\mathbf{x}' dt' \\ \Leftrightarrow &= 2\lambda C \int_0^t H(\mathbf{x}, t') dt' \end{aligned} \quad (15)$$

Eq. 15 can be used in Eq. 12:

$$\frac{\partial}{\partial t} \left( 2\lambda C \int_0^t H(\mathbf{x}, t') dt' \right) = 2D2\lambda C\Delta \int_0^t H(\mathbf{x}, t') dt' + 2\lambda C\delta(\mathbf{x}) \quad (16)$$

$$\Leftrightarrow \int_0^t \left( \frac{\partial H(\mathbf{x}, t')}{\partial t'} - 2D\Delta H(\mathbf{x}, t') \right) dt' = \delta(\mathbf{x}) \quad (17)$$

$$\Leftrightarrow \int_0^t \delta(\mathbf{x})\delta(t') dt' = \delta(\mathbf{x}) \quad (18)$$

which is true.

A solution for the Green's function using  $\mathcal{L} = \partial_t - 2D\Delta$  in 2 dimensions is

$$H(r, t) = \frac{1}{4\pi 2Dt} \exp\left(\frac{-r^2}{4 \times 2Dt}\right)$$

$G(r, t)$  can then be computed:

$$G(r, t) = 2\lambda C \left[ \frac{E_1\left(\frac{r^2}{8Dt}\right)}{8D\pi} \right]_0^t \quad (19)$$

where  $E_1$  is the exponential integral. Using  $G(r, 0) = C^2$  and  $\lim_{x \rightarrow +\infty} E_1 = 0$  in Eq. 19,

$$G(r, t) = 2\lambda C \frac{E_1\left(\frac{r^2}{8Dt}\right)}{8D\pi} + C^2 \quad (20)$$

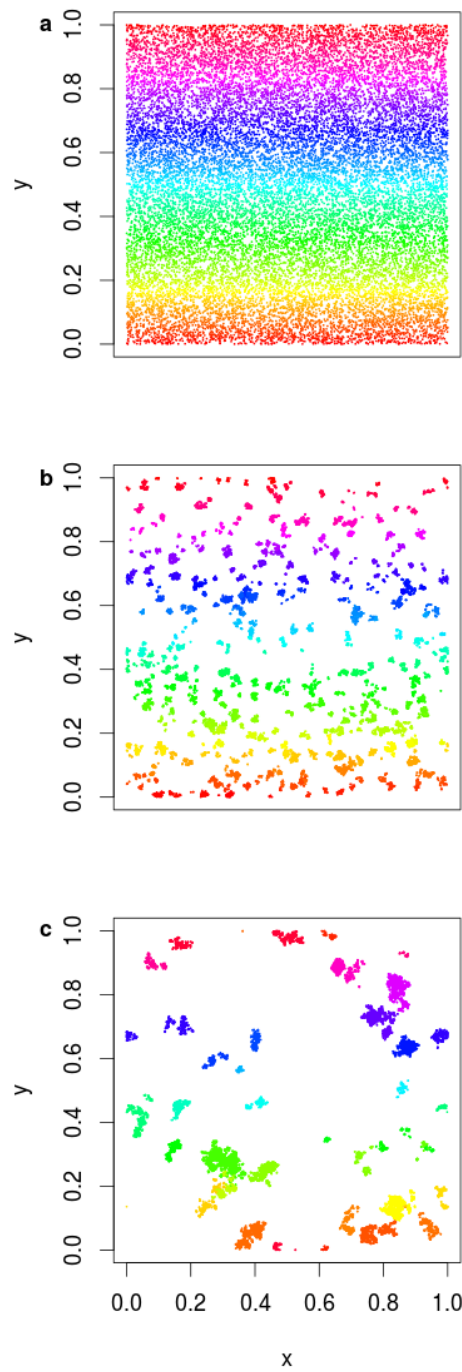
$$\Leftrightarrow g(r, t) = \frac{2\lambda}{C} \frac{E_1\left(\frac{r^2}{8Dt}\right)}{8D\pi} + 1 \quad (21)$$

## Results

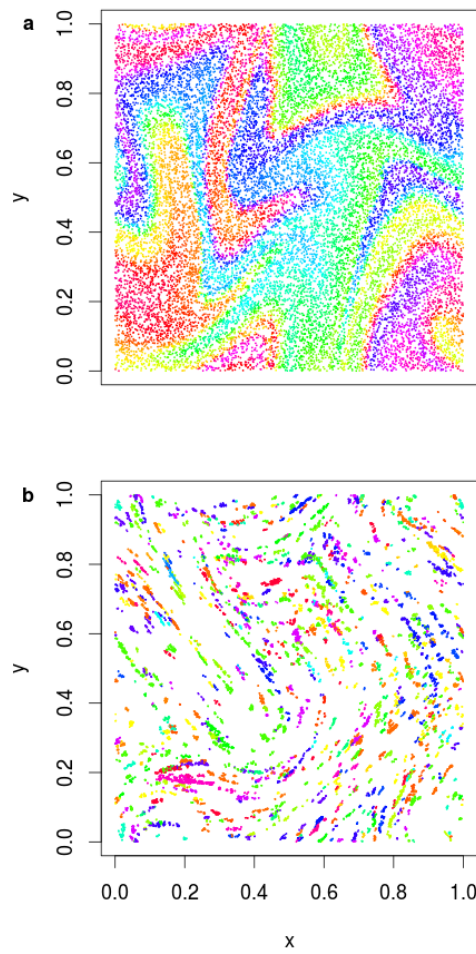
We were able to reproduce the spatial distributions of brownian bugs in each case described in [1].

