

The Uncertainty in Neutron Diffraction Results Caused by Solving Systems of Linear Equations to Compute the Partial Structural Features

Imre Bakó^a, Tamás Grósz^b, Szabolcs Bálint^b, and Gábor Pálinkás^c

^a Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Pusztaszeri út 59-67, H-1025 Budapest, Hungary

^b Institute of Molecular Pharmacology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Pusztaszeri út 59-67, H-1025 Budapest, Hungary

^c Research Centre for Natural Sciences, Hungarian Academy of Sciences, Pusztaszeri út 59-67, H-1025 Budapest, Hungary

Reprint requests to I. B.; E-mail: bako.imre@tk.mta.hu

Z. Naturforsch. **68a**, 85–90 (2013) / DOI: 10.5560/ZNA.2012-0103

Received September 3, 2012 / published online February 15, 2013

Dedicated to Professor Alfred Klemm on the occasion of his 100th birthday

Neutron diffraction plays an important role in structural chemistry. In order to reveal the solution structure, the partial radial distribution functions have to be determined by using isotope substitution technique yielding different diffraction pattern while the structural parameters remain unchanged. The extraction of parameters from the series of measurements thus reduces to solving a system of linear equations that is affected by experimental errors. In this article, we give an estimation of the size of this error and also directions on how to minimize this effect by properly selecting the systems to be studied.

Key words: Neutron Diffraction; Experimental Error; Linear Equation System; Euclidean Norm.

1. Introduction

Neutron diffraction combined with isotopic substitution (NDIS) [1–3] is an extremely powerful method; it claims to yield unambiguous information about the local atomic structure in disordered materials. It has been applied successfully for many years, obtaining the partial structure factors of a wide range of liquids (i. e. water [4–7], ethanol [8], phenol [9], glucose [9], formic acid [10], ethanediol [11], liquid mixtures [12–14], aqueous solutions [15], and polymer electrolytes [16, 17] as well as of glassy materials. The radial correlation or structure functions extracted from NDIS experiments are very complicated. However, many of their features can be assigned and interpreted with the assistance of molecular dynamics (MD) or Monte Carlo (MC) simulations of the same systems under similar conditions. At the same time, the structure functions obtained can be used to test the quality of potential model applied in MD or MC simulation.

The general principles of isotopic substitution in neutron scattering are rather simple. Samples with dif-

ferent isotopic compositions (with markedly different coherent scattering lengths) yield different diffraction patterns while the underlying structural features remain unchanged. This method provides a convenient way for separating the partial terms in the composite structure factor. It has been applied for binary systems RX_n , where R denotes the central part of this system without substitution, and X stands for the isotope to be substituted. For example, in the case of water or methanol, the oxygen atom or the CD_3O group can be denoted as R , respectively.

The underlying mathematical formalism of this method is very straightforward. In order to determine three different partial structure factors or radial distribution functions (RR , RX , XX) in a binary system, at least three independent experiments have to be performed with different average scattering lengths for X . In practice, this means that one has at the end to solve a system of linear equations that has the following form:

