A STEREOLOGICAL UNFOLDING METHOD FOR THE STUDY OF THE MITOCHONDRIAL NETWORK

Daniel Christopher Flynn[1], James Darwin Nulton[2], Peter Salamon[2], Terrence Gene Frey[3], Avinoam Rabinovitch[4], Mei Guo Sun[3] and Arlette Rose Cornelia Baljon[1]

[1]Department of Physics, San Diego State University, San Diego, California, USA; [2]Department of Mathematical Sciences, San Diego State University; [3]Department of Biology, San Diego State University; [4]Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva, Israel
e-mail: mdacflynn@umich.edu, jnulton@mail.sdsu.edu, salamon@math.sdsu.edu, tfrey@sciences.sdsu.edu, sun.mei@gene.com, abaljon@mail.sdsu.edu
(Accepted March 6, 2009)

ABSTRACT

When the mitochondrial network within the cell is in a fragmented state, it resembles a spatial distribution of prolate spheroids of various shapes, sizes and orientations. This paper presents a maximum-likelihood scheme for inferring the distribution of spheroidal shapes, sizes and orientations from the observed distribution of elliptical sections. We also present a parametric bootstrap suitable for inferring confidence intervals for the parameters describing the shape and size distribution of the spheroids.

Keywords: bootstrap, maximum likelihood, mitochondria, spheroid, statistics, unfolding problem.

INTRODUCTION

The motivation for this paper comes from the study of conformational changes in mitochondria associated with cell death. During a late stage of apoptosis (programmed cell death) the cell's mitochondrial network fragments into individual mitochondria. Cross-sections of these particles observed in electron micrographs suggest that they have an oblong shape. The particles originate in a process that is reminiscent of the way sausage is made, in that the fragments apparently result from the pinching off (fissioning) the cylindrical tubes that make up the mitochondrial network. Since the tubes are cylindrical, we expect the fragments, being sacks of fluid, to adjust to a roughly spheroid shape. In our study we approximate these particles by prolate spheroids. Low resolution visual evidence from three-dimensional confocal microscopy (Frank *et al.*, 2001) argues against the presence of oblate structures. Moreover, in section "A sample application to mitochondria" we show that the sectioned data itself argues similarly.

These particles are believed to undergo changes over time in shape and size due to disruption of inner membrane function and possible resultant swelling. While confocal light microscopy is able to partially document this process, the resolution is limited to values comparable to the particle size. Quantitative information on size and shape is obtained from electron micrographs, which usually show only 2-dimensional cross-sections of the particles.

Nevertheless, assuming the particles are prolate spheroids, it is possible to infer 3-dimensional information from the statistics of the 2-dimensional sections.

It is this stereological problem that we propose to address here. Wicksell (1925; 1926) dubbed this "the corpuscle problem." He was motivated by the study of secondary follicles embedded in the lymphatic tissue of the spleen. More recently the problem has become known as the "unfolding problem." An excellent history and set of references can be found in Stoyan *et al.* (1987). A complete treatment of stereological methods with an emphasis on practical problems can be found in the two volumes of Weibel (1979; 1980). Most recently, the general problem, involving 3D size, shape and orientation, was treated by Sato *et al.* (1996).

The basic mathematical problem is to relate the distributions of observable quantities in 2D cross-sections to the distributions of desired quantities that are descriptive of the 3D particles. Typically this takes the form of an integral transform expressing, for example, the distribution Ψ of cross-sectional diameters in terms of the distribution Φ of particle diameters. The practical problem is then to use this machinery to work from data in the form of a sample from Ψ to a description of Φ. Two approaches were proposed in Wicksell's original papers and have been taken up by subsequent authors (De Hoff, 1962; Cruz-orive, 1976; Franklin, 1977; Sato *et al.*, 1996). First, the integral transform can be used to derive a recursive

scheme that expresses the moments of Φ in terms of those of Ψ. Secondly, in some cases, it is possible to invert analytically the transform and to solve for Φ in terms of Ψ and its derivatives. Both of these approaches suffer when the data set is not large enough to provide a clear picture of Ψ. In the first approach, statistical moments from the data are only estimates of Ψ's true moments, and the associated statistical error is propagated through the transform. Similarly, in the second approach, estimates of the derivatives of Ψ by finite difference methods introduce uncertainties that are difficult to quantify.

Four distinct additional approaches have been identified (Blödner *et al.*, 1984); among these is the parametric method that we propose to apply. The parametric approach has been used to predict a very precise distribution of sphere diameters from circular cross-sections (Keiding and Jensen, 1972). In the present paper we describe a parametric approach, which does not seem to have been used by previous authors in the context of the full unfolding problem described by Sato, involving simultaneous predictions of the distributions of size, shape and orientation. We propose parametric models for desired distributions for Φ, and use a maximum-likelihood approach to optimize the parameters. The *method of parametric bootstrap* (Efron and Tibshirani, 1993) was used to determine confidence regions for the parameters.

In the next section we set out the geometrical framework, define the quantities studied, and state the relations among them. In the section following that we apply the method of Sato to derive the machinery expressing the distributions of cross-sectional size and shape in terms of the hypothetical or desired distributions of particle size, shape, and orientation. The three sections beginning with the Maximum-likelihood Scheme present a description of our method. We describe the parametric models used for the hypothetical distributions and the maximum-likelihood scheme used to determine values for the parameters. We also describe how to set up a parametric bootstrap for determining confidence regions for the values of the parameters. The last section presents the results of applying the method to a set of data derived from electron micrographs.

GEOMETRICAL PRELIMINARIES

In this section we summarize the geometrical relations among the slicing plane, the spheroid, and the elliptical cross-section. See Fig. 1.
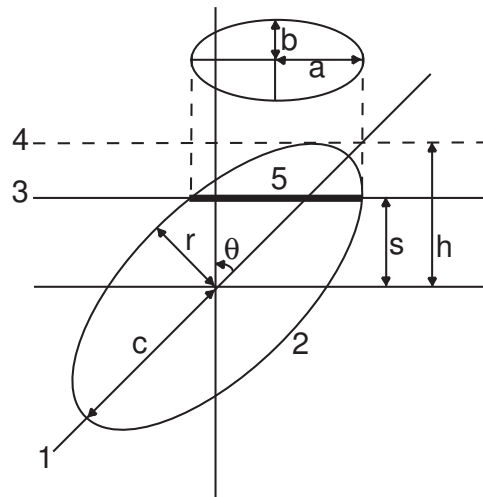

Fig. 1. *Drawing of sliced spheroid: 1) Axis of spheroid. 2) Prolate spheroid. 3) Slicing plane. 4) Tangent plane. 5) Elliptical cross section.*

