Maximum entropy method for diffractive imaging

Hiroyuki Shioya[1,*] and Kazutoshi Gohara[2]

[1]*Department of Computer Science and Systems Engineering, Muroran Institute of Technology, 27-1 Mizumoto-cho, Muroran 050-8585, Japan*

[2]*Division of Applied Physics, Graduate School of Engineering, Hokkaido University, N13, W8, Kita-ku, Sapporo 060-8628, Japan*

**Corresponding author: shioya@csse.muroran-it.ac.jp*

Received March 18, 2008; revised September 8, 2008; accepted September 8, 2008; posted September 15, 2008 (Doc. ID 93717); published October 24, 2008

Based on the minimization of the Lagrange formula, which is composed of two kinds of information measure, the maximum entropy method (MEM) is derived for diffractive imaging contaminated by quantum noise. This gives a suitable object corresponding to the maximum entropy principle with an iterative procedure. The MEM-based iterative phase retrieval algorithm with the initial process of the hybrid input–output (HIO-MEM) is presented, and a simple numerical example shows that the algorithm is effective for Poisson noise added to Fourier intensity. The relationship between the newly derived MEM for diffractive imaging and the conventional MEM for structure analysis based on crystallography is revealed. © 2008 Optical Society of America

OCIS codes: 100.5070, 100.3008.

1. INTRODUCTION

A phase problem generally arises in the measurement of waves; that is, the phase is generally missing while amplitude is observed as intensity. Sayre pointed out that the missing phase can be retrieved by measuring the intensity only in terms of Shannon's sampling theorem [1]. An iterative algorithm for phase retrieval was first presented by Gerchberg and Saxton [2]; then Fienup derived an error reduction (ER) algorithm based on the steepest-descent method and presented the hybrid input–output (HIO) algorithm [3]. These algorithms have been used in the fields of astronomy, optics, and both x-ray and electron microscopy. In materials science, the first instance of imaging using a soft x-ray diffraction pattern was presented by Miao *et al.*[4] This epochal result opens the door to structure analysis of noncrystalline materials without an objective lens. Following that, many related experiments were presented using different sources, such as x-ray [5–8], electron microscope [9–11], and tabletop light sources of laser [12]. These recent achievements are summarized in [13].

In the phase problem of x-ray crystallography, technical and theoretical methods for determining the periodic structures of materials have been introduced and developed. Although there are some resemblances between crystallography and general optical imaging, Millane pointed out in [14], p. 394, that "phase-retrieval theory and algorithms in crystallography and in general imaging have been developed fairly independently." As the methodology, the maximum entropy method (MEM) for crystallography introduced by Collins [15] was originally focused on reconstruction from incomplete noisy data [16,17]. This made the MEM a practical method for x-ray crystallography [18–22].

However, the relationship between the MEM for crystallography and the phase retrieval algorithms is still vague: the objectives of both are the same in terms of reconstructing a suitable object image, but the MEM for crystallography is used for superresolution from the noise data and not for phase retrieval directly. The relationship between an iterative algorithm for phase retrieval and the information-theoretic measures was pointed out in [23], revealing the importance of the MEM for diffractive imaging. One such information-theoretic measure was introduced for phase retrieval [24].

The MEM from incomplete and noisy data brings a feasible result that is helpful for diffractive imaging, because the noise in the Fourier intensity measurement generally influences the process of phase retrieval. Poisson noise due to quantum noise is inevitable in the measurement of diffraction intensity. Thus, the possibility of a phase retrieval algorithm providing an optimal solution to this noise is important for diffractive imaging.

In this paper, we introduce a Lagrange formalism with two function spaces defined on the object and Fourier domains, clarify the relationship between phase retrieval and the MEM, and for the first time, to our knowledge, present a MEM-based iterative phase retrieval algorithm with the initial process of the hybrid input–output (HIO-MEM). A simple numerical example using Fourier intensity contaminated by Poisson noise is also presented.

2. LAGRANGE'S METHOD FOR PHASE RETRIEVAL ALGORITHMS

Phase retrieval is represented as a corrective diagram of the Fourier and inverse Fourier transforms between the object domain and the Fourier domain, which is presented in the Gerchberg–Saxton iterative algorithm [2]. The reconstruction of the Fourier phase using the intensity measurement of the Fourier domain is presented as a cyclic transform. The prior object ρ is transformed into F by the Fourier transform FT; F is replaced by F' (the amplitude is given by the experiment in the Fourier domain

and the phase of F' is the same as that of F, while the replaced amplitude is the constraint in the Fourier domain); ρ' is obtained by the inverse Fourier transform FT^{-1} of F'; and ρ' is replaced by the updated object as the next ρ using some constraints in the object domain.