In Fig. 1, the model has been run without the advection component and we can see the clumping of organisms due to demographic processes. In Fig. 2 a), the model has been run without its demographic component, but with advection and diffusion, confirming that hydrodynamics alone cannot ensure cluster formation, while in Fig. 2 b), all processes are taken into account, in which case organisms form aggregates.

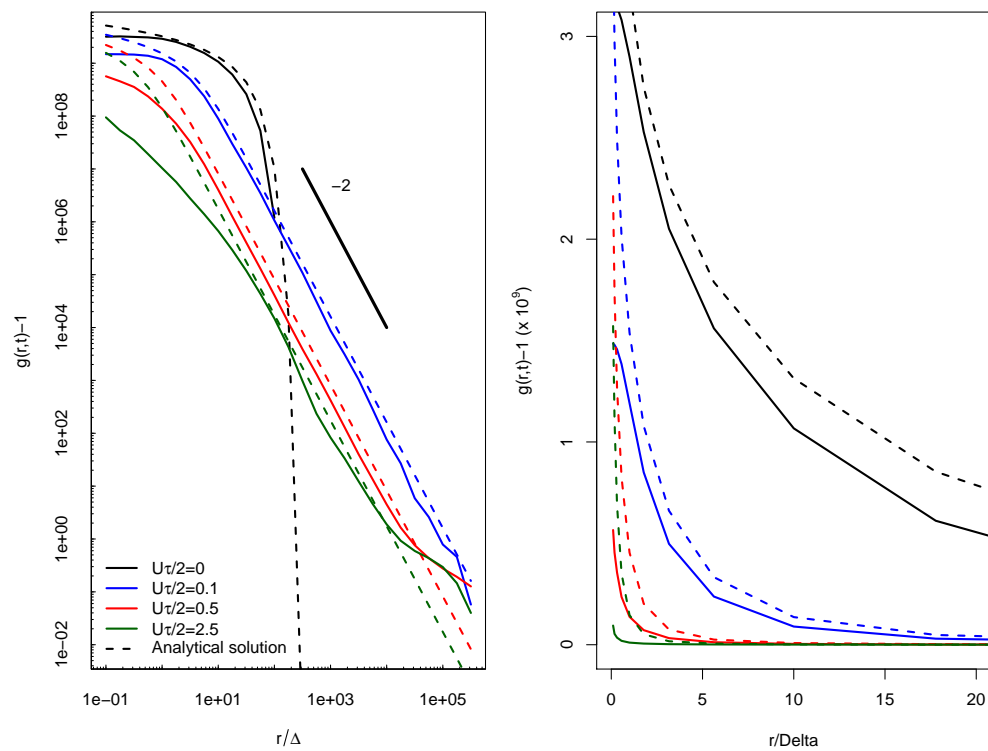
Fig. 3 proved much more challenging. Retrieving the analytical solutions of Eq. 4 could hardly be done without the inputs of WY. However, we also encountered issues when computing the pair correlation functions on simulations: for large values of  $r/\Delta$ , we observed zero values (absent pairs) of the pcf when  $U = 0$ . Despite the missing values, we can confirm that simulated and analytical pcf match (Fig. 3), with a slight underestimation by the simulations. The pcf also got closer to 1 than the analytical predictions for larger values of  $r$ .



**Figure 1.** Distribution of brownian bugs at different times in a simulation with  $\Delta = 10^{-3}$  and  $U = 0$ : initial conditions with a Poisson spatial distribution (a),  $t = 100\tau$  (b) and  $t = 1000\tau$  (c). Each particle is identified by a color which corresponds to the initial position on the y-axis of its ancestor at  $t = 0$ .



**Figure 2.** Distribution of brownian bugs with different processes in a simulation with  $\Delta = 10^{-3}$  and  $U\tau/2 = 0.1$ : without demographic processes at  $t = 30\tau$  (a), and with demographic processes at  $t = 1000\tau$  (b). Each particle is identified by a color which corresponds to the initial position on the y-axis of its ancestor at  $t = 0$ .



**Figure 3.** Logarithmic (a) and linear (b) plots of  $g(r, t)$  versus  $r/\Delta$ , with  $\Delta = 10^{-7}$  and  $U\tau/2 = 0, 0.1, 0.5, 2.5$  at  $t = 1000\tau$ . Solid lines result from simulations, dotted lines correspond to analytical solutions and the solid grey line indicates the  $r^{-2}$  scaling predicted by Eq. 3.

## Discussion

We successfully replicated both the numerical results and analytical solutions of [1]. Even though stochasticity prevents us from replicating exactly the same spatial point patterns as those seen in Fig. 1 and Fig. 2, we considered the patterns to be close enough to validate the simulations. Fig. 3 was also very close to the one shown in the original article, despite a slight underestimation.

