

Local intraspecific aggregation in phytoplankton model communities: spatial scales of occurrence and implications for coexistence

Coralie Picoche^{1,*}, William R. Young², Frédéric Barraquand^{1,*}

¹Institute of Mathematics of Bordeaux, University of Bordeaux and CNRS, Talence, France

²Scripps Institution of Oceanography, La Jolla, California, USA

Abstract

The coexistence of multiple phytoplankton species despite their reliance on similar resources is often explained with mean-field models assuming mixed populations. In reality, observations of phytoplankton indicate spatial aggregation at all scales, including at the scale of a few individuals. Local spatial aggregation can hinder competitive exclusion since individuals then interact mostly with other individuals of their own species, rather than competitors from different species. To evaluate how microscale spatial aggregation might explain phytoplankton diversity maintenance, an individual-based, multispecies representation of cells in a hydrodynamic environment is required. We formulate a three-dimensional and multispecies individual-based model of phytoplankton population dynamics at the Kolmogorov scale. The model is studied through both simulations and the derivation of spatial moment equations, in connection with point process theory. The spatial moment equations show a good match between theory and simulations. We parameterized the model based on phytoplankters' ecological and physical characteristics, for both large and small phytoplankton. Defining a zone of potential interactions as the overlap between nutrient depletion volumes, we show that local species composition—within the range of possible interactions—depends on the size class of phytoplankton. In small phytoplankton, individuals remain in mostly monospecific clusters. Spatial structure therefore favours intra- over inter-specific interactions for small phytoplankton, contributing to coexistence. Large phytoplankton cell neighbourhoods appear more mixed. Although some small-scale self-organizing spatial structure remains and could influence coexistence mechanisms, other factors may need to be explored to explain diversity maintenance in large phytoplankton.

Keywords: aggregation; coexistence; individual-based model; phytoplankton; spatial moment equations; spatial point process

*correspondence to: frederic.barraquand@u-bordeaux.fr & cpicoche@gmail.com

26 Introduction

27 Phytoplankton communities are among the most important photosynthetic groups on Earth,
28 being at the bottom of the marine food chain, and responsible for approximately half the
29 global primary production (Field *et al.*, 1998). Their contribution to ecosystem functions
30 is only matched by their contribution to biodiversity. Indeed, phytoplankton communities
31 are characterized by a surprisingly high number of species. For example, a single sample
32 as small as a few mL can contain up to seventy species (REPHY, 2017; Widdicombe &
33 Harbour, 2021). This observation is usually called the “paradox of the plankton” (a term
34 coined by Hutchinson, 1961), which refers to the conflict between the observed diversity of
35 species competing for similar resources in a seemingly homogeneous environment, and models
36 predicting that only a few species will persist by outcompeting the others (MacArthur &
37 Levins, 1964; Huisman & Weissing, 1999; Schippers *et al.*, 2001). Phytoplankton models for
38 coexistence are now almost as diverse as their model organisms (Record *et al.*, 2014), but they
39 often describe only a handful of species, which does not correspond to the diversity observed in
40 the field. When modeling rich communities (> 10 species), classical answers to the plankton
41 paradox involving temporal fluctuations (e.g., Li & Chesson, 2016; Chesson, 2018) are not
42 sufficient to maintain a realistic diversity. For instance, we found that a phytoplankton
43 community dynamics model with environmental fluctuations and storage effect still requires
44 extra niche differentiation for coexistence, which manifests in stronger intraspecific than
45 interspecific interactions (Picoche & Barraquand, 2019). However, it is not clear that we
46 should resort to hidden niches to explain phytoplankton coexistence, as most models also
47 make hidden simplifying assumptions that could be relaxed. One that we relax here is mean-
48 field dynamics at the microscale. Indeed, field observations have revealed phytoplankton
49 patchiness for decades, with early records in the past centuries (Bainbridge, 1957; Stocker,
50 2012), from the macro- to the micro-scale (Leonard *et al.*, 2001; Doubell *et al.*, 2006; Font-
51 Muñoz *et al.*, 2017).

52 Phytoplankton patchiness can at least be partly explained by the hydrodynamics of their
53 environment: the size of these organisms is mostly below the size of the smallest eddy (i.e.,
54 the Kolmogorov scale). In a typical aquatic environment such as the ocean, phytoplankton
55 individuals are embedded in viscous micro-structures (Peters & Marrasé, 2000) while phy-
56toplankton populations are displaced by a turbulent flow at slightly larger scales (Martin,
57 2003; Prairie *et al.*, 2012). Phytoplankton organisms therefore live in an environment where
58 fluid viscosity dominates at the scale of an individual but turbulent dispersion dominates on
59 length scales characteristic of a small population of those individuals (Estrada *et al.*, 1987;
60 Prairie *et al.*, 2012).

61 This leads us to consider demography in the context of this environmental variation cre-
62 ated by hydrodynamic processes. Individual-based models provide a convenient depiction of
63 population dynamics and movement at the microscale (Hellweger & Bucci, 2009). In this
64 framework, population growth is a result of individual births and deaths. Aggregation of
65 individuals can emerge from local reproduction coupled with limited dispersal, which can
66 happen in a fluid where turbulence and diffusion are not strong enough to disperse kin ag-
67 gregates (Young *et al.*, 2001). The resulting local aggregation can then affect the community
68 dynamics at larger spatial scales, even when all competitors are equivalent (i.e., with equal
69 interaction strengths irrespective of species identity). Indeed, the combination of local dis-
70 persal after reproduction and local interactions leads to stronger intraspecific interactions
71 than interspecific interactions at the population level (Detto & Muller-Landau, 2016). This
72 mechanism stabilizes the community, as a high intra-to-interspecific interaction strength ratio
73 makes a species control its abundance more than it controls the abundance of other species,
74 which is associated with coexistence in theoretical models (Levine & HilleRisLambers, 2009;
75 Barabás *et al.*, 2017) and often observed in the field at the population level (Adler *et al.*,
76 2018; Picoche & Barraquand, 2020). Therefore, the microscale spatial distribution of indi-
77 viduals likely affects the interaction structure within a community, and may sustain diversity
78 (Haegeman & Rapaport, 2008).

79 Existing models of phytoplankton populations near the Kolmogorov scale — between 1
80 mm and 1 cm in an oceanic environment (Barton *et al.*, 2014) — focus on a single species and
81 the clustering of its individuals (Young *et al.*, 2001; Birch & Young, 2006; Bouderbala *et al.*,
82 2018; Breier *et al.*, 2018). These models share similarities to dynamic point process models
83 (Law *et al.*, 2003; Bolker & Pacala, 1999; Plank & Law, 2015) developed initially with larger
84 organisms in mind. When phytoplankton individual-based models consider multiple types of
85 organisms, they focus for now on how organisms with opposite characteristics (e.g., increase
86 versus decrease in density with turbulence in Borgnino *et al.*, 2019; Arrieta *et al.*, 2020)
87 segregate spatially, or on coexistence of species that have contrasting trait values (e.g., size
88 in Benczik *et al.*, 2006). Such models are useful as an explanation of how species with marked
89 differences might coexist. The difficulty of the coexistence problem, however, is that we also
90 have to explain how closely related species or genera (e.g., within diatoms), many of whom
91 have similar size, buoyancy, chemical composition, etc., manage to coexist within a single
92 trophic level. This requires modelling *similar* species in a spatially realistic environment and
93 objectively quantifying whether they aggregate or segregate in space.

94 To do so, we build a multispecies version of the Brownian Bug Model (BBM) of Young
95 *et al.* (2001), an individual-based model which includes an advection process mimicking a
96 turbulent fluid flow, passive diffusion of organisms, as well as stochastic birth and death pro-

cesses. The initial version of this model (Young *et al.*, 2001) coupled limited dispersal and local reproduction with ocean-like microscale hydrodynamics, and showed spatial clusters of individuals of the same species. The original BBM was limited to a single species and was illustrated with two-dimensional simulations. The model was not strongly quantitative (Picoche *et al.*, 2022) in the sense that parameters were not informed by current knowledge on phytoplankton biology (numbers of cells per liter, diffusion characteristics, etc.). As phytoplankton organisms live in a three-dimensional environment, informing the model with more realistic parameters requires us to shift to three dimensions. We also extend the model to multiple species, and consider two size classes for our phytoplankton communities, which are either made of nanophytoplankton ($3 \mu\text{m}$ diameter, $\approx 10^6$ cells L^{-1}) or microphytoplankton ($50 \mu\text{m}$, $\approx 10^4$ cells L^{-1}). We populate each community with 3 to 10 different species.

The Brownian Bug Model (in its original single-species form as in the multispecies version considered here) is related to spatial branching processes. Without advection, it combines a continuous-time, discrete-state model for population growth and a continuous-time, continuous-space Brownian motion for particle diffusion (Birch & Young, 2006). It is further complexified by a turbulent flow in Young *et al.* (2001); Picoche *et al.* (2022) as well as here. In spite of this complexity, it remains possible to derive the dynamics of pair density functions, which quantify the degree of intra- and interspecific clustering of organisms, via correlations between positions of organisms (see next section). Thus we can understand emergent spatial structures in analytic detail and compare these predictions to the results from three-dimensional simulations. Furthermore, because we do not consider direct interactions between organisms, the multispecies spatial point process that represents the stable state of the BBM is a random superposition of spatial point processes for each species (Illian *et al.*, 2008). This enables us to derive, in addition to pair correlation functions, analytical formulas for the species composition in the neighbourhood of an individual, which are more readily ecologically interpreted than pair density or correlation functions.

Model and spatial statistics

Brownian Bug Model

The Brownian Bug Model (BBM) describes the dynamics of individuals in a turbulent and viscous environment, including demographic processes. The model is continuous in space and time. Here we extend the mostly two-dimensional, monospecific version in Young *et al.* (2001), to three dimensions and S species.

Each individual is characterized by its species identity i and its position $\mathbf{x}^T = (x, y, z)$.

130 The population dynamics are modelled by a linear birth-death process with birth rate λ_i
 131 and death rate μ_i . Each individual independently follows a Brownian motion with diffusivity
 132 D_i , and is advected by a common stochastic and chaotic flow modelling turbulence. The
 133 model applies in the Batchelor regime, which means that the separation $s(t)$ between two
 134 individuals k and l grows exponentially with time with stretching parameter γ , i.e. $s(t) =$
 135 $\ln(|\mathbf{x}_k - \mathbf{x}_l|(t)) \propto 3\gamma t$ (Kraichnan, 1974; Young *et al.*, 2001).

