Contrast mechanisms in scanning transmission x-ray microscopy

Pierre Thibault,^{*} Martin Dierolf,[†] Cameron M. Kewish, Andreas Menzel, and Oliver Bunk
Paul Scherrer Institut, 5232 Villigen PSI, Switzerland

Franz Pfeiffer

Department of Physics, Technical University of Munich, 85748 Garching, Germany

(Received 17 June 2009; published 14 October 2009)

We present a derivation of the main contrast mechanisms accessible in scanning transmission x-ray microscopy with a pixel-array detector. We consider the effect of the probe defocus and show that it can produce strong differential absorption contrast. The effect of noise is derived and used to compare the relative merits of absorption and differential contrast imaging in various experimental conditions. We illustrate the main results with an experiment that combines a through-focus series with the near-edge signal of a cerium oxide sample. The measurements are seen to follow closely the derived contrast expressions, including the defocus-dependent differential absorption contrast. The analysis includes additional considerations about the application of principal component analysis on a through-focus image series.

DOI: [10.1103/PhysRevA.80.043813](https://doi.org/10.1103/PhysRevA.80.043813)

PACS number(s): 41.50.+h, 68.37.Yz, 61.05.cp, 78.70.Dm

I. INTRODUCTION

Scanning transmission x-ray microscopy (STXM) is routinely used to produce maps of the local projected absorbance of a specimen. These maps are obtained by measuring the total transmitted intensity of a focused x-ray beam as the specimen is raster-scanned. This type of absorption mapping has been combined with various forms of spectroscopy, for instance, by measuring the same map repeatedly over a range of x-ray energies [1,2]. From the resulting “data cube,” one can extract spatial distributions of elemental concentrations or chemical states.

It has been known for a long time that a different detection configuration allows the measurement of additional information on a specimen. Information is best preserved if the complete spatial distribution of the far-field wave is measured. In optimal conditions, coherent diffractive imaging (CDI) methods, such as ptychography [3–5], can be used. However, the requirements for a complete inversion of the scattering problem (high degree of coherence, high signal-to-noise ratio, etc.) make such an approach not always possible in practice. Even in suboptimal cases, the measured diffracted intensities provide alternative contrast mechanisms, such as dark-field and differential phase contrast (DPC) [6–8]. The latter, obtained from a measurement of the deflection of the beam by the specimen, is especially relevant to hard x-ray imaging as the real part of the refractive index is typically much larger than its imaginary part in the keV regime. Segmented detectors, first introduced in scanning transmission electron microscopy (STEM) [9–11] and more recently for x rays [12,13], can give access to the DPC signal from an intensity difference between areas of the detector. In addition, the progress in fast-framing detectors and the in-

creasing data storage capacity of computers have promoted the recent development of DPC scanning instruments based on pixel-array detectors [14,15]. Quantitative phase measurements should allow extending absorption spectroscopy to “refraction spectroscopy,” where the real part of the index of refraction gives chemical information [16].

We present in this paper a formal derivation of the contrast mechanisms accessible with a STXM apparatus equipped with a pixel-array detector. In addition to the DPC signal, we show that absorption gradients contribute to the total differential contrast. A differential absorption contrast (DAC) contribution was described recently by de Jonge *et al.* [17] (as well as Hawkes [11] in the context of scanning transmission electron microscopy) and found to be negligible most of the time. Here we consider in addition the effect of defocus and find that it can be very strong. Among the implications of this additional contrast mechanism are the possibility of contamination of a DPC signal with DAC and potential gains in signal-to-noise ratio. We illustrate our results with a through-focus spectromicroscopy measurement over the L_{III} edge of a CeO_2 test sample.

II. CONTRAST MECHANISMS IN A SCANNING MICROSCOPE

In a STXM setup, a portion of the incident beam is selected and focused down onto the specimen using a Fresnel zone plate [see Fig. 1(a)]. Because all higher orders of the zone plate are removed, the illumination incident on the specimen, called the probe, is very well modeled with that produced by an ideal thin lens. After interacting with the specimen, the wave diverges and propagates freely to the detector plane, where its intensity distribution is measured.

Figure 1(b) shows an example of intensity distribution measured in the far field by a pixel-array detector. This diffraction pattern, drawn from the experimental data set described below, does not satisfy the oversampling and coherence requirements for phase retrieval. In this situation, a simple approach is to extract information on the specimen

^{*}Present address: Department of Physics, Technical University of Munich, 85748 Garching, Germany; pierre.thibault@ph.tum.de

[†]Present address: Department of Physics, Technical University of Munich, 85748 Garching, Germany.