The most challenging part of the replication was not actually to replicate the numerical results, but to find back the analytical expression of the pair density function  $G(r, t)$  from first principle (this was briefly explained in words in the original article). WY provided the necessary mathematical steps to find the analytical solutions shown in Fig. 3. We hope that such additional mathematical derivations will help readers through both the original and replication articles.

The original article did not provide quantitative values exactly matching marine microbes ecology; we thus wondered about the time and spatial scales that could be used for a realistic phytoplankton model. The length of the square side,  $L$ , is defined roughly as the Kolmogorov scale, the scale at which viscosity starts dominating turbulence. In the Batchelor-Kolmogorov regime, parent and offspring are separated by a distance lower than  $k^{-1}$ , where  $k$  is the wavenumber for the flow. Here, we consider the smallest wavenumber corresponding to the largest length scale  $L$ , i.e.  $k = 2\pi/L$ . In this case, we can write



$$Re = \frac{U}{k\nu} \quad (22)$$

$$\Rightarrow 1 = \frac{UL}{2\pi\nu} \quad (23)$$

$$\Leftrightarrow U = \frac{2\pi\nu}{L} \quad (24)$$

where  $\nu = 10^{-6} \text{ m}^2.\text{s}^{-1}$  is the kinematic viscosity for water.  
In this case,  $U = 6.3 \times 10^{-4} \text{ m.s}^{-1}$ .

To determine diffusivity of small organisms, we can use the Stokes-Einstein equation [8]:

$$D = \sqrt{\frac{RT}{N_A} \frac{1}{6\pi\eta a}} \quad (25)$$

where  $R = 8.314 \text{ J.K}^{-1}.\text{mol}^{-1}$  is the molar gas constant,  $T = 293 \text{ K}$  is the temperature of the environment,  $N_A = 6.0225 \times 10^{23}$  is Avogadro's number,  $\eta = 10^{-3} \text{ m}^{-1}.\text{kg.s}^{-1}$  is the viscosity of water and  $a$  is the radius of the organism considered.

We now apply this formula to microphytoplankton organisms of diameter  $50\mu\text{m}$ , keeping  $\tau$  outside of the equation for now.

$$\Delta = \sqrt{2D\tau} \quad (26)$$

$$= \sqrt{\frac{RT}{N_A} \frac{\tau}{3\pi\eta a}} \quad (27)$$

$$= \sqrt{\frac{8.314 \times 293}{6.0225 \times 10^{23}} \frac{1}{3\pi \times 10^{-3} \times 25 \times 10^{-6}}} \sqrt{\tau} \quad (28)$$

$$= 1.3 \times 10^{-7} \sqrt{\tau} \text{ m} \quad (29)$$

To compute  $\tau$ , we can consider a phytoplankton doubling rate of  $1 \text{ d}^{-1}$  [9], which means, with  $p = 0.5$ , that  $\tau = 0.5 \text{ d}$ .

This leads to  $U\tau/2 \approx 5.4 \times 10^3 \text{ cm.d}^{-1}$  and  $\Delta \approx 5 \times 10^{-5} \text{ cm}$ . These two values are much higher than the value used in Fig. 3 ( $0.1 < U\tau/2 < 2.5$  and  $\Delta = 10^{-7}$ ). A thorough discussion of the parameters is therefore necessary before extrapolating these results to real phytoplanktonic systems.

As the brownian bug model is currently fairly theoretical in its 2D formulation, a logical next step would be to consider similar dynamics in a 3D-model, which would render the comparison to real data easier. Using actual concentrations of phytoplanktonic organisms (e.g. diatoms), between  $10^3$  and  $10^6 \text{ C/L}$ , this would lead to 1 to  $10^3$  organisms if we kept  $L = 1 \text{ cm}$ . We might therefore need to increase the size of the considered square, or apply the model to small bacteria only. With a closer match between field and simulated concentrations, the model could better inform the likely fine-scale spatial structure of phytoplanktonic organisms.

## References

1. W. R. Young, A. J. Roberts, and G. Stuhne. "Reproductive pair correlations and the clustering of organisms." In: *Nature* 412.6844 (2001), pp. 328–331.

2. S. Lovejoy. "Universal multifractals and ocean patchiness: phytoplankton, physical fields and coastal heterogeneity." In: **Journal of Plankton Research** 23.2 (2001), pp. 117–141.
3. B. Pinel-Alloul and A. Ghadouani. "Spatial heterogeneity of planktonic microorganisms in aquatic systems." In: **The spatial distribution of microbes in the environment**. Springer, 2007, pp. 203–310.
4. J. S. Font-Muñoz, A. Jordi, I. Tuval, J. Arrieta, S. Anglès, and G. Basterretxea. "Advection by ocean currents modifies phytoplankton size structure." In: **J. R. Soc. Interface** 14.130 (2017), p. 20170046.
5. J. R. Seymour, Marcos, and R. Stocker. "Resource Patch Formation and Exploitation throughout the Marine Microbial Food Web." In: **The American Naturalist** 173.1 (2009), p. 15.
6. F. Peters and C. Marrasé. "Effects of turbulence on plankton: an overview of experimental evidence and some theoretical considerations." In: **Marine Ecology Progress Series** 205 (2000), pp. 291–306.
7. R. Pierrehumbert. "Tracer microstructure in the large-eddy dominated regime." In: **Chaos, Solitons & Fractals** 4.6 (1994), pp. 1091–1110.
8. A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." In: **Annalen der physik** 4 (1905).
9. J. E. Bissinger, D. J. S. Montagnes, J. Harples, and D. Atkinson. "Predicting marine phytoplankton maximum growth rates from temperature: Improving on the Eppley curve using quantile regression." In: **Limnology and Oceanography** 53.2 (2008), pp. 487–493.
10. R. H. Kraichnan. "Convection of a passive scalar by a quasi-uniform random straining field." In: **J. Fluid Mech.** 64.4 (1974), pp. 737–762.

