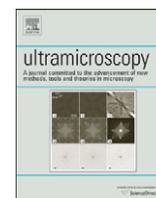




ELSEVIER

Contents lists available at ScienceDirect

## Ultramicroscopy

journal homepage: [www.elsevier.com/locate/ultramic](http://www.elsevier.com/locate/ultramic)

# Throughput maximization of particle radius measurements through balancing size versus current of the electron probe

W. Van den Broek<sup>a,\*</sup>, S. Van Aert<sup>a</sup>, P. Goos<sup>b,c</sup>, D. Van Dyck<sup>a</sup>

<sup>a</sup> EMAT, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

<sup>b</sup> Faculty of Applied Economics and StatUa Center for Statistics, University of Antwerp, Prinsstraat 13, B-2000 Antwerp, Belgium

<sup>c</sup> Erasmus School of Economics, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands

## ARTICLE INFO

## Keywords:

Bayesian experimental design

HAADF STEM

Maximum throughput

Electron probe

## ABSTRACT

In this paper we investigate which probe size maximizes the throughput when measuring the radius of nanoparticles in high angle annular dark field scanning transmission electron microscopy (HAADF STEM). The size and the corresponding current of the electron probe determine the precision of the estimate of a particle's radius. Maximizing throughput means that a maximum number of particles should be imaged within a given time frame, so that a prespecified precision is attained. We show that Bayesian statistical experimental design is a very useful approach to determine the optimal probe size using a certain amount of prior knowledge about the sample. The dependence of the optimal probe size on the detector geometry and the diameter, variability and atomic number of the particles is investigated. An expression for the optimal probe size in the absence of any kind of prior knowledge about the specimen is derived as well.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

In this paper we investigate the optimal probe settings for high angle annular dark field scanning transmission electron microscopy (HAADF STEM). It is common practice to optimize the resolution of the coherent probe contribution in some respect, see for example Ref. [1] for a derivation of the Scherzer settings or Ref. [2] for a derivation of the settings when the limiting aberrations are of higher order. In these studies the incoherent probe contribution due to a finite source size is ignored, assuming a purely coherent point source. However, a point source emits no current, so that no electrons would be present for the actual imaging. Introducing a finite source size broadens the probe in a way that is well parameterized by Barth and Kruit in Ref. [3]. The optimal probe size will depend on a trade-off between the probe width and the beam current. A large width increases the beam current and augments the signal-to-noise ratio, although at the expense of reduced resolution. In practice, operators balance these two effects by adjusting the so-called spot size of the microscope. However, this choice may be somewhat subjective and therefore operator dependent. In this paper we provide a sound theoretical basis for this choice. The problem we investigate, is maximizing the throughput when measuring the radii  $R$  of spherical nanoparticles deposited on a uniform support. That is, given a prespecified precision of the estimates of  $R$ , we seek the probe size that yields

the minimum required recording time needed to reach that precision.

The images are considered as data planes from which structural information has to be estimated quantitatively. For this we use a model for the object and for the imaging process, including electron-object interaction, microscope transfer and image detection. This model describes the expectations of the intensity observations and it contains the parameters that have to be measured. These parameters are determined by fitting the model to the experimental data by the use of a criterion of goodness-of-fit, such as least squares or maximum likelihood. In this way structure determination becomes a statistical parameter estimation problem. The precision with which structure parameters can be estimated is limited by the presence of noise. Use of the Fisher information [4] allows to derive an expression for the best attainable precision with which the structure parameters can be estimated. This expression, which is called the Cramér-Rao lower bound (CRLB), is a function of the object parameters, the microscope parameters and the electron dose.

Statistical optimal design is a discipline that if applied to electron microscopy, searches the set of microscope parameters that yields the highest attainable precision on the estimates of one or several of the structure parameters of the sample. In Ref. [5] for example, the CRLB on the variance with which atom column positions can be estimated is used as a performance measure in the optimization of STEM experiments. This methodology has been applied to optimize the design of other microscopy experiments as well, see Refs. [6–9] for examples. In this article, a lower bound  $\sigma_{CR}^2$  on the variance of the estimate of  $R$  is derived. In our current problem,  $\sigma_{CR}^2$  is a function of  $R$ , the parameter that has to be

\* Corresponding author. Tel.: +32 3 265 32 46; fax: +32 3 265 33 18.  
E-mail address: wouter.vandenbroek@ua.ac.be (W. Van den Broek).

estimated and thus is unknown a priori. This problem is common for any optimal experimental design model involving a non-linear statistical model [10]. For this reason, we adopt a Bayesian approach in which we use a prior probability distribution,  $p(R)$ , which reflects the distribution of  $R$  and seek the probe size that is optimal over the entire distribution of radii. In order to derive an overall optimal probe size, we define  $\sigma_B^2$  as the average  $\sigma_{CR}^2(R)$  weighted by  $p(R)$ . It is explained in Section 4 that the probe size optimizing  $\sigma_B^2$  also accomplishes maximum throughput. Therefore, in the remainder of this paper we will look for the probe optimizing  $\sigma_B^2$  with constant recording time per unit area whilst keeping in mind that this same probe also maximizes throughput.

The outline of this paper is as follows. In Section 2, we introduce the models for the probe, the particles and the support and combine them into a model for the images. In Section 3, we specify the joint probability function, and explain how it leads to the Fisher information matrix, and the CRLB. In Section 4, we introduce the Bayesian optimality criterion, show the equivalence between minimum  $\sigma_B^2$  and maximum throughput, give analytical results and rules of thumb for the optimal probe sizes, and carry out a simulation study to check the influence of the particles' atomic numbers, the detector geometry, and the mean and variance of the prior distribution  $p(R)$ . In Section 5, we summarize the final conclusions.

## 2. The image model

### 2.1. Probe model

In Ref. [3], Barth and Kruit propose a root-power-sum algorithm that relates the probe current  $I$  to the probe size  $d_p$ , where  $d_p$  is the diameter of the disc containing  $p\%$  of the total probe current. In this paper, we choose  $d_{50}$  as a resolution measure, as suggested in Refs. [3,11].

