Postfragmentation density function for bacterial aggregates in laminar flow

Erin Byrne
Department of Applied Mathematics, University of Colorado at Boulder, Boulder 80309, USA

Steve Dzul and Michael Solomon
Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA

John Younger
Department of Emergency Medicine, University of Michigan, Ann Arbor, Michigan 48109, USA

David M. Bortz*
Department of Applied Mathematics, University of Colorado at Boulder, Boulder 80309, USA
(Received 1 February 2010; published 15 April 2011; publisher error corrected 28 April 2011)

The postfragmentation probability density of daughter flocs is one of the least well-understood aspects of modeling flocculation. We use three-dimensional positional data of *Klebsiella pneumoniae* bacterial flocs in suspension and the knowledge of hydrodynamic properties of a laminar flow field to construct a probability density function of floc volumes after a fragmentation event. We provide computational results which predict that the primary fragmentation mechanism for large flocs is erosion. The postfragmentation probability density function has a strong dependence on the size of the original floc and indicates that most fragmentation events result in clumps of one to three bacteria eroding from the original floc. We also provide numerical evidence that exhaustive fragmentation yields a limiting density inconsistent with the log-normal density predicted in the literature, most likely due to the heterogeneous nature of *K. pneumoniae* flocs. To support our conclusions, artificial flocs were generated and display similar postfragmentation density and exhaustive fragmentation.

DOI: 10.1103/PhysRevE.83.041911 PACS number(s): 87.18.Fx

I. INTRODUCTION

In modeling the flocculation dynamics of bacterial and other biological and cellular aggregates in flowing aqueous environments, three important phenomena arise: growth, aggregation, and fragmentation [1]. Though much attention has been given to the dynamics of floc aggregation since von Smoluchowski published his coagulation equations in 1916 [2], the research into floc breakup has been less successful as the complex, stochastic nature of shear-force-induced fragmentation makes the mathematics of floc breakup significantly more challenging than that of aggregation [3]. Simulations of the breakup process have been completed (e.g., Refs. [4–6]), though the focus of these studies were on aggregates held together through short-range van der Waals attractions. The extension and rupture of polymer networks differs fundamentally from the breakup of flocs whose structure is due to these short-range attractions [7]. Direct observation of the fragmentation process is challenging, though it has been visualized for flocs in extensional [8–11] swirling and contraction [12] flows. Various strategies have also been proposed to infer floc strength indirectly including microscale cantilever beams [13], atomic force microscopy [14], and traditional particle sizer techniques [15,16] (augmented by modeling).

This paper considers a representative example of a bacterial species, *Klebsiella pneumoniae*, that commonly grows in an aggregate community. *Klebsiella pneumoniae* is an aggregate and biofilm forming organism common to fresh water environments and is also frequently encountered as a human pathogen [17,18]. In human disease it is an important cause of biofilm-based intravenous catheter infections and therefore a common cause of disseminated bloodstream infections. *Klebsiella* is a typical member of the family *Enterobacteriaceae*, to which belong a number of problematic species for mankind, including *Escherichia coli* (causing food borne diarrheal and urinary tract illnesses), *Salmonella* and *Shigella* (causing water borne diarrheal illness), *Yersinia* (plague), and *Erwinia* (an important group of agricultural pathogens). Relevant to the current discussion, a recent study (by the Younger and Solomon labs) carried out high-resolution confocal microscopic imaging analysis of *K. pneumoniae* aggregates (data on file). It is from these results that the current analysis proceeds.

A vital component of the fragmentation kernel for mathematically modeling *K. pneumoniae* flocculation dynamics is the probability density of floc sizes after a fragmentation event [the postfragmentation probability density, $\Gamma(x;y)$], which defines the density function for producing a daughter floc of size x from a floc of size y. Many different functional forms of the postfragmentation probability density are used in the literature and we direct the interested reader to the extensive reviews in Refs. [19–22]. The same forms commonly appear for use in modeling both animate (e.g., algae [23,24], bacteria [1], yeast [25]) and inanimate (polymer [19,22]) material.

The two most prominently used functional forms are binary and log-normal densities. The binary density typically assumes that any fragmentation event results in two equally sized daughter flocs, i.e., the mean field approximation to a β fragmentation process. The log-normal is derived from assuming that an unlimited cascade of fragmentation events will result in many small and a few large daughter flocs.

*dmbortz@colorado.edu

Kolmogoroff [26] showed the log-normal density to be the limiting density for repeated, exhaustive fragmentation, and Lu and Spielman [27] and Pandya and Spielman [28] adopted this result for repeated floc erosion. One assumption which yields a log-normal density is homogeneity and does not hold for our bacterial flocs. The predictions from our computations suggest that neither are binary fragmentation symmetric nor does exhaustive fragmentation yield a log-normal size density.

The focus of this paper is to construct a postfragmentation particle size probability density function for bacterial aggregates based on high-resolution structural measurements of *K. pneumoniae*. We focus on particles in linear, laminar flow primarily for analytic simplicity, but note the potential relevance of this particular flow regime to one of *Klebsiella*'s most troublesome features, the tendency to invade the bloodstream of critically ill patients.

Section II describes the method used to construct the postfragmentation density function from the three-dimensional (3D) positional data. Section III reports the results of the simulated floc breakup and discusses the implications of the findings. Section III also supports our results by comparing them with those of artificially generated flocs. Section IV summarizes the conclusions of the work.

II. MODEL AND METHODS

Our goal is to develop an analytic strategy for identifying breakage locations, calculating the associated stress due the hydrodynamic forces at work, and then using that information to predict if a fragmentation event will occur. This process must consider both the geometry and the material properties of the floc as well as the hydrodynamic environment. Aggregates are discretized into a countable number of bacteria that are assumed to be nonbreakable. *Klebsiella pneumoniae* is a cylindroid, roughly twice as long as wide, and we assume that this shape does not significantly impact our results. While we have data on the 3D locations and orientation of the bacteria that constitute a given floc, the density and extent of the *extracellular polysaccharides* (EPS), or *biofilm*, are unknown. We assume an EPS layer of uniform thickness encapsulating each bacterium, and no additional strength or structure is attributed to any overlapping regions. Last, note that in our analysis we do not consider the possibility of an EPS fragment (without bacteria) separating from the mother floc.

A. Description of the data

The 3D structural data of 39 aggregates were collected and imaged in the Younger and Solomon labs [29]. In brief, bacterial aggregates were obtained by growing organisms in defined media under gentle hydrodynamic conditions. These structures were fluorescently stained and then imaged with confocal microscopy to produce 3D image volumes containing each bacterium in an aggregate. Image analysis software [30] identified the Cartesian coordinates of the center of mass of each cell. Note that we use all 39 aggregates in the estimation of $\Gamma(x; y)$.