136 Within a given community (the set of all individuals of the S species), all species share the
 137 same parameters: λ_i , μ_i and D_i values can change between communities, as we later consider
 138 small and large phytoplankton, but are set to common values within a community. On the
 139 contrary, γ describes the environment and is not community-specific, i.e., all individuals
 140 are displaced by the same turbulent stirring. For numerical simulations, time needs to be
 141 discretized (this is required for diffusion and advection modelling). The approximated model
 142 advances through time in small steps of duration τ . During each interval, events unroll as
 143 follows:

- 144 1. Demography: each individual can either reproduce with probability $p_i = \lambda_i\tau$ (forming
 145 a new individual of the same species i at the same position \mathbf{x} as the parent), die with
 146 probability $q_i = \mu_i\tau$, or remain unchanged with probability $1 - p_i - q_i$.
- 147 2. Diffusion: each individual moves to a new position $\mathbf{x}(t') = \mathbf{x}(t) + \delta\mathbf{x}(t)$, with $t < t' <$
 148 $t + \tau$. The random displacement $\delta\mathbf{x}(t)$ is drawn from a Gaussian distribution $\mathcal{N}(0, \Delta_i^2)$
 149 with $D_i = \Delta_i^2/2\tau$ the diffusivity. This diffusive step separates the initially coincident
 150 pairs produced by reproduction in step 1 above.
- 151 3. Turbulence: each individual is displaced by a turbulent flow, modelled with the Pier-
 152 rehumbert map (Pierrehumbert, 1994), adapted to three dimensions following Ngan &
 153 Vanneste (2011). Thus given the position at time t' the updated position at time $t + \tau$
 154 is

$$\begin{aligned}
 x(t + \tau) &= x(t') + \frac{U\tau}{3} \cos(ky(t') + \phi(t)) \\
 y(t + \tau) &= y(t') + \frac{U\tau}{3} \cos(kz(t') + \theta(t)) \\
 z(t + \tau) &= z(t') + \frac{U\tau}{3} \cos(kx(t + \tau) + \psi(t)).
 \end{aligned} \tag{1}$$

155 Above, U is the velocity of the flow, $k = 2\pi/L_s$ is the wavenumber for the flow at the length
 156 scale L_s (see below) and $\phi(t)$, $\theta(t)$, $\psi(t)$ are random phases drawn from a uniform distribution
 157 between 0 and 2π ; these phases remain constant during the interval between t and $t + \tau$.
 158 The shift from continuous to discrete-time turbulence modelling is described in Section S1 in
 159 the Supplementary Information. The velocity U is related to γ . As the separation between
 160 two points grows exponentially with parameter 3γ due to turbulence, the exponent γ can be

161 estimated as the slope of $1/3 \langle \ln(s(t)) \rangle = f(t)$ in the absence of diffusion and demography
162 (Young *et al.*, 2001; Picoche *et al.*, 2022).

163 Individuals are distributed in a cube of side length L , with periodic boundary conditions.
164 The cube dimensions are determined to balance computing costs and realistic concentrations
165 of individuals; they represent the accumulation of a few volumes of scale L_s .

166 Characterization of the spatial distribution

167 Let W be the observation window (in our case, the whole cube, which we never subsample
168 hereafter). The state of the system at time t can be described as a collection of S populations,
169 where the population of species i is made of n_i individuals randomly distributed in W ,
170 with positions $\mathbf{X}_i(t) = [\mathbf{x}_{1,i}(t), \mathbf{x}_{2,i}(t), \dots, \mathbf{x}_{n_i,i}(t)]$. $\mathbf{X}(t) = [\mathbf{X}_1(t), \dots, \mathbf{X}_S(t)]$ arises from
171 a stochastic and spatial individual-based model changing through time, but can also be
172 analyzed as a spatial point process at time t . We note that the point distributions remain
173 the same for all spatial translations $\boldsymbol{\xi}$ (i.e., the point process described by the set $\mathbf{X} =$
174 $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$ is the same as $\mathbf{X}_{\boldsymbol{\xi}} = [\mathbf{x}_1 + \boldsymbol{\xi}, \mathbf{x}_2 + \boldsymbol{\xi}, \dots, \mathbf{x}_k + \boldsymbol{\xi}]$): the process is stationary.

175 A useful method to characterize a spatial point process is the use of spatial moments (il-
176 lustrated in Section S2 of the SI for simple spatial point processes). These can be theoretically
177 derived and used to check simulations. The spatial moments of a process are, however, merely
178 statistical indicators which then need to be related to more easily ecologically interpretable
179 quantities. This is the role of the dominance index, which we present below.

180 Spatial moments

181 The first-order moment is the intensity of the process, or mean concentration of individuals,
182 whose empirical estimate is $C_i = \frac{\widehat{N_i(W)}}{V(W)}$, where $\widehat{N_i(W)}$ is the empirical number of individuals
183 of species i in the cube W and $V(W) = L^3$ is the volume of the cube; it does not give any
184 information regarding the spatial distribution of individuals, and their spatial correlations.

185 The second-order product density, or pair density $G(r, t)$, is the expected density of pairs
186 of points separated by a distance r (Law *et al.*, 2003). A similar statistic can be used for
187 marked spatial point process. In our case, the marks are the species' identities, and we can
188 define $G_{ij}(r, t)$, so that $G_{ij}(r, t)d\mathbf{x}_A d\mathbf{x}_B$ is the probability of finding an individual of species
189 i in volume $d\mathbf{x}_A$ and an individual of species j in volume $d\mathbf{x}_B$, with the distance between
190 the centers of $d\mathbf{x}_A$ and $d\mathbf{x}_B$ equal to r (pages 219 and 325 in Illian *et al.*, 2008). We define
191 $\boldsymbol{\xi}$ as the vector connecting the center of $d\mathbf{x}_A$ to the center of $d\mathbf{x}_B$, while $r = |\boldsymbol{\xi}|$ is the radial
192 distance. We show in Picoche *et al.* (2022) that the intraspecific pair density $G_{ii}(r, t)$, in

193 three dimensions, is a solution of

$$\frac{\partial G_{ii}}{\partial t}(r,t) = \frac{2D_i}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G_{ii}}{\partial r} \right) + \frac{\gamma}{r^2} \frac{\partial}{\partial r} \left(r^4 \frac{\partial G_{ii}}{\partial r} \right) + 2(\lambda_i - \mu_i)G_{ii} + 2\lambda_i C_i \delta(\boldsymbol{\xi}). \quad (2)$$

194 The pair correlation function $g_{ij}(r,t)$, or pcf, can be derived from the pair density and is
195 defined as

$$g_{ij}(r,t) = \frac{G_{ij}(r,t)}{C_i C_j}. \quad (3)$$

196 The pcf is equal to one when the spatial distribution of species i individuals is random relative
197 to species j individuals. To compute the intraspecific pcf $g_{ii}(r,t)$ at steady state, considering
198 a population at equilibrium, we integrate Eq. 2 (see Appendices, Eqs. 19-30) with $\lambda_i = \mu_i$
199 and obtain

$$g_{ii}(r) = 1 + \frac{\lambda_i}{4\pi D_i C_i \ell_{B,i}} \left(\frac{\ell_{B,i}}{r} + \arctan \left(\frac{r}{\ell_{B,i}} \right) - \frac{\pi}{2} \right), \quad (4)$$

200 where $\ell_{B,i} = \sqrt{2D_i/\gamma}$ approximates the Batchelor scale for species i .

201 The system converges rapidly to the solution in Eq. 4 in the presence of advection.
202 However, when there is no turbulent advection, convergence is much slower, to the point
203 that an equilibrium assumption requires unrealistically long timeframes (see Section S3 in
204 the SI). We therefore need a time-dependent formula for the pcf in the absence of advection,
205 which can be obtained in the case where $\gamma = 0$ using a Green's function (see derivation in
206 the Appendices, Eqs. 31-37),

$$g_{ii}(r,t) = 1 + \frac{\lambda_i}{4\pi r D_i C_i} \left\{ 1 - \operatorname{erf} \left(\frac{r}{\sqrt{8D_i t}} \right) \right\}. \quad (5)$$

207 The above equations match when $\gamma \rightarrow 0$ and $t \rightarrow +\infty$.

208 As populations of different species do not directly interact, each population is an inde-
209 pendent realization of a point process, which means that the distribution of all individuals
210 within the community at time t is a random superposition of stationary point processes and
211 thus $g_{ij}(r,t) = 1$ if $i \neq j$ (Illian *et al.*, 2008, p. 326, eq. 5.3.13).

212 Related to the pair correlation function is Ripley's K -function $K(r)$. Using its marked
213 version, $C_j K_{ij}(r)$ is the average number of points of species j surrounding an individual of
214 species i within a sphere of radius r (Illian *et al.*, 2008), i.e.,

$$\forall r \geq 0, K_{ij}(r) = \frac{1}{C_j} \mathbb{E}_i (N_j(b(o,r) \setminus \{o\})), \quad (6)$$

215 where \mathbb{E}_i is the expectation with respect to individuals of species i and $N_j(b(o,r) \setminus \{o\})$ is

216 the number of individuals of species j in the sphere of radius r centered on individual o , not
217 counting individual o itself. $K_{ij}(r)$ is related to $g_{ij}(r)$ as

$$g_{ij}(r) = \frac{K'_{ij}(r)}{4\pi r^2}. \quad (7)$$

218 Combining Eq. 7 and, when $U > 0$, Eq. 4, we can show that (see Appendices, Eqs. 38-44)

$$K_{ii}(r) = \frac{4}{3}\pi r^3 + \frac{\lambda_i r^3}{3D_i C_i \ell_{B,i}} \left(\frac{\ell_{B,i}}{r} + \frac{\ell_{B,i}^3 \log\left(\frac{r^2}{\ell_{B,i}^2} + 1\right)}{2r^3} + \arctan\left(\frac{r}{\ell_{B,i}}\right) - \frac{\pi}{2} \right). \quad (8)$$

219 When $U = 0$, we need a time-dependent solution corresponding to our simulation dura-
220 tion, i.e. (see Appendices, Eq. 46-51)

$$K_{ii}(r, t) = \frac{4}{3}\pi r^3 + \frac{\lambda_i r^2}{C_i D_i} \left(\frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{r}{\sqrt{8D_i t}}\right) \left(1 - \frac{4D_i}{r^2}t \right) - \frac{\sqrt{2D_i t}}{\sqrt{\pi}r} e^{-\frac{r^2}{8D_i t}} \right). \quad (9)$$

221 For random superposition of stationary point processes, $K_{ij}(r, t) = \frac{4}{3}\pi r^3$ if $i \neq j$ (Illian
222 et al., 2008, p. 324, eq. 5.3.5).

223 Dominance index

224 The dominance index (defined in Table S1 in the Supporting Information of Wiegand *et al.*,
225 2007) is the ratio between the number of conspecifics and the number of individuals of all
226 species surrounding a given individual.

227 Let $M_{ij}(r)$ be the average number of individuals of species j within a circle of radius
228 r around an individual of species i , which can also be written with Ripley's K -function
229 as $M_{ij}(r) = C_j K_{ij}(r)$. $M_{ii}(r)$ corresponds to the conspecific neighbourhood and $M_{io}(r) =$
230 $\sum_{j=1, j \neq i}^S M_{ij}(r)$ corresponds to individuals of all other species. We can then define \mathcal{D}_i as

$$\begin{aligned} \mathcal{D}_i(r) &= \frac{M_{ii}(r)}{M_{ii}(r) + M_{io}(r)} \\ &= \frac{C_i K_{ii}(r)}{\sum_{j=1}^S C_j K_{ij}(r)}. \end{aligned} \quad (10)$$

231 When individuals of the same species i tend to cluster, $\mathcal{D}_i(r)$ tends to 1 while it tends to
232 the proportion of individuals of species i in the whole community when the distribution is
233 uniform (Section S2 of the SI).