We take the slicing plane to be horizontal, a distance s from the center of the spheroid. We take θ to be the angle between the axis of the spheroid and the vertical. The spheroid is described by its semi-minor axis r, its semi-major axis c, and its eccentricity e_0. Similarly, the elliptical section, is described by axes b and a, and eccentricity e. The axes and eccentricities are related by $e_0^2 = 1 - r^2/c^2$ and $e^2 = 1 - b^2/a^2$. The figure shows a cross-section through the spheroid's axis. The following equations are easily established and summarize the relationships among the variables required for the analysis in the next section.

$$e = e_0 \sin \theta \, , \tag{1}$$

$$b = r\sqrt{1 - \frac{s^2}{h^2}} \, , \tag{2}$$

$$h = r\frac{\sqrt{1 - e_0^2 \sin^2 \theta}}{\sqrt{1 - e_0^2}} \, . \tag{3}$$

Geometrically θ can be considered to vary from 0 to π, and s, from $-\infty$ to $+\infty$[1]. Nevertheless, the distribution of θ is understood to be symmetric about the value $\pi/2$. Similarly, s is distributed symmetrically about 0. These symmetry considerations allow us, in the following analysis, to restrict θ to $[0, \pi/2]$ and s to $[0, \infty)$.

RELATIONS AMONG THE DISTRIBUTIONS

Our goal is to extract information on the distributions of e_0, r, and θ from observed distributions

---

[1]Of course an elliptical slice is obtained only when $|s| < h$.

of e and b. To that end we propose a model for the joint density function $\Omega(b,e,s,\theta,r,e_0)$ from which a marginal distribution

$$\Psi(b,e) = \int_0^1 \int_0^\infty \int_0^{\pi/2} \int_0^\infty \Omega(b,e,s,\theta,r,e_0)\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}e_0 \ . \quad (4)$$

Apart from the geometrical constraints given in Eqs. 1-3, the principal element of our statistical model is the assumption that the variables e_0, r, and θ are distributed independently. This allows us to factor Ω as follows:

$$\Phi_0(\theta)\,\Phi_1(r)\,\Phi_2(e_0)\,\Sigma(s\,|\,\theta,r,e_0)\,\Lambda(b,e\,|\,s,\theta,r,e_0)\ , \quad (5)$$

where Σ and Λ are conditional densities. We will eventually propose simple parametric models for Φ_0, Φ_1, and Φ_2. See Eqs. 18-21.

Some remarks are in order to justify the assumptions of independence. First, there are a priori reasons for believing that r and e_0 are independent. The mitochondrial fragments arise as a result of the fissioning of long tubular networks of roughly constant width or diameter. Consequently, the resulting fragments would be expected to have roughly the same minor axis, or, equivalently, that the distribution of the minor axis would be relatively sharply peaked. Any correlation that might exist between this size parameter and shape would be very small and difficult to tease out of the data by the indirect methods we are proposing. Secondly, a consequence of the assumption of independence of r and e_0 is Eq. 14 below, which says that b and e should also be uncorrelated. This was found to be true of the data before any substantive analysis was attempted. It seemed likely that if our data showed no discernable correlation between b and e, then, even if there were a slight correlation between r and e_0, our maximum-likelihood scheme applied to our data was not likely to reveal it.

It is difficult to imagine what mechanism would correlate orientation with size or shape. In any case, the assumption of independence here is made, frankly, as a matter of convenience.

Returning now to the analysis of Eq. 4, it is convenient to separate out the integration over s and θ to obtain

$$\Psi(b,e) = \int_0^1 \int_0^\infty I(b,e,r,e_0)\,\Phi_1(r)\,\Phi_2(e_0)\,\mathrm{d}r\,\mathrm{d}e_0\ , \quad (6)$$

where

$$I(b,e,r,e_0) = \int_0^{\pi/2} \int_0^\infty \Phi_0(\theta)\,\Sigma(s\,|\,\theta,r,e_0)\,\Lambda(b,e\,|\,s,\theta,r,e_0)\mathrm{d}s\,\mathrm{d}\theta\ . \quad (7)$$

The integral I can now be simplified. Geometrical considerations will dictate the form of the conditional densities Σ and Λ. Consider Σ first. We consider equally likely all positions s that actually result in an intersection, so we may take

$$\Sigma(s\,|\,\theta,r,e_0) \propto \begin{cases} 1\,, & 0 \le s \le h(\theta,r,e_0)\,, \\ 0\,, & \text{else.} \end{cases} \quad (8)$$

Next consider Λ. Once s, e_0, r, and θ are specified, Eqs. 1-3 show that the values of e and b are determined, so we may write, using Dirac's delta function

$$\Lambda(b,e\,|\,s,\theta,r,e_0) = \delta(e - e_0\sin(\theta)) \cdot \delta\left(b - r\sqrt{1 - \frac{s^2}{h^2}}\right)\ . \quad (9)$$

Incorporating Eqs. 8 and 9 into Eq. 7, we have

$$I(b,e,r,e_0) = \int_0^{\pi/2} \Phi_0(\theta)\delta(e - e_0\sin(\theta)) \times \int_0^h \delta\left(b - r\sqrt{1 - \frac{s^2}{h^2}}\right)\mathrm{d}s\,\mathrm{d}\theta\ . \quad (10)$$

The integration over s can be carried out explicitly. For $0 < b < r$ the result is

$$\int_0^h \delta\left(b - r\sqrt{1 - \frac{s^2}{h^2}}\right)\mathrm{d}s = \frac{b\,h(\theta,r,e_0)}{r\sqrt{r^2 - b^2}}$$
$$= \frac{b\sqrt{1 - e_0^2\sin^2(\theta)}}{\sqrt{r^2 - b^2}\sqrt{1 - e_0^2}}\ . \quad (11)$$

For $r \le b$ the integral is zero. The last equality uses Eq. 3. Eq. 10 becomes

$$I(b,e,r,e_0) = \frac{b}{\sqrt{r^2 - b^2}}\int_0^{\pi/2}\Phi_0(\theta)\delta(e - e_0\sin(\theta)) \times \frac{\sqrt{1 - e_0^2\sin^2(\theta)}}{\sqrt{1 - e_0^2}}\mathrm{d}\theta\ , \quad (12)$$

for $0 < b < r$, and zero for $r \leq b$. Finally, the integration over θ can be carried out, producing, for $r > b$ and $e_0 > e$,

$$I(b,e,r,e_0) = \frac{b\sqrt{1-e^2}\,\Phi_0(\sin^{-1}(e/e_0))}{\sqrt{r^2-b^2}\sqrt{1-e_0^2}\sqrt{e_0^2-e^2}}\,, \qquad (13)$$

and the value zero for all other values of r and e_0.

