Supplementary Materials: Size and shape selective classification of nanoparticles

Cornelia Damm, Danny Long, Johannes Walter, and Wolfgang Peukert

S1. Code for creating the figures

All of the figures created in the main text (aside from the ones reproduced from other work) can be reproduced by downloading the associated Matlab files from https://github.com/dklong-csu/nd_psds_final_report. There is a Matlab script for each figure generated and a few additional files to define helpful functions.

S2. Details regarding multivariate normal and lognormal distributions

In the article, we present a formula for the multi-dimensional lognormal distribution (equation 6). This formula is written concisely for publication through linear algebra notation, however we acknowledge that not all readers will be familiar with this notation. For those readers whose background has not involved mathematics heavily, we would like to walk through a few examples of what the verbose formulas are. Moreover, since many readers may not be familiar with lognormal distributions or, more generally, multivariate distributions, we discuss the connection to the more familiar normal distribution.

S2.1. Relating the normal distribution to the lognormal distribution

The univariate normal distribution is widely used in science. The probability density function is given by the formula

$$q(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\}, \tag{1}$$

where μ is the mean value and σ is the standard deviation. The normal distribution is also described by its cumulative distribution function

$$Q(x) = \int_{-\infty}^{x} q(z)\,\mathrm{d}z. \tag{2}$$

In the context of a particle size distribution, $Q(x)$ provides the percentage of particles smaller than size x (based on the weighting of the distribution; see Section S3 for more on this). However, the normal distribution has an issue with interpretation: negative values (i.e., negative sizes!) are included for x. Hence, the lognormal distribution is commonly used to avoid this unphysical behavior.

The lognormal distribution is defined such that the natural logarithm of the independent variable (size) follows a normal distribution. Vice versa, if data x follow a normal distribution, then $y = \exp(x)$ follow a lognormal distribution. We can derive the equation for a lognormal distribution through calculus properties. We begin with the cumulative distribution function:

$$Q(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2 \right\} \mathrm{d}z. \tag{3}$$

We now substitute variables with $y = \exp(z) \Rightarrow \ln(y) = z$. However, we are substituting within an integral and wish for the integral to evaluate to the same value, hence we need to utilize u-substitution (or, in this case, y-substitution). The calculus procedure for this is to replace z and the differential increment $\mathrm{d}z$ by noting

$$\begin{aligned} z &= \ln(y), \\ \frac{\mathrm{d}}{\mathrm{d}y}(z) &= \frac{\mathrm{d}}{\mathrm{d}y}(\ln(y)), \\ \frac{\mathrm{d}z}{\mathrm{d}y} &= \frac{1}{|y|}, \\ &\Rightarrow \\ \mathrm{d}z &= \frac{\mathrm{d}y}{|y|}. \end{aligned} \tag{4}$$

Hence, in (3), we exchange every z with $\ln(y)$ and every $\mathrm{d}z$ with $\mathrm{d}y/|y|$, which yields

$$Q(\hat{x} = \ln(x)) = \int_0^{\hat{x}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln(y) - \mu}{\sigma}\right)^2\right\} \frac{\mathrm{d}y}{|y|}. \tag{5}$$

The integral bounds are computed as follows:

$$\begin{aligned} &\text{Lower:} \quad z = -\infty \to y = \lim_{z \to -\infty^+} \exp(z) = 0, \\ &\text{Upper:} \quad z = x \to y = \ln(z) = \ln(x) = \hat{x}. \end{aligned}$$

Therefore, by the fundamental theorem of calculus, the density function $q(y)$ is equal to the integrand of (5):

$$q(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln(y) - \mu}{\sigma}\right)^2\right\}, \tag{6}$$

where we remove the absolute value from y in the preceding fraction since it is now understood that y only takes positive values.

We can similarly understand the multivariate lognormal distribution through the multivariate normal distribution. First of all, the multivariate normal distribution – where N is the number of dimensions – is defined as

$$q(\vec{x}) = \left(\frac{1}{2\pi}\right)^{N/2} \left(\frac{1}{|\Sigma|}\right)^{1/2} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^\intercal \Sigma^{-1} (\vec{x} - \vec{\mu})\right). \tag{7}$$

Here $|\Sigma|$ is the matrix determinant, $\vec{\mu}$ is a vector containing the average value for each dimension, and Σ is the covariance matrix, which contains the variance and covariance terms.