234 Using Eq. 9 and 10, we obtain the formula for the dominance index in the presence of

235 advection as

$$\mathcal{D}_i(r) = -\frac{\frac{\lambda_i}{3D_i\ell_{B,i}} \left(\frac{\ell_{B,i}}{r} + \frac{\ell_{B,i}^3 \log\left(\frac{r^2}{\ell_{B,i}^2} + 1\right)}{2r^3} + \arctan\left(\frac{r}{\ell_{B,i}}\right) - \frac{\pi}{2} \right) + \frac{4}{3}\pi C_i}{\frac{\lambda_i}{3D_i\ell_{B,i}} \left(\frac{\ell_{B,i}}{r} + \frac{\ell_{B,i}^3 \log\left(\frac{r^2}{\ell_{B,i}^2} + 1\right)}{2r^3} + \arctan\left(\frac{r}{\ell_{B,i}}\right) - \frac{\pi}{2} \right) + \sum_{j=1}^S \frac{4}{3}\pi C_j}. \quad (11)$$

236 In the absence of advection ($U = 0, \gamma = 0$), we use the time-dependent dominance index,
237 computed similarly:

$$\mathcal{D}_i(r, t) = \frac{\frac{\lambda_i}{D_i r} \left(\frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{r}{\sqrt{8D_i t}}\right) \left(1 - \frac{4D_i}{r^2}t\right) - \frac{\sqrt{2D_i t}}{\sqrt{\pi r}} e^{-\frac{r^2}{8D_i t}} \right) + \frac{4}{3}\pi C_i}{\frac{\lambda_i}{D_i r} \left(\frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{r}{\sqrt{8D_i t}}\right) \left(1 - \frac{4D_i}{r^2}t\right) - \frac{\sqrt{2D_i t}}{\sqrt{\pi r}} e^{-\frac{r^2}{8D_i t}} \right) + \sum_{j=1}^S \frac{4}{3}\pi C_j}. \quad (12)$$

238 Parameters

239 We model two types of organisms: microphytoplankton (defined by a diameter between 20
240 and 200 μm , here 50 μm) and nanophytoplankton (defined by a diameter between 2 and 20
241 μm , here 3 μm). These two groups are characterized respectively by a low diffusivity, slow
242 growth and lower concentration vs. high diffusivity, fast growth and higher concentration.
243 Organisms are displaced by a turbulent fluid whose velocity defines the time scale of the
244 discretized model: we give here the reasoning behind parameter values, keeping in mind that
245 our model parameters are only approximate. Main parameter definitions and values are given
246 in Table 1.

247 Advection

248 We first consider the advection process, due to the turbulence of the environment. We
249 only consider the Batchelor-Kolmogorov regime, i.e., the size of the volume W is below the
250 size of the smallest eddy, but above the smallest length scale of fluctuations in nutrient
251 concentrations. The defining scale of the environment therefore corresponds to a Reynolds
252 number

$$\operatorname{Re} = \frac{U}{k\nu} \approx 1 \quad (13)$$

253 where $\nu = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ is the kinematic viscosity for water. The smallest wavenumber k
254 corresponds to the largest length scale L_s (Kolmogorov scale), i.e., $k = 2\pi/L_s$, with $L_s \approx 1$

255 cm in the ocean (Barton *et al.*, 2014). The definition of the Reynolds number leads to

$$\begin{aligned} 1 &\approx \frac{UL_s}{2\pi\nu} \\ \Leftrightarrow U &\approx \frac{2\pi\nu}{L_s}. \end{aligned} \quad (14)$$

256 This means that $U = 6.3 \times 10^{-4} \text{ m s}^{-1} = 5.4 \times 10^3 \text{ cm d}^{-1}$. Using $U\tau/3 = 0.5 \text{ cm}$ as in
 257 Young *et al.* (2001), we have $\tau = 2.8 \times 10^{-4} \text{ d} = 24 \text{ s}$. When $U\tau/3 = 0$, the environment is
 258 only diffusive, we keep the same value for τ . For $U\tau/3 = 0.5 \text{ cm}$, we estimate $\gamma = 1231 \text{ d}^{-1}$.

259 Diffusion

260 If we use the Stokes-Einstein equations (Einstein, 1905, cited from Dusenberry, 2009), diffu-
 261 sivity can be computed with

$$D_i = \frac{RT}{N_A} \frac{1}{6\pi\eta a_i} \quad (15)$$

262 where $R = 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$ is the molar gas constant, $T = 293 \text{ K}$ is the temperature,
 263 $N_A = 6.0225 \times 10^{23}$ is Avogadro's number, $\eta = 10^{-3} \text{ m}^{-1} \text{ kg s}^{-1}$ is the dynamic viscosity of
 264 water and a_i is the radius of the organism considered.

265 Using $D_i = \frac{\Delta_i^2}{2\tau}$, we find that

$$\begin{aligned} \Delta_i &= \sqrt{2\tau D_i} \\ \Leftrightarrow \Delta_i &= \sqrt{\frac{RT}{N_A} \frac{\tau}{3\pi\eta a_i}}. \end{aligned} \quad (16)$$

266 We consider $a_n = 1.5 \mu\text{m}$ for nanophytoplankton individuals and $a_m = 25 \mu\text{m}$ for micro-
 267 phytoplankton individuals, which allows us to compute Δ_n and Δ_m (see Table 1).

268 Ecological processes

269 We study the community at equilibrium, with the birth rate equal to the death rate, i.e.,
 270 $p_i = q_i \forall i$. We use a microphytoplankton doubling rate of 1 d^{-1} (Bissinger *et al.*, 2008) and
 271 consider the fastest-growing nanophytoplankton species, corresponding to a diameter of $3 \mu\text{m}$
 272 (Bec *et al.*, 2008), for which the doubling rate is between 2 and 3 d^{-1} (set to 2.5 d^{-1}
 273 here).

Parameter	Definition	Value
p_m, q_m	Probability of reproducing/dying for microphytoplankton	2.8×10^{-4}
p_n, q_n	Probability of reproducing/dying for nanophytoplankton	6.9×10^{-4}
U	Turbulent advection speed	$\{0, 0.06\} \text{ cm.s}^{-1}$
Δ_m	Diffusion parameter for microphytoplankton	$6.4 \times 10^{-5} \text{ cm}$
Δ_n	Diffusion parameter for nanophytoplankton	$2.6 \times 10^{-4} \text{ cm}$

Table 1: Definitions and values of the main parameters used in the three-dimensional BBM, assuming the duration of a time step τ is 24 seconds.

274 Range of interaction

275 As we examine individual aggregation and its potential effects on interactions between species,
 276 we have to ascertain the volume in which an individual can be affected by the presence of other
 277 individuals, or affect other individuals. We only consider here interactions due to competition
 278 for nutrients, and therefore need to define a nutrient depletion volume. We approximate this
 279 volume as the sphere of radius r where $C(r) \leq 90\%C_\infty$ with C_∞ the background concentration
 280 of the nutrient and $C(a_i) = 0$ (perfect absorption at the cell surface). The radius of this
 281 nutrient depletion volume is maximized when the individual is in stagnant water so that
 282 diffusion is the only hydrodynamic process. In this case, the depletion radius corresponds
 283 to 10 times the radius of the individual (Jumars *et al.*, 1993; Karp-Boss *et al.*, 1996). We
 284 define the maximum distance which allows for potential interactions (due to competition for
 285 resources) between two individuals of radius a_i and a_j as $d_{\text{threshold}}$, and the corresponding
 286 volume of potential interactions around an organism as $V_{\text{int}} = 4/3\pi d_{\text{threshold}}^3$ with

$$d_{\text{threshold}} = 10a_i + 10a_j. \quad (17)$$

287 We consider this maximum value as our baseline, keeping in mind that turbulence reduces
 288 the size of the nutrient depletion volume and increases the nutrient flux to the cell (Arnott
 289 *et al.*, 2021). We caution that determination of the shape of the nutrient depletion volume
 290 in the presence of turbulence is too complex to be addressed here (Karp-Boss *et al.*, 1996).

291 We consider a total volume of 1000 cm^3 for microphytoplankton and 10 cm^3 for nanophy-
 292 toplankton (volumes are adapted to balance realistic concentrations and computation time)
 293 with periodic boundary conditions. Individuals are uniformly distributed in the cube at
 294 the beginning of the simulation. We run an idealized simulation with 3 species with an
 295 even abundance distribution of about $10^4 \text{ cells L}^{-1}$ for microphytoplankton (Picoche & Bar-
 296 raquand, 2020) and $10^6 \text{ cells L}^{-1}$ for nanophytoplankton individuals (Edwards, 2019). We
 297 then model a more realistic community with 10 species having a skewed abundance dis-
 298 tribution (between 55,000 and 400 cells L^{-1} for microphytoplankton, according to obser-

299 vations of field abundance distributions in Picoche & Barraquand, 2020, and multiplied
300 by 10^2 for nanophytoplankton). All simulations are run for 1000 time steps of duration
301 τ (corresponding to approximately 6h40 of phytoplankton time—note that runtimes can
302 be much longer). The computation of g and K for simulated distributions is explained in
303 Section S4 of the SI. The code for all simulations and analyses can be found at https://github.com/CoraliePicoche/brownian_bug_3D/.
304

305 Results

306 We show an example of nanophytoplankton spatial distributions with and without advection
307 at the end of a simulation in Fig. 1: clustering is not visible to the naked eye, even when
308 zooming in on the observation volume, in the presence of advection, but removing turbulence
309 helps visualising small aggregates of conspecifics. Microphytoplankton distributions are not
310 straightforward to interpret as no clusters can be detected visually (although they may ac-
311 tually be present), whether advection is included or not (Section S5 of the SI). Statistics are
312 therefore needed to go further in detecting patterns of aggregation.

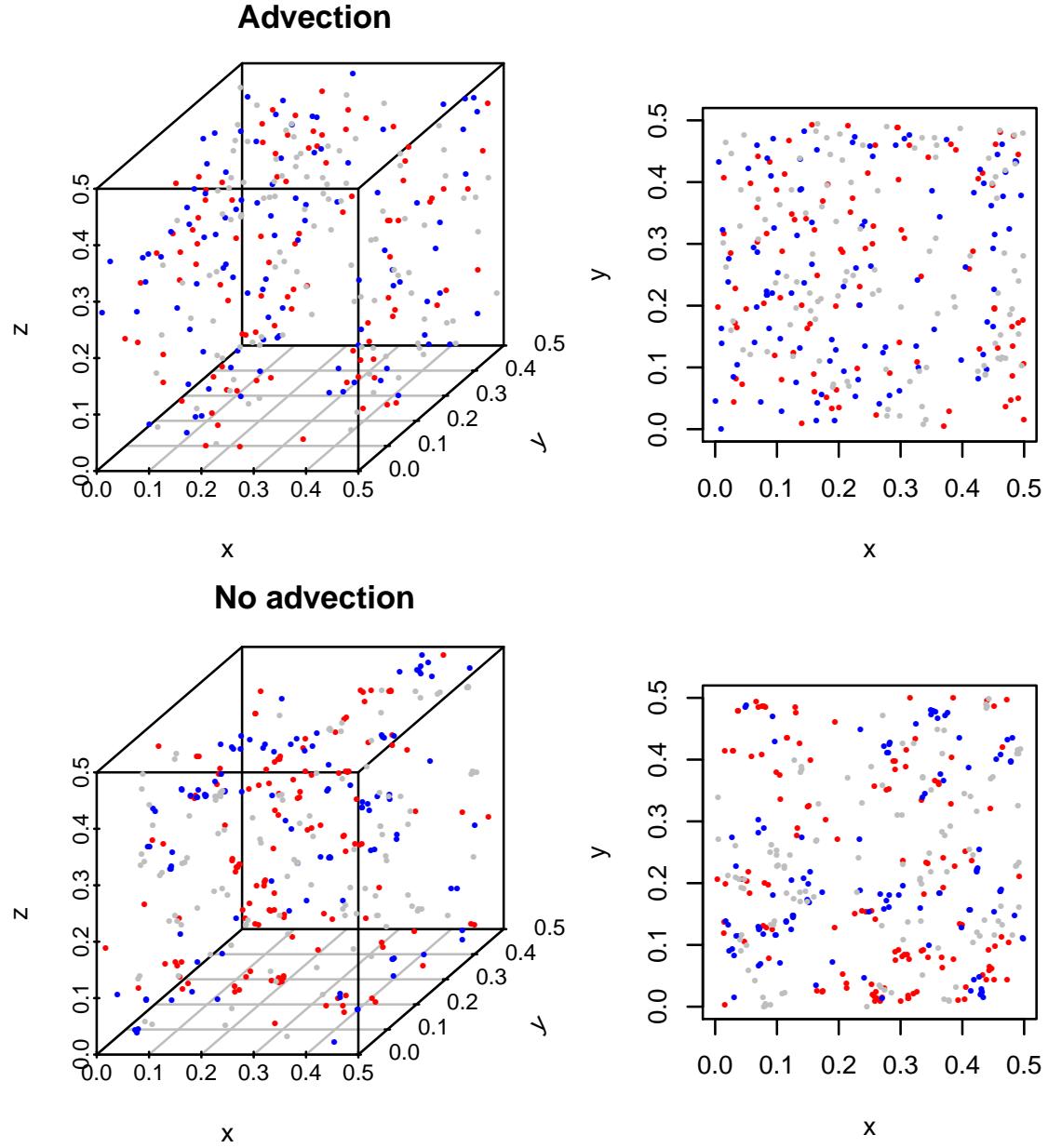


Figure 1: Spatial distributions of a 3-species community of nanophytoplankton with and without advection with density $C = 10^3$ cells cm^{-3} after 1000 time steps. Each color corresponds to a different species. On the left-hand side, only a zoom on a $0.5 \times 0.5 \times 0.5 \text{ cm}^3$ cube is shown, and its projection on the x-y plane is shown on the right-hand side.