This result, together with the limits on the variables r and e_0, can be substituted into Eq. 6 and the resulting double integral factored to give

$$\Psi(b,e) \propto \Psi_1(b) \cdot \Psi_2(e)\,, \qquad (14)$$

where

$$\Psi_1(b) = \int_b^\infty \frac{b\,\Phi_1(r)}{\sqrt{r^2-b^2}}\mathrm{d}r\,, \qquad (15)$$

and

$$\Psi_2(e) = \int_e^1 \frac{\sqrt{1-e^2}\,\Phi_0(\sin^{-1}(e/e_0))\,\Phi_2(e_0)}{\sqrt{1-e_0^2}\,\sqrt{e_0^2-e^2}}\,\mathrm{d}e_0\,. \qquad (16)$$

We caution the reader that Ψ_1 and Ψ_2 are not probability densities, because they have not been normalized. Normalization is not required to set up the likelihood functions described in the next section. However, as we will see later, these functions are the basic tools for simulating a slicing experiment. In that case they must be normalized before they can play that role.

THE MAXIMUM-LIKELIHOOD SCHEME AND PARAMETRIC MODELS

We have yet to assign statistical models for the densities $\Phi_0(\theta)$, $\Phi_1(r)$, and $\Phi_2(e_0)$. Once that is done the parameters in those models become the basis for varying the hypothetical distribution of spheroids so as to optimally fit the 2D data. In this section we first formulate the likelihood of a dataset given the hypothetical set of parameter-dependent Φ's. Maximizing the likelihood optimizes the fit and completes the prediction. We then turn to a description of and rationale for the statistical models recommended for the Φ's. These were chosen for the application described in the last section.

Suppose we have collected data in the form of m elliptical cross-sections. For each section we have

recorded the semi-minor axis b and the eccentricity e: $(b_1,e_1),(b_2,e_2),\ldots,(b_m,e_m)$. The data is separated into $D_1 = \{b_1,b_2,\ldots,b_m\}$ and $D_2 = \{e_1,e_2,\ldots,e_m\}$[2]. D_1 will be used to infer $\Phi_1(r)$ through Eq. 15, and D_2 will be used separately to infer $\Phi_0(\theta)$ and $\Phi_2(e_0)$ through Eq. 16.

If the distributions Φ_0, Φ_1, and Φ_2 in Eqs. 15 and 16 were a true description of the population of spheroids, then D_1 would be a collection of m independent choices from the distribution $\Psi_1(b)$, and, similarly, D_2, a collection from $\Psi_2(e)$. Therefore, we can formulate the likelihood of each data set as a product:

$$L_1(D_1) = \prod_{j=1}^m \Psi_1(b_j), \qquad L_2(D_2) = \prod_{j=1}^m \Psi_2(e_j). \qquad (17)$$

Maximizing these likelihoods over the space of parameters on which the densities $\Phi_0(\theta)$, $\Phi_1(r)$, and $\Phi_2(e_0)$ depend determines values of those parameters that optimally fit the observed data. This completes the algorithm by which predictions for the distribution of spheroids is made. For example, Φ_1 will be modelled by a normal distribution with the usual parameters μ and σ. Then, for the given data D_1, L_1 becomes a function of μ and σ. The values μ_* and σ_* that maximize L_1 then describe the most likely normal model for Φ_1 that is consistent with the data D_1. A similar process with L_2 is used to obtain predictions of $\Phi_0(\theta)$ and $\Phi_2(e_0)$ from the data D_2.

We now describe and discuss the parameter-dependent statistical models recommended for the Φ's. In each case the simplest appropriate model was chosen. We do not regard these models as strictly valid, or as having been justified by some independent experiment. Rather, they should be regarded as probes of the first- and second-order features of the statistics we are trying to study.

The statistical model suitable for probing the first two moments of an unknown statistics is the normal or Gaussian density; we find that appropriate for $\Phi_1(r)$. The usual 2-parameter choice for modelling the statistics of a naturally bounded variable is the beta distribution or density; we find that appropriate for both $\Phi_0(\theta)$ and $\Phi_2(e_0)$, since theta is confined to the interval $[0,\pi]$ and e_0 is confined to the interval $[0, 1]$. Φ_0 requires further comments, given below. There is a symmetry that sets the mean of θ at $\pi/2$. As a result Φ_0 becomes a 1-parameter density. The remaining parameter characterizes the spread, which, as explained below, can be interpreted as a measure of the deviation from isotropic orientation.

---

[2]As we have remarked, for the data we collected, there was no discernable correlation between b and e. That result is consistent with equation (14), and, thus, with our assumption that r and e_0 for the spheroids are independent. Hence, the separation of the data.

Concretely, the models that we propose are as follows. Φ_1 is modeled by a normal density.

$$\Phi_1(r|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\frac{(r-\mu)^2}{\sigma^2}\right). \quad (18)$$

Φ_2 is modelled by a beta density,

$$\Phi_2(e_0|p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}e_0^{p-1}(1-e_0)^{q-1}, \quad (19)$$

for which the mean and variance are given by

$$\mu = \frac{p}{p+q}, \qquad \sigma^2 = \frac{pq}{(p+q)^2(p+q+1)}. \quad (20)$$

For the density Φ_0, describing the orientations of the spheroids, we use a simple 1-parameter beta density scaled to the interval $[0,\pi]$:

$$\Phi_0(\theta|\alpha) = \frac{2\Gamma(2\alpha+2)}{\Gamma(\alpha+1)^2\pi^{2\alpha+1}}\theta^\alpha(\pi-\theta)^\alpha. \quad (21)$$

The case of θ, the angle between the spheroid's major axis and the vertical, requires more discussion because of the special role it plays in the type of data we are studying. First of all, the mitochondrial spheroids are presumably floating freely in the cytoplasm, so why assume anything other than a purely isotropic orientation for their major axes? The micrographs are horizontal sections through a cell that is adsorbed to a horizontal plane surface. Therefore we have no reason to expect anisotropy in the horizontal plane. However, there is reason to expect anisotropy in the vertical orientation. The basis of this expectation is related to the origin of the spheroids. The spheroids result from the fragmentation of an extensive tubular network as explained in the introduction. Since the cell is strongly attached to a horizontal surface, there is a flattening of the cell from its normal state. Confocal microscopy, our own as well as that of Frank *et al.* (2001), demonstrate the flattening and indicate that the tubes in the network show some preference for the horizontal. As a result the tubes in the network may show some preference for the horizontal. When the tubes fission, producing the spheroids, a trace of that preference may linger in the orientation of the spheroids. For that reason we structure our model so that it has the potential to detect and quantify any anisotropy that might exist.

