

© [2009] IEEE. Reprinted, with permission, from Lei, Gang., Shao, Ke., Li, Yabin., Yang, Guangyuan., Guo, Youguang., Zhu, Jianguo. & Lavers J 2009, Bayesian Inversion Method and Its Information Determination for the Estimation of Particle Size Distribution in Ferrofluids', IEEE Transactions On Magnetics, Vol. 45, no. 10, pp. 3981-3984. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Technology, Sydney's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Bayesian Inversion Method and its Information Determination for the Estimation of Particle Size Distribution in Ferrofluids

Gang Lei¹, K. R. Shao¹, Y. B. Li², G. Y. Yang¹, Youguang Guo³, Jianguo Zhu³, *Senior Member, IEEE*, and J. D. Lavers⁴, *Fellow, IEEE*

¹College of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China

²School of Electronic Information, Zhongyuan University of Technology, Zhengzhou, 450007, China

³Faculty of Engineering, University of Technology, Sydney, N.S.W. 2007, Australia

⁴Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada

Using Gaussian numerical integration formula, the problem of estimating the particle size distribution (PSD) in ferrofluids can be converted into an electromagnetic inverse problem. Then we present two Bayesian analytical estimators, minimum mean square error estimator and maximum a posterior estimator, to reconstruct the PSD of magnetic particles. In the implementation, weighted minimum norm approach, maximum likelihood estimator and weighted least square estimator are employed to determine prior information for the unknown parameter. And we also present two methods to provide the noise information for the error term. Finally, using Monte Carlo method, we give a ferrofluid example to illustrate the efficiency of the proposed methods.

Index Terms—Bayesian estimator, ferrofluid, noise information, prior information, particle size distribution.

I. INTRODUCTION

A ferrofluid is a colloidal suspension of very fine magnetic particles. Since the particles are not uniform in size, the magnetic property of ferrofluid is strongly affected not only by the average particle size, but also by the particle size distribution (PSD) [1], [2]. So the PSD estimation problem is an important issue in fundamental research and industry application as well.

Many devices have been employed as a direct extraction technique to estimate the PSD of magnetic particles, such as transmission electron microscopy (TEM), X-ray diffraction and atomic force microscopy [3]. Meanwhile, many indirect extraction techniques are also widely used to solve this problem. And some of them are discussed with the consideration of magnetization data and curve. Generally, all these indirect methods can be classified according to whether some prior information are given or not. Singular value decomposition (SVD) technique is a classical inference method, which is discussed without any prior information [2].

However, TEM is complicated and time consuming. SVD technique is highly sensitive to the noise [2]. So some prior information about the PSD (such as lognormal distribution) are widely introduced in many researches [4]-[7]. And that prior information has been considered in the framework of Bayesian statistical inference method [6], [7].

As we know, the efficiency of Bayesian method depends a lot on the determination of the prior and noise information. But some prior information selection methods are not very reasonable from the physical point of view, such as maximum entropy method [6], [7]. And there are very few discussions about the determination of noise information. Unlike other papers, we first give a systemic discussion about the PSD estimation problem within the framework of Bayesian theory

in this work. Then we give some methods to provide prior and noise information for the Bayesian inversion process.

II. BAYESIAN ESTIMATION METHOD

According to the Langevin theory of superparamagnetism, the magnetization of ferrofluid can be expressed as an integral equation with respect of the given PSD [1]-[3]. However, there are many unavoidable errors in the actual measurement process, so the magnetization (M) in the external field (H) can be expressed as

$$M / M_s = \int L(x)p(D)dD + N(H), \quad (1)$$

where D is the diameter of the particle, $L(x) = \coth(x) - 1/x$ is the Langevin function ($x = I_s \pi D^3 H / 6kT$), k is the Boltzmann constant, T is the absolute temperature, I_s is the saturation magnetization of the bulk material, and M_s is the saturation magnetization of the particle system, $N(H)$ represents the error term in the external field H .