313 Ripley's K -functions extracted from numerical simulations match theoretical formula
 314 (Fig. 2) for both types of organisms, which also indicates that dominance indices extracted
 315 from the simulations match theoretical expectations.

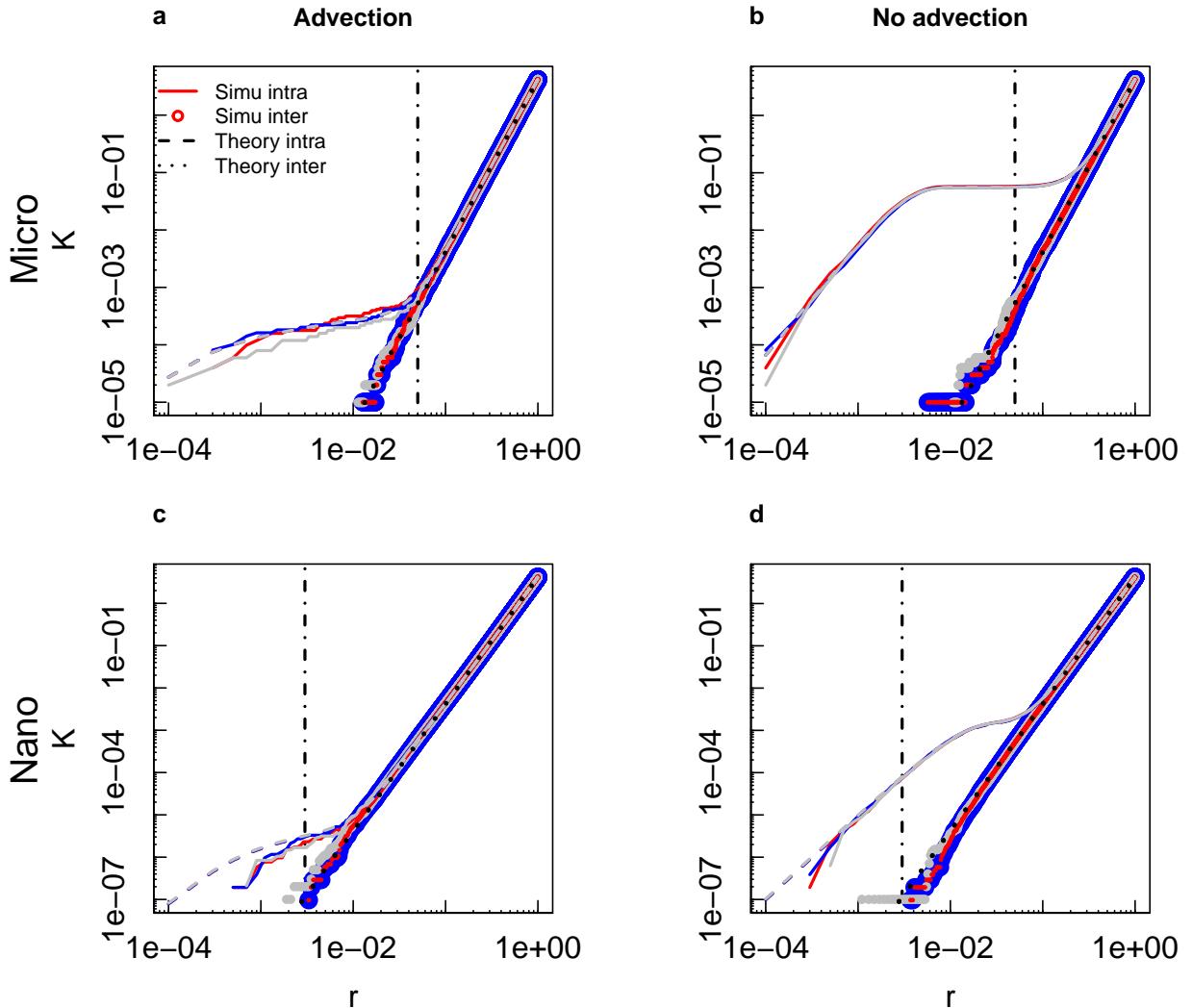


Figure 2: Comparison of theoretical and simulated Ripley’s K -functions as a function of distance (in cm) for microphytoplankton (a-b) and nanophytoplankton (c-d) in a 3-species community with even abundance distributions after 1000 timesteps, with (a, c) and without (b, d) advection. Each color represents a different species. Intraspecific K -functions are shown with dashed (theoretical values) and solid (simulated values) lines. Interspecific K -functions are shown with dotted lines (theoretical values) and circles (simulated values). The black dash-dotted line corresponds to the threshold considered as the maximum distance for nutrient-based competition.

316 Dominance indices all follow a similar pattern (Fig. 3 and 4). The dominance index is
 317 close to 1 for small distances: there is always a scale at which an organism is surrounded
 318 almost only by conspecifics. The index then decreases sharply to converge at large distances
 319 (close to 1 cm) to the proportion of the focus species in the whole community, as it would for
 320 a uniform spatial distribution. Patterns differ at intermediate ranges of distances between

321 organisms.

322 In the presence of advection, the dominance index starts decreasing for a distance ap-
323 proximately 10 times smaller than when advection is absent, which indicates that organisms
324 are closer to heterospecifics when their environment is turbulent. A quasi-uniform distribu-
325 tion is also reached for smaller distances with advection than without. Microphytoplankton
326 species start mixing for distances larger than for nanophytoplankton species irrespective of
327 the hydrodynamic regime surrounding them.

328 In a 3-species community with the same initial abundances, in the presence of advection,
329 microphytoplankton dominance indices are between 0.37 and 0.47 at the distance thresh-
330 old for potential interactions, while they are between 0.80 and 0.94 for nanophytoplankton
331 species. In the absence of turbulence, dominance indices are all above 0.98 when the distance
332 threshold is reached (Fig. 3). Microphytoplankton organisms are therefore as likely to share
333 their depletion volume with conspecifics as they are with heterospecifics, but only when tur-
334 bulent advection is accounted for, whereas nanophytoplankton organisms always have almost
335 only conspecifics around them.

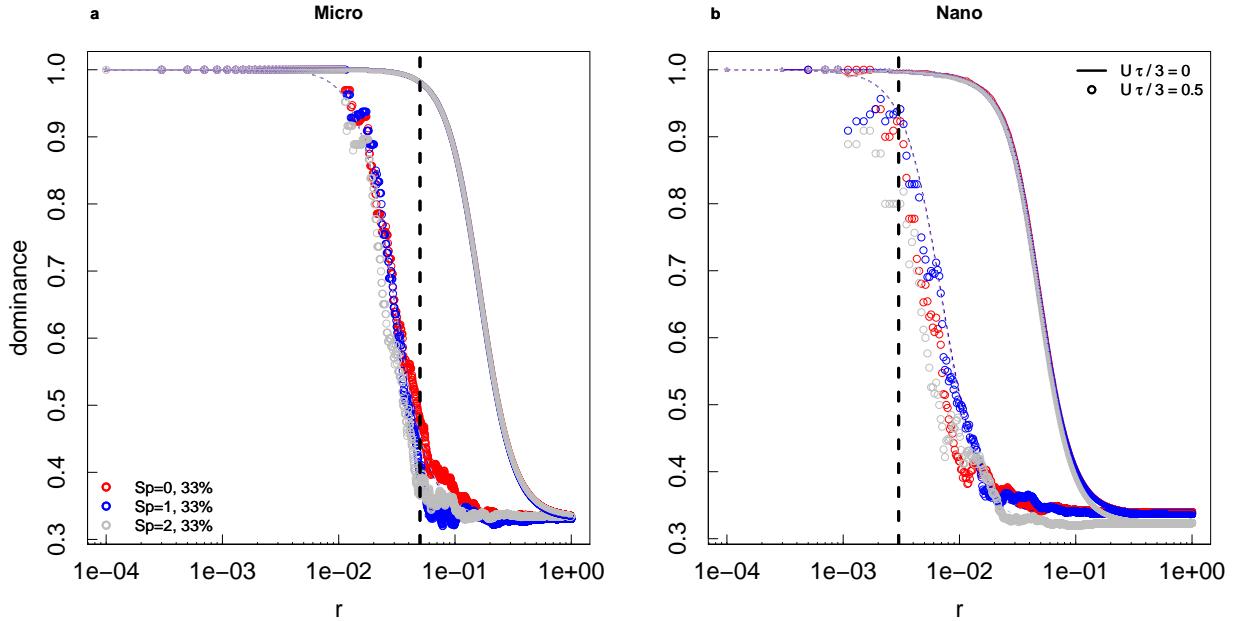


Figure 3: Dominance indices as a function of distance (in cm) for microphytoplankton (a) and nanophytoplankton (b) in a 3-species community with even abundance distributions (final proportions in the community are indicated in the figure) after 1000 timesteps, with (circles) and without (lines) advection. Each color represents a different species. The grey dashed curve represents the analytical solution. The black dashed line corresponds to the threshold considered as the maximum distance for nutrient-based competition.

More mixing in microphytoplankton than nanophytoplankton, and more mixing with advection, also holds when considering a 10 species-community with a skewed abundance distribution (Fig. 4), but dominance indices are overall lower in communities with more species and with less even abundances. In the presence of advection, microphytoplankton dominance indices at the distance threshold are between 0.34 (for the most abundant species) and 0.033 (for one of the least abundant species), while they are between 0.90 and 0.85 when advection is not taken into account. Nanophytoplankton species, too, are more mixed than in the 3 species-community: dominance indices vary between 0.54 and 0.2 when the depletion threshold is reached (with an exception of 0 for one particular species which had no conspecific for distances below 10^{-2} cm) when organisms are displaced by turbulence, while the same quantity is between 1 and 0.97 when they are only subject to diffusion.