FIG. 1. (Color online) (a) Schematic of a typical STXM setup. A Fresnel zone plate (A) produces a focal spot near or close to the sample plane (C). An aperture (B), together with a central beam stop in the plane of the zone plate, filters all unwanted diffraction orders. The intensity is measured in the far field (D). (b) An example of measured intensity distribution.

through a moment analysis of the measured intensity. In the following, we compute the zeroth and first moments of the intensity with minimal assumption on the illumination conditions [18]

$$M_0 = \int |\psi(\mathbf{q})|^2 d\mathbf{q}, \quad (1)$$

$$M_\alpha = \int q_\alpha |\psi(\mathbf{q})|^2 d\mathbf{q}, \quad \alpha = x, y, \quad (2)$$

$$M_{\alpha\beta} = \int q_\alpha q_\beta |\psi(\mathbf{q})|^2 d\mathbf{q}, \quad \alpha = x, y \quad \beta = x, y, \quad (3)$$

where $\psi(\mathbf{q})$ is the wave in the far field at the spatial frequency $\mathbf{q}$. These moments, essentially expectation values, can be expressed elegantly in the Dirac notation [19]

$$M_0 = \langle \psi | \psi \rangle, \quad (4)$$

$$M_\alpha = \langle \psi | q_\alpha | \psi \rangle, \quad \alpha = x, y, \quad (5)$$

$$M_{\alpha\beta} = \langle \psi | q_\alpha q_\beta | \psi \rangle, \quad \alpha = x, y \quad \beta = x, y. \quad (6)$$

In Eqs. (1)–(6) and in the following calculations, we limit the derivation to fully coherent illumination. In principle, partial coherence can be included rigorously in a similar formalism using density matrices (i.e., $M_0 = \text{Tr } \rho$, $M_\alpha = \text{Tr } q_\alpha \rho$, ...). The full coherence assumption is expected to remain mostly valid even for partial coherent cases, the main effect being an additional blurring of the STXM maps caused by a convolution with the image of the extended source.

Let $|P\rangle$ be the wave field incident on the specimen. For small-angle scattering, the exit wave is well approximated by $|\psi\rangle = O|P\rangle$, with

$$O(\mathbf{r}) = \exp \left[ik \int (n-1) dz \right], \quad (7)$$

where n is the sample's index of refraction, which for x rays is commonly written explicitly in terms of its real and imaginary parts: $n = 1 - \delta - i\beta$. We define the integrated phase and absorption as

$$\phi(\mathbf{r}) = k \int \delta dz \quad \text{and} \quad \mu(\mathbf{r}) = 2k \int \beta dz. \quad (8)$$

A. Absorption contrast

With the quantities defined above, the zeroth moment is written as

$$M_0 = \langle P | O^\dagger O | P \rangle = \langle P | e^{-\mu(\mathbf{r})} | P \rangle. \quad (9)$$

Explicitly writing the dependence of M_0 on each point $\mathbf{R}$ of the raster scan on which it is measured, one finds

$$M_0(\mathbf{R}) = \int |\langle \mathbf{r} - \mathbf{R} | P \rangle|^2 e^{-\mu(\mathbf{r})} d\mathbf{r} = |P(\mathbf{r})|^2 * e^{-\mu(\mathbf{r})}, \quad (10)$$

where $*$ is the convolution operation, recovering the known fact that the point-spread function (PSF) of a STXM absorption map is equal to the squared amplitude of the probe. Assuming the total intensity of the probe $I_0 = \langle P | P \rangle$ is known, a map of the absorption by the specimen, at a resolution limited by the extent of the PSF, is given by

$$\mu(\mathbf{R}) = -\ln[M_0(\mathbf{R})/I_0]. \quad (11)$$

B. Differential contrast

Computation of the first moment, given by

$$M_\alpha = \langle P | O^\dagger q_\alpha O | P \rangle, \quad (12)$$

will in general require a model for the probe. However, the special case of a uniformly absorbing object can be considered even if $|P\rangle$ is kept general. When $\mu = \mu_0$, one finds

$$O^\dagger q_\alpha O = O^\dagger (O q_\alpha - O \partial_\alpha \phi) = e^{-\mu_0} (q_\alpha - \partial_\alpha \phi). \quad (13)$$

If the probe is centrosymmetric (i.e., it is an even function in x and y), one finds

$$M_\alpha(\mathbf{R}) = -e^{-\mu_0} |P(\mathbf{r})|^2 * \partial_\alpha \phi(\mathbf{r}). \quad (14)$$

When absorption variations are negligible, the first moment of the intensity distribution yields a map of the gradient of the phase part of O , with the same PSF as the absorption. This imaging mechanism defines differential phase contrast. The equivalent form,

$$M_\alpha(\mathbf{R}) = -e^{-\mu_0} \partial_\alpha |P(\mathbf{r})|^2 * \phi(\mathbf{r}), \quad (15)$$

allows the interpretation of $M_\alpha(\mathbf{R})$ as a map of the phase convolved with the antisymmetric PSFs given by $\partial_\alpha |P(\mathbf{r})|^2$.

We now consider the general situation where the absorption is not uniform. We will assume that the probe is formed by the free-space propagation over a distance z of the wave field produced by an ideal converging lens of focal length f and of centrosymmetric pupil aperture $a(\mathbf{r}) = \langle \mathbf{r} | a \rangle$. Furthermore, we assume that O varies sufficiently slowly in the region illuminated by the probe. μ and ϕ are then well approximated by their first-order expansion about a given point in the raster scan (which we set to zero to keep the expressions simple)

$$O(\mathbf{r}) = \exp \left[-i\phi(\mathbf{r}) - \frac{1}{2}\mu(\mathbf{r}) \right] \approx \exp \left[-i\mathbf{g}_\phi \cdot \mathbf{r} - \frac{1}{2}\mu_0 - \frac{1}{2}\mathbf{g}_\mu \cdot \mathbf{r} \right]. \quad (16)$$

The zeroth order term in the expansion of ϕ is omitted since

it always vanishes in the expectation value calculation.

