Optimal surface estimation and thresholding of confocal microscope images of biofilms using Beer's Law

Authors: Albert E. Parker, J. A. Christen, Lindsey A. Lorenz, and Heidi J. Smith

© This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Parker, A.E., Christen, J.A., Lorenz, L., & Smith, H. (2020). Optimal surface estimation and thresholding of confocal microscope images of biofilms using Beer's Law. *Journal of Microbiological Methods*, 174, 105943. <https://doi.org/10.1016/j.mimet.2020.105943>.

Made available through Montana State University's [ScholarWorks](https://scholarworks.montana.edu)
scholarworks.montana.edu

1 Optimal surface estimation and thresholding of confocal
2 microscope images of biofilms
3 using Beer's Law

4 A. E. Parker^{a,b,*}, J. A. Christen^c, L. Lorenz^a, H. Smith^{a,d}

5 ^aCenter for Biofilm Engineering, Montana State University, Bozeman, MT, USA

6 ^bDepartment of Mathematical Sciences, Montana State University, Bozeman, MT, USA

7 ^cCentro de Investigación en Matemáticas, Jalisco S/N, Valenciana, Guanajuato, GTO,
8 36023, MEXICO

9 ^dDepartment of Microbiology and Immunology, Montana State University, Bozeman,
10 MT, USA

11 **Abstract**

Beer's Law explains how light attenuates into thick specimens, including thick biofilms. We use a Bayesian optimality criterion, the maximum of the posterior probability distribution, and computationally efficiently fit Beer's Law to the 3D intensity data collected from thick living biofilms by a confocal scanning laser microscope. Using this approach the top surface of the biofilm and an optimal image threshold can be estimated. Biofilm characteristics, such as bio-volumes, can be calculated from this surface. Results from the Bayesian approach are compared to other approaches including the method of maximum likelihood or simply counting bright pixels. Uncertainty quantification (i.e., error bars) can be provided for the parameters of interest. This approach is applied to confocal images of stained biofilms of a common lab strain of *Pseudomonas aeruginosa*, stained biofilms of *Janthinobacterium* isolated from the Antarctic, and biofilms of *Staph aureus* that have been genetically modified to fluoresce green.

*Corresponding author
Preprint submitted to Elsevier
Email address: parker@math.montana.edu (A. E. Parker)
URL: www.math.montana.edu/parker (A. E. Parker)

May 8, 2020

12 *Keywords:* Attenuation, thresholding, Maximum likelihood, Beer-Lambert
13 Law, Bayesian, confocal microscope image analysis

14 **1. Introduction**

15 Confocal Scanning Laser Microscopy (CSLM) has become an essential
16 instrument in biofilm research because of the ability to generate 3D visu-
17 alizations of hydrated biofilms noninvasively and in real-time [e.g., 25, 11].
18 Fluorophores are introduced into the microbes by either staining or geneti-
19 cally modifying the species to express a fluorescent protein. CSLM utilizes
20 the same lens to excite and to detect emitted light from these fluorophores
21 [20]. The pinhole system enables the removal of out-of-focus light necessary
22 for the generation of high resolution 3D images (10^8 pixels or more) at spatial
23 resolutions less than $1\mu\text{m}$ (see Figure 1).

24 When imaging thick biofilms, the laser's intensity attenuates markedly
25 as it passes through the top layers of the biofilm so that the bacteria deep
26 in the biofilm do not fluoresce at all [23]. Hence, CSLM image resolution is
27 restricted by the depth of sample penetration which has made the quantita-
28 tive assessment of CSLM data challenging. With conventional CSLM, light
29 begins to scatter severely between $60\text{-}80\mu\text{m}$ [4]. The penetration depth can
30 be greater for more advanced CSLM, e.g. multiphoton confocal microscopes,
31 but attenuation still occurs for thick enough specimens. This phenomenon
32 is predicted by *Beer's Law* (or the Beer-Lambert Law), a physics model that
33 describes how light transmission decays as the microscope's laser is focused
34 deeper into a specimen [6, 26]. Our investigations confirm that this model
35 also well explains, quantitatively, the phenomenon in biofilms.

Figure 1: CSLM images of thick biofilms ($> 100\mu\text{m}$) visualized with Imaris. These images are analyzed in this paper. A: Using a 25x objective, a $620\mu\text{m} \times 620\mu\text{m}$ field of view over $1024 \times 1024 \times 163$ pixels of a $192\mu\text{m}$ thick stained *Pseudomonas aeruginosa* biofilm. B: Using a 25x objective, a $620\mu\text{m} \times 620\mu\text{m}$ field of view over $512 \times 512 \times 17$ pixels of a $112\mu\text{m}$ thick biofilm of GFP *Staph. aureus*. The hollowed out appearance of the biofilm in the side-panels is due to the attenuation of light because of Beer's Law. C: When using a 63x objective, a $246\mu\text{m} \times 246\mu\text{m}$ field of view over $1024 \times 1024 \times 286$ pixels of a $141\mu\text{m}$ thick stained *Janthinobacterium* biofilm isolated from Antarctica.

36 We present a novel approach for the analysis of CSLM data that fits
 37 Beer’s Law to the intensities in these 3D images. Our approach considers
 38 a *Beer’s threshold*, that identifies which bright pixels (with fluorescence in-
 39 tensities above Beer’s threshold) are associated with the top surface of the
 40 biofilm. This is in contrast to a conventional image threshold that specifies
 41 which bright pixels (with intensities above the threshold) are associated with
 42 biomass regardless of the spatial location. Our approach considers a range of
 43 possible Beer’s threshold values provided by the user (e.g., the microscopist
 44 or data analyst). For each possible Beer’s threshold, a linear regression model
 45 that describes Beer’s Law is fit to *all* of the intensity data below the surface
 46 by least squares. Reported are two Beer’s thresholds that are optimal com-
 47 pared to any other candidate in the range of possible thresholds: (1) the
 48 maximum likelihood estimate (MLE) that minimizes the mean squared er-
 49 ror (or, equivalently, maximizes the R^2 as for any regression model); and
 50 (2) the maximum of the Bayesian posterior distribution (MAP) that max-
 51 imizes the probability of the threshold given the data. Importantly, from
 52 the biofilm’s surface, our approach can quantify other biofilm characteris-
 53 tics, such as bio-volume, from a CSLM image. Methods that require the user
 54 to subjectively choose a single threshold can result in quantitative outcomes,
 55 like bio-volumes, that are biased [e.g., see 18].

56 There are a number of automatic threshold detection schemes already
 57 available in image analysis software such as COMSTAT (implemented in
 58 MATLAB) [14], COMSTAT2 (implemented in ImageJ) [27], Imaris [16], and
 59 daime [8]. Otsu’s method for threshold selection, used by the COMSTAT im-
 60 plementations in MATLAB and ImageJ, maximizes the intra-class correlation

[19], where each class is composed of pixels with intensities either above or below the threshold. Imaris implements a method due to [7] that iteratively reduces the threshold until the correlation of the image using only pixels with intensities below the threshold reaches 0 [16, p. 241]. The Objective Threshold Selection (OTS) method [28], based on Robust Automated Threshold Selection (RATS-L), is implemented by the software daime [8] and uses an edge matrix to estimate the threshold, then keeps modifying the threshold until bio-volume, defined as the number of bright pixels above the threshold, stabilizes. A variant of OTS, the bio-volume elasticity method [18], uses a heuristic approach to choose a threshold based on the rate of decrease of the bio-volume as the threshold is increased. Due to Beer's Law, quantifying bio-volume simply by counting bright pixels is unbiased only if the biofilm being imaged is thin (e.g., less than $100\mu\text{m}$). For thick biofilms (Figure 1), counting bright pixels would underestimate the bio-volume [22]. We compare our approach to Otsu's (via COMSTAT2 in ImageJ), the method due to [7] (via Imaris) and RATS-L (via daime).