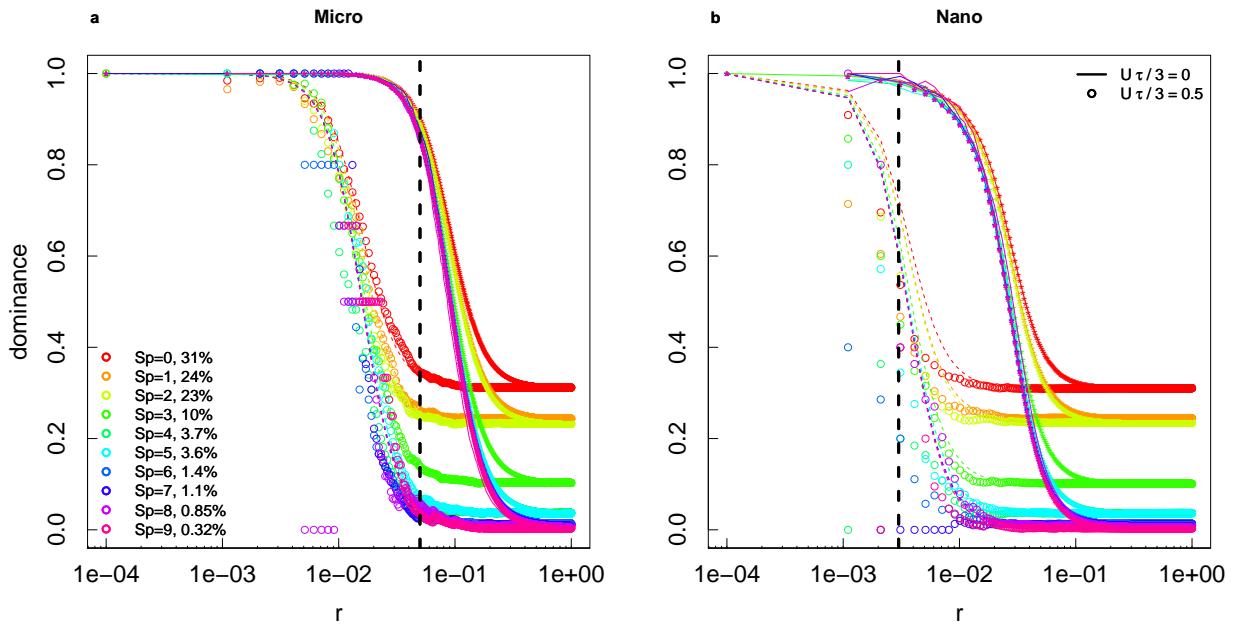


Figure 4: Dominance indices as a function of distance (in cm) for microphytoplankton (a) and nanophytoplankton (b) in a 10-species community with a skewed abundance distribution (final proportions in the community are indicated in the figure) after 1000 timesteps, with (circles) and without (lines) advection. Each color represents a different species. The coloured dashed curves (advection) and small stars (no advection) represent the analytical solution. The black dashed line corresponds to the threshold considered as the maximum distance for nutrient-based competition.

Differences in spatial distributions are not only due to organism sizes, which determine their demographic and hydrodynamic properties, but also to their abundances (here set through initial values). In the presence of turbulence, the threshold distance at which domi-

nance falls below 95% is smaller for more abundant species (Fig. 5 a-b). Abundant species tend to be present nearly everywhere when they are mixed in the environment. Therefore, they are also more likely to be close to a heterospecific, but still have more conspecifics close to them than the less abundant species ($\mathcal{D}(d_{\text{threshold}})$ increases with abundance, Fig. 5 c-d). However, this increase is less clear for nanophytoplankton than for microphytoplankton (Fig. 5 c-d). When turbulence is absent, the relationships with abundance are unclear, possibly affected by sampling effects, and we refrain from interpreting them.

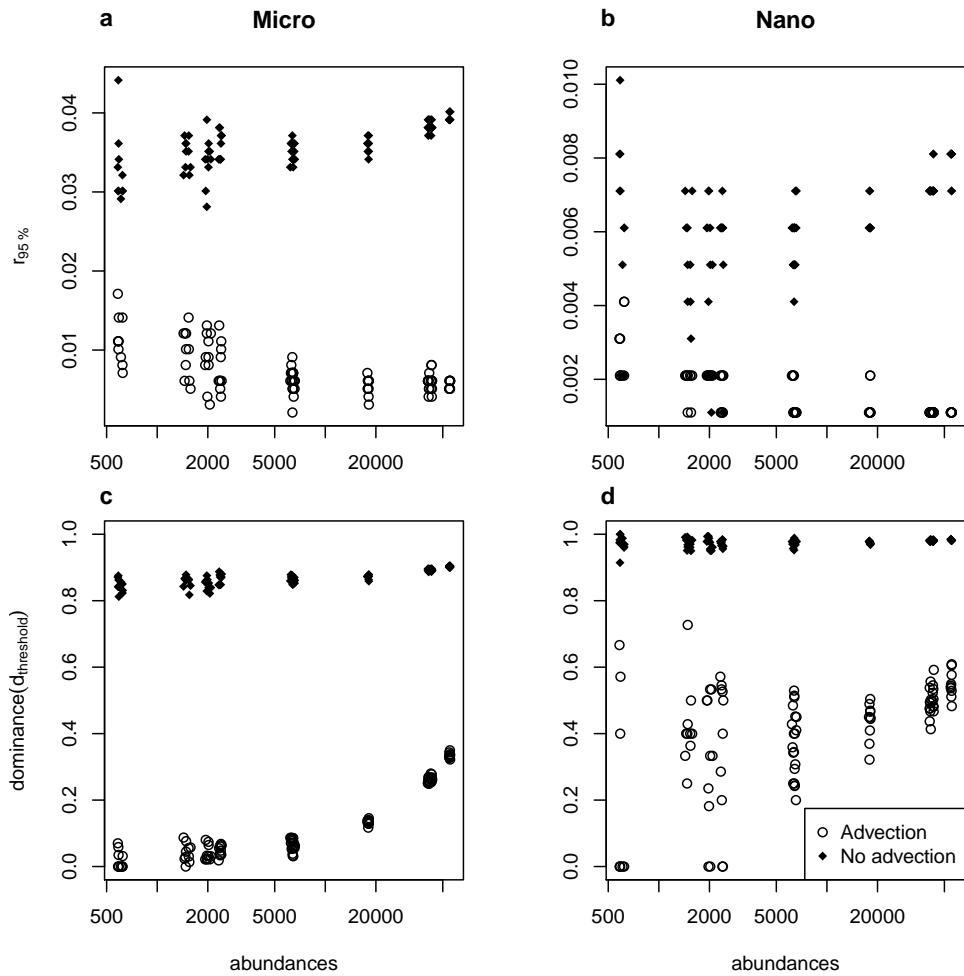


Figure 5: Minimum distances (in cm) between points for dominance to drop below 95% (a and b) and dominance at a distance corresponding to the threshold for competition (c and d) as a function of abundances (note the logarithmic scale on the x-axis) for microphytoplankton and nanophytoplankton. We consider cases with and without advection in a 10-species community with a skewed abundance distribution. These have been obtained combining 10 sets of simulations.

357 **Discussion**

358 We designed a stochastic, three-dimensional, individual-based model of the spatial distribu-
359 tion of multiple species in a viscous and turbulent flow. We conducted both mathematical
360 analyses and numerical simulations to quantify spatial correlations in the distribution of or-
361 ganisms. We focused on the pair correlation function and Ripley's K -function, for which nu-
362 mercial and theoretical analyses showed a good agreement, and extracted a more ecologically-
363 oriented metric from them, i.e., the dominance index. This statistic is the *local* average ratio
364 of conspecifics, i.e., the number of organisms of the focal species in the neighbourhood of an
365 individual of the same species, divided by the total number of organisms in that neighbour-
366 hood. Intraspecific clustering corresponds to a dominance index close to 1, which decreases
367 when interspecific mixing increases. The choice of this index was motivated by two reasons:
368 (1) it is at its core a proportion of a focal species in a certain volume, i.e. a scale-dependent,
369 localized metric bounded between 0 and 1 as opposed to other statistics whose values are less
370 directly interpreted, and (2) it is easy to relate to coexistence theory as it describes the envi-
371 ronment of an organism in terms of heterospecifics and conspecifics, which can, under certain
372 assumptions that we discuss below, be related to interspecific and intraspecific interactions.
373 Comparing the distributions of organisms of different sizes, we showed that the presence of
374 turbulence always increased mixing (results are robust to slight modifications in the com-
375 putation of advection velocity U , shown in Section S6 of the SI). The species composition
376 around an organism depended on its size, which mechanically determines its hydrodynamic
377 properties (diffusivity), and is linked with its ecological characteristics (growth rate and den-
378 sity). Microphytoplankters (20 to 200 μm), larger cells with lower diffusivity, growth rate and
379 abundance, were on average further away from other cells, due to their lower concentrations
380 (Fig. S11 of the SI), than nanophytoplankters (2 to 20 μm). They were surrounded by more
381 heterospecifics than conspecifics within a volume of potential interactions, whose radius is
382 defined as the maximum distance for which nutrient depletion volumes of two different indi-
383 viduals may overlap. If we consider that interactions between species (not modelled directly
384 here because of timescale issues, see below) could occur with equal probability at all distances
385 within the volume of potential interactions, we would conclude that microphytoplankters are
386 more likely to interact with individuals from other species than with individuals of their own
387 species. This affirmation is, however, conditional upon interactions at 10 cell diameters from
388 an individual being equally likely than at 1 diameter from an individual. If we keep in mind
389 that interactions are more likely or stronger at very short distances, or that the maximal
390 radius of interaction could be shorter than our estimation and advection velocity U lower (SI
391 Section S6), microphytoplankters may still experience more frequent effects of conspecifics

392 than heterospecifics.

393 To see this, let us first focus on the smallest distances between organisms. The nearest
394 neighbour of an organism was always an organism of the same species, and the minimum
395 distance between conspecifics was always lower than expected for a uniform distribution (Sec-
396 tion S7 of the SI). The dominance index remained close to 1 for distances below 10^{-2} cm or
397 10^{-3} cm for microphytoplankton and nanophytoplankton respectively. There was therefore
398 always *some* intraspecific aggregation, i.e. conspecifics were always closer than heterospecifics
399 at the smallest distances. This is due to the prevalence of demographic processes at indi-
400 vidual scales, because an individual acts as a source point for other organisms of the same
401 species, and hydrodynamic processes do not separate conspecifics fast enough to prevent ag-
402 gregation. This remains true if we add an initial separation distance between mother and
403 daughter cells upon birth (additional simulations, see code repository). If we consider that
404 interaction strengths are a smoothly decaying function of distance, a common assumption
405 in spatial coexistence models (e.g., Bolker & Pacala, 1999; Law *et al.*, 2003), this implies
406 that population-level intraspecific interactions could be stronger than interspecific interac-
407 tions due to intraspecific micro-scale aggregation. However, the mechanisms of competition
408 at this scale are poorly known, likely relying on multiple types of resources with different
409 distributions in the environment, effects on the cell, uptakes, etc. Rather than weighting
410 much more heavily the potential interactions with the closest neighbour(s) through an in-
411 teraction kernel, we therefore chose conservatively to define a maximum distance for two
412 organisms to possibly affect the concentrations of elements in the environment of each other,
413 assuming perfect absorption on the cell surface. We consider that, at all distances below this
414 threshold, interactions could happen between organisms. We continue the discussion with
415 that simplification in mind, and explicitly mention when it is relaxed.

