Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Bayesian approach to the inverse problem in a light scattering application

Fernando A. Otero[ab], Helcio R. Barreto Orlande[c], Gloria L. Frontini[ab] & Guillermo E. Eliçabe[a]

[a] Institute of Materials Science and Technology (INTEMA), University of Mar del Plata and National Research Council (CONICET), J. B. Justo 4302, 7600 Mar del Plata, Argentina

[b] Department of Mathematics, College of Engineering, University of Mar del Plata, J. B. Justo 4302, 7600 Mar del Plata, Argentina

[c] PEM/COPPE-UFRJ, Universidade Federal de Rio de Janeiro, Caixa Postal 68503, Rio de Janeiro 21941-972, Brazil

Published online: 23 Dec 2014.

To cite this article: Fernando A. Otero, Helcio R. Barreto Orlande, Gloria L. Frontini & Guillermo E. Eliçabe (2015) Bayesian approach to the inverse problem in a light scattering application, Journal of Applied Statistics, 42:5, 994-1016, DOI: 10.1080/02664763.2014.993370

To link to this article: http://dx.doi.org/10.1080/02664763.2014.993370

Journal of Applied Statistics, 2015
Vol. 42, No. 5, 994–1016, http://dx.doi.org/10.1080/02664763.2014.993370

Taylor & Francis
Taylor & Francis Group

Bayesian approach to the inverse problem in a light scattering application

Fernando A. Otero[a,b], Helcio R. Barreto Orlande[c], Gloria L. Frontini[a,b] and Guillermo E. Eliçabe[a*]

[a]*Institute of Materials Science and Technology (INTEMA), University of Mar del Plata and National Research Council (CONICET), J. B. Justo 4302, 7600 Mar del Plata, Argentina;* [b]*Department of Mathematics, College of Engineering, University of Mar del Plata, J. B. Justo 4302, 7600 Mar del Plata, Argentina;* [c]*PEM/COPPE-UFRJ, Universidade Federal de Rio de Janeiro, Caixa Postal 68503, Rio de Janeiro 21941-972, Brazil*

(*Received 11 August 2014; accepted 26 November 2014*)

In this article, static light scattering (SLS) measurements are processed to estimate the particle size distribution of particle systems incorporating prior information obtained from an alternative experimental technique: scanning electron microscopy (SEM). For this purpose we propose two Bayesian schemes (one parametric and another non-parametric) to solve the stated light scattering problem and take advantage of the obtained results to summarize some features of the Bayesian approach within the context of inverse problems. The features presented in this article include the improvement of the results when some useful prior information from an alternative experiment is considered instead of a non-informative prior as it occurs in a deterministic maximum likelihood estimation. This improvement will be shown in terms of accuracy and precision in the corresponding results and also in terms of minimizing the effect of multiple minima by including significant information in the optimization. Both Bayesian schemes are implemented using Markov Chain Monte Carlo methods. They have been developed on the basis of the Metropolis–Hastings (MH) algorithm using Matlab® and are tested with the analysis of simulated and experimental examples of concentrated and semi-concentrated particles. In the simulated examples, SLS measurements were generated using a rigorous model, while the inversion stage was solved using an approximate model in both schemes and also using the rigorous model in the parametric scheme. Priors from SEM micrographs were also simulated and experimented, where the simulated ones were obtained using a Monte Carlo routine. In addition to the presentation of these features of the Bayesian approach, some other topics will be discussed, such as regularization and some implementation issues of the proposed schemes, among which we remark the selection of the parameters used in the MH algorithm.

Keywords: Bayesian estimation; particle size distribution; inverse problem; Metropolis–Hastings; static light scattering

*Corresponding author. Email: elicabe@fi.mdp.edu.ar

1. Introduction

Inverse problems arise in many fields of science and engineering. Typical examples include tomography [16,21], geophysics [23], heat transfer [1,22] and image processing [35] among many others. The ill-posedness of inverse problems requires an effective use of some sort of process to introduce additional information. Traditionally, classical deterministic approaches such as regularized least-squares using initially Tikhonov regularization [31] and more recently total variation regularization [32] have been used to obtain a single solution to the problem. However, instead of obtaining a unique solution, it might be more valuable to achieve a confidence interval for the model parameters. In this case, it is worth recalling that once we have regularized a least-squares problem, we lose the ability to obtain statistically valid confidence intervals for the parameters. This is because regularization brings about the appearance of a bias in the solution [2]. A different point of view is developed in the Bayesian approach, in it all the parameters and variables of the models are modeled as random variables and the solution of the inverse problem is a probability distribution of these model parameters. This approach makes the calculation of the respective confidence intervals a straightforward task. Another characteristic of the Bayesian approach is the inclusion of additional information as a prior probability distribution of the unknown variables. However, as Lopes and Tobias stated in [20], to say that the Bayesian approach is unique or clearly differentiated from the classical approach to inference because of its reliance on prior information is misleading. In fact, from a Bayesian point of view, many classical regularization techniques are equivalent to imposing certain prior distributions on model parameters and, furthermore, the Bayesian approach can be effectively used for the selection of the so-called regularization parameter as can be seen in some articles [19,36]. It is also important to remark that although the Bayesian approach allows us to compute confidence intervals, single values for the model parameters can also be obtained, such as the mode of the posterior distribution, that is, the maximum a posteriori (MAP), and the median which has been suggested as a robust estimator in the bibliography [28,29]. Fraley and Raftery showed in [8] the advantage of computing the MAP solution instead of the traditional maximum likelihood estimation (MLE) in order to avoid singularities or degeneracies. Finally, a fourth advantage of the Bayesian approach is that it explicitly addresses the uncertainty in the model.

Algorithms for solving the inverse problems whether they are based on a Bayesian approach or not, have two main issues to be considered: robustness and efficiency. Even when these attributes are problem-specific, we may understand by robustness of the algorithm its ability to attain the best estimate of the model parameters starting with much wider guess values of them. Efficiency, on the other hand, is attributed to the degree of compute-intensiveness [27]. In this sense, most of the studies using the Bayesian approach to inverse problems are related to the Bayesian robustness, that is, the sensitivity of Bayesian answers to the uncertain inputs: the model, the prior distribution or some combination thereof [3]. In particular, the selection of the prior distribution is a matter of continuous debate. In fact, one of the major critics to the Bayesian approach is due to de Finetti's subjectivist conception of Bayesianism which has led to the personal choice of prior. However, there also have been several tries to formalize the choice of priors. For instance, non-informative (NI) priors were considered to yield robust answers since the time of Laplace, but it was after a century later when formal rules were introduced for its construction as it can be seen in the excellent review by Kass and Wasserman in reference [18]. On the other side, there are many times when the state of knowledge before the data recollection is far from being a state of ignorance and some robustness has to be sacrificed in order to effectively use this prior information. In this sense, one of the biggest challenges in statistical inversion theory is the design of methods that incorporate all prior information [17]. Furthermore, it has been seen that for data with a high signal-to-noise ratio, a Bayesian analysis can frequently yield many orders

of magnitude improvement in model parameter estimation, through the incorporation of relevant prior information [13].

In this paper many of the features described above will be shown when we examine a particular inverse problem in the area of static light scattering (SLS). Several articles have been written using a Bayesian approach to estimate the particle size distribution (PSD) of different materials [7,37]. A prior distribution acting as a regularization term will be seen as mentioned in a previous paragraph and corresponding results will show how the Bayesian approach can help to solve a multiple local minima problem. An alternative procedure will be detailed also for the consideration of model errors, equivalent to the one presented in [17]. Finally, the potentiality of the Bayesian approach in the improvement in model parameter estimation will be presented during a comparison of the results between a NI prior and an informative prior obtained from a previous experiment.

2. The inverse problem of estimating the PSD using SLS

2.1 *General description*

There are numerous materials appearing in diverse industrial applications composed of particles either suspended in a fluid or embedded in a solid matrix. In many cases, these particle systems can be characterized through their PSD's. The estimation of the PSD can be achieved using different particle sizing techniques. Optical techniques stand out because they are non-destructive. One of the most important optical techniques is the so-called SLS which has been widely used in practice for particles from about 50 nm to several micrometers [12]. The basic process consists in illuminating a sample of particles by a laser light and measuring the intensity of the light scattered at different angles.

Estimation of the PSD using SLS measurements requires solving an ill-conditioned inverse problem using a predefined model. In this article, the estimated PSD corresponds to a number-PSD. The inverse problem can be formulated as estimating the PSD represented by a function $f(R)$ (considering that particles are spherical and therefore the PSD can be expressed as a function of a radius R) by using the intensity of light scattered as data, noted by $I_s(q)$, where the variable q makes reference to the magnitude of the scattering vector and it is computed as

$$q = \frac{4\pi n_m}{\lambda_0} \sin\frac{\theta}{2}, \tag{1}$$

where n_m is the refractive index of the suspending media, λ_0 is the wavelength of the incident light in vacuum and θ is the scattering angle.

When the inverse problem is solved and the used model makes it possible, also the volume fraction of particles, η, can be estimated, and in some cases, previous information about the statistical parameters of the PSD can be incorporated in the estimation process. For these cases, the inverse problem is solved by combining the SLS measurements with the results of scanning electron microscopy (SEM) which brings the previous information mentioned above through the analysis of micrographs of the studied systems. Examples of these micrographs can be seen in Figure 1(a) and 1(b), while the corresponding SLS intensities are shown in Figure 1(c) and 1(d). These examples, labeled as 30PIB5 and 50PIB25, correspond to samples of polymer particles embedded in a solid polymer matrix. The details of these samples will be given in a later section. It must also be pointed out that in the process of solving the inverse problem other parameters proper to the corresponding used models must also be estimated as we will see in the following section.