The ability to quantify this anisotropy is not the most important reason for including the distribution of spheroid orientation in our scheme. If anisotropy is present and we do not account for it in our likelihood machinery, our results on the distribution of spheroid shape will be distorted. Eq. 1 reveals the intimate connection between the shape of the elliptical section and the shape and *orientation* of the spheroid. Eq. 16 shows how their distributions are intertwined in the unfolding process from 2D to 3D.

We now explain our choice of Φ_0 and how we propose to quantify the anisotropy. For a purely isotropic distribution of points on a sphere, the colatitudes (corresponding to the angle θ) of the points have the density function $\Phi(\theta) = \frac{1}{2}\sin\theta$ on $[0,\pi]$. This density has an easily calculated variance of $(\pi^2-8)/4$. A variance smaller than this is indicative of something called a "girdle" or "equatorial" distribution. As presented in the literature (Selby, 1964; Watson, 1965), it is a distribution of points with a higher concentration near the equator, and one that is symmetric about the equator. This is the situation we find when considering the orientation of spheroid particles with a slight preference for the horizontal. In our case, the symmetry is a built-in artifice, i.e., a spheroid whose axis is not perfectly horizontal is represented both by the angle θ and $\pi-\theta$.

We expect our model Φ_0 to quantify the anisotropy in the following way. Let σ^2 be the variance of the particular Φ_0 determined by our maximum-likelihood algorithm. We define an index of isotropy:

$$I_{iso} = \frac{\sigma}{\sigma_{iso}} = \frac{2\sigma}{\sqrt{\pi^2-8}}. \quad (22)$$

This has a ready interpretation as the ratio of the observed spread in the polar angle to the isotropic spread. It has a value ranging from 0 to 1 for equatorial distributions.

All one needs in order to capture this anisotropy is a fairly all-purpose, 1-parameter, symmetric density. As reported above our choice was a beta distribution. Other 1-parameter models for equatorial distributions have been proposed, and would probably serve equally well as models for Φ_0. For example, one prominent choice among these is the Dimroth-Watson model (Mardia, 1972; Batschelet, 1981):

$$\Phi(\theta) = Ce^{-\kappa\cos^2\theta}\sin\theta, \quad (23)$$

where C is a suitable normalization constant For $\kappa \geq 0$ it represents equatorial distributions. Here the parameter κ can be regarded as a measure of anisotropy. However, except for $\kappa = 0$, which corresponds to isotropic, its numerical value has no compelling interpretation. One could claim that this model has the advantage that the isotropic case is included in the family. However, that is not a practical advantage over Eq. 21. Remarkably, for $\alpha = 1.16072$, Eq. 21 is virtually indistinguishable from $\sin\theta$, uniformly over the interval $[0,\pi]$, and, hence includes the isotropic case as well.

THE PREDICTOR AND THE SIMULATOR

In the first three paragraphs of the previous section we have described an algorithm for which the *input* is data $(b_1, e_1), \ldots, (b_m, e_m)$ in the form of observed elliptical sizes and shapes, and for which the *output* is a set of predicted values $(\mu_*, \sigma_*, p_*, q_*, \alpha_*)$ for the parameters that describe the most likely population of spheroids to have produced the input data in a slicing experiment. Let us call this algorithm the PREDICTOR. As a result of the factorization in Eq. 14, there are actually two algorithms: PREDICTOR$_1$ that predicts the parameters (μ_*, σ_*) for the size distribution of prolate spheroids from the data $D_1 = (b_1, \ldots, b_m)$, and PREDICTOR$_2$ that predicts the parameters (p_*, q_*, α_*) for the shape and orientation distributions from the data $D_2 = (e_1, \ldots, e_m)$.

Note that the statistical role of the PREDICTOR is to provide a most likely *point estimate* for the unknown parameters consistent with a given set of data. A complete statistical analysis requires, in addition, some measure of the uncertainty of the method in the form of a region, containing the point estimate, within which the true values of the parameters are believed to fall with some specified level of confidence. For that purpose it will be important to define a second algorithm, the SIMULATOR, which uses the same apparatus as the PREDICTOR, but which, in a sense, reverses the process.

The geometry and statistics of the ideal slicing experiment has already been formulated in the mathematical machinery of the two previous sections. Suppose that we know (or hypothesize) that the spheroid population is described by the parameter values $(\mu_0, \sigma_0, p_0, q_0, \alpha_0)$[3]. These values specify the distributions $\Phi_0(\theta)$, $\Phi_1(r)$, and $\Phi_2(e_0)$ by way of Eqs. 18-21. Eqs. 15 and 16 then uniquely determine the distributions $\Psi_1(b)$ and $\Psi_2(e)$, which describe how we expect any data obtained from the hypothesized population to be distributed. As remarked earlier, Ψ_1 and Ψ_2 are presented in non-normalized form, which makes them suitable for a likelihood function, but not suitable as probability densities. Let $\hat{\Psi}_1(b)$ and $\hat{\Psi}_2(e)$ represent these distributions after they have been normalized so that their integrals are unity. These densities are the tools by which the ideal slicing experiment can be simulated. A simulation producing m elliptical sections is effected by choosing numbers $(b_1, \ldots, b_m)$ according to the distribution $\hat{\Psi}_1(b)$, and numbers $(e_1, \ldots, e_m)$ from the distribution $\hat{\Psi}_2(e)$. Thus, the SIMULATOR is the algorithm whose input is a set of hypothetical values $(\mu_0, \sigma_0, p_0, q_0, \alpha_0)$ and whose output is a set of data $(b_1, \ldots, b_m)$ and $(e_1, \ldots, e_m)$ representing the result of an ideal slicing experiment with the population of spheroids whose distribution is described by the parameters $(\mu_0, \sigma_0, p_0, q_0, \alpha_0)$.

To see how the SIMULATOR bears on the problems of reliability and the accuracy of our method, consider the following. Suppose the SIMULATOR is applied to the hypothetical values $(\mu_0, \sigma_0, p_0, q_0, \alpha_0)$ N times (a large number) in a Monte Carlo scheme. Each of the N simulated data sets can be fed to the PREDICTOR to produce N different predictions $(\mu_1, \sigma_1, p_1, q_1, \alpha_1), \ldots, (\mu_N, \sigma_N, p_N, q_N, \alpha_N)$. If our method is *highly accurate* we would expect the points $(\mu_1, \sigma_1, p_1, q_1, \alpha_1), \ldots, (\mu_N, \sigma_N, p_N, q_N, \alpha_N)$ to be closely clustered about the point $(\mu_0, \sigma_0, p_0, q_0, \alpha_0)$ representing the "true" distribution. Indeed, the uncertainty of our method will be measured by quantifying this clustering.