416 Dominance indices began to decrease at distances above 10^{-3} cm, still below the maximum
417 distance for interactions. At this distance and above, the balance between heterospecifics and
418 conspecifics was much more sensitive to different phytoplankters' demographic and hydro-
419 dynamic traits. The species composition of an organism's neighbourhood depended on its
420 size: nanophytoplankton organisms mainly shared their volume of potential interactions with
421 conspecifics (the dominance index remained close to 1, even near the distance threshold, i.e.
422 the maximum distance for the overlap of nutrient depletion volumes) while microphytoplank-
423 ton organisms could affect both conspecifics and heterospecifics (the dominance index was
424 often below 0.5 at the distance threshold, i.e. an individual's depletion zone probably over-
425 lapped with more heterospecifics' than conspecifics'). Microphytoplankters were therefore
426 more likely to share their depletion volume with heterospecifics than nanophytoplankters.
427 The rate of production of new microphytoplankton conspecifics was not sufficient to com-

pensate for the mixing induced by turbulence and diffusivity, even though the diffusivity of microphytoplankters was smaller than that of nanophytoplankters. There may therefore be different mechanisms at play at the community level for microphytoplankton and nanophytoplankton to maintain coexistence. For nanophytoplankton, the spatial structure likely leads to more interactions between conspecifics than between heterospecifics. The spatial distribution of microphytoplankton species, on the contrary, encourages more interactions between heterospecifics. If we consider that local interaction strengths are equal within the volume of potential interactions, scaling to the population level, we would likely observe stronger intra- over interspecific interactions for nanophytoplankton (a key factor in coexistence theory, Barabás *et al.*, 2017) but not necessarily so for microphytoplankton. Using a timescale separation argument, we show in Section S8 in the SI how stronger interactions at population level than individual level may arise in a Lotka-Volterra model whose spatial structure is summed up by the dominance indices evidenced here. Stronger intra- than interspecific competition may arise at population level even when assuming that all local interaction strengths between individuals are equal, regardless of the identity of competitors.

All of the above discussion is based on a microphytoplankter's neighbourhood in its nutrient depletion volume. To simplify the computation, we used maximum volumes of potential interactions, corresponding to a diffusive-only flow of nutrient particles. But when fluid turbulence increases, nutrient uptake increases, and the size of the depletion zone decreases (Karp-Boss *et al.*, 1996). The proportion of change in the depletion volume increases with the size of organisms: a 10 μm -diameter organism might not experience any change, while the uptake of a 100 μm -diameter organism would increase by at least 50% (Karp-Boss *et al.*, 1996). Therefore the volume of potential interactions shrinks in the presence of turbulence for microphytoplankton, but not necessarily for nanophytoplankton. An additional reason why microphytoplankters might still be surrounded by conspecifics at ecologically meaningful distances and interacting more frequently with them is imperfect absorption of nutrients: if nutrient concentration at the cell surface is not zero but C_0 , then the radius of interaction is $10a_i(1 - C_0/C_\infty)$.

Up to now, we have only focused on the dominance index, a localized proportion of conspecifics. However, interactions also depend on the absolute densities of individuals. Mechanically, when density decreases, the distances between neighbours increase, which explains that the distances between the low-abundance microphytoplankters tended to be greater than distances between the more abundant nanophytoplankters (Section S7 of the SI). Explicit mathematical models using pair densities to express interaction rates (e.g. Law *et al.*, 2003; Plank & Law, 2015) may be able to incorporate those effects; however, as we highlight below, the timescales and spatial correlations that are seen in such models may not necessarily

464 represent faithfully phytoplankton community dynamics.

465 Contrary to other similar models (e.g., Birch & Young, 2006; Bouderbala *et al.*, 2018),
466 we did not consider explicit effects of local density on survival and fertility rates. Outside of
467 simply maintaining analytical tractability, we had another, more biological reason to do so:
468 we cannot be sure that these local density-dependencies make sense in our phytoplankton
469 context. To understand why, consider that even if a species abundance is locally tripled,
470 competition might not directly ensue at the time scales covered by our model (≈ 7 h),
471 if nutrient depletion has not had time to set in yet. Even if we considered longer time
472 frames, we would need lagged local density-dependencies, which are to our knowledge not
473 leading to tractable spatial branching or dynamic point processes. We could, of course,
474 directly model nutrients, perhaps as resource “points” with a dynamics of their own (Murrell,
475 2005; North & Ovaskainen, 2007), which in turn change the reproduction or death rate
476 of individuals. If the resource points risk being depleted, this entails a negative spatial
477 correlation between organisms and their resources (Murrell, 2005; Barraquand & Murrell,
478 2012). And that is where such models might be inadequate. The phycosphere, a micro-
479 environment at the periphery of a phytoplankton organism where communities of bacteria
480 interact (Seymour *et al.*, 2017), can also impact phytoplankton fitness, both positively (cross-
481 feeding) and negatively (algicidal activities of bacteria). This can sometimes lead to an
482 accumulation of key resources close to the phytoplankter. This will lead to positive spatial
483 correlations between consumers and their resources, and we currently do not have theoretical
484 models to represent this process (short of modelling precisely the spatial distribution of these
485 bacteria).

486 Our model should be viewed as a first model of spatial distributions of multiple phyto-
487 plankton species in a realistic, three-dimensional environment at the microscale, describing
488 only basic hydrodynamic and demographic processes. Using this model, we were able to
489 predict whether phytoplankters could be in contact with individuals of their own or other
490 species, and form reasonable conjectures regarding potential intra vs interspecific interac-
491 tions between species, emerging at the population level through spatial distributions (Detto
492 & Muller-Landau, 2016). It is worthwhile to keep in mind that there are many remaining
493 features of phytoplankton physiology and life histories which we do not address here, but
494 which may affect spatial distributions. Many phytoplankters are able to move actively in
495 three dimensions, which can favour cluster formation (Breier *et al.*, 2018). Even those who
496 are believed to move passively actually often move along the vertical dimension by regulat-
497 ing their buoyancy (Reynolds, 2006), and can at times aggregate to form pairs (Font-Muñoz
498 *et al.*, 2019). Finally, a part of spatial structure is explained by the partially colonial nature
499 of microphytoplankton (Kiørboe *et al.*, 1990). This clearly calls for viewing our model as

500 a null model to which more complex mechanistic models and their spatial outputs can be
 501 compared.

502 **Acknowledgements**

503 FB and CP were supported by the grant ANR-20-CE45-0004. CP was supported by a PhD
 504 grant from the French Ministry of Research. We are thankful for constructive reviewer
 505 feedback.

506 **Appendices**

507 **Derivation of the spatial characteristics of the Brownian Bug Model**

508 We show here how to compute the monospecific pair correlation function and Ripley's K -
 509 function of the Brownian Bug Model (see Young *et al.*, 2001 and Picoche *et al.*, 2022 for a
 510 detailed derivation from the master equation). As these formula only apply to intraspecies
 511 pairs, we ignore species' index in the following for the sake of clarity. Similar formula for
 512 well-known spatial point processes are given in the Supplementary Information, for readers
 513 who want to understand better the properties of these spatial statistics.

514 **Proof of Eq. 4 and Eq. 5**

515 In three dimensions, when the birth rate λ is the same as the mortality rate μ , the pair
 516 density $G(r)$ is a solution of

$$\frac{\partial G}{\partial t} = \frac{2D}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r^2} \frac{\partial}{\partial r} \left(r^4 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\xi). \quad (18)$$

Steady-state solution We first compute the steady-state solution, *i.e.*

$$\begin{aligned} 0 &= \frac{2D}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r^2} \frac{\partial}{\partial r} \left(r^4 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\xi) \\ 0 &= 4\pi r^2 \left(\frac{2D}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + \frac{\gamma}{r^2} \frac{\partial}{\partial r} \left(r^4 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\xi) \right) \\ 0 &= 4\pi \left(2D \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + \gamma \frac{\partial}{\partial r} \left(r^4 \frac{\partial G}{\partial r} \right) \right) + 4\pi r^2 2\lambda C \delta(\xi). \end{aligned} \quad (19)$$

We can then integrate Eq. 18 over a small sphere centered on an individual, with radius ρ . Let us first note that in an isotropic environment the 3D-Dirac relates to the radial one as

$$\delta(\boldsymbol{\xi}) = \frac{1}{4\pi r^2} \delta(r) \quad (20)$$

with $4\pi r^2$ the surface of the sphere of radius r . Using Eq. 19 and 20,

$$\begin{aligned} 0 &= 4\pi \left(2Dr^2 \frac{\partial G}{\partial r} + \gamma r^4 \frac{\partial G}{\partial r} \right) + 2\lambda C \\ \Leftrightarrow \frac{\partial G}{\partial r} &= -\frac{1}{4\pi} \frac{2\lambda C}{2Dr^2 + \gamma r^4}. \end{aligned} \quad (21)$$

We can integrate Eq. 21 between ρ and ∞ . As $G(\infty) = C^2$,

$$C^2 - G(\rho) = -\frac{\lambda C}{2\pi} \int_{\rho}^{\infty} \frac{1}{2Dr^2 + \gamma r^4} dr. \quad (22)$$

We first compute the primitive $A = \int \frac{1}{2Dr^2 + \gamma r^4} dr$.

$$A = \int \frac{1}{r^2(2D + \gamma r^2)} dr \quad (23)$$

$$= \int \frac{1}{2Dr^2} - \frac{\gamma}{2D(2D + \gamma r^2)} dr \quad (24)$$

$$= -\frac{1}{2Dr} - \frac{\gamma}{2D} \int \frac{1}{2D \left(1 + \left(\sqrt{\frac{\gamma}{2D}} r \right)^2 \right)} dr. \quad (25)$$

⁵¹⁷ With a change of variable $u = \sqrt{\frac{\gamma}{2D}} r$, using $\int \frac{1}{1+u^2} = \arctan(u)$, we have

$$A = -\frac{1}{2Dr} - \frac{\sqrt{\gamma} \arctan \left(\frac{\sqrt{\gamma}r}{\sqrt{2D}} \right)}{2\sqrt{2D}\sqrt{D}} + K \quad (26)$$

⁵¹⁸ where K is a constant. We can now compute $B = [A]_{\rho}^{\infty}$.

$$B = -\frac{\sqrt{\gamma}\pi}{4\sqrt{2D}\sqrt{D}} + \frac{1}{2D\rho} + \frac{\sqrt{\gamma} \arctan \left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}} \right)}{2\sqrt{2D}\sqrt{D}}. \quad (27)$$

This leads to

$$G(\rho) = C^2 + \frac{\lambda C}{2\pi} B \quad (28)$$

$$= C^2 + \frac{\lambda C}{2\pi} \left[\frac{1}{2D\rho} + \frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}}\right)}{2\sqrt{2D}\sqrt{D}} - \frac{\sqrt{\gamma}\pi}{4\sqrt{2D}\sqrt{D}} \right]. \quad (29)$$

⁵¹⁹ Finally, the pair correlation function $g = G/C^2$ is defined as

$$g(\rho) = \frac{\lambda}{4\pi CD} \left(\frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}}\right)}{\sqrt{2D}} + \frac{1}{\rho} - \frac{\pi\sqrt{\gamma}}{2\sqrt{2D}} \right) + 1. \quad (30)$$

⁵²⁰ **Time-dependent solution** In the absence of advection by turbulent diffusion ($U = 0$, $\gamma = 0$), convergence to the steady-state solution can be very slow (more than a week, see Section ⁵²¹ S3 in the SI). In order to keep a realistic timeframe, we need to compute a time-dependent ⁵²² solution. We can get back to Eq. 18 with $\gamma = 0$, which yields ⁵²³

$$\frac{\partial G}{\partial t} = \frac{2D}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + 2\lambda C \delta(\xi). \quad (31)$$

⁵²⁴ Assuming an isotropic environment, this means

$$\frac{\partial G}{\partial t} - 2D\Delta G = 2\lambda C \delta(\xi) \quad (32)$$