B. Breakage location identification

Our analysis of possible fragmentation locations is based entirely on the relative positions of the bacteria that constitute a given floc. The discretized nature of the bacteria allows us to characterize the aggregate as an undirected graph where the centers of mass of the individual bacteria are the nodes and the pairwise connections between the centers of mass are the edges. Each edge is given a weight corresponding to the distance between the two bacteria it connects. A *spanning tree* of a graph is any subset of edges with no closed loops where the edges connect all the nodes in the graph. The *minimum spanning tree* (MST) is the spanning tree that minimizes the sum of the weights of the constituent edges and is unique for graphs with distinct edge weights. Further information on the MST and its properties can be found in Ref. [31].

By finding the MST using distances as the edge weights, we identify the smallest total distance needed to connect all the bacteria in the floc. We assume the longest edge in the MST is the most likely to fragment first because it indicates the largest distance between neighboring bacteria. Conceptually, if the uniform EPS layer were reduced in thickness until a separation occurred, the first separation would be located on the longest edge of the MST.

We use Kruskal's algorithm [32] for finding the MST and consider each edge in the MST as a candidate for fragmentation. The fragmentation plane is taken to be the perpendicular bisector of the edge and separates the floc into two daughter flocs. A sample floc from our data set, its bacterial centers of mass, and the corresponding MST are depicted in Fig. 1(a), while an alternative view with a candidate fragmentation plane $P_\perp$ is depicted in Fig. 1(b).

C. Simplification of the model

Flocs come in all shapes and sizes, and this spatial variation makes calculating the exact hydrodynamic forces at work on the effective surface of the floc challenging. We therefore use an approximation which facilitates calculation while still capturing the general shape of the floc.

Optical observations [8] find that simple shear flow leads to rotation of the floc, and its motion can be understood by the behavior of a solid ellipsoid. We approximate the aggregate by an ellipsoid tumbling in a linear flow field as is assumed in Refs. [8,9,12] and validated in Refs. [15,16]. To find a hydrodynamically equivalent ellipsoid, we use principle components analysis (PCA) on the spatial data for centers of mass of the bacteria in the floc. PCA yields an orthogonal linear transformation that maps data to a new coordinate system such that each principle component sequentially accounts for as much of the variability as possible. On spatial data, PCA simply performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance. Further information on PCA can be found in Ref. [33].

The floc is rotated to align the longest semiaxis (with length a_1) with the x axis, the next longest (with length a_2) with the y axis, and the shortest (with length a_3) with the z axis. For visualization purposes, we depict in Fig. 2 the rotated centers of mass and semiaxes for an equivalent ellipsoid. We chose the magnitude of the semiaxes as twice the standard deviation along the principal components, though note that the calculation of the hydrodynamic forces depends only upon the ratio of the semiaxis lengths and not upon the lengths themselves (see Sec. II D). The surface of the aligned

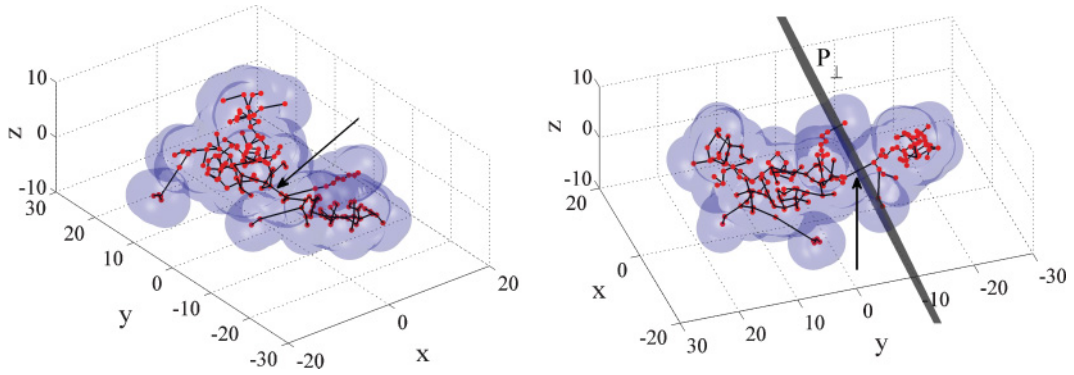

FIG. 1. (Color online) (a) Sample floc, enveloped by the effective surface of a uniform EPS layer (3D reconstruction). The red circles are the centers of mass of the bacteria in the floc and the blue lines are the edges that constitute the minimum spanning tree (MST) of the centers of mass. (b) Alternate view of the same floc shown with a sample fragmentation plane $P_\perp$ determined by the edge in the MST identified by an arrow. Axes are in μm.

ellipsoid is denoted D and described by D: $\frac{x^2}{a_1^2}+\frac{y^2}{a_2^2}+\frac{z^2}{a_3^2}=1$, where $a_1 \geqslant a_2 \geqslant a_3$.

Our focus is primarily on medium to large flocs, but for completeness small flocs are included in our analysis. Our results take the size of the mother floc into consideration, and we note the impact of finite-size effects in small flocs. When any semiaxis length is found to be less that 1 μm, the value is replaced with 1 μm to allow for the presence of the EPS. When only two bacteria make up a floc, using PCA to find the equivalent ellipsoid fails due to the lack of any variance of the data in the second and third principle directions. In this case, we assume the length of the first semiaxis to be half the distance between the two bacteria plus an additional 1-μm buffer for the EPS, and we assume the length of the second and third semiaxes to be 1 μm.

D. Hydrodynamic forces

The forces on the surface of an ellipsoid in a linear flow field are well understood [34,35] and our criteria for floc breakage is determined by comparing floc strength to hydrodynamic forces. While the energy-based approach advocated in Refs. [36,37] is more general, we consider only a laminar flow environment and not one with turbulent mixing. Thus the energy dissipation rates and the potential energy associated with the rupture need not be formulated for our case.

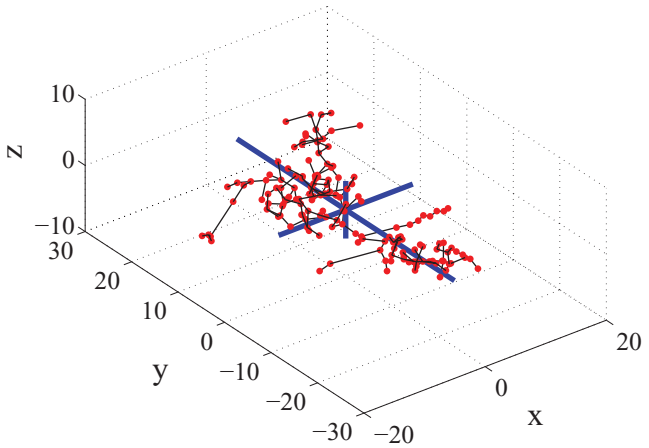

FIG. 2. (Color online) Sample floc depicting the centers of mass of the bacteria in the floc, the corresponding MST, and the semiaxes of a hydrodynamically equivalent ellipsoid as determined using PCA. Axes are in μm.