We will address this uncertainty in the next section, but first, let us point out that, just as the PREDICTOR splits into two algorithms, the same is true of the SIMULATOR. That is, SIMULATOR$_1$ takes as input hypothetical values (μ_0, σ_0) for the size distribution of spheroids, and, in a Monte Carlo scheme, produces a large number N of new sets of data, which in turn produce N predictions $(\mu_1, \sigma_1), \ldots, (\mu_N, \sigma_N)$. This can be pictured as a two-dimensional scattergram of points whose center of mass should closely approximate the point (μ_0, σ_0). Fig. 8 shows that scattergram for the example in the last section. Similarly, SIMULATOR$_2$ in a Monte Carlo scheme produces a three-dimensional scatter of points $(p_1, q_1, \alpha_1), \ldots, (p_N, q_N, \alpha_N)$ from the hypothetical description (p_0, q_0, α_0) of the distribution of shapes and orientations of the spheroids.

QUANTIFYING THE UNCERTAINTY

As remarked above, we must establish a region within which we believe the true values of the parameters will fall with some specified level of confidence. The method described here is called a *parametric bootstrap* (Efron and Tibshirani, 1993); the essential idea is as follows. Numerous samples (of size m) are simulated from the ideal distribution represented by our point estimates, μ_*, σ_*, etc., of the parameters. Maximum-likelihood estimates for μ and

---

[3]In the next section we will take these hypothetical values to be the maximum-likelihood values μ_*, σ_*, etc., predicted from the observed data.

σ, etc. are then calculated for each of these samples. The statistics of these estimates provide a basis for determining approximate confidence regions for the original point estimates.

An ordinary variant of bootstrapping may also be used to obtain robust confidence regions. In that case, the numerous samples required are obtained not from the ideal distribution, but by sampling with replacement many times from the original empirical sample. We chose to use the parametric variant outlined above and described below.

Although the method can be applied to any confidence level, for the sake of definiteness, we take that level to be 90% confidence for this discussion and for the example presented in the last section.

Also, for purposes of describing the method we will focus on the size problem, using PREDICTOR$_1$ and SIMULATOR$_1$ in a Monte Carlo scheme to describe the two-dimensional confidence region in the space (μ, σ). The case of shape and orientation is conceptually similar, but the discussion in the example will reflect the three-dimensional character of the region.

Now the only tool we have for assessing a given prediction is the likelihood function L_1 in Eq. 17 obtained from our real set of data $(b_1, \ldots, b_m)$. L_1 is regarded as a function of the parameters (μ, σ) of the prediction. It is convenient to standardize this as a likelihood ratio $L_1(\mu_*, \sigma_*)/L_1(\mu, \sigma)$, where (μ_*, σ_*) is the maximum-likelihood prediction. As a practical matter in the implementation of the method we use the logarithm of the likelihood function, so we define the log-likelihood-difference (*LLD*) by

$$LLD(\mu, \sigma) = \ln\left(\frac{L_1(\mu_*, \sigma_*)}{L_1(\mu, \sigma)}\right)$$
$$= \ln(L_1(\mu_*, \sigma_*)) - \ln(L_1(\mu, \sigma)). \quad (24)$$

We can now define the 90% confidence region in the (μ, σ)-plane. Our analysis is of a single data set $(b_1, \ldots, b_m)$ for which we have a single maximum-likelihood prediction (μ_*, σ_*). We consider that the experiment of collecting this data set is a relatively expensive affair. Nevertheless, suppose, contrary to fact, that we were able to collect many, say N, real data sets, all of the same size m, collectively producing many predictions $(\mu_1, \sigma_1), \ldots, (\mu_N, \sigma_N)$. Let x represent the 90th percentile of all the numbers $LLD(\mu_1, \sigma_1), \ldots, LLD(\mu_N, \sigma_N)$. Clearly, the inequality $LLD(\mu, \sigma) \leq x$ establishes a region R of the (μ, σ)-plane with the property that, with probability 0.90, any repetition of our experiment will result in a prediction

that falls within the region R. This is precisely what we mean by a 90% confidence region.

Unfortunately, x, as defined above, is not accessible without repeating our expensive experiment N times. Nevertheless, we can simulate the experiment as many times as we like by taking the most likely prediction (μ_*, σ_*) from our single experiment as the input to SIMULATOR$_1$. Since in this case we are taking only our most likely prediction (μ_*, σ_*) as our estimate of the "true" values — (μ_0, σ_0), this will produce only an approximation of the confidence region.

To summarize, our approximation to the 90% confidence region is determined as follows. The maximum-likelihood prediction (μ_*, σ_*) from the data is supplied as input to SIMULATOR$_1$ to produce N sets of simulated data. PREDICTOR$_1$ then produces a prediction for each of those sets of data: $(\mu_1, \sigma_1), \ldots, (\mu_N, \sigma_N)$. The 90% confidence region is then determined by the inequality $LLD(\mu, \sigma) \leq x$, where x is the 90th percentile of the numbers $LLD(\mu_1, \sigma_1), \ldots, LLD(\mu_N, \sigma_N)$.

A SAMPLE APPLICATION TO MITOCHONDRIA

In the section on the Maximum-likelihood Scheme we discussed our reasons for believing that the 3D structures we are studying exhibit anisotropy in vertical orientation. However our evidence for it comes from three-dimensional confocal microscopy and is visual rather that quantitative. It would have been desirable to have sectioning from a number of different angles (besides just horizontal) in order to quantify the anisotropy directly from the sectioned data. However, the cells whose mitochondria are being sectioned are grown on a flat surface and are very thin. Non-horizontal sectioning in this circumstance is technically difficult; in fact for this study, prohibited. Nevertheless, the parameter α in our statistical model provides us with an indirect confirmation of the presence of anisotropy as well as a quantitative measure.

