

**Super-Resolution Image Reconstruction by Nonlocal Means
applied to High-Angle Annular Darkfield Scanning Transmission
Electron Microscopy (HAADF–STEM)**

Peter Binev,^{*} Francisco Blanco-Silva,[†] Douglas Blom,[‡]
Wolfgang Dahmen,[§] Robert Sharpley,[¶] and Tom Vogt^{**}

*Interdisciplinary Mathematics Institute,
University of South Carolina, Columbia, SC 29208^{††}*

(Dated: September 12, 2009)

Abstract

We outline a new systematic approach to extracting high resolution information from HAADF–STEM images which will be beneficial to the characterization of beam sensitive materials. The idea is to treat *several, possibly many* low electron dose images with specially adapted digital image processing concepts at a minimum allowable spatial resolution. Our goal is to keep the overall cumulative electron dose as low as possible while still staying close to an acceptable level of physical resolution. We shall present the main conceptual imaging concepts and restoration methods that we believe are suitable for carrying out such a program and, in particular, allow one to correct special acquisition artifacts which result in blurring, aliasing, rastering distortions and noise.

PACS numbers: 07.05.Pj,68.37.Ma

Keywords: High resolution imaging for STEM, iterated nonlocal means

Modern electron microscopic imaging has reached resolutions significantly better than 100pm which allows for unprecedented measurements of the composition and structure of materials [1–3]. However, one faces several severe obstacles to fully exploiting the information provided by aberration-corrected instruments. On the one hand, one needs to constantly remediate and reduce environmental perturbations such as air flow, acoustic noise, floor vibrations, AC and DC magnetic fields, and temperature fluctuations. On the other hand, high resolution and a good signal to noise ratio requires a high density of electrons per square nanometer. Unfortunately, soft materials are very susceptible to beam damage, and can only be visualized with low dose beams, resulting in poor resolution and a prohibitively low signal to noise ratio. Imaging matter using electron microscopes, in particular STEM (scanning transmission electron microscopes, see [4], [5]), will become increasingly important in the near future, especially in biology.

Our goal is therefore to compensate for the required lower dose by more sophisticated image processing techniques. In what follows we outline a strategy addressing this question. The guiding aspects for our approach can be summarized as follows: Rastering of the beam across the sample enables certain electron imaging and spectroscopic techniques such as mapping by energy dispersive X-ray (EDX) spectroscopy, electron energy loss spectroscopy (EELS) and annular dark-field imaging (ADF). These signals can be obtained simultaneously, allowing direct correlation of image and spectroscopic data. By using a STEM and a high-angle annular detector, it is possible to obtain atomic resolution images where the contrast is directly related to the atomic number ($\approx Z^2$). This is in contrast to conventional high resolution electron microscopy, which uses phase-contrast, and therefore produces results which need simulation to aid in interpretation. As for beam sensitivity, a critical issue in electron microscopy is the amount of dose needed to produce an image. Higher dose scans can damage the specimen while lower dose scans result in very low signal to noise ratio. In STEM mode, the electron dose onto the sample can be controlled in a variety of ways. The number of electrons per unit time can be varied by changing the demagnification of the electron source through the strength of the first condenser lens. The dwell time of the probe is typically varied between $7\mu\text{s}$ and $64\mu\text{s}$ per pixel in practice, although a much larger range is possible. The size of the image can be varied from a very small number of pixels in a frame (256×256) to over 64 million pixels per image (8192×8192). Finally, the magnification of the image sets the area of the specimen exposed to the electrons and

thereby affects the dose per unit area onto the specimen.

Images produced by electron microscopes offer only an indirect reflection of reality. One measures the distribution of the intensity of electron scattering at a detector. These intensities depend upon the structure and composition of the sample, the information transfer properties of the microscope as well as uncontrolled perturbations by external stimuli.

This has two major conceptual consequences. The first one concerns the overall goal. We do *not* attempt to develop techniques that aim at reaching a resolution that is higher than the one permitted by the hardware. When talking about *superresolution* we always mean to recover the level of resolution set by the microscope, but by using a *time series* of low resolution - viz. low dose - images. Let us briefly recall the standard way of producing high resolution reconstructions from a series of low resolution/noisy images. Several observation models that relate the original high-resolution image to the observed low-resolution micrographs have been proposed in the literature [6]. These are classically formulated as a global model (with local noise \mathbf{n}) of the form

$$\mathbf{y}_t = (D \cdot B_t \cdot M_t)\mathbf{x} + \mathbf{n}_t, \tag{1}$$

where \mathbf{x} is the desired high-resolution image which is assumed constant during the acquisition of the multiple micrographs, except for any motion and degradation allowed by the model. Therefore, the observed low-resolution images result from warping (M_t), blurring (B_t), and subsampling (D) operators performed on \mathbf{x} . It is also assumed that each micrograph is corrupted by additive noise \mathbf{n}_t only. We shall see below that this paradigm is not applicable in this form.