The dependence of  $d_{50}$  on the microscope settings is given by

$$d_{50}^2 = (d_l^{1.3} + (d_A^4 + d_s^4)^{1.3/4})^{2/1.3} + d_c^2 \quad (1)$$

with

$$d_l = \frac{21}{\pi \alpha} \sqrt{\frac{I}{B_r E_0}},$$

$$d_A = 0.54 \frac{\lambda}{\alpha},$$

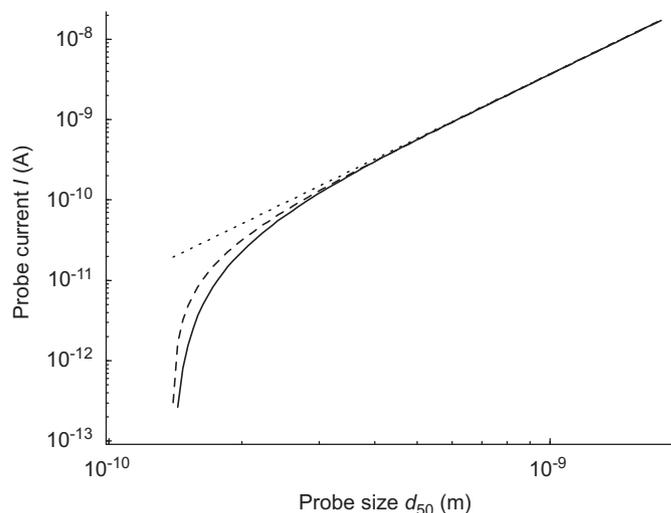
$$d_s = 0.18 C_s \alpha^3$$

and

$$d_c = 0.34 C_c \frac{\delta E}{E_0} \alpha.$$

In these expressions,  $B_r$  is the reduced brightness [12] of the electron gun,  $E_0$  is the acceleration voltage,  $\alpha$  is the semi-angle of the aperture selecting the spot size,  $\lambda$  the electron wavelength,  $C_s$  the spherical aberration,  $C_c$  the chromatic aberration, and  $\delta E$  the full width at half the maximum of the electrons' energy distribution.

Eq. (1) can be used to produce  $(d_{50}, I)$ -curves by fixing  $I$  at different values and minimizing  $d_{50}$  numerically with respect to  $\alpha$  for each of the  $I$  values. The solid curve in Fig. 1 was calculated in this way, with the microscope parameters used throughout this paper and given in Table 1. This approach is somewhat unwieldy. Therefore, we will derive an analytical, albeit approximate, expression for the  $(d_{50}, I)$ -curves, providing more insight in the problem at hand, that is, the maximization of throughput of particle radius measurements through optimization of the probe size  $d_{50}$ .



**Fig. 1.** Three curves relating the beam current  $I$  to  $d_{50}$ , with microscope parameters summarized in Table 1. The solid curve is derived from Eq. (1) with  $C_c=1$  mm. The dotted curve depicts the approximation in Eq. (3), and the dashed curve shows the approximation given by Eq. (5) and used throughout the paper.

**Table 1**

Microscope parameters used in the simulations.

$E_0$	$\lambda$	$C_s$	$B_r$
300 kV	1.97 pm	1 mm	$5 \times 10^7 \text{ A m}^{-2} \text{ sr}^{-1} \text{ V}^{-1}$

Various terms in Eq. (1) can be neglected for certain microscope settings. This is shown, for instance, in Ref. [12], where various approximations for low values of  $E_0$  are given. In this paper, a similar derivation is given. The acceleration voltage  $E_0$  equals 300 keV, while  $\delta E$  is only 0.4 eV, as a consequence the  $d_c^2$ -term can be neglected. In addition, numerical calculations showed that the angle  $\alpha$  minimizing  $d_{50}$  for a given  $I$  increases monotonically with  $d_{50}$ . This suggests that, for large  $I$  and large  $d_{50}$ , the  $d_A^4$ -term can be dropped from the model as well. The model then simplifies to

$$d_{50}^{1.3} = d_l^{1.3} + d_s^{1.3}. \quad (2)$$

It can be shown analytically that, in this case,  $d_{50}$  is minimized with respect to  $\alpha$  if

$$I = 0.25 \pi^2 B_r E_0 C_s^{-2/3} d_{50}^{8/3}. \quad (3)$$

This function is shown as the dotted line in Fig. 1. It is clear that for large probe sizes, it provides an excellent approximation to the original model. For smaller probe sizes, however, the simplified model behaves qualitatively different from the exact model.

A better approximation is obtained by incorporating the geometrically limited probe size  $d_{g,50}$ . It is defined as the probe size in the limit of zero probe current. For higher values of  $E_0$ , it is approximated by

$$d_{g,50}^4 = d_A^4 + d_s^4.$$

Analytically minimizing  $d_{g,50}$  with respect to  $\alpha$  yields

$$d_{g,50} = 0.47 C_s^{1/4} \lambda^{3/4}. \quad (4)$$

By definition it holds that  $I$  tends to 0 if  $d_{50}$  approaches  $d_{g,50}$ . We therefore propose to incorporate  $d_{g,50}$  in Eq. (3) in the following way:

$$I = 0.25 \pi^2 B_r E_0 C_s^{-2/3} (d_{50}^{8/3} - d_{g,50}^{8/3}). \quad (5)$$

This results in an approximation that is excellent for large probe sizes and behaves qualitatively correct for small probe sizes too. This is shown by the dashed curve in Fig. 1.

We model the shape of the probe  $h(r)$  using a Gaussian function with resolution  $d_{50}$  and with an integrated intensity equal to the total number of electrons [13]. The required function is

$$h(r) = \frac{I\tau}{e} \frac{1}{2\pi\sigma_{d_{50}}^2} \exp\left(\frac{-r^2}{2\sigma_{d_{50}}^2}\right) \quad (6)$$

with

$$\sigma_{d_{50}}^2 = \frac{d_{50}^2}{8\ln 2}, \quad (7)$$

$r$  the radial distance from the centre of the probe,  $\tau$  the dwell time and  $e$  the elementary charge.

## 2.2. Specimen model

We model the particles as homogeneous spheres with radius  $R$ . The projection  $q(r)$  of such a sphere is

$$q(r) = 2\gamma_1 \sqrt{R^2 - r^2} \quad \text{for } r \leq R, \\ = 0 \quad \text{for } r > R,$$

where  $r$  is the radial distance from the centre of the particle's projection and  $\gamma_1$  is a proportionality factor accounting for the percentage of electrons in the beam that are counted at the detector. This means that  $\gamma_1$  accounts for the particle's atomic number, the detector geometry and the detector's efficiency.

In order to gain physical insight we approximate the image model to make it analytically tractable, but note that an exact and purely numerical treatment is possible as well. The function  $q(r)$  is therefore approximated by a Gaussian function  $q(r)$  that has the same variance, and the same total intensity  $4\pi\gamma_1 R^3/3$ . This is achieved using

$$q(r) = \frac{2}{3}\gamma_1 \frac{R^3}{\sigma_R^2} \exp\left(\frac{-r^2}{2\sigma_R^2}\right) \quad (8)$$

with

$$\sigma_R^2 = \frac{1}{5}R^2. \quad (9)$$

The particles are deposited on an amorphous support that is modeled as a constant additive background  $\gamma_0$  that accounts for the support's thickness and atomic number, and the detector's geometry and efficiency.