In the introduction we promised a direct empirical argument in favor of the prolate hypothesis for the shape of the spheroid particles. Weibel (1979) states two rough empirical tests that can be performed on the sectioned data to decide between prolate and oblate. As a first test, consider the diameter, d, of the profile of largest area among the 5% most circular profiles. For comparison, consider the largest diameter, $2a$, and the smallest diameter, $2b$, of the profile of largest area among the 5% most elongated profiles. Weibel agrues

that if $d \approx 2b$, the 3D structures are most likely prolate; if $d \approx 2a$, they are most likely oblate. In our case, $d = 340$ nm, $2a = 890$ nm, and $2b = 270$ nm, pointing to the prolate hypothesis. As a second test, Weibel points out that, in the case of prolate spheroids, we expect the near circular profiles to be more numerous than the highly elliptical profiles. This can be checked using the axis ratio, r, of an elliptical profile, which is defined as the ratio of the largest diameter to the smallest. In our case, for the sectioned data, the median value of r is 1.52. When this is compared with the minimum (1.02) and maximum (5.70) values of r, the test again confirms the prolate hypothesis.

As discussed in the introduction, the motivation for this research derives from the desire to analyze conformational changes, particularly changes in shape of mitochondria undergoing apoptosis. The motivation is to construct a tool for quantifying the progression of shapes after the tubular mitochondrion splits into fragments which eventually become spherical. Our method could be used to monitor those changes quantitatively. We apply our method to a group of HeLa cells in H_2O_2-induced apoptosis.

As it turns out, the "proof of concept" presented here is not an idle exercise. Some remarks are in order concerning the *a priori* feasibility of the task that we have set out. There is a sense in which the mathematical problem we have formulated has a fundamental ambiguity that may rear its ugly head in certain contexts. Consider two extreme populations of prolate spheroids: (*i*) a distribution of nearly perfect spheres with their major axes oriented horizontally, and (*ii*) a distibution of highly elongated spheroids with all of their major axes nearly vertical. In both cases (since our sectioning plane is defined to be horizontal) the 2D sections are nearly perfect circles. There is thus no information in the shape of the sections that can be used to infer the 3D shape of the spheroids. This example illustrates the type of problem that can arise when trying to unfold 3D information on both shape and orientation from 2D sections. The presence of such an ambiguity is revealed in the behavior of the likelihood function that is computed from the data. The likelihood function could have several peaks that are not well separated and/or that are of roughly the same height, or it may have an ill defined peak. In these cases this method as it is formulated cannot be applied without bringing in additional information about the population of spheroids that would permit a resolution of the ambiguity.

For the data that we analyzed, a simple constraint of $\alpha \leq 5$ for the parameter in $\Phi_0(\theta|\alpha)$ was found sufficient to guarantee a single well-defined interior peak for the likelihood functions L_1 and L_2. This constraint rules out any spheroid population with a standard deviation of θ less than 25.0 degrees. That is, if more than about 68% of the spheroids are so flattened that their axes rise no more than 25.0 degrees above horizontal, then the results we have obtained may not be valid. While we consider this highly unlikely, it has to be recognized as a limitation of this method.

We generated a set of micrographs of the group of cells described above and in the Ph.D. thesis of one of the authors (Sun, 2007). The outer membranes of a total of 231 mitochondrial sections were traced in these micrographs using the program ImageJ, which ellipticized each trace yielding minor and major axes, from which a semi-minor axis b and an eccentricity e could be determined[4]. Fig. 2 shows a typical electron micrograph used in our tracings.
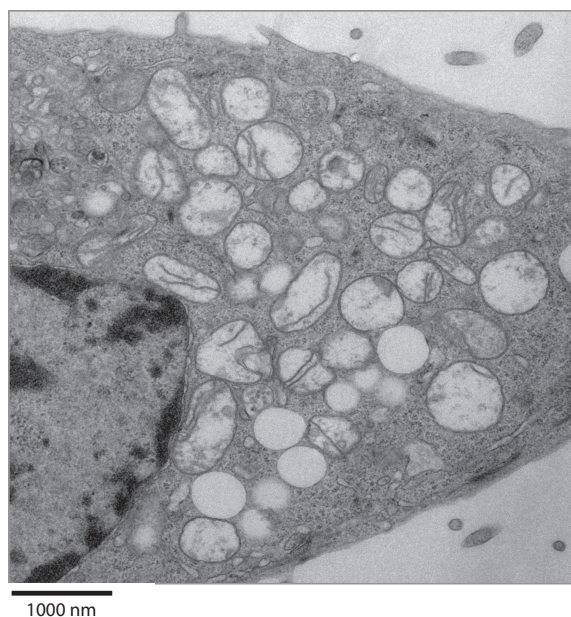

1000 nm

Fig. 2. *Typical electron micrograph of a HeLa cell having undergone H_2O_2 induced apoptosis.*

Sun *et al.* (2007) contains a description of how the thin sections were prepared. The thickness of the sections, 80 nm, represents about 15% of a typical particle diameter. Any thoroughgoing effort to use our method would need to address the standard stereological problems of "lost caps" and overprojection associated with the finite thickness of the sections (Weibel, 1979; 1980). We did not do so in this study for two reasons. First, we anticipate that

---

[4]ImageJ, (Image Processing and Analysis in Java, http://rsb.info.nih.gov/ij/) uses the "best fit ellipse" algorithm to match the moment of inertia of the region traced to determine an equivalent ellipse.

in our case the corrections would have a minor effect on the results. Secondly, the scope of our enterprise did not extend beyond laying out the theoretical apparatus and testing the feasibility of the full-scale unfolding problem of size, shape and orientation of spheroids. Handling the corrections above for nonspherical particles is not well developed and would have been, for us in this effort, a theoretical distraction.

We now recall in outline the procedure described in the previous three sections for handling the data, and we summarize the results as applied to the data derived from our micrographs. The predicted size distribution $\Phi_1(r)$ of the prolate spheroids is dependent on the parameters μ_r, σ_r, which describe the most likely normal distribution of the semi-minor axis, r. To produce a likelihood in terms of μ_r, σ_r, Eqs. 15 and 17 are combined to give

$$L_1(D_1|\mu_r,\sigma_r) = \prod_{j=1}^{231} \int_{b_j}^{\infty} \frac{b_j\,\Phi_1(r|\mu_r,\sigma_r)}{\sqrt{r^2-b_j^2}}\mathrm{d}r \quad (25)$$

in terms of the 231 semi-minor elliptical axes $b_1,\ldots,b_{231}$ obtained from the micrographs.

