Unwarping GISAXS data

Jiliang Liu and Kevin G. Yager*

Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973, USA. *Correspondence e-mail: kyager@bnl.gov

Grazing-incidence small-angle X-ray scattering (GISAXS) is a powerful technique for measuring the nanostructure of coatings and thin films. However, GISAXS data are plagued by distortions that complicate data analysis. The detector image is a warped representation of reciprocal space because of refraction, and overlapping scattering patterns appear because of reflection. A method is presented to unwarp GISAXS data, recovering an estimate of the true undistorted scattering pattern. The method consists of first generating a guess for the structure of the reciprocal-space scattering by solving for a mutually consistent prediction from the transmission and reflection sub-components. This initial guess is then iteratively refined by fitting experimental GISAXS images at multiple incident angles, using the distorted-wave Born approximation (DWBA) to convert between reciprocal space and detector space. This method converges to a high-quality reconstruction for the undistorted scattering, as validated by comparing with grazing-transmission scattering data. This new method for unwarping GISAXS images will broaden the applicability of grazing-incidence techniques, allowing experimenters to inspect undistorted visualizations of their data and allowing a broader range of analysis methods to be applied to GI data.

1. Introduction

Grazing-incidence small-angle X-ray scattering (GISAXS) is a powerful and broadly applicable method for quantifying structural order in interfaces, coatings, thin films and multi-layered nanostructures (Smilgies *et al.*, 2002; Müller-Buschbaum, 2003; Cristofolini, 2014). The scattering intensity in GISAXS is greatly enhanced because of beam projection along the sample surface (Smilgies, 2009), as well as waveguide effects that can increase the effective path-length through the sample and localize the photon field within the film (Jiang *et al.*, 2011). The geometry of GISAXS probes both in-plane (q_x) and out-of-plane (q_z) components, allowing the structure of ordered thin films to be quantified, including the material organization in the film-normal direction (Lu *et al.*, 2013). As a result, GISAXS has become an extremely popular technique for measuring three-dimensional order in thin films (Renaud *et al.*, 2009; Hexemer & Müller-Buschbaum, 2015; Müller-Buschbaum, 2016). However, experimental GISAXS detector images exhibit a number of complications because of the grazing-incidence geometry; correspondingly, data analysis is one of the main challenges in GISAXS experiments. The detector image is a (non-linearly) distorted version of reciprocal space caused by refraction of the incident X-ray beam as it enters a thin film, and refraction of scattered rays as they exit the film. Reflections of the X-ray beam off of the film and substrate interfaces will interfere with each other; the corresponding X-ray reflectivity curve modulates the

Computational unwarping

https://doi.org/10.1107/S2052252518012058 **737**
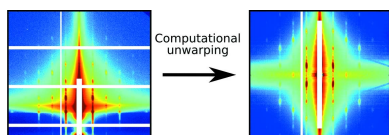
observed scattering intensity. Finally, a GISAXS image has multiple superimposed scattering patterns with different apparent origins, since scattering can occur from both the incident and reflected beams. Together, these effects conspire to yield a non-trivially warped rendition of the sample's 'true' reciprocal-space scattering, giving rise to the shifts and doubling effects that are characteristic of GISAXS detector images.

This data distortion is typically handled by fitting GISAXS data using a model incorporating the above-noted effects. The most popular model is the distorted-wave Born approximation (DWBA), which accounts for multiple scattering effects, such as scattering from the reflected beam (Sinha *et al.*, 1988; Lazzari, 2002; Lazzari *et al.*, 2007; Renaud *et al.*, 2009). Quasi-kinematic approximations (Heitsch *et al.*, 2010; Smilgies *et al.*, 2012) and integral potential methods (Wu, 1993, 1994) have also been described. When iteratively fitting experimental data with these methods, the GISAXS complications must be computed repeatedly, since on each iteration a candidate real-space model must be converted into reciprocal space, and then this must be warped into the experimental detector space.