Based on the above cyclic transformation, let us define the Lagrange formalism for phase retrieval. We restrict ourselves to the case in which the object function is real and nonnegative; that is, the real positive condition for one of the object-domain constraints is assumed. As a mathematical presentation, let $\mathbf{F}_O$ be a function space generated by the set of all real functions with finite volume on the object domain O, and let $\mathbf{F}_K$ be a function space generated by the set of all finite-norm functions on the Fourier domain K. The Gerchberg–Saxton diagram is regarded as the transform back and forth between the two function spaces, $\mathbf{F}_O$ and $\mathbf{F}_K$ (Fig. 1). Thus, to express the changes of the transforms in the diagram, a criterion is needed for both spaces, simultaneously. Particularly, it is needed to introduce an information measure into the object domain for the estimation of a target object from uncertain conditions. To start, we introduce the minimization of the following Lagrange formula composed of two kinds of information measures for $\mathbf{F}_O$ and $\mathbf{F}_K$ as

$$\mathcal{L} = D + \lambda E, \tag{1}$$

where D and E are distance-like measures denoting the discrimination between the two elements of $\mathbf{F}_O$ and of $\mathbf{F}_K$, respectively, and λ is a coefficient.

Various information measures denoting the complexity or discrimination have been used in information theory [25–27]. If the similarity of two elements corresponds to the value of some information measure, it is suitable for the formulation of $D(\rho,\rho')$ and $E(F,F')$. In information theory, a distance-like measure between two probability distributions is called *information divergence*, and it is generalized for nonnegative real functions with finite volume [28].

As for two elements, we introduce the nth object ρ_n and $(n+1)$th object ρ_{n+1}, which are real nonnegative functions in $\mathbf{F}_O$, and the norm $|\rho_{n+1}-\rho_n|$ is assumed to be sufficiently small. This assumption is an underlying condition for mathematically deriving this algorithm. The relationship between the nth and $(n+1)$th objects denotes an iterative transform in the function space $\mathbf{F}_O$. In the function space $\mathbf{F}_K$, we introduce F_{n+1} and F'_{n+1}, where F_{n+1} denotes the Fourier transform of ρ_{n+1}, $|F'_{n+1}|$ is the observation $|F_{\mathrm{obs}}|$ of the Fourier domain, and the phase of F'_{n+1} is the same as that of F_{n+1}.

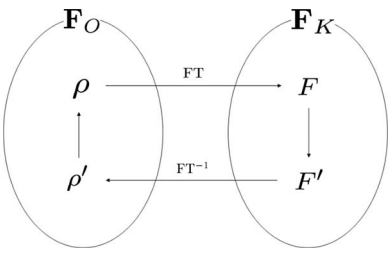

Fig. 1. Gerchberg–Saxton iterative diagram of phase retrieval based on two function spaces, $\mathbf{F}_O$ and $\mathbf{F}_K$.

Concerning the representation of the discrimination D, the I-divergence is well used in information theory [28], and the squared loss function is suitable for E, because the Fourier intensity is obtained and the phase of the Fourier domain is lost. From the above setting for D and E, the Lagrange formula Eq. (1) is concretely represented as

$$L = I(\rho_{n+1},\rho_n) + \lambda E(F_{n+1},F'_{n+1}), \tag{2}$$

where

$$I(\rho_{n+1},\rho_n) = \sum_{\boldsymbol{r}} \rho_{n+1}(\boldsymbol{r}) \ln \frac{\rho_{n+1}(\boldsymbol{r})}{\rho_n(\boldsymbol{r})} + \sum_{\boldsymbol{r}} \rho_n(\boldsymbol{r}) - \sum_{\boldsymbol{r}} \rho_{n+1}(\boldsymbol{r}), \tag{3}$$

$$E(F_{n+1},F'_{n+1}) = \frac{1}{M}\sum_{\boldsymbol{k}} ||F_{n+1}(\boldsymbol{k})| - |F'_{n+1}(\boldsymbol{k})||^2, \tag{4}$$

$F_{n+1}=|F_{n+1}|\exp(i\psi_{n+1})$, $F'_{n+1}=|F_{\mathrm{obs}}|\exp(i\psi_{n+1})$, and M is the cardinality of the domain K.