The mean vector $\vec{\mu}$ is a simple extension of the mean in the univariate case:

$$\begin{aligned} \vec{\mu} &= \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_N \end{bmatrix}^\intercal \\ \mu_i &= \int_{\vec{x} \in \mathbb{R}^N} x_i q(\vec{x}) \, \mathrm{d}\vec{x}. \end{aligned} \tag{8}$$

For estimating the mean from sample data, the formula is also similar to the univariate case

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_i^{(j)}, \tag{9}$$

where $x_i^{(j)}$ indicates the ith dimension of the jth data point and n is the number of data points.

The covariance matrix is defined such that the element in row i and column j is

$$\begin{aligned}\sigma_{ij} &= E\left[(x_i - \mu_i)(x_j - \mu_j)\right] \\ &= \int_{\vec{x}\in\mathbb{R}^N} (x_i - \mu_i)(x_j - \mu_j) q(\vec{x})\,\mathrm{d}\vec{x}.\end{aligned} \tag{10}$$

If $i = j$, then (10) is the traditional variance formula. In addition, $\sigma_{ij} = \sigma_{ji}$ since the multiplication order in (10) does not matter. This leads to

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{bmatrix}. \tag{11}$$

If estimating the covariance from sample data, the formula for each entry of the matrix is

$$\sigma_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} \left(x_i^{(k)} - \bar{x}_i\right)\left(x_j^{(k)} - \bar{x}_j\right). \tag{12}$$

We can then derive the formula for a multivariate lognormal distribution in a similar manner to the univariate case. Here, we perform the log transform for each dimension $x_i = \ln(y_i)$ and perform u-substitution for each to find

$$\begin{aligned} Q(\vec{x}) &= \int_{\vec{0}}^{\vec{x}} \left(\frac{1}{2\pi}\right)^{N/2} \left(\frac{1}{|\Sigma|}\right)^{1/2} \exp\left(-\frac{1}{2}(\vec{z}-\vec{\mu})^{\intercal}\Sigma^{-1}(\vec{z}-\vec{\mu})\right) \mathrm{d}\vec{z}, \\ \text{Substitute:} &\quad \left(\vec{z} = \ln(\vec{y}) \Rightarrow \mathrm{d}\vec{z} = \frac{\mathrm{d}\vec{y}}{\prod_{i=1}^{N} |y_i|}\right), \\ Q(\vec{w} = \ln(\vec{x})) &= \int_{\vec{0}}^{\vec{w}} \prod_{i=1}^{N} y_i^{-1} \left(\frac{1}{2\pi}\right)^{N/2} \left(\frac{1}{|\Sigma|}\right)^{1/2} \exp\left(-\frac{1}{2}(\ln(\vec{y})-\vec{\mu})^{\intercal}\Sigma^{-1}(\ln(\vec{y})-\vec{\mu})\right) \mathrm{d}\vec{y}. \end{aligned} \tag{13}$$

Therefore, the density function is

$$q(\vec{y}) = \prod_{i=1}^{N} y_i^{-1} \left(\frac{1}{2\pi}\right)^{N/2} \left(\frac{1}{|\Sigma|}\right)^{1/2} \exp\left(-\frac{1}{2}(\ln(\vec{y})-\vec{\mu})^{\intercal}\Sigma^{-1}(\ln(\vec{y})-\vec{\mu})\right). \tag{14}$$

S2.1.1. Specific formulas for two-dimensional distributions

To make things more explicit, we can write the formula for the two-dimensional density functions without vectors and matrices. Let

$$\begin{aligned} \vec{y} &= \begin{bmatrix} y_1 & y_2 \end{bmatrix}^{\intercal}, \\ \vec{\mu} &= \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix}^{\intercal}, \\ \Sigma &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}. \end{aligned} \tag{15}$$

The determinant of a 2×2 matrix has a simple formula, so we have

$$|\Sigma| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2. \tag{16}$$

Similarly, the inverse of a 2×2 matrix can be written easily as

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}. \tag{17}$$

With these facts, we can write the equation for a two-dimensional lognormal density with

$$q(y_1,y_2) = \frac{1}{2\pi y_1 y_2}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2-\sigma_{12}^2}}\exp\left(-\frac{1}{2}\frac{\sigma_2^2(\ln(y_1)-\mu_1)^2-2\sigma_{12}(\ln(y_1)-\mu_1)(\ln(y_2)-\mu_2)+\sigma_1^2(\ln(y_2)-\mu_2)^2}{\sigma_1^2\sigma_2^2-\sigma_{12}^2}\right). \tag{18}$$

