Supplementary Information

Measurement and mitigation of free convection in microfluidic gradient generators

Yang Gu[1], Varun Hegde[1], and Kyle J. M. Bishop*[1]

[1]Department of Chemical Engineering, Columbia University

Contents

*Address correspondence to kyle.bishop@columbia.edu

Data Analysis

Velocity Measurements

Likelihood Function. Particle tracking provides data on the positions (x_{nk}, y_{nk}) of the n^{th} particle at the k^{th} time step. For simplicity, we focus our analysis on particle motions in the x direction; motions in the y direction can be treated in similar fashion. During imaging, we capture particle motions within a region of finite thickness centered on the focal plane z_0. Consequently, there is a distribution of particle velocities within the imaging region as characterized by a mean velocity μ_U and a standard deviation σ_U. In particular, we make the convenient assumption that the particle velocities are normally distributed as

$$p(U \mid \mu_U, \sigma_U) = \frac{1}{\sqrt{2\pi}\sigma_U} \exp\left(-\frac{(U-\mu_U)^2}{2\sigma_U^2} \right), \tag{S1}$$

where $p(\ )$ denotes the probability density. During the k^{th} time interval ($\tau = t_k - t_{k-1}$), the displacement of a single particle ($\Delta_k = x_k - x_{k-1}$) is also normally distributed with mean $\mu_\Delta = U\tau$ and standard deviation $\sigma_\Delta = \sqrt{2D_p\tau}$, where D_p is the particle diffusivity,

$$p(\Delta_k \mid \mu_\Delta, \sigma_\Delta) = \frac{1}{\sqrt{2\pi}\sigma_\Delta} \exp\left(-\frac{(\Delta_k-\mu_\Delta)^2}{2\sigma_\Delta^2} \right). \tag{S2}$$

The probability of observing a series of K independent displacements $\{\Delta_k\}$ is given by the product

$$p(\{\Delta_k\} \mid \mu_\Delta, \sigma_\Delta) = \prod_{k=1}^{K} p(\Delta_k \mid \mu_\Delta, \sigma_\Delta). \tag{S3}$$

This distribution is conditioned on the particle velocity (i.e., on $\mu_\Delta = U\tau$), which is itself uncertain. A more useful distribution is obtained by marginalizing over the particle velocity as

$$p(\{\Delta_k\} \mid \mu_U, \sigma_U, \sigma_\Delta) = \int p(\{\Delta_k\} \mid \mu_\Delta, \sigma_\Delta) p(U \mid \mu_U, \sigma_U) \mathrm{d}U. \tag{S4}$$

Note that this distribution is conditioned on the three constant parameters μ_U, σ_U, and σ_Δ. Carrying out the integration, we obtain

$$p(\{\Delta_k\} \mid \mu_U, \sigma_U, \sigma_\Delta) = \frac{1}{(2\pi\sigma_U^2)^{1/2}} \frac{1}{(2\pi\sigma_\Delta^2)^{K/2}} \sqrt{\frac{\pi}{A}} \exp\left(\frac{B^2}{4A} - C \right), \tag{S5}$$

where the terms A, B, and C are given by

$$A = \frac{1}{2\sigma_U^2} + \frac{K\tau^2}{2\sigma_\Delta^2}, \quad B = -\frac{\mu_U}{\sigma_U^2} - \frac{\tau}{\sigma_\Delta^2} \sum_{k=1}^{K} \Delta_k, \quad C = \frac{\mu_U^2}{2\sigma_U^2} + \frac{1}{2\sigma_\Delta^2} \sum_{k=1}^{K} \Delta_k^2. \tag{S6}$$

Finally, the likelihood of obtaining the observed data for all N independent particles is

$$p(\{\Delta_{nk}\} \mid \mu_U, \sigma_U, \sigma_\Delta) = \prod_{n=1}^{N} \frac{1}{(2\pi\sigma_U^2)^{1/2}} \frac{1}{(2\pi\sigma_\Delta^2)^{K_n/2}} \sqrt{\frac{\pi}{A_n}} \exp\left(\frac{B_n^2}{4A_n} - C_n \right). \tag{S7}$$

Posterior Distribution. We are now prepared to estimate the unknown parameters μ_U, σ_U, and σ_Δ from the data $\{\Delta_{nk}\}$ by application of Bayes theorem as

$$p(\mu_U, \sigma_U, \sigma_\Delta \mid \{\Delta_{nk}\}) \propto p(\{\Delta_{nk}\} \mid \mu_U, \sigma_U, \sigma_\Delta) p(\mu_U, \sigma_U, \sigma_\Delta). \tag{S8}$$