One finds, to first order in $\mathbf{g}_\mu$ and $\mathbf{g}_\phi$ (see Appendix A),

$$M_\alpha = I_0 e^{-\mu_0} g_\alpha, \quad (17)$$

with $I_0 = \langle a|a \rangle = \langle P|P \rangle$ the integrated incident intensity and

$$\mathbf{g} = -\mathbf{g}_\phi - \mathbf{g}_\mu \frac{(z-f) \langle P|\mathbf{q}^2|P \rangle}{2k I_0} - \mathbf{g}_\mu \frac{f \langle a|\mathbf{q}^2|a \rangle}{2k I_0}. \quad (18)$$

The measured differential contrast has three main contributions. The first term is the DPC signal, consistent with the phase object assumption above. The second and third terms depend on the gradient of the absorption part and therefore contribute to the DAC signal. Both brackets in this expression have a straightforward interpretation. The first one, $\langle P|\mathbf{q}^2|P \rangle$, is the second moment (the variance) of the intensity distribution measured in the far field in the absence of a specimen and is thus readily computed from measured data. The second bracket, $\langle a|\mathbf{q}^2|a \rangle$, is the second moment of the intensity distribution in the focal plane. Interestingly, this number is in principle infinite for an ideal circular pupil (it is the variance of an Airy pattern). Related to the spot size at the focal plane, it is reasonable in practice to assume that $\langle a|\mathbf{q}^2|a \rangle$ is the same order as the integral of the square of the central lobe of $|\langle \mathbf{q}|a \rangle|$. Considering for a moment the ideal case of a circular pupil of radius R , one finds that $\langle a|\mathbf{r}^2|a \rangle = I_0 R^2/2$, which suggests the definition of an effective numerical aperture

$$\mathcal{N}_A^2 = \frac{R^2}{f^2} = \frac{1}{f^2} \frac{2 \langle a|\mathbf{r}^2|a \rangle}{I_0}. \quad (19)$$

A geometric optics approximation gives

$$\langle P|\mathbf{q}^2|P \rangle \approx \frac{k^2}{f^2} \langle a|\mathbf{r}^2|a \rangle = \frac{I_0 k^2 \mathcal{N}_A^2}{2}. \quad (20)$$

Using $\langle a|\mathbf{q}^2|a \rangle \approx I_0 c_0^2 / R^2$, where $c_0 = 3.83$ is the first zero of the Bessel function of order 1, one finds

$$\langle a|\mathbf{q}^2|a \rangle \approx \frac{I_0 c_0^2}{f^2 \mathcal{N}_A^2} \quad (21)$$

so that Eq. (18) becomes

$$\mathbf{g} \approx -\mathbf{g}_\phi - \mathbf{g}_\mu \frac{k \Delta z \mathcal{N}_A^2}{4} - \mathbf{g}_\mu \frac{c_0^2/2}{k f \mathcal{N}_A^2}, \quad (22)$$

where $\Delta z = z - f$. Another convenient formulation involves the depth of focus of the lens, which will be here defined as [20]

$$z_{\text{DOF}} = \frac{1}{k \mathcal{N}_A^2}, \quad (23)$$

giving

$$\mathbf{g} \approx -\mathbf{g}_\phi - \mathbf{g}_\mu \frac{\Delta z}{2 z_{\text{DOF}}} - \mathbf{g}_\mu \frac{c_0^2}{4 f}. \quad (24)$$

In this form, it is clear that the last term, described recently by de Jonge *et al.* [21], can often be neglected. With a proper tuning of the defocus, the second term can be made large,

thus opening the way to a different absorption contrast mechanism.

C. Noise

To compute the uncertainty on μ_0 and $\mathbf{g}$, we assume that the measured intensities follow the noise model

$$I_{\text{measured}} = (1 + s)(I + n_m). \quad (25)$$

where I is the true intensity, n_m is the error introduced in the measurement, and s is a systematic fluctuation in the incoming beam intensity, with mean zero and variance $\text{var}(s) = \sigma^2$. The variance of n_m typically has two contributions: one from the Poisson statistics that arise in photon counting and one from the instrumental fluctuations, such as read-out noise and electronic background. The Poisson noise contribution depends on $\mathbf{q}$ through I , but we assume here that the other term is constant over the detector area

$$\text{var}[n_m(\mathbf{q})] = I(\mathbf{q}) + c. \quad (26)$$

The calculation of the error propagation is carried out on Eqs. (11) and (17), which can be written as

$$\mu_0 = -\ln(M_0/I_0), \quad (27)$$

$$g_\alpha = M_\alpha/M_0. \quad (28)$$

It is however necessary to cast Eqs. (1) and (2) in a discrete form, where the integrals are replaced with sums over the pixel coordinates of the detector area

$$M_0 = \sum_{\mathbf{q}} I_{\text{meas}}(\mathbf{q}),$$

$$M_\alpha = \sum_{\mathbf{q}} q_\alpha I_{\text{meas}}(\mathbf{q}).$$

The variance of μ_0 is computed using standard error propagation

$$\begin{aligned} (\Delta\mu_0)^2 &= \sum_{\mathbf{q}} \left(\frac{\partial \mu}{\partial n_m(\mathbf{q})} \right)^2 \text{var}[n_m(\mathbf{q})] + \left(\frac{\partial \mu}{\partial s} \right)^2 \text{var}(s) \\ &= \frac{1}{I_0^2} \sum_{\mathbf{q}} \text{var}[n_m(\mathbf{q})] + \text{var}(s) \\ &\approx \frac{M_0}{I_0^2} + \frac{Nc}{I_0^2} + \sigma^2, \end{aligned} \quad (29)$$