We consider the floc immersed in a linear shear flow field, represented by a 3×3 matrix, $\mathbf{G}$, where the (1,2) element is γ and all other elements are zero. The parameter γ is the constant fluid shear rate along the y axis.

The center of mass of the floc is assumed to be moving with the flow, so the only hydrodynamic forces acting on the ellipsoid are due to rotation (Jeffery [38] provides a formulation for this force). We assume the fluid's pressure on the ellipsoid is negligible. The rotational force on the surface of the ellipsoid is given by

$$\mathbf{f_R} = 4\mu \left(\mathbf{A} - \sum_{k=1}^{3} \chi_k A_{k,k} \mathbf{I} \right) \mathbf{n}, \tag{1}$$

where μ is the viscosity of the fluid, $\mathbf{n}$ is the normal vector to the surface of the ellipsoid, χ_k are elliptic integrals involving the semiaxis lengths, and $\mathbf{A}$ is a matrix linearly dependent on $\mathbf{G}$ and the semiaxis lengths. The formulations for χ_k and $\mathbf{A}$ can be found in the equations in Appendix A, adapted from Blaser [35]. Last, note that f_R depends only on the ratios of the semiaxis lengths and not on the semiaxis lengths themselves.

If we constrain the net hydrodynamic torque to be zero, we can use the *Jeffery equations* to find the angular velocity, ω, of the ellipsoid. We assume the semiaxis a_3 and the vorticity vector of the shear flow are constant, reducing the Jeffery equations to a single equation for the angular velocity around the z axis:

$$\omega_3 = \Omega_{1,2} + \frac{a_1^2 - a_2^2}{a_1^2 + a_2^2} E_{1,2}, \tag{2}$$

where $E_{i,k}$ and $\Omega_{i,k}$ are the (i,k)-th entries of the rate-of-strain tensor $\mathbf{E} = \frac{1}{2}(\mathbf{G} + \mathbf{G}^{\mathrm{T}})$ and the vorticity tensor $\Omega = \frac{1}{2}(\mathbf{G} - \mathbf{G}^{\mathrm{T}})$.

The formulation for the force is given in the body-fixed frame and the velocity gradient given in the laboratory frame

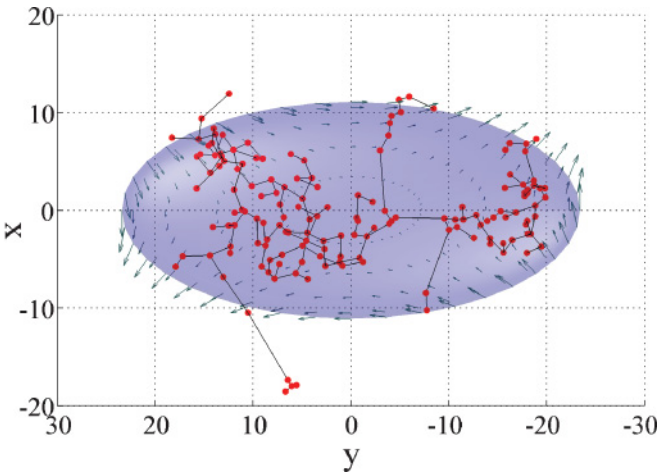

FIG. 3. (Color online) Sample floc depicting the centers of mass of the bacteria in the floc, the corresponding MST, a hydrodynamically equivalent ellipsoid, and sample ellipsoid surface force vectors (in Newtons) per unit area (longer vectors denote larger magnitude forces). Note that this figure is a projection of the data onto the xy plane. The force vectors have no z component because of the ellipsoid's alignment with the flow and the largest forces occur around the ellipsoid's equator (the $a_1 \times a_2$ plane). Axes are in μm.

$\mathbf{G}'$ must be rotated using the transformation $\mathbf{G} = \mathbf{R}\,\mathbf{G}'\,\mathbf{R}^{\mathrm{T}}$. Since we assume only rotation about the z axis, the rotation matrix is given by

$$\mathbf{R} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

where ϕ is the counterclockwise angle of rotation measured from the positive $x$axis. The motion of the ellipsoid is then related to the angular velocity by the equation $\phi' = \omega_3$.

The forces calculated through this method are vector quantities of stress (force per unit area) on the surface of the ellipsoid and are visualized in Fig. 3. Note that this formulation accounts for all shear-flow-induced forces on the surface of the ellipsoid.

E. Construction of a postfragmentation probability density function

This section describes the procedure we use to construct the postfragmentation probability density function, $\Gamma(x; y)$, from the data. After finding the MST for one floc (Sec. II B), we find a hydrodynamically equivalent ellipsoid (Sec. II C) and calculate the stress on the surface of the ellipsoid (Sec. II D).

The longest edge in the MST is the largest distance between neighboring clusters of bacteria, and thus the perpendicular bisector ($P_\perp$) of that edge is a candidate separation plane for the floc. This plane divides the ellipsoid into two pieces, and integration of the forces normal to $P_\perp$ over each piece of the ellipsoid yields the total tensile force:

$$\mathbf{F}_\perp = \left| \iint_{D>P_\perp} \mathbf{f}_\mathrm{R} \cdot \mathbf{n}_\perp \, dA \right| + \left| \iint_{D<P_\perp} \mathbf{f}_\mathbf{R} \cdot \mathbf{n}_\perp \, dA \right|, \tag{4}$$

where $\mathbf{n}_\perp$ is the unit normal to $P_\perp$. Because the ellipsoid is in a laminar shear field, it rotates and we calculate the force over one period of motion, using the maximum value in the calculation of the maximum tensile stress, $\sigma_{\max}$.

When the *rupture stress*, σ_r, of the EPS is exceeded, we say a fragmentation event has occurred. If $\sigma_{\max}$ for the longest edge in the MST does not exceed σ_r, we recursively consider $\sigma_{\max}$ for the next longest edge in the MST as a candidate for separation until σ_r is exceeded. When a fragmentation event occurs, the edge under consideration is removed from the MST, creating two disconnected graphs which represent the two daughter flocs formed by the fragmentation. Each daughter floc is then considered to be a new floc for potential fragmentation and analyzed as such. If σ_r is never exceeded along any edge in the MST, we say the floc does not fragment.