In the numerical calculation of the integral equation (1), Newton formula is always used as the numerical integration algorithm. However, we use Gauss formula in this work because its results can reach the same precision with less discrete points. And more importantly, the dimension of matrix which needs to be inverted can be reduced with Gauss formula. Then (1) has a discrete form as

$$M / M_s = B \sum_{i=1}^m [G_m L(I_s \pi D_i^3 H / 6kT) p(D_i)] + N(H), \quad (2)$$

where B is an integral coefficient, G_m s are the weight coefficient of Gauss formula. For n sample data, (2) can be simplified as a matrix notation.

$$\mathbf{y} = \mathbf{A}\mathbf{w} + \mathbf{n}, \quad (3)$$

where \mathbf{y} is a $n \times 1$ standardization magnetization vector, \mathbf{w} is a $n \times 1$ error vector, $\mathbf{\theta}$ is a $m \times 1$ unknown vector of PSD, \mathbf{A} is a $n \times m$ observation matrix with respect to $L(x)$, D and H .

$$\begin{aligned} \mathbf{y} &= [y_1, y_2, \dots, y_n]^T = [M_1, M_2, \dots, M_n]^T / M_s, \\ \boldsymbol{\theta} &= [\theta_1, \theta_2, \dots, \theta_m]^T = [p(D_1), p(D_2), \dots, p(D_m)]^T, \\ \mathbf{w} &= [w_1, w_2, \dots, w_n]^T = [N(H_1), N(H_2), \dots, N(H_n)]^T, \\ \mathbf{A} &= \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} f(D_1; H_1) & \dots & f(D_m; H_1) \\ \vdots & \ddots & \vdots \\ f(D_1; H_n) & \dots & f(D_m; H_n) \end{bmatrix}, \end{aligned}$$

where $f(D_i; H_j) = L(I_s \pi D_i^3 H_j / 6kT) \Delta D$, ΔD is the step size.

Given the experimental data \mathbf{A} and \mathbf{y} , the PSD and its statistical property can be estimated with the following two Bayesian estimators.

A. Minimum mean square error estimator

Proceeding to the minimum mean square error estimator, we assume that $\boldsymbol{\theta}$ is a random vector with probability density function $N(\boldsymbol{\mu}_0, \mathbf{C}_0)$, \mathbf{w} is a noise vector independent of $\boldsymbol{\theta}$ and have probability density function $N(\mathbf{0}, \mathbf{C}_w)$, that is

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^n [\det(\mathbf{C}_w)]}} \exp \left\{ -\frac{(\mathbf{y} - \mathbf{A}\boldsymbol{\theta})^T \mathbf{C}_w^{-1} (\mathbf{y} - \mathbf{A}\boldsymbol{\theta})}{2} \right\}. \quad (4)$$

Then from Bayesian theory, the posterior probability density function of $\boldsymbol{\theta}$ is

$$p(\boldsymbol{\theta}|\mathbf{y}) \propto p(\boldsymbol{\theta}) p(\mathbf{y}|\boldsymbol{\theta}). \quad (5)$$

The minimum mean square error estimator of $\boldsymbol{\theta}$ is the mathematical expectation of $p(\boldsymbol{\theta}|\mathbf{y})$. Then we have

$$\hat{\boldsymbol{\theta}}_{\text{MMSE}} = \mathbf{C}_0 \mathbf{A}^T (\mathbf{A} \mathbf{C}_0 \mathbf{A}^T + \mathbf{C}_w)^{-1} \mathbf{y}, \quad (6)$$

and covariance is

$$\text{Cov}(\hat{\boldsymbol{\theta}}_{\text{MMSE}}) = \mathbf{C}_0 \mathbf{A}^T (\mathbf{A} \mathbf{C}_0 \mathbf{A}^T + \mathbf{C}_w)^{-1} \mathbf{A} \mathbf{C}_0. \quad (7)$$

As we know, the above estimator is the minimum variance unbiased estimator, which is the one whose variance for each component is minimum among all unbiased estimators [8]. It is an optimal estimator in the statistical sense.