where N is the number of terms in the sum, in practice equal to the number pixels in the detector. The last step is an approximation since the sum over $I(\mathbf{q})$ was replaced by its estimator M_0 . The variance of $\mathbf{g}$ is calculated in a similar way

$$\begin{aligned} (\Delta g_\alpha)^2 &= \sum_{\mathbf{q}} \left(\frac{\partial g_\alpha}{\partial n_m(\mathbf{q})} \right)^2 \text{var}[n_m(\mathbf{q})] \\ &= \sum_{\mathbf{q}} \left(\frac{q_\alpha}{M_0} - \frac{M_\alpha}{M_0^2} \right)^2 \text{var}[n_m(\mathbf{q})] \\ &\approx \frac{M_{\alpha\alpha}}{M_0^2} + \frac{c}{M_0^2} \sum_{\mathbf{q}} q_\alpha^2. \end{aligned} \quad (30)$$

In the last step, second-order terms M_α^2 were dropped and the sums over the true intensities $I(\mathbf{q})$ were again replaced with their estimators. Evaluating the second term requires additional assumption on the detector configuration. Here, we assume that the detector is a square array $\sqrt{N} \times \sqrt{N}$ and that its angular extent matches the extent covered by the pupil, that is $|q_\alpha| \leq k\mathcal{N}_A^2$. In this situation, the sum is readily computed

$$\sum_{\mathbf{q}} q_\alpha^2 \approx \frac{1}{3} N k^2 \mathcal{N}_A^2. \quad (31)$$

Finally, using $M_{\alpha\alpha} \approx \langle P | q_\alpha^2 | P \rangle$, Eq. (20), and letting κ be the ratio of instrumental variance to the mean counting statistics variance,

$$\kappa = \frac{c}{M_0/N} = \frac{\langle \text{var } n_m \rangle}{\langle I \rangle} - 1, \quad (32)$$

gives

$$(\Delta\mu_0)^2 = \frac{1}{M_0} (1 + \kappa) + \sigma^2, \quad (33)$$

$$(\Delta g)^2 = \frac{\langle P | \mathbf{q}^2 | P \rangle}{M_0} \left(1 + \frac{2\kappa}{3\sqrt{N}} \right). \quad (34)$$

It is worth noting that the uncertainty on the differential signal is independent of σ , indicating that random fluctuations in the incident intensity do not degrade the signal of $\mathbf{g}$.

In the case of an ideal photon-counting detector ($c=0$) and if the absorption is low ($M_0 \approx I_0$), the uncertainties are reduced to

$$\Delta\mu_0 \approx \sqrt{\frac{1}{I_0} + \sigma^2}, \quad (35)$$

$$\Delta g \approx k\mathcal{N}_A^2 \sqrt{\frac{2}{I_0}}, \quad (36)$$

where I_0 is the integrated intensity of the incoming beam expressed in number of photons and $\Delta g = \Delta g_x = \Delta g_y$. The relative reliability of absorption maps compared to differential contrast depends heavily on the amplitude of σ [22]. It may be convenient to set

$$\rho = \sqrt{1 + \sigma^2 I_0} \quad (37)$$

so that $\Delta\mu_0 = \rho \sqrt{1/I_0}$.

D. Discussion

We now consider in more detail the special situation where the specimen is made of a single chemical species. The integrated phase and absorbance ϕ and μ are then given by

$$\phi = k\delta t, \quad (38)$$

$$\mu = 2k\beta t, \quad (39)$$

where t is the thickness of the specimen integrated in the propagation direction. Neglecting the weak focus-

independent DAC term, the differential contrast is simplified to

$$\mathbf{g} = -k \nabla t \left(\delta + \beta \frac{\Delta z}{z_{\text{DOF}}} \right), \quad (40)$$

where again z_{DOF} is to be understood as an effective value obtained through Eq. (19). Contrast inversion occurs upstream from the focal plane (for x rays) at a distance

$$\Delta z = -z_{\text{DOF}} \frac{\delta}{\beta}. \quad (41)$$

The simplified Eq. (40) can be used to evaluate the merit of differential contrast relative to absorption contrast. However, this cannot be done in complete generality since absorption depends on the specimen's thickness while differential contrast depends on the thickness variations. Since the modulation transfer function (MTF) associated with the DPC mode is 0 at the origin, a slab of constant thickness has 0 signal-to-noise ratio (S/N). On the other hand, a thin object having spatial fluctuation on the same scale as the probe can have a strong DPC signal while absorbing very little. Let r_{abs} , r_{DPC} , and r_{DAC} be the signal-to-noise ratios of the absorption signal, differential phase contrast, and differential absorption contrast, respectively. Using Eqs. (35), (36), (38), and (39) leads to the following ratios:

$$\frac{r_{\text{DAC}}}{r_{\text{abs}}} = \frac{\rho \Delta z |\nabla t| \mathcal{N}_A}{\sqrt{2} t}. \quad (43)$$