Fienup's relation gives the following representation of object functions [3]:

$$\frac{\partial E(F_{n+1},F'_{n+1})}{\partial \rho_{n+1}} = \frac{2}{M}(\rho_{n+1} - \rho'_{n+1}). \tag{5}$$

Under the constraint that $|\rho_{n+1}-\rho_n|$ be sufficiently small, $\partial L/\partial\rho_{n+1}=0$ gives the following iterative formula with a positive real constant ξ [23]:

$$\rho_{n+1}(\boldsymbol{r}) = \rho_n(\boldsymbol{r})\exp\{\xi(\rho'_n(\boldsymbol{r}) - \rho_n(\boldsymbol{r}))\}. \tag{6}$$

This algorithm gives a suitable object corresponding to the maximum entropy principle with an iterative procedure.

In this section, we have presented an iterative MEM phase retrieval algorithm based on the Lagrange method. However, there is no example using the iterative MEM algorithm for diffractive imaging. We next propose in Section 3 an algorithm using the MEM in order to establish its effective use for diffractive imaging.

3. NUMERICAL EXAMPLE

In this section, a numerical example using Poisson noise in the measurement of Fourier intensity is presented. When the difference between the nth object and the $(n+1)$th object is large, Fienup's HIO [3] or charge flipping [29] is often used; when the difference is small, the ER algorithm is used. Here we present a sample simulation using the HIO-MEM:

$$\rho_{n+1}(\boldsymbol{r}) = \begin{cases} \rho_n'(\boldsymbol{r}) & \boldsymbol{r} \notin S, \\ & n \leqslant n_0, \\ \rho_n(\boldsymbol{r}) - \beta\rho_n'(\boldsymbol{r}) & \boldsymbol{r} \notin S, \\ \rho_n(\boldsymbol{r})\exp\{\xi(\rho_n'(\boldsymbol{r}) - \rho_n(\boldsymbol{r}))\} & \boldsymbol{r} \notin S, \\ & n > n_0, \\ 0 & \boldsymbol{r} \in S, \end{cases} \tag{7}$$

where the HIO is used for an initial process with a suitable number of iterations n_0 and a pseudo-suitable object is obtained, and the MEM phase retrieval algorithm Eq. (6) is then used for the finishing process. Here S is the set of the points violating the object-domain constraint, and β and ξ are the coefficients. As a comparison, the HIO-ER is also presented as

$$\rho_{n+1}(\boldsymbol{r}) = \begin{cases} \rho_n'(\boldsymbol{r}) & \boldsymbol{r} \notin S, \\ & n \leqslant n_0, \\ \rho_n(\boldsymbol{r}) - \beta\rho_n'(\boldsymbol{r}) & \boldsymbol{r} \notin S, \\ \rho_n'(\boldsymbol{r}) & \boldsymbol{r} \notin S, \\ & n > n_0. \\ 0 & \boldsymbol{r} \in S, \end{cases} \tag{8}$$

The setting of the numerical example can be described as follows. The target ρ_{target} consists two same-height pillars shown at the upper left of Fig. 2 (the object domain is N pixels (N=256)). The observation with the Poisson noise is generated by the formula

$$\text{Poisson}\{|F^c_{\text{target}}(\boldsymbol{k})|^2\} \sim c^2|F^{\text{poisson}}_{\text{obs}}(\boldsymbol{k})|^2, \tag{9}$$