Uncertainty quantification via a Bayesian probability interval (also called a credible interval), or a confidence interval, can be provided by our approach for all parameters of interest. For example, our approach provides the uncertainty in the Beer's threshold, the location of the top surface of the biofilm, and the bio-volume from a single image. In contrast, uncertainty quantification for characteristics from single images are not provided by Costes' approach (in Imaris) or by Otsu's (COMSTAT2 in ImageJ). OTS determines a confidence interval for the bio-volume by generating a threshold for each column in the image and then considering the distribution of threshold values

86 [29].

87 To illustrate that our technique can be applied successfully to many dif-
88 ferent types of microbiological samples, we analyze CSLM images from three
89 individual organisms: (1) a lab strain of *P. aeruginosa* (ATCC #15442),
90 likely the most studied biofilm forming bacterium ever; (2) a genetically
91 modified green fluorescing *Staphylococcus aureus*; and (3) an environmen-
92 tal *Janthinobacterium* isolate (Figure 1). To demonstrate the accuracy of
93 our method, results are compared to other available image analysis software
94 packages when possible.

95 2. Methods

96 2.1. Biofilm growth and image acquisition

97 *Pseudomonas aeruginosa* biofilms (ATCC 15442, <https://www.atcc.org/products/all/15442.aspx>)
98 were grown in the high shear environment in the CDC Biofilm Reactor using
99 ASTM method E2542 [13]. A biofilm sample was then stained with 200 μ L
100 of a diluted 1:1 mixture of Syto 9 and Propidium Iodide (LIVE/DEAD Ba-
101 cLight Bacterial Viability Kit stain (Invitrogen, catalog #L7012)) and then
102 imaged.

103 Green-fluorescing protein (GFP) *Staphylococcus aureus* AH2547 were kindly
104 provided by Dr. Alexander Horswill [21]. These GFP *S. aureus* biofilms were
105 grown in a separate CDC Biofilm Reactor. The growth medium was supple-
106 mented with 10 μ g/mL of chloramphenicol to retain the GFP plasmid. A
107 biofilm sample was imaged over time using the Treatment Imaging Flow Cell
108 (BioSurface Technologies Corporation, catalog #FC 30). The time series was
109 analyzed by [22]; here we analyze a single image from that series.

110 Biofilm formation was also investigated for Antarctic isolate *Janthinobac-*
 111 *terium* sp. strain CG23_2, a psychrotolerant, aerobic, violacein-pigmented,
 112 rod-shaped, Gram-negative, catalase-positive organism [24]. *Janthinobac-*
 113 *terium* biofilms were grown in the CDC Biofilm Reactor following the same
 114 procedure as above [13] with the following modifications. The reactor was in-
 115 oculated with a second-generation mid exponential phase bacterial inoculum
 116 grown in a minimal R2A medium (BD Difco, purchased from Fisher Scien-
 117 tific, catalog number DF186-07-3) at 15 degrees C. Bacteria were grown in
 118 full strength R2A media for 48 hours to allow for initial colonization. After
 119 48 hours the pump was turned on with a flow rate of 1.67 mL/min, the inflow
 120 media was a 1 to 10 dilution of R2A to induce biofilm production. Through-
 121 out the reactor run, biofilms were subjected to a constant shear of 80 rpm.
 122 Following seven days of growth coupons were removed from the rods and
 123 stained with a 1:1 mixture of Syto 9 and Propidium Iodide (LIVE/DEAD
 124 BacLight Bacterial Viability Kit stain (Invitrogen, catalog #L7012)).

125 The fully intact and hydrated biofilm samples described above were im-
 126 aged on an upright Leica TCS-SP5 CSLM using using either a 25X/0.95NA
 127 2.5mmWD water objective or a 63X/0.9NA 2.2mmWD water objective. Flu-
 128 orophore excitation lasers and emission bandwidths were as follows: Syto9
 129 (live) 488 nm excitation, 500–550 nm emission collection; Propidium iodide
 130 (dead), 561 nm excitation, 580–700 nm emission collection.

131 Gray-scale (8-bit) images were generated using Imaris [16]. Only the
 132 “LIVE” (green) channel was analyzed for each of the 3 bugs. The analysis
 133 methodology presented here can be updated for the simultaneous analysis of
 134 multiple florescence channels. We investigate this possibility elsewhere.

135 2.2. Model description

136 2.2.1. Image data

137 The CSLM images in Figure 1 were collected over different lattice sizes
138 $n_x \times n_y \times n_z$, chosen by the microscopist. In our context, a lattice is an
139 ordered array of points describing the arrangement of pixels that form a 3D
140 CSLM image. The x - and y -dimensions of the lattice are associated with the
141 horizontal dimensions parallel to the surface where the biofilm is attached.
142 The z -dimension is associated with the vertical dimension orthogonal to the
143 surface. A CSLM constructs 3D images by a z -stack; that is, by first col-
144 lecting a single $n_x \times n_y$ planar image called a z -slice at each vertical location
145 or distance (z) into the biofilm and then stacking the z -slices together. The
146 number of x and y pixels are usually the same, $n_x = n_y$. Each z -slice is
147 generated by a raster laser scanner that causes microbes and/or other mate-
148 rial in the biofilm that have been stained or genetically modified to fluoresce
149 when excited by the laser.

150 At each spatial location corresponding to a pixel in the image, the CSLM
151 records the intensity of the biofilm’s fluorescence as an 8-bit integer (i.e., an
152 integer value between 0 and 255). In our examples, the horizontal (xy) field
153 of view for each planar z -slice is captured by either a 25x or 63x objective
154 lens resulting in either a $620\mu m \times 620\mu m$ or $246\mu m \times 246\mu m$ field of view
155 respectively. Because microscopists desire high spatial resolutions to achieve
156 appealing visualizations, they typically choose high dimensional pixelations,
157 $n_x = n_y = 1024$ (Figure 1A and 1C). However, when generating a movie of
158 images [e.g., see 22], the microscopist may back off on the spatial resolution,
159 say $n_x = n_y = 512$ (Figure 1B) to increase the temporal resolution of the time

lapse image frames in the movie. In our examples, the vertical dimension (z) ranges from $100\mu\text{m}$ to $200\mu\text{m}$. The microscopist sets the number (n_z) of z -slices over this range depending on the thickness of the biofilm. In the analyses that follow, each z -slice is identified with an integer value between $z = 1$ (where, ideally, the biofilm is attached to a substratum) and $z = n_z$ (the z -slice at the very top of the image). In a similar way each pixel in the horizontal x and y dimensions can also be identified with an integer so that each pixel in the 3D image can be identified by the coordinates (x, y, z) where $1 \leq x \leq n_x$, $1 \leq y \leq n_y$ and $1 \leq z \leq n_z$.

2.2.2. Beer's Law

When there is uniform attenuation of light throughout the depth of the biofilm, then Beer's Law takes the simple form

$$I = I_0 e^{-\psi c \ell} \quad (1)$$

[26] where

- I_0 is the intensity of light received (from the laser) by a single z -slice in the CSLM
- I is the intensity of light transmitted by a single z -slice in the CSLM
- ψ is the attenuation cross section of the fluorophore in the z -slice
- c is the three-dimensional volumetric number density of the fluorophore in the z -slice. This is an important feature of the model for biofilms because it is known that the bacteria and the matrix that glues them together are spatially heterogeneous, and so we expect c to vary throughout the biofilm.