We assume a uniform prior distribution over the relevant region of parameter space,

$$p(\mu_U, \sigma_U, \sigma_\Delta) = \text{constant}. \tag{S9}$$

The logarithm of the posterior distribution, $p(\mu_U, \sigma_U, \sigma_\Delta \mid \{\Delta_{nk}\})$, can then be written as

$$\mathcal{L}(\mu_U, \sigma_U, \sigma_\Delta) = \text{constant} - \sum_{n=1}^{N} \left[\ln \sigma_U + K_n \ln \sigma_\Delta + \frac{1}{2} \ln A_n - \frac{B_n^2}{4A_n} + C_n \right]. \tag{S10}$$

The most probable parameter values are those which maximize this function. We used numerical optimization methods to identify these most probable parameters. Additionally, the reliabilities of our estimates were evaluated by approximating the posterior distribution as a multivariate normal distribution and examining the corresponding covariance matrix. Specifically, we expanded the logarithm of the posterior in a second order Taylor expansion about the most probable value; the desired covariance matrix is equal to the inverse of the computed Hessian matrix.

Special Case ($K_n = K$). When the number of displacements is the same for each particle, it is possible to derive simple expressions for the most probable parameters values and their reliabilities, which provide useful insight into the sources of uncertainty our estimates. These expressions are conveniently expressed in terms of the sample means and standard deviations of the respective particle displacements,

$$\mu_n = \frac{1}{K} \sum_k \Delta_{nk}, \quad \sigma_n^2 = \frac{1}{K-1} \sum_k (\Delta_{nk} - \mu_n)^2. \tag{S11}$$

Using these statistics, the most probable parameter values are

$$\mu_{U,o} = \frac{1}{\tau N} \sum_n \mu_n, \tag{S12}$$

$$\sigma_{\Delta,o}^2 = \frac{1}{N} \sum_n \sigma_n^2 \tag{S13}$$

$$\sigma_{U,o}^2 = \frac{1}{N} \sum_n \left(\frac{\mu_n}{\tau} - \mu_{U,o} \right)^2 - \frac{\sigma_{\Delta,o}^2}{\tau^2}. \tag{S14}$$

As noted above, the reliability of these estimates can be obtain by approximating the posterior distribution as a multivariate normal distribution about these most probable values. The respective variances of the marginal distributions for each of the three parameters are

$$\delta^2(\mu_U) = \frac{\sigma_{U,o}^2}{N} \left[1 + \frac{1}{K} \left(\frac{\sigma_{\Delta,o}}{\tau \sigma_{U,o}} \right)^2 \right], \tag{S15}$$

$$\delta^2(\sigma_\Delta) = \frac{\sigma_{\Delta,o}^2}{2N(K-1)}, \tag{S16}$$

$$\delta^2(\sigma_U) = \frac{\sigma_{U,o}^2}{N} \left[\frac{1}{2} + \frac{1}{K} \left(\frac{\sigma_{\Delta,o}}{\tau \sigma_{U,o}} \right)^2 + \frac{1}{2K(K-1)} \left(\frac{\sigma_{\Delta,o}}{\tau \sigma_{U,o}} \right)^4 \right]. \tag{S17}$$

These uncertainty estimates are useful in choosing the duration of the particle tracks $K\tau$ used in experiment. Specifically, they suggest that tracks longer that $K\tau \sim D_p/\sigma_U^2$ are increasingly ineffective in reducing the uncertainty in the parameters μ_U and σ_U. For example, for 500 nm tracer particles ($D_p = 10^{-12}$ m^2/s in water) and velocity variations of $\sigma_U = 0.1$ μm/s, this condition suggests that tracks should be ca. 100 s long; slower flows require longer tracks.

Fitting the Velocity Profiles

We used Bayesian data analysis[1] combined with Markov Chain Monte Carlo (MCMC) sampling to evaluate the proposed model and infer the model parameters. We focused our analysis on the mean particle velocity within the focal region μ_U (see previous section) as a function of height z within the channel. This average velocity includes contributions from particles moving at different heights within the channel (not just those in the focal plane). We model this average velocity by the convolution

$$\tilde{v}_x(z) = \int_0^H g(z' - z) v_x(z') \mathrm{d}z', \tag{S18}$$

where the function $g(z - z')$ describes the contribution of particles at height z' to the measured velocity at height z. In general, this function is expected to depend on the distribution of tracer particles in the channel and on the details of particle imaging and tracking. Here, we assume that particles withing a characteristic distance w contribute to the image as

$$g(z' - z) = \frac{C(z)}{\sqrt{2\pi w^2}} \exp\left(-\frac{(z - z')^2}{2w^2}\right). \tag{S19}$$