We can also write (18) in terms of the correlation for a more intuitive form. By definition, the correlation ρ satisfies $\sigma_{12}=\rho\sigma_1\sigma_2$. Therefore,

$$\sigma_1^2\sigma_2^2-\sigma_{12}^2 = \sigma_1^2\sigma_2^2-(\rho\sigma_1\sigma_2)^2 = \sigma_1^2\sigma_2^2(1-\rho^2), \tag{19}$$

and by substituting (19) into (18), we yield

$$q(y_1,y_2) = \frac{1}{2\pi y_1 y_2\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\left(\frac{\ln(y_1)-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{\ln(y_1)-\mu_1}{\sigma_1}\right)\left(\frac{\ln(y_2)-\mu_2}{\sigma_2}\right)+\left(\frac{\ln(y_2)-\mu_2}{\sigma_2}\right)^2\right)\right]. \tag{20}$$

S3. Weighting particle size distributions differently

In the main article, a few nice relationships were made concerning transforming particle size distributions to be weighted differently (e.g., number, surface, or volume weighted). It is worth showing how these relationships are derived, but these proofs are tedious and would not contribute much to the main text. Instead, we provide our derivations here.

Let the dimension of the particle size vector $\vec{x}$ be N; that is, $\vec{x}\in\mathbb{R}^N$. The most general result is as follows. If q_r is lognormal with mean μ and covariance Σ, then the generalized moment conversion to q_k with

$$\kappa(\vec{x}) = C\prod_{i=1}^{N} x_i^{b_i}$$

results in q_r being lognormally distributed as well. If $LN(\mu,\Sigma)$ denotes the probability density function of the lognormal distribution and

$$\vec{b} := \begin{bmatrix} b_1 & b_2 & \cdots & b_N\end{bmatrix}^\intercal,$$

then it is the case that

$$\begin{aligned} q_r(\vec{x}) &= LN(\mu,\Sigma) \\ &\implies \text{(by definition)} \\ q_k(\vec{x}) &= \frac{\kappa(\vec{x})q_r(\vec{x})}{\int_0^\infty \kappa(\vec{x})q_r(\vec{x})} \\ &\implies \text{(via my derivations below)} \\ q_k(\vec{x}) &= LN(\mu+\Sigma\vec{b},\Sigma). \end{aligned} \tag{21}$$

In the case that

$$q_0(\vec{x}) = LN(\mu,\Sigma),$$

and the particles are cylinders. We have that

$$\vec{x} = \begin{bmatrix} d & \ell\end{bmatrix}^\intercal$$

and the conversion $q_0\to q_3$ is conducted via

$$\kappa(\vec{x}) = \frac{\pi}{4}d^2\ell.$$

Therefore

$$\vec{b} = \begin{bmatrix} 2 & 1 \end{bmatrix}^\intercal \tag{22}$$

and hence

$$\boxed{\begin{aligned} q_0(\vec{x}) &= LN(\mu, \Sigma) \\ &\implies \\ q_3(\vec{x}) &= LN\left(\mu + \Sigma\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \Sigma\right). \end{aligned}} \tag{23}$$

If the surface area of a cylinder is used as κ, then $q_2(\vec{x})$ is constructed instead using

$$\kappa(\vec{x}) = \pi d\ell + \frac{1}{2}\pi d^2$$

and as a result we find a weighted sum of two lognormal distributions

$$\boxed{\begin{aligned} q_0(\vec{x}) &= LN(\mu, \Sigma) \\ &\implies \\ q_2(\vec{x}) &= w_1 LN\left(\mu + \Sigma\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \Sigma\right) + w_2 LN\left(\mu + \Sigma\begin{bmatrix} 2 \\ 0 \end{bmatrix}, \Sigma\right) \\ &\text{where} \\ w_1 &= \frac{\pi C_1}{\pi C_1 + \frac{1}{2}\pi C_2} \\ &\text{and} \\ w_2 &= \frac{\frac{1}{2}\pi C_2}{\pi C_1 + \frac{1}{2}\pi C_2} \\ &\text{for constants} \\ C_1 &= \exp\left\{\frac{1}{2}\begin{bmatrix} 1 & 1 \end{bmatrix}\left(\Sigma\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2\mu\right)\right\} \\ &\text{and} \\ C_2 &= \exp\left\{\frac{1}{2}\begin{bmatrix} 2 & 0 \end{bmatrix}\left(\Sigma\begin{bmatrix} 2 \\ 0 \end{bmatrix} - 2\mu\right)\right\}. \end{aligned}} \tag{24}$$