182 • ℓ is the path length of the beam of light through the z -slice

183 We fit this physical model to the intensities in each 3D image.

184 2.2.3. Regression model

185 To fit Beer’s Law to data it is crucial to know where the top surface of
 186 the biofilm is. A key feature in the modeling process is that an edge detector
 187 is applied to the 3D light intensity image data to estimate the location of the
 188 surface of the biofilm. The simple edge detection scheme that we use was
 189 described previously in [22] and is summarized as follows. First the data is
 190 thresholded so that pixels with an intensity value less than some candidate
 191 Beer’s threshold θ are set to 0, and other pixels are set to 1. The value θ is
 192 called a *Beer’s threshold* because, unlike conventional thresholding, it is only
 193 used to find the biofilm’s top surface. The biofilm’s surface, z_{surf} is estimated
 194 in each column of the image (at each (x, y) location) by applying the edge
 195 detector to the thresholded image data, starting at the top ($z = n_z$) and
 196 moving toward the bottom of the image ($z = 1$) until an edge is detected (at
 197 $z = z_{surf}$). Thresholding is a common feature to finding edges in any kind of
 198 image. Any edge detector can be used, but we use a simple computationally
 199 inexpensive scheme that identifies the set of z values for which there are two
 200 adjacent z -slices with non-zero pixels in the thresholded data. The biofilm’s
 201 surface or thickness, z_{surf} , is then set to the largest value of z in this identified
 202 set. At each planar spatial location (x, y) , the top surface of the biofilm is
 203 denoted by $z_{surf}(x, y|\theta)$, that emphasizes the dependence on the candidate
 204 Beer’s threshold θ .

205 Hence, there is a set of candidate biofilm surfaces found by the edge

Figure 2: Visualization of the quadratic and linear regression model for Beer's Law fit to a single column from the CSLM image of the *P. aeruginosa* biofilm in Figure 1A. A: Fit of the model to the column of intensity data (I) at $(x = 300, y = 100)$ when the parabola models a single region of high fluorescence (i.e., $\beta_3 < 0$). B: Fit of the model to the column of intensity data (I) at $(x = 300, y = 200)$ when the parabola models two regions of high fluorescence (i.e., $\beta_3 > 0$).

206 detector, one for each candidate threshold θ in the range supplied by the
 207 user (for integer valued $\theta \in [\theta_{\min}, \theta_{\max}]$): $z_{surf}|\theta = \theta_{\min}$; $z_{surf}|\theta = (\theta_{\min} + 1)$;
 208 ...; $z_{surf}|\theta = \theta_{\max}$.

209 Given a candidate Beer's threshold (θ) and an estimate of where the
 210 surface of the biofilm is (z_{surf}), then a linear regression model is fit to the
 211 log-transformed intensity data:

$$\ln(I(x, y, z|\theta)) = \begin{cases} \beta_0(x, y|\theta) + \varepsilon(x, y, z|\theta) & \text{for } z > z_{surf}(x, y|\theta) \\ f(\beta(x, y|\theta)) + \varepsilon(x, y, z|\theta) & \text{for } z \leq z_{surf}(x, y|\theta) \end{cases} \quad (2)$$

212 Note that the regression equation is different for pixels above the biofilm

213 versus pixels below the biofilm's surface. The pixel at (x, y, z) for which the
 214 z -coordinate is $z > z_{surf}(x, y)$ is above the biofilm; otherwise the pixel is
 215 inside the biofilm.

216 The regression equation in (2) specifies a set of regression coefficients for
 217 each column in the 3D CSLM image (at location (x, y)) for each value of
 218 the threshold θ , so we write $\beta := \beta(x, y|\theta)$. We will drop the dependence
 219 on (x, y) for brevity in the equations below. The $\varepsilon(x, y, z|\theta)$ term is the
 220 measurement error of the log-transformed intensities at a (x, y, z) location
 221 that is described further in section 2.3.1. The regression function f in (2)
 222 describes the intensities generated by the biofilm excited by the laser:

$$f(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) = \begin{cases} \beta_1 + \beta_2 z + \beta_3 z^2 & \text{for } I(x, y, z) \geq \theta \\ \beta_4 + \beta_5 z & \text{for } I(x, y, z) < \theta \end{cases}. \quad (3)$$

223 The terms in the linear regression model in (2) and (3) are interpreted as
 224 follows (cf. (1)):

- 225 • The surface of the biofilm is at z_{surf} (see the vertical pink line in Figure
 226 2);
- 227 • $\beta_0(x, y)$ is the noise level of the intensities above the biofilm (see the
 228 horizontal black dashed line in Figure 2);
- 229 • $\beta_1(x, y)$, $\beta_2(x, y)$ and $\beta_3(x, y)$ are the regression parameters for the
 230 parabola that models the log intensities of the brightly fluorescing
 231 biofilm just below the surface (see the green dashed-dotted parabola in
 232 Figure 2)
- 233 • the parameters $\beta_4(x, y)$ and $\beta_5(x, y)$ model the exponential decay of

234 the light intensities deep in the biofilm as prescribed by Beer’s Law
 235 (see the solid blue line in Figure 2)

236 In summary, according to Beer’s Law, the intensities of the 3D image of
 237 the biofilm depends on an estimate of the location of the biofilm’s surface;
 238 and the location of the biofilm’s surface in turn depends on identifying a
 239 Beer’s threshold θ of the data. In contrast to other methodologies, the Beer’s
 240 threshold does not indicate that we drop any pixels from consideration in
 241 our analysis. Instead, the threshold only aids in fitting the correct regression
 242 model to all of the intensities in the 3D image by identifying: (1) the top
 243 surface of the biofilm via an edge detector; (2) which pixels are “close to the
 244 surface” to be modeled as one or two bright parabolic regions (see Figure 2);
 245 and (3) which pixels are “deep in the biofilm” to be modeled by exponential
 246 decay as predicted by Beer’s Law.

247 *2.3. Inverse problem*

248 *2.3.1. Finding optimal values for Beer’s Law parameters for a single Beer’s* 249 *threshold*

250 Given CSLM image data, the inverse problem is to estimate the param-
 251 eters for the model (i.e., equations (2) and (3)) that describe the data. The
 252 parameters of interest are the Beer’s threshold θ and the regression coeffi-
 253 cients $\{\beta_i\}_{i=0}^5$. To formulate the inverse problem, we consolidate the notation
 254 from equations (2) and (3):

$$d = F_{\theta}\beta_{\theta} + \varepsilon$$

255 where

- 256 • The data $d = \ln(I)$ is a vector of length $n_x n_y n_z$ that includes all inten-
257 sities over all (x, y, z) spatial locations.
- 258 • β_θ is a vector of length $6n_x n_y$ because there are 6 regression coefficients
259 $(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ per column of image data. The θ subscript in β_θ
260 is a reminder that a set of $6n_x n_y$ regression coefficients is found for
261 each θ .
- 262 • F_θ is the *model matrix* (with dimensions $n_x n_y n_z \times 6n_x n_y$) so that $F_\theta \beta_\theta$
263 represents the simple model in (2) and (3), conditioned on the Beer's
264 threshold θ .
- 265 • ε is the vector (of length $n_x n_y n_z$) of measurement errors, $\varepsilon \sim N(0, \sigma^2 I)$.
266 That is, the measurement error is assumed to be independent and iden-
267 tically normally distributed with variance σ^2 . A correlated structure
268 may be incorporated into the likelihood, i.e., $\varepsilon \sim N(0, \sigma^2 \Omega(\gamma))$, where
269 the parameter(s) γ can be estimated as in spatial statistics. Or, as in
270 [15, 22], one may assume non-correlated errors in the likelihood as we
271 do while the spatial correlation is imposed via an informative (normal)
272 prior on β_θ or z_{surf} , e.g. with a scaled Laplacian as the prior precision
273 matrix. The results presented here using non-correlated errors in the
274 likelihood appear satisfactory. Imposing a correlation structure will
275 require far more computations, as discussed more in the Discussion in
276 section 4.