Some experimental approaches have been described which circumvent the distortions of GISAXS. A thin-film sample can be measured in transmission (SAXS) geometry, with sample rotation used to progressively reconstruct the (q_x, q_z) reciprocal space (Jones *et al.*, 2003). This rotational SAXS method, often called CD-SAXS since it is used to quantify a critical dimension (Hu *et al.*, 2004), has been used with great success to reconstruct lithographic (Wang *et al.*, 2007; Settens *et al.*, 2014; Sunday *et al.*, 2015), nano-imprinted (Jones, Hu *et al.*, 2006; Jones, Soles *et al.*, 2006) and self-assembled (Sunday *et al.*, 2013, 2014; Khaira *et al.*, 2017) patterns. However, this method can be time-consuming since the desired reciprocal space is probed slice by slice. An alternative is the recently described grazing-incidence transmission small-angle X-ray scattering (GTSAXS), where the X-ray beam is directed towards the downstream edge of a sample at an incident angle well above the critical angle (Lu *et al.*, 2013; Mahadevapuram *et al.*, 2013; Weidman *et al.*, 2015). This geometry allows the sub-horizon scattering to exit through the substrate with minimal absorption; in the resultant data, the refraction distortions and multiple scattering effects are nearly eliminated. While this allows a minimally distorted (q_x, q_z) slice to be measured in a single exposure, the signal-to-noise ratio (SNR) (compared with GISAXS) is diminished because of the reduced beam projection along the sample and substrate attenuation. Moreover, this geometry requires the sample of interest to be present at the edge of the substrate, which may not always be convenient. Despite the availability of these techniques, in many cases GISAXS remains the preferred experimental technique because of the strong scattering signal that can be measured, the rapid data collection and the simplicity of the required sample and experimental protocols.

Here, we investigate whether GISAXS data can be computationally unwarped into an undistorted (SAXS-like) scattering pattern. This is a difficult inverse problem, since GISAXS data have multiple superimposed scattering contributions, all of which are convolved with distortions that change q positions and intensities. Moreover, this is a ill-posed mathematical problem, because of the one-to-many mapping from reciprocal space to the detector image, and the finite range of q space that is probed experimentally. Given the challenges in performing GISAXS data unwarping, researchers to date have instead focused on correctly approximating the effects of a GISAXS experiment, such that data can be iteratively fit using a forward model. However, there are potential advantages to unwarping GISAXS data into an estimate for the 'true' (undistorted) reciprocal-space scattering. Firstly, such a data representation is easier for human scientists to interpret. Common errors that are made when interpreting GISAXS data (such as mis-identifying a reflection-mode peak as a transmission-mode peak) could be avoided. Scientists can more easily identify canonical scattering patterns when unencumbered by GISAXS complications. Secondly, the speed of data fitting could be improved. When iteratively fitting GISAXS with a forward model, the grazing-incidence distortions must be computed repeatedly. If one could unwarp GISAXS data, then this undistorted scattering could be used for iterative fitting with a much lower computational cost. Thirdly, a broad range of existing scattering models (Pedersen, 1997; Förster *et al.*, 2005; Székely *et al.*, 2010; Li *et al.*, 2011; Yager, Zhang *et al.*, 2014; Senesi & Lee, 2015; Croset, 2017) and fitting software (Kline, 2006; Förster *et al.*, 2010), which are currently only intended to handle transmission SAXS data, could be used in a GISAXS context if data unwarping were available. Fourthly, a range of sophisticated data-analysis methods could be more easily applied to GISAXS data if they were remapped to an undistorted reciprocal space. For example, modern developments in correlation methods such as angular correlation analysis (Wochner *et al.*, 2009; Altarelli *et al.*, 2010; Lehmkühler *et al.*, 2014, 2018; Lhermitte *et al.*, 2017), fluctuation scattering (Chen *et al.*, 2012; Malmerberg *et al.*, 2015; Martin, 2017), or variance scattering (Yager & Majewski, 2014; Gommes, 2016) are not currently used in a GISAXS context. Finally, we note that healing data can be a useful pre-processing step (Liu *et al.*, 2017), allowing such data to be used with existing data-analysis pipelines, or input into modern machine-learning methods (Kiapour *et al.*, 2014; Wang *et al.*, 2016, 2017; Meister *et al.*, 2017).