Here, $C(z)$ is a normalization factor for the truncated normal distribution on the domain $0 < z < H$,

$$C(\tilde{z}) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{1 - \bar{z}}{\sqrt{2}\bar{w}}\right) + \mathrm{erf}\left(\frac{\bar{z}}{\sqrt{2}\bar{w}}\right)\right], \tag{S20}$$

where $\bar{z} = z/H$ and $\bar{w} = w/H$. For buoyancy driven flow, the profile for the average velocity $\tilde{v}_x(z)$ is derived using equation (6) in the main text to obtain

$$\frac{\tilde{v}_x(\tilde{z})}{U_B} = \frac{4\tilde{w}}{C(\tilde{z})}\sqrt{\frac{2}{\pi}}\left[(2(1-\bar{z})^2 - (1-\bar{z}) + 4\bar{w}^2)e^{-\frac{\bar{z}^2}{2\bar{w}^2}} - (2z^2 - z + 4\bar{w})e^{-\frac{(1-\bar{z})^2}{2\bar{w}^2}}\right] + 8(2\bar{z} - 1)(3\bar{w}^2 - (1-\bar{z})\bar{z}), \tag{S21}$$

Using this model, we inferred the magnitude of the velocity U_B and the width of the averaging function w assuming Gaussian noise in our measurements of the average velocity profiles. MCMC sampling was performed in Python using the pymc3 package.[2]

Hydrodynamic Model

Numerical Solution (Pe = 0)

Scaling lengths by the channel height H, velocity by $V = \beta g G H^3/\nu$, and pressure by $\mu U/H$, the hydrodynamic equations in the main text can be written in dimensionless form as

$$0 = \frac{\partial v_x}{\partial x} + \frac{\partial v_x}{\partial x}, \tag{S22}$$

$$0 = -\frac{\partial p}{\partial x} + \frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_x}{\partial z^2}, \tag{S23}$$

$$0 = -\frac{\partial p}{\partial z} + \frac{\partial^2 v_z}{\partial x^2} + \frac{\partial^2 v_z}{\partial z^2} + x. \tag{S24}$$

The no-slip conditions at the boundaries of the channel imply that $v_x = v_z = 0$ for $x = 0, W$ and $z = 0, 1$. This incompressible 2D flow can be expressed in terms of the stream function ψ where $v_x = \partial\psi/\partial z$ and $v_z = -\partial\psi/\partial x$. Subtracting the z-derivative of equation (S23) from the x-derivative of equation (S24), we obtain

$$\frac{\partial^4 \psi}{\partial x^4} + 2\frac{\partial^4 \psi}{\partial x^2 \partial z^2} + \frac{\partial^4 \psi}{\partial z^4} = \nabla^4 \psi = 1, \tag{S25}$$

subject to the following boundary conditions

$$\partial_x \psi(0,0) = \partial_x \psi(0,1) = \partial_x \psi(W,0) = \partial_x \psi(W,1) = 0, \tag{S26}$$

$$\partial_z \psi(0,0) = \partial_z \psi(0,1) = \partial_z \psi(W,0) = \partial_z \psi(W,1) = 0. \tag{S27}$$

For a given channel width W, this above equation was solved numerically using the finite element method in COMSOL. In the numerical implementation, it was convenient to decompose the 4th order equation above into three 2nd order equations as

$$P = \frac{\partial^2 \psi}{\partial x^2}, \quad Q = \frac{\partial^2 \psi}{\partial z^2}, \quad \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial z^2} + \frac{\partial^2 Q}{\partial z^2} = 1. \tag{S28}$$

Analytical Solution for $H/W \ll 1$

In the limit of short channel, flow is approximately unidirectional in the x-direction; equations (S23) and (S24) can be simplified as

$$0 = -\frac{\partial p}{\partial x} + \cancel{\frac{\partial^2 v_x}{\partial x^2}} + \frac{\partial^2 v_x}{\partial z^2}, \tag{S29}$$

$$0 = -\frac{\partial p}{\partial z} + \cancel{\frac{\partial^2 v_z}{\partial x^2}} + \cancel{\frac{\partial^2 v_z}{\partial z^2}} + x. \tag{S30}$$

Integrating equation (S36), we find that the pressure is of the form

$$p = xz + f(x), \tag{S31}$$

where $f(x)$ is an unknown function of x. In order that equation (S35) not depend on position x along the channel, this function must be linear $f(x) = Ax$ where A is a constant to be determined. Substituting the pressure into equation (S35) and integrating, we find

$$v_x(z) = \frac{z^3}{6} + A\frac{z^2}{2} + Bz + C. \tag{S32}$$

The constants A, B, and C are uniquely determined by the boundary conditions, $v_x(0) = v_x(1) = 0$, and by the condition of no net flow along the channel

$$\int_0^1 v_x(z)dz = 0. \tag{S33}$$

The resulting solution is

$$v_x(z) = \frac{z^3}{6} - \frac{z^2}{4} + \frac{z}{12}. \tag{S34}$$

The maximum (dimensionless) velocity is $v_{\max} = (3-\sqrt{3})/6 = 0.0080$ and occurs at $z = (3\pm\sqrt{3})/6$.

