Supplementary Materials for
CompuCell3D Model of Cell Migration Reproduces Chemotaxis

Pedro C. Dal-Castel[1], Gilberto L. Thomas[1],
Gabriel C. Perrone[1], and Rita M.C. de Almeida[1,2,3]*

[1] *Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil*

[2] *Instituto Nacional de Ciência e Tecnologia: Sistemas Complexos, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil and*

[3] *Programa de Pós-Graduação em Bioinformática, Instituto Metrópole Digital, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil*

(Dated: March 24, 2023)

* pdalcastel@gmail.com; glt@if.ufrgs.br; gabriel.perrone@if.ufrgs.br; rita@if.ufrgs.br

Os parâmetros das simulações (tabelas) não estão referenciadas no texto principal! E devem ser referenciadas com Table S1, S2...

I. INPUT PARAMETERS AND RESULTS

Table S1 presents the results of MSD fits and drift speed for cell radius $R = 10$, zero chemotactic response $\mu = 0$, protrusion strength $\lambda_{F-actin} = \{125, 150, 175\}$ and Lamel volume fraction $\phi_F = \{0.05, 0.10, 0.20\}$. The MSD fit yields slow diffusion coefficient $D[R^2/\text{MCS}]$, the S parameter and the persistence time $P[\text{MCS}]$. S parameter is the time interval that separates the first diffusive regime and the ballistic-like regime in natural units. It also relates to the fast diffusion coefficient $A = \frac{2DS}{1-S}$. The drift speed V_d is obtained from the average of velocity component parallel to polarization for smallest time interval, as we defined in Eq. (11).

Table S2 presents the results of MSD fits, drift speed, terminal velocity and chemotactic efficiency for cell radius $R = 10$, saturated chemotactic response $\mu = 10^6$, protrusion strength $\lambda_{F-actin} = \{125, 150, 175\}$ and Lamel volume fraction $\phi_F = \{0.05, 0.10, 0.20\}$. We find $B[R^2/\text{MCS}^2]$, the parameter of the last ballistic regime associated to chemotactic response. The terminal velocity $V_T[R/\text{MCS}] = \sqrt{B}$ is the cell speed for large time intervals during chemotactic response, and chemotactic efficiency $\varepsilon = \frac{V_T}{V_d}$ is the ratio between terminal velocity and drift speed.

Table S1. Results for simulations with parameter inputs: $R = 10$, $\mu = 0$, $\lambda_{F-actin} = \{125, 150, 175\}$, and $\phi_F = \{0.05, 0.10, 0.20\}$.

$R = 10 \quad \mu = 0$					
ϕ_f	$\lambda_{F-actin}$	$D[R^2/\text{MCS}]$	$P[\text{MCS}]$	S	$V_d[R/\text{MCS}]$
0.05	125	4.18E-04	1.30E+03	8.39E-01	5.32E-04
0.05	150	1.88E-03	3.52E+03	9.21E-01	7.31E-04
0.05	175	3.16E-03	7.20E+03	9.54E-01	7.37E-04
0.1	125	1.84E-04	1.05E+03	8.08E-01	3.81E-04
0.1	150	1.86E-03	2.90E+03	9.06E-01	8.06E-04
0.1	175	5.40E-03	7.18E+03	9.54E-01	9.65E-04
0.2	125	–	–	–	–
0.2	150	7.18E-04	2.49E+03	8.92E-01	4.86E-04
0.2	175	3.17E-03	4.60E+03	9.29E-01	8.64E-04

Table S2. Results for simulations with parameter inputs: $R = 10$, $\mu = 10^6$, $\lambda_{F-actin} = \{125, 150, 175\}$, and $\phi_F = \{0.05, 0.10, 0.20\}$.

$R = 10 \quad \mu = 10^6$								
ϕ_f	$\lambda_{F-actin}$	$D[R^2/\text{MCS}]$	$P[\text{MCS}]$	S	$B[R^2/\text{MCS}^2]$	$V_d[R/\text{MCS}]$	$V_T[R/\text{MCS}]$	ε
0.05	125	3.20E-04	1.20E+03	5.90E-02	1.07E-07	5.75E-04	3.28E-04	5.70E-01
0.05	150	1.45E-03	3.15E+03	1.33E-02	1.00E-07	7.61E-04	3.17E-04	4.16E-01
0.05	175	4.79E-03	9.27E+03	3.97E-03	3.97E-08	7.47E-04	1.99E-04	2.67E-01
0.1	125	1.85E-04	9.95E+02	9.78E-02	2.32E-08	4.49E-04	1.52E-04	3.39E-01
0.1	150	1.73E-03	2.76E+03	1.11E-02	4.61E-08	7.83E-04	2.15E-04	2.74E-01
0.1	175	6.80E-03	7.15E+03	2.80E-03	2.88E-08	9.72E-04	1.70E-04	1.74E-01
0.2	125	–	–	–	–	–	–	–
0.2	150	7.54E-04	2.58E+03	2.51E-02	9.69E-09	5.14E-04	9.84E-05	1.91E-01
0.2	175	3.19E-03	3.69E+03	5.87E-03	1.49E-08	9.13E-04	1.22E-04	1.34E-01

II. 4-REGIMES MSD FIT STEP-BY-STEP WITH EXAMPLE

Fitting an MSD curve can be difficult due to the multiple regimes for very different time scales. Some parameters depend on each other and it can be hard to separate them. A direct implementation of a gradient descent algorithm may not give the best result, even if you weight the data appropriately.