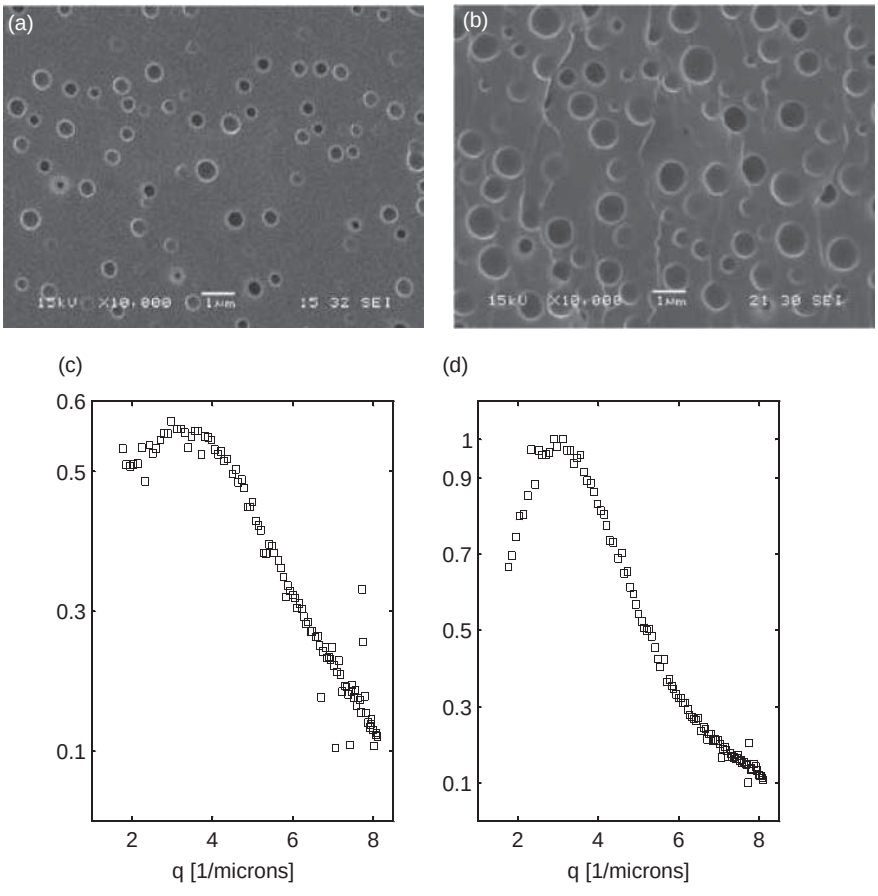

Figure 1. SEM micrographs for the experimental examples: (a) 30PIB5, and (b) 50PIB25; and their corresponding SLS measurements: (c) 30PIB5, and (d) 50PIB25.

2.2 *SLS models*

Estimation of the PSD using SLS measurements requires solving an ill-conditioned inverse problem using a predefined model. The well-known model derived under the so-called Mie regime for one particle can be straightforwardly extended to a system of many particles only if the assumption of independent single scattering is valid, which means that the system is diluted [5,15]. However, for concentrated particle systems which cannot be controlled by dilution, approximated models based on the Rayleigh–Gans theory have been developed [4,14]. In the works of Vrij [33,34] a model is formulated for finite mixtures of particles treated as impenetrable spheres which cannot be overlapped in space, that is, hard spheres (HS). Vrij's HS model provides an analytical solution to the SLS problem of a concentrated system of spherical particles in the so-called Percus–Yevick approximation. Vrij's HS model provides a matrix form for the calculation of the scattered intensities. In this model, the observable magnitude is the differential scattering cross section (DSCS), proportional to the intensity of the scattered light $I_s(q)$

$$\frac{d\sum(q)}{d\Omega} = \frac{I_s(q)A_d}{E_0^2(A_d/D^2)} = K_i I_s(q), \tag{2}$$

where $\mathrm{d}\sum(q)/\mathrm{d}\Omega$is the DSCS per unit volume, E_0 is the intensity of the incident electric field, D is the distance from the sample to the detector, A_d is the area of the detector and K_i is an instrumental constant which merges all involved parameters.

The complete formulation for Vrij's HS model can be seen in Appendix 1. Estimated parameters on this model are in addition to the PSD, the volume fraction η, a proportionality constant C between particle radii and HS radii and a global constant K', a constant which is reciprocal to K_i.

Vrij's HS model constitutes an exact solution when particles can be modeled by 'hard spheres'. However, this model, which has a strong nonlinear dependence on the PSD, has several problems, including high error propagation and a very time-consuming computational implementation when used in an inverse problem. Thus, some approximate models were built to be used instead. Pedersen [24,25] developed the local monodisperse approximation (LMA) assuming that the particles are spatially distributed according to their sizes. Under this assumption the scattered light intensity becomes mathematically related in a linear fashion to the PSD and can be written as

$$I_\mathrm{s}(q) = K \int_0^\infty f(R) S(p, q, R) F^2(q, R)\, \mathrm{d}R, \tag{3}$$

where K is a global constant, $f(R)$ is the PSD and $F^2(q, R)$ is the form factor of each particle of radius R. The factor $S(p, q, R)$ is the so-called structure factor where p is an effective parameter of the model, equivalent to the volume fraction in monodisperse systems but losing its physical meaning as polydispersity increases. Analytical expressions for $F(q, R)$ and $S(p, q, R)$ used in this article correspond to spherical particles and they are described from Equations (A9)–(A14) in Appendix 2.

Parameters to be estimated in the LMA are besides the PSD, the effective parameter p and the global constant K.

3. Description of the developed Bayesian schemes

Given SLS measurements (simulated or experimental) as the problem data, the inverse problem pursuits the objective of determining the PSD and also sometimes the volume fraction of particles. In order to make this information retrieval possible, all the remaining unknowns involved in each considered model must also be estimated. The estimated unknowns depend on both the model and the scheme which is used. Once a Bayesian approach has been developed these unknowns are defined by means of their probability density functions (pdfs). In this work, parametric and non-parametric Bayesian schemes (NPBSs) are implemented through the Metropolis–Hastings (MH) algorithm. Although this algorithm has a relatively simple implementation, its correct performance is related to the correct selection of several parameters. In the same way, even when the Bayes theorem is the keystone of the Bayesian approach, explicit forms of this theorem and its specific application differ in every developed scheme explained below.

3.1 *Parametric Bayesian scheme*

The first proposed scheme is applied to both SLS models described above and is called the parametric Bayesian scheme (PBS) because it makes use of a parametrical distribution to represent the PSD. In this article, a log-normal distribution (with PSD mean radius $\bar{R}$ and standard deviation) was utilized because it allows a relative skewness of the distribution. It is important to point out that under this scheme, there is no need of an additional regularization term for the PSD since its representation in a parametrical family converges already to a stable solution. In the PBS, the

Bayes theorem is applied in the form of Equation (4)

$$\pi_{\text{posterior}}(\mathbf{P}) = \pi\,(\mathbf{P}/I_\varepsilon) = \frac{\pi_{\text{prior}}(\mathbf{P})\pi\,(I_\varepsilon/\mathbf{P})}{\pi\,(I_\varepsilon)}, \tag{4}$$

where $\pi_{\text{posterior}}(\mathbf{P})$ is the so-called posterior pdf, $\pi_{\text{prior}}(\mathbf{P})$ is the prior pdf (which refers to the parameters information previous to measurements), $\pi\,(I_\varepsilon/\mathbf{P})$ is the likelihood function and $\pi\,(I_\varepsilon)$ is the measurements pdf where I_ε corresponds to noisy scattered light intensities (simulated or experimental) while $\mathbf{P}$ represents a vector of estimated parameters. When Vrij's HS model is used in the inversion stage, the vector of parameters is given by $\mathbf{P} = [\bar{R}, \sigma, \eta, C, K']$. Where parameter C appears as a result of considering that the 'hard spheres' radius can be modeled as proportional to the particle radius by a proportionality factor named C; and parameter K' is obtained as $K' = 1/K_i$ since K' appears as a proportionality factor between the scattered intensities and the DSCS in the form of Equation (5):

$$I_{\text{s}}(q) = \frac{1}{K_i}\frac{\text{d}\sum(q)}{\text{d}\Omega} = K'\frac{\text{d}\sum(q)}{\text{d}\Omega} = -K'D(q)\Delta(q)^{-1}. \tag{5}$$

When the LMA model is used instead, the corresponding parameter vector is given by $\mathbf{P} = [\bar{R}, \sigma, p, K]$.

All prior information available before the SLS measurements is incorporated into $\pi_{\text{prior}}(\mathbf{P})$. In this work, NI and partially informative priors were used and statistical independence and normality in informative priors are considered as it will be explained in the prior information subsection inside the implementation section.

A simplified analytical expression can be derived for the likelihood function $\pi\,(I_\varepsilon/\mathbf{P})$ when measurements are modeled as a discrete variable and noise as a normal random variable:

$$\pi\,(I_\varepsilon/\mathbf{P}) = \frac{1}{(\sqrt{2\pi})^M\sqrt{|W|^{-1}}}\exp\left\{-\frac{1}{2}[I_\varepsilon - I_{\text{s}}(P)]^{\text{T}}W[I_\varepsilon - I_{\text{s}}(P)]\right\}, \tag{6}$$

where $\mathbf{I}_{\text{s}}(\mathbf{P}) = [I_{\text{s}}(q_1, \mathbf{P})I_{\text{s}}(q_2, \mathbf{P})\cdots I_{\text{s}}(q_{\text{M}}, \mathbf{P})]^{\text{T}}$, M is the number of measurements, $\mathbf{W}$ is the inverse of the covariance matrix of the measurements and $I_{\text{s}}(q_i, \mathbf{P})$ are SLS intensities generated by the considered model according to Equations (3) and (A9)–(A14) or Equations (A1)–(A7) for the LMA model and Vrij's HS model, respectively.