## 2.3. The image model

In Ref. [14] it is stated that the image formation in HAADF STEM can be assumed incoherent. Therefore, the expected image intensity  $\lambda(r)$  is a convolution of the Gaussian probe with the Gaussian particle projection. The result is a new Gaussian function with a variance equal to the sum of the variances of  $h(r)$  and  $q(r)$ , and an integrated intensity that is the product of those of  $h(r)$  and  $q(r)$ :

$$\lambda(r) = \gamma_0 \frac{I\tau}{e} + \frac{2}{3}\gamma_1 \frac{I\tau}{e} \frac{R^3}{\sigma_{d_{50}}^2 + \sigma_R^2} \exp\left(\frac{-r^2}{2(\sigma_{d_{50}}^2 + \sigma_R^2)}\right) \quad (10)$$

with  $\sigma_{d_{50}}$  depending on  $d_{50}$  according to Eq. (7) and  $\sigma_R$  depending on  $R$  according to Eq. (9).

In practice, one rarely has prior knowledge about the parameters  $\gamma_0$ ,  $\gamma_1$  or  $I$ . Therefore, we carry out the coordinate transformation

$$A = \gamma_0 \frac{I\tau}{e}, \quad (11)$$

$$B = \frac{2}{3}\gamma_1 \frac{I\tau}{e} \frac{R^3}{\sigma_{d_{50}}^2 + \sigma_R^2}, \quad (12)$$

$$R = R, \quad (13)$$

such that the model in Eq. (10) simplifies to

$$\lambda(r) = A + B \exp\left(\frac{-r^2}{2(\sigma_{d_{50}}^2 + \sigma_R^2)}\right). \quad (14)$$

$A$  and  $B$  are nuisance parameters that have to be estimated together with  $R$ , but which are not of primary interest. The interdependence between the three variables  $R$ ,  $A$  and  $B$  is ignored and will not be used in the estimation process. This reflects the common case that one has too little knowledge to extract information out of the absolute value of  $A$  or any ratio between  $R$ ,  $A$  and  $B$ . One is therefore forced to estimate  $R$ ,  $A$  and  $B$  separately, as if they are independent. For estimation of the model parameters, typically a finite region around the particle's projection, i.e. the field of view (FOV), is selected. In this paper, the FOV is chosen as a square with sides equal to  $4R_{\text{eff}}$ , where  $R_{\text{eff}}$  is the effective radius of the model, defined analogously to Eq. (9):

$$R_{\text{eff}}^2 = 5(\sigma_{d_{50}}^2 + \sigma_R^2). \quad (15)$$

In STEM, the probe scans the specimen in a raster fashion. The image is thus recorded as a function of the probe position. Without loss of generality, the image magnification is ignored. Therefore the probe position directly corresponds to an image pixel at the same position. The image is assumed to consist of  $M \times M$  equidistant pixels of area  $\Delta^2$ , where  $\Delta$  is the probe sampling distance. Without loss of generality, the FOV is chosen to be centred about the centre of the particle's projection. It is convenient to index the pixels with a single number  $i$  that enumerates them column-wisely. If  $r_i$  is the distance between the centre of pixel  $i$  and the centre of the FOV, then the model becomes

$$\lambda(r_i) = A + B \exp\left(\frac{-r_i^2}{2(\sigma_{d_{50}}^2 + \sigma_R^2)}\right). \quad (16)$$

In Refs. [5,15,16] it is stated that the precision is independent of the sample distance  $\Delta$ , if  $\Delta$  is small compared to the Gaussian distribution's width and if the number of counts per unit area is kept constant. To ensure that  $\Delta$  is small enough for all values of  $d_{50}$  and  $R$ , we use a fixed number of pixels  $M^2$  inside the  $4R_{\text{eff}} \times 4R_{\text{eff}}$  region of interest. As a consequence,  $\Delta$  will vary as a function of  $d_{50}$  and  $R$  according to

$$\Delta^2 = \frac{(4R_{\text{eff}})^2}{M^2} = \frac{80}{M^2} (\sigma_{d_{50}}^2 + \sigma_R^2), \quad (17)$$

where Eq. (15) is used. Since in this paper the recording time per unit area is kept fixed, we must adjust the dwell time accordingly,

$$\tau = \tau_0 \Delta^2, \quad (18)$$

where  $\tau_0$ , the recording time per unit area, is a fixed constant. Evidently, in experimental practice a fixed sampling distance must be chosen. Since a sampling distance small enough for a particular radius will also be sufficiently small for larger radii, specimen containing particles with different radii can be treated by setting the sampling distance small enough for the smallest particle present.

## 3. The Cramér–Rao lower bound

The Cramér–Rao lower bound (CRLB) is a lower bound on the variance of any unbiased estimator. Suppose that an experimentalist wants to estimate the radius  $R$  of a particle. Then he can choose between many estimators, such as least squares, least

absolute values or maximum likelihood estimators. The precision of an estimator is represented by the variance or the standard deviation of the estimates. In general, different estimators will have different precisions. However, the variance of an unbiased estimator will never be lower than the CRLB. Fortunately, there exists a class of estimators, including the maximum likelihood estimator, that achieves this bound at least asymptotically, that is, for the number of observations going to infinity, see [17] for details.

The pixel values in the projections are electron counts and are recorded consecutively, they thus suffer from Poisson noise and are uncorrelated. Therefore, the probability  $P(\mathbf{g}; \theta)$  of obtaining a set of measurements  $\mathbf{g} = (g_1, \dots, g_{M^2})^T$ , with parameters defined by the vector  $\theta = (R, A, B)^T$ , is given as

$$P(\mathbf{g}; \theta) = \prod_{i=1}^{M^2} \frac{\lambda(r_i)^{g_i}}{g_i!} \exp(-\lambda(r_i)). \quad (19)$$

This function is called the joint probability density function of the observations. It is a function of the observations  $\mathbf{g}$ . The parameters  $\theta$  to be estimated enter  $P(\mathbf{g}; \theta)$  via  $\lambda(r_i)$ .

Next, the so-called Fisher information matrix  $\mathbf{F}$  can be computed. The  $(u, v)$ th element of  $\mathbf{F}$  is defined as [4,5,17]

$$F_{uv} = -E \left[ \frac{\partial^2 \ln P(\mathbf{g}; \theta)}{\partial \theta_u \partial \theta_v} \right], \quad (20)$$

where  $E$  denotes the expectation operator, and  $\theta_u$  the  $u$ th element of  $\theta$ . From Eqs. (19) and (20), it follows that

$$F_{uv} = \sum_{i=1}^{M^2} \frac{1}{\lambda(r_i)} \frac{\partial \lambda(r_i)}{\partial \theta_u} \frac{\partial \lambda(r_i)}{\partial \theta_v}. \quad (21)$$

If the pixel size  $\Delta$  is small enough, and if the summation over the  $4R_{\text{eff}} \times 4R_{\text{eff}}$  square is replaced by an integral over a disc with the same area, then  $\mathbf{F}$  can be approximated by

$$F_{u,v} \simeq \frac{2\pi}{\Delta^2} \int_0^{r_{\text{fov}}} \frac{1}{\lambda(r)} \frac{\partial \lambda(r)}{\partial \theta_u} \frac{\partial \lambda(r)}{\partial \theta_v} r \, dr, \quad (22)$$

where  $\pi r_{\text{fov}}^2 = (4R_{\text{eff}})^2$ .