where F^c_{target} is obtained by the Fourier transform of $c\rho_{\text{target}}$ and c=0.0001. The expectation of the Poisson distribution is the Fourier intensity $|F^c_{\text{target}}|^2$. An initial object is a random object. In the first step, the HIO with a tight support is used, and its parameters are settled as β=0.5 and n_0=1000) for the HIO-MEM and the HIO-ER, respectively; next, two kinds of algorithms, the ER and a MEM phase retrieval algorithm (the parameter ξ=0.01), are used, and the HIO-ER and the HIO-MEM are compared. Five thousand iterations are provided after the initial process of these algorithms. The number of iterations for these algorithms is sufficient to obtain a suitable result. All parameters in the HIO-ER and the HIO-MEM including the iteration numbers n_0 and n are chosen as one of the feasible settings with some trials. The typical objects ρ_{mem} and ρ_{er} obtained by the HIO-MEM and the HIO-ER, respectively, are presented in Fig. 2. The shape of ρ_{er} shows two pillars; however, the influence of the noise is remarkable. The shape of ρ_{mem} is favorable and the noise is smaller than for ρ_{er} not only on the tops of the pillars but also between them. Generally, the noise in $|F_{\text{target}}|^2$ is an obstacle in obtaining the target object; thus a standard phase retrieval algorithm using the HIO and the ER does not give a feasible object. As a comparison between the HIO-ER and the HIO-MEM, the MEM-based algorithm with a sufficient number of iterations and suitable parameter settings gives a better result than that using the ER for the Fourier intensity contaminated by Poisson noise.

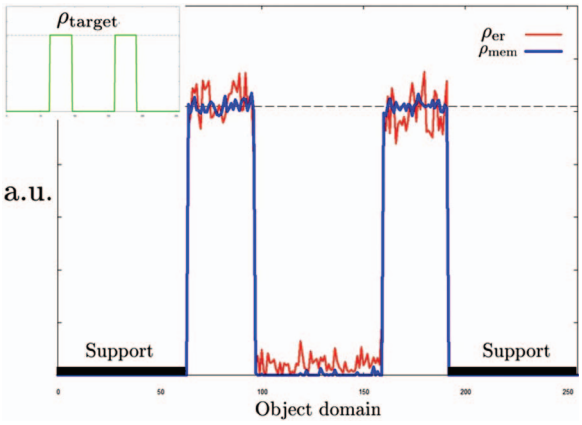

Fig. 2. ρ_{er} and ρ_{mem} are obtained by the HIO-ER and HIO-MEM, respectively. The figure on the upper left represents ρ_{target}.

Figure 3 shows schematic discrimination between ρ_{target} and ρ_{mem} and between ρ_{target} and ρ_{er} using I-divergence in the object-function space F_O. The I-divergence is well used as a distance-like measure between two positive real object functions [28]. Concerning the objects shown in Fig. 2, the ratio $I_{\text{mem}}/I_{\text{er}}$ is 0.054, where $I_{\text{mem}}=I(\rho_{\text{target}},\rho_{\text{mem}})$ and $I_{\text{er}}=I(\rho_{\text{target}},\rho_{\text{er}})$. Thus the I-divergence I_{mem} is less than I_{er} with a suitable setting of the HIO-MEM parameters. Needless to say, a large noise in the Fourier intensity does not give the characteristic result in the comparison between the HIO-ER and the HIO-MEM.

4. DISCUSSION

We focus now on the relationship between the MEM for diffractive imaging and the conventional MEM for structure analysis based on crystallography. The MEM was established in terms of statistical physics by Jaynes [30,31] and was introduced into the field of crystallography to produce electron density maps [15–22]. To investigate the relationship between the MEMs for crystallography conventionally used and those for phase retrieval derived, as far as we know for the first time, in this paper, let us re-

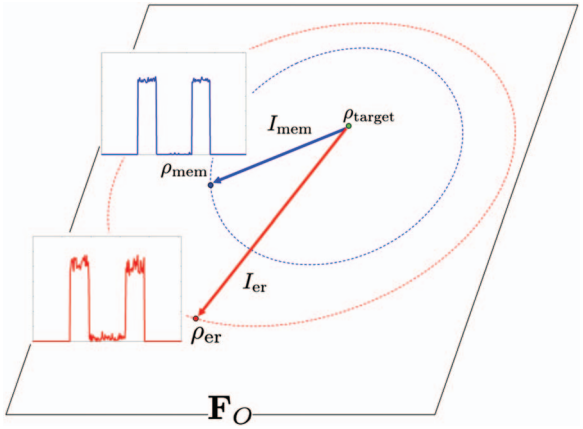

Fig. 3. Discrimination based on I-divergence between two object functions is shown in the function space F_O. The I-divergence I_{mem} (from ρ_{target} to ρ_{mem}) is less than I_{er} (from ρ_{target} to ρ_{er}) with a suitable setting of the parameters of the HIO-MEM.