It is important to explain a particular consideration involving Equation (6). At this point it seems important to remark differences between treatments with Vrij's HS model and the LMA model. In cases where Vrij's HS model has been used or when the LMA model has been used where modeling errors can be neglected (i.e. low concentration or diluted or quasi-monodisperse particle systems), light scattering measurements can be modeled as

$$I_\varepsilon(q_i) = I_{\text{s}}(q_i, \mathbf{P}) + \varepsilon(q_i), \tag{7}$$

where $I_{\text{s}}(q_i, \mathbf{P})$ is the scattered intensity generated according to the corresponding model, and $\varepsilon(q_i)$ is a gaussian additive noise. However, when the LMA model is used and modeling errors cannot be neglected, then we can approximate light scattering measurements as

$$I_\varepsilon(q_i) = I_{\text{s}}(q_i, \mathbf{P}) + E(q_i, \mathbf{P}) + \varepsilon(q_i), \tag{8}$$

where $E(q_i, \mathbf{P})$corresponds now to the errors between LMA model and Vrij's HS model once again for some set of parameters $\mathbf{P}$. This approach for the approximation error was introduced in [17].

Matrix **W** of Equation (6) must be estimated before solving the inverse problem for some previously obtained set of parameters $\mathbf{P} = \mathbf{P_0}$. If measurements are modeled as in Equation (7), **W** should be estimated from the error between $\mathbf{I}_\varepsilon$ and $\mathbf{I}_s(\mathbf{P})$ for $\mathbf{P} = \mathbf{P_0}$. On the other side, if measurements are modeled as in Equation (8), it is also necessary to add the errors between the LMA and Vrij's HS models taking a predefined set of parameters $\mathbf{P_0}$. In the developed implementation, the predefined set $\mathbf{P_0}$ used for the estimation of **W** was computed from a previously performed deterministic least-squares estimation.

3.2 *Non-parametric Bayesian scheme*

The NPBS has only been developed for the LMA model since it requires an important amount of computational resources and the use of the analytical Vrij's HS model would imply an excessive cost. In the NPBS scheme the corresponding weights to each particle size proportional to the PSD are estimated. The global constant K is absorbed in these weights. The corresponding PSD can be obtained from a simple normalization of these weights and it can be represented as a vector $\mathbf{f} = [f(R_1), f(R_2), \ldots, f(R_N)]$ where the respective particle system is composed of N different sizes. It is convenient that, as it will be presented later, the total number of particles N_p is also included in the estimation process. It can be noticed that the mathematical model described in Equation (3) displays a linear relation between $f(R)$ and $I_s(q)$ for a given value of parameter p, which is limited to the interval [0,1]. Taking advantage of this situation, an alternative version of Bayes theorem can be written as

$$\pi(\mathbf{f}/\mathbf{I}_\varepsilon, p) = \frac{\pi(\mathbf{I}_\varepsilon/\mathbf{f}, p)\pi(\mathbf{f}/p)}{\int_f \pi(\mathbf{I}_\varepsilon/\mathbf{f}, p)\pi(\mathbf{f}/p)\mathrm{d}\mathbf{f}}, \tag{9}$$

where $\pi(\mathbf{f}/\mathbf{I}_\varepsilon, p)$ is the conditional pdf for the PSD given the measurements $\mathbf{I}_\varepsilon$ and the parameter p and $\pi(\mathbf{f}/p)$ is the prior pdf for the PSD given the parameter p; $\pi(\mathbf{I}_\varepsilon/\mathbf{f}, p)$ is the likelihood function which is similar to Equation (6) and the integral is a normalizing factor.

When no assumption on the shape of $f(R)$ is considered, as in this case, an explicit regularization is needed. This regularization is carried out by including in the prior pdf the so-called smoothness condition defined as the following pdf:

$$\pi_s(\mathbf{f}) = \frac{1}{(\sqrt{2\pi})^N} \exp\left(-\frac{1}{2}\gamma \mathbf{f}^T \mathbf{H}^T \mathbf{H}\mathbf{f}\right), \tag{10}$$

where N is the number of components of the discretized PSD, **H** is a regularization matrix and γ is an adjustable regularization parameter.

It is also worth to say that besides the regularization term, any additional prior information related to the PSD statistical parameters can be included in $\pi(\mathbf{f}/p)$ as in the PBS. This will be explained later in the implementation section of the NPBS when the form will be detailed.

Finally, an iterative Bayesian method (IBM) using the Bayes theorem in the form of Equation (9) is proposed to solve the NPBS. It can be described in the following steps:

(1) Choose a value for the parameter p and a seed function for the PSD.
(2) Compute a value for γ using some method such as generalized cross validation (GCV), L-curve (LC) and the principle of discrepancy (PD).
(3) Estimate the pdf of $f(R)$, applying MH algorithm in the form of Equation (9).
(4) Repeat the first three steps for the whole range of possible values of p.
(5) Select the pdf of $f(R)$with the maximum product.

4. Implementation

4.1 *MH algorithm*

The PBS as well as the NBPS have been implemented using the MH algorithm; its choice is based on the versatility shown in many applications. It also allows us to create a simplified program with the assumed considerations. The goal of the MH algorithm is to build an adequate Markov chain $\mathbf{X}$ which simulates a distribution that has a density $\pi(\mathbf{P})$, defining a density $q_{\mathrm{T}}(\mathbf{P}^{(t+1)}/\mathbf{P}^{(t)})$, called candidate-generating density, between a state of the Markov chain at a time t (i.e. a value of the set of parameters $\mathbf{P}^{(t)}$) and the next one $\mathbf{P}^{(t+1)}$. Then, the MH algorithm is defined by two steps: a first step in which a proposed value is drawn from the candidate-generating density and a second step in which the proposed value $\mathbf{P}^*$ is accepted as the next iterate in the Markov chain according to the probability $\alpha_{\mathrm{MH}}(\mathbf{P}^*/\mathbf{P}^{(t)})$ defined as

$$\alpha_{\mathrm{MH}}\left(\mathbf{P}^*/\mathbf{P}^{(t)}\right) = min\left(\frac{\pi(\mathbf{P}^*/\mathbf{I}_\varepsilon)q_{\mathrm{T}}(\mathbf{P}^*/\mathbf{P}^{(t)})}{\pi(\mathbf{P}^{(t)}/\mathbf{I}_\varepsilon)q_{\mathrm{T}}(\mathbf{P}^{(t)}/\mathbf{P}^*)}, 1\right). \tag{11}$$

If $\alpha_{\mathrm{MH}}(\mathbf{P}^*/\mathbf{P}^{(t)}) > \upsilon$, where υ is a random variable of uniform distribution $U[0,1]$, then the drawn sample $\mathbf{P}^{(t+1)}$ is accepted and $\mathbf{P}^* = \mathbf{P}^{(t+1)}$; if $\alpha_{\mathrm{MH}}(\mathbf{P}^*/\mathbf{P}^{(t)}) \leq \upsilon$, $\mathbf{P}^{(t+1)}$ is rejected and $\mathbf{P}^* = \mathbf{P}^{(t)}$. An important aspect of the algorithm is that the acceptance probability given by $\alpha_{\mathrm{MH}}(\mathbf{P}^*/\mathbf{P}^{(t)})$ avoids the computation of $\pi(\mathbf{I}_\varepsilon)$. A detailed analysis of the MH algorithm can be found elsewhere in the literature [6]. As it was mentioned before, a correct performance of the MH algorithm requires an appropriate selection of its own parameters. Such parameters include initial sample, acceptance ratio, candidate-generating density, total length of the Markov chain and samples considered and discarded in this chain. For a full discussion on these parameters see also [6].

Convergence tests have been performed for different initial samples in order to determine whether differences between obtained chains are significant. These tests must determine also if the influence of the initial samples is reduced as long as the chains grow. Furthermore, the MH algorithm cannot warranty the solution when multiple local minima problems are present. An alternative solution is to slightly modify the algorithm by including concepts from the simulated annealing (SA) algorithm [11]. In this case the parameter $\alpha_{\mathrm{MH}}(\mathbf{P}^*/\mathbf{P}^{(t)})$ of Equation (11) is changed to

$$\alpha_{\mathrm{MH}}\left(\mathbf{P}^*/\mathbf{P}^{(t)}\right) = \min\left(\left[\frac{\pi(\mathbf{P}^*/\mathbf{I}_\varepsilon)}{\pi(\mathbf{P}^{(t)}/\mathbf{I}_\varepsilon)}\right]^t \frac{q_{\mathrm{T}}(\mathbf{P}^*/\mathbf{P}^{(t)})}{q_{\mathrm{T}}(\mathbf{P}^{(t)}/\mathbf{P}^*)}, 1\right), \tag{12}$$

which accelerates, for convergence purposes, the achievement of the maximum of the distribution, that is, the MAP solution. After the MAP solution is reached using Equation (A9) when the parameter vector $\mathbf{P}$ converges, a new application of the MH algorithm using Equation (A8) is performed starting from this MAP solution for computing the respective confidence intervals.

Acceptance ratio is the percentage of times a new sample is accepted; so, if a too small value is selected, then, the chain slowly covers (or even does not cover) the whole distribution. However, choosing a too large value produces jumps very often. Best results in all cases showed that acceptance ratio should be set between 0.23 and 0.33, in agreement with a previous work [10].