The algorithm is outlined by the following pseudocode to facilitate understanding. A more detailed version of the algorithm can be found in Appendix B.

```
(1)  INPUT = Flocs = List of flocs
(2)  WHILE i ⩽ (No. of flocs in Flocs)
(3)     Find MST for Flocs(i)
(4)     Find equivalent ellipsoid for Flocs(i)
(5)     FOR = k = 1 : (No. of edges in MST)
(6)        Find P⊥ for k^th longest edge (separation plane)
(7)        σmax = max tensile stress for separation plane
(8)        IF = σmax ⩾ σr
(9)           → Fragmentation Event
(10)          Add daughters to Flocs
(11)          GOTO 14
(12)       END IF
(13)    END FOR
(14)    i = i + 1
(15) END WHILE
```

We repeat this process for each floc and construct a normalized histogram of the daughter floc sizes relative to the mother floc. This relative size is determined from the fraction of the number of bacteria in the mother floc that are in each daughter floc. To allow for future laboratory comparison, we apply the algorithm to selected combinations of fluid shear rates and feasible biofilm rupture stresses. We allow the biofilm rupture stress (σ_r) to vary between 1 and 150 Pa and the fluid shear rate (γ) to vary between 10 and 500 s^{-1}.

III. RESULTS AND DISCUSSION

We investigated the postfragmentation density functions, $\Gamma(x; y)$, for various combinations of fluid shear rate and biofilm rupture stress. In all cases we explored, the resultant density functions predominantly exhibit fragmentations resulting in one daughter floc significantly larger than the other. This suggests that the primary type of fracture occurring is the erosion of small clumps of bacteria off the mother floc, where we define *erosion* to be any fracture where one daughter floc is at least threefold the size of the other daughter floc. In the following sections we investigate the dependence of the postfragmentation density function on the size of the

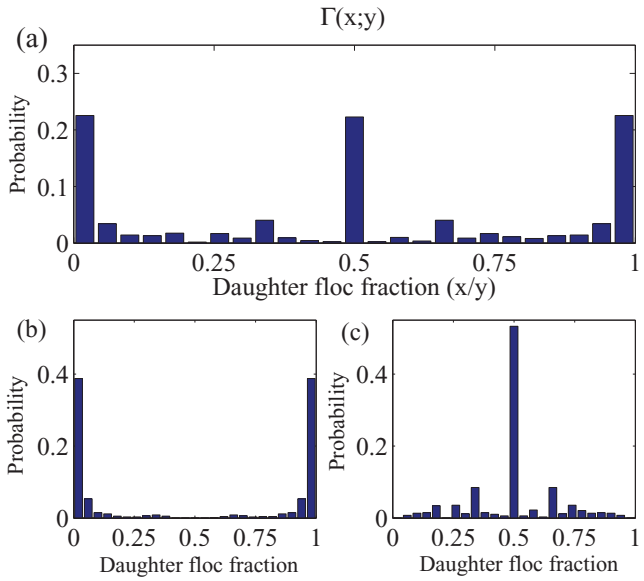

FIG. 4. Sample postfragmentation density function $\Gamma(x; y)$ ($\sigma_r = 10$ Pa, $\gamma = 100$ s^{-1}) considering (a) all mother flocs, (b) only mother flocs with 15 or more bacteria, and (c) only mother flocs with fewer than 15 bacteria. The x axis is the size of a daughter floc as a fraction of the size of the mother floc.

mother floc and different combinations of fluid shear rate and biofilm rupture stress. We also examine the density of particle sizes after *exhaustive fragmentation*, which we define to be the result of repeated fragmentation until the hydrodynamic forces are insufficient to fragment any remaining flocs. Finally, we support our results by comparing then with artificially generated flocs with statistical properties similar to flocs acquired in the laboratory.

A. Dependence on mother floc size

The density $\Gamma(x; y)$ depends heavily on the size of the mother floc since it is dramatically different when 1 or 2 bacteria fragment from a mother floc with 100 bacteria than from a mother floc with 5 bacteria. According to our definition, erosion is only possible for flocs with 4 or more bacteria and flocs consisting of 2 or 3 bacteria are unable to have any fragmentation event result in erosion. An example of a postfragmentation density function using mother flocs of all sizes ($\sigma_r = 10$ Pa, $\gamma = 100$ s^{-1}) is shown in Fig. 4(a). The same density is also shown having been separated into mother flocs with 15 or greater bacteria [Fig. 4(b)] and those with fewer than 15 bacteria [Fig. 4(c)].

If we consider only mother flocs with 15 or more bacteria, over 90% of the fractures can be classified as erosion. Even for mother flocs as small as 5 bacteria, over 80% of the fractures are due to erosion. The resulting shape is distinctive, with the highest probability occurring on either end of the domain and the lowest occurring in the center [similar to Fig. 4(b)]. For smaller mother flocs, the density is more uniform, with spikes at the even fractions where it is possible to produce daughter flocs. For example, there are spikes at $\frac{1}{3}$ and $\frac{2}{3}$, reflecting mother flocs with 3 bacteria fracturing into daughter flocs with 1 and 2 bacteria, respectively.

The inclusion of mother floc size as a parameter in the density is common practice in the literature. For example, a uniform density function for producing a daughter floc of size x from a mother floc of size y can be represented by $\Gamma(x; y) = y^{-1}$ for $x \in [0, y]$ (see Ref. [19] for a summary of forms).

We examine the heavy dependence on the mother floc size more closely by considering it as a parameter in the postfragmentation density function. While the number of mother flocs of a specific size varies in our data, there is enough resolution to see the general shape of the density evolve with the size of the mother floc [Fig. 5(a)]. For small mother flocs, we observe large spikes at fractions of the mother floc size. As the mother floc size increases, the largest of these spikes moves toward the edges of the domain, similar to several of the experimental results in Ref. [13] for densely packed flocs. This migration is indicative of the trend for daughter flocs with a small number of bacteria to erode from the mother floc [Figs. 5(b)–5(d)].

B. Dependence on fluid shear rate and biofilm rupture stress

As mentioned, there are slight differences in the density functions for different combinations of fluid shear rate and biofilm rupture stress. Figure 6 depicts sample combinations of γ and σ_r. A combination of high rupture stress and low fluid shear rate (depicted in Fig. 6 upper right, $\sigma_r = 15$ Pa, $\gamma = 50$ s^{-1}) reduces the number of small mother flocs and thus the corresponding Γ more closely resembles the density for only larger mother flocs seen in Fig. 4(b). Fewer fragmentations occur because the hydrodynamic forces acting on the floc are weaker. Conversely, a combination of low rupture stress and high fluid shear rate (depicted in Fig. 6 lower left, $\sigma_r = 5$ Pa, $\gamma = 150$ s^{-1}) increases the number of fragmentations of small flocs, resulting in a density resembling Γ for only small mother flocs [Fig. 4(c)].