277 Given a threshold θ , it is well known that the maximum likelihood esti-
278 mator (MLE) for the $6n_x n_y$ regression parameters is

$$\hat{\beta}_{MLE}|\theta = (F_\theta^T F_\theta)^{-1} F_\theta^T d \quad (4)$$

279 [e.g., see 17]. The standard deviations $SD(\hat{\beta}_{MLE}|\theta)$ are equal to the square
 280 root of the diagonal of the variance-covariance matrix

$$Var(\hat{\beta}_{MLE}|\theta) = MSE_{\theta} \times (F_{\theta}^T F_{\theta})^{-1}; \quad (5)$$

the mean squared error is

$$MSE_{\theta} = \frac{1}{n_x n_y n_z - 6n_x n_y} \sum_{i=1}^{n_x n_y n_z} (d_i - \hat{d}_i)^2;$$

281 and the predicted log intensities are $\hat{d} = F_{\theta} \times (\hat{\beta}_{MLE}|\theta)$ (here $\times$ refers to
 282 matrix-vector multiplication) [17]. The equation for MSE uses the fact that
 283 the degrees of freedom for error is $DFE = n_x n_y n_z - 6n_x n_y$ as is typical for
 284 regression.

The Bayesian MAP estimator for the regression coefficients maximizes the posterior probability distribution $(\pi(\beta|\theta, d))$ of the regression coefficients given a threshold θ and the data [2]. When uniform priors are chosen for the regression coefficients $\{\beta_i\}$ as we do here, then

$$\hat{\beta}_{MAP}|\theta = \hat{\beta}_{MLE}|\theta$$

285 [3, 5, 15]. Moreover, in this case $\pi(\beta|\theta, d)$ is a t -distribution with mean
 286 equal to $\hat{\beta}_{MAP}|\theta$ and variance given by (5) (see section 2.3.3). Using these
 287 results, uncertainty quantification of the regression coefficients, including how
 288 to construct probability intervals, is provided in section 2.3.3.

289 2.3.2. Finding the optimal Beer's threshold

We use a brute force technique to find the Beer's threshold $\hat{\theta}_{MAP}$, over the range of possible integer threshold values supplied by the user, that maximizes the Bayesian marginal posterior distribution,

$$\hat{\theta}_{MAP} = \operatorname{argmax}_{\theta \in [\theta_{\min}, \theta_{\max}]} \pi(\theta|d)$$

290 (see section 2.3.3 for the definition of $\pi(\theta|d)$).

The MLE for the Beer's threshold minimizes the mean squared error,

$$\hat{\theta}_{MLE} = \operatorname{argmin}_{\theta \in [\theta_{\min}, \theta_{\max}]} MSE_{\theta}.$$

291 We use a brute force approach to find the MLE by first finding the MLE for
 292 the regression parameters via (4) for each integer threshold value in the user
 293 provided range, and then calculating MSE_{θ} .

294 The optimal MAP $\hat{\theta}_{MAP}$ and MLE $\hat{\theta}_{MLE}$ for the Beer's threshold are
 295 each associated with a biofilm surface (section 2.2.3). The biofilm surface
 296 $z_{surf}|\theta = \hat{\theta}_{MAP}$ maximizes the probability of the threshold given the model
 297 for Beer's Law and the data. The biofilm surface $z_{surf}|\theta = \hat{\theta}_{MLE}$ minimizes
 298 the mean squared error of the fit of Beer's Law to the data. The optimal
 299 MAP and MLE threshold and surfaces may be different as we investigate in
 300 the examples.

301 2.3.3. Uncertainty Quantification of threshold and regression parameters

302 Uncertainties are quantified for the parameters (θ and β) by the Bayesian
 303 posterior probability distribution. One advantage of the Bayesian approach
 304 is that it allows any prior information that the researcher has available about
 305 the parameters to be incorporated into the model by probability statements
 306 (via a probability distribution $\pi(\cdot)$). The prior probability distributions for
 307 the parameters in (2) and (3) are:

$$\pi(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5) = UNIF(-\infty, \infty) \quad (6)$$

$$\pi(\theta) = UNIF(\text{all integers from } \theta_{\min} \text{ to } \theta_{\max}) \quad (7)$$

$$\pi(\sigma^2) \propto 1/\sigma^2. \quad (8)$$

308 The prior for the Beer’s Law parameters (β_i) is given by a flat uniform distri-
 309 bution over the real numbers in (6). Such a flat uniform prior is sometimes
 310 described as “non-informative” because it does not prefer any values for the
 311 β_i over any other. If we wanted to constrain some parameter β_i to be pos-
 312 itive, then we might impose a prior $\pi(\beta_i) = UNIF(0, \infty)$. Or if there was
 313 some preferred value for β_i , say β_i^* , then we might impose a normal prior
 314 centered at β_i^* , that is, $\pi(\beta_i) = N(\beta_i^*, \sigma_\beta^2)$.

315 The prior for the Beer’s threshold in (7) is informed by the range, $[\theta_{\min}, \theta_{\max}]$,
 316 provided by the user. The uniform prior over the integers in this range en-
 317 forces the constraint that the Beer’s threshold is constrained to be between
 318 $\theta_{\min}$ and $\theta_{\max}$ but with no preference for any of the values in this range over
 319 any other.

320 The prior specification (8) for the variance of the measurement error
 321 prefers smaller values.

322 The results in the rest this section presume the use of the priors (6)-(8).
 323 Using non-uniform priors for either θ or β would encapsulate more specific
 324 knowledge of the value of the threshold or of the regression parameters de-
 325 scribing Beer’s Law and could, for example, result in $\hat{\beta}_{MAP}|\theta \neq \hat{\beta}_{MLE}|\theta$ (c.f.
 326 section 2.3.1) and change the form of the posterior distributions in equations
 327 (9) and (10) below (see section 5).

328 The posterior probability distribution of interest that incorporates the
 329 information from the prior distributions with the likelihood of the data is:

$$\begin{aligned} \pi(\beta, \theta|d) &\propto \pi(\beta|\theta, d) \times \pi(\theta|d) \\ &\propto t(\hat{\beta}_{MLE}|\theta, Var(\hat{\beta}_{MLE}|\theta), df = DFE) \times \pi(\theta|d) \end{aligned} \quad (9)$$

330 where

$$\pi(\theta|d) = c \sqrt{\frac{(MSE_\theta/2)^{n_x n_y (6-n_z)}}{\det(F_\theta^T F_\theta)}} \quad (10)$$

331 and c is a normalization constant (see section 5 and Figure 4).