A candidate-generating density must also be specified, typically from a family of distributions with tuning parameters such as scale and location. The corresponding implemented densities will be detailed in the following subsections. Finally, the length of the generated chain must be large enough to be a good approximation to the final distribution and first samples until reaching stationary regime should be discarded. These specific aspects related to both schemes, PBS and NPBS, will be discussed in the implementation sections of each developed scheme.

4.2 *Prior information*

Selection of a prior distribution is a critical and controversial factor in the Bayesian approach and depends on the state of knowledge at the time of measurements. In this work, we consider non-informative priors and partially informative priors. We include in the partially informative prior distributions based on alternative SEM experiments to study their effects on estimations. In this context, SEM micrographs were obtained from experiments or simulations. SEM simulations have been generated with a Monte Carlo routine. Corresponding prior densities for both types of micrographs (simulated and experimental) for $\bar{R}$ and σ were approximated to be normal-shaped according to the large-sample confidence interval for a size of sample $N_s \geq 30$ [9]. As a consequence of this, $\bar{R}$ has been described by a normal density where its mean corresponds to the sample mean $\mu_{\bar{R}}$ of the mean radius and its standard deviation is $s/\sqrt{N_s}$ where s is the sample standard deviation of the PSD, which leads to

$$\pi_{\text{prior}}(\bar{R}) = \frac{1}{(\sqrt{2\pi})} \exp\left\{-\frac{1}{2}[\bar{R} - \mu_{\bar{R}}]^{\mathrm{T}} \frac{N_s}{s^2}[\bar{R} - \mu_{\bar{R}}]\right\}. \tag{13}$$

In a similar manner, σ can be described by

$$\pi_{\text{prior}}(\sigma) = \frac{1}{(\sqrt{2\pi})} \exp\left\{-\frac{1}{2}[\sigma - \mu_{\sigma}]^{\mathrm{T}} \frac{2N_s}{s^2}[\sigma - \mu_{\sigma}]\right\}, \tag{14}$$

where μ_{σ} is the sample mean of the PSD standard deviation.

The prior information used for the other involved parameters will be detailed in the next sections for each developed scheme. In all cases we have made the assumption that priors for each parameter are statistically independent.

4.3 *Parametric Bayesian scheme*

Implementation of the PBS in the MH algorithm is a straightforward task once all the MH parameters are selected. As it was specified in the general description of the implementation, a modification of the algorithm as expressed by Equation (12) was necessary in order to improve convergence to the MAP and avoid local minima. $\pi_{\text{prior}}(\mathbf{P})$ has been built as an independent combination of priors for each parameter. For both models prior information about PSD statistical parameters is presented as the normal densities of Equations (13) and (14). In addition to this, in Vrij's HS model, parameter η has uniform distribution in the range [0,1] according to all possible values, while parameter C has been chosen to follow an uniform distribution in the range [1,3], based on its physical meaning and according to some previous experimental results, and parameter K' has a NI prior. On the other side, in the LMA model, parameter p has a uniform distribution in the range [0,1] and K has a NI prior. The candidate-generating density follows an independent generation for parameters $\bar{R}$ and σ, while the so-called random walk was chosen for the rest of the parameters in each model. After studying simulated examples we conclude that around 20,000 samples must be taken to reach the stationary regime and that the length of the chain which is created to build the final distribution must be of the order of 50,000 samples.

4.4 *Non-parametric Bayesian scheme*

In general, in the studied examples the NPBS has shown worse convergence problems than the PBS because of the large quantity of variables involved in the iteration process. This situation requires the use of the SA algorithm as explained before, where in this case Equation (12) takes

the following form:

$$\alpha_{\mathrm{MH}}(\mathbf{f}^*/\mathbf{f}^{(t)}) = \min\left(\left[\frac{\pi(\mathbf{f}^*/\mathbf{I}_\varepsilon, p)}{\pi(\mathbf{f}^{(t)}/\mathbf{I}_\varepsilon, p)}\right]^t \frac{q_T(\mathbf{f}^*/\mathbf{f}^{(t)})}{q_T(\mathbf{f}^{(t)}/\mathbf{f}^*)}, 1\right). \tag{15}$$

The prior density involved in the NBPS, $\pi(\mathbf{f}/p)$, can be expressed just as the smoothness condition, $\pi_s(\mathbf{f})$, when no prior information is included, and as a product of $\pi_s(\mathbf{f})$ and the priors for mean radius $\pi_{\mathrm{prior}}(\bar{R})$ and for PSD standard deviation $\pi_{\mathrm{prior}}(\sigma)$ when information from the micrographs is used. The computation of these two last densities $\pi_{\mathrm{prior}}(\bar{R})$ and $\pi_{\mathrm{prior}}(\sigma)$ for every PSD computed at each iteration in the IBM is performed using the corresponding number density, since

$$\bar{R} = \frac{1}{N}\sum_{k=1}^{N} N_p^{(k)} R^{(k)}, \tag{16}$$

$$\sigma = \left[\frac{1}{N}\sum_{k=1}^{N} N_p^{(k)} (R^{(k)} - \bar{R})^2\right]^{1/2}. \tag{17}$$

The candidate-generating density used in the NBPS was the random walk. However, this method of generation of samples limits the number of variables as it will be discussed in the next paragraph. Final distributions obtained from simulated examples were generated from 100,000 samples after reaching the stationary regime.

The NPBS has some other aspects which make its implementation more complicated. These include discretization of the PSD, selection of the regularization matrix **H** and selection of a method for computing the regularization parameter γ. The random walk process was used to generate the samples of the chain. In this process, each candidate is generated changing one parameter at a time. However, even when the implementation of this method is very simple, it has performance and convergence problems when there is a large number of components of the PSD. Thus, a maximum number of around 40 points for the PSD were considered to make the problem tractable. A few traditional regularization matrices were studied including the identity matrix, and first and second derivative matrices. Best results in simulations were obtained for the second derivative matrix. The computation of the regularization parameter was an important point to analyze since it has a great influence on the computed solution. In this work three methods were analyzed: GCV, LC and PD. It has been shown that GCV usually produces data overfitting [36] and hence it may have a bad performance on systems were modeling errors are important. LC, on the other hand, produces an over-regularization of the solution loosing a good part of information provided by the data. Finally, PD seemed to bring an intermediate solution. This method requires an estimation of the noise level which has to be according to the modeling of measurements proposed in Equations (7) and (8).

5. Examples, results and discussion

We have selected three simulated examples of particle systems described exactly by a log-normal distribution where all measurements were generated using Vrij's HS model. The first example has been analyzed with the PBS using Vrij's HS model and the two last examples have been analyzed with both schemes using the LMA model. The first example (Example 1) corresponds to a concentrate particle system with parameters $\bar{R} = 0.35\,\mu$m, $\sigma = 0.1\,\mu$m, $\eta = 0.1$, $C = 1.8$ and $K' = 1$. The second example (Example 2) corresponds to a sample with a PSD having $\bar{R} = 0.2\mu$, $\sigma = 0.02\mu$, $\eta = 0.01$, $C = 1$ and $K' = 1$. The third example (Example 3) corresponds to a sample with a PSD of $\bar{R} = 0.35\mu$ and $\sigma = 0.10\mu$, $\eta = 0.03$ and once again $C = 1$ and

$K' = 1$. Three corresponding SEM micrographs of 50, 100 and 40 particles where simulated for Examples 1–3, respectively.

We also have analyzed two experimental examples. Experimental examples were taken from a previously reported publication [30]. In this work, blends of polyisobutilene (PIB) labeled PIB5 and PIB025 in isobornyl methacrylate (IBoMA) were reacted until vitrification. The blends contained 50% of PIB025 (50PIB025) and 30% of PIB5 (30PIB5). During reaction, the PIB is phase separated forming micron-sized particles. Thus, the resulting final sample is made of polymer particles embedded in the solid polymer matrix. Refractive indices were 1.51 for PIB5 and PIB025 and 1.48 for the IBoMA. SLS measurements were performed using a flat cell light scattering apparatus which consists of a linear array of photodiodes that detects the light scattered by a thin sample illuminated by a 17 mW He–Ne laser with random polarization. SEM micrographs were obtained using a Jeol JSM 6460 LV device and they can be seen in Figure 1(a) and 1(b). The example 30PIB5 was used with Vrij's HS model using the PBS, while the example 50PIB25 was used with the LMA model using both schemes.

Corresponding results are presented from Tables 1–6 and from Figures 2–10. We include only the results of the parameters of interest ($\bar{R}$, σ and eventually η) and the corresponding 95% confidence intervals for the simulated SEM micrographs and for the chains obtained using both a NI prior and SEM micrograph information as discussed in the previous section. We also compute the median and the MAP solution. In the presented figures, with the exception of Figure 8 in

Table 1. Results for Example 1 ($\bar{R} = 0.35\mu$ and $\sigma = 0.1\mu$, $\eta = 0.1$, $N_s = 50$) using the PBS.

		SEM	SLS (NI prior)	SLS (SEM prior)
Mean radius	95% C.I.	[0.3331–0.3857]	[0.3031–0.3757]	[0.3280–0.3802]
	Median		0.3397	0.3535
	MAP		0.3402	0.3422
PSD std. deviation	95% C.I.	[0.0762–0.1134]	[0.0831–0.1125]	[0.0890–0.1066]
	Median		0.0978	0.0980
	MAP		0.0971	0.0988
Volume fraction	95% C.I.		[0.0761–0.1055]	[0.0923–0.1045]
	Median		0.0908	0.0973
	MAP		0.0911	0.0932

Table 2. Results for Example 2 ($\bar{R} = 0.2\mu$ and $\sigma = 0.02\mu$, $\eta = 0.01$, $N_s = 100$) using the PBS and NPBS.