Here we present a step-by-step fitting procedure of Eq. S1

$$\langle |\Delta\vec{r}|^2 \rangle = 2D(\Delta t - P(1 - e^{-\Delta t/P})) + A\Delta t + B\Delta t^2 \quad . \tag{S1}$$

over a simulated MSD data that presents all four regimes of movement (short time diffusive, intermediate ballistic, long term diffusive and long term ballistic), as shown in Fig. S1.

This method yields good fits, although it is more complicated than simply hitting "the fit button" on a data analysis software, but it is faster and less arbitrary than the brute force iterative fitting in each curve region. You only need access to basic data manipulation and basic gradient descent fitting routine.

We start with a MSD vs Δt in log-log scale, as shown in Fig. S1. This curve is the particular case $R = 10$, $\phi_F = 0.10$, $\lambda_{F-actin} = 150$. We obtained the given curve by averaging the MSDs of 10 different cell trajectories.

The first parameter to extract is B. It is the easiest since the squared term is dominant for large Δt. We take advantage of this fact: by differentiating our MSD function given in

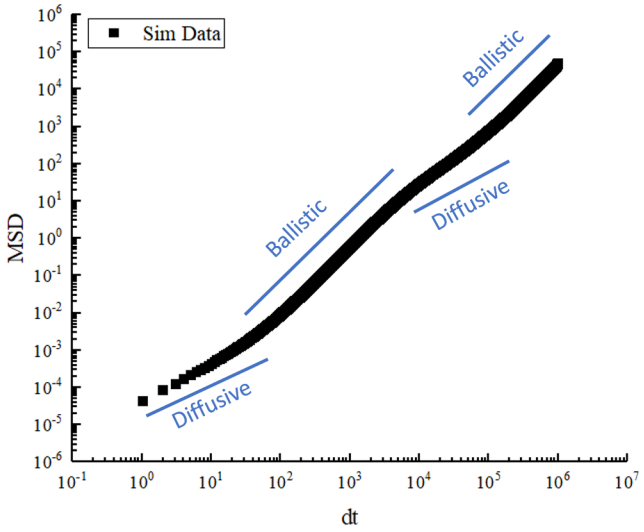

Figure S1. Plot of MSD vs Δt in log-log scale.

Eq. S1 and taking the limit for large Δt:

$$MSD_t = \frac{d}{d(\Delta t)}\langle|\Delta\vec{r}|^2\rangle = 2D(1 - e^{-\Delta t/P}) + A + 2B\Delta t \quad ,$$
$$\lim_{\Delta t\to\infty} MSD_t = 2B\Delta t \quad , \tag{S2}$$

Only parameter B remains. Then, we take the numerical derivative of our MSD simulated data and plot it linearly, as we show in Fig. S2. From the graph, we identify the straight-line region, whose slope must be equal to $2B$. A linear fit of this region, shown by a red line, determines B.

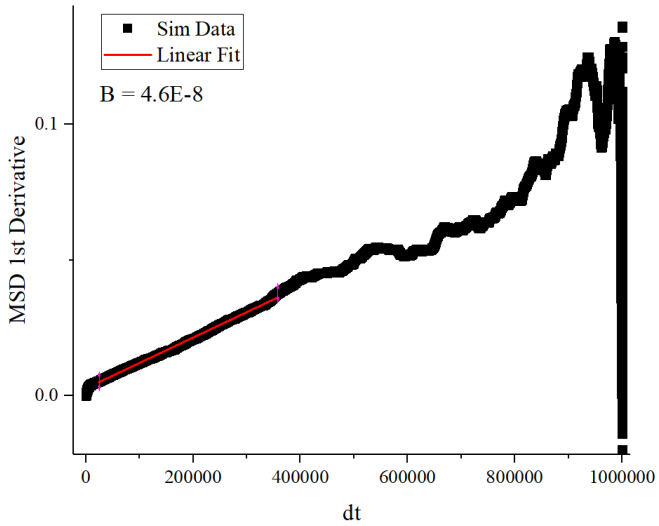

Figure S2. Plot and linear fit of the linear region in the 1st differentiated MSD curve (linear scale).