⁵²⁵ where $\Delta = \nabla^2$ is the Laplacian operator. We therefore have

$$\mathcal{L}G(\xi, t) = 2\lambda C \delta(\xi) \quad (33)$$

⁵²⁶ where \mathcal{L} is the linear differential operator $\partial_t - 2D\Delta$. Therefore $G(y) = \int H(y, s)2\lambda C \delta(s)ds$ ⁵²⁷ where $H(y, s) = H(y - s)$ is the Green kernel (heat kernel). We can therefore write

$$\begin{aligned} G(\xi, t) &= 2\lambda C \int_{\mathbb{R}^3} \int_0^t H(\xi - \xi', t') \delta(\xi') d\xi' dt' \\ \Leftrightarrow G(\xi, t) &= 2\lambda C \int_0^t H(\xi, t') dt'. \end{aligned} \quad (34)$$

⁵²⁸ A solution for the Green's function using $\mathcal{L} = \partial_t - 2D\Delta$ in three dimensions is $H(r, t) =$ ⁵²⁹ $\left(\frac{1}{8\pi Dt}\right)^{3/2} \exp\left(-\frac{r^2}{8Dt}\right)$. $G(r, t)$ can then be computed as

$$G(r, t) = 2\lambda C \left(\frac{-\operatorname{erf}\left(\frac{r}{\sqrt{8tD}}\right)}{8\pi Dr} + K \right) \quad (35)$$

⁵³⁰ where erf is the error function. Using $G(r, 0) = C^2$ and $\lim_{x \rightarrow +\infty} \text{erf}(x) = 1$ in Eq. 35,

$$\begin{aligned} C^2 &= 2\lambda C \left(\frac{-1}{8\pi Dr} + K \right) \\ \Leftrightarrow \frac{C}{2\lambda} + \frac{1}{8\pi Dr} &= K. \end{aligned} \quad (36)$$

⁵³¹ We can finally compute $G(r, t)$:

$$\begin{aligned} G(r, t) &= 2\lambda C \left(-\frac{\text{erf}\left(\frac{r}{\sqrt{8Dt}}\right)}{8\pi Dr} + \frac{C}{2\lambda} + \frac{1}{8D\pi r} \right) \\ &= \frac{\lambda C}{4\pi Dr} \left\{ 1 - \text{erf}\left(\frac{r}{\sqrt{8Dt}}\right) \right\} + C^2 \\ \Leftrightarrow g(r, t) &= \frac{\lambda}{4D\pi r C} \left\{ 1 - \text{erf}\left(\frac{r}{\sqrt{8Dt}}\right) \right\} + 1. \end{aligned} \quad (37)$$

⁵³² Proof of Eq. 8 and Eq. 9

⁵³³ We can integrate thepcf formula to compute Ripley's K -function, as $g(r) = \frac{K'(r)}{4\pi r^2}$.

⁵³⁴ **Steady-state solution** From Eq. 30,

$$K(\rho) = 4\pi \int_0^\rho r^2 + \frac{\lambda}{2\pi C} \left[\frac{r}{2D} + \frac{\sqrt{\gamma}r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right)}{2\sqrt{2D}\sqrt{D}} - \frac{\sqrt{\gamma}\pi r^2}{4\sqrt{2D}\sqrt{D}} \right] dr. \quad (38)$$

⁵³⁵ We define $A = \int_0^\rho r^2 dr$, $B = \int_0^\rho \frac{r}{2D} dr$, $C = \int_0^\rho r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right) dr$ and $E = \int_0^\rho \frac{\sqrt{\gamma}\pi r^2}{4\sqrt{2D}\sqrt{D}} dr$.

$$\begin{aligned} A &= \frac{1}{3}\rho^3. \\ B &= \frac{\rho^2}{4D}. \\ E &= \frac{\sqrt{\gamma}\pi\rho^3}{12\sqrt{2D}\sqrt{D}}. \end{aligned} \quad (39)$$

⁵³⁶ We can also compute $C = \int_0^\rho r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right) dr$. We first change variable, with $u = \frac{r}{\sqrt{2D}}$,
⁵³⁷ $dr = \sqrt{2D}du$, and obtain

$$C = (2D)^{3/2} \int_0^{\rho/\sqrt{2D}} u^2 \arctan(\sqrt{\gamma}u) du. \quad (40)$$

⁵³⁸ We can integrate by parts, with $f = \arctan(\sqrt{\gamma}u)$ and $g' = u^2$, which leads to

$$C = (2D)^{3/2} \left(\frac{\rho^3}{3(2D)^{3/2}} \arctan\left(\sqrt{\frac{\gamma}{2D}}\rho\right) - \frac{\sqrt{\gamma}}{3} \int_0^{\rho/\sqrt{2D}} \frac{u^3}{(\gamma u^2 + 1)} du \right). \quad (41)$$

⁵³⁹ We then substitute $v = \gamma u^2 + 1$, $du = \frac{1}{2\gamma u} dv$, and have

$$\begin{aligned} \int_0^{\rho/\sqrt{2D}} \frac{u^3}{(\gamma u^2 + 1)} du &= \frac{1}{2\gamma^2} \int_1^{\gamma\rho^2/2D+1} \frac{v-1}{v} dv \\ &= \frac{1}{2\gamma^2} \int_1^{\gamma\rho^2/2D+1} 1 - \frac{1}{v} dv \\ &= \frac{1}{2\gamma^2} \left(\gamma \frac{\rho^2}{2D} - \log\left(\gamma \frac{\rho^2}{2D} + 1\right) \right). \end{aligned} \quad (42)$$

⁵⁴⁰ Going back to C, we obtain

$$\begin{aligned} C &= \frac{\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}} \rho)}{3} - (2D)^{3/2} \frac{\sqrt{\gamma}}{3} \frac{1}{2\gamma^2} \left(\frac{\gamma}{2D} \rho^2 - \log\left(\gamma \frac{\rho^2}{2D} + 1\right) \right) \\ &= \frac{\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}} \rho)}{3} - \frac{\sqrt{2D}}{6\sqrt{\gamma}} \rho^2 + \frac{\sqrt{2D}^{3/2}}{3\gamma^{3/2}} \log\left(\gamma \frac{\rho^2}{2D} + 1\right). \end{aligned} \quad (43)$$

⁵⁴¹ Combining all equations,

$$\begin{aligned} K(\rho) &= \frac{4}{3}\pi\rho^3 + \frac{2\lambda}{C} \left(\frac{\rho^2}{4D} + \frac{\sqrt{\gamma}\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}} \rho)}{6\sqrt{2D}^{3/2}} - \frac{\rho^2}{12D} + \frac{\log\left(\gamma \frac{\rho^2}{2D} + 1\right)}{6\gamma} - \frac{\sqrt{\gamma}\pi\rho^3}{12\sqrt{2D}\sqrt{D}} \right) \\ &= \frac{4}{3}\pi\rho^3 + \frac{\lambda}{3C} \left(\frac{\rho^2}{D} + \frac{\sqrt{\gamma}\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}} \rho)}{\sqrt{2D}^{3/2}} + \frac{\log\left(\gamma \frac{\rho^2}{2D} + 1\right)}{\gamma} - \frac{\sqrt{\gamma}\pi\rho^3}{2\sqrt{2D}\sqrt{D}} \right). \end{aligned} \quad (44)$$

⁵⁴² Note that in the absence of advection,

$$\begin{aligned} g(r) &= \frac{\lambda}{4\pi CD r} + 1 \\ \Rightarrow K'(r) &= \frac{\lambda r}{CD} + 4\pi r^2 \\ \Leftrightarrow K(r) &= \frac{\lambda r^2}{2CD} + \frac{4}{3}\pi r^3. \end{aligned} \quad (45)$$

⁵⁴³ **Time-dependent solution** In the absence of advection ($U = 0, \gamma = 0$), we need to compute a time-dependent solution. From eq. 37,

$$\begin{aligned} K(\rho) &= \int_0^\rho r \frac{\lambda}{DC} \left\{ 1 - \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right) \right\} + 4\pi r^2 dr \\ &= \frac{\lambda}{CD} \left(\frac{\rho^2}{2} - \int_0^\rho r \times \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right) dr \right) + \frac{4}{3}\pi\rho^3. \end{aligned} \quad (46)$$

⁵⁴⁵ We first compute the primitive for $\int_0^\rho r \times \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right) dr$. We define $u = \frac{r}{\sqrt{8Dt}}$, $dr = \sqrt{8Dt} du$,
⁵⁴⁶ then

$$\int_0^\rho r \times \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right) dr = 8Dt \int_0^{\rho/\sqrt{8Dt}} u \times \operatorname{erf}(u) du. \quad (47)$$

⁵⁴⁷ We can integrate by parts, with $f = \operatorname{erf}(u)$ and $g' = u$, and obtain

$$8Dt \int_0^{\rho/\sqrt{8Dt}} u \times \operatorname{erf}(u) du = 8Dt \left(\frac{\rho^2}{2} \frac{1}{8Dt} \operatorname{erf}\left(\frac{\rho}{\sqrt{8Dt}}\right) - \frac{1}{\sqrt{\pi}} \int_0^{\rho/\sqrt{8Dt}} u^2 e^{-u^2} du \right). \quad (48)$$

548 We integrate by parts again, this time with $f = u$ and $g' = ue^{-u^2}$, which leads to

$$\int u^2 e^{-u^2} du = -\frac{ue^{-u^2}}{2} + \frac{1}{2} \int e^{-u^2} du = -\frac{ue^{-u^2}}{2} + \frac{\sqrt{\pi} \operatorname{erf}(u)}{4}. \quad (49)$$

549 If we use Eq. 49 in Eq. 48,

$$\begin{aligned} 8Dt \int_0^{\rho/\sqrt{8Dt}} u \times \operatorname{erf}(u) du &= 8Dt \left(\frac{\rho^2}{2} \frac{1}{8Dt} \operatorname{erf}\left(\frac{\rho}{\sqrt{8Dt}}\right) - \frac{\operatorname{erf}\left(\frac{\rho}{\sqrt{8Dt}}\right)}{4} + \frac{1}{2\sqrt{\pi}} \frac{\rho}{\sqrt{8Dt}} e^{-\frac{\rho^2}{8Dt}} \right) \\ \Leftrightarrow \int_0^{\rho} r \times \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right) dr &= \frac{1}{2} \operatorname{erf}\left(\frac{\rho}{\sqrt{8Dt}}\right) (\rho^2 - 4Dt) + \frac{\sqrt{2Dt}}{\sqrt{\pi}} \rho e^{-\frac{\rho^2}{8Dt}}. \end{aligned} \quad (50)$$

550 We can now compute $K(\rho)$:

$$K(\rho) = \frac{\lambda}{CD} \left(\frac{\rho^2}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{\rho}{\sqrt{8Dt}}\right) (\rho^2 - 4Dt) - \frac{\sqrt{2Dt}\rho}{\sqrt{\pi}} e^{-\frac{\rho^2}{8Dt}} \right) + \frac{4}{3}\pi\rho^3. \quad (51)$$

551 References

- 552 Adler, P.B., Smull, D., Beard, K.H., Choi, R.T., Furniss, T., Kulmatiski, A., Meiners, J.M.,
553 Tredennick, A.T. & Veblen, K.E. (2018). Competition and coexistence in plant communities:
554 intraspecific competition is stronger than interspecific competition. *Ecology Letters*,
555 21, 1319–1329.
- 556 Arnott, R.N., Cherif, M., Bryant, L.D. & Wain, D.J. (2021). Artificially generated turbulence:
557 a review of phycological nanocosm, microcosm, and mesocosm experiments. *Hydrobiologia*,
558 848, 961–991.
- 559 Arrieta, J., Jeanneret, R., Roig, P. & Tuval, I. (2020). On the fate of sinking diatoms: the
560 transport of active buoyancy-regulating cells in the ocean. *Philosophical Transactions of
561 the Royal Society A*, 378, 20190529.
- 562 Bainbridge, R. (1957). The size, shape and density of marine phytoplankton concentrations.
563 *Biological Reviews*, 32, 91–115.
- 564 Barabás, G., Michalska-Smith, M.J. & Allesina, S. (2017). Self-regulation and the stability
565 of large ecological networks. *Nature Ecology & Evolution*, 1, 1870–1875.
- 566 Barraquand, F. & Murrell, D.J. (2012). Evolutionarily stable consumer home range size in
567 relation to resource demography and consumer spatial organization. *Theoretical Ecology*,
568 5, 567–589.