When considering a constant rupture stress value and allowing the shear rate to increase, we see similar fragmentation behavior among all stresses. The shift from large mother floc to small mother floc densities occurs near $\gamma = 10\sigma_r$. We also notice a large increase in the number of fragmentation events that occur, depicted for various (σ_r,γ) combinations in Fig. 7. For high rupture stresses, the number of fragmentations can increase by a factor of 1000 as the shear rate increases from 10 to 500 s^{-1}. For low rupture stresses, there is little difference since even small stresses will cause a fragmentation to occur and the number of fragmentations is limited primarily by the number of bacteria present. Whenever more than 1000 fragmentation events occur, the density function is well represented by the example shown in Fig. 5(a), and typically the number of fragmentations exceeds 1000 when $\gamma > 10\sigma_r$.

C. Particle density after exhaustive fragmentation

Many different functional forms of the postfragmentation probability density have been used in the literature. As mentioned earlier, the two most common are binary and log-normal. For exhaustive fragmentation, our results do not convincingly match the log-normal density. We denote the exhaustive density as Γ_{EX}, the size density of daughter flocs

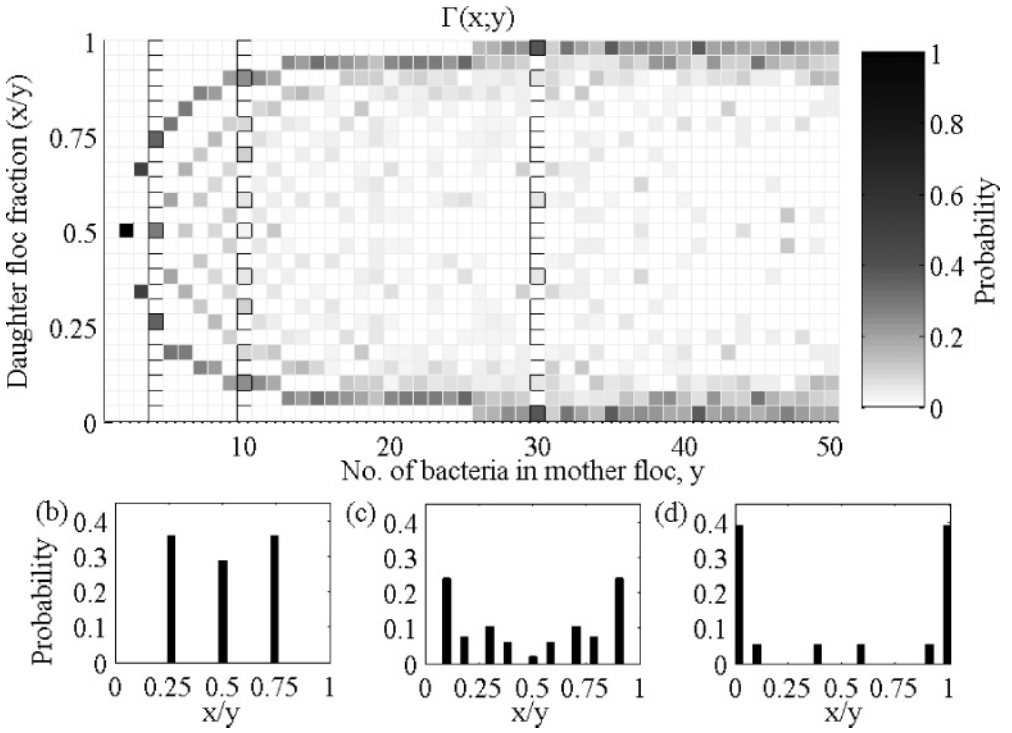

FIG. 5. Sample postfragmentation density function $\Gamma(x; y)$, depicting dependence on mother floc size, y. (a) Contour plot of the density. The y axis is the mother floc size and the x axis is the size of a daughter floc as a fraction of the size of the mother floc. Also shown are (b) $\Gamma(x; 4)$, (c) $\Gamma(x; 10)$, and (d) $\Gamma(x; 30)$.

after repeated fragmentation. Our computations suggest that Γ_{EX} exhibits many small and few large flocs. Figure 8(a) depicts a sample comparison between Γ_{EX} and a discretized log-normal density fit to Γ_{EX} for $\sigma_r = 10$ Pa and $\gamma = 150\ s^{-1}$.

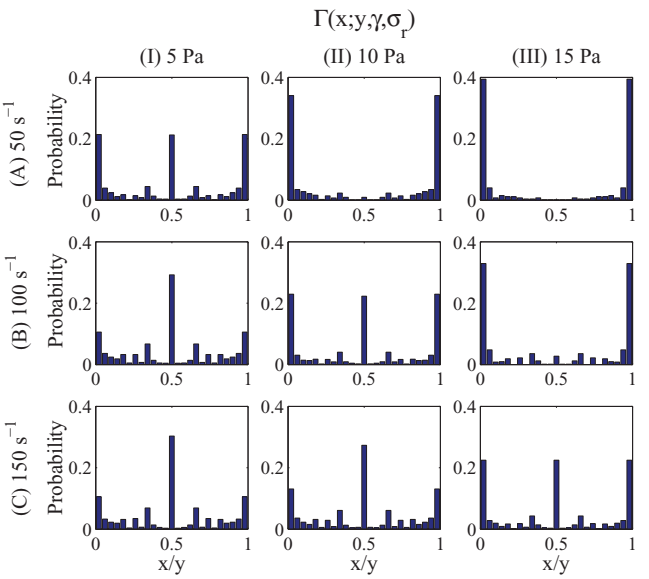

FIG. 6. Postfragmentation densities for various combinations of fluid shear rates and biofilm rupture stresses. As fluid shear rate increases, the number of fragmentations increases (Fig. 7) resulting in many small flocs fragmenting and a more centralized density function. Increasing rupture stress has the opposite effect, decreasing the number of fragmentations and resulting primarily in erosion.

The Hellinger distance [39] between Γ_{EX} and the fitted log-normal density ranges between 0.13 and 0.74. While the asymptotic tails of Γ_{EX} and the log-normal densities do match, the majority of the probability in Γ_{EX} is not captured by log-normal.