After we found the parameter B, we replot MSD $- B\Delta t^2$ in Fig. S3. We can see that this subtraction eliminates the long-term ballistic regime, and the plotted curve resembles the MSD in the absence of the chemotactic field.

The procedure for fitting the remaining parameters have already been published by Fortuna and collaborators in 2020 [1]. For the sake of completeness, we will reproduce it here.

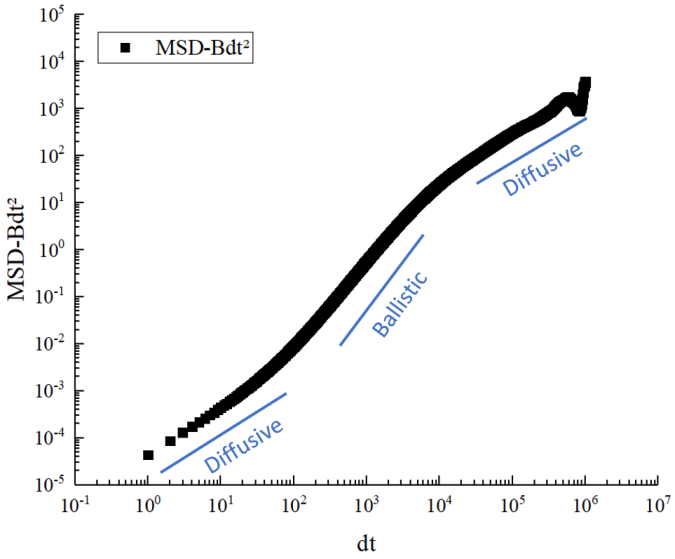

Figure S2. MSD without the ballistic term $B\Delta t^2$.

The reduced MSD equation now is

$$\langle|\Delta\vec{r}|^2\rangle = 2D(\Delta t - P(1 - e^{-\Delta t/P})) + A\Delta t \quad , \tag{S3}$$

and we have to find the parameters A, D, and P.

We differentiate Eq. S3 twice in respect to Δt to get rid of the linear term $A\Delta t$, resulting

$$\mathrm{MSD}_{tt} = \frac{d^2}{d(\Delta t)^2}\langle|\Delta\vec{r}|^2\rangle = \frac{2D}{P}e^{-\Delta t/P} . \tag{S4}$$

So, the $\log(\mathrm{MSD}_{tt}) \times \Delta t$ plot should show a curve region with linear behavior. The linear fitting of this region will determine parameters D and P.

Fig S4 presents the time-differentiated simulated data with black squares. The blue line is the smoothed curve for the data. It helps us to identify the linear region. In this region, we perform a nonlinear curve fit using Eq. S4 to find parameters D and P.

Subtracting from original MSD both Fürth and chemotactic terms, we are left with the

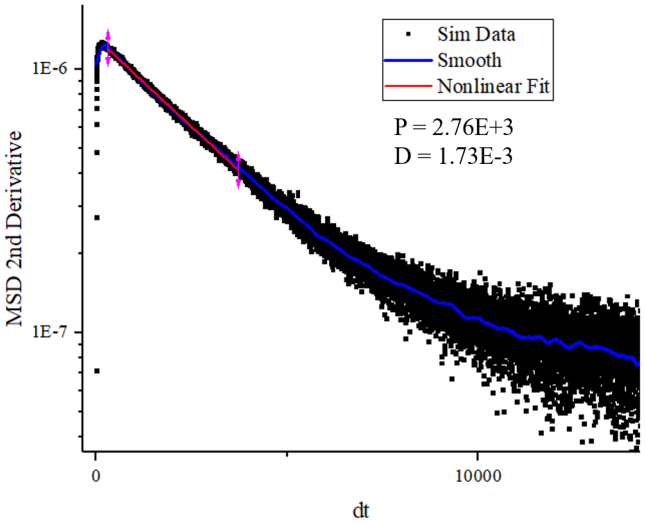

Figure S4. Plot and nonlinear fit (using Eq. S4) of MSD 2nd derivative to find parameters D and P. Graph scales are log-linear.

linear term $A\Delta t$, *i.e.*,

$$A\Delta t = \text{MSD} - 2D(\Delta t - P(1 - e^{-\Delta t/P})) - 2B\Delta t^2 \ . \tag{S5}$$

This subtraction in the simulated MSD data gives us the curve presented with black squares in Fig. S5.