## Supplementary Material

### Derivation of $G(r,t)$

**Diffusion and birth/death processes** – In this part, we aim to find back Eq. 2 in [1], i.e. Eq. 3 in our manuscript. We will first focus on the diffusion and birth/death processes, that is:

$$\frac{\partial G}{\partial t} = 2Dr^{1-d} \frac{\partial}{\partial r} \left( r^{d-1} \frac{\partial G}{\partial r} \right) + 2(\lambda - \mu)G + 2\lambda C\delta(\mathbf{x}) \quad (30)$$

We first define an ensemble of  $k$  identical brownian bugs in a  $d$ -dimension space. The bug number  $p$  is located at  $\mathbf{x}_p = [x_1, x_2, \dots, x_d]$ . At time  $t$ , the space is defined by (a) the number of brownian bugs  $k$  and (b) the vector of their locations  $\mathbf{X}_k = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$ . This is also called the Fock space.

The probability distribution over the state space is given by the functions  $\mathcal{P}_k(\mathbf{X}_k, t)$  such that:

$$\mathcal{P}_k(\mathbf{X}_k, t)d\mathbf{X}_k = \Pr\{k \text{ bugs, with a bug in } d\mathbf{x}_1, \text{ a bug in } d\mathbf{x}_2, \text{ etc.}\} \quad (31)$$

As bugs are indistinguishable, we can exchange  $\mathbf{x}_p$  and  $\mathbf{x}_q$  (permutation symmetry):

$$\mathcal{P}_k(\mathbf{x}_1, \dots, \mathbf{x}_p, \dots, \mathbf{x}_q, \dots, \mathbf{x}_k, t) = \mathcal{P}_k(\mathbf{x}_1, \dots, \mathbf{x}_q, \dots, \mathbf{x}_p, \dots, \mathbf{x}_k, t) \quad (32)$$

The normalization is:

$$\mathcal{P}_0(t) + \int \mathcal{P}_1(\mathbf{X}_1, t)d\mathbf{x}_1 + \int \int \mathcal{P}_2(\mathbf{X}_2, t)d\mathbf{x}_1 d\mathbf{x}_2 + \dots + \int_{\mathbb{R}^{2k}} \mathcal{P}_k(\mathbf{X}_k, t)d\mathbf{X}_k = 1 \quad (33)$$

We define  $b_k(\mathbf{x}, t) = \int \mathcal{P}_k(\mathbf{x}, \mathbf{X}_{k-1}, t)d\mathbf{X}_{k-1}$ , i.e.  $b_k(\mathbf{x}, t)d\mathbf{x}$  is the probability that there are  $k$  bugs and bug number 1 is in  $d\mathbf{x}$ .

The density of points is defined as:

$$\rho(\mathbf{x}, t) = \sum_{k=1}^{\infty} k b_k(\mathbf{x}, t) = \sum_{k=1}^{\infty} k \int \mathcal{P}_k(\mathbf{x}, \mathbf{X}_{k-1}, t) d\mathbf{X}_{k-1} \quad (34)$$

The pair correlation function is then:

$$G(\mathbf{x}, \mathbf{y}, t) = \sum_{k=2}^{\infty} k(k-1) \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} \quad (35)$$

We define the two particle distribution functions  $c_k(\mathbf{x}, \mathbf{y}, t) = \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2}$ . Note that  $G(\mathbf{x}, \mathbf{y}, t) = \sum_{k=2}^{\infty} k(k-1) c_k(\mathbf{x}, \mathbf{y}, t)$ .

**Proposition**

We show that

$$\frac{\partial c_k}{\partial t}(\mathbf{x}, \mathbf{y}, t) = D \nabla^2 c_k \quad (36a)$$

$$- k(\lambda + \mu) c_k \quad (36b)$$

$$+ (k + 1) \mu c_{k+1} \quad (36c)$$

$$+ 2 \frac{\lambda}{k} \delta(\mathbf{x} - \mathbf{y}) b_{k-1}(\mathbf{x}) + \lambda \frac{(k-2)(k+1)}{k} c_{k-1} \quad (36d)$$

**Proof**

We write the evolution of  $\mathcal{P}_k(\mathbf{X}_k, t)$ :

$$\frac{\partial \mathcal{P}_k}{\partial t} = D \nabla_k^2 \mathcal{P}_k \quad (37a)$$

$$- k(\lambda + \mu) \mathcal{P}_k \quad (37b)$$

$$+ (k + 1) \mu \int \mathcal{P}_{k+1}(\mathbf{X}_k, \mathbf{y}, t) d\mathbf{y} \quad (37c)$$

$$+ \frac{\lambda}{k} \sum_{p=1}^k \sum_{q=1, q \neq p}^k \delta(\mathbf{x}_p - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{X}_{k|p}, t) \quad (37d)$$

where  $\nabla_k^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \dots + \frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2}$  and  $\mathbf{X}_{k|p} = \mathbf{X}_k$  without  $\mathbf{x}_p$ .