view the MEM formalism of Collins [15]. τ is the obtained electron density by the structure factors $|F_{\mathrm{obs}}|\exp\{i\psi_{\mathrm{model}}\}$, where $|F_{\mathrm{obs}}|$ and ψ_{model} are the observed and the model-based phase, respectively. F_{cal} is the Fourier transform of ρ, and $|\hat{F}(0)|$ $(=|F_0|)$ is the number of electrons in a unit cell. Let us consider the maximization of

$$Q(\lambda) = -\sum_{\boldsymbol{r}} \rho'(\boldsymbol{r}) \ln \frac{\rho'(\boldsymbol{r})}{\tau'(\boldsymbol{r})} - \frac{\lambda}{2}\sum_{\boldsymbol{k}} \frac{|F_{\mathrm{cal}}(\boldsymbol{k}) - |F_{\mathrm{obs}}(\boldsymbol{k})|\exp[i\psi_{\mathrm{model}}(\boldsymbol{k})]|^2}{\sigma(\boldsymbol{k})^2} \quad (10)$$

by setting $\partial Q(\lambda)/\partial\rho(\boldsymbol{r}) = 0$, where the summation over $\boldsymbol{r}$ includes a unit cell; $\rho' = \rho/\Sigma\rho(\boldsymbol{r})$ and $\tau' = \tau/\Sigma\tau(\boldsymbol{r})$ are used in the entropy formula; the summation $\boldsymbol{k}$ is taken over the Fourier domain K; $\sigma(\boldsymbol{k})$ is a standard deviation of $|F_{\mathrm{obs}}(\boldsymbol{k})|$; and λ is a constant. The maximum entropy mapping is obtained as [15]

$$\rho(\boldsymbol{r}) = \exp\left\{ \ln \tau(\boldsymbol{r}) + \lambda F_0 \times \sum_{\boldsymbol{k}} \frac{(|F_{\mathrm{obs}}(\boldsymbol{k})|\exp[i\psi_{\mathrm{model}}(\boldsymbol{k})] - F_{\mathrm{cal}}(\boldsymbol{k}))}{\sigma(\boldsymbol{k})^2} \times \exp\{-i2\pi\boldsymbol{k}\cdot\boldsymbol{r}\} \right\}. \quad (11)$$

Thus the above equation presents the relationship between the prior object τ and the posterior object ρ; however, the procedure does not update the phase of the structure factors ψ_{model} introduced as the model-based phase. In crystallography, the phase-update procedure is not needed after the process of determining the model-based phase. Thus the relationship between the MEM and phase retrieval has not been treated, and this is a reason why the relationship between phase retrieval in general optics and in crystallography has not been discussed in the literature.

The variance $\sigma(\boldsymbol{k})$ is used in Eq. (11). However, some uncertain factors are contained in the observed $|F_{\mathrm{obs}}|$. Then $\sigma(\boldsymbol{k})$ is redundant in the Lagrange formulation. The relationship between two function spaces is introduced into our formalism for a MEM phase retrieval algorithm; thus such a harmonious formulation between the two function spaces gives the fundamental formalism for phase retrieval.

From this discussion, we conclude the following. The MEM for crystallography has been used to reconstruct images from incomplete and noisy data; focusing on such use, we express the MEM for diffractive imaging in terms of an iterative algorithm based on the Lagrange formalism for phase retrieval. Important future work will include the application of a MEM-based phase retrieval algorithm using experimental data. The numerical example in our paper is a simple setting in order to emphasize the effectiveness of the proposed algorithm for phase retrieval with Poison noise. The imaging of multiwall carbon nanotubes by diffraction microscopy using a 20 kV electron beam has been achieved experimentally by our research group [11], and we will apply the proposed algorithm to such experimental data in future research. Theoretical extension to complex object functions based on an information-theoretic analysis and introduction of the algorithm into the newly developed method using an ensemble of the estimated objects will also be subjects of our future work.