$$f_i = W_{RR,i}RR + W_{RH,i}RX + W_{HH,i}XX, \quad (1)$$

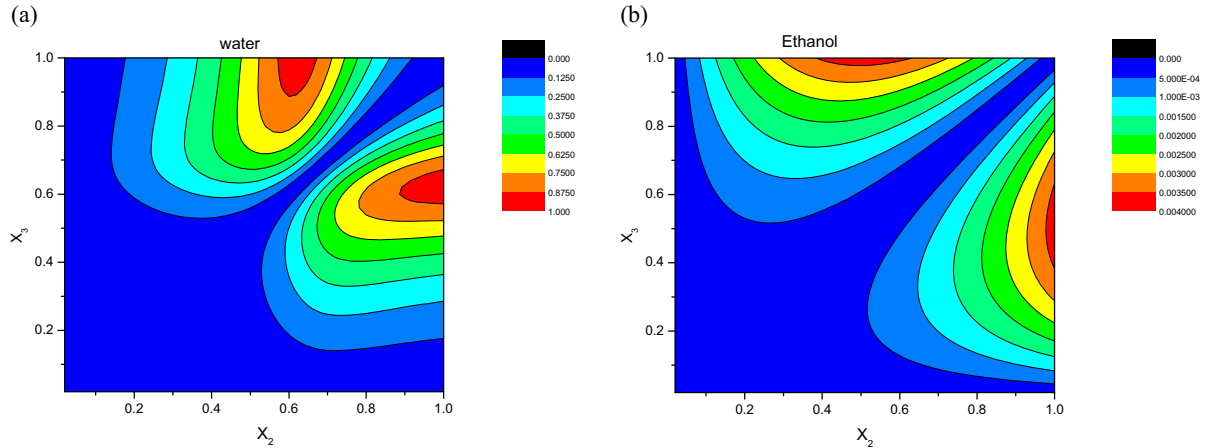


Fig. 1 (colour online). Contour plot of $\text{abs}(\det(W^{\text{norm}}))$ for water (a) and for ethanol (b).

where f_i is the measured data set of i th isotopic substitution, and W_{ij} are corresponding elements of the weighting matrix, e. g. for $X = H/D$:

$$\begin{aligned} W_{RR,i} &= b_{RR}^2, \\ W_{HH,i} &= (x_{H,i} \cdot b_H + (1 - x_{H,i}) \cdot b_D)^2, \\ W_{RH,i} &= 2 \cdot b_{RR} \cdot (x_{H,i} \cdot b_H + (1 - x_{H,i}) \cdot b_D), \end{aligned} \quad (2)$$

where $x_{H,i}$ is the mole fraction of the light hydrogen in the exchangeable part of i th isotopic substitution of a chemical substance.

Equation (1) can be written in matrix form as

$$f = W \times x, \quad (3)$$

where $x = (RR, RH, HH)^T$, T denoting the transposed matrix.

The aim of this study is to investigate how the solution of system (3) is affected by the magnitude of the experimental relative errors. These considerations can lead us to select the optimal parameter set for isotope substitution experiments.

2. Application of the Theory

It has been shown that a careful perturbation analysis is capable of providing a realistic assessment of the uncertainty and reliably indicates the sensitivity of the solution to experimental errors [18–21]. The following inequality holds between the uncertainty in the solution and the relative error of the experiment ($\| \delta X \|$

notes the corresponding vector and matrix norms):

$$\frac{\|\delta X\|}{\|X\|} \leq \|W\| \|W^{-1}\| \frac{\|\delta f\|}{\|f\|}. \quad (4)$$

The parameter κ_n ($\kappa_n = \|W\| \times \|W^{-1}\|$) is the so-called condition number of the equation; it quantifies the nearness to singularity: For values significantly larger than 1, the matrix is ill conditioned. Equation (4) shows that the condition number can be seen as a relative error magnification factor. The relative error in the computed solution X (not to be confused with the X above, which designates the isotope) may be much larger than the relative error in f if the condition number κ_n is large.

In practice, the most common vector norms are the l_1 , l_2 , and l_∞ (Manhattan, Euclidean, and Chebyshev norms). The vector norms can be defined in the following way:

$$\begin{aligned} \|x\|_1 &= \sum_i \text{abs}(x_i), \quad \|x\|_2 = \left(\sum_i x_i^2 \right)^{0.5}, \\ \|x\|_\infty &= \max(\text{abs}(x_i)) \end{aligned} \quad (5)$$

and the corresponding matrix norms have the form

$$\begin{aligned} \|W\|_1 &= \max_j \sum_i \text{abs}(W_{ij}), \quad \|W\|_2 = \left(\sum_{i,j} W_{ij}^2 \right)^{0.5}, \\ \|W\|_\infty &= \max_i \sum_j \text{abs}(W_{ij}) \end{aligned} \quad (6)$$

(matrix 1-norm, Frobenius norm, matrix ∞ -norm).

