**Soot Aggregate Sizing through Inverse Analysis of Light Scattering Data**

K. J. Daun and D. W. Burr
Department of Mechanical and Mechatronics Engineering
University of Waterloo
200 University Avenue W., Waterloo, ON, N2L 3G1

In most combustion processes pyrolyzed fuel forms nanospheres called primary particles that agglomerate into micro-scaled polydisperse fractal soot aggregates, as shown in Fig. 1. The impact of these aggregates on human health and the environment are functions of their motility and optical characteristics, which in turn depend on the number of primary particles per aggregate, $N_p$. Accordingly, there is a critical need for instruments that quickly assess the size of aggregates within a soot-laden aerosol, which is complicated by the fact that soot aggregate sizes are rarely monodisperse; instead, due to aggregation dynamics they obey an unknown distribution function $P(N_p)$.

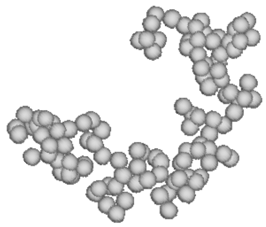

**Figure 1.** The transport and optical properties of soot aggregates depends on the number of primary particles, $N_p$, within the aggregate.

Multiangle elastic light scattering is a promising technique for characterizing $P(N_p)$. In this procedure, shown schematically in Fig. 2, a collimated light source is shone through a soot-laden aerosol and the scattered light is measured over a set of angles. Plotting the scattered light intensity versus the scattering angle reveals the mean particle size and a ratio of distribution moments, and by assuming a distribution type (most often lognormal) one can then infer the distribution parameters from this information [1]. The presumed distribution type may not necessary be correct, however.

An alternate way to recover $P(N_p)$ is by deconvolving the governing integral equation

$$b(\theta) = C\int_1^\infty P(N_p)K(\theta, N_p)\,dN_p \tag{1}$$

where the kernel, $K(\theta, N_p)$, is derived from light scattering physics [2] and $C$ depends on the experimental apparatus and the aggregate number density, $N_{agg}$, in the aerosol. If these parameters are known Eq. (1) is a Fredholm integral equation of the first-kind, which is ill-posed due to the smoothing properties of the kernel. In particular, since the intensity of light scattered by soot aggregates is roughly proportional to $N_p^2$ the contribution made to $b(\theta)$ by small aggregates is overwhelmed by scattering from larger aggregates.

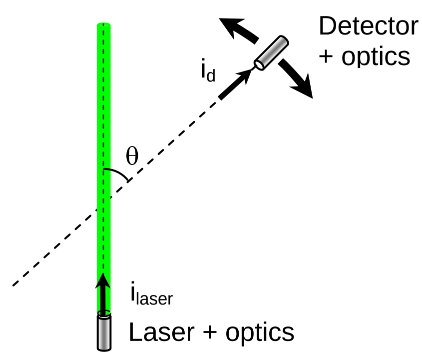


**Figure 2.** Schematic of a multiangle elastic light scattering experiment

This problem is solved by transforming Eq. (1) into a matrix equation, $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x}$ is a discrete form of $P(N_p)$ and $\mathbf{b}$ contains the angular scattering data. The ill-posedness of Eq. (1) means that $\mathbf{A}$ is ill-conditioned, so the light scattering data must be augmented with additional equations (called *priors*) that promote assumed attributes of $P(N_p)$. Analyzing the angular scattering intensity data with an assumed distribution type as described above is one way to do this, although this procedure ignores the underlying ill-posedness of the problem since it only finds one of many candidate solutions that could explain the data.

A better way to incorporate priors is through maximum a posteriori (MAP) inference, which finds the $\mathbf{x}$ that maximizes the posterior probability $P(\mathbf{x}|\mathbf{b})$, which through Bayes' theorem is given by

$$P\left(\mathbf{x}\middle|\mathbf{b}\right) = \frac{P\left(\mathbf{b}\middle|\mathbf{x}\right)}{P\left(\mathbf{b}\right)} P_{model}\left(\mathbf{x}\right) \tag{2}$$

where $P(\mathbf{b}|\mathbf{x})$ is the probability of the data in **b** occurring for a given **x**, $P(\mathbf{b})$ is the marginal probability of the data, and $P_{model}(\boldsymbol{x})$ is the probability of the solution being correct based on assumed priors. Since $P(\mathbf{b})$ only scales $P(\mathbf{x}|\mathbf{b})$, it can be ignored.