The Cramér–Rao inequality states that for any unbiased estimator  $\hat{\theta}$  of  $\theta$ ,

$$\text{cov}(\hat{\theta}, \hat{\theta}) \geq \mathbf{F}^{-1}, \quad (23)$$

where  $\text{cov}(\hat{\theta}, \hat{\theta})$  is the  $3 \times 3$  covariance matrix of  $\hat{\theta}$  [18]. The matrix  $\mathbf{F}^{-1}$  is called the Cramér–Rao lower bound (CRLB) on the variance of  $\hat{\theta}$ . Eq. (23) expresses that the difference between the left-hand and right-hand member is positive semi-definite. A property of a positive semi-definite matrix is that its diagonal elements cannot be negative. Therefore, the diagonal elements of  $\mathbf{F}^{-1}$  define lower bounds on the variances of the elements of  $\hat{\theta}$ ,

$$\text{var}(\hat{\theta}_u) \geq [\mathbf{F}^{-1}]_{uu}. \quad (24)$$

#### 4. Optimal experimental design

Our objective is to find the probe size  $d_{50}$  that ensures the highest throughput. As we explain below, that probe size also allows the most precise measurement of the particle radius  $R$ .

##### 4.1. Optimality criterion

The precision with which  $R$  can be estimated, is represented by the first diagonal element of the CRLB, i.e. by  $[\mathbf{F}^{-1}]_{1,1}$ , which we denote by  $\sigma_{\text{CR}}^2$ . The variance  $\sigma_{\text{CR}}^2$  depends on  $d_{50}$  and  $R$ . Therefore, minimizing  $\sigma_{\text{CR}}^2$  with respect to  $d_{50}$  yields the probe size that allows the most precise estimate of  $R$  for that particular value of  $R$ .

We now face the problem that, in general,  $R$  is not known a priori—to the contrary, it is the parameter that we want to estimate. We circumvent this problem with a Bayesian approach [10], in which any prior information is summarized in a so-called prior distribution  $p(R)$ . That distribution expresses our uncertainty about the exact value of  $R$  prior to the data collection. Rather than minimizing  $\sigma_{\text{CR}}^2$ , the Bayesian approach minimizes the expected value of  $\sigma_{\text{CR}}^2$  over the prior distribution. In this paper, we therefore seek the probe size that minimizes

$$\sigma_{\text{B}}^2 = \int_0^\infty \sigma_{\text{CR}}^2(R) p(R) \, dR. \quad (25)$$

We refer to the probe size that minimizes  $\sigma_{\text{B}}^2$  as the Bayesian optimal probe size  $d_{B,50}$ . The Bayesian approach to optimal experimental design for non-linear models is common practice, for instance, in industrial experimentation, medical trials and choice experiments in marketing [19,20].

The main challenge in the Bayesian approach is to specify a suitable prior distribution  $p(R)$ . It turns out that, for the present problem, it is not hard to find a sensible prior distribution. In Ref. [21], it is argued that the radii of ground particles typically follow a Weibull distribution, which is well approximated by a log-normal distribution. Often, particles produced with aerosol techniques follow a log-normal distribution as well [22]. Therefore, it is natural to choose a log-normal distribution for the prior  $p(R)$ :

$$p(R) = \frac{1}{\sqrt{2\pi R \ln \sigma}} \exp\left(-\frac{\ln^2(R/\rho)}{2 \ln^2(\sigma)}\right), \quad (26)$$

where  $\rho$  is the geometric mean radius and  $\sigma$  is the geometric standard deviation, with  $\rho > 0$  and  $\sigma > 1$  [22]. The geometric mean and standard deviation are defined as

$$\ln(\rho) = \int_0^\infty \ln(R) p(R) \, dR,$$

$$\ln^2(\sigma) = \int_0^\infty \ln^2(R/\rho) p(R) \, dR.$$

In the remainder of this paper the terms “mean” and “standard deviation” refer to the geometric mean and geometric standard deviation as defined here. For a log-normal distribution, the interval  $[\rho/\sigma, \rho\sigma]$  contains 68.3% of the total intensity, while the interval  $[\rho/\sigma^3, \rho\sigma^3]$  contains 99.7%. The log-normal distribution becomes sharply peaked around  $\rho$  as  $\sigma$  tends to 1. To avoid numerical difficulties in our computations, we used 1.01 as the smallest value for  $\sigma$ .

As the integral in Eq. (25) cannot be solved analytically, we approximate it numerically. This is customary in Bayesian optimal experimental design [10,19,20]. We evaluate the integrand in 21 points in the interval  $[\rho/\sigma^3, \rho\sigma^3]$ , these points are chosen to be equidistant on the logarithmic scale. Integration is carried out by approximating the integrand by a piecewise linear function. Twenty-one points were sufficient for convergence.

##### 4.2. Maximum throughput

The Bayesian optimal probe size  $d_{B,50}$  does not depend on the recording time per unit area  $\tau_0$ . This can be seen as follows. Eqs. (10) and (18) show that the expectation model  $\lambda(r)$  is proportional to  $\tau_0$ . Therefore, it follows from Eq. (21) that all elements of the Fisher information matrix are proportional to  $\tau_0$  as well. As a result,  $\sigma_{\text{CR}}^2$  and  $\sigma_{\text{B}}^2$  are inversely proportional to  $\tau_0$ , which means that  $\tau_0$  affects the absolute value of  $\sigma_{\text{B}}^2$ , but not the position of its minimum, i.e.  $d_{B,50}$ .

This is an important result since it means that the probe size  $d_{B,50}$  that minimizes  $\sigma_{\text{B}}^2$  also maximizes the throughput. Suppose one minimizes  $\sigma_{\text{B}}^2$  with respect to  $d_{50}$  while keeping  $\tau_0$  fixed. Then,

one can tune  $\tau_0$  such that  $\sigma_B^2(d_{B,50})$  equals a prespecified variance on the estimates of the radii. Since  $\sigma_B^2$  attains a minimum in  $d_{B,50}$ , other probe sizes will not attain the required variance. If  $\tau_0$  would be reduced, even  $\sigma_B^2(d_{B,50})$  would be too large. Hence,  $d_{B,50}$  is the probe diameter at which  $\sigma_B^2$  attains the prespecified precision with the lowest  $\tau_0$ . Therefore, in the remainder of this paper we will look for the probe minimizing  $\sigma_B^2$  whilst keeping in mind that this same probe also maximizes throughput.