332 A very convenient consequence of (9) is that the marginal posterior for
333 the regression coefficients, that encapsulates uncertainty in the thresholding
334 value is

$$\pi(\beta|d) = \sum_i t(\hat{\beta}_{MLE}|\theta_i, Var(\hat{\beta}_{MLE}|\theta_i)) \times \pi(\theta_i|d) \quad (11)$$

335 where the mean and covariance of each multivariate t is calculated via (4)
336 and (5), and $\pi(\theta_i|d)$ is calculated via (10). Equation (11) shows that $\pi(\beta|d)$
337 is a mixture of t distributions with mixture weights $\pi(\theta_i|d)$.

338 Uncertainty quantification for θ and β , via probability statements and
339 interval estimates, can be constructed using (10) and (11) respectively. An
340 immediate consequence is that probability intervals for $\{\beta_i\}$ cannot be con-
341 structed using the standard $t(DFE)$ distribution unless $\pi(\theta|d)$ has most of
342 its probability mass at $\hat{\theta}_{MAP}$ as in the examples we present here (see Figure
343 4).

344 2.3.4. Optimally Estimating Bio-volumes

For a fixed threshold, the volume V is estimated by integrating under the
biofilm's surface for a fixed threshold,

$$\hat{V}_\theta = \sum_{x,y} z_{surf}(x, y|\theta) \Delta_{xyz}$$

345 where Δ_{xyz} gives the voxel ("3D pixel") size in cubic microns (μm^3). Over
346 all thresholds, the biofilm's volume can be estimated from the mean and

standard deviation (SD) of the marginal posterior for the volume (see section 5)

$$\begin{aligned} Mean(V|d) &= \sum_i \hat{V}_{\theta_i} \times \pi(\theta_i|d); \\ SD(V|d) &= \sqrt{\sum_i (n_x n_y \hat{\sigma}_{surf}^2 | \theta_i) \times \pi(\theta_i|d)^2} \end{aligned}$$

where $\pi(\theta_i|d)$ is calculated via (10); and $\hat{\sigma}_{surf}^2 | \theta_i$ is calculated via (14) in the Appendix.

3. Results

3.1. Image Analysis

We use our optimal Bayesian (MAP) and likelihood (MLE) approaches to analyze the biofilm images in Figure 1.

Figure 2 illustrates the fit of Beer’s Law to two different columns of the *P. aeruginosa* image data (see the description of the figure in the context of the model in the text after equation (3)). This plot makes clear that the model is fit to the intensity data for every pixel in the image; no data were excluded prior to the model fit. Figure 3A shows the surface of the *P. aeruginosa* biofilm over all of the 1024 columns in a single x -slice and Figure 3B shows the biofilm’s surface over the entire $620\mu m \times 620\mu m$ field of view. Figure 4A shows the marginal posterior (in (10)) for the Beer’s threshold parameter (maximized by $\hat{\theta}_{MAP} = 7$) alongside the likelihood (maximized by $\hat{\theta}_{MLE} = 7$). In this example, the MAP and MLE are the same. Because the log-transformed posterior at the MAP is orders of magnitude larger than at other values, the Beer’s threshold value is 7 with probability close to 1.

Figure 3: Visualizations of the surface of the $192\mu\text{m}$ thick biofilm of *P. aeruginosa* from Figure 1A based on the optimal fit of Beer's Law to the intensity data. A: An estimate of the biofilm's surface in a single x -slice ($x = 300$). Pixels are colored according to the value of the log-intensity. B: An estimate of the surface over the entire field of view. The vertical dimension is colored according to the biofilm's thickness.

Figure 4: The log marginal posterior, $\pi(\theta|d)$, for the Beer’s threshold (θ) compared to the likelihood $\pi(d|\theta, \beta, \sigma^2)$ given CSLM data (d). A. For *P. aeruginosa*, the MLE and the MAP are clearly the same at $\theta = 7$. B. For *S. aureus*, the MLE is at $\theta = 6$, whereas the MAP is at $\theta = 8$.

Figure 4B shows different results for the MAP and the MLE for the Beer’s threshold for the *S. aureus* image data. With probability close to 1, the Beer’s threshold is equal to the MAP at $\hat{\theta}_{MAP} = 8$, whereas with probability close to 0 it is equal to the MLE at $\hat{\theta}_{MLE} = 6$. The associated MLE and MAP estimates for the surface of the *S. aureus* biofilm are presented in Figure 5. The differences in the two surface estimates are subtle. Upon closer inspection (compare the black rectangular region in both panes), the MAP estimate in Figure 5B illustrates a smoother fit to the data whereas the MLE surface in Figure 5A is more noisy (i.e., it does not fit the data as well). Inspection of residual plots (similar to Figure 3A, not shown) confirm this conclusion.

Figure 6 shows results when our Bayesian image analysis approach is ap-

Figure 5: Visualizations of two different surface estimates of the $112\mu\text{m}$ thick biofilm of *Staph. aureus* based on two optimal fits of Beer's Law to the intensity data. The vertical dimension is colored according to the biofilm's thickness. The black rectangle in each pane highlights a region of the biofilm's surface where the MLE is much noisier than the MAP. A: The biofilm's surface estimated via the MLE at $\hat{\theta}_{MLE} = 6$. B: The surface estimated via the Bayesian MAP for the Beer's threshold at $\hat{\theta}_{MAP} = 8$.

Figure 6: Visualizations of the surface of the $141\mu\text{m}$ thick biofilm of *Janthinobacterium* from Figure 1C based on the optimal fit of Beer's Law to the intensity data. A: The estimate of the biofilm's surface in an x slice ($x = 325$). Pixels are colored according the value of the log-intensity. B: The surface of the biofilm using the Bayesian MAP for the Beer's threshold, $\hat{\theta}_{MAP} = 26$. The vertical dimension is colored according to the biofilm's thickness.

379 plied to CSLM data collected from an environmental isolate from the Antarc-
 380 tic, *Janthinobacterium*. The x -slice in Figure 6A shows a biofilm architecture
 381 that is much different than the stained *P. aeruginosa* biofilm or the GFP *S.*
 382 *aureus* biofilm. Still, the x -slice suggests that our approach provides a rea-
 383 sonable estimate of the biofilm’s surface. Figure 6B shows an estimate of
 384 the surface over all of the x -slices. The MLE is $\hat{\theta}_{MLE} = 24$ and the MAP
 385 is $\hat{\theta}_{MAP} = 26$ (the posterior and likelihood curves similar to Figure 4 not
 386 shown). However, unlike the differing results for *S. aureus* (Figure 5), there
 387 is little qualitative difference in the MLE and MAP estimates of the surface
 388 for this example.

389 3.2. Assessing model fit

390 Figures 2, 3A and 6A function like residual plots that can be generated
 391 after any regression analysis. These plots allow the data analyst to assess
 392 the fit of the model to the data. Unfortunately the other image analyses
 393 discussed here (COMSTAT in MATLAB or ImageJ, Imaris and daime) do
 394 not provide residual diagnostics.

395 For example, there are two artifacts in the estimated surfaces for *P. aerug-*
 396 *inosa* that would be missed without assessing the fit of the model to the data.
 397 First, sometimes the technician may inadvertently collect image data into the
 398 substratum where the biofilm is attached, as occurs for the bottom-most z -
 399 slices in Figure 3A. This can easily be remedied by excluding these z -slices
 400 from the analysis. Second, there are some surface model misfits, identified in
 401 Figure 3A by the high frequency in the black curve at the substratum where
 402 there is little biofilm. One approach to dealing with such misfit is to smooth
 403 the estimate of the biofilm’s surface. This possibility is addressed further in

404 section 4.