Relying solely on the multiangle scattering data without incorporating additional priors is equivalent to maximizing $P(\mathbf{b}|\mathbf{x})$ by itself. Assuming the data in **b** is contaminated with Gaussian-distributed error having a variance $\sigma^2$,

$$P\left(\mathbf{b}\middle|\mathbf{x}\right) \propto \exp\left[-\frac{1}{2\sigma^2}\left\|\mathbf{Ax}-\mathbf{b}\right\|_2^2\right] \tag{3}$$

which is minimized by solving $\mathbf{Ax} = \mathbf{b}$. Since **A** is ill-conditioned, however, there exists a large set of solutions that "almost" minimizes $||\mathbf{Ax}-\mathbf{b}||_2^2$, so it is necessary to add additional priors through $P_{model}(\mathbf{x})$ to obtain a robust estimate for **x**. Based on the physics of aggregation we expect **x** to be smooth, which can be promoted through a Gibbs-type prior,

$$P_{Gibbs}\left(\mathbf{x}\right) \propto \exp\left(-\beta\left\|\mathbf{Lx}\right\|_2^2\right) \tag{4}$$

where **L** is a first-order smoothing matrix defined so that Eq. (4) is minimized by a uniform solution. By definition **x** must also be strictly nonnegative, which is enforced by

$$P_{nonneg}\left(\mathbf{x}\right) = \prod_{j=1}^{n} \mathrm{H}\left(x_j\right) \tag{5}$$

where the Heaviside function, $\mathrm{H}(x)$, is zero if $x < 0$ and otherwise equals unity. Assembling Eqs. (2-5), letting $\lambda^2 = 2\sigma^2\beta$, and taking the logarithm shows that $P(\mathbf{x}|\mathbf{b})$ is maximized by the solution of the constrained linear least squares problem

$$\mathbf{x}^* = \arg\min f\left(\mathbf{x}\right) = \arg\min \left\| \begin{bmatrix} \mathbf{A} \\ \lambda\mathbf{L} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} \right\|_2^2 \ s.t.\ \mathbf{x} \geq 0 \tag{6}$$

which is equivalent to constrained Tikhonov regularization [3, 4].

We demonstrate this technique by deconvolving artificial data generated with $C = 1$ and a specified $P(N_p)$ found by fitting a lognormal distribution to a histogram approximation of $P(N_p)$ derived from transmission electron microscopy (TEM) of soot sampled from a well-characterized flame [5]. The artificial data is contaminated with Gaussian noise corresponding to a standard deviation of 3%, which is typical of experimentally-collected data. Figure 3 (*a*) shows the solution obtained with a value of $\lambda$ corresponding to the location of maximum curvature on the L-curve in Fig. 3 (*b*) [4]. The reconstructed distribution agrees with the imposed distribution except at small $N_p$, which is expected given the $N_p^2$ dependence on scattering intensity described above.

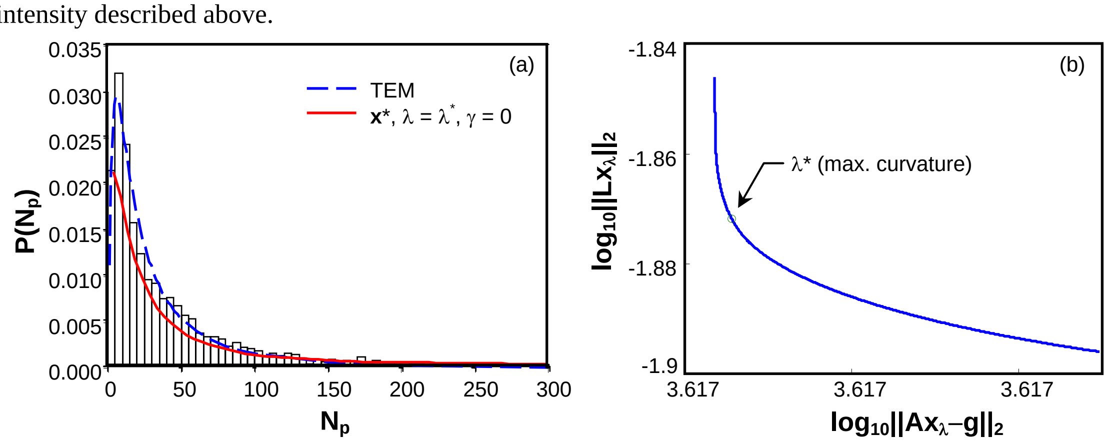


**Figure 3.** (*a*) Reconstructed artificial data solved using MAP with smoothness and nonnegativity priors, and (*b*) the L-curve used to find $\lambda$.

The problem is substantially more challenging if $C$ is unknown, which is often the case since $N_{agg}$ is difficult to measure independently. In this case, it would be desirable to treat $C$ as an unknown parameter to be solved by MAP inference. We evaluate the feasibility of this approach by analyzing angular scattering data found under identical experimental conditions used to obtain the TEM-histogram [6].