Here, Eq. 37a is the diffusion part of the process, Eq. 37b corresponds to the rate at which realizations with  $k$  bugs lose a bug by mortality or gain a bug through birth, Eq. 37c corresponds to the rate at which realization with  $k + 1$  bugs lose a bug. Finally, a realization with  $k$  bugs can also be produced by a birth in a realization with  $k - 1$  bugs.

Using the derivation of  $\mathcal{P}_k(\mathbf{X}_k, t)$  and the definition of  $c_k$ , we have:

$$\frac{\partial c_k}{\partial t}(\mathbf{x}, \mathbf{y}, t) = D \int \nabla_k^2 \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} \quad (38a)$$

$$- k(\lambda + \mu) \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} \quad (38b)$$

$$+ (k + 1) \mu \int \mathcal{P}_{k+1}(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, \mathbf{z}, t) d\mathbf{X}_{k-2} d\mathbf{z} \quad (38c)$$

$$+ \frac{\lambda}{k} \int \sum_{p=1}^k \sum_{q=1, q \neq p}^k \delta(\mathbf{x}_p - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{X}_{k|p}, t) \quad (38d)$$

We will treat one term after the other.

**Diffusion term (38a)**

$$D \int \nabla_k^2 \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} = D \nabla_k^2 \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} \quad (39a)$$

$$= D \nabla_2^2 \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} \quad (39b)$$

$$= D \nabla_2^2 c_k \quad (39c)$$

because we already integrate over  $k - 2$  coordinates.

**Second term (38b)**

$$k(\lambda + \mu) \int \mathcal{P}_k(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, t) d\mathbf{X}_{k-2} = k(\lambda + \mu) c_k \quad (40)$$

**Death term (38c)**

$$(k+1)\mu \int \mathcal{P}_{k+1}(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-2}, \mathbf{z}, t) d\mathbf{X}_{k-2} d\mathbf{z} = (k+1)\mu \int \mathcal{P}_{k+1}(\mathbf{x}, \mathbf{y}, \mathbf{X}_{k-1}, t) d\mathbf{X}_{k-1} \quad (41a)$$

$$= (k+1)\mu c_{k+1} \quad (41b)$$

**Birth term (38d)**

In this section, we assume  $\mathbf{x} = \mathbf{x}_1$  and  $\mathbf{y} = \mathbf{x}_2$  (the reasoning is the same for different positions of  $\mathbf{x}$  and  $\mathbf{y}$  due to permutation symmetry).

We can decompose the double sum, starting with  $p = 1$ .

$$\sum_{q=2}^k \delta(\mathbf{x}_1 - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) = \delta(\mathbf{x}_1 - \mathbf{x}_2) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) \quad (42a)$$

$$+ \delta(\mathbf{x}_1 - \mathbf{x}_3) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) \quad (42b)$$

$$+ \dots \quad (42c)$$

$$+ \delta(\mathbf{x}_1 - \mathbf{x}_k) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) \quad (42d)$$

We integrate over the last  $k - 2$  coordinates.

$$\int \sum_{q=2}^k \delta(\mathbf{x}_1 - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (43a)$$

$$= \int \delta(\mathbf{x}_1 - \mathbf{x}_2) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (43b)$$

$$+ \int \delta(\mathbf{x}_1 - \mathbf{x}_3) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k + \dots \quad (43c)$$

$$+ \int \delta(\mathbf{x}_1 - \mathbf{x}_k) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (43d)$$

$$(43e)$$

As  $\int \delta(\mathbf{x}_p - \mathbf{x}_q) F(\mathbf{x}_q) d\mathbf{x}_q = F(\mathbf{x}_p)$ , we note here that

$$\int_{\mathbf{x}_3} \int_{\dots} \int_{\mathbf{x}_q} \int_{\dots} \int_{\mathbf{x}_k} \delta(\mathbf{x}_p - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_q, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k = \quad (44)$$

$$\int_{\mathbf{x}_3} \int_{\dots} \int_{\mathbf{x}_{q-1}} \int_{\mathbf{x}_{q+1}} \int_{\dots} \int_{\mathbf{x}_k} \mathcal{P}_{k-2}(\mathbf{x}_3, \dots, \mathbf{x}_p, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (45)$$

$$\sum_{q=2}^k \delta(\mathbf{x}_1 - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) = \delta(\mathbf{x}_1 - \mathbf{x}_2) \int \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (46a)$$

$$+ \int \mathcal{P}_{k-3}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_k, t) d\mathbf{x}_4 \dots d\mathbf{x}_k + \dots \quad (46b)$$

$$+ \int \mathcal{P}_{k-3}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_1, t) d\mathbf{x}_3 \dots d\mathbf{x}_1 \quad (46c)$$

$$= \delta(\mathbf{x}_1 - \mathbf{x}_2) \int \mathcal{P}_{k-1}(\mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (46d)$$

$$+ \int \mathcal{P}_{k-3}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{X}_{k-3}, t) d\mathbf{X}_{k-3} + \dots \quad (46e)$$

$$+ \int \mathcal{P}_{k-3}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{X}_{k-3}, t) d\mathbf{X}_{k-3} \quad (46f)$$

$$= \delta(\mathbf{x}_1 - \mathbf{x}_2) b_{k-1}(\mathbf{x}_2) + (k-2)c_{k-1} \quad (46g)$$

By symmetry, if  $p = 2$ , we obtain  $\delta(\mathbf{x}_1 - \mathbf{x}_2) b_{k-1}(\mathbf{x}_1) + (k-2)c_{k-1}$ .