		SEM	SLS (NI prior)	SLS (SEM prior)
PBS				
Mean radius	95% C.I.	[0.2004–0.2106]	[0.1875–0.2083]	[0.1941–0.2019]
	Median		0.1976	0.1978
	MAP		0.1984	0.1981
PSD std. deviation	95% C.I.	[0.0183–0.0285]	[0.0121–0.0243]	[0.0167–0.0225]
	Median		0.0183	0.0196
	MAP		0.0174	0.0204
NPBS				
Mean radius	95% C.I.	[0.2004–0.2106]	[0.1856–0.2138]	[0.1956–0.2079]
	Median		0.1972	0.1997
	MAP		0.1840	0.2040
PSD std. deviation	95% C.I.	[0.0183–0.0285]	[0.0129–0.0340]	[0.0158–0.0288]
	Median		0.0208	0.0194
	MAP		0.0238	0.0269

which we have used the MAP values of the parameters, and to improve visualization we only have included PSDs and data fitting computed from the medians of the parameters.

The discussion may start with the analysis of the chosen implementation of the priors. This problem has been solved using NI priors and partially informative priors; in this last case we have specified informative priors on the parameters of interest and NI priors on the nuisance parameters. We have selected normal priors for $\bar{R}$ and σ mainly as we have mentioned before, with the purpose of simplifying the implementation. Nevertheless, this selection of informative priors has robustness problems. In fact, the price to be paid for utilization of inherently robust procedures is computational [3]. As Berger pointed out in [3], standard choices such as the normal prior can lead to the non-robust design in several ways: models can be very sensitive to outliers in the data

Table 3. Results for Example 3 ($\bar{R} = 0.35\mu$ and $\sigma = 0.1\mu$, $\eta = 0.03$, $N_s = 40$) using the PBS and NPBS.

		SEM	SLS (NI prior)	SLS (SEM prior)
PBS				
Mean radius	95% C.I.	[0.3111–0.3805]	[0.3377–0.4019]	[0.3121–0.3749]
	Median		0.3699	0.3439
	MAP		0.3705	0.3636
PSD std. deviation	95% C.I.	[0.0874–0.1364]	[0.0665–0.1049]	[0.0843–0.1027]
	Median		0.0857	0.0941
	MAP		0.0855	0.0853
NPBS				
Mean radius	95% C.I.	[0.3111–0.3805]	[0.3471–0.4291]	[0.3357–0.3655]
	Median		0.3869	0.3496
	MAP		0.4027	0.3425
PSD std. deviation	95% C.I.	[0.0874–0.1364]	[0.1034–0.1300]	[0.0965–0.1149]
	Median		0.1168	0.1050
	MAP		0.1159	0.1013

Table 4. Multiple minima found in 30PIB5 using deterministic MLE.

$\bar{R}$	σ	η	C	K'	Residue
0.2702	0.0622	0.0152	1.84	19.04	0.072
0.2037	0.0788	0.0933	3.24	8.42	0.073
0.3071	0.045	0.0125	1.82	20.14	0.074
0.2318	0.0948	0.2441	2.26	5.25	0.074

Table 5. Results for 30PIB5 using the PBS.

		SEM	SLS (NI prior)	SLS (SEM prior)
Mean radius	95% C.I.	[0.1980–0.2220]	[0.1934–0.3470]	[0.1978–0.2218]
	Median		0.2654	0.2099
	MAP		0.2702	0.2133
PSD std. deviation	95% C.I.	[0.0456–0.0624]	[0.0295–0.0949]	[0.0489–0.0693]
	Median		0.0643	0.0593
	MAP		0.0622	0.0703
Volume fraction	95% C.I.		[0.0054–0.0212]	[0.0076–0.0177]
	Median		0.0134	0.0106
	MAP		0.0152	0.0142

Table 6. Results for 50PIB25 using the PBS and NPBS.

		SEM	SLS (NI prior)	SLS (SEM prior)
PBS				
Mean radius	95% C.I.	[0.2035–0.2565]	[0.2066–0.4587]	[0.2047–0.2576]
	Median		0.2876	0.2301
	MAP		0.3263	0.2734
PSD Std. deviation	95% C.I.	[0.1039–0.1420]	[0.0871–0.1912]	[0.1030–0.1298]
	Median		0.1372	0.1173
	MAP		0.1390	0.1161
NPBS				
Mean radius	95% C.I.	[0.2035–0.2565]	[0.2168–0.4006]	[0.2254–0.2410]
	Median		0.3293	0.2350
	MAP		0.3293	0.2378
PSD Std. deviation	95% C.I.	[0.1039–0.1420]	[0.0963–0.1737]	[0.1070–0.1268]
	Median		0.1438	0.1173
	MAP		0.1430	0.1161

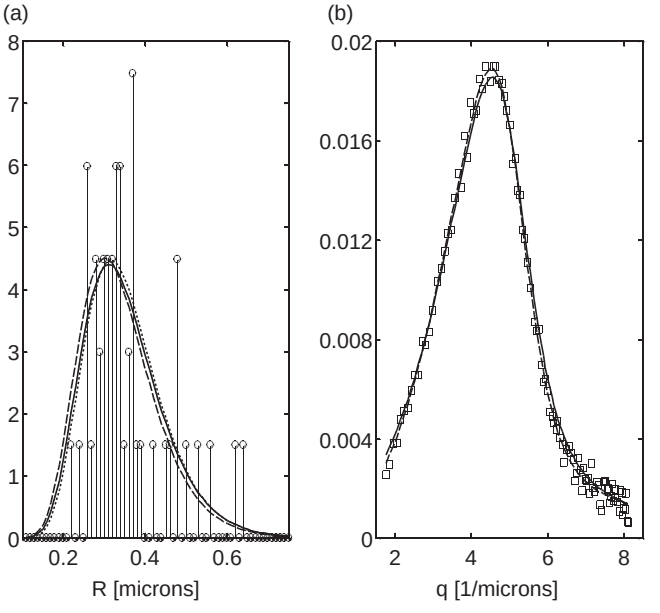

Figure 2. (a) Retrieved PSDs for Example 1 using the PBS and the parameters medians: actual PSD, full line; PSD from SEM simulations, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, dotted line. (b) SLS model fitting to simulated measurements for Example 1 using the PBS and the parameters medians: simulated measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

(this effect will be seen in some results) and conjugate priors can have a pronounced effect on the answers even if data are in conflict with the specified prior information. However, this last is not always a problem, in fact in this work this situation is preferred when the LMA model has been used and modeling errors are significant or when noise in measurements is really important. In this case, when data are being analyzed with an approximate model or prior information is sufficiently strong compared with measurement errors, data fitting does not play the main role in the inverse problem solution.

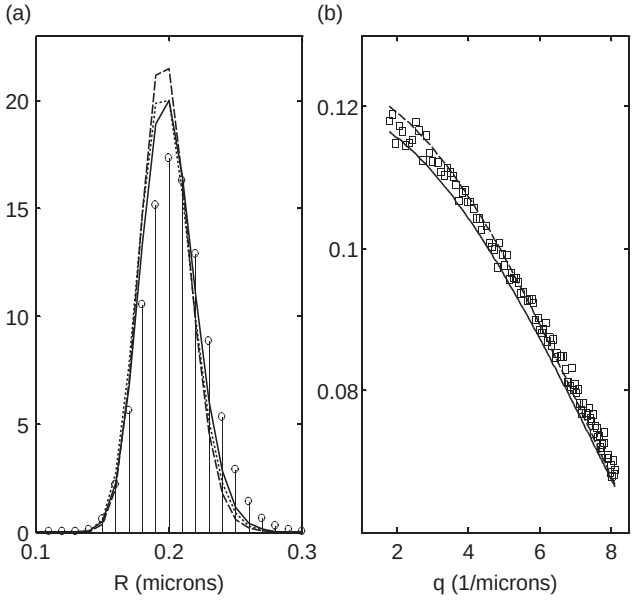

Figure 3. (a) Retrieved PSDs for Example 2 using the PBS and the parameters medians: actual PSD, full line; PSD from SEM simulations, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, dotted line. (b) SLS model fitting to simulated measurements for Example 2 using the PBS and the parameters medians: simulated measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

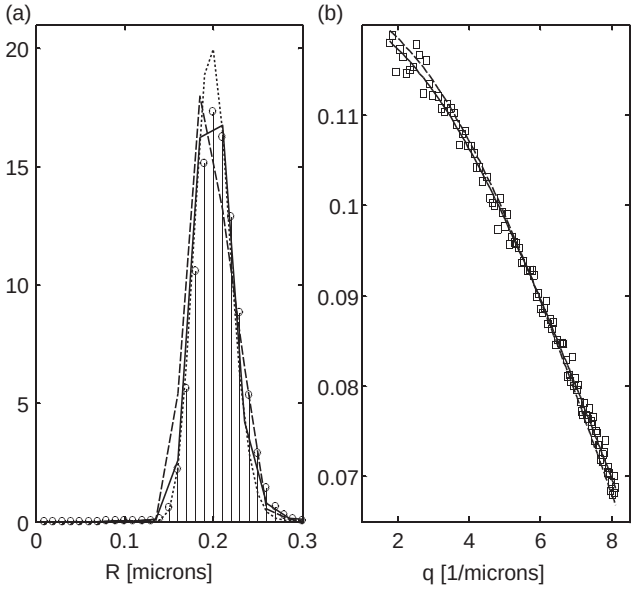

Figure 4. (a) Retrieved PSDs for Example 2 using the NPBS and the PSDs ordinate medians: actual PSD, full line; PSD from SEM simulations, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, dotted line. (b) SLS model fitting to simulated measurements for Example 2 using the NPBS and the PSDs ordinate medians: simulated measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