B. Maximum a posterior estimator

Maximum a posterior estimator (MAP) is a very commonly estimator in Bayesian statistics. It is derived from choosing a feasible parameter to maximize the posterior probability density function. That is to say,

$$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} \{ p(\boldsymbol{\theta}|\mathbf{y}) \}, \quad (8)$$

Substitute the probability density function and we have

$$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} \left\{ \frac{1}{\sqrt{(2\pi)^n [\det(\mathbf{C}_w)]}} \exp \left\{ -\frac{(\mathbf{y} - \mathbf{A}\boldsymbol{\theta})^T \mathbf{C}_w^{-1} (\mathbf{y} - \mathbf{A}\boldsymbol{\theta})}{2} \right\} \right\}. \quad (9)$$

Because the peak location of the normal probability density function is equal to its mean, the MAP estimator is identical to (6). It should be noted that both methods have the same solution for this problem. But their starting points are very different; the former is the mean of posterior probability density function; while the latter is obtained from maximizing the posterior probability density function.

III. DETERMINATION OF PRIOR INFORMATION

There are two important issues in Bayesian inference method. One is the selection of prior information for the unknown parameter $\boldsymbol{\theta}$; the other is the determination of noise information for \mathbf{w} . In this section, we present three methods to provide the prior information.

A. Weighted minimum norm approach

Weighted minimum norm (WMN) approach has been successfully used in biomedical electromagnetic inverse problems, such as neuronal current sources localization and imaging [9]. It typically finds solutions which match the experimental data while minimizing a weighted 2-norm on the solution vector. Namely, it involves solving the constrained optimization problem as

$$\hat{\boldsymbol{\theta}}_{\text{WMN}} = \arg \min_{\boldsymbol{\theta}} \{ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 \}, \text{ subject to } \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 = 0. \quad (10)$$

where \mathbf{C}_0 ($m \times m$) is an arbitrary symmetric positive definite matrix, and the solution is

$$\hat{\boldsymbol{\theta}}_{\text{WMN}} = (\mathbf{A} \mathbf{C}_0 \mathbf{A}^T)^{-1} \mathbf{y}. \quad (11)$$

Now we give some discussion about the correlation of this approach and SVD technique. Using Cholesky decomposition of $\mathbf{C}_0 = \mathbf{W}\mathbf{W}^T$, then the WMN solution can also be expressed as the SVD form [2].

$$\hat{\boldsymbol{\theta}}_{\text{WMN}} = \sum_{i=1}^n (\mathbf{v}_i / s_i) \mathbf{u}_i, \quad (12)$$

where s_i , \mathbf{u}_i and \mathbf{v}_i are the i th singular value, left and right singular vectors of $\mathbf{A}\mathbf{W}$, respectively [9].

However, exact matching of the data in (10) results in many problems, such as the solution is ill-conditioning and highly sensitive to the noise. Fortunately, regularized forms of WMN methods can lower the noise sensitivity and one commonly used technique is termed as truncated SVD. In this method, the original problem (10) is converted into an unconstrained minimization of a combination of the residual error norm and weighted norm of the solution vector by a regularization method, and it has the form as

$$\hat{\boldsymbol{\theta}}_{\text{TSVD}} = \arg \min_{\boldsymbol{\theta}} \{ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 + \lambda \|\boldsymbol{\theta}\|^2 \}. \quad (13)$$

For any value of the regularization parameter λ , we can find the solution of (13) as

$$\hat{\boldsymbol{\theta}}_{\text{TSVD}} = \mathbf{A}_0 (\mathbf{A} \mathbf{C}_0 \mathbf{A}^T + \lambda \mathbf{I})^{-1} \mathbf{y}. \quad (14)$$

The optimal value of λ can be derived from some other methods, such as L-curve method. Furthermore, assuming that $\boldsymbol{\mu}_0 = \mathbf{0}$ and $\mathbf{C}_w = \lambda \mathbf{I}$ in (6), we can get the same results in (14).

From the above discussion, we can see that the results of SVD technique are the equivalent form of WMN estimators. Furthermore, WMN estimators can be expressed as a more succinct form. As all WMN estimators are the special cases of Bayesian estimators, we only discuss it in theory in this work.