Now, we need to use  $p \geq 3$ .

$$\int \sum_{q=1, q \neq p}^k \delta(\mathbf{x}_p - \mathbf{x}_q) \mathcal{P}_{k-1}(\mathbf{x}_1, \dots, \mathbf{x}_{p-1}, \mathbf{x}_{p+1}, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (47a)$$

$$= \int \delta(\mathbf{x}_p - \mathbf{x}_1) \mathcal{P}_{k-1}(\mathbf{x}_1, \dots, \mathbf{x}_{p-1}, \mathbf{x}_{p+1}, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (47b)$$

$$+ \int \delta(\mathbf{x}_p - \mathbf{x}_2) \mathcal{P}_{k-1}(\mathbf{x}_1, \dots, \mathbf{x}_{p-1}, \mathbf{x}_{p+1}, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k + \quad (47c)$$

$$+ \int \delta(\mathbf{x}_p - \mathbf{x}_3) \mathcal{P}_{k|p}(\mathbf{x}_1, \dots, \mathbf{x}_{p-1}, \mathbf{x}_{p+1}, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k + \dots \quad (47d)$$

$$+ \int \delta(\mathbf{x}_p - \mathbf{x}_k) \mathcal{P}_{k|p}(\mathbf{x}_1, \dots, \mathbf{x}_{p-1}, \mathbf{x}_{p+1}, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_k \quad (47e)$$

$$= \int \delta(\mathbf{x}_p - \mathbf{x}_1) d\mathbf{x}_p \int \mathcal{P}_{k-1}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_{p-1} d\mathbf{x}_{p+1} \dots d\mathbf{x}_k \quad (47f)$$

$$+ \int \delta(\mathbf{x}_p - \mathbf{x}_2) d\mathbf{x}_p \int \mathcal{P}_{k-1}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, t) d\mathbf{x}_3 \dots d\mathbf{x}_{p-1} d\mathbf{x}_{p+1} \dots d\mathbf{x}_k \quad (47g)$$

$$+ \int \mathcal{P}_{k|p}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_p, \dots, \mathbf{x}_k, t) d\mathbf{x}_4 \dots d\mathbf{x}_k + \dots \quad (47h)$$

$$+ \int \mathcal{P}_{k|p}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p, t) d\mathbf{x}_3 \dots d\mathbf{x}_{k-1} \quad (47i)$$

$$= 2 \int \mathcal{P}_{k-1}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{X}_{k-3}, t) d\mathbf{X}_{k-3} \quad (47j)$$

$$+ (k-3) \int \mathcal{P}_{k-1}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{X}_{k-3}, t) d\mathbf{X}_{k-3} \quad (47k)$$

$$= 2c_{k-1} + (k-3)c_{k-1} \quad (47l)$$

$$= (k-1)c_{k-1} \quad (47m)$$

Thus,  $\sum_{p=3}^k \sum_{q \neq p} \delta(\mathbf{x}_p - \mathbf{x}_q) \mathcal{P}_{k-1} d\mathbf{X}_{k|p} = (k-2)(k-1)c_{k-1}$ .

Finally, the birth term is

$$\frac{\lambda}{k} (2\delta(\mathbf{x} - \mathbf{y})b_{k-1}(\mathbf{x}) + 2(k-2)c_{k-1} + (k-2)(k-1)c_{k-1}) = 2\frac{\lambda}{k}\delta(\mathbf{x}-\mathbf{y})b_{k-1}(\mathbf{x}) + \lambda \frac{(k-2)(k+1)}{k}c_{k-1} \quad (48)$$

Combining all terms, we obtain:

$$\frac{\partial c_k}{\partial t}(\mathbf{x}, \mathbf{y}, t) = D\nabla_2^2 c_k \quad (49a)$$

$$- k(\lambda + \mu)c_k \quad (49b)$$

$$+ (k+1)\mu c_{k+1} \quad (49c)$$

$$+ 2\frac{\lambda}{k}\delta(\mathbf{x} - \mathbf{y})b_{k-1}(\mathbf{x}) + \lambda \frac{(k-2)(k+1)}{k}c_{k-1} \quad (49d)$$

### Proposition

In cartesian coordinates, the pair density admits the following evolution equation:

$$\frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}, t) = D\nabla_2^2 G + 2(\lambda - \mu)G + 2\lambda\delta(\mathbf{x} - \mathbf{y})\rho(\mathbf{x}) \quad (50)$$