- 569 Barton, A.D., Ward, B.A., Williams, R.G. & Follows, M.J. (2014). The impact of fine-scale
570 turbulence on phytoplankton community structure. *Limnology and Oceanography: Fluids*
571 and Environments
- 572 Bec, B., Collos, Y., Vaquer, A., Mouillot, D. & Souchu, P. (2008). Growth rate peaks at inter-
573 mediate cell size in marine photosynthetic picoeukaryotes. *Limnology and Oceanography*,
574 53, 863–867.
- 575 Benczik, I.J., Károlyi, G., Scheuring, I. & Tél, T. (2006). Coexistence of inertial competitors
576 in chaotic flows. *Chaos*, 16, 043110.
- 577 Birch, D.A. & Young, W.R. (2006). A master equation for a spatial population model with
578 pair interactions. *Theoretical Population Biology*, 70, 26–42.
- 579 Bissinger, J.E., Montagnes, D.J.S., Harples, J. & Atkinson, D. (2008). Predicting marine
580 phytoplankton maximum growth rates from temperature: improving on the Eppley curve
581 using quantile regression. *Limnology and Oceanography*, 53, 487–493.
- 582 Bolker, B.M. & Pacala, S.W. (1999). Spatial moment equations for plant competition: under-
583 standing spatial strategies and the advantages of short dispersal. *The American Naturalist*,
584 153, 575–602.
- 585 Borgnino, M., Arrieta, J., Boffetta, G., De Lillo, F. & Tuval, I. (2019). Turbulence induces
586 clustering and segregation of non-motile, buoyancy-regulating phytoplankton. *Journal of*
587 *the Royal Society Interface*, 16, 20190324.
- 588 Bouderbala, I., El Saadi, N., Bah, A. & Auger, P. (2018). A 3D individual-based model
589 to study effects of chemotaxis, competition and diffusion on the motile-phytoplankton
590 aggregation. *Acta Biotheoretica*, 66, 257–278.
- 591 Breier, R.E., Lalescu, C.C., Waas, D., Wilczek, M. & Mazza, M.G. (2018). Emergence of
592 phytoplankton patchiness at small scales in mild turbulence. *Proceedings of the National*
593 *Academy of Sciences*, 115, 12112–12117.
- 594 Chesson, P. (2018). Updates on mechanisms of maintenance of species diversity. *Journal of*
595 *Ecology*, 106, 1773–1794.
- 596 Detto, M. & Muller-Landau, H.C. (2016). Stabilization of species coexistence in spatial mod-
597 els through the aggregation-segregation effect generated by local dispersal and nonspecific
598 local interactions. *Theoretical Population Biology*, 112, 97–108.

- 599 Doubell, M.J., Seuront, L., Seymour, J.R., Patten, N.L. & Mitchell, J.G. (2006). High-
600 resolution fluorometer for mapping microscale phytoplankton distributions. *Applied and*
601 *Environmental Microbiology*, 72, 4475–4478.
- 602 Dusenberry, D. (2009). *Living at the microscale*. Harvard University Press.
- 603 Edwards, K.F. (2019). Mixotrophy in nanoflagellates across environmental gradients in the
604 ocean. *Proceedings of the National Academy of Sciences*, p. 201814860.
- 605 Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte
606 bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4.
- 607 Estrada, M., Alcaraz, M. & Marrasé, C. (1987). Effects of turbulence on the composition
608 of phytoplankton assemblages in marine microcosms. *Marine Ecology Progress Series*, 38,
609 267–281.
- 610 Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski, P. (1998). Primary production
611 of the biosphere: integrating terrestrial and oceanic components. *Science*, 281, 237–240.
- 612 Font-Muñoz, J.S., Jeanneret, R., Arrieta, J., Anglès, S., Jordi, A., Tuval, I. & Basterretxea,
613 G. (2019). Collective sinking promotes selective cell pairing in planktonic pennate diatoms.
614 *Proceedings of the National Academy of Sciences*, 116, 15997–16002.
- 615 Font-Muñoz, J.S., Jordi, A., Tuval, I., Arrieta, J., Anglès, S. & Basterretxea, G. (2017).
616 Advection by ocean currents modifies phytoplankton size structure. *Journal of the Royal*
617 *Society Interface*, 14, 20170046.
- 618 Haegeman, B. & Rapaport, A. (2008). How flocculation can explain coexistence in the
619 chemostat. *Journal of Biological Dynamics*, 2, 1–13.
- 620 Hellweger, F.L. & Bucci, V. (2009). A bunch of tiny individuals – individual-based modeling
621 for microbes. *Ecological Modelling*, 220, 8–22.
- 622 Huisman, J. & Weissing, F.J. (1999). Biodiversity of plankton by species oscillations and
623 chaos. *Nature*, 402, 407–410.
- 624 Hutchinson, G.E. (1961). The paradox of the plankton. *The American Naturalist*, 95, 137–
625 145.
- 626 Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical analysis and modelling*
627 *of spatial point patterns*. vol. 70. John Wiley & Sons.

- 628 Jumars, P.A., Deming, J., Hill, P., Karp-Boss, L., Yager, P. & Dade, W. (1993). Physical
629 constraints on marine osmotrophy in an optimal foraging context. *Marine Microbial Food*
630 *Webs*, 7, 121–159.
- 631 Karp-Boss, L., Boss, E. & Jumars, P.A. (1996). Nutrient fluxes to planktonic osmotrophs in
632 the presence of fluid motion. *Oceanography and Marine Biology: An Annual Review*, 34,
633 71–107.
- 634 Kiørboe, T., Andersen, K.P. & Dam, H.G. (1990). Coagulation efficiency and aggregate
635 formation in marine phytoplankton. *Marine Biology*, 107, 235–245.
- 636 Kraichnan, R.H. (1974). Convection of a passive scalar by a quasi-uniform random straining
637 field. *Journal of Fluid Mechanics*, 64, 737–762.
- 638 Law, R., Murrell, D.J. & Dieckmann, U. (2003). Population growth in space and time: spatial
639 logistic equations. *Ecology*, 84, 252–262.
- 640 Leonard, C.L., Bidigare, R.R., Seki, M.P. & Polovina, J.J. (2001). Interannual mesoscale
641 physical and biological variability in the North Pacific Central Gyre. *Progress in Oceanog-*
642 *rphy*, 49, 227–244.
- 643 Levine, J.M. & HilleRisLambers, J. (2009). The importance of niches for the maintenance
644 of species diversity. *Nature*, 461, 254–257.
- 645 Li, L. & Chesson, P. (2016). The effects of dynamical rates on species coexistence in a
646 variable environment: the paradox of the plankton revisited. *The American Naturalist*,
647 188, E46–E58.
- 648 MacArthur, R. & Levins, R. (1964). Competition, habitat selection and character displace-
649 ment in a patchy environment. *Proceedings of the National Academy of Sciences*, 51,
650 1207–1210.
- 651 Martin, A.P. (2003). Phytoplankton patchiness: the role of lateral stirring and mixing.
652 *Progress in Oceanography*, 57, 125–174.
- 653 Murrell, D. (2005). Local spatial structure and predator-prey dynamics: counterintuitive
654 effects of prey enrichment. *The American Naturalist*, 166, 354–367.
- 655 Ngan, K. & Vanneste, J. (2011). Scalar decay in a three-dimensional chaotic flow. *Physical*
656 *Review E*, 83, 056306.

- 657 North, A. & Ovaskainen, O. (2007). Interactions between dispersal, competition, and land-
658 scape heterogeneity. *Oikos*, 116, 1106–1119.
- 659 Peters, F. & Marrasé, C. (2000). Effects of turbulence on plankton: an overview of experi-
660 mental evidence and some theoretical considerations. *Marine Ecology Progress Series*, 205,
661 291–306.
- 662 Picoche, C. & Barraquand, F. (2019). How self-regulation, the storage effect, and their
663 interaction contribute to coexistence in stochastic and seasonal environments. *Theor Ecol*,
664 12, 489–500.
- 665 Picoche, C. & Barraquand, F. (2020). Strong self-regulation and widespread facilitative
666 interactions in phytoplankton communities. *Journal of Ecology*, 108, 2232–2242.
- 667 Picoche, C., Young, W.R. & Barraquand, F. (2022). [Re] Reproductive pair correlations and
668 the clustering of organisms. *ReScience C*, 8.
- 669 Pierrehumbert, R.T. (1994). Tracer microstructure in the large-eddy dominated regime.
670 *Chaos, Solitons & Fractals*, 4, 1091–1110.
- 671 Plank, M.J. & Law, R. (2015). Spatial point processes and moment dynamics in the life
672 sciences: a parsimonious derivation and some extensions. *Bulletin of Mathematical Biology*,
673 77, 586–613.
- 674 Prairie, J.C., Sutherland, K.R., Nickols, K.J. & Kaltenberg, A.M. (2012). Biophysical inter-
675 actions in the plankton: a cross-scale review. *Limnology and Oceanography: Fluids and*
676 *Environments*, 2, 121–145.
- 677 Record, N.R., Pershing, A.J. & Maps, F. (2014). The paradox of the “paradox of the plank-
678 ton”. *ICES Journal of Marine Science*, 71, 236–240.
- 679 REPHY (2017). *REPHY dataset - French Observation and Monitoring program*
680 *for Phytoplankton and Hydrology in coastal waters. 1987-2016 Metropolitan data.*
681 <https://www.seanoe.org/data/00361/47248/>.
- 682 Reynolds, C.S. (2006). *The ecology of phytoplankton*. Cambridge University Press.
- 683 Schippers, P., Verschoor, A.M., Vos, M. & Mooij, W.M. (2001). Does “supersaturated coex-
684 istence” resolve the “paradox of the plankton”? *Ecology Letters*, 4, 404–407.
- 685 Seymour, J.R., Amin, S.A., Raina, J.B. & Stocker, R. (2017). Zooming in on the phycosphere:
686 the ecological interface for phytoplankton–bacteria relationships. *Nature Microbiology*, 2,
687 17065.

- 688 Stocker, R. (2012). Marine microbes see a sea of gradients. *Science*, 338, 628–633.
- 689 Widdicombe, C. & Harbour, D. (2021). Phytoplankton taxonomic abundance and biomass
690 time-series at Plymouth Station L4 in the Western English Channel, 1992–2020.
- 691 Wiegand, T., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N. & Huth, A. (2007). How individual
692 species structure diversity in tropical forests. *Proceedings of the National Academy of
693 Sciences*, 104, 19029–19033.
- 694 Young, W.R., Roberts, A.J. & Stuhne, G. (2001). Reproductive pair correlations and the
695 clustering of organisms. *Nature*, 412, 328–331.