### 4.3. Analytical results

To gain physical insight and to give rules of thumb, we carry out an analytical evaluation. The details are given in Appendix A. To make the problem analytically tractable, the background term  $A$  in Eq. (14) is neglected. This assumption is often justified, especially for heavy element particles deposited on a light element support since in HAADF STEM the image intensity approximately scales with the atomic number raised to a power of typically 1.7 [14]. Ignoring the background term results in a model with only two unknown parameters,  $\theta = (R, B)^T$ . The Fisher information matrix  $\mathbf{F}$  is then given by

$$\mathbf{F} = \begin{pmatrix} \frac{16\pi B}{25\Delta^2} \frac{R^2}{\sigma_{d_{50}}^2 + \sigma_R^2} & \frac{4\pi R}{5\Delta^2} \\ \frac{4\pi R}{5\Delta^2} & \frac{2\pi(\sigma_{d_{50}}^2 + \sigma_R^2)}{B\Delta^2} \end{pmatrix}. \quad (27)$$

This matrix can be analytically inverted, yielding the following expression for  $\sigma_{CR}^2$ :

$$\sigma_{CR}^2 = [\mathbf{F}^{-1}]_{1,1} = \frac{75}{16\pi} \frac{(\sigma_{d_{50}}^2 + \sigma_R^2)^2}{R^5} \frac{1}{\gamma_1} \frac{e}{I\tau_0}. \quad (28)$$

This expression is used to compute  $\sigma_B^2$  in two limiting cases: (i) a sharply peaked prior distribution  $p(R)$ , and (ii) a very wide prior distribution.

For a sharply peaked prior distribution, the optimal probe size  $d_{B,50}$  is independent of the particles' atomic number and of the microscope settings. In that case,  $d_{B,50}$  linearly increases with the mean particle size  $\rho$ ,

$$d_{B,50} = 1.49\rho, \quad \text{for narrow } p(R). \quad (29)$$

For a wide prior distribution,  $d_{B,50}$  is independent of the mean radius of the particles and their atomic number, and depends only on the microscope parameters. It is proportional to the geometrically limited probe size  $d_{g,50}$ :

$$d_{B,50} = 1.51d_{g,50}, \quad \text{for broad } p(R). \quad (30)$$

This analytical approach reveals the functional dependence of the optimal probe size on the microscope parameters and on the specimen properties. It provides rules of thumb (Eqs. (29) and (30)) and proportionality constants that have the right order of magnitude. However, if exact values are needed, for distributions of arbitrary width, numerical simulations like in Section 4.4 are required.

### 4.4. Computational results

We carried out three simulation studies. First, we used a Monte Carlo simulation to verify that the variance of the estimated radii reaches  $\sigma_B^2$ . Second, we investigated the influence of the particle's atomic number and the detector geometry on the optimal probe size. Third, we studied the relation between the mean of the distribution of the radii and the optimal probe size.

#### 4.4.1. Setting the model constants

The modeled systems are spheres of aluminium (Al), zinc (Zn) and gold (Au). For each of these, we assume a log-normal distribution for the radii with mean  $\rho$  of 2.5, 5.0 or 10 nm. The particles are dispersed on a 10 nm thick uniform support of carbon (C). The microscope parameters in the simulations are summarized in Table 1.

The model constants  $\gamma_0$ ,  $\gamma_1$  and  $\tau_0$  are chosen such that they reflect the physical properties of the specimen. They are set such that the expected background count is 1 if one scans a particle with a radius equal to the mean radius  $\rho$ , laying on a support with a thickness  $t$  of 10 nm, using a probe of diameter  $d_{50}$  equal to  $1.25d_{g,50}$ , and with a dwell time  $\tau$  of 100  $\mu$ s. The constants are thus set by the following equations:

$$\tau_0 = \tau/\Delta^2, \quad (31)$$

$$\gamma_0 = \frac{e}{I\tau}, \quad (32)$$

and

$$\gamma_1 = \gamma_0 \frac{3}{5t} \left( \frac{Z_p}{Z_s} \right)^{1.7}. \quad (33)$$

The equations have been derived as follows. Eq. (31) follows from Eq. (18). Eq. (32) results from setting the  $A$ -term in Eq. (14) to (1). Eq. (33) expresses that in HAADF STEM the intensity is proportional to the projection of the specimen's atomic number raised to a power of typically 1.7. The ratio between the maximum of the intensity of the projected particle,  $q(0)$ , and the background,  $\gamma_0$ , is thus  $q(0)/\gamma_0 = (2R/t)(Z_p/Z_s)^{1.7}$ , with  $Z_p$  and  $Z_s$  the atomic numbers of the particle and the support, respectively.

#### 4.4.2. Monte Carlo simulation

The CRLB is the lower bound on the variance of any unbiased estimator. This means that in general the variance of an unbiased estimator will be higher than the CRLB. Optimal experimental design by minimizing the CRLB therefore only makes sense if there exists an estimator of which the variance equals the CRLB. The maximum likelihood (ML) estimator has this property at least asymptotically, that is, for the number of observations going to infinity. In this section we verify if this asymptotical property holds for the finite sample sizes one is confined to in practice.

The ML method for estimating the parameters  $\theta$  is as follows. The available observations  $\mathbf{g}$  are substituted in the probability density function in Eq. (19). Since the observations are numbers, the resulting expression depends only on the elements of the parameter vector  $\theta$ . The elements of  $\theta$ , the hypothetical true values, are now considered to be variables. To express this, they are replaced by  $\mathbf{t}$ . The logarithm of the resulting function,  $\ln P(\mathbf{g}; \mathbf{t})$ , is called the log-likelihood function. The ML-estimate  $\hat{\theta}_{ML}$  of the parameter vector  $\theta$  is defined as the vector that maximizes the log-likelihood

$$\hat{\theta}_{ML} = \underset{\mathbf{t}}{\operatorname{argmax}} \ln P(\mathbf{g}; \mathbf{t}), \quad (34)$$

i.e. it is the parameter vector that is most likely to have produced the observed measurement vector  $\mathbf{g}$ .