B. Maximum likelihood estimator

Maximum likelihood estimation is an analytic maximization approach and a popular statistical method used for fitting a mathematical model to some data. Under the model

assumption of (3) and (4), we can find the maximum likelihood estimator (MLE) of θ by minimizing

$$L(\theta) = \frac{1}{2} (\mathbf{y} - \mathbf{A}\theta)^T \mathbf{C}_w^{-1} (\mathbf{y} - \mathbf{A}\theta). \quad (15)$$

This is a quadratic function of θ , and \mathbf{C}_w^{-1} is a positive definite matrix. Upon setting the gradient equal to zero and we can get the MLE as

$$\hat{\theta}_{\text{MLE}} = (\mathbf{A}^T \mathbf{C}_w^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_w^{-1} \mathbf{y}. \quad (16)$$

C. Weighted least square estimator

Least square method is a very popular technique used to find or estimate the optimal values of the parameters to fit a function to a set of data. It is perhaps the most widely used technique in engineering data analysis. Unlike maximum likelihood method presented above, it can be applied to any problem for which we don't know the general form of the joint probability density function. In practical application, weighted least square estimator (WLSE) is always used as an extension of least square estimator. Generally, we can get the WLSE of θ by minimizing

$$J(\theta) = (\mathbf{U}\mathbf{y} - \mathbf{U}\mathbf{A}\theta)^T (\mathbf{U}\mathbf{U}^T)^{-1} (\mathbf{U}\mathbf{y} - \mathbf{U}\mathbf{A}\theta), \quad (17)$$

where \mathbf{U} is a weighted matrix, and the solution is

$$\hat{\theta}_{\text{WLSE}} = (\mathbf{U}\mathbf{A}^T \mathbf{U}\mathbf{U}^T)^{-1} \mathbf{U}\mathbf{A}^T \mathbf{U}\mathbf{U}^T \mathbf{y}. \quad (18)$$

Obviously, if we let $\mathbf{U} = \mathbf{C}_w^{-1}$ in (18), we can get the same results of MLE in (16).

IV. DETERMINATION OF NOISE INFORMATION

In this section, we give two methods to offer the noise information for Bayesian inverse process. One is the unbiased estimation within the framework of statistical regression model; the other is the correlation matrix from Kriging method.

Firstly, assuming that all the noise terms are independent and have the same sampling distribution, that is to say,

$$N(H_i) \sim N(0, \sigma^2), \quad i = 1, 2, \dots, n. \quad (19)$$

In this assumption, model (3) is a normal multivariable linear regression model. \mathbf{C}_w is a diagonal matrix with variance σ^2 and zero covariance. The unbiased estimator of variance can be obtained from maximum likelihood method and it has the form as

$$\hat{\sigma}^2 = [(\mathbf{y} - \mathbf{A}\hat{\theta})^T \mathbf{C}_w^{-1} (\mathbf{y} - \mathbf{A}\hat{\theta})] / (n - m). \quad (20)$$

Moreover, we give an alternative method for some readers. This method is derived from Kriging method in which $\mathbf{C}_w = [c_{ij}]$ is a nonzero covariance matrix and $c_{ij} = \sigma^2 \mathbf{R}[r(z_i, z_j)]$, where \mathbf{R} is the correlation matrix, r is the user-specified correlation function [10]. It should be noted that this solution is more accurate but more complex than the former. As the determination of prior information for Bayesian method is more important than that of noise information, we focus our attention on the selection of prior information. In the following experiment, we use the former method to estimate the noise information.

V. EXPERIMENTS AND RESULTS

In order to illustrate the efficiency of the proposed methods, we give an example of the application to the PSD estimation in a ferrofluid. It follows the assumption that the PSD is a lognormal distribution, which is the most widely used in the research of ferrofluids.

This sample was chemically synthesized at room temperature from an aqueous solution containing the metal nitrates by coprecipitation method [5]. And TEM indicates that the PSD of this sample appears to be a lognormal distribution with a mean particle size of 10 nm and a standard deviation of 0.24 nm. It can be abbreviated as LN (10.0, 0.24). With the same experiment parameters (M_s is 19.2 emu/g, I_s is 84 emu/g), we can get the normalized magnetization data and curve.