Analytical Solution for $H/W \gg 1$

In the limit of tall channels, flow is approximately unidirectional in the z-direction; equations (S23) and (S24) can be simplified as

$$0 = -\frac{\partial p}{\partial x} + \cancel{\frac{\partial^2 v_x}{\partial x^2}} + \cancel{\frac{\partial^2 v_x}{\partial z^2}}, \tag{S35}$$

$$0 = -\frac{\partial p}{\partial z} + \frac{\partial^2 v_z}{\partial x^2} + \cancel{\frac{\partial^2 v_z}{\partial z^2}} + x. \tag{S36}$$

Assuming the pressure has the form $p = Az$, we can integrate equation (S36) to obtain the solution

$$v_z(x) = -W^3 \left[\frac{1}{6}\left(\frac{x}{W}\right)^3 - \frac{1}{4}\left(\frac{x}{W}\right)^2 + \frac{1}{12}\left(\frac{x}{W}\right)\right]. \tag{S37}$$

The peak velocity scales as $v_{\max} \sim W^3$ as illustrated in Figure 3b.

Simulations of Bacterial Chemotaxis

To illustrate the effects of gradient driven flows on the chemotaxis of bacterial populations, we investigate a simple model that describes (1) the (active) diffusive motion of individual bacteria, (2) their systematic migration in the solute gradient, (3) their convective transport with the fluid, and (4) their sedimentation under gravity. In particular, we consider a suspension of bacteria moving in a microfluidic glucose gradient similar to that shown in Figure 2. In the model, the bacteria concentration $c_b(x, z, t)$ is governed by the following conservation equation

$$\frac{\partial c_b}{\partial t} + (\boldsymbol{v} + \boldsymbol{v}_c + \boldsymbol{v}_s) \cdot \nabla c_b = D_b \nabla^2 c_b, \tag{S38}$$

where D_b is the bacteria diffusivity, $\boldsymbol{v} = v_x(z)\boldsymbol{e}_x$ is the fluid velocity, $\boldsymbol{v}_c = U_c \boldsymbol{e}_x$ is the (constant) chemotactic velocity, and $\boldsymbol{v}_s = U_s \boldsymbol{e}_z$ is the (constant) sedimentation velocity. The fluid velocity is described by hydrodynamic model described above. Initially, the bacteria concentration is assumed to be uniform throughout the gradient channel

$$c_b(x, z, 0) = c_b^o. \tag{S39}$$

We assume that there is no flux of bacterial through the boundaries of the channel. The transient concentration profile is computed numerically in COMSOL.

Figure (S1) shows the result of two simulations showing chemotaxis without (left) and with (right) buoyancy-driven convective flows. In the simulations, the bacteria swim to the left towards higher solute concentrations (positive chemotaxis) and sediment into the bottom half of the channel. Buoyancy driven flows along the bottom half of the channel drive the bacteria in the opposite direction. As a result, the apparent chemotactic velocity is significantly reduced by the presence of convective flows. At the same time, buoyancy driven flows do not change the qualitative behavior of the bacteria and might therefore be overlooked in quantifying bacteria chemotaxis. This numerical example corresponds to experimentally relevant estimates of bacterial motion such as a chemotactic velocity $U_c = 0.5$ μm/s, bacterial diffusivity $D_b = 25$ μm^2/s, and sedimentation velocity $U_s = 0.5$ μm/s.[3] The channel height $H = 100$ μm and buoyancy driven flow velocity $U_B = 1$ μm/s are similar to those observed in our experiments.

References

(1) Sivia, D. S.; Skilling, J., *Data Analysis: A Bayesian Tutorial*, 2nd; Oxford University Press: New York, 2006.

(2) Salvatier, J.; Wiecki, T.; Fonnesbeck, C. *PeerJ Comput. Sci.* **2016**, *2*, e55.

(3) Son, K.; Menolascina, F.; Stocker, R. *Proc. Natl. Acad. Sci.* **2016**, *113*, 8624–8629.