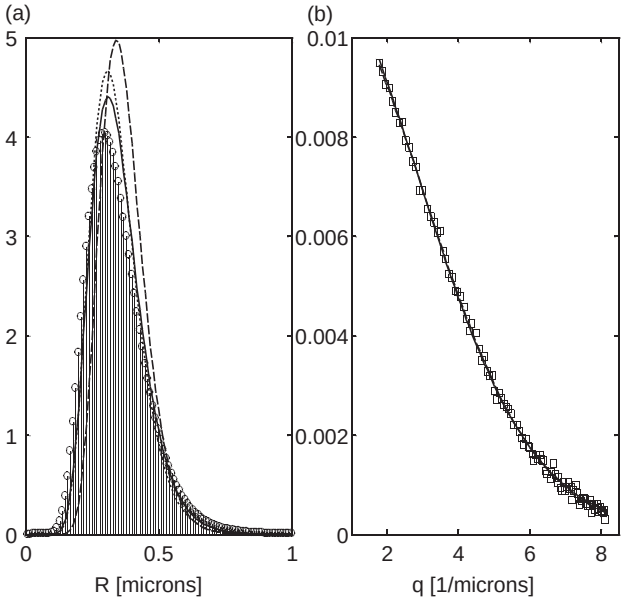

Figure 5. (a) Retrieved PSDs for Example 3 using the PBS and the parameters medians: actual PSD, full line; PSD from SEM simulations, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, dotted line. (b) SLS model fitting to simulated measurements for Example 3 using the PBS and the parameters medians: simulated measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

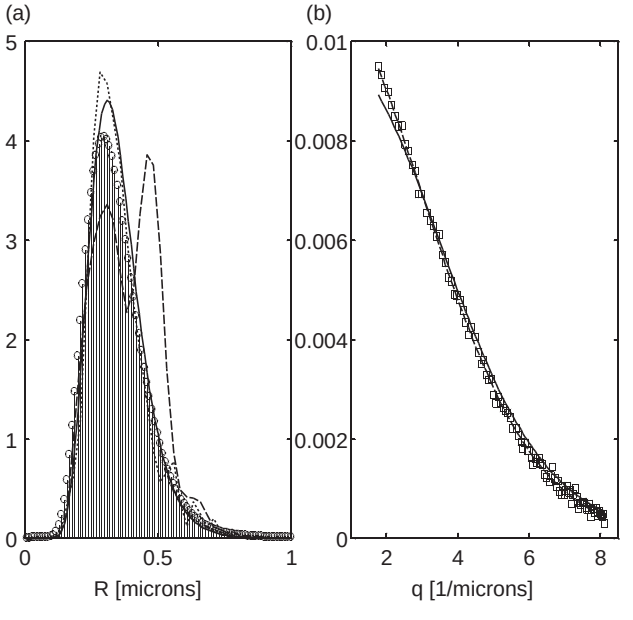

Figure 6. (a) Retrieved PSDs for Example 3 using the NPBS and the PSDs ordinate medians: actual PSD, full line; PSD from SEM simulations, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, dotted line. (b) SLS model fitting to simulated measurements for Example 3 using the NPBS and the PSDs ordinate medians: simulated measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

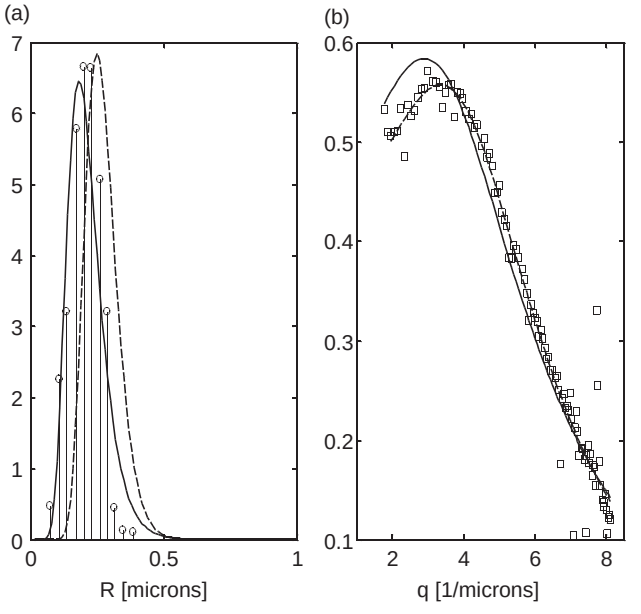

Figure 7. (a) Retrieved PSDs for Experimental Example 30PIB5 using the PBS and the parameters medians: PSD from SEM measurements, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, full line. (b) SLS model fitting to measurements for Experimental Example 30PIB5 using the PBS and the parameters medians: SLS measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

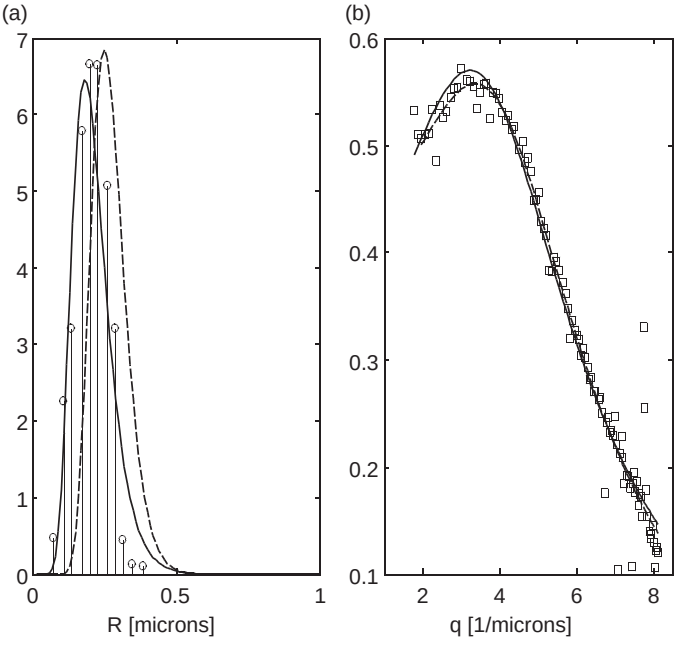

Figure 8. (a) Retrieved PSDs for Experimental Example 30PIB5 using the PBS and the parameters MAPs: PSD from SEM measurements, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, full line. (b) SLS model fitting to measurements for Experimental Example 30PIB5 using the PBS and the parameters MAPs: SLS measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

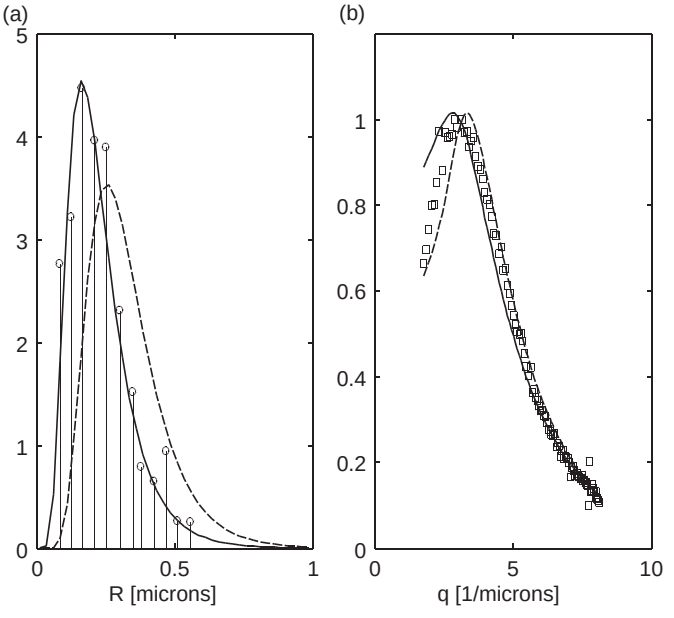

Figure 9. (a) Retrieved PSDs for Experimental Example 50PIB25 using the PBS and the parameters medians: PSD from SEM measurements, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, full line. (b) SLS model fitting to measurements for Experimental Example 30PIB5 using the PBS and the parameters medians: SLS measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

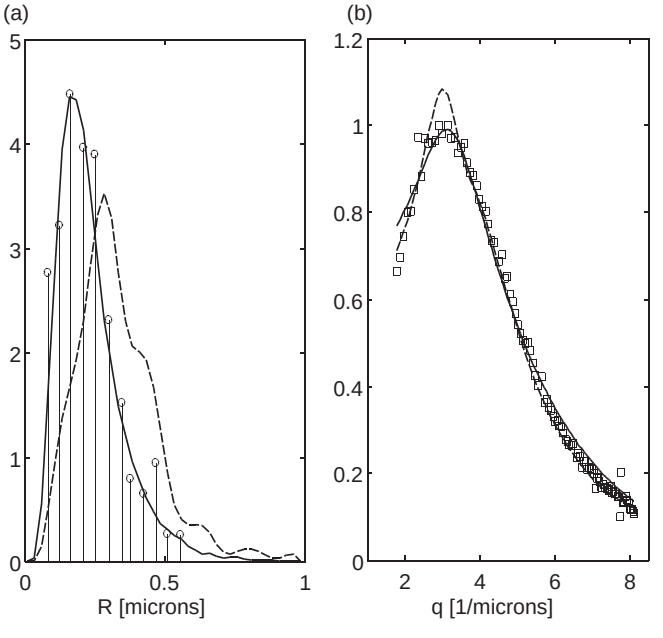

Figure 10. (a) Retrieved PSDs for Experimental Example 50PIB25 using the NPBS the PSDs ordinate medians: PSD from SEM measurements, plot bars; PSD from SLS with NI prior, dashed line; and PSD from SLS with SEM prior, full line. (b) SLS model fitting to measurements for Experimental Example 30PIB5 using the PBS and the PSDs ordinate medians: SLS measurements, unfilled squares; model fitting with a NI prior, dashed line; and model fitting with SEM prior, full line.