The finite extent of the PSF puts a higher bound on the gradient of the integrated thickness, namely, $|\nabla t| < t/d$, where d is the width of the spot size. For a disk pupil, the nominal spot size is $d = c_0/(k\mathcal{N}_A)$ and thus

$$\frac{r_{\text{DPC}}}{r_{\text{abs}}} \lesssim \frac{\delta}{\beta} \frac{\rho}{\sqrt{2} c_0}, \quad (44)$$

$$\frac{r_{\text{DAC}}}{r_{\text{abs}}} \lesssim \frac{\sqrt{2} \rho}{c_0} \frac{\Delta z}{z_{\text{DOF}}}. \quad (45)$$

Differential phase contrast offers better signal-to-noise ratio than absorption almost as soon as $\delta > \beta$. DPC also offers enhanced signal-to-noise ratio in cases where the incoming beam intensity has large fluctuations, which increases the value of ρ . Large intensity fluctuations also make differential absorption contrast valuable, though generally at the price of a decrease in resolution caused by the defocus of the probe.

In the next section, we will present x-ray measurements that illustrate Eq. (40) and its behavior as the incident x-ray energy is varied across one of the specimen's absorption edges. Whereas δ and β vary quickly as function of energy, all other wavelength-dependent quantities can be expanded linearly in the energy difference. If the range of energies scanned is written $(1 + \epsilon)E$, with $|\epsilon| \ll 1$, one finds

FIG. 2. STXM maps. (a) Transmission $E=5724$ eV, $\Delta z=150$ μm . (b) Transmission $E=5740$ eV, $\Delta z=0$. (c) Differential contrast $E=5724$ eV, $\Delta z=150$ μm . (d) Differential contrast $E=5740$ eV, $\Delta z=0$. Both the specimen's displacement along the beam and the change in energy caused significant drifts. (e) First eigenvector of the PCA analysis on the transmission data. All scale bars are 2 μm .

$$\mathbf{g}(\epsilon) \approx -k \nabla t \left[\delta(\epsilon)(1 + \epsilon) + \beta(\epsilon) \frac{1}{z_{\text{DOF}}} (\Delta z - \epsilon f) \right]. \quad (46)$$

In this expression, k , z_{DOF} , Δz , and f are the values at a reference energy $\epsilon=0$. The term ϵf comes from the assumption that the lens is a diffractive optic (a zone plate), whose focal distance is proportional to the energy. Fitting this equation against differential contrast experimental data makes possible the recovery of both δ and β , up to a multiplicative constant since $k \nabla t$ is generally unknown. Even when the focal plane position is not well known, Eq. (46) always allows a quantitative evaluation of the ratio δ/β .

III. EXPERIMENTAL DEMONSTRATION

We illustrate the theory developed above with experimental data measured at the cSAXS beamline at the Swiss Light Source (Switzerland). The x-ray energy was tuned around the CeO_2 L_{III} edge (5723 eV) using a double-crystal Si(111) monochromator. A Fresnel zone plate of radius 50 μm with 100 nm outermost zonewidth was used to produce the focus. The specimen was a thin layer of CeO_2 nanoparticles formed by letting a water solution dry on a 1- μm -thin silicon nitride window. The thickness variations required for differential contrast occurred naturally in the form of desiccation cracks similar to those formed in the clay of arid regions (see Fig. 2). The intensity distribution past the specimen was measured with the PILATUS 2M [23], a single-photon counting pixel-array detector having no read-out noise and no point-spread function beyond a single pixel.

The measured STXM data, absorption μ , and differential contrast $\mathbf{g}$ are elements of four-dimensional quantities. Each STXM map is a two-dimensional image that covers a 10×10 μm^2 area in 51×51 points, each exposed for 20 ms. The two additional dimensions are spanned by defocus and energy. We measured images for 13 different defocus values covering a range of 600 μm and for 20 x-ray energies between 5698 and 5759 eV. Figure 2 shows a small selection of the data set.

Both the change in energy and in defocus caused significant drift of the beam position relative to the sample. Ac-

cordingly, realigning all images is the first step of the data analysis. This is accomplished through cross correlation and refined iteratively. Once stacked, the realigned images form “data cubes” from which spectroscopic information can be extracted. Analysis can be accomplished in a variety of ways. We use principal component analysis (PCA), a standard numerical analysis procedure used to find a set of orthogonal directions along which the variance of the data is highest. Essentially a dimensionality reduction procedure, PCA defines a model for the data cubes $\mu(\mathbf{r}, E, \Delta z)$ and $\mathbf{g}(\mathbf{r}, E, \Delta z)$ as

$$\mu(\mathbf{r}, E, \Delta z) \approx \sum_i^N \rho_i^\mu(\mathbf{r}) s_i^\mu(E, \Delta z), \quad (47)$$

$$\mathbf{g}(\mathbf{r}, E, \Delta z) \approx \sum_i^N \rho_i^g(\mathbf{r}) s_i^g(E, \Delta z), \quad (48)$$

where N is the reduced number of dimensions, adjusted to model the most significant variations in the data. The model that optimally represents the data is such that the arrays $\rho_i(\mathbf{r})$ are the eigenvectors of the effective covariance matrix of μ and $\mathbf{g}$.

As a result of the drift, the region common to all individual maps shrank to about 3% of the area covered by a single map. Instead of eliminating all but a minute portion of the measured data, we adopted a log-likelihood formulation of PCA that can accommodate missing data regions [24]. Regions where data are missing for multiple defocus and energy values can still cause unwanted bias. For the analysis, we have kept only the area covered by at least one measurement for each energy values. In this area, which covers about 50% of a single scan, the completeness of the data varies from 25% to 100%.