### Proof

Using the definition of  $G(\mathbf{x}, \mathbf{y}, t)$  and Eq. 36a-36d:

$$\frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}, t) = \sum_{k=2}^{\infty} k(k-1)D\nabla_2^2 c_k \quad \equiv T1 \quad (51a)$$

$$- \sum_{k=2}^{\infty} k(k-1)k(\lambda + \mu)c_k \quad \equiv T2 \quad (51b)$$

$$+ \sum_{k=2}^{\infty} k(k-1)(k+1)\mu c_{k+1} \quad \equiv T3 \quad (51c)$$

$$+ \sum_{k=2}^{\infty} k(k-1)2\frac{\lambda}{k}\delta(\mathbf{x} - \mathbf{y})b_{k-1}(\mathbf{x}) \quad \equiv T4 \quad (51d)$$

$$+ \sum_{k=2}^{\infty} \lambda k(k-1)\frac{(k-2)(k+1)}{k}c_{k-1} \quad \equiv T5 \quad (51e)$$

$$T1 \quad T1 = D\nabla_2^2 \sum_{k=2}^{\infty} k(k-1)c_k = D\nabla_2^2 G_k$$

T3

$$T3 = \sum_{k=2}^{\infty} k(k-1)(k+1)\mu c_{k+1} \quad (52a)$$

$$= \sum_{k'=3}^{\infty} (k'-1)(k'-2)k'\mu c_{k'} \quad \text{with } k' = k+1 \quad (52b)$$

$$= \sum_{k'=2}^{\infty} (k'-1)(k'-2)k'\mu c_{k'} \quad \text{if } k' = 2, k' - 2 = 0 \quad (52c)$$

T4

$$T4 = 2\lambda\delta(\mathbf{x} - \mathbf{y}) \sum_{k=2}^{\infty} (k-1)b_{k-1}(\mathbf{x}) \quad (53a)$$

$$= 2\lambda\delta(\mathbf{x} - \mathbf{y}) \sum_{k''=1}^{\infty} k''b_{k''}(\mathbf{x}) \quad \text{with } k'' = k-1 \quad (53b)$$

$$= 2\lambda\delta(\mathbf{x} - \mathbf{y})\rho(\mathbf{x}) \quad (53c)$$

T5

$$T5 = \lambda \sum_{k=2}^{\infty} k(k-1) \frac{(k-2)(k+1)}{k} c_{k-1} \quad (54a)$$

$$= \lambda \sum_{k''=1}^{\infty} k''(k''-1)(k''+2)c_{k''} \quad \text{with } k'' = k-1 \quad (54b)$$

$$= \lambda \sum_{k''=2}^{\infty} k''(k''-1)(k''+2)c_{k''} \quad \text{if } k'' = 1, k''-1 = 0 \quad (54c)$$

T2+T3+T5

$$T2 + T3 + T5 = \sum_{k=2}^{\infty} -k(k-1)k(\lambda + \mu)c_k + (k-1)(k-2)k\mu c_k + k(k-1)(k+2)\lambda c_k \quad (55a)$$

$$= \sum_{k=2}^{\infty} k(k-1)c_k(-k(\lambda + \mu) + (k-2)\mu + (k+2)\lambda) \quad (55b)$$

$$= 2(\lambda - \mu) \sum_{k=2}^{\infty} k(k-1)c_k \quad (55c)$$

$$= 2(\lambda - \mu)G_k \quad (55d)$$

T1+T2+T3+T4+T5 Combining all terms, we have

$$\frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}, t) = D\nabla_2^2 G \quad (56a)$$

$$+ 2(\lambda - \mu)G_k \quad (56b)$$

$$+ 2\lambda\delta(\mathbf{x} - \mathbf{y})\rho(\mathbf{x}) \quad (56c)$$

**Advection process** – With clarifications from WY on the Kraichnan equations [10], we prove that, taking only advection into account:

$$\frac{\partial G}{\partial t} = \gamma r^{1-d} \frac{\partial}{\partial r} \left( r^{d+1} \frac{\partial G}{\partial r} \right) \quad (57)$$

Let  $\mathbf{r}(t)$  being the distance between two points as a function of time, which follows a geometric brownian motion, and  $\mathbf{q}(t) = \log(\mathbf{r}(t))/\log(\mathbf{r}(0))$ . Kraichnan defines  $Q(\mathbf{q})$  as the probability distribution of  $\mathbf{q}$ . We have  $Q = r^d G$ . From Kraichnan and Fokker-Planck equations, we know that:

$$\frac{\partial Q}{\partial t} = \gamma \frac{\partial^2 Q}{\partial q^2} - \gamma d \frac{\partial Q}{\partial q} \quad (58)$$



where  $\gamma$  is the stretching parameter, i.e. the diffusivity and  $\gamma d$  is the drift.