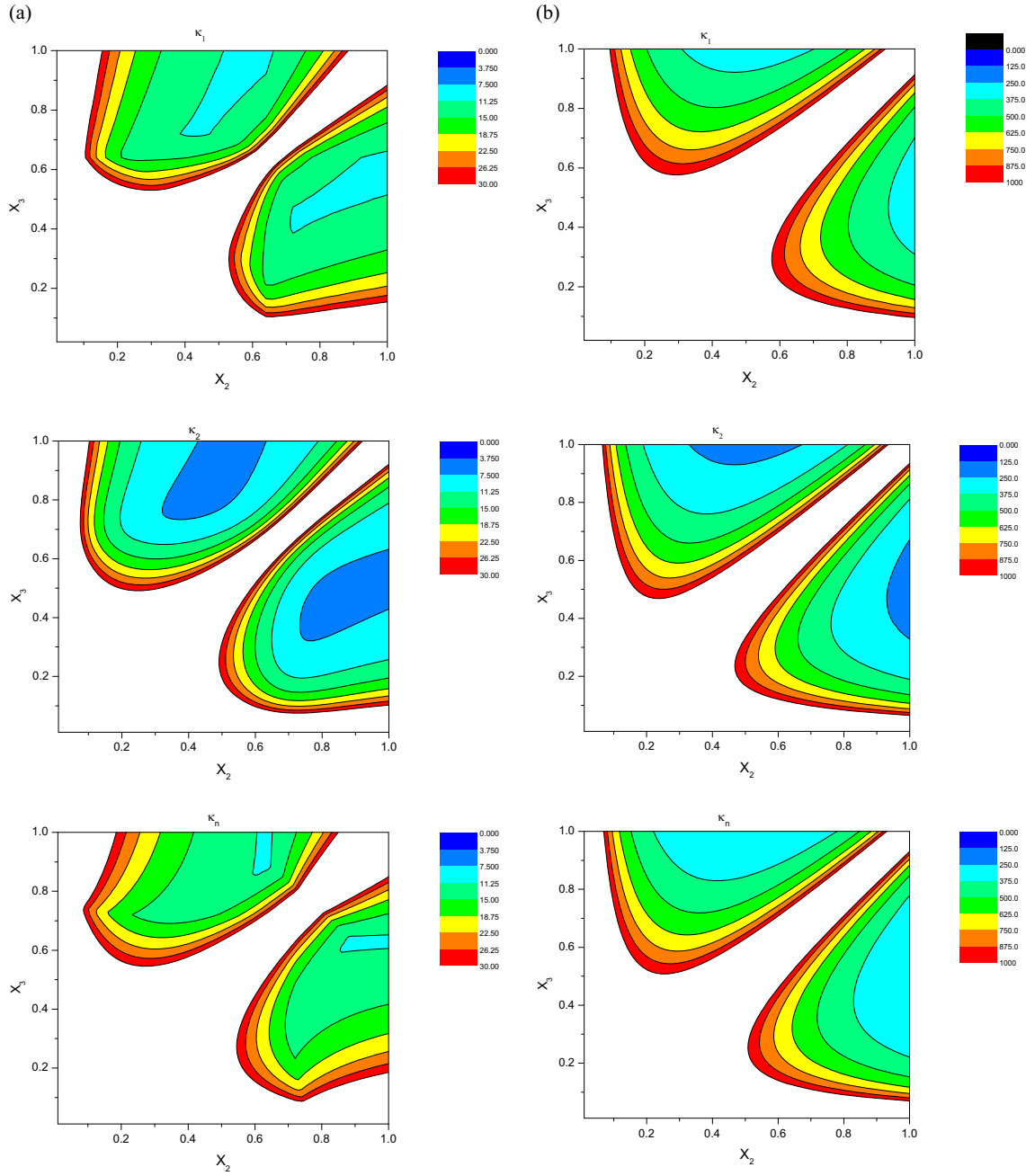


Fig. 2 (colour online). Contour plot of the κ_1 , κ_2 , and κ_∞ condition number for water (a) and for ethanol (b).

Furthermore, the condition number for the 2-norm can be expressed in the form

$$\kappa_2 = \frac{\sigma(\max)}{\sigma(\min)}, \quad (7)$$

where $\sigma(\max)$ and $\sigma(\min)$ are the largest and smallest singular values of the weighting matrix W . For overdetermined (more equations than unknown variables) systems the expression (7) is still valid while expression (4) holds for square matrices only.