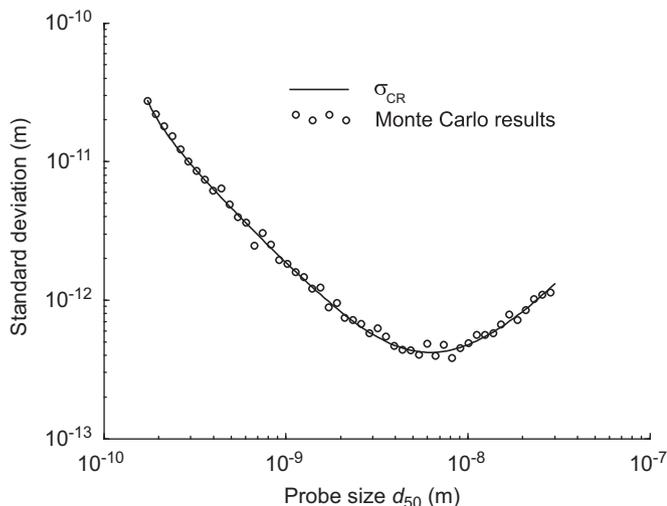
This is tested with a Monte Carlo (MC) simulation for Au particles with radius  $\rho = 5$  nm dispersed on a carbon support of 10 nm thickness. The settings for this experiment are given in the last column of Table 2. The variance  $\operatorname{var}(\hat{R}_{ML})$  is calculated out of 100 estimates for 50 values of  $d_{50}$  between 0.173 and 30 nm. As can be seen from Fig. 2 the experimental variance follows  $\sigma_{CR}^2$  over the entire range of probe sizes.

The absolute values of  $\sigma_{CR}$  in Fig. 2 are quite low. This is because an expected count of 1 in the background at the lowest probe diameter will yield very high counts at larger probes because the

**Table 2**

Values for the constants in the model for three different elements when the mean of the radii equals 5.0 nm.

	Al	Zn	Au
$Z_p$	13	30	79
$(Z_p/Z_s)^{1.7}$	3.72	15.4	80.0
$\tau_0$ (s m <sup>-2</sup> )	$6.50 \times 10^{14}$	$6.50 \times 10^{14}$	$6.50 \times 10^{14}$
$\gamma_0$	$1.03 \times 10^{-4}$	$1.03 \times 10^{-4}$	$1.03 \times 10^{-4}$
$\gamma_1$ (m <sup>-1</sup> )	$2.30 \times 10^4$	$9.50 \times 10^4$	$49.4 \times 10^4$



**Fig. 2.** The solid line shows  $\sigma_{CR}$  as a function of  $d_{50}$ . The circles give the standard deviation on the ML-estimate, as found by a MC simulation. The variance on  $\hat{R}_{ML}$  follows  $\sigma_{CR}^2$  over the entire range of probe sizes.

probe intensity relates to the probe diameter according to a power of 8/3, as follows from Eq. (5). However, here we are concerned with showing that the variance of a ML-estimate equals  $\sigma_{CR}^2$ , irrespective of the absolute values.

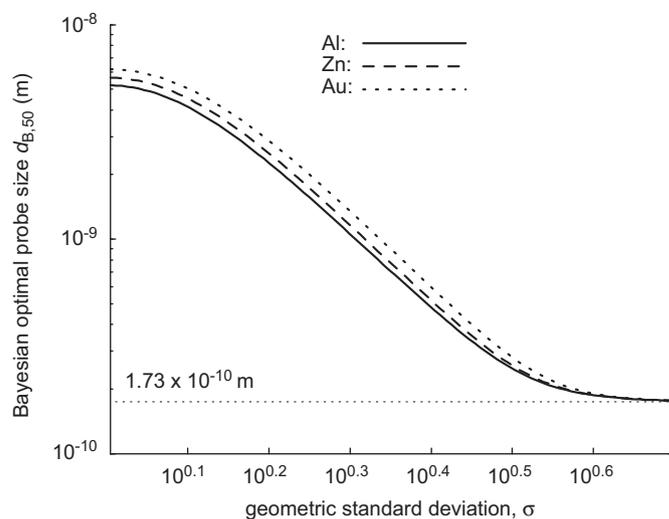
#### 4.4.3. Influence of the atomic number

The relation between the standard deviation  $\sigma$  of the prior distribution  $p(R)$ , and the Bayesian optimal probe size  $d_{B,50}$  is found by minimizing  $\sigma_B^2$  with respect to  $d_{50}$  for values of  $\sigma$  between 1.01 and 5. The mean particle radius  $\rho$  is set at 5.0 nm. In Fig. 3, the results for Al, Zn and Au are displayed. The behavior of the curves depends little on the atomic number of the particles. For sharply peaked distributions ( $\sigma = 1.01$ ),  $d_{B,50}$  is approximately equal to the mean particle radius: 5.22, 5.66 and 6.21 nm for Al, Zn and Au, respectively. For very wide distributions, the optimal diameter is constant and equal to 0.173 nm, or 1.25 times the geometrically limited probe size  $d_{g,50}$ . The input parameters are summarized in Table 2.

These results are in qualitative correspondence with the approach in Section 4.3, where the functional dependence of the optimal probe size on the microscope parameters and on the specimen properties has been revealed analytically. Moreover, the proportionality constants found here, are in the same order of magnitude as those given by the analytical rules of thumb. However, for exact values, for distributions of arbitrary width, numerical simulations like in this section are required.

#### 4.4.4. Influence of the detector geometry

The results in Section 4.4.3 can be extended to treat the influence of the detector geometry as well. In Ref. [14] it is shown that the image intensity in HAADF STEM scales with the atomic



**Fig. 3.** Relation between  $\sigma$ , the width of  $p(R)$ , and the Bayesian optimal probe size  $d_{B,50}$ . For a narrow  $p(R)$ ,  $d_{B,50}$  approximately equals the mean radius, while for broad  $p(R)$   $d_{B,50}$  is 1.25 $d_{g,50}$ . The atomic number of the particles is only of minor importance.

**Table 3**

Values for the constants in the model for three different means of the radii when the spheres are made of Al.

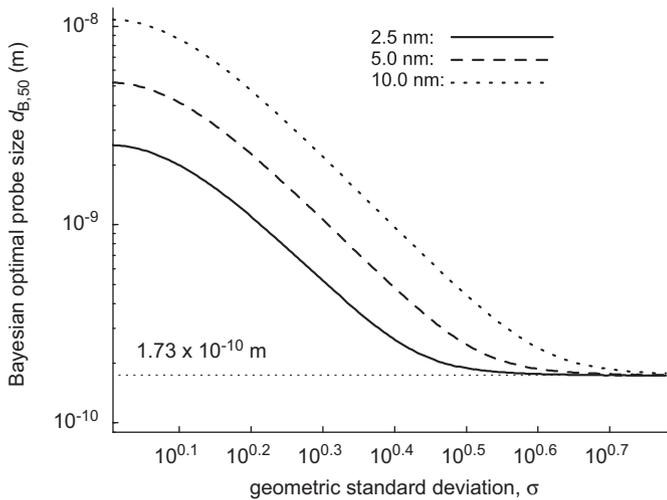
$\rho$ (nm)	2.5	5.0	10.0
$\tau_0$ (s m <sup>-2</sup> )	$25.9 \times 10^{14}$	$6.50 \times 10^{14}$	$1.63 \times 10^{14}$
$\gamma_0$	$1.03 \times 10^{-4}$	$1.03 \times 10^{-4}$	$1.03 \times 10^{-4}$
$\gamma_1$ (m <sup>-1</sup> )	$2.30 \times 10^4$	$2.30 \times 10^4$	$2.30 \times 10^4$

number raised to a power of  $\nu$ , where  $\nu$  depends on the detector's inner and outer radius. For a wide range of inner and outer radii,  $\nu$  lays between 1.5 and 1.9. In Eq. (33) we have chosen  $\nu$  equal to the mean of these two extremes, i.e.  $\nu = 1.7$ .