REFERENCES

1. D. Sayre, "Some implications of a theorem due to Shannon," Acta Crystallogr. **5**, 843 (1952).
2. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik (Stuttgart) **35**, 237–246 (1972).
3. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
4. J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," Nature (London) **400**, 342–344 (1999).
5. J. Miao, T. Ishikawa, B. Johnson, E. H. Anderson, B. Lai, and K. O. Hodgson, "High resolution 3D x-ray diffraction microscopy," Phys. Rev. Lett. **89**, 088303 (2002).
6. J. Miao, T. Ishikawa, E. H. Anderson, and K. O. Hodgson, "Phase retrieval of diffraction patterns from noncrystalline samples using the oversampling method," Phys. Rev. B **67**, 174104 (2003).
7. Y. Nishino, J. Miao, and T. Ishikawa, "Image reconstruction of nanostructured nonperiodic objects only from oversampled hard x-ray diffraction intensities," Phys. Rev. B **68**, 220101(R) (2003); http://www.nature.com/nphys/journal/v2/n12/index.html (October 2008).
8. H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege *et al.*, "Femtosecond diffractive imaging with a soft-x-ray free-electron laser," Nat. Phys. **2**, 839–843 (2006).
9. U. Weierstall, Q. Chen, J. C. H. Spence, M. R. Howells, M. Isaacson, and R. R. Panepucci, "Image reconstruction from electron and X-ray diffraction patterns using iterative algorithms: experiment and simulation," Ultramicroscopy **90**, 171–195 (2002).
10. J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara, "Atomic resolution imaging of a carbon nanotube from diffraction intensities," Science **300**, 1419–1421 (2003).
11. O. Kamimura, K. Kawahara, T. Doi, T. Dobashi, T. Abe, and K. Gohara, "Diffraction microscopy using 20 kV electron beam for multiwall carbon nanotubes," Appl. Phys. Lett. **92**, 024106 (2008).
12. R. L. Sandberg *et al.*, "Lensless diffractive imaging using tabletop coherent high-harmonic soft-x-ray beams," Phys. Rev. Lett. **99**, 098103 (2007).
13. J. C. H. Spence, "Diffractive (Lensless) imaging," in *Science of Microscopy*, P. W. Hawkes and J. C. H. Spence, eds. (Springer, 2007), Chap. 19.
14. R. P. Millane, "Phase retrieval in crystallography and optics," J. Opt. Soc. Am. A **7**, 394–411 (1990).
15. D. M. Collins, "Electron density images from imperfect data by iterative entropy maximization," Nature (London) **298**, 49–51 (1982).
16. S. F. Gull and G. J. Daniell, "Image reconstruction from incomplete and noisy data," Nature (London) **272**, 686–690 (1978).
17. S. F. Gull, A. K. Livesey, and D. S. Sivia, "Maximum entropy solution of a small centrosymmetric crystal structure," Acta Crystallogr. **43**, 112–117 (1987).
18. B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," J. Opt. Soc. Am. **62**, 511 (1972).
19. O. E. Piro, "Information theory and the 'phase problem' in crystallography," Acta Crystallogr. **A39**, 61–68 (1983).
20. W. Wei, "Application of the maximum entropy method to electron density determination," J. Appl. Crystallogr. **18**, 442–445 (1985).

21. A. Podjarny, D. Moras, J. Navaza, and P. M. Alizari, "Low-resolution phase extension and refinement by maximum entropy," Acta Crystallogr. **A44**, 545–551 (1988).
22. M. Sakata and M. Sato, "Accurate structure analysis by the maximum-entropy method," Acta Crystallogr. **46**, 63 (1990).
23. H. Shioya and K. Gohara, "Generalized phase retrieval algorithm based on information measures," Opt. Commun. **266**, 88–93 (2006).
24. K. Choi and A. Lanterman, "Phase retrieval from noisy data based on minimization of penalized I-divergence," J. Opt. Soc. Am. A **24**, 34–49 (2007).
25. C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. **27**, 379–423, 623–656 (1948).
26. S. Kullback and R. A. Leibler, "On information and sufficiency,"Ann. Math. Stat. **22**, 79–86 (1951).
27. I. Csiszár, "On topological properties of *f*-divergences," Stud. Sci. Math. Hung. **2**, 329–339 (1967).
28. I. Csiszár, "Why least squares and maximum entropy?An axiomatic approach to inverse problems," Ann. Stat. **19**, 2033–2066 (1991).
29. G. Oszlanyi and A. Suto, "*Ab initio* structure solution by charge flipping," Acta Crystallogr. **60**, 134–141 (2004).
30. E. T. Jayes, "Information theory and statistical mechanics," Phys. Rev. **106**, 620–630 (1957).
31. E. T. Jayes, "Information theory and statistical mechanics. II," Phys. Rev. **108**, 171–175 (1957).