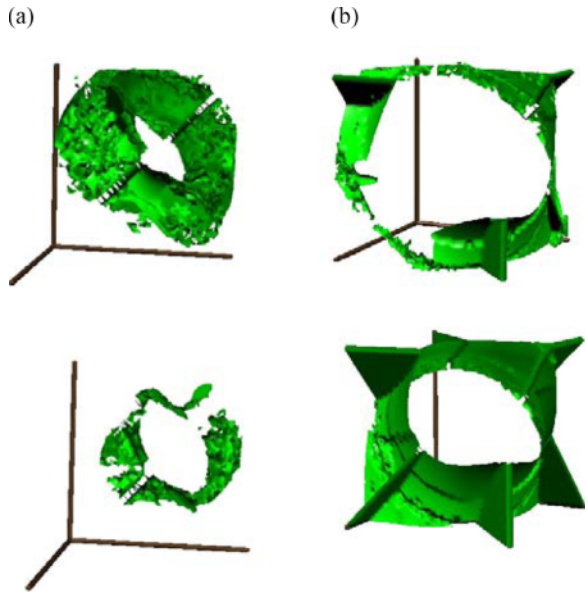


Fig. 3 (colour online). Three dimensional representation of the optimum values obtained for x_H . Points plotted represent the optimal values and those within a 10% (lower part) and 20% (upper part) range of deviance for water (a) and ethanol (b), respectively. The ends of the rods represent unity in the direction of the three axes ($x_{H,i}$).

According to our knowledge, these quantities have never been used to determine the reliability of the partial structure factors or the radial distribution functions of molecular liquids. McGreevy and Pusztai [22] applied the matrix norm corresponding to the l_1 vector norm for quantifying the relative information content of different total structure factors for glassy materials. In an earlier study Edwards et al. [23] defined the de-

terminant of the normalised weighting matrix as quantifying the conditioning of the equation.

An element of normalised matrix is defined in the form

$$W_{ij}^{\text{norm}} = \frac{W_{ij}}{(\sum_j W_{ij}^2)^{0.5}}. \quad (8)$$

An equation system is well-conditioned if the value of $\text{abs}(\det(W^{\text{norm}}))$ is close to 1; if this value is significantly smaller than 1, it is regarded as ill-conditioned.

The condition number corresponding to the l_2 Euclidean norm has already been applied for testing the accuracy of anomalous X-ray scattering by several authors [24–27]. We study here the properties of the κ_1 , κ_2 , and κ_∞ condition numbers and of $\text{abs}(\det(W^{\text{norm}}))$ in the case of liquid water, methanol, ethanol, formic acid, and 1,2-ethane-diol as examples of substances available for isotopic substitution on exchangeable (acidic) hydrogen.

3. Numerical Calculation and Results

Model calculations have been performed to obtain parameters which allow to optimise the x_H values for NDIS experiments on the substances mentioned above. The calculations were done using an in-house written Fortran programme. Three nested loop (for the three experiments) are used to determine numerically the κ_1 , κ_2 , κ_∞ , and $\text{abs}(\det(W^{\text{norm}}))$ quantities for each x_H set. The resolution of the grid defined by the loops was 0.01 ($0.00 \leq x_{H,i} \leq 1.00$ with 0.01 step size). From this dataset those x_H sets were selected where the corresponding conditional numbers and $\text{abs}(\det(W^{\text{norm}}))$

Table 1. Characteristic values of the numerical experiment: neutron scattering lengths (B in 10^{-14} m), optimal experimental compositions (x_H), extremum values of the condition numbers defined according to the various norms (κ_i), and $\text{abs}(\det(W^{\text{norm}}))$.