Note that to allow comparison between Γ_{EX} and the log-normal, we discretized the log-normal density over a partition of the domain chosen to match the Γ_{EX} bins. Naturally, this

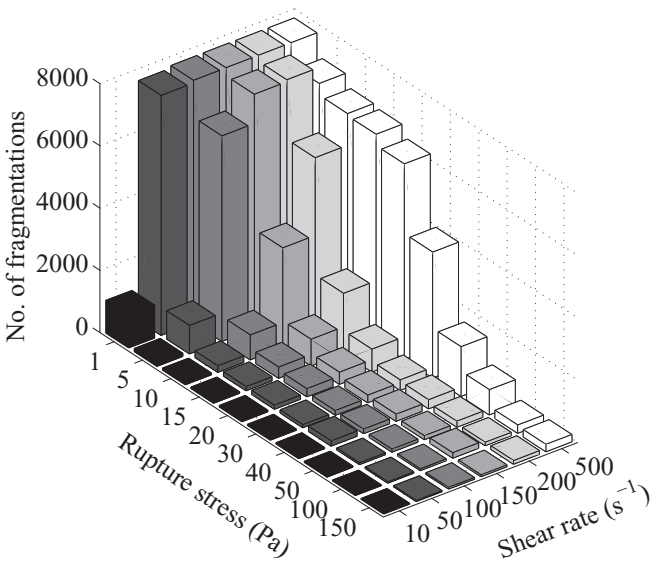

FIG. 7. Total number of fragmentations that occur for different values of biofilm rupture stress (σ_r) and fluid shear rate (γ).

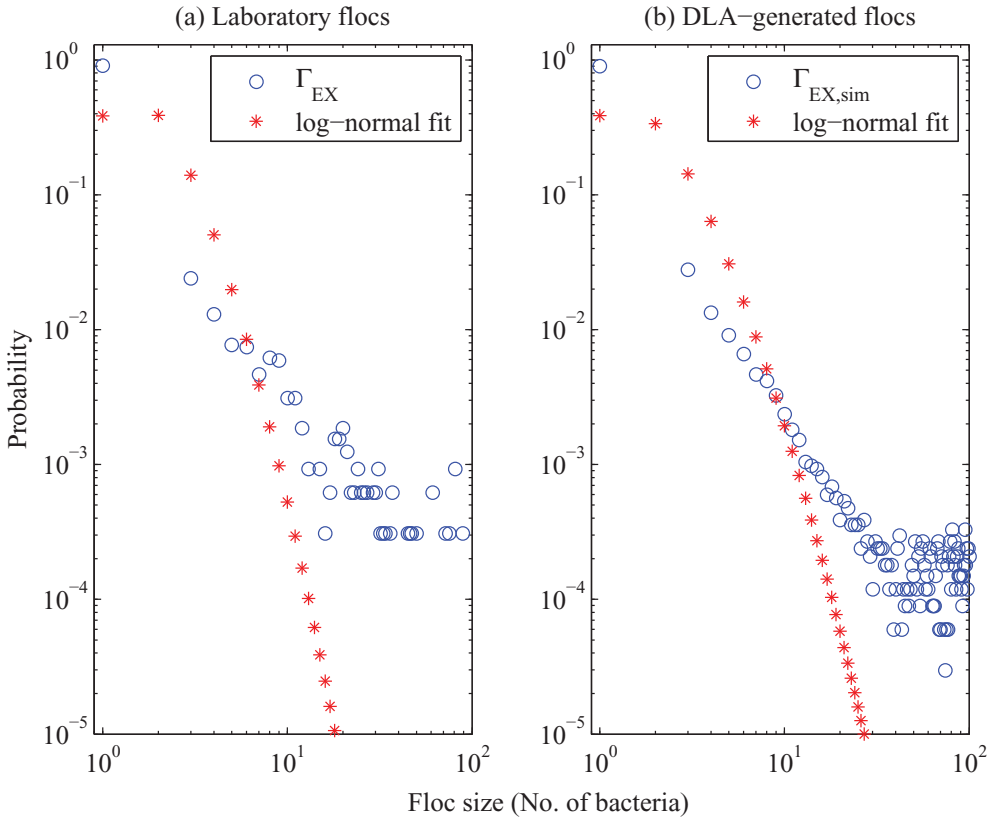

FIG. 8. (Color online) Sample floc size density ($\sigma_r = 10$ Pa, $\gamma = 150$ s^{-1}) after exhaustive fragmentation Γ_{EX} (blue circles), along with fitted, discretized log-normal densities (red stars). Densities are shown for (a) laboratory-acquired flocs and (b) DLA-generated flocs. The x axis is the number of bacteria in a remaining floc and the y axis is probability. Γ_{EX} reaches a lower bound near flocs of size 20 in the laboratory-acquired flocs and size 40 in the DLA-generated flocs due to the finite number of flocs present. Note the underprediction of single bacteria flocs and flocs larger than 8, and the overprediction of flocs between these values.

does introduce some error, but it is on the order of the bin widths. Furthermore the derivation of the log-normal density assumes homogeneity of structure. As can be seen in Fig. 1, *K. pneumoniae* flocs are heterogeneous; the fractal dimensions for our flocs range between 1.8 and 2.4.

D. Comparison with similar diffusion-limited-aggregation-generated aggregates

To explore the possible effects on our results of having a sample size of 39 flocs, we artificially generated flocs for analysis and comparison with the laboratory-acquired flocs. An extension of Witten's and Sanders's *diffusion limited aggregation* (DLA) algorithm [40] was implemented to construct these additional flocs. The DLA algorithm was modified in two ways. First, the restriction that particles be placed on a simple cubic lattice was removed, allowing new particles to be placed freely in space. Second, we allow the distance between particles as they are placed to vary, with the specific distance for each new placement being drawn from the distribution of edge lengths in the MSTs of the original 39 flocs. We do not consider the spatial distribution of edge lengths, allowing long edges to appear throughout the floc as is observed in our laboratory flocs. For further details on DLA, we direct the interested reader to Refs. [40,41].

From more than 10 000 DLA-generated aggregates, we chose 1159 for analysis by comparing the distribution of edge lengths in their MST to that of our laboratory-acquired flocs. We found the Hellinger distance between the edge-length distributions of each DLA-generated floc and the laboratory flocs. The range of Hellinger distances was 0.16 to 0.36, and we selected the flocs where the distance was less than 0.21. The fractal dimension was also checked to ensure it lay in the range of fractal dimensions present in the laboratory-acquired flocs. The selected DLA-generated aggregates were then used in our algorithm for generating a postfragmentation density function.

The resulting density function depicted in Fig. 9(a) is consistent with the density function based on our laboratory-acquired flocs. All DLA-generated flocs showed the predominant fragmentation mechanism to be erosion. The Hellinger distance between the postfragmentation density function for the laboratory-acquired flocs and that for the DLA-generated flocs for each mother floc size is shown in Fig. 9(b). The increase in this value as the mother floc size increases can be attributed to the small number of mother flocs of a specific size in our laboratory data. The DLA-generated flocs clearly have the same fragmentation behavior as the laboratory-acquired flocs and support our results.