Fig. 3 shows the contour of L_1 that surrounds the 90% confidence region for μ_r, σ_r, as well as the point marking the maximum-likelihood values. Fig. 4 displays the maximum-likelihood distribution in black. The 90% confidence region is suggested by the green region surrounding the black curve. It is the result of superimposing all of the distributions whose parameter values lie in the 90% confidence region shown in Fig. 3.
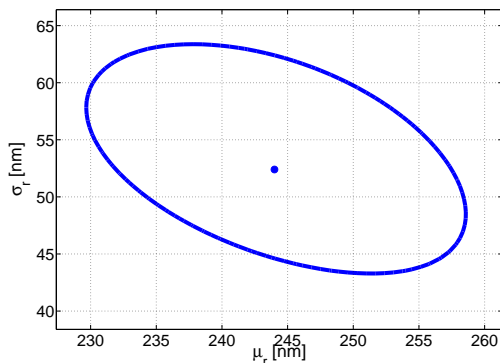

Fig. 3. *The maximum likelihood point and the 90% confidence region for the mean and standard deviation of the semi-minor axis, r, of the spheroid.*
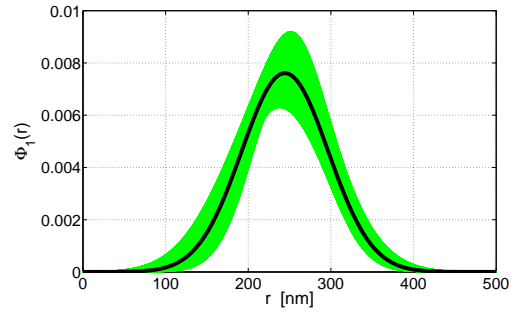

Fig. 4. *The maximum likelihood spheroid size distribution (black) and the 90% confidence region (green).*

The predicted spheroid shape and orientation are found in a similar manner. Eqs. 16, 17, 19, and 21 are combined to produce the likelihood $L_2(D_2|p,q,\alpha)$, which is then optimized over p,q and α. The parameters p and q describe a Beta distribution, while α determines the orientation distribution. By applying Eq. 20 the optimal p, q yield the eccentricity statistics $\mu_{e_0} = 0.875$ and $\sigma_{e_0} = 0.091$. Fig. 5 shows the projection of the 90% confidence region. Fig. 6 illustrates the optimal Beta distribution (in black) as well as the continuum of Beta distributions found within the 90% confidence region (in green).
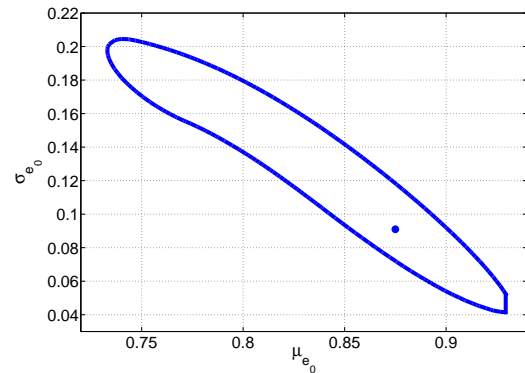

Fig. 5. *The maximum likelihood point and the 90% confidence region for the mean and standard deviation of the spheroid eccentricity, e_0.*
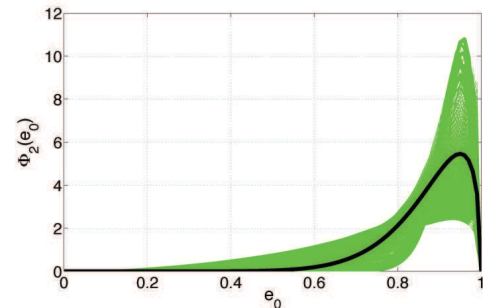

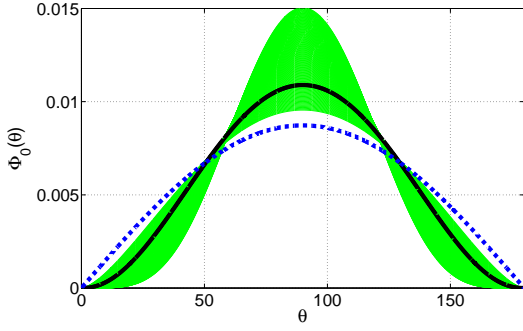Fig. 6. *The maximum likelihood spheroid eccentricity distribution (black) and the 90% confidence region (green).*

Fig. 7. *The maximum likelihood spheroid orientation distribution (black), a uniform distribution (dashed-blue) and the 90% confidence region (green).*

The predicted orientation distribution $\Phi_0(\theta|\alpha)$ occurred for $\alpha = 2.26$. This value of α corresponds to $\sigma_\theta = 32.82$ degrees and an isotropy index, from Eq. 22, of $I_{iso} = .838$. The predicted distribution of orientation is seen in Fig. 7. Also shown for comparison is the distribution corresponding to perfect isotropy. The 90% confidence analysis shows that our measured anisotropy is statistically significant.

The 90% confidence regions shown in Figs. 3 and 5 were determined by performing a parametric bootstrap as described in the previous section. Explicitly, in the case of spheroid size (semi-minor axis, r, Fig. 3) the maximum-likelihood values $\mu_* = 244$ nm and $\sigma_* = 52.4$ nm were taken to determine the model distribution $\Phi_1(r)$ from Eq. 18 and the associated predicted distribution for elliptical cross-section size, $\hat{\Psi}_1(b)$ after normalizing Eq. 15. A Monte Carlo scheme to simulate the slicing experiment was iterated 5000 times ($N = 5000$). Each iteration involved sampling 231 ($m = 231$) values of b from the distribution $\hat{\Psi}_1(b)$ and interpreting that sample as data for a new prediction of μ and σ. These 5000 predictions were then processed as explained in the previous section. Fig. 8 shows the raw collection of predictions in the (μ, σ)-plane, clustering nicely about the original maximum-likelihood prediction (shown in the center).

The blue curve in Fig. 9 is an empirical plot, for all 5000 iterations, of *LLD* versus rank in the form of percentile. As explained in Sec. 6, the value of *LLD* (= 3.3) determines the boundary of the 90% confidence region. This is the contour shown in Fig. 3.
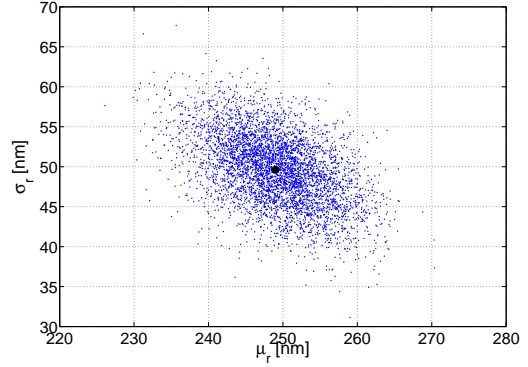