Section 4.4.3 shows that the optimal probe size depends little on the atomic number of the particle. This is mathematically equivalent with stating that it depends little on the exact value of the factor  $(Z_p/Z_s)^{1.7}$  in Eq. (33), since it is only here that the Z-dependence is taken into account. Furthermore, it makes no difference whether this factor changes due to a change in atomic number, or a change in the exponent. Therefore, the optimal probe size depends little on the detector geometry. Consequently, one should choose the inner radius as small as possible to maximize the electron count, but big enough to permit incoherent image formation. In the remainder of this paper  $\nu$  equals 1.7.

#### 4.4.5. Influence of the mean radius

In this section we investigate the influence of the mean radius  $\rho$  of the particles. The mean radii equal 2.5, 5.0 and 10.0 nm. In Section 4.4.3, it was shown that the atomic number of the particles is only of minor influence, so the Bayesian optimal probe size  $d_{B,50}$  is determined for values of  $\sigma$  between 1.01 and 6 for Al only. In Table 3, the parameters are summarized. It can be seen that the recording time per unit area,  $\tau_0$ , is different for different mean radii, making the absolute values of  $\sigma_B^2$  not comparable. However, as pointed out in Section 4.2,  $d_{B,50}$  is independent of  $\tau_0$ . The result is depicted in Fig. 4. For sharply peaked  $p(R)$ , the optimal probe size is approximately equal to the particle radius, i.e. 2.52, 5.22 and 10.8 nm for  $\rho$  equal to 2.50, 5.00 and 10.0 nm, respectively. For wide prior distributions  $p(R)$  (large  $\sigma$ ), the probe size converges to



**Fig. 4.** Optimal probe sizes as a function of  $\sigma$ . For a narrow  $p(R)$ , the Bayesian optimal probe size  $d_{B,50}$  approximately equals the mean radius, while for broad  $p(R)$   $d_{B,50}$  is  $1.25d_{g,50}$ . A comparison with Fig. 3 shows that the mean radius of  $p(R)$  is far more important than the atomic number.

**Table 4**

Bayesian optimal probe sizes  $d_{B,50}$  (nm) for three different elements and three different mean radii  $\rho$ . The standard deviation of  $p(R)$  is 1.01.

$\rho$ (nm)	2.5	5.0	10.0
Al	2.52	5.22	10.8
Zn	2.72	5.66	11.8
Au	2.99	6.21	12.8

$0.173 \text{ nm} = 1.25d_{g,50}$ , and becomes independent of the actual width of  $p(R)$ .

A wide distribution  $p(R)$  is equivalent to having little prior information about the sample. So by taking the limit of infinitely broad distributions, we find the optimal probe size for an unknown sample. This is reflected in the fact that then the optimal probe size is proportional to  $d_{g,50}$ , and thus depends on the microscope parameters  $\lambda$  and  $C_s$  only, and not on the mean radius or atomic number of the particles.

In Table 4, it is confirmed that the atomic number only has a minor influence on the optimal probe size. Values for  $d_{B,50}$  are listed for Al, Zn and Au in the case of  $\sigma = 1.01$ . In all cases  $d_{B,50}$  is approximately equal to  $\rho$ .

## 5. Conclusion

We used Bayesian experimental design to determine the probe size at which the throughput for measuring particle radii is maximized. The dependence of the probe size on the current determining the signal-to-noise ratio is explicitly taken into account, using the algorithm given in Ref. [3]. In addition, we provide a closed form approximation for this existing algorithm in the case of high acceleration voltages. The particles are modeled as spheres on a uniform support, and their radii are assumed to follow a log-normal distribution.

In case of a narrow distribution of radii, the Bayesian optimal probe size is proportional to the geometric mean radius. An analytical investigation predicted no influence of the element's atomic number. The numerical calculations revealed a marginal dependence on atomic number and detector geometry.

The case of broad radius distributions can be interpreted as having no prior knowledge about the specimen. Analytical investigation showed that in that case the Bayesian optimal probe size is proportional to only the geometrically limited probe size, and that it is independent of the atomic number of the particles, or the mean or width of the distribution of radii. Numerical computations confirmed this and showed that the optimal probe size is independent of detector geometry and equal to 1.25 times the geometrically limited probe size.

The solution of the problem of throughput maximization outlined in this paper, is well suited for an adaptive approach. If an unknown sample is inserted, one can start with measuring the particle radii with a probe of 1.25 times the geometrically limited probe. As one's knowledge of the radius distribution gradually increases, the probe size can be adapted continuously during the measurements. Adaptive approaches have been successfully used in experimental design for non-linear functions, see for example Ref. [23]. Investigating the added value of such approaches is an interesting avenue for future research.

## Acknowledgements

This work is carried out as part of the Condor project, a project under the supervision of the Embedded Systems Institute (ESI) and with FEI company as the industrial partner. This project is partially supported by the Dutch Ministry of Economic Affairs under the BSIK program.

S. Van Aert, P. Goos, and D. Van Dyck (project no. G.0188.08) gratefully acknowledge financial support from the Fund for Scientific Research—Flanders (FWO).

## Appendix A. Analytical calculations

In this appendix we give analytical approximations for the optimal probe diameter  $d_{B,50}$  in the limiting cases of (i) a sharply peaked prior distribution  $p(R)$ , and (ii) a very wide distribution.

### A.1. Derivation of the Fisher information matrix and the CRLB

In order to derive the results in Eqs. (27) and (28), we assume that the support's intensity is negligible with respect to the particle's, i.e.  $A$  in Eq. (14) is assumed to be 0. In the second row of Table 2 the ratios of these intensities are listed, and this indicates that this assumption is reasonable, especially for particles with a high atomic number. The model then becomes

$$\lambda(r) = B \exp \frac{-r^2}{2(\sigma_{d_{50}}^2 + \sigma_R^2)}. \quad (\text{A.1})$$

The dimension of the Fisher information matrix  $\mathbf{F}$  now is  $2 \times 2$  and  $\boldsymbol{\theta} = (R, B)^T$ . We use Eq. (22) to calculate the elements of  $\mathbf{F}$ . Since a Gaussian density function tends to zero quickly, the upper bound  $r_{\text{FOV}}$  of the integral in Eq. (22) is replaced by infinity, such that we can use the equality

$$\int_0^\infty x^n \exp(-ax^2) dx = \frac{k!}{2a^{k+1}}, \quad (\text{A.2})$$

for  $n = 2k + 1$ ,  $k$  an integer and  $a > 0$  [24]. The resulting Fisher information matrix is

$$\mathbf{F} = \begin{pmatrix} \frac{16\pi B}{25\Delta^2} \frac{R^2}{\sigma_{d_{50}}^2 + \sigma_R^2} & \frac{4\pi R}{5\Delta^2} \\ \frac{4\pi R}{5\Delta^2} & \frac{2\pi(\sigma_{d_{50}}^2 + \sigma_R^2)}{B\Delta^2} \end{pmatrix}. \quad (\text{A.3})$$

Inverting  $\mathbf{F}$  analytically yields

$$\sigma_{\text{CR}}^2 = [F^{-1}]_{1,1} = \frac{75}{16\pi} \frac{(\sigma_{d_{50}}^2 + \sigma_R^2)^2}{R^5} \frac{1}{\gamma_1} \frac{e}{I\tau_0}. \quad (\text{A.4})$$

We use this expression to compute  $\sigma_B^2$  with Eq. (25).