To keep the experimental demonstration simple, the sample was made of one chemical species. Variations in the data set are thus concentrated along a single axis. The first eigenvector $\rho_1(\mathbf{r})$ resulting from the PCA calculation on the transmission data is shown in Fig. 2(e). The main deviation from the first principal axis, given by ρ_2 , is caused by the variation of the point-spread function as a function of defocus. As described in Appendix B, this effect can be used to provide a correction term for the linear relationship of the signal with Δz in Eq. (46).

The spectra, given by s_i in the above expressions, are shown in Fig. 3. The defocus-dependent correction term (see Appendix B) was needed only in the differential contrast case and was obtained by a least-square minimization of $s_1 + \alpha s_2$ on a model given by Eq. (46), giving $\alpha \approx 0.56$. The parameters obtained from the least-square fit to the differential contrast data shown in Fig. 3(b) give, up to scaling and offset, the real and imaginary parts of the sample's index of refraction as shown in Fig. 4. Keeping in mind the few adjustable parameters, the retrieved refractive index is in good agreement with published experimental data [25].

As emphasized above, the main purpose of this measurement was to illustrate the behavior of the transmission and differential terms as a function of defocus and energy. In

FIG. 3. (a) The spectrum extracted from the absorption data μ . As expected, it is independent of defocus. (b) The spectrum extracted from the differential contrast signal g . The dashed and continuous lines show the fit of Eq. (46) on the data, shown as circles. The y-axis scale is arbitrary in both (a) and (b).

practical applications, decoupling the DPC from the DAC terms could require as few as two defocus values. The main source of errors in our results appears to be related to the incompleteness of the data set resulting from the large drifts between the individual maps. In future spectromicroscopy applications, better results are expected by reducing drift and scanning a larger area.

IV. CONCLUSION

In this paper, we have derived the main contrast mechanisms present in a STXM apparatus equipped with a pixel-array detector. We have introduced explicitly the effect of the probe defocus and shown that it can produce a strong differential absorption contrast signal. Expressions for the expected signal-to-noise ratio of various contrast mechanisms were derived. These can be used to decide the most efficient way of extracting spectroscopic information from a sample. Finally, we have presented the result of a test experiment illustrating a situation where the differential absorption contrast can be strong. The contrast variations as a function of energy and defocus were found to correspond well to the theoretical predictions.

In situations where the coherent flux is sufficient, reconstruction of spatial distribution of the refractive index should be attempted using ptychography, which offers higher resolution than the probe size and is not sensitive to defocus inaccuracies. When coherent imaging is not possible, this paper provides a framework for future attempts at combining differential contrast with spectroscopic measurements.

FIG. 4. Real and imaginary parts of the index of refraction as a function of energy. The circles indicate the fit on the measured data. The continuous curve for β comes from a measurement by Nachimuthu *et al.* [25]. The curve and the points have been shifted in amplitude and rescaled to coincide approximately with the published data by Henke *et al.* [26] away from the near-edge region. Data from Henke and Nachimuthu were spliced together and Hilbert transformed to generate the curve for δ , following a method similar to that presented by Jacobsen *et al.* [16]. In addition to the offset and the scaling, the data points for δ were adjusted with a third free parameter which corresponds to a correction in the position of the focal plane in Eq. (46).

Future quantitative studies of differential contrast images will often require that absorption contributions be taken into account. The effect of differential absorption contrast will become increasingly important as x-ray optics improve, leading to smaller spot sizes and shorter depths of focus.

ACKNOWLEDGMENTS

We are grateful to Martin de Jonge and Graeme Morrison for fruitful discussions on the differential absorption contrast. We thank Jérôme Rose for suggesting this research avenue and for providing the cerium oxide samples. Christian David and Joan Vila Comamala provided the focusing optics. Julian Viereck carried a part of the data analysis. P.T. acknowledges financial support from the Fonds Québécois de la recherche sur la nature et les technologies (FQRNT).

APPENDIX A: DERIVATION OF THE DIFFERENTIAL CONTRAST TERM

The wave field in the specimen plane is written as

$$|P\rangle = T_z C_f |a\rangle, \quad (A1)$$

where T_z is the free-space propagation operator,

$$T_z = \exp\left(\frac{-iz}{2k} \mathbf{q}^2\right), \quad (A2)$$

C_f is the curved wave front produced by an ideal lens,

$$C_f = \exp\left(\frac{-ik}{2f}\mathbf{r}^2\right), \quad (\text{A3})$$

and $|a\rangle$ is the wave field at the exit of the aperture normalized such that $\langle a|a\rangle$ is equal to I_0 , the number of incident photons on the specimen. Because the incident wave field is assumed to be planar, $\langle \mathbf{r}|a\rangle$ is a real-valued function. For a circular pupil of radius R , one has for instance

$$\langle \mathbf{r}|a\rangle = \sqrt{\frac{I_0}{\pi R^2}} \begin{cases} 1 & \text{if } |\mathbf{r}| < R \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A4})$$

The first moment is now given by

$$M_\alpha = \langle a|C_f^\dagger T_z^\dagger O^\dagger q_\alpha O T_z C_f|a\rangle. \quad (\text{A5})$$