The second consequence, further enhanced by the fact that new concepts are needed, is that the methods developed are first backed by experiments with materials that exhibit very little beam sensitivity. In particular, inorganic materials allow us to validate our methods by comparing a reconstruction from low resolution images with a high resolution counterpart of the same object. Therefore we focus first on inorganic materials which we understand well and that have proven to be stable under HAADF-STEM conditions (see e.g. [7]). In particular, the M1 catalyst (Fig. 1), a Mo-V-Te-Nb-oxide, has various properties that lend themselves to our initial investigations: (1) it has well-understood contrast variations along the $\langle 001 \rangle$ projection, (2) beam-sensitive Te contained in pores of the metal oxide framework can be used to monitor electron beam-induced damage over time series while the surrounding

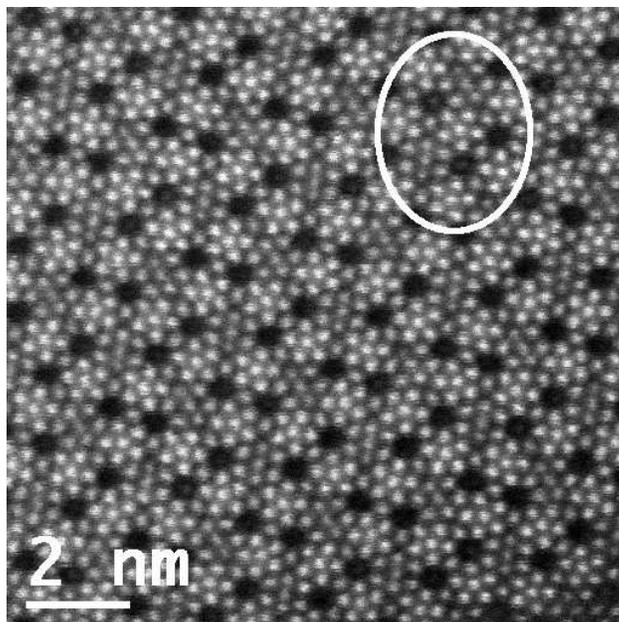


Figure 1: High-resolution HAADF STEM micrograph of the M1 catalyst.

structure does not deteriorate, and (3) defects that can be used as fiducials.

An example: In Fig. 1, the white oval shows pores in the metal oxide framework containing Te whose evaporation can be used to monitor long term exposure to electron beams. Thus, measuring time series of M1 at lower resolutions allows us to compare the reconstructions with micrographs taken at higher resolutions and thereby validate our algorithms and theoretical approaches which guide the treatment of more and more beam sensitive materials.

Of course, one would be able to reduce beam damage (in expectation) if the total accumulated dose used to produce several low resolution images could be kept even below the dose needed for a single high resolution image while still recovering the same information from the low resolution images. But even if in both scenarios the same total dose was necessary, the damage due to heating effects would clearly be smaller when taking successive low dose images. Whether a temporal stretching also has a beneficial relaxation effect on the other sources of beam damage is an open question on which the intended research may actually shed some light.

Such principle advantages come at a price however. The scanning process takes time during which the specimen moves due to electromagnetic, mechanical, or acoustic perturbations. The overall resulting motion is significant, even for a single frame, but all the more

so when taking several images of the same specimen.

Moreover, this motion is very complex. A global drift is typically overlaid by jitter. (For illustration, see Fig. 4 and the discussion below.) Tracking and estimating this motion by a sufficiently accurate model M_t in (1) based on low resolution possibly very noisy data, is not feasible. Thus, standard superresolution concepts as described above that are based on registration and motion tracking are not applicable.

We wish to propose an alternative strategy that is motivated by the above observations and can be summarized as follows.

(i) *Time series:* As before a high resolution image is to be recovered from a timeseries of HAADF STEM micrographs \mathbf{y}_t of the same object, where the “time” t is the frame index and runs through a finite set.