#### A.2. Optimal probe size $d_{B,50}$ in two limiting cases

To calculate  $\sigma_B^2$  for a narrow  $p(R)$ , we approximate  $p(R)$  in Eq. (25) by a Dirac delta function centred on  $\rho$  such that the sift-property leads to

$$\sigma_B^2 \simeq \frac{75}{16\pi} \frac{(\sigma_{d_{50}}^2 + \sigma_\rho^2)^2}{\rho^5} \frac{1}{\gamma_1} \frac{e}{I\tau_0}. \quad (\text{A.5})$$

If Eq. (5) is used, then Eq. (A.5) is minimized if

$$\frac{\partial}{\partial d_{50}} \left[ \frac{(\sigma_{d_{50}}^2 + \sigma_\rho^2)^2}{d_{50}^{8/3} - d_{g,50}^{8/3}} \right] = 0, \quad (\text{A.6})$$

or

$$\frac{1}{16\ln 2} d_{50}^{8/3} - \sigma_\rho^2 d_{50}^{2/3} - \frac{3}{16\ln 2} d_{g,50}^{8/3} = 0. \quad (\text{A.7})$$

The last term on the left-hand side is negligible such that

$$d_{B,50} \simeq 4\sqrt{\frac{\ln 2}{5}} \rho, \quad (\text{A.8})$$

$$= 1.49\rho. \quad (\text{A.9})$$

To describe the behavior of  $\sigma_B^2$  for wide  $p(R)$ , we note in Figs. 3 and 4 that the exact shape of  $p(R)$  is of no importance for wide prior distributions, and we approximate  $p(R)$  by a uniform distribution between  $\rho/\sigma$  and  $\rho\sigma$ . Evaluating Eq. (25) analytically and keeping only the highest order term in  $\sigma$  then yields

$$\sigma_B^2 \simeq \frac{75}{64} \frac{1}{\gamma_1 \pi \rho \tau_0 I} \frac{\sigma_{d_{50}}^4}{\rho^4} \sigma^3. \quad (\text{A.10})$$

This is minimized if

$$\frac{\partial}{\partial d_{50}} \left[ \frac{\sigma_{d_{50}}^4}{d_{50}^{8/3} - d_{g,50}^{8/3}} \right] = 0, \quad (\text{A.11})$$

where again Eq. (5) is used. Solving Eq. (A.11) yields

$$d_{B,50} = 3^{3/8} d_{g,50}, \quad (\text{A.12})$$

$$= 1.51 d_{g,50}. \quad (\text{A.13})$$

#### References

- [1] E.J. Kirkland, *Advanced Computing in Electron Microscopy*, Plenum Press, New York, 1998.
- [2] V. Intaraprasong, H.L. Xin, D.A. Muller, *Ultramicroscopy* 108 (2008) 1454.
- [3] J.E. Barth, P. Kruit, *Optik* 101 (1996) 101.
- [4] B.R. Frieden, *Physics from Fisher Information—A Unification*, Cambridge University Press, Cambridge, 1998.
- [5] S. Van Aert, A.J. den Dekker, D. Van Dyck, A. van den Bos, *Ultramicroscopy* 90 (2002) 273.
- [6] S. Van Aert, D. Van Dyck, A.J. den Dekker, *Optics Express* 14 (2006) 3830.
- [7] S. Van Aert, A.J. den Dekker, D. Van Dyck, *Micron* 35 (2004) 425.
- [8] S. Van Aert, A.J. den Dekker, A. van den Bos, D. Van Dyck, *Statistical experimental design for quantitative atomic resolution transmission electron microscopy*, in: P.W. Hawkes (Ed.), *Advances in Imaging and Electron Physics*, Academic Press, San Diego, vol. 130, pp. 1–164.
- [9] A.J. den Dekker, S. Van Aert, D. Van Dyck, A. van den Bos, P. Geuens, *Ultramicroscopy* 89 (2001) 275.
- [10] A.C. Atkinson, A.N. Donev, *Optimum Experimental Designs*, Clarendon Press, Oxford, 1992.
- [11] J.E. Barth, M.D. Nykerk, *Nuclear Instruments and Methods in Physics Research A* 427 (1999) 86.
- [12] P. Kruit, M. Bezuijen, J.E. Barth, *Journal of Applied Physics* 99 (2006) 024315-1.
- [13] S. Van Aert, D. Van Dyck, *Philosophical Magazine B* 81 (2001) 1833.
- [14] P. Hartel, H. Rose, C. Dinges, *Ultramicroscopy* 63 (1996) 93.
- [15] A.J. den Dekker, J. Sijbers, D. Van Dyck, *Journal of Microscopy* 194 (1999) 95.
- [16] E. Bettens, D. Van Dyck, A.J. den Dekker, J. Sijbers, A. van den Bos, *Ultramicroscopy* 77 (1999) 37.
- [17] A.J. den Dekker, S. Van Aert, A. van den Bos, D. Van Dyck, *Ultramicroscopy* 104 (2005) 83.
- [18] M.G. Kendall, A. Stuart, *The Advanced Theory of Statistics—Inference and Relationship*, second ed., vol. 2, Charles Griffin and Company Limited, London, 1967.
- [19] H.A. Dror, D.M. Steinberg, *Technometrics* 48 (4) (2006) 520.
- [20] J. Yu, P. Goos, M. Vandebroek, *Transportation Research Part B: Methodological* 44 (10) (2010) 1268.
- [21] W.K. Brown, K.H. Wohletz, *Journal of Applied Physics* 78 (4) (1995) 2758.
- [22] T.T. Kostas, M.J. Hampden-Smith, *Aerosol Processing of Materials*, Wiley-VCH, New York, 1999.
- [23] H.A. Dror, D.M. Steinberg, *Journal of American Statistical Association* 103 (2008) 288.
- [24] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, Elsevier Academic Press, 2007.