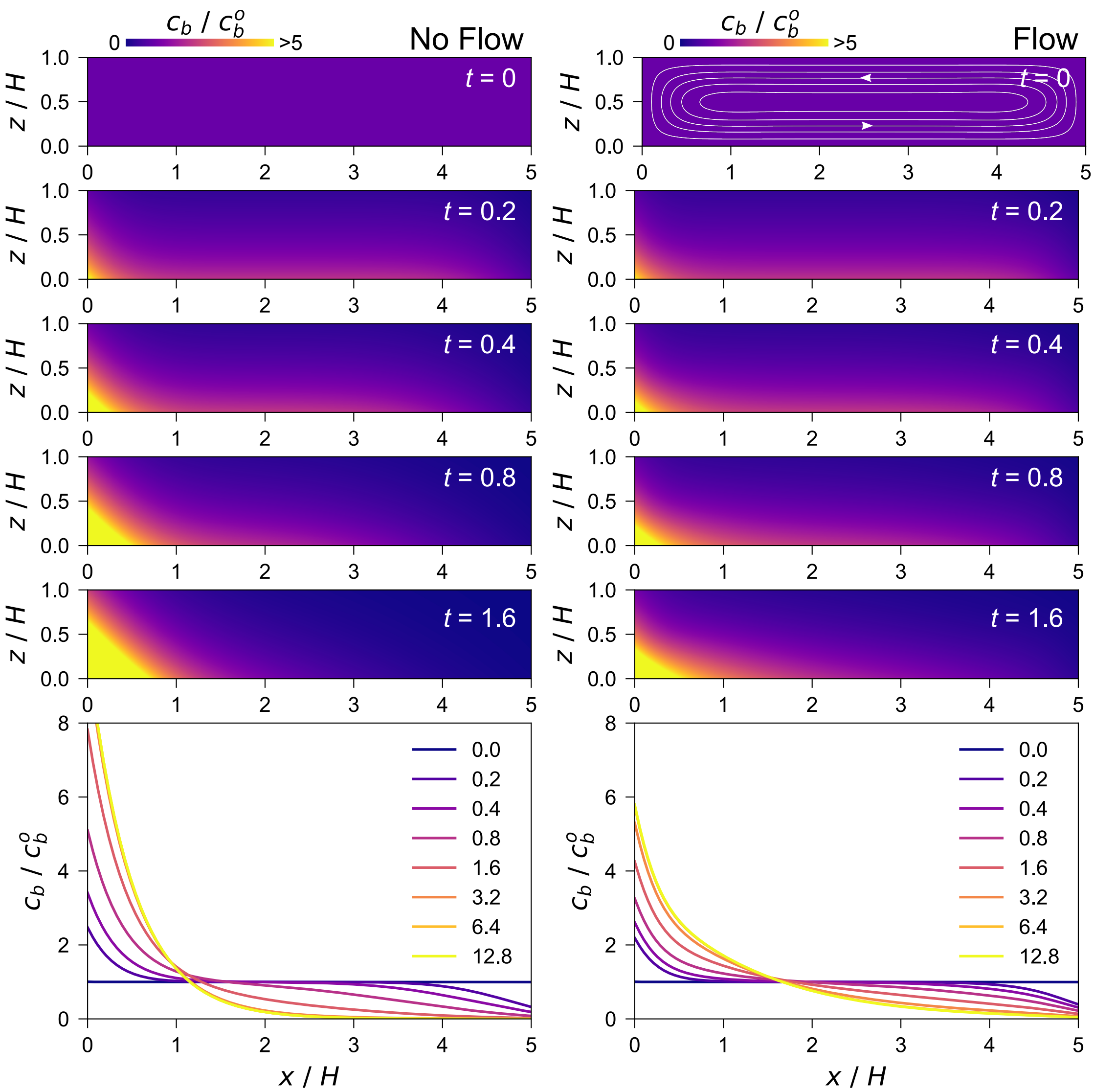

Figure S1: Transient concentration profile $c_b(x, z, t)$ for a population of chemotactic bacteria moving in a concentration gradient without (left) and with (right) buoyancy driven flows. In the simulations, length is scaled by the channel height H, time by H^2/D_b, velocity by D_b/H, and concentration by c_b^o. The chemotactic velocity is $U_c = -2(D_b/H)$, the sedimentation velocity $U_s = -2(D_b/H)$, and the channel width $W = 5H$. On the right, the characteristic magnitude of the buoyancy driven flows is $U_B = 4(D_b/H)$. The plots below show the concentration profiles integrated over the height of the channel—that is, $\int_0^H c_b(x, z, t)\mathrm{d}z$.