(ii) *The key tool:* Denoising based on time series is based on averaging the same specimen portion appearing in different frames. As explained above it is impossible to identify such portions from the low resolution frames. It is therefore crucial to employ a technique that avoids an explicit registration and motion tracking. The concept of *nonlocal means*, developed by Buades, Coll and Morel in [8] as a denoising algorithm, offers exactly this property. The key point is to average image portions whose intensity distributions are close to each other. This can be done as follows. With every pixel position p we associate a neighborhood N_p centered at p in as well as a (small) patch $R(p, t)$ in the frame \mathbf{y}_t also centered at p . We wish to produce an updated (target) value $z(p, t)$ at position p in the frame at t from source values $y(p', t')$ at positions p' in a (usually different) frame at t' by computing

$$z(p, t) = \frac{\sum_{p' \in N_p, t' \in N_t} w(p, p', t, t') y(p', t')}{\sum_{p' \in N_p, t' \in N_t} w(p, p', t, t')} \quad (2)$$

where N_t denotes a “time neighborhood” of t , i.e. a collection of timewise neighboring frames that are to be taken into account for the averaging process. Here the *weights* $w(p, p', t, t')$ have the form

$$w(p, p', t, t') := \exp \left\{ - \frac{\text{dist} (R(p, t), R(p', t'))^2}{2\sigma^2} \right\}, \quad (3)$$

where σ is another tuning parameter that stands for the variance of the data and the *distance* $\text{dist} (R(p, t), R(p', t'))$ between two patches is to quantify the *similarity* between those patches. The distance notion is a crucial parameter of such a scheme which, in particular, allows us to incorporate knowledge about data acquisition and special artifacts. We postpone the discussion of this issue and are content for the time being with the perhaps

simplest version which views the patch $R(p, t)$ as a vector of intensity values and applies the Euclidean norm to compare two patches

$$\text{dist} (R(p, t), R(p', t')) := \|R(p, t) - R(p', t')\|^2. \quad (4)$$

A few comments on the rationale of such schemes are in order. Obviously, in principle the weight assigned to a source value $y(p', t')$ is larger as the distance between the corresponding intensities for the respective patches is smaller, regardless of the spatial distance between the respective pixel positions. Thus, in contrast to conventional averaging techniques, closeness in the range is emphasized rather than in the domain, thereby enabling tracking of local jitter (see Fig. 2). The search for similar patches is only limited by the search neighborhood N_p which could be chosen as a complete frame. However, this would significantly increase the computational cost. On the other hand, too many elements participating in the averaging will increase blurring effects. Therefore, a judicious (data dependent) choice of the search neighborhoods as well as of the variance σ is important regarding quality and efficiency.