Unfortunately $f(\mathbf{x})$ derived from the smoothness and nonnegativity priors lacks a strong minimum, so it is necessary to impose further priors to find robust solutions for both $C$ and $\mathbf{x}$. The TEM-derived histogram shows that $P(N_p)$ qualitatively resembles a log-normal distribution so we define an additional prior to promote this distribution shape

$$P_{dist}(\mathbf{x}) = \exp\left[-\alpha \left\| \mathbf{x} - \mathbf{x}_{dist}(\boldsymbol{\theta}^*) \right\|_2^2\right] \tag{7}$$

where $\boldsymbol{\theta}^*$ contains the parameters of the lognormal distribution that best matches the current estimate of $\mathbf{x}$, found by minimizing the Kolmogorov-Smirnov goodness-of-fit statistic, and $\mathbf{x}_{dist}(\boldsymbol{\theta}^*)$ is that distribution evaluated at the same discrete $N_p$ values that correspond to the elements of $\mathbf{x}$. Combining this prior with Eqs. (3-5), defining $\gamma^2 = 2\sigma^2\alpha$, and taking the logarithm shows that $P(\mathbf{x}|\mathbf{b})$ is now maximized by

$$\mathbf{x}^* = \arg\min f(\mathbf{x}) = \arg\min \left\| \begin{bmatrix} \mathbf{A} \\ \lambda \mathbf{L} \\ \gamma \mathbf{I} \end{bmatrix} x - \begin{bmatrix} \mathbf{b} \\ 0 \\ \gamma \mathbf{x}_{dist}(\boldsymbol{\theta}^*) \end{bmatrix} \right\|_2^2 \quad s.t.\ x \geq 0 \tag{8}$$

A distinct solution for $C^*$ emerges for $\gamma > 1$, although $f(\mathbf{x})$ increases with $\gamma$ because the distribution prior, Eq. (7), contradicts the smoothness prior, Eq. (4). Figure 4 (*a*) shows that the recovered solution does indeed match the TEM-derived histogram, and substituting $\mathbf{x}^*$ and $C^*$ back into Eq. (1) recovers the experimentally-observed data as shown in Fig. 4 (*b*).

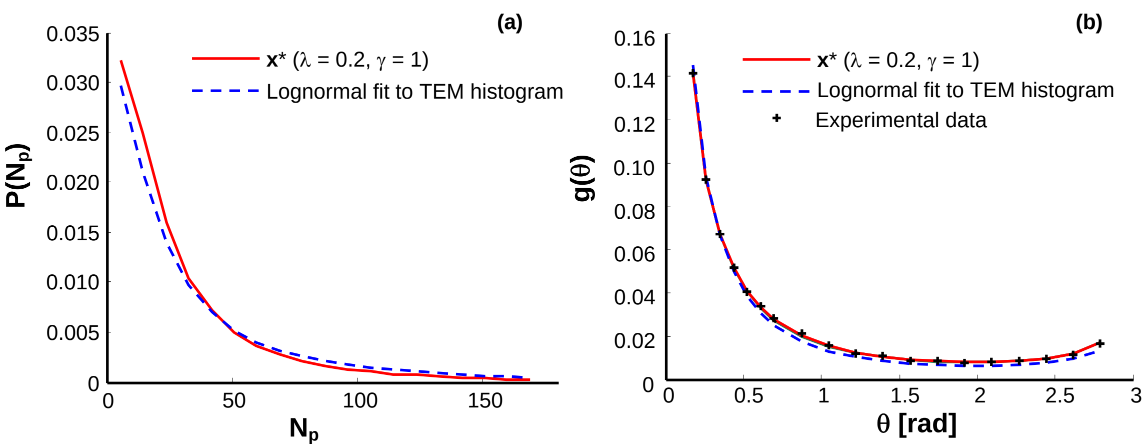


**Figure 4.** (*a*) MAP analysis of experimental data with smoothness, nonnegativity, and distribution priors recovers a distribution that closely matches the TEM histogram; substituting both distributions back into Eq. (1) (*b*) reproduces the experimental data.


[1] Sorensen, C. M., *Aerosol Sci. Tech.*, **30** (1999) 1379-1389.
[2] Dobbins, R. A., and Megaridis, C. M., *Applied Optics*, **30** (1991) 4747-4754.
[3] Tikhonov, A. N., and Arsenin, Y. V., *Solutions of Ill-Posed Problems*, V. H. Winston and Sons, Washington DC, 1977.
[4] Hansen, P. C., *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia PA, 1998.
[5] Tian, K., Thomson, K. A., Liu, F., Snelling, D. R., Smallwood G. J., and Wang, D., *Combustion and Flame*, **144** (2006) 782-791.
[6] Link, O., Snelling, D. R., Thomson, K. A., and Smallwood, G. J., "Development of Absolute Intensity Multi-Angle Light Scattering for the Determination of Polydisperse Soot Aggregate Properties," *Proc. 33rd International Combustion Symposium*, Beijing China, August 1-6, 2010.