Example 1 corresponds to a concentrated and polydisperse system analyzed with Vrij's HS model. Results from applying the PBS appear in Table 1. In this table, it may be seen how including a prior information from SEM slightly improve the predicted confidence interval in the parameter σ and how barely more accurate estimations are obtained using SEM priors for the computed median as well as for the MAP solution. This improvement is in this case, not particularly remarkable in the PSD retrieval as can also be seen in Figure 2. In this figure, it may also be observed that differences between data fitting using a NI prior and a partially informative prior are not significant.

Example 2 corresponds to a narrow PSD where SLS modeling errors can be neglected and both techniques, SLS and SEM, obtain good results. For this example it can also be seen in Table 2 how more accurate are estimations of both statistical parameters of the PSD and how combining prior SEM information with the SLS measurements reduces the 95% confidence interval with respect to the Bayesian approach with a NI prior. It can also be observed how biased information from the prior SEM, which has the true mean radius value (0.20μ) out of the 95% confidence interval, is improved in the corresponding posterior confidence intervals after applying both Bayesian schemes, even when no greater precision is achieved in this case.

Corresponding results from retrieved PSDs and histograms from the micrographs are shown in Figures 3(a) and 4(a) while Figures 3(b) and 4(b) show a bigger discrepancy in the data fitting than the one observed in Example 1.

Example 3 corresponds to a wider PSD with a more vague SEM prior information and considerable SLS modeling errors. Results of combining both techniques show in addition to Example 1, a remarkable improvement in the resulting confidence interval for the PSD standard deviation σ and also for the mean radius $\bar{R}$ obtained in the NPBS. In this case, biased estimations, a consequence from the use of an approximate model when a NI prior has been used, are considerably improved using prior information as it can be seen when comparing the corresponding estimated medians and MAP solutions in Table 3. While improvement in the estimated medians may be clearly seen in Figure 5(a), no significant differences in data fitting are observed in Figure 5(b). Significant differences between using a NI prior and SEM information can be seen using the NPBS in Figure 6(a) with a respective difference in the data fitting observed in Figure 6(b).

When analyzing experimental cases, the example 30PIB5 shows an important issue observed in Vrij's HS model, which is the appearance of multiple local minima with a very similar value of residue. This can be observed in Table 4 when a deterministic MLE approach has been performed. Multiple local minima are the result of a high correlation between some of the model parameters. It has also been seen that the global character of a minimum cannot be guaranteed in the experiments when no additional information is included. As a proof of this behavior, Monte Carlo simulations of noisy measurements from known model parameters have shown that global minima and local minima can change from a simulated experiment to another. This exchange between locality and globality has been also observed when addressing a model error, specifically by modifying the relation between particle radii and 'hard sphere' radii as $R_{\text{hs}}^{(k)} = (C + \Delta C)R^{(k)}$ when ΔC is a normal random variable of zero mean and standard deviation of 0.01. Applying the PBS in this example with a NI prior takes into account such uncertainty with a large confidence interval including all the local minima. As an alternative to this issue, the inclusion of parameter information in a statistical manner solves the problem in a similar way to that of a constrained optimization, but in a more flexible performance allowing a direct calculus of the confidence intervals. In this case, as expected, results of statistical parameters of the PSD are close to those from the SEM micrograph, considering that SLS measurements are particularly noisy and the prior information term is stronger than the likelihood function. In fact no particular improvement in the confidence interval has been made. Corresponding medians for the obtained PSDs can be seen in Figure 7(a). It is interesting to notice in this example how this influence from the prior is

clearly showed in the median which has not a good fitting to the experimental measurements in Figure 7(b). However, taking the MAP solution instead improves remarkably the data fitting as it can be seen in Figure 8(b) while it still conserves much influence from the prior as it can be observed in Figure 8(a). The MAP solution shows a greater difference with respect to the prior SEM information. This is reflected in its estimated value of the PSD standard deviation which falls out of the 95% confidence interval. This situation is a consequence of choosing a prior in conflict with the scattered intensity measurements. The considerable improvement in data fitting using the MAP solution (when the used model is more accurate) may lead to a preference of this solution over the median, however this is not always the case as we have seen before in Example 1. In general, the MAP solution is only a 'good' estimate to pick up the best value without caring how far of the confidence interval it is. The posterior median, on the other side, penalizes in a proportional manner to the distance to the 'true' value. In this problematic case, an alternative procedure, for instance changing the shape of the prior for a type more robust or achieving additional information, can be useful to solve this difficulty.

Finally, we consider a last experimental example (50PIB25) which was solved using the LMA model and both schemes. In this case, using prior information from SEM micrographs only shows some improvement in confidence intervals for the parameter σ, when once again the prior has a stronger influence than the likelihood function as it can be seen in Table 6 and also in Figure 9(a). In this case, important differences between median and MAP solutions are observed when the PBS has been used with both a NI prior and SEM information. The sum of model error, noisy data and a conflictive prior (in the case this one has been utilized) leads to a poor fitting in Figure 9(b). However, when the NPBS has been performed, results of the MAP solution and the median are very similar and data fitting is acceptable as it can be observed in Figure 10(b). The considerable improvements in the 95% confidence interval are shown in Table 6 and once again using a partially informative prior produces a retrieved PSD very similar to the SEM which is presented in Figure 10(a).

6. Conclusions

We have developed two Bayesian schemes for solving an inverse problem in SLS, a PBS and an NPBS both implemented through the MH algorithm. In this work we have shown the potentials of the Bayesian approach within the framework of these two proposed schemes in the studied problem and also have presented some limitations found in the development of the methodology. The Bayesian approach may improve estimations combining deductive and inductive logic, specially in cases where informative priors are used to complete the uncertainty in the likelihood function and weights of both prior and likelihood are similar (as was detailed when discussing results from Example 3). An application of the method was also presented in terms of decreasing the effect of multiple minima, reducing in this form the estimated confidence intervals when using informative priors on the parameters of interest (as was described in the experimental example 30PIB5). The use of an approximate model using the Bayesian approach made possible a simplified treatment of modeling errors in both schemes with a very simple implementation in practice.

However, some limitations have been also found. A considerable difficulty appears when the term of prior information is much stronger than the likelihood function. In those cases, applying the Bayesian approach (especially when priors are normal shaped and conflicted with data) may result in controversial estimations, for instance, median and MAP solutions may result quite different, this last solution even out of the 95% confidence interval (as it has been seen in the example 30PIB5). In those cases, some limitations of the performed implementation using the MH algorithm may also be mentioned, related to many involved factors. While the PBS has a straightforward implementation, the NPBS has many significant issues, such as the tuning of

parameters some of them proper to the MH algorithm (such as the candidate-generating density) and some other parameters proper to the scheme such as the selection of the regularization matrix and the regularization parameter. Another difficulty involving the implementation, specifically the selection of normal prior densities on the parameters of interest, leads to failures of robustness, which can be improved by changing the shape of the priors, as it has been discussed before. It is also important to notice that the chosen implementation using Markov Chain Monte Carlo methods (in this specific case, through the MH algorithm) is not essential to the developed schemes and therefore it can be changed. In fact, when working with normal distributions, the optimal estimation method (OEM) can be used instead [26]. The OEM is also based on the Bayes theorem and searches to transform the mathematical model into a conditional probability and model parameters and measurements into probability distributions by assuming normal statistics and empirically determined covariance matrices. However, OEM has its own issues too. In the two proposed schemes OEM has to consider a nonlinear problem in the PBS and a linear problem which has to be iteratively solved with a cost function hard to build in the NPBS.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] O.M. Alifanov, *Inverse Heat Transfer Problems*, Springer Verlag, Berlin, 1994.

[2] R. Aster, B. Borchers, and C. Thuerber, *Parameter Estimation and Inverse Problems*, Elsevier Academic Press, New York, 2005.

[3] J. Berger, *An overview of robust Bayesian analysis*, Technical Report #93–53C, Department of Statistics, Purdue University, West Lafayette, 1993.

[4] W.K. Bertram, *Correlation effects in small-angle neutron scattering from closely packed spheres*, J. Appl. Cryst. 29 (1996), pp. 682–685.

[5] C. Bohren and D. Huffman, *Absorption and Scattering of Light by Small Particles*, John Wiley, New York, 1983.

[6] S. Chib and E. Greenberg, *Understanding the Metropolis–Hastings algorithm*, Amer. Statist. 49 (1995), pp. 327–335.

[7] L.A. Clementi, J.R. Vega, L.M. Gugliotta, and H.R.B. Orlande, *A Bayesian inversion method for estimating the particle size distribution of latexes from multiangle dynamic light scattering measurements*, Chemometr. Intell. Lab. Syst. 107 (2011), pp. 165–173.

[8] C. Fraley and A.E. Raftery, *Bayesian regularization for normal mixture estimation and model-based clustering*, J. Classif. 24 (2007), pp. 155–181.

[9] J. Freund, *Modern Elementary Statistics*, Prentice Hall, New Jersey, 1967.

[10] A. Gelman, W.R. Gilks, and G.O. Roberts, *Weak convergence and optimal scaling of random walk Metropolis algorithms*, Ann. Appl. Probab. 7 (1997), pp. 110–120.

[11] C.J. Geyer and E.A. Thompson, *Annealing Markov chain Monte Carlo with applications to ancestral inference*, J. Amer. Statist. Assoc. 90 (1995), pp. 909–920.