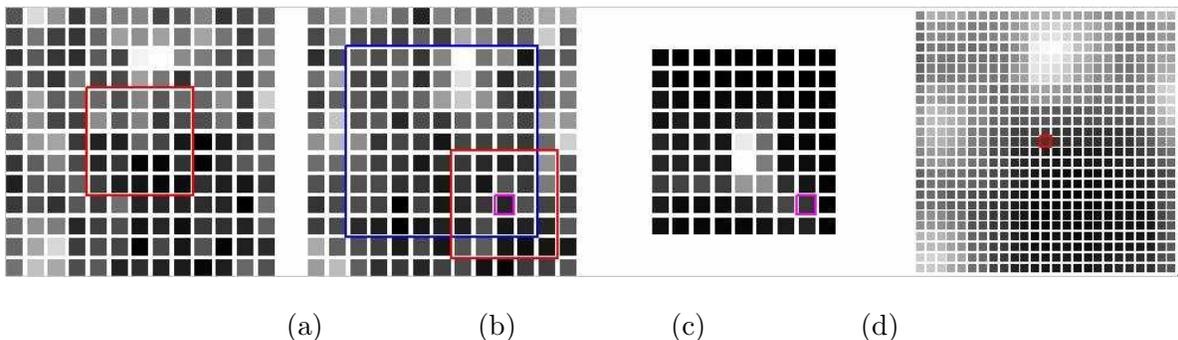
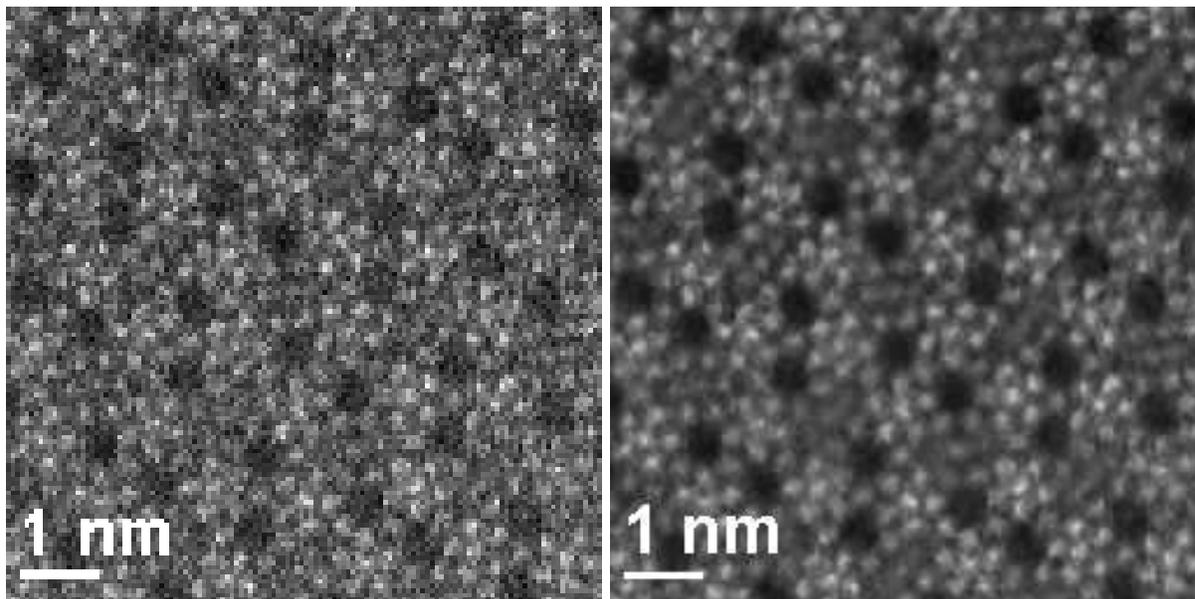


Figure 2: (Color online) (a) patch around central pixel (in red); (b) neighborhood (in blue) of central pixel hosting comparison patches; (c) support of weight function for the comparison patches which equals the neighborhood in (b); (d) position (in red) of updated pixel value in upscaled image.

While the main issue is to get rid of noise caused by low dose, a limited range of increased spatial resolution can be incorporated as well. The basic premise for increasing the spatial resolution is the availability of multiple images taken from the same object (which in typical applications represent different views of the same scene). The low-resolution images are subsampled (aliased) as well as distorted with subpixel precision. If the images have aliasing and different subpixel distortions, the new information contained in each low-resolution micrograph can be exploited to obtain the desired high-resolution image.

Of course, a concrete scheme based on nonlocal means requires a proper specification of all the parameters (patch shape/size, spatial neighborhood size, time neighborhood size, variance, distance) which will be elaborated on and analyzed in a forthcoming paper[9]. Figure 3 shows a typical result for the catalyst M1.



Original from series

Normalized (before deblurring)

Figure 3: Super-resolution reconstruction by nonlocal means of a timeseries of low resolution HAADF STEM micrographs of the catalyst M1 taken with exposure of $7\mu\text{s}$ per pixel.

(iii) *A two-stage approach:* Nonlocal means alone does not yet provide super-resolution. On the one hand, it should be complemented by a deblurring step which we shall briefly comment on below. On the other hand, in order to exploit the full potential of nonlocal means we propose an *iterated averaging procedure* which is outlined next.

Recall that the warping that occurs during the image acquisition in HAADF STEM may contain global and local translations, rastering distortion, local rotations, and so on, whose effect may grow over time and hamper the feature identification in subsequent images. As mentioned earlier a very low signal to noise ratio, as well as an unknown complex motion, prevent a reliable identification of alike patches in different frames. However, since the scanning of a low resolution frame takes relatively little time, one expects correspondingly milder motion effects within each single frame. Moreover, structures like the M1 catalyst exhibit a stable base pattern.