We demonstrate that GISAXS data can be successfully unwarped to yield an estimate of the undistorted scattering. This method iteratively fits multiple input GISAXS images, using estimates of the sample's transmission and reflectivity curves to account for refraction and multiple scattering effects. This reconstruction is found to be qualitatively correct, in general, even when the underlying assumptions are imperfect. In cases where high-quality estimates for reflectivity are available, the reconstruction is both quantitative and has a higher SNR than the corresponding GTSAXS data. Overall, the proposed method is a robust means of viewing the true reciprocal-space scattering for materials measured in GISAXS geometry.

2. Methods

Our reconstruction method consists of two phases. In the first phase, a highly approximate guess for the undistorted pattern is generated. As will be shown below, even an imperfect guess (even as extreme as random noise) will ultimately yield a valid reconstruction during the second phase. However, an appropriate initial guess greatly reduces the computation time in the second phase. We deploy an efficient guessing method that estimates the ratio of transmitted and reflected signals based on the mismatch between the predictions of these two signals (described in detail in Section 3). In the second phase, the initial guess is iteratively refined in an optimization loop. On each loop, a candidate undistorted scattering is converted into GISAXS data using the DWBA theory (and assumed forms for the reflection and transmission curves). By minimizing the mismatch between the computed and experimental GISAXS data, the candidate scattering pattern converges towards the 'true' undistorted scattering. In our implementation, we simultaneously fit multiple experimental GISAXS images (taken on the sample but at different incident angles) in order to improve coverage of reciprocal space, and compensate for untrustworthy data in any particular GISAXS image (Fig. 1). Moreover, this reconstruction can be combined with GTSAXS data (if available) to further improve data quality. While GTSAXS intrinsically records a nearly undistorted image, the proposed reconstruction takes advantage of the much stronger scattering intensity of GISAXS to improve the SNR in the data.

2.1. GISAXS geometry

In conventional transmission X-ray scattering (TSAXS/TWAXS) the detector image can be directly related to reciprocal space. More precisely, for scattering with X-rays of wavelength λ, elastic scattering arises from the surface of the Ewald sphere in reciprocal space (a sphere with radius $k = 2\pi/\lambda$). Measurements at scattering angle 2θ probe the sample's reciprocal space at the coordinate $Q = (4\pi/\lambda)\sin\theta$; one can qualitatively say that the measured scattering arises from the slice where the Ewald sphere intersects the sample's reciprocal space. In grazing-incidence experiments, refraction of incident and scattered rays shifts and warps reciprocal space, leading to a distorted version of the data on the detector (Toney & Brennan, 1989; Busch *et al.*, 2006; Lazzari *et al.*, 2007; Breiby *et al.*, 2008).

Multiple scattering causes multiple 'copies' of the scattering pattern to be superimposed on the detector, with intensities modulated non-monotonically (Lee, Park, Yoon *et al.*, 2005). In this work, we use Q to denote coordinates in the sample's intrinsic reciprocal space (undistorted scattering) and use q to denote coordinates computed from the position on the detector image. Correspondingly, we differentiate between intensity of the sample's reciprocal-space scattering as $I_R(Q)$, and experimentally measured intensity on the detector as $I_d(q)$.