Calculations are made straightforward using the commutation relations readily obtained from Eqs. (A2) and (A3)

$$[r_\alpha, O] = 0 \quad [q_\alpha, O] = -i\partial_\alpha O,$$

$$[r_\alpha, T_z] = q_\alpha \frac{z}{k} T_z \quad [q_\alpha, T_z] = 0,$$

$$[r_\alpha, C_f] = 0 \quad [q_\alpha, C_f] = -r_\alpha \frac{k}{f}.$$

The first simplification step is to evaluate Eq. (A5) using the first-order expansion (16). To first order in $\mathbf{g}_\mu$ and $\mathbf{g}_\phi$, we find

$$M_\alpha \approx e^{-\mu_0} \langle a| \left[-g_{\phi,\alpha} + \frac{i}{2}g_{\mu,\alpha} + q_\alpha + \frac{z}{f}\mathbf{g}_\mu \cdot (\mathbf{r}q_\alpha + \mathbf{q}r_\alpha) - \mathbf{g}_\mu \cdot \mathbf{r}q_\alpha - \frac{z}{k}\mathbf{g}_\mu \cdot \mathbf{q}q_\alpha - \frac{k}{f}r_\alpha - \frac{k(z-f)}{f^2}\mathbf{g}_\mu \cdot \mathbf{r}r_\alpha \right] |a\rangle, \quad (\text{A6})$$

where it should be reminded that $\mathbf{g}_\mu$ and $\mathbf{g}_\phi$ are constant expansion coefficients of $O(\mathbf{r})$ around a given scanning point. Assuming that $\langle \mathbf{r}|a\rangle$ is centrosymmetric reduces this expression to

$$M_\alpha \approx e^{-\mu_0} \langle a| \left[-g_{\phi,\alpha} + \frac{i}{2}g_{\mu,\alpha} + \frac{z}{f}g_{\mu,\alpha}(r_\alpha q_\alpha + q_\alpha r_\alpha) - g_{\mu,\alpha}r_\alpha q_\alpha - \frac{z}{k}g_{\mu,\alpha}q_\alpha^2 - \frac{k(z-f)}{f^2}g_{\mu,\alpha}r_\alpha^2 \right] |a\rangle. \quad (\text{A7})$$

Finally, if $\langle \mathbf{r}|a\rangle$ is real-valued, it can be shown that $\langle a|r_\alpha q_\alpha|a\rangle = \frac{i}{2}\langle a|a\rangle$, leading to the cancellation of all but three terms

$$M_\alpha \approx e^{-\mu_0} \langle a| \left[-g_{\phi,\alpha} - \frac{z}{k}g_{\mu,\alpha}q_\alpha^2 - \frac{k(z-f)}{f^2}g_{\mu,\alpha}r_\alpha^2 \right] |a\rangle. \quad (\text{A8})$$

Equation (18) is recovered using

$$\begin{aligned} \langle P|q_\alpha^2|P\rangle &= \langle a|C_f^\dagger T_z^\dagger q_\alpha^2 T_z C_f|a\rangle \\ &= \langle a|q_\alpha^2|a\rangle + \frac{k^2}{f^2}\langle a|r_\alpha^2|a\rangle - \frac{k}{f}\langle a|q_\alpha r_\alpha + r_\alpha q_\alpha|a\rangle \\ &= \langle a|q_\alpha^2|a\rangle + \frac{k^2}{f^2}\langle a|r_\alpha^2|a\rangle, \end{aligned} \quad (\text{A9})$$

where the real-valuedness of $\langle \mathbf{r}|a\rangle$ was used again for the cancellation of the last term.

APPENDIX B: PCA OVER A THROUGH-FOCUS SERIES

Contrast variations as a function of energy are not relevant to the following discussion. We consider a data set made of a sequence of images resulting of a range of defocus values. Since PCA is a linear procedure, the problem can be re-expressed in terms of the Fourier transform of the images, where the effect of defocus is simply a change in the width of the MTF. Inspired by the properties of a Gaussian MTF, we consider the case of a general MTF $T(q^2/\sigma^2)$, where the effect of defocus is accounted for by a rescaling of σ . Expanding T in $x = 1 - \sigma/\sigma_0$, one finds

$$T(q^2/\sigma^2) \approx T(q^2/\sigma_0^2) + 2x \frac{q^2}{\sigma_0^2} T'(q^2/\sigma_0^2) + O(x^2), \quad (\text{B1})$$

where T' is the first derivative of T with respect to its argument. For small-enough x , applying PCA on a data set following (C1) is expected to return mainly two orthogonal directions, related to the two linearly independent terms in the equation above. If $x \ll 1$ for all images in the data set, the first component will be close to the first term

$$\rho_1(q) = A_1 T(q^2/\sigma_0^2), \quad (\text{B2})$$

where A_1 is a normalization constant. The second component is found using Gram-Schmidt

$$\rho_2(q) = A_2 \left[T(q^2/\sigma_0^2) + 2 \frac{q^2}{\sigma_0^2} T'(q^2/\sigma_0^2) \right]. \quad (\text{B3})$$

A PCA calculation therefore will expand the MTF of any image in the data set as a linear combination

$$T(q^2/\sigma^2) \approx s_1 \rho_1(q) + s_2 \rho_2(q), \quad (\text{B4})$$

with

$$s_1 = \frac{1-x}{A_1} \quad \text{and} \quad s_2 = \frac{x}{A_2}. \quad (\text{B5})$$