The DLA-generated flocs were also used to construct the density of particle sizes after exhaustive fragmentation, and

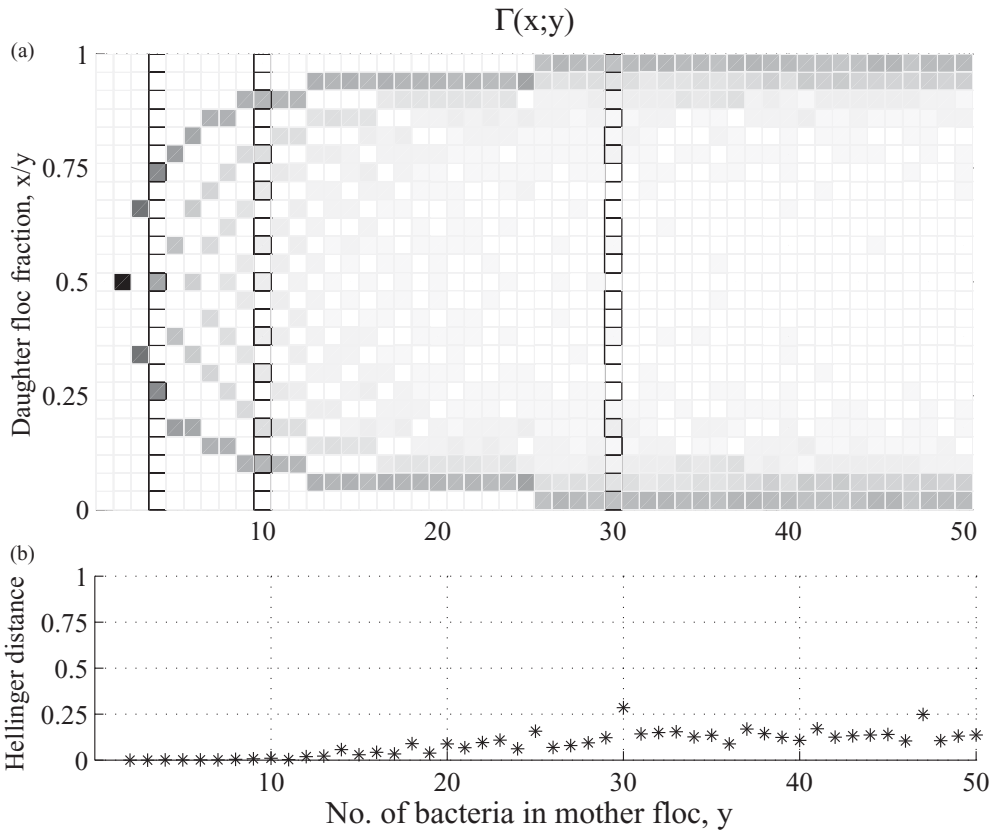

FIG. 9. (a) Contour plot of the postfragmentation density function constructed by our algorithm for 1024 DLA-generated flocs. The y axis is the mother floc size and the x axis is the size of a daughter floc as a fraction of the size of the mother floc. (b) The Hellinger distance between the postfragmentation density functions of the laboratory-acquired flocs and the DLA-generated flocs as a function of mother floc size. The y axis is the mother floc size and the x axis is the Hellinger distance.

these results are consistent with those of the laboratory-acquired flocs [Fig. 8(b)]. For the example densities shown in Fig. 8, the Hellinger distance between the data and the fitted log-normal curve is 0.29 for the laboratory data and 0.26 for the artificial data, while the Hellinger distance between the densities for the laboratory and artificial data is merely 0.01.

IV. CONCLUDING REMARKS

We propose a methodology to predict bacterial floc fragmentation in laminar flow. We apply our method to *K. pneumoniae* flocs and construct a postfragmentation density. Numerical results suggest that the primary fragmentation mechanism for medium to large biological flocs is erosion and not the splitting of the original floc into two similarly sized daughter flocs, as is commonly assumed in the literature.

We also investigate the limiting density of floc sizes Γ_{EX} after exhaustive fragmentation and find our results to be inconsistent with the log-normal density frequently employed in the literature [27,28]. Our results are supported by applying our algorithm to artificially generated flocs with statistical properties similar to the flocs acquired in the laboratory.

Much of the previous work on floc fragmentation is based on the breakup of flocs whose structure is due to short-range van der Waals attractions. Conversely, the structure of bacterial flocs is based on polymer networks secreted by the bacteria themselves, and the deformation and rupture of polymer networks is fundamentally different [7]. We find that rupture stress, shear rate, and floc shape all have minimal impact on the postfragmentation density function. While we acknowledge that our treatment of the flocs as brittle is simplistic, our proposed methodology forms the foundation for more accurate modeling of fragmentation. Accordingly, future efforts will be directed toward experimental validation and mathematical investigation of the density identification inverse problem.

APPENDIX A: FORMULATIONS OF χ_k AND A

The following formulations for χ_k and $\mathbf{A}$ were adapted from Blaser [35]. For the ellipsoid D: $\frac{x^2}{a_1^2}+\frac{y^2}{a_2^2}+\frac{z^2}{a_3^2}=1$, where $a_1 > a_2 > a_3$, let $r_1 = \frac{a_1}{a_3}$, $r_2 = \frac{a_2}{a_3}$, and $r_3 = 1$. Then the quantity χ_k is given by

$$\chi_k = \int_0^\infty \frac{u\,du}{\left(r_k^2+u^2\right)\sqrt{\left(r_1^2+u^2\right)\left(r_2^2+u^2\right)\left(r_3^2+u^2\right)}}. \quad \text{(A1)}$$

To define the matrix **A**, let us first define the additional quantities χ_k' and χ_k'':

$$\chi_k' = \int_0^\infty \frac{u\left(r_k^2+u^2\right)du}{\left[\left(r_1^2+u^2\right)\left(r_2^2+u^2\right)\left(r_3^2+u^2\right)\right]^{3/2}}, \quad \text{(A2)}$$

$$\chi_k'' = \int_0^\infty \frac{u^3\left(r_k^2+u^2\right)du}{\left[\left(r_1^2+u^2\right)\left(r_2^2+u^2\right)\left(r_3^2+u^2\right)\right]^{3/2}}. \quad \text{(A3)}$$

Using these definitions, the elements of **A** are given by

$$A_{i,i} = \frac{1}{d}\sum_{k=1}^{3}\alpha_k\,\chi_k''\,E_{k,k}$$

$$A_{i,j} = \frac{-\chi_j\,E_{i,j} + r_i^2\,\chi_k'\left(\Omega_{i,j}+\omega_p^k\right)}{2\chi_k'\left(r_i^2\chi_i + r_j^2\chi_j\right)}, \quad \text{(A4)}$$

$$k \in \{1,2,3\}, \quad k \neq i,j,$$

where

$$\alpha_k = \begin{cases} 2, & k = i, \\ -1, & k \neq i, \end{cases}$$

$$d = 6\left(\chi_1''\,\chi_2'' + \chi_2''\,\chi_3'' + \chi_1''\,\chi_3''\right).$$