S3.1. Derivations

S3.1.1. General product of size components

First some notation. The following computations are easier to understand using inner-product notation from Linear Algebra. Inner product for real-valued vectors is the dot product

$$\langle x, y \rangle := x^\intercal y.$$

The following properties also hold for x, y, z, w being real-valued vectors, a being a scalar, and A being a real-valued matrix:

$$
\begin{aligned}
\text{Inner product is linear} &\implies \langle x+y, z+w\rangle = \langle y,z\rangle + \langle y,w\rangle + \langle x,z\rangle + \langle x,w\rangle \\
\text{Inner product is linear} &\implies \langle ax, y\rangle = a\langle x,y\rangle \\
\text{Matrix multiplication is linear} &\implies \langle x, A(y+z)\rangle = \langle x, Ay\rangle + \langle x, Az\rangle \\
\text{Inner product has (conjugate) symmetry} &\implies \langle x,y\rangle = \langle y,x\rangle \\
\text{If } A \text{ is symmetric} &\implies \langle Ax, y\rangle = \langle x, Ay\rangle.
\end{aligned}
\tag{25}
$$

The covariance matrix Σ in a lognormal distribution is symmetric and positive-definite. This implies that Σ^{-1} exists and is also symmetric.

Let a particle size distribution (PSD) be lognormally distributed. Denote this as $q(\vec{x})$ and let the parameters of the lognormal distribution be μ and Σ. We then want to re-weight $q(\vec{x})$ to a different PSD $q_k(\vec{x})$ through the generalized moment method

$$
q_k(\vec{x}) = \frac{q(\vec{x})\kappa(\vec{x})}{\int_0^\infty q(\vec{x})\kappa(\vec{x})\mathrm{d}\vec{x}}. \tag{26}
$$

Consider the special case of

$$
\kappa(\vec{x}) = \prod x_i^{b_i}.
$$

The PSD $q(\vec{x})$ is lognormal and therefore – if the dimension of $\vec{x}$ is N – we have

$$
q(\vec{x}) = (2\pi)^{-N/2}|\Sigma|^{-1/2}\left(\prod_{i=1}^{N} x_i^{-1}\right)\exp\left\{\frac{-1}{2}\left\langle \ln(\vec{x}) - \mu, \Sigma^{-1}(\ln(\vec{x}) - \mu)\right\rangle\right\}.
$$

Therefore,

$$
q(\vec{x})\kappa(\vec{x}) = \left(\prod_{i=1}^{N} x_i^{b_i}\right)(2\pi)^{-N/2}|\Sigma|^{-1/2}\left(\prod_{i=1}^{N} x_i^{-1}\right)\exp\left\{\frac{-1}{2}\left\langle \ln(\vec{x}) - \mu, \Sigma^{-1}(\ln(\vec{x}) - \mu)\right\rangle\right\}. \tag{27}
$$

Note that

$$
x_i^{b_i} = \exp\left(\ln\left(x_i^{b_i}\right)\right) = \exp(b_i \ln(x_i))
$$

and hence

$$
\prod_{i=1}^{N} x_i^{b_i} = \prod_{i=1}^{N} \exp\{b_i \ln(x_i)\} = \exp\left\{\langle \vec{b}, \ln(\vec{x})\rangle\right\}.
$$

This can be substituted into (27) and the exponential terms can be combined to find

$$
q(\vec{x})\kappa(\vec{x}) = (2\pi)^{-N/2}|\Sigma|^{-1/2}\left(\prod_{i=1}^{N} x_i^{-1}\right)\exp\left\{\frac{-1}{2}\left\langle \ln(\vec{x}) - \mu, \Sigma^{-1}(\ln(\vec{x}) - \mu)\right\rangle + \langle \vec{b}, \ln(\vec{x})\rangle\right\}. \tag{28}
$$