The first stage - Iterated nonlocal averaging - from spatial width to time depth: Therefore

a first normalization pass of nonlocal means should use a small time neighborhood N_t and a relatively large spatial neighborhood N_p with a simple weights and distance notion such as (3, 4). As a result one obtains a new time series of smoothed frames, see Fig. 3. These frames are now better suited for identifying alike image portions in consecutive frames and corresponding local distortions caused by small rotations or shearing. Now it makes sense to employ more subtle distance notions adapted to the specific features of STEM imaging. Namely, to replace the neighborhood $N_p \times N_t$, from which (p', t') is selected, with a (smaller) domain $\mathcal{N}(p, t)$ that respects the frame-to-frame motion detected in the first pass. For a detailed elaboration of this issue we refer to [9]. Iterating further passes of nonlocal means with improved similarity criteria, one can gradually decrease the size of spatial neighborhoods while increasing time neighborhoods so as to average eventually only image patches that correspond to each other. It is important to stress though that these iterative passes will always apply to the original data, just using upgraded information concerning *local* registration extracted from the intermediate frames. In a way such an iterative procedure may be viewed as gradually refining the image formation in HAADF STEM and modeling the distortions encountered during the imaging process. Moreover, from the possible change of the weights over time one may be able to learn more about beam damage.

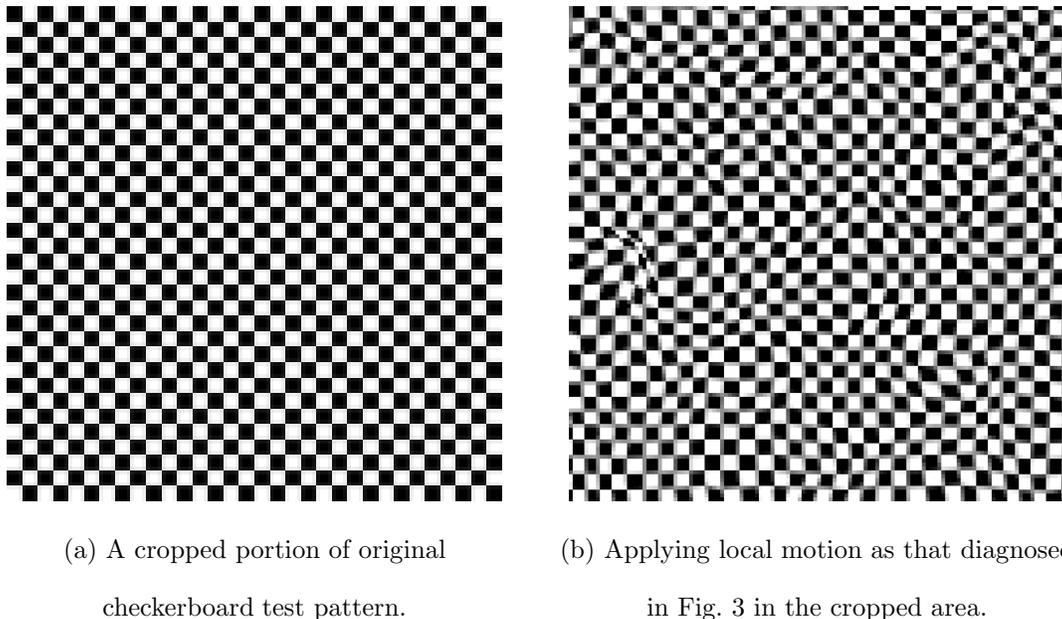


Figure 4: Effect of diagnosed local motion estimated by non-local means analysis when applied to a test pattern. The weights are those used for the images in Fig. 3.

The second stage - Deblurring: Finally the iteration should be concluded with a deblurring step, which aims at compensating the averaging effect introduced in the first stage. Here it is natural to employ a point spread function generated from the data and our first experiments look promising. Moreover, we plan to use *sparsity recovering techniques* based on regularization with ℓ_1 or total variation penalties (see [10]).