APPENDIX B: ALGORITHM PSEUDOCODE

The following pseudocode is a more detailed description of the algorithm from Sec. II E used to construct the postfragmentation density functions from the 3D positional data for the flocs.

```
(1)  INPUT = Flocs = list of flocs
(2)    γ = fluid shear rate
(3)    σ_r = biofilm rupture stress
(4)  i = 0
(5)  WHILE i ⩽ (No. of flocs in Flocs)
(6)    X = centers of mass for bacteria in Flocs(i)
(7)    dist = √(X * X^T)
(8)    Edges = MST(X,dist)
(9)    [A, var] = PCA(X)
(10)   semiaxes = √var;
(11)   D = ellipsoid(semiaxes)
(12)   f_R = forces(D, γ)
(13)   FOR = k = 1 : (No. of edges in Edges)
(14)     Edge_k = k^th longest edge
(15)     P_⊥ = ⊥ bisector of Edge_k
(16)     n_⊥ = normal to P_⊥
(17)     F_⊥ = | ∬_{D>P_⊥} f_R · n_⊥ dσ | + | ∬_{D<P_⊥} f_R · n_⊥ dσ |
(18)     σ_max = F_⊥ / area(D ∩ P_⊥)
(19)     IF = σ ⩾ σ_r
(20)       REMOVE Edge_k from Edges
(21)       ADD daughter flocs TO Flocs
(22)       BREAK
(23)     END IF
(24)     IF k = (No. of edges in Edges)
(25)       ADD Flocs(i) TO Unfragmented Flocs
(26)     END IF
(27)   END FOR
(28)   i = i + 1
(29) END WHILE
```


[1] D. M. Bortz, T. L. Jackson, K. A. Taylor, A. P. Thompson, and J. G. Younger, Bull. Math. Biol. **70**, 745 (2008).
[2] M. Smoluchowski, Phys. Z **17**, 557 (1916).
[3] P. Somasundaran, V. Runkanan, and P. Kapur, in *Coagulation and Flocculation*, 2nd ed., edited by H. Stechemesser and B. Dobias (CRC Press, Boca Raton, FL, 2005), Chap. 11, pp. 767–803.
[4] A. T. Potanin, J. Colloid Interface Sci. **157**, 399 (1993).
[5] K. Higashitani and K. Iimura, J. Colloid Interface Sci. **204**, 320 (1998).
[6] M. L. Eggersdorfer, D. Kadau, H. J. Herrmann, and S. E. Pratsinis, J. Colloid Interface Sci. **342**, 261 (2010).
[7] R. C. Sonntag and W. B. Russell, J. Colloid Interface Sci. **115**, 378 (1987).
[8] S. Blaser, J. Colloid Interface Sci. **225**, 273 (2000).
[9] D. F. James, N. Yogachandran, M. R. Loewen, H. Liu, and A. M. J. Davis, J. Pulp Pap. Sci. **29**, 377 (2003).
[10] V. A. Tolpekin, M. H. G. Duits, D. van den Ende, and J. Mellema, Langmuir **20**, 2614 (2004).
[11] N. D. Vassileva, D. van den Ende, F. Mugele, and J. Mellema, Langmuir **22**, 4959 (2006).
[12] S. Blaser, Colloids Surf. A **166**, 215 (2000).
[13] A. K. C. Yeung and R. Pelton, J. Colloid Interface Sci. **184**, 579 (1996).
[14] M. Hermawan, G. C. Bushell, V. S. J. Criag, W. Y. Teoh, and R. Amal, Langmuir **20**, 6450 (2004).
[15] M. Kobayashi, Colloids Surf. A **235**, 73 (2004).
[16] M. Kobayashi, Water Res. **39**, 3723 (2005).
[17] R. Podschun and U. Ullman, Clin. Microbiol. Rev. **11**, 589 (1998).
[18] J. M. Matsen, J. A. Spindler, and R. O. Blosser, Appl. Microbiol. **28**, 672 (1974).
[19] B. Han, S. Akeprathumchai, S. Wickramasinghe, and X. Qian, AIChE J. **49**, 1687 (2003).
[20] D. Ramkrishna, *Population Balances: Theory and Applications to Particulate Systems in Engineering* (Academic Press, San Diego, 2000).
[21] D. N. Thomas, S. J. Judd, and N. Fawcett, Water Res. **33**, 1579 (1999).
[22] P. T. Spicer and S. E. Pratsinis, AIChE J. **42**, 1612 (1996).
[23] A. S. Ackleh, Nonlinear Anal. Theory Methods Appl. **28**, 837 (1997).
[24] I. N. McCave, Deep-Sea Res. **31**, 329 (1984).
[25] R. A. Speers, T. D. Durance, M. A. Tung, and J. Tou, Biotechnol. Prog. **9**, 267 (1993).
[26] A. N. Kolmogoroff, Dokl. Akad. Nauk SSSR **31**, 99 (1941).
[27] C. F. Lu and L. A. Spielman, J. Colloid Interface Sci. **103**, 95 (1985).

[28] J. D. Pandya and L. A. Spielman, J. Colloid Interface Sci. **90**, 517 (1982).
[29] S. Dzul, M. Thornton, D. Hohne, D. Bortz, M. Solomon, and J. Younger (unpublished).
[30] A. Mohraz and M. Solomon, Langmuir **21**, 5298 (2005).
[31] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. (MIT Press/McGraw-Hill, Cambridge, MA/New York, 2001).
[32] J. Kruskal, Proc. Am. Math. Soc. **7**, 48 (1956).
[33] I. Jolliffe, *Principal Component Analysis* (Springer, New York, 2002), 2nd ed.
[34] H. Lamb, *Hydrodynamics* (Cambridge University Press, Cambridge, UK, 1932).
[35] S. Blaser, Chem. Eng. Sci. **57**, 515 (2002).
[36] D. H. Bache, Chemical Eng. Sci. **59**, 2521 (2004).
[37] D. H. Bache, Water Sci. Technol. **12**, 55 (2004).
[38] G. Jeffery, Proc. R. Soc. London A **102**, 161 (1922).
[39] A. W. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, Cambridge, UK, 1998).
[40] T. A. Witten and L. M. Sander, Phys. Rev. Lett. **47**, 1400 (1981).
[41] S. Havlin, *Fractals in Physics* (North-Holland, Amsterdam, 1986).