405 3.3. Comparisons with Imaris, COMSTAT2 and daime

406 In addition to the Bayesian and MLE approaches that we present, the
407 analysis of each of the three biofilm images in Figure 1 was attempted using
408 three other methods: Imaris, COMSTAT2 (in ImageJ) [27], and daime [8].
409 Table 1 compares the thresholds and volumes estimated by the different
410 methods.

411 Unfortunately, neither COMSTAT2 nor daime were able to analyze all 3
412 images. Only the *P. aeruginosa* image was able to be analyzed with COM-
413 STAT2, as the file sizes of the other two images were too large for the program
414 to handle. Even though the pixelation was less for the *Staph* data, the image
415 included here is part of a larger file containing a movie of the biofilm over
416 time [22]. Only the *Staph* GFP image could be analyzed with the daime
417 software, as the other two images contained a z-stack that was too thick for
418 the program to handle. Our Bayesian and likelihood approaches were able to
419 process all 3 images because, by design, they are computationally efficient.
420 For example, it took 40s to consider a single threshold for the *Pseudomonas*
421 data, about 1s to consider a single threshold for *Staph*, and 2m 45s to consider
422 a single threshold for *Janth* using a machine with an Intel (R) Core i7-6900K
423 CPU at 3.2 GHz. Therefore, our approach, Imaris and COMSTAT2 (in Im-
424 ageJ) can only be compared for the *P. aeruginosa* data; and our approach,
425 Imaris and daime can only be compared for the *Staph* data.

426 Based on the information provided in Table 1, the COMSTAT2 output
427 for the volume of the *P. aeruginosa* biofilm is similar to the Bayesian MAP
428 (although the thresholds are different). This is because COMSTAT2 cal-

Method	Software	Characteristic	Biofilm Image from Figure 1		
			<i>P. aeruginosa</i>	<i>S. aureus</i>	<i>Janth.</i>
Beer's MAP	MATLAB	Threshold	7	8	26
		Bio-volume	2.55×10^7	1.21×10^7	6.51×10^6
Beer's MLE	MATLAB	Threshold	7	6	24
		Bio-volume	2.55×10^7	1.38×10^7	6.55×10^6
Costes'	Imaris	Threshold	23	58	50
		Bio-volume	2.8×10^6	3.6×10^6	5.4×10^5
Otsu's	COMSTAT2	Threshold	3	NC	NC
		Bio-olume	2.32×10^7	NC	NC
RATS-L	daime	Threshold	NC	NA	NC
		Bio-volume	NC	6.66×10^5	NC

Table 1: Comparison of image analysis results for 5 different approaches applied to *P. aeruginosa*, *S. aureus*, and *Janthinobacterium* biofilms. Reported thresholds are intensity values ranging from 0 to 255; Biofilm bio-volumes reported as μm^3 . Approaches presented herein are listed as Beer's MAP and Beer's MLE. COMSTAT2 was implemented in ImageJ. NA indicates that daime did not provide the thresholding value. NC indicates quantities that could not be calculated due to restrictions on the file type and/or file size.

429 culates a “connected volume” which, similar to our approach, includes all
 430 pixels under the top of the biofilm in the volume calculation. The Imaris
 431 and daime analyses underestimate the volume compared to our Bayesian
 432 and MLE approaches. This is not surprising since Imaris [16, p. 235-8] and
 433 daime [8] calculate the volume by considering only bright pixels (above the
 434 image threshold). Unfortunately, it is challenging to determine specifically
 435 what the other software programs are doing in the absence of diagnostics of
 436 model fit as in Figures 2, 3 and 6.

437 A last crucial point is that the Bayesian MAP is the only approach to
 438 provide uncertainty quantification for the threshold. The MAP estimate
 439 for the Beer’s threshold for each microbial species is preferred over the other
 440 possible threshold values with probability close to 1 (Figure 4). Furthermore,
 441 the MAP and MLE approaches are the only two that provide uncertainty
 442 quantification for the bio-volume for each of the 3 species. The Bayesian
 443 MAP for the bio-volumes in Table 1 are associated with $SD(V|d)$ of $1.64 \times$
 444 10^4 , 2.35×10^4 and 1.61×10^3 (see section 2.3.4). The MLE bio-volumes
 445 in Table 1 had SDs of 1.64×10^4 , 2.42×10^4 and 1.57×10^3 . From these,
 446 because most of the probability mass for $\pi(\theta|d)$ is at $\hat{\theta}_{MAP}$ (Figure 4), then
 447 probability or confidence intervals can be calculated using the $t(df = n_x n_y -$
 448 $1)$ distribution by $\hat{V} \pm t_{df} SD$.

449 4. Discussion

450 We analyze spatial intensity data generated by CSLM. This technology
 451 is advancing. In addition to the intensity data that we focus on here, state-
 452 of-the-art Raman CSLMs can also provide high dimensional spectroscopy

453 information for chemical species at every point in the high dimensional spatial
454 lattice. Our regression approach is flexible enough to allow the user to add
455 covariates and/or factors to the regression equations to investigate spatial
456 relationships between the biofilm species and chemical species. Instead of
457 regression models, one could use ODE or PDE models of these relationships
458 (e.g., to model reaction and diffusion of substrates into the biofilm), but at
459 an increase in computational expense.

460 Previous Bayesian statistical analyses of CSLM images of thin layers of
461 human cells considered less severe attenuation effects [1]. [26] looked at much
462 thicker specimens of starfish oocytes and empirically assessed the attenuation
463 effects due to Beer’s Law using different CSLM objectives lenses (e.g., 25x
464 and 63x) and media (e.g., oil or water). [22] used an edge detector as we do
465 here to find a Bayesian estimate of the surface of the biofilm, but they used
466 a fixed threshold. We expand upon those results by considering a range of
467 user-supplied thresholds and selecting the one that optimizes the fit of Beer’s
468 Law to the data.

469 Our parabolic regression model for Beer’s Law explicitly accounts for one
470 or two bright regions just below the biofilm’s surface (Figure 2). It is not
471 clear what is occurring in these highly fluorescing regions of the image, but
472 three plausible yet speculative explanations are: higher cell densities; higher
473 levels of cell activity (especially when imaging GFP organisms); or a spatial
474 architecture of the biofilm that allows the fluorescent signal to be readily
475 collected.

476 The use of the edge detector is crucial. When developing our approach
477 we attempted fitting quadratic and quartic polynomial regression models

478 directly to each column of the data (i.e., without identification of the surface).
 479 Unfortunately, these approaches were adversely influenced by either the noisy
 480 low intensity pixels above the biofilm or the attenuating low intensity pixels
 481 in the depth of the biofilm. Our approach minimizes this limitation by using
 482 an edge detector to find the biofilm’s top surface then fitting a regression
 483 model that depends on the surface. Furthermore, our approach provides the
 484 optimal Beer’s threshold that gives the “best” fit model. The “best” Beer’s
 485 thresholds we report are the Bayesian maximum of the posterior (MAP) and
 486 the maximum likelihood estimate (MLE). In our *Staph* example, the MAP
 487 appeared to better fit the data, whereas for the *Janth* example there was
 488 little difference. In all of the examples, the MAP for the Beer’s threshold
 489 was preferred with probability close to 1 over any other threshold value.

490 In contrast to the other automatic threshold methods that we discuss, the
 491 threshold we find is based on the optimal fit of Beer’s Law to the data, not
 492 on a heuristic that requires tweaking the threshold until some biofilm char-
 493 acteristic like bio-volume “stabilizes.” Furthermore, our approach is compu-
 494 tationally efficient, taking between 1 second and 3 minutes to assess a single
 495 threshold on a single image in the examples presented, whereas some of the
 496 other methods presented could not process the large CSLM image data sets
 497 at all (Table 1).