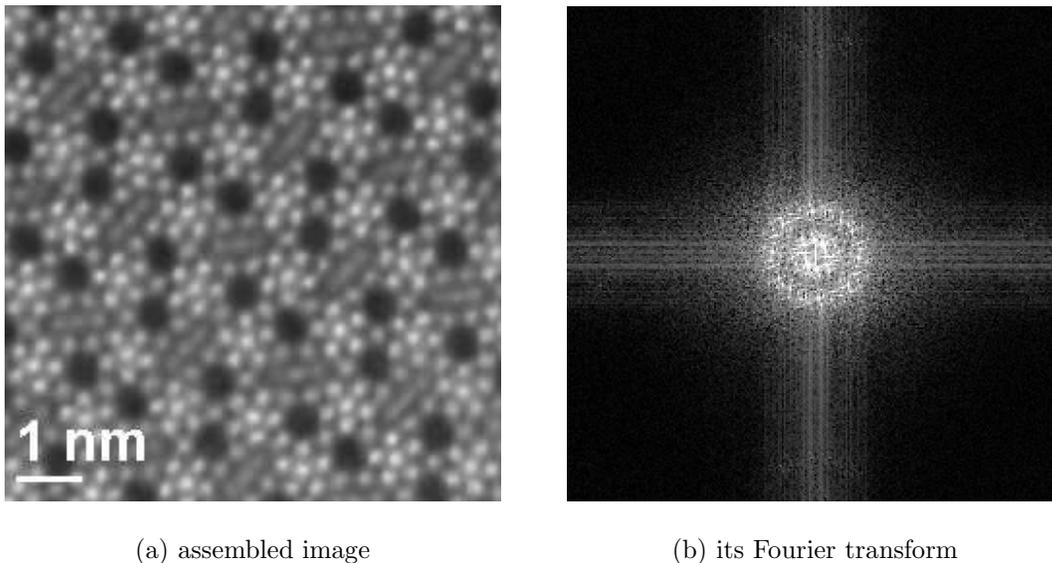


Figure 5: Result of assembling 11 images.

As an example of the success of our routine we present the result of the first stage assembling 11 consecutive low dose micrographs of M1 catalyst in Fig. 5. In (a) only the portion of the image corresponding to the ones from Fig. 3 is shown. It is remarkable that the Fourier transform of the result exhibits the same characteristics as the ones of the originals which cannot be claimed for vast majority of the image processing approaches.

We have sketched a new approach to processing STEM images so as to obtain high resolution information from time series of low resolution/low dose frames. Current research focuses on analyzing the effects and identifying suitable choices of the involved scheme parameters. The scheme will then be applied to more and more beam sensitive materials beginning with zeolites. Moreover, we emphasize that the method offers various diagnostic tools. For instance, the variation of the weights over time may shed some light on beam damage mechanisms and their causes. Applying the weights to simple checkerboard test patterns helps visualizing the motion of the specimen during the imaging process which is hoped to better understand it. This work will be reported in [9].

* binev@math.sc.edu; Department of Mathematics

† blanco@math.sc.edu; Department of Mathematics

‡ Doug.Blom@sc.edu; NanoCenter; Electron Microscopy Center

§ dahmen@igpm.rwth-aachen.de; Institut für Geometrie und Praktische Mathematik, RWTH Aachen

¶ sharpley@math.sc.edu; Department of Mathematics

** tvogt@mailbox.sc.edu; NanoCenter; Department of Chemistry and Biochemistry

†† This research was supported in part by the College of Arts and Sciences at the University of South Carolina, the Leibniz program of the German Research Foundation, MURI ARO Grant # W911NF-07-1-0185, and NSF Grant # DMS-0915104

- [1] C. Kisielowski, B. Freitag, M. Bischoff, H. van Lin, S. Lazar, G. Knippels, P. Tiemeijer, M. van der Stam, S. von Harrach, M. Steckelenburg, et al., *Microscopy and Microanalysis* **14**, 469 (2008).
- [2] R. Erni, M. D. Rossell, C. Kisielowski, and U. Dahmen, *Physical Review Letters* **102**, 096101 (2009).
- [3] H. Sawada, Y. Tanishiro, N. Ohashi, T. Tomita, F. Hosokawa, T. Kaneyama, Y. Kondo, and K. Takayanagi, *J Electron Microsc (Tokyo)* p. dfp030 (2009).
- [4] M. Haider, S. Uhlemann, E. Schwan, H. Rose, B. Kabius, and K. Urban, *Nature* **392**, 768 (1998).
- [5] P. E. Batson, N. Dellby, and O. L. Krivanek, *Nature* **418**, 617 (2002).
- [6] T. Huang and R. Tsai, *Adv. Comput. Vis. Image Process.* **1**, 317 (1984).
- [7] W. D. Pyrz, D. A. Blom, T. Vogt, and D. J. Buttrey, *Angewandte Chemie International Edition* **47**, 2788 (2008).
- [8] A. Buades, B. Coll, and J. M. Morel, *Multiscale Modeling and Simulation* **4**, 490 (2005).
- [9] P. Binev, F. Blanco-Silva, D. A. Blom, W. Dahmen, R. Sharpley, and T. Vogt, preprint (2009).
- [10] M. Protter, M. Elad, H. Takeda, and P. Milanfar, *Image Processing, IEEE Transactions on* **18**, 36 (2009), ISSN 1057-7149.