Focus in on the expression inside the exponential. Using the properties (25), we perform a multivariate "complete the squares" operation as follows:

$$
\begin{aligned}
&-\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle + \langle \vec{b}, \ln(\vec{x})\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle -2\vec{b}, \ln(\vec{x})\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle -\vec{b}, \ln(\vec{x})\rangle - \frac{1}{2}\langle -\vec{b}, \ln(\vec{x})\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle -\Sigma^{-1}\Sigma\vec{b}, \ln(\vec{x})\rangle - \frac{1}{2}\langle \ln(\vec{x}), -\Sigma^{-1}\Sigma\vec{b}\rangle.
\end{aligned}
$$

Now let $-\Sigma\vec{b} := \vec{c}$. Then:

$$
\begin{aligned}
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})\rangle - \frac{1}{2}\langle \ln(\vec{x}), \Sigma^{-1}\vec{c}\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})\rangle - \frac{1}{2}\langle \ln(\vec{x}), \Sigma^{-1}\vec{c}\rangle - \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu)\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})\rangle - \frac{1}{2}\langle \ln(\vec{x})-\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle, \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\mu+\vec{c})\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle.
\end{aligned}
$$

To simplify notation, let $\mu - \vec{c} := \hat{\mu}$. With this substitution, we find:

$$
\begin{aligned}
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle - \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \ln(\vec{x})-\hat{\mu}\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle - \frac{1}{2}\langle \vec{c}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\mu+\vec{c}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\hat{\mu}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle + \frac{1}{2}\langle -\mu, \Sigma^{-1}\vec{c}\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\hat{\mu}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, -\hat{\mu}-\mu\rangle \\
&= -\frac{1}{2}\Big\langle \ln(\vec{x})-\hat{\mu}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle + \frac{1}{2}\langle \Sigma^{-1}\vec{c}, \vec{c}-2\mu\rangle.
\end{aligned}
$$

Therefore, we can express (28) as

$$
\begin{aligned}
q(\vec{x})\kappa(\vec{x}) &= \exp\left\{\frac{1}{2}\langle \Sigma^{-1}\vec{c}, \vec{c}-2\mu\rangle\right\}(2\pi)^{-N/2}|\Sigma|^{-1/2}\left(\prod_{i=1}^{N} x_i^{-1}\right)\exp\left\{\frac{-1}{2}\Big\langle \ln(\vec{x})-\hat{\mu}, \Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\Big\rangle\right\} \\
&= \exp\left\{\frac{1}{2}\vec{b}^{\intercal}(\Sigma\vec{b}-2\mu)\right\}\underbrace{(2\pi)^{-N/2}|\Sigma|^{-1/2}\left(\prod_{i=1}^{N} x_i^{-1}\right)\exp\left\{\frac{-1}{2}(\ln(\vec{x})-\hat{\mu})^{\intercal}\Sigma^{-1}(\ln(\vec{x})-\hat{\mu})\right\}}_{\text{Lognormal with mean } \hat{\mu} \text{ and covariance } \Sigma \to LN(\hat{\mu},\Sigma)}.
\end{aligned}
$$

This is then substituted into the definition of $q_k(\vec{x})$ (26) to find

$$
\begin{aligned}
q_k(\vec{x}) &= \frac{\exp\left\{\frac{1}{2}\vec{b}^\intercal(\Sigma\vec{b}-2\mu)\right\}LN(\hat{\mu},\Sigma)}{\exp\left\{\frac{1}{2}\vec{b}^\intercal(\Sigma\vec{b}-2\mu)\right\}\underbrace{\int_0^\infty LN(\hat{\mu},\Sigma)\mathrm{d}\vec{x}}_{=1}} \\
&= LN(\hat{\mu},\Sigma) \\
&= LN(\mu+\Sigma\vec{b},\Sigma).
\end{aligned}
\tag{29}
$$

Therefore, we have shown that (21) holds.