498 A major benefit of our approach is that we are able to assess the fit of
 499 the model using diagnostic plots, analogous to checking residuals for any
 500 regression analysis [17]. Because such assessments are not available from
 501 other software packages, it is not clear to the user what other packages are
 502 doing, or how to assess what they are doing.

503 One approach to dealing with data that have detrimental influence on the
 504 fit of the regression model (e.g., identified by Figure 3) is to smooth the sur-
 505 face. Smoothing is a consequence of imposing a spatial correlation structure
 506 on the surface or on the intensities [15]. While implementing smoothing is
 507 easy (e.g., by convolution), we recommend: (1) Using a Bayesian statistical
 508 process to choose the amount of smoothing (i.e., depending on posterior prob-
 509 abilities) as opposed to a data analyst choosing; (2) generating some measure
 510 of uncertainty (e.g., probability intervals) for the smoothing parameter; and
 511 (3) incorporating the uncertainty in the smoothing into the uncertainty for
 512 all downstream quantities of interest. Not surprisingly, implementing these
 513 steps is computationally expensive. Because of the expense, unfortunately,
 514 it is common to let a data analyst choose the level of smoothing and to ig-
 515 nore the uncertainty in the smoothing term in biofilm CSLM image analysis
 516 [e.g., see 18]. Our approach is flexible enough to allow these steps to be
 517 implemented although at increased computational cost [22]. We investigate
 518 simultaneous threshold estimation with computationally efficient approaches
 519 to choosing the smoothing parameter (e.g., following [9] or [10]) elsewhere.

520 One drawback of our approach is that it does not model holes or overhang-
 521 ing features within the biofilm. That is, our computationally efficient model
 522 may be too simplistic for biofilms that have many holes and/or overhang-
 523 ing features. Independent analyses (i.e., cryosectioning, optical coherence
 524 tomography or inverted CSLM) suggest that in many cases, biofilms contain
 525 viable bacteria all the way through [22]. The diagnostic graphical tools dis-
 526 played in Figures 3, 5 and 6 allow the user to assess the appropriateness of
 527 the model presented here. We are currently developing a more computation-

ally expensive technique that implements a more complex statistical model
that can deal with these types of features.

Acknowledgements

We thank the Industrial Associates of the CBE for partially supporting this work. JAC is partially funded by CONACYT CB-2016-01-284451, RDECOMM and ONRG grants.

Data Availability

The data and MATLAB computer code are available on request from the authors.

5. Appendix

By Bayes Rule, the full posterior probability distribution for all of the parameters in our problem is

$$\pi(\beta, \sigma^2, \theta|d) = \frac{\pi(d|\beta, \sigma^2, \theta)\pi(\beta|\sigma^2, \theta)\pi(\sigma^2|\theta)\pi(\theta)}{\pi(d)}.$$

The first three factors in the numerator on the right hand side are proportional to $\pi(\beta, \sigma^2|\theta, d)$:

$$\pi(\beta, \sigma^2|\theta, d) = \frac{\pi(d|\beta, \sigma^2, \theta)\pi(\beta|\sigma^2, \theta)\pi(\sigma^2|\theta)}{\pi(d|\theta)}. \quad (12)$$

The specification for the model $d = F_\theta\beta_\theta + \varepsilon$ in section 2.3.1 implies that the likelihood $\pi(d|\beta, \sigma^2, \theta)$ is normal. The specification of the priors in section

2.3.3 implies that $\pi(\beta, \sigma^2 | \theta, d)$ is a normal-gamma distribution [3, 15]. It follows that integrating out σ^2 from (12) gives the multivariate t ,

$$\pi(\beta | \theta, d) = t(\hat{\beta}_{MLE} | \theta, Var(\hat{\beta}_{MLE} | \theta))$$

[12, 3]. This substantiates the claim that the first factor in equation (9) is a multivariate t . We derive the second factor $\pi(\theta | d)$ in (9) and (10) next.

Rearranging equation (12) we get that

$$\pi(d | \theta) = \frac{\pi(d | \beta, \sigma^2, \theta) \pi(\beta | \sigma^2, \theta) \pi(\sigma^2 | \theta)}{\pi(\beta, \sigma^2 | \theta, d)}$$

which is of the form

$$\pi(d | \theta) = \frac{(\text{normal likelihood}) \times \text{priors}}{(\text{normalgamma})}.$$

After cancellation of like terms [3] we get:

$$\pi(d | \theta) \propto \frac{\left((d^T d - \hat{\beta}_{MLE}^T F_\theta^T F_\theta \hat{\beta}_{MLE}) / 2 \right)^{-DFE/2}}{\det(F_\theta^T F_\theta)^{1/2}}$$

where $DFE = n_x n_y n_z - 6 n_x n_y$. Similar calculations to derive $\pi(d | \theta)$ were performed by [10] although they assumed normal instead of uniform priors on the $\{\beta_i\}$. If we let $\hat{d} = F_\theta \hat{\beta}_{MLE}$, then the numerator can be simplified by noting that $DFE \cdot MSE_\theta = (d - \hat{d})^T (d - \hat{d}) = d^T d - 2d^T \hat{d} + \hat{d}^T \hat{d}$. Because $d^T \hat{d} = \sqrt{d^T d} \cdot \sqrt{\hat{d}^T \hat{d}} \cos(\angle(d, \hat{d}))$ and $\cos(\angle(d, \hat{d})) = \sqrt{\hat{d}^T \hat{d}} / \sqrt{d^T d}$ it follows that $DFE \cdot MSE_\theta = d^T d - \hat{d}^T \hat{d}$. So

$$\pi(d | \theta) \propto \sqrt{\frac{(DFE \cdot MSE_\theta / 2)^{-DFE}}{\det(F_\theta^T F_\theta)}}. \quad (13)$$

549 The determinant calculation $\det(F_\theta^T F_\theta)$ needed to evaluate (13) is straight-
 550 forward even for our high dimensional problem because F_θ is composed of
 551 $n_x n_y$ blocks (one for each column), each of size $n_z \times 6$. Hence $\det(F_\theta^T F_\theta)$ is
 552 the product of the determinant of $n_x n_y$ blocks, each of size 6×6 .

By Bayes Rule, the marginal posterior probability distribution for the Beer's threshold is

$$\pi(\theta|d) = \frac{\pi(d|\theta)\pi(\theta)}{\sum_{\theta'} \pi(d|\theta')\pi(\theta')}$$

where $\pi(d|\theta)$ is calculated via (13). This equation shows how to incorporate prior information for θ for any prior $\pi(\theta)$. Under the assumption of a flat prior over the possible integer values for θ (as in section 2.3.3), then

$$\pi(\theta|d) = \frac{\pi(d|\theta)}{\sum_{\theta'} \pi(d|\theta')}$$

553 where $\pi(d|\theta)$ is calculated via (13). This substantiates equation (10).