It should be noted that the PSD reconstruction quality depends drastically on the magnitude (standard deviation σ) of experimental error as the ill-conditioned nature of this problem. And this is a very important issue in the fundamental research. So we use Monte Carlo method to illustrate the dependence of reconstruction quality on the experimental error magnitude σ . Three cases of error magnitude are considered, which are $\sigma = 0.1\%$, $\sigma = 0.5\%$ and $\sigma = 1\%$. Now we give the detailed reconstruction processes for each case of σ .

1) Use Monte Carlo method to generate 10^4 sets of experiment data (H and M) with an σ .

2) Set the range of D . In order to ensure the reconstruction quality of Bayesian method, the range of D should satisfy that $p(D) > 0.02$; otherwise the matrix may be singular in the matrix inversion process. Here we let D be in the range 6–15 nm.

3) Determine MLE and WLSE. As the special property of Langevin function, the sample points of D should not be too many to avoid the matrix to be singular [7]. Five points Gaussian formula was used in (2) in this work. Meanwhile, the same weighted matrix ($\sigma^2 \mathbf{I}$) is chosen in MLE and WLSE, and parameter σ^2 can be estimated by (20). It should be noted that the result given by MLE and WLSE is a multidimensional normal distribution. But the estimating PSD is a one dimensional lognormal distribution, so a conversion must be done here.

4) Obtain Bayesian estimator. Bayesian estimator can be gained with MLE and WLSE as its prior distribution. As the matrix which needs to be inverted in Bayesian estimation is less singular compared with that in MLE, we use seven points Gaussian formula in (2) to get the Bayesian estimator. Then we can easily obtain the Bayesian solution from (6). Noting that this solution is also a multidimensional normal distribution, a conversion was also done here.

From the above discussion, we can get the PSD estimation with different methods for each error magnitude σ . Actually, MLE and WLSE are also widely considered as classic statistical estimators. Now we give a comparison of all estimators.

Table I shows the root mean square error (RMSE) of these estimators. Here error means the absolute error between the

results of TEM and our estimators. Fig. 1 shows the RMSE of these estimators.

From the table and the figure, we can see that the RMSE of mean and standard deviation increases with the increase of error magnitude σ . And Bayesian estimators are superior to those of classic estimators. Obviously, the curves of Bayesian estimators are smoother compared with classic estimators, so the Bayesian estimators are less sensitive to the noise compared with classic estimators.

TABLE I
RMSE data of classic and Bayesian estimators
(CE means classic estimators; BE means Bayesian estimators)

| RMSE | Unit | $\sigma = 0.1\%$ | $\sigma = 0.5\%$ | $\sigma = 1\%$ |
|--------------------------|------|------------------|------------------|----------------|
| Mean of CE | nm | 0.0838 | 0.3706 | 0.6897 |
| Standard deviation of CE | nm | 0.0165 | 0.0713 | 0.1069 |
| Mean of BE | nm | 0.0478 | 0.1113 | 0.2000 |
| Standard deviation of BE | nm | 0.0073 | 0.0198 | 0.0308 |