S3.1.2. Surface area of cylinder

Given

$$q_0(d,\ell) = LN(\mu,\Sigma)$$

and cylindrical particles, the conversion to $q_2(d,\ell)$ is done with the surface area formula

$$\kappa(d,\ell) := S(d,\ell) = \pi d\ell + \frac{1}{2}\pi d^2.$$

Thus we have

$$
\begin{aligned}
q_2(d,\ell) &= \frac{S(d,\ell)q_0(d,\ell)}{\int_0^\infty S(d,\ell)q_0(d,\ell)\mathrm{d}d\mathrm{d}\ell} \\
&= \frac{\pi d\ell LN(\mu,\Sigma)+\frac{1}{2}\pi d^2 LN(\mu,\Sigma)}{\pi\int_0^\infty d\ell LN(\mu,\Sigma)\mathrm{d}d\mathrm{d}\ell+\frac{1}{2}\pi\int_0^\infty d^2 LN(\mu,\Sigma)\mathrm{d}d\mathrm{d}\ell}.
\end{aligned}
\tag{30}
$$

From the previous derivation we know that

$$
\begin{aligned}
d\ell LN(\mu,\Sigma) &= \underbrace{\exp\left\{\frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\left(\Sigma\begin{bmatrix}1\\1\end{bmatrix}-2\mu\right)\right\}}_{C_1} LN\left(\mu+\Sigma\begin{bmatrix}1\\1\end{bmatrix},\Sigma\right) \\
d^2 LN(\mu,\Sigma) &= \underbrace{\exp\left\{\frac{1}{2}\begin{bmatrix}2 & 0\end{bmatrix}\left(\Sigma\begin{bmatrix}2\\0\end{bmatrix}-2\mu\right)\right\}}_{C_2} LN\left(\mu+\Sigma\begin{bmatrix}2\\0\end{bmatrix},\Sigma\right)
\end{aligned}
$$

and thus we can substitute these into (30) to find

$$
\begin{aligned}
q_2(d,\ell) &= \frac{\pi C_1 LN\left(\mu+\Sigma\begin{bmatrix}1\\1\end{bmatrix},\Sigma\right)+\frac{1}{2}\pi C_2 LN\left(\mu+\Sigma\begin{bmatrix}2\\0\end{bmatrix},\Sigma\right)}{\pi C_1\int_0^\infty LN\left(\mu+\Sigma\begin{bmatrix}1\\1\end{bmatrix},\Sigma\right)\mathrm{d}d\mathrm{d}\ell+\frac{1}{2}\pi C_2\int_0^\infty LN\left(\mu+\Sigma\begin{bmatrix}2\\0\end{bmatrix},\Sigma\right)\mathrm{d}d\mathrm{d}\ell} \\
&= \frac{\pi C_1 LN\left(\mu+\Sigma\begin{bmatrix}1\\1\end{bmatrix},\Sigma\right)+\frac{1}{2}\pi C_2 LN\left(\mu+\Sigma\begin{bmatrix}2\\0\end{bmatrix},\Sigma\right)}{\pi C_1+\frac{1}{2}\pi C_2} \\
&= w_1 LN\left(\mu+\Sigma\begin{bmatrix}1\\1\end{bmatrix},\Sigma\right)+w_2 LN\left(\mu+\Sigma\begin{bmatrix}2\\0\end{bmatrix},\Sigma\right),
\end{aligned}
$$

where

$$w_1 = \frac{\pi C_1}{\pi C_1 + \frac{1}{2}\pi C_2},$$
$$w_2 = \frac{\frac{1}{2}\pi C_2}{\pi C_1 + \frac{1}{2}\pi C_2}.$$

Hence, we find the result (24).

S3.2. Equations for distributions in Figure 2

In Figure 2 of the main text, a number-weighted particle size distribution is adjusted to be surface- and volume-weighted. For demonstration purposes, we carry out the explicit calculation here. The parameters of the number-weighted distribution $q_0(d, \ell)$ are chosen such that the marginal distribution for the diameter has a median value of 20 nm and the marginal distribution for the length has a median value of 100 nm. These choices lead to the parameters of the lognormal distribution being

$$\mu_0 = \begin{bmatrix} \ln(20) \\ \ln(100) \end{bmatrix} \approx \begin{bmatrix} 3.0 \\ 4.6 \end{bmatrix}. \tag{31}$$

Then, we let

$$\Sigma_0 = \begin{bmatrix} \ln(1.1) & 0.3\ln(1.1)\ln(1.1) \\ 0.3\ln(1.1)\ln(1.1) & \ln(1.1) \end{bmatrix} \approx \begin{bmatrix} 0.1 & 0.003 \\ 0.003 & 0.1 \end{bmatrix}. \tag{32}$$