To estimate bio-volume (V) based on the surface, we consider the following model of the biofilm's surface

$$z_{surf}(x, y|\theta) = \tau(x, y|\theta) + \epsilon$$

where τ is the true thickness of the biofilm and $\epsilon \sim N(0, \sigma_{surf}^2 I)$. Because z_{surf} depends on the Beer's threshold θ (through the edge detector and the regression equations (2) and (3) describing Beer's Law), then, similar to the calculations leading to (11),

$$\pi(V|d) = \sum_i t(\hat{V}_{MAP}|\theta_i, n_x n_y \hat{\sigma}_{surf}^2|\theta_i, df = n_x n_y - 1) \times \pi(\theta_i|d)$$

554 where the bio-volume for a single Beer's threshold θ_i is estimated by $\hat{V}_{MAP}|\theta_i =$
 555 $\sum_{x,y} z_{surf}(x, y|\theta_i) \Delta_{xyz};$

$$\hat{\sigma}_{surf}^2|\theta_i = (z_{surf} - \frac{\hat{V}_{MAP}|\theta_i}{n_x n_y})^T (z_{surf} - \frac{\hat{V}_{MAP}|\theta_i}{n_x n_y}) / (n_x n_y - 1); \quad (14)$$

556 and $\pi(\theta_i|d)$ is calculated by (10). This substantiates the equations for $\text{Mean}(V|d)$
557 and $\text{SD}(V|d)$ in section 2.3.4.

558 References

- 559 [1] Al-Awadhi, F., Hurn, M., and Jennison, C. (2011). Three-dimensional
560 Bayesian analysis and confocal microscopy. *Journal of Applied Statistics*,
561 38(1):29–46.
- 562 [2] Bernardo, J. and Smith, A. (1994). *Bayesian Theory*. John Wiley and
563 Sons.
- 564 [3] Broemeling, L. D. (1985). *Bayesian Analysis of Linear Models*. Marcel-
565 Dekker.
- 566 [4] Bruns, T., Schickinger, S., Wittig, R., and Schneckenburger, H. (2012).
567 Preparation strategy and illumination of three-dimensional cell cultures
568 in light sheet-based fluorescence microscopy. *Journal of biomedical optics*,
569 17(10).
- 570 [5] Calvetti, D. and Somersalo, E. (2007). *Introduction to Bayesian Scientific*
571 *Computing*. Springer.
- 572 [6] Castlemen, K. R. (2002). Multiple color fluorescence imaging in mi-
573 croscopy. In Diaspro, A., editor, *Confocal and Two-Photon Microscopy*,
574 pages 285–299. Wiley, New York.
- 575 [7] Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G., and
576 Lockett, S. (2004). Automatic and quantitative measurement of protein-
577 protein colocalization in live cells. *Biophysical Journal*, 86:3993–4003.

- 578 [8] Daims, H., Lucker, S., and Wagner, M. (2006). daime, a novel image anal-
579 ysis program for microbial ecology and biofilm research. *Environmental*
580 *Microbiology*, 8:200–213.
- 581 [9] Feron, O., Orieux, F., and Giovannelli, J. (2016). Gradient scan Gibbs
582 sampler: an efficient algorithm for high dimensional Gaussian distribu-
583 tions. *IEEE Journal Of Selected Topics In Signal Processing*, 10(2):343–
584 352.
- 585 [10] Fox, C. and Norton, R. (2016). Fast sampling in a linear-Gaussian in-
586 verse problem. *SIAM/ASA Journal of Uncertainty Quantification*, 4:1191–
587 1218.
- 588 [11] Franklin, M. J., Chang, C., Akiyama, T., and Bothner, B. (2015). New
589 technologies for studying biofilms. *Microbiology Spectrum*, 3.
- 590 [12] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995).
591 *Bayesian Data Analysis*. Chapman & Hall.
- 592 [13] Goeres, D., Loetterle, L., Hamilton, M., Murga, R., Kirby, D., and Don-
593 lan, R. (2005). Statistical assessment of a laboratory method for growing
594 biofilms. *Microbiology*, 151:757–762.
- 595 [14] Heydorn, A., Nielsen, A. T., Hentzer, M., Sternberg, C., Givskov, M.,
596 Ersboll, B. K., and Molin, S. (2000). Quantification of biofilm structures
597 by the novel computer program COMSTAT. *Microbiology*, 146:2395–2407.
- 598 [15] Higdon, D. (2006). A primer on space-time modelling from a Bayesian
599 perspective. In Finkenstadt, B., Held, L., and Isham, V., editors, *Statis-*

- 600 tics of *Spatio-Temporal Systems*, pages 217–279, New York. Chapman &
601 Hall/CRC.
- 602 [16] Imaris (2018). Imaris 9.2 reference manual. Bitplane.
- 603 [17] Kutner, M., Hachtsheim, C., Neter, J., and Li, W. (2005). *Applied*
604 *Linear Statistical Models*. McGraw-Hill/Irwin, Boston.
- 605 [18] Luo, T. L., Eisenberg, M. C., Hayashi, M. A. L., Gonzalez-Cabezas,
606 C., Foxman, B., Marrs, C. F., and Rickard, A. H. (2018). A sensitive
607 thresholding method for confocal laser scanning microscope image stacks
608 of microbial biofilms. *Scientific Reports*, (13013).
- 609 [19] Otsu, N. (1979). A threshold selection method from gray-level his-
610 tograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–
611 66.
- 612 [20] Pampaloni, F., Ansari, N., and Stelzer, E. H. (2013). High-resolution
613 deep imaging of live cellular spheroids with light-sheet-based fluorescence
614 microscopy. *Cell and tissue research*, 352(1):161–177.
- 615 [21] Pang, Y. Y., Schwartz, J., Thoendel, M., Ackermann, L. W., Horswill,
616 A. R., and Nauseef, W. M. (2010). agr-dependent interactions of *Staphy-*
617 *lococcus aureus* USA300 with human polymorphonuclear neutrophils. *J.*
618 *Innate Immun.*, 2:546–559.
- 619 [22] Parker, A., Pitts, B., Lorenz, L., and Stewart, P. (2018). Polynomial
620 accelerated solutions to a LARGE Gaussian model for imaging biofilms: in
621 theory and finite precision. *Journal of the American Statistical Association*.

- 622 [23] Pitts, B. and Stewart, P. (2008). Confocal laser microscopy on biofilms:
623 Successes and limitations. *Microscopy Today*, 16(4):18–21.
- 624 [24] Smith, H. J., Foreman, C. M., Akiyama, T., Franklin, M. J., Devitt,
625 N. P., and Ramaraj, T. (2016). Genome sequence of *Janthinobacterium*
626 sp. CG23, a violacein-producing isolate from an Antarctic supraglacial
627 stream. *Genome Announc.*, 4(1):e01468–15.
- 628 [25] Stewart, P. S., Davison, W. M., and Steenbergen, J. N. (2009). Dap-
629 tomycin rapidly penetrates a *Staphylococcus epidermidis* biofilm. *Antimi-
630icrobial Agents and Chemotherapy*, 53:3505–3507.
- 631 [26] Terasaki, M. (2006). Quantification of fluorescence in thick specimens,
632 with an application to cyclin B-GFP expression in starfish oocytes. *Biol.
633 Cell*, 98:245–252.
- 634 [27] Vorregaard, M. (2008). Comstat2 - a modern 3D image analysis environ-
635 ment for biofilms, in informatics and mathematical modelling. Technical
636 University of Denmark, Kongens Lyngby, Denmark.
- 637 [28] Xavier, J., Schnell, A., Wuertz, S., Palmer, R., White, D., and Almeida,
638 J. (2001). Objective threshold selection procedure (OTS) for segmentation
639 of scanning laser confocal microscope images. *Journal of Microbiological
640 Methods*, 47(2):169 – 180.
- 641 [29] Xavier, J., White, D., and Almeida, J. (2003). Automated biofilm mor-
642 phology quantification from confocal laser scanning microscopy imaging.
643 *Water Science and Technology*, 47(5):31–37.