	Water	Methanol	Ethanol	Formic acid	Ethane diol
B_R	0.583	3.248	5.246	2.497	3.828
B_H	-0.774	-0.374	-0.374	-0.374	-0.774
B_D	1.334	0.667	0.667	0.667	1.334
$x_{H,1}$	0.00 0.00 0.00	0.00	0.00	0.00	0.00
$x_{H,2}$	0.62 0.44 0.64	0.50	0.50	0.50	0.50
$x_{H,3}$	1.0 0.81 0.89	1.0	1.0	1.0	1.0
κ_1	10.76	124.8	316	75.1	46.4
κ_2	6.77	83.07	215	49.2	29.2
κ_∞	10.80	113.44	258	74.5	49.2
$\text{abs}(\det(W^{\text{norm}}))$	0.0, 0.948, 0.62, 1.00	0.00, 0.014, 0.50, 1.0	0.00, 0.0037, 0.50, 1.0	0.00, 0.020, 0.50, 1.00	0.00, 0.061, 0.50, 1.0
x_H (4 data)	0.00, 0.52, 0.58, 0.88	0.00, 0.20, 0.78, 1.0	0.00, 0.22, 0.76, 1.0	0.00, 0.16, 0.72, 1.0	0.00, 0.66, 0.76, 1.0
$\kappa_{2/4}$	2.65	5.49	8.94	4.11	3.42

have their minimum and maximum values, respectively.

In this simplest case, the sizes of the f and x_H vectors of (3) are the same, i. e. three for the binary systems. In Table 1, the so-determined minimum values of κ_1 , κ_2 , κ_∞ , and the maximum value of $\text{abs}(\det(W^{\text{norm}}))$ as well as the corresponding hydrogen atomic mole fractions (x_H) are listed.

The three matrix norms used in the calculations resulted in three different hydrogen atomic mole fractions for the extreme values of κ in the case of water. For all other samples, the different norms yielded the same atomic mole fractions at the extreme values ($x_H = 1.00, 0.50, 0.00$). For water, it can be clearly seen that the optimal set of x_H for the Manhattan norm and the $\text{abs}(\det(W^{\text{norm}}))$ is the same: 1.00, 0.62 (close to the null-mixture: $x_{H,i} = 0.64$) and 0.00. The magnitude of conditional values increases by an order from water to methanol and ethanol.

Figures 1, 2a, and 2b represent the sensitivity of $x_{H,i}$ to the κ and $\text{abs}(\det(W^{\text{norm}}))$, respectively. It can be seen that the calculated conditional quantities are rather insensitive in a certain $x_{H,i}$ range. Outside this area of stability, these values change rapidly.

It is clear, that the size of the relative error of x_H can be lowered either by performing longer NDIS experiments or by decreasing the conditional number. Unfortunately, the beam-time is rather limited, thus it is more effective to determine the partial distribution functions from more than three experiments (overdetermined system).

A numerical calculation with an f vector of size four and x_H of size three has been performed. In the case of the NDIS experiment, this represent four differently substituted sample. The corresponding norm of this calculation is $\kappa_{2/4}$, calculated according to (7). It should be noted here that one of the four points is held fixed in the calculation, namely that of the purely deuterated sample ($x_{H,1} = 0.00$). The value of the con-

ditional number decreases by an order of one magnitude compared to the previous calculations. On the other hand, the addition of one more sample leads to a decrease of the large difference between relative errors for the various chemical substances. This result suggests that trying to decrease the conditional number is a more efficient strategy compared to longer NDIS measurements.

In Figure 3a and Figure 3b, the isosurfaces of $\kappa_{2/4,\text{min}} \cdot 1.05$ (lower part) and $\kappa_{2/4,\text{min}} \cdot 1.20$ (upper part) for water and ethanol are plotted, respectively. The length of the axes is unity and the free axes are representing $x_{H,i}$. Comparing the two parts of the figures, it can be seen that the sensitivity of $\kappa_{2/4}$ to the $x_{H,i}$ values is low.

For methanol, ethanol, and formic acid the minimum values of $\kappa_{2/4}$ resulted to be around 0.00, 0.20, 0.75 and 1.00 for x_H . For water and ethane diol, this x_H set turned out to be different, as can be seen in Table 1.

4. Conclusion

The main message of this work is as follows: It is recommended to perform more experiments with a reasonable amount of statistics rather than the minimum number of experiments with very good statistics. It should be mentioned that in this work the effect of correlation possibly existing between NDIS experimental data points is not taken into account. Thus, the relative error in the partials is only proportional to the calculated quantities.