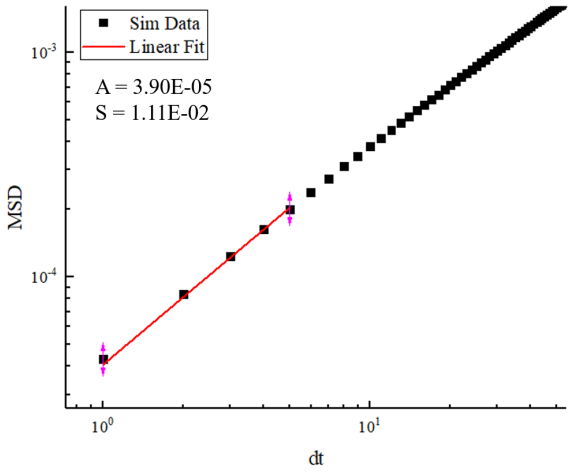

Figure S5. MSD curve only with the remaining $A\Delta t$ term. The relation between S and A is $A = \frac{2DS}{1-S}$.

Fixing the intercept coefficient at zero, we perform a linear fit for the small Δt region. The resulting fitting is presented by the red line in Fig. S5. Note that for this fitting, it does not matter if the graph is put in log-log or linear-linear scales (it looks like a straight line

in both choices), provided that you perform a linear fit in the original coordinate system.

With all 4 parameters, we can plot both MSD simulated data points and the final fit curve. This is showed in Fig. S5. As you may have noticed, not all steps are user independent, different decisions will impact the final result, for example: 1) what exactly is the range of data points I have to set in each fit step? 2) do I ignore this noisy part? 3) do I ignore this curvy part in my supposedly linear region? 4) in which order do I fit the parameters? 5) do I repeat the process for some of the parameters? The first consideration is that no model will perfectly fit the data. To be practical, we don't care about obtaining precise parameter values as much as we care about the relationship between the fitting parameters and cell behavior. For example: cells with clear and stable polarization should present higher P relative to those with less stable polarization; cells that walk longer distances over same periods of time should present a higher D relative to those that do not move as much; cells that present higher fluctuations in its structure should present a higher A; and cells that walk faster towards a chemical field source should present a higher B. Precise estimates are not the goal and may even not be possible to achieve in any way. The real requirements are: the procedure is 1) standardized, 2) reproducible, and 3) the results must correlate with observed cell behavior.

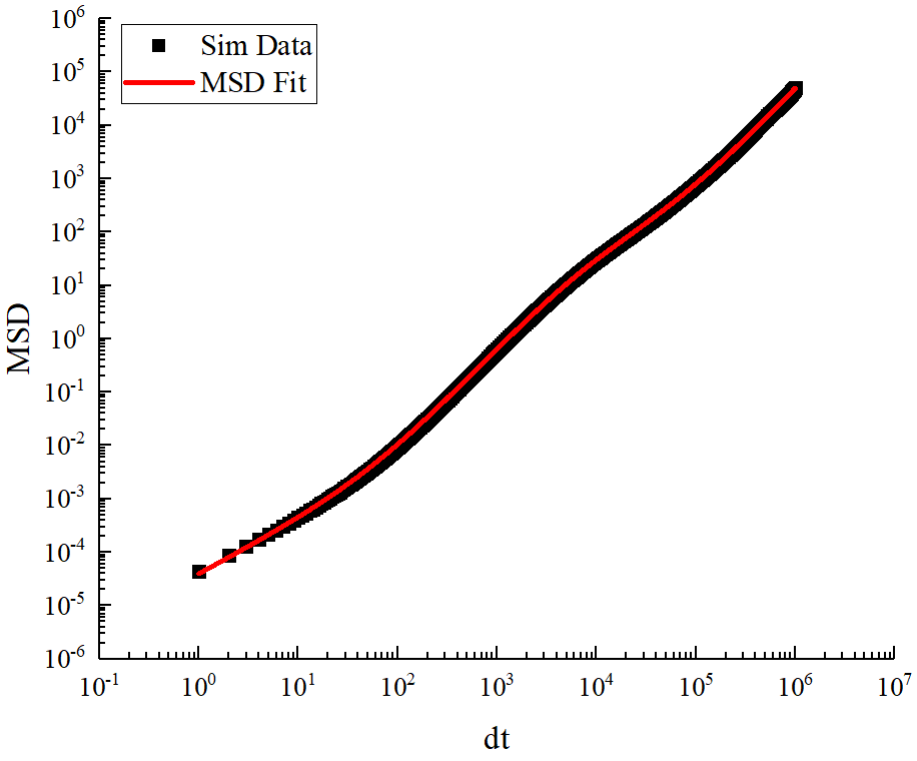

Figure S5. Final MSD fit in log-log scale.

BIBLIOGRAPHY

[1] Ismael Fortuna, Gabriel C Perrone, S Krug, Monique, Eduarda Susin, Julio M Belmonte, Gilberto L Thomas, Glazier James A, and Rita M C de Almeida. Compucell3d simulations reproduce mesenchymal cell migration on flat substrates. *Biophysical J.*, 118(11):2801–2815, 2020.