We have:

$$\frac{\partial Q}{\partial q} = \frac{\partial(r^d G)}{\partial r} \frac{\partial r}{\partial q} \quad (59a)$$

$$= r \left( dr^{d-1} G + r^d \frac{\partial G}{\partial r} \right) \quad (59b)$$

$$= dr^d G + r^{d+1} \frac{\partial G}{\partial r} \quad (59c)$$

Therefore,

$$\frac{\partial^2 Q}{\partial q^2} = \frac{\partial}{\partial q} \left( \frac{\partial(r^d G)}{\partial r} \frac{\partial r}{\partial q} \right) \quad (60a)$$

$$= r \frac{\partial}{\partial r} \left( dr^d G + r^{d+1} \frac{\partial G}{\partial r} \right) \quad (60b)$$

$$= r \frac{\partial}{\partial r} (r^{d+1} \frac{\partial G}{\partial r}) + rd \frac{\partial(r^d G)}{\partial r} \quad (60c)$$

Using Eq. 58,

$$\frac{\partial(r^d G)}{\partial t} = \gamma r \frac{\partial}{\partial r} (r^{d+1} \frac{\partial G}{\partial r}) + \gamma dr \frac{\partial(r^d G)}{\partial r} - \gamma dr \frac{\partial(r^d G)}{\partial r} \quad (61a)$$

$$r^d \frac{\partial G}{\partial t} = \gamma r \frac{\partial}{\partial r} (r^{d+1} \frac{\partial G}{\partial r}) \quad (61b)$$

$$= \gamma r^{1-d} \frac{\partial}{\partial r} (r^{d+1} \frac{\partial G}{\partial r}) \quad (61c)$$

### Stretching parameter $\gamma$

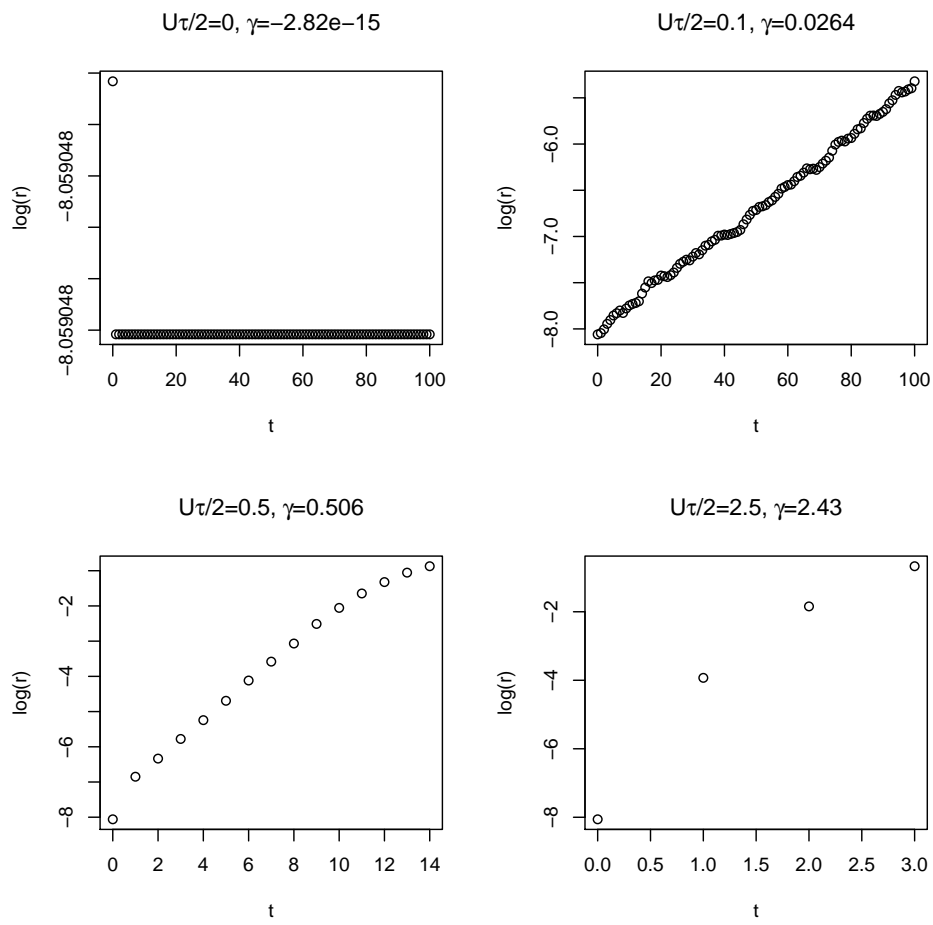
$\gamma$  is computed with simulations, with the formula  $r(t) \propto \exp(\gamma dt) \rightarrow \frac{1}{2} \ln(r(t)) = \gamma t$  if  $d = 2$  with  $r$  being the separation between pairs of particles.  $\gamma$  is estimated as the slope of

$$\frac{1}{2} \langle \ln(r(t)) \rangle = f(t)$$

with  $\langle \ln(r(t)) \rangle$  being the average obtained from 800 pairs of particles.

$$\forall t, \langle \ln(r(t)) \rangle = \frac{1}{800} \sum_{p=1}^{800} \ln(r(\mathbf{x}_{1,p}(t) - \mathbf{x}_{2,p}(t)))$$

where  $r(\mathbf{x}_{1,p}(t) - \mathbf{x}_{2,p}(t))$  is the distance between a particle  $1p$  at position  $\mathbf{x}_{1,p}$  and its counterpart  $2p$ , initialized with  $r(0) = 10^{-7} \forall p$  (see Fig. 4 for  $\gamma$  estimates).



**Figure 4.** Estimates of  $\gamma$  for different  $U\tau/2$