A deeper investigation with more results will be the matter of a forthcoming article.

Acknowledgements

This work was supported by the Hungarian OTKA grant K83889. The authors thank Philippe A. Bopp for helpful discussion.

- [1] H. E. Fischer, A. C. Barnes, and P. S. Salmon, *Rep. Prog. Phys.* **69**, 233 (2006).
- [2] A. K. Soper and P. A. Egelstaff, *Mol. Phys.* **42**, 399 (1981).
- [3] J. Z. Turner, A. K. Soper, and J. L. Finney, *Mol. Phys.* **70**, 679 (1990).
- [4] A. K. Soper, *J. Phys. Condens. Matter* **19**, 335206 (2007).
- [5] A. K. Soper, *Chem. Phys.* **258**, 121 (2000).
- [6] L. Pusztai, *Phys. Rev. B* **60**, 11851 (1999).
- [7] A. K. Soper and M. G. Phillips, *Chem. Phys.* **107**, 47 (1986).
- [8] K. S. Sidhu, J. M. Goodfellow, and J. Z. Turner, *J. Chem. Phys.* **110**, 7943 (1999).
- [9] Y. Kameda, H. Ebata, T. Usuki, O. Uemura, and M. Misawa, *Bull. Chem. Soc. Jpn.* **67**, 3159 (1994).

- [10] I. Bakó, G. Schubert, T. Megyes, G. Pálinkás, G. I. Swan, J. Dore, and M. C. Bellisent-Funel, *Chem. Phys.* **306**, 241 (2004).
- [11] I. Bakó, T. Grósz, G. Pálinkás G, and M. C. Bellisent-Funel, *J. Chem. Phys.* **118**, 3215 (2003).
- [12] P. E. Mason, G. W. Neilson, C. E. Dempsey, D. L. Price, M. L. Saboungi, and J. W. Brady, *J. Phys. Chem. B* **114**, 5412 (2010).
- [13] I. Bakó, G. Pálinkás, J. C. Dore, and H. Fisher, *Chem. Phys. Lett.* **303**, 315 (1999).
- [14] S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney, and A. K. Soper, *Nature* **416**, 829 (2002).
- [15] G. W. Neilson and J. E. Enderby, *J. Phys. Chem.* **100**, 1317 (1996).
- [16] G. M. Mao, M.-L. Saboungi, D. L. Price, M. B. Armand, and W. S. Howells, *Phys. Rev. Lett.* **84**, 5536 (2000).
- [17] G. M. Mao, M.-L. Saboungi, D. L. Price, Y. S. Badyal, and H. E. Fischer, *Europhys. Lett.* **54**, 347 (2001).
- [18] J. J. Ducroz and N. J. Highham, *IMA J. Num. Anal.* **12**, 1 (1992).
- [19] L. El Ghaoui, *Lin. Alg. Appl.* **343**, 171 (2002).
- [20] W. W. Hager, *SIAM J. Sci. Stat. Comput* **5**, 31 (1984).
- [21] G. Strang, *Linear Algebra and its Applications*, Academic Press, New York 1980.
- [22] R. L. McGreevy and L. Pusztai, *J. Neutron Research.* **3**, 125 (1996).
- [23] F. G. Edwards, J. E. Enderby, R. A. Howe, and D. I. Page, *J. Phys. C. Solid State* **8**, 3483 (1975).
- [24] K. F. Ludwig, W. K. Warburton, L. Wilson, and A. I. Bienenstock, *J. Chem. Phys.* **87**, 604 (1987).
- [25] R. G. Munro, *Phys. Rev. B* **25**, 5037 (1982).
- [26] A. Zeidler, P. S. Salmon, R. A. Martin, T. Usuki, P. E. Mason, G. J. Cuello, S. Kohara, and H. E. Fischer, *Phys. Rev. B* **82**, 104208 (2010).
- [27] J. C. de Lima, D. Raoux, Y. Charriere, and M. Maurer, *J. Phys. Condens. Matter* **20**, 115103 (2008).