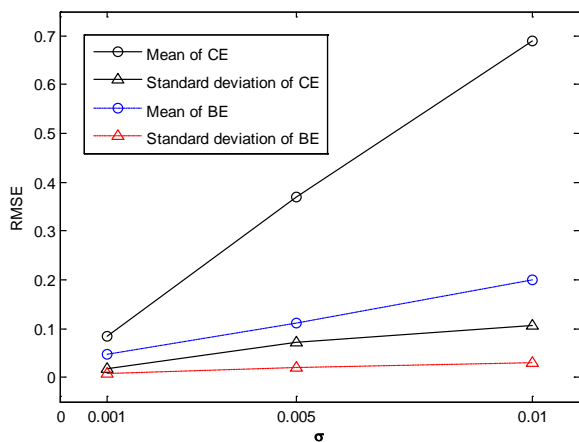


Fig. 1. RMSE curves with error magnitude

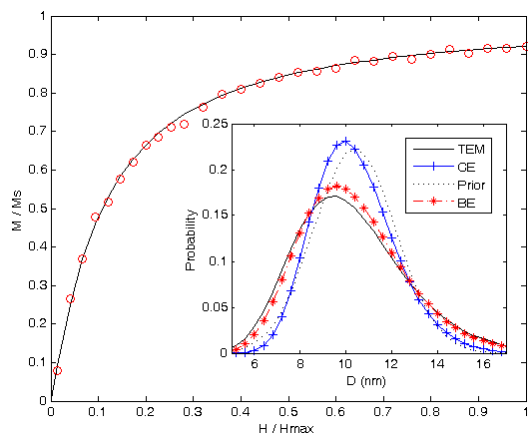


Fig. 2. Magnetization curve and PSD curve (insert)

Finally, as an example, we give an illustration for the case of $\sigma = 1\%$. Fig. 2 is the magnetization curve and magnetization

values (red o) of a sample. Obviously, there are some points which are not fitting the curve very well, and this is very reasonable for real experiments. With the proposed methods, the obtained PSD of MLE and WLSE is $LN(9.19, 0.36)$. After a conversion, prior distribution of μ_0 is $N(9.81, 3.68)$, it is a normal distribution. Then the Bayesian estimator is $LN(9.87, 0.25)$. Obviously, it is better than MLE and WLSE compared with the result of TEM. The PSD distribution curves are inserted in Fig.2.

VI. CONCLUSIONS

This work presents two Bayesian estimators to reconstruct the PSD of magnetic particles based on magnetization data. Meanwhile, we give some estimators to provide the prior and noise information, and we have presented the correlation of these estimators. Experiment results demonstrate that Bayesian method can be easily implemented and can induce satisfied results. And Bayesian estimators are superior to classical estimators, including the accuracy and the sensitivity to the noise. All these methods can be seen as effective direct extraction procedures of PSD from the magnetization curve and experimental data.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 50877029.

REFERENCES

- [1] S. Thurm and S. Oderbach, "Particle size distribution as key parameter for the flow behavior of ferrofluids," *Physics of Fluids*, vol. 15, no. 6, pp. 1658-1664, 2003.
- [2] D.V. Berkov, P. Gornert, N. Buske, et al, "New method for the determination of the particle magnetic moment distribution in a ferrofluid," *J. Phys. D: Appl. Phys.* vol.33, 331-337, 2000.
- [3] B. M. Lacava, R. B. Azevedo, L. P. Silva and Z. G. M. Lacava, "Particle sizing of magnetite-based magnetic fluid using atomic force microscopy: A comparative study with electron microscopy and birefringence," *Appl. Phys. Lett.*, vol.77, pp.1876-1878, 2000.
- [4] R. W. Chantrell, J.Popplewell and S. W. Charles, "Measurement of particle size distribution parameters in ferrofluids," *IEEE Trans. Magn.*, vol.14, no.5, pp. 975-977, 1978.
- [5] T. Kim and M. Shima, "Reduced magnetization in magnetic oxide nanoparticles," *J. Appl. Phys.* 101, 09M516,2007.
- [6] D. S. Xue and M. S. Si, "Bayesian inference approach to particle size distribution estimation in ferrofluids," *IEEE Trans. Magn.*, vol.42, no.11, 3657-3660, 2006
- [7] Gang Lei, K. R. Shao, G. Y. Yang, Jun Zhao, "Linear unbiased estimators for particle size distribution of magnetic nanoparticles," *J. Phys. D: Appl. Phys.* vol.41, 065002, 2008.
- [8] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall USA, 1993.
- [9] J. W. Phillips, R. M. Leahy and J. C. Mosher, "MEG-based imaging of focal neuronal current sources," *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 338-348, 1997.
- [10] L. Lebensztajn, C. A. R. Marretto, M. C. Costa, and J. L. Coulomb, "Kriging: a useful tool for electromagnetic device optimization," *IEEE Trans. Magn.*, vol. 40, no. 2, pp. 1196-1199, 2004.