As x increases (i.e., as the width of the MTF decreases), a part of the amplitude of the first component is transferred into the second one. The amplitude of T is thus recovered from s_1 and s_2 ,

$$T(0) \propto s_1 + \alpha s_2, \quad (\text{B6})$$

where the ratio $\alpha = A_2/A_1$ is a number depending on T . In the case of a Gaussian MTF, ρ_1 and ρ_2 are recognized at once as the zeroth and second Hermite functions and calculation can be carried out analytically, giving

$$\alpha = \frac{1}{\sqrt{2}}. \quad (\text{B7})$$

In the case of differential phase contrast, Eq. (14) implies that the MTF is antisymmetric and the calculations can be redone, replacing $T(q^2/\sigma^2)$ with $\mathbf{q}T(q^2/\sigma^2)$. Again modeling T with a Gaussian function, we find that

$$\alpha = \sqrt{\frac{2}{3}}. \quad (\text{B8})$$

In general, α will be a number of order unity, provided that the defocus range is sufficiently narrow.

-
- [1] H. Ade, J. Kirz, S. Hulbert, E. Johnson, E. Anderson, and D. Kern, *Appl. Phys. Lett.* **56**, 1841 (1990).
- [2] H. Ade, X. Zhang, S. Cameron, C. Costello, J. Kirz, and S. Williams, *Science* **258**, 972 (1992).
- [3] R. Hegerl and W. Hoppe, *Ber. Bunsen. Phys. Chem* **74**, 1148 (1970).
- [4] J. M. Rodenburg and R. H. T. Bates, *Philos. Trans. R. Soc. London, Ser. A* **339**, 521 (1992).
- [5] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, *Science* **321**, 379 (2008).
- [6] A. Huizer and P. Van Toorn, *J. Phys. D* **15**, 747 (1982).
- [7] G. Morrison and J. Chapman, *Optik (Stuttgart)* **64**, 1 (1983).
- [8] H. Chapman, C. Jacobsen, and S. Williams, *Rev. Sci. Instrum.* **66**, 1332 (1995).
- [9] H. Rose, *Optik (Stuttgart)* **39**, 416 (1974).
- [10] N. Dekkers and H. de Lang, *Optik (Stuttgart)* **41**, 452 (1974).
- [11] P. Hawkes, *J. Opt.* **9**, 235 (1978).
- [12] M. Feser, B. Hornberger, C. Jacobsen, G. De Geronimo, P. Rehak, P. Holl, and L. Strüder, *Nucl. Instrum. Methods Phys. Res. A* **565**, 841 (2006).
- [13] B. Hornberger, M. Feser, and C. Jacobsen, *Ultramicroscopy* **107**, 644 (2007).
- [14] A. Menzel *et al.* (unpublished).
- [15] A. Gianoncelli, G. Morrison, B. Kaulich, D. Bacescu, and J. Kovac, *Appl. Phys. Lett.* **89**, 251117 (2006).
- [16] C. Jacobsen, S. Wang, W. Yun, and S. Frigo, *Proc. SPIE*, **5538**, 23 (2004).
- [17] M. D. de Jonge, B. Hornberger, C. Holzner, D. Legnini, D. Paterson, I. McNulty, C. Jacobsen, and S. Vogt, *Phys. Rev. Lett.* **100**, 163902 (2008).
- [18] The second moment will also be required below to evaluate the uncertainty on the first moment.
- [19] The analogy with quantum mechanics is complete as the paraxial approximation transforms the Helmholtz equation for a monochromatic wave into a form identical to the two-dimensional (2D) time-dependent Schrödinger equation.
- [20] Mostly for aesthetic reasons, we have adopted a definition for the depth of focus that differs by a factor $2/\pi$ with the somewhat more common definition [27] of $z_{\text{DOF}} = \lambda/\mathcal{N}_A^2$.
- [21] M. de Jonge, S. Vogt, D. Legnini, I. McNulty, C. Rau, D. Paterson, B. Hornberger, C. Holzner, and C. Jacobsen, *Nucl. Instrum. Methods Phys. Res. A* **582**, 218 (2007).
- [22] Assuming that the noise model (25) is valid and given a sequence of intensity measurements $\{I\}$ over an empty area, σ can be evaluated using
- $$\sigma^2 = \frac{\langle I^2 \rangle}{\langle I \rangle} - 1,$$
- or
- $$\rho \approx \sqrt{1 + \frac{\text{var } I}{\langle I \rangle}} \quad \text{for } I \gg 1.$$
- [23] C. Brönnimann *et al.*, *J. Synchrotron Radiat.* **13**, 120 (2006).
- [24] S. Roweis, *NIPS '97: Proceedings of the 1997 Conference on Advances in Neural Information Processing Systems 10* (MIT Press, Cambridge, MA, 1998), pp. 626-632.
- [25] P. Nachimuthu, W. Shih, R. Liu, L. Jang, and J. Chen, *J. Solid State Chem.* **149**, 408 (2000).
- [26] B. L. Henke, E. M. Gullikson, and J. C. Davis, *At. Data Nucl. Data Tables* **54**, 181 (1993).
- [27] J. Kirz and D. T. Attwood, *Zone Plates* (Lawrence Berkeley National Laboratory, University of California, Berkeley, 2001), pp. 4-28-4-32.