In (32), the value of 0.3 simply provides a correlation between d and ℓ. The values of $\ln(1.1)$ indicate that in the univariate case of a lognormal distribution, 95% of the particles would be within the interval

$$\left[\text{median}/1.1^2, \quad 1.1^2 \times \text{median}\right].$$

In other words, 95% of the diameters within

$$\left[20/1.1^2 = 16.5, \quad 1.1^2 \times 20 = 24.2\right]$$

and 95% of the lengths within

$$\left[100/1.1^2 = 82.6, \quad 1.1^2 \times 100 = 121\right].$$

The correlation term of 0.3 distorts these ranges in the multivariate distribution.

The volume-weighted distribution $q_3(d, \ell)$ is computed with the formula in (23). Thus we can compute

$$\begin{aligned} \mu_3 &= \mu_0 + \Sigma_0 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \ln(20) \\ \ln(100) \end{bmatrix} + \begin{bmatrix} \ln(1.1) & 0.3\ln(1.1)\ln(1.1) \\ 0.3\ln(1.1)\ln(1.1) & \ln(1.1) \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \ln(20) + 2\ln(1.1) + 0.3\ln(1.1)\ln(1.1) \\ \ln(100) + 0.6\ln(1.1)\ln(1.1) + \ln(1.1) \end{bmatrix} \\ &\approx \begin{bmatrix} 3.2 \\ 4.7 \end{bmatrix} \\ &\Rightarrow \\ q_3(d, \ell) &= LN\left(\begin{bmatrix} 3.2 \\ 4.7 \end{bmatrix}, \begin{bmatrix} 0.1 & 0.003 \\ 0.003 & 0.1 \end{bmatrix} \right) \end{aligned} \tag{33}$$

The surface-weighted distribution $q_2(d, \ell)$ is computed with the formulas in (24). This gives

$$\begin{aligned}
\Sigma \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} \ln(1.1) + 0.3 \ln(1.1) \ln(1.1) \\ \ln(1.1) + 0.3 \ln(1.1) \ln(1.1) \end{bmatrix} \\
\Sigma \begin{bmatrix} 2 \\ 0 \end{bmatrix} &= \begin{bmatrix} 2 \ln(1.1) \\ 0.6 \ln(1.1) \ln(1.1) \end{bmatrix} \\
&\Rightarrow \\
C_1 &= \exp\left\{ \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} \ln(1.1) + 0.3 \ln(1.1) \ln(1.1) \\ \ln(1.1) + 0.3 \ln(1.1) \ln(1.1) \end{bmatrix} - 2 \begin{bmatrix} \ln(20) \\ \ln(100) \end{bmatrix} \right) \right\} \approx 0.0005 \\
C_2 &= \exp\left\{ \frac{1}{2} \begin{bmatrix} 2 & 0 \end{bmatrix} \left(\begin{bmatrix} 2 \ln(1.1) \\ 0.6 \ln(1.1) \ln(1.1) \end{bmatrix} - 2 \begin{bmatrix} \ln(20) \\ \ln(100) \end{bmatrix} \right) \right\} \approx 0.003 \\
w_1 &= \frac{\pi C_1}{\pi C_1 + \frac{1}{2} \pi C_2} = \frac{1}{4} \\
w_2 &= \frac{\pi C_2}{\pi C_1 + \frac{1}{2} \pi C_2} = \frac{3}{4} \\
&\Rightarrow \\
q_2(d, \ell) &= \frac{1}{4} LN\left(\mu_0 + \Sigma_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \Sigma_0 \right) + \frac{3}{4} LN\left(\mu_0 + \Sigma_0 \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \Sigma_0 \right) \\
&= \frac{1}{4} LN\left(\begin{bmatrix} 3.1 \\ 4.7 \end{bmatrix}, \begin{bmatrix} 0.1 & 0.003 \\ 0.003 & 0.1 \end{bmatrix} \right) + \frac{3}{4} LN\left(\begin{bmatrix} 3.2 \\ 4.6 \end{bmatrix}, \begin{bmatrix} 0.1 & 0.003 \\ 0.003 & 0.1 \end{bmatrix} \right).
\end{aligned}$$