Fig. 8. *Graphical depiction of uncertainty in the method, showing predictions of μ_r and σ_r obtained from 5000 simulations of the sampling experiment (231 sections each) based on a population of spheroids with parameters corresponding to the data-based, maximum-likelihood prediction (shown in center).*
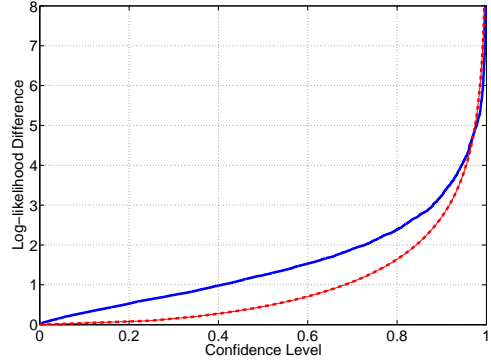

Fig. 9. *The log likelihood difference as a function of the confidence level. The dashed red curve is obtained from the likelihood ratio test. The solid blue curve is obtained from the parametric bootstrap.*

The second curve in Fig. 9 (red, dashed) is shown for comparison. If we had chosen not to perform the bootstrap, *i.e.* not to collect statistics on the performance of our method, we could have used the so-called Likelihood Ratio Test to establish confidence regions (Kendall and Stuart, 1983; Hilborn and Mangel, 1997). This is an asymptotic result (large m) stating that under certain conditions $LLD/2$ is distributed like a chi-square statistic of one degree of freedom. The associated percentile curve is shown in red. The discrepancy between these two curves illustrates the necessity, in the case of our method, to base uncertainty results on actual performance statistics, rather than assuming the conditions required for the Likelihood Ratio Test.

A similar parametric bootstrap was performed for results on spheroid eccentricity and orientation. In this case the parameter space $(\mu_{e_0}, \sigma_{e_0}, \alpha)$ is three-dimensional and so is the 90% confidence region.

Again, 5000 simulations based on the maximum-likelihood parameters were carried out and the *LLD* values ranked in an empirical curve. This curve fixed the *LLD* value, and, hence, the contour for the bounding level surface of the confidence region. Fig. 5 shows the projection of this region onto the $(\mu_{e_0}, \sigma_{e_0})$-plane, this projection giving the 90% confidence region for those parameters.

The results from our method may be summarized as follows. At the apoptotic state measured, the mitochondria exhibited a mean eccentricity of 0.875 with a standard deviation of 0.091. In addition the semi-minor axis, *r*, of the spheroid had a mean value of 244 nm with a standard deviation of 52.4 nm. This corresponds to a semi-major axis of 504 nm. Our most likely prolate spheroid is depicted in Fig. 10. Finally, the principal axes of the mitochondria showed a statistically significant directional anisotropy. Their directions showed an isotropic index of 0.838. That is, their spread about the horizontal was only 84% of that expected in an isotropic distribution.
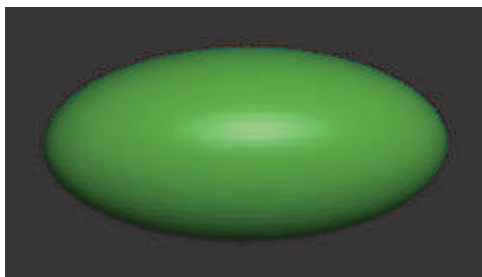

Fig. 10. *Spheroid with mean dimensions predicted by the method: semi-minor axis 244 nm, semi-major axis 504 nm, eccentricity 0.875.*

CONCLUSION

We developed a novel stereological method that utilizes a maximum-likelihood scheme to infer the distribution of spheroidal shapes, sizes and orientations from the observed distribution of elliptical sections. The maximum-likelihood method is applied to a parametrized set of distributions appropriate for studying the shape of mitochondria undergoing apoptosis. A parametric bootstrap for estimating confidence intervals for the parameters in our distributions is described and illustrated for an example. We expect that future applications to resolving time variations of apoptotic mitochondria will provide insights into apoptosis. We also expect our method, with possible modifications of the parametrization, to serve as a useful template for other specialized applications of the unfolding problem.

REFERENCES

Batschelet E (1981). Circular statistics in biology. London: Academic Press.

Blödner R, Mühlig P, Nagel W (1984). The comparison by simulation of solutions of Wicksell's problem. J Microsc 135:61-74.

Cruz-Orive LM (1976). Particle size-shape distributions: the general spheroid problem. J Microsc 107:233-53.

De Hoff RT (1962). The determination of the size distribution of ellipsiodal particles from measurements made on random plane sections. Trans Metall Soc AIME 224:474-7.

Efron B, Tibshirani RT (1993). An introduction to the bootstrap. New York: Chapman and Hall.

Frank S, Gaume B, Bergman-Leitner ES, Leitner WW, Robert EG, Catez F, Smith CL, Youle RJ (2001). The roles of dynamin-related protein 1, a mediator of mitochondrial fission in apoptosis. Dev Cell, 1:515-25.

Franklin JN (1977). Some stereological principles in morphometric cytology. SIAM J Appl Math 33:267-78.

Hilborn R, Mangel M (1997). The Ecological Detective: Confronting Models with Data. Princeton, NJ: Princeton University Press.

Keiding N and Jensen ST (1972). Maximum likelihood estimation of the size distribution of liver cell nuclei from the observed distribution in a plane section. Biometrics 28:813-829.

Kendall M and Stuart A (1983). The Advanced Theory of Statistics, Vol. 2. London: C. Griffin.

Mardia KV (1972). Statistics of directional data. London: Academic Press.

Sato E, Kondo N, Wakai F (1996). Particle size, shape and ori- entation distributions: general spheroid problem and application to deformed Si_3N_4 microstructures. Philos Mag A 74:215-28.

Selby B (1964). Girdle distributions on a sphere. Biometrika 51:381-92.

Stoyan D, Kendall WS, Mecke J (1987). Stochastic geometry and its applications. Berlin: John Wiley & Sons.

Sun MG (2007). Mitochondrial Structure During Apoptosis. Ph.D. thesis, San Diego State University, San Diego.

Sun MG, Williams J, Munoz-Pinedo C, Perkins GA, Brown JM, Ellisman MH, Green DR, Frey TG (2007). Correlated three-dimensional light and electron microscopy reveals transformation of mitochondria during apoptosis. Nature Cell Biol 9:1057-65.

Watson GS (1965). Equatorial distributions on a sphere. Biometrika 52:193-201.

Weibel ER (1979). Stereological methods, vol. 1. London: Academic Press.

Weibel ER (1980). Stereological methods, vol. 2. London: Academic Press.

Wicksell SD (1925). The corpuscle problem. A mathematical study of a biometric problem. Biometrika 17: 84-99.

Wicksell SD (1926). The corpuscle problem. Second Memoir. Case of ellipsoidal corpuscles. Biometrika 18:152-72.