[12] O. Glatter and M. Hofer, *Interpretation of elastic light scattering data. III. Determination of size distributions of polydisperse systems*, J. Coll. Interface Sci. 122 (1988), pp. 496–506.

[13] P.C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support*, Cambridge University Press, Cambridge, 2005.

[14] E.K. Hobbie and L. Sung, *Rayleigh–Gans scattering from polydisperse colloidal suspensions*, Am. J. Phys. 64 (1996), pp. 1298–1303.

[15] H.C. van de Hulst, *Light Scattering by Small Particle*, Dover Publications, New York, 1981.

[16] J.P. Kaipio, V. Kohlemainen, E. Somersalo, and M. Vauhkonen, *Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography*, Inverse Probl. 16 (2000), pp. 1487–1522.

[17] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems, Series: Applied Mathematical Sciences*, Vol. 160, Springer, New York, 2005.

[18] R.E. Kass and L. Wasserman, *The selection of prior distributions by formal rules*, J. Amer. Statist. Assoc. 91 (1996), pp. 1343–1370.

[19] S. Konishi, T. Ando, and S. Imoto, *Bayesian information criteria and smoothing parameter selection in radial basis function networks*, Biometrika 91 (2004), pp. 27–43.
[20] H.F. Lopes and J.L. Tobias, *Confronting prior convictions: On issues of prior sensitivity and likelihood robustness in Bayesian analysis*, Annu. Rev. Econ. 3 (2011), pp. 107–31.
[21] V.A. Markel and J.C. Schotland, *Inverse problem in optical diffusion tomography. I. Fourier-Laplace inversion formulas*, J. Opt. Soc. Am. A 18 (2001), pp. 1336–1347.
[22] M.N. Osizik and H.R.B. Orlande, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor & Francis, New York, 2000.
[23] R.L. Parker, *Geophysical Inverse Theory*, Princeton University Press, Princeton, NJ, 1994.
[24] J.S. Pedersen, *Small-angle scattering from precipitates: Analysis by use of a polydisperse hard-sphere model*, Phys. Rev. B 47 (1993), pp. 657–665.
[25] J.S. Pedersen, *Determination of size distribution from small-angle scattering data for systems with effective hard-sphere interactions*, J. Appl. Cryst. 27 (1994), pp. 595–608.
[26] C.D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, Singapore, 2000.
[27] I.G. Roy, *Iteratively adaptive regularization in inverse modeling with Bayesian outlook – application on geophysical data*, Inverse Probl. Sci. Eng. 13 (2005), pp. 655–670.
[28] P.K. Sen, J.P. Keating, and R.L. Mason, *Pitman's Measure of Closeness: A Comparison of Statistical Estimators*, SIAM, Philadelphia, 1993.
[29] D. Sorensen and D. Gianola, *Likelihood, Bayesian and MCMC Methods in Quantitative Genetics*, Springer, New York, 2002.
[30] E.R. Soulé and G.E. Elicabe, *Determination of size distributions of concentrated polymer particles embedded in a solid polymer matrix*, Part. Part. Syst. Charact. 25 (2008), pp. 84–91.
[31] N. Tikhonov and V.Y. Arsenin, *Solution of Ill-Posed Problems*, Winston & Sons, Washington, 1977.
[32] C.R. Vogel and M.E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput. 17 (1996), pp. 227–238.
[33] A. Vrij, *Light scattering of a concentrated multicomponent system of hard spheres in the Percus-Yevick approximation*, J. Chem. Phys. 69 (1978), pp. 1742–1747.
[34] A. Vrij, *Mixtures of hard spheres in the Percus-Yevick approximation. Light scattering at finite angles*, J. Chem. Phys. 71 (1979), pp. 3267–3275.
[35] G. Wang, J. Zhang, and G.-W. Pan, *Solution of inverse problems in image processing by wavelet expansion*, IEEE Trans. Image Process. 4 (1995), pp. 579–593.
[36] H. Wang, R. Li, and C.-L. Tsai, *Tuning parameter selectors for the smoothly clipped absolute deviation method*, Biometrika 94 (2007), pp. 553–568.
[37] D.S. Xue and M.S. Si, *Bayesian inference approach to particle size distribution estimation in ferrofluids*, IEEE Trans. Magnet. 42 (2006), pp. 3657–3660.

Appendix 1. Vrij's HS model

Vrij's HS model gives an analytic formula for the DSCS per unit volume for a mixture of spherical particles of different sizes, which is expressed as:

$$\frac{\mathrm{d}\sum(q)}{\mathrm{d}\Omega} = -D(q)\Delta(q)^{-1}, \tag{A1}$$

where the expressions for the computation of $D(q)$ and $\Delta(q)$ can be obtained using Equations (A2)–(A8)

$$\begin{aligned} -\frac{\pi}{6}(1-\langle d^3\rangle)^4 D(q) = {} & \langle b(q)^2\rangle T_1 T_1^* + \langle d^6\phi^2\rangle T_2 T_2^* + 9\langle d^4\psi^2\rangle T_3 T_3^* \\ & + \langle b(q)d^3\phi\rangle(T_1 T_2^* + T_1^* T_2) + 3\langle b(q)d^2\psi\rangle(T_1 T_3^* + T_1^* T_3) \\ & + 3\langle d^5\phi\psi\rangle(T_2 T_3^* + T_2^* T_3) \end{aligned} \tag{A2}$$

and

$$\Delta(q) = \frac{1}{(1-\langle d^3\rangle)^4}(F_{11}F_{22} - F_{12}F_{21})(F_{11}^* F_{22}^* - F_{12}^* F_{21}^*) \tag{A3}$$

with

$$T_1(q) = F_{11}F_{12} - F_{12}F_{21}$$
$$T_2(q) = F_{21}\langle db(q)\, e^{iX}\rangle - F_{12}\langle b(q)\, e^{iX}\rangle \tag{A4}$$
$$T_3(q) = F_{12}\langle b(q)\, e^{iX}\rangle - F_{11}\langle db(q)\, e^{iX}\rangle,$$
$$F_{11}(q) = 1 - \langle d^3\rangle + \langle d^3\phi\, e^{iX}\rangle$$
$$F_{12}(q) = \langle d^4\phi\, e^{iX}\rangle \tag{A5}$$
$$F_{22}(q) = 1 - \langle d^3\rangle + 3\langle d^3\psi\, e^{iX}\rangle$$
$$F_{21}(q) = \tfrac{1}{2}(1 - \langle d^3\rangle)iq - 3\langle d^2\rangle + 3\langle d^3\psi\, e^{iX}\rangle$$
$$d = 2R_{hs}$$
$$X = qR_{hs}$$
$$\psi = \frac{\sin X}{X} \tag{A6}$$
$$\phi = \frac{3}{X^3}(\sin X - X\cos X),$$
$$b(q) = \Delta\rho\left(\frac{4\pi}{3}R^3\right)\left[\frac{3}{(qR)^3}(\sin qR - qR\cos qR)\right]. \tag{A7}$$

The used brackets indicate in the general case that

$$\langle A_1A_2\cdots A_n\rangle = \frac{\pi}{6}\sum_{k=1}^{N} N_p^{(k)}A_1^{(k)}A_2^{(k)}\cdots A_n^{(k)},$$

where $A_j^{(k)}$ is the value of the function A_j evaluated for the kth particle, with $N_p^{(k)}$ as the number of particles per unit volume of particles of kth size; N is the number of different sizes in the sample; $\Delta\rho$ is the difference of length scattering densities between medium and particles. On the other side, $R^{(k)}$ and $R_{hs}^{(k)}$, make reference, respectively, to the radii of the particles and to the radii of the 'hard spheres' both corresponding to the kth size, with $R^{(k)} \le R_{hs}^{(k)}$. This last consideration is due to the nature of Vrij's HS model which states that interaction between particles is given by spheres of greater size, that is, the 'hard spheres'. Physically, the radii of both particles and 'hard spheres' can be related in two main forms: $R_{hs}^{(k)} = CR^{(k)}$ and $R_{hs}^{(k)} = R^{(k)} + \Delta R$. It seems important to remark that for computation of the so-called form factor $b(q)$ the true radii of the different particles is used as it can be seen in Equation (9). Finally, the so-called volume fraction of the particles can be calculated according to Equation (10):

$$\eta = \frac{4\pi}{3}\sum_{k=1}^{N} N_p^{(k)}(R^{(k)})^3. \tag{A8}$$

In a similar manner, η_{hs}, the volume fraction of the 'hard spheres' can be obtained replacing $R^{(k)}$ by $R_{hs}^{(k)}$ in Equation (10).

Appendix 2. Local monodisperse approximation

Expressions for $F(q, R)$ and $S(p, q, R)$ used in Equation (3) are displayed below:

$$F(q,R) = \frac{1}{q}\int_0^R r\sin(qr)\,dr = \frac{1}{q^3}[\sin(qR) - qR\cos(qR)], \tag{A9}$$

$$S(p,q,R)^{-1} = 1 - 24p\left[\frac{\alpha+\beta+\delta}{u^2}\cos(u) - \frac{\alpha+\beta+\delta}{u^3}\sin(u) - \frac{2(\beta+6\delta)}{u^4} + \frac{2\beta}{u^4} + \frac{24\delta}{u^5}\sin(u) + \frac{24\delta}{u^6}[\cos(u) - 1]\right], \tag{A10}$$

where

$$\alpha = \frac{(1+2p)^2}{(1-p)^4}, \tag{A11}$$

$$\beta = -6p\frac{(1+p/2)^2}{(1-p)^4}, \tag{A12}$$

$$\delta = \frac{p(1+2p)^2}{2(1-p)^4}, \tag{A13}$$

$$u = 2qR. \tag{A14}$$