Refraction of both the incident and scattered X-ray beams occurs owing to the different refractive indices (for X-rays) of the ambient medium (typically a vacuum), and the thin film being studied. Although for X-rays such refractive index differences are small ($<10^{-4}$), at the shallow incident angles used in GISAXS ($\theta_i < 1°$) beam refraction becomes non-negligible (Fig. 3 quantifies the shifts of beam positions and scattering features for the typical case of a polymer film coated onto an inorganic substrate). For an incident beam impinging
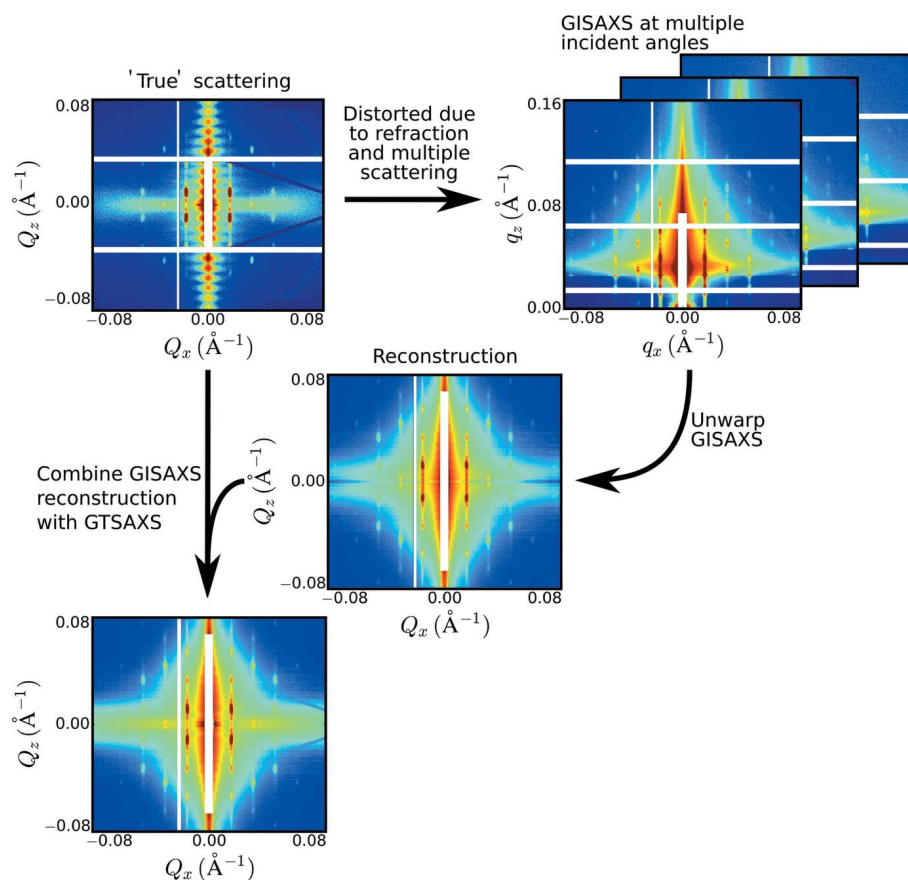

Figure 1
In GISAXS, the sample's 'true' scattering is distorted because of refraction of incident and scattered rays, and multiple scattering effects, which lead to superimposed scattering features and non-monotonic modulation of intensities. These effects greatly complicate the analysis of GISAXS data. Here, we present a method to combine multiple GISAXS images measured at different incident angles, thereby reconstructing a faithful estimate of the unwarped reciprocal-space scattering. This reconstruction can be combined with grazing-transmission measurements (GTSAXS) if available, to further improve data quality. Example data shown in the figure were collected in GISAXS and GTSAXS geometry on a thin film of shear-aligned block copolymer. The reconstruction is difficult to assess near $Q_x \simeq 0$ because of the beamstop. However, the structural peaks resolved elsewhere demonstrate the quality of the reconstruction.

on a thin film at angle θ_i (measured with respect to the sample plane), the beam is refracted to a smaller angle of α_i within the film (refer to Fig. 2 for a diagram of GISAXS geometry),

$$n_0 \cos \theta_i = n_1 \cos \alpha_i, \qquad (1)$$

which is simply the cosine form of Snell's law (with angles measured relative to the film plane), where n_j is the refractive index of the layer j, and we follow the convention that the ambient medium is layer 0, the thin film being studied is layer 1, and the substrate is layer 2. For X-rays, a vacuum has a refractive index of $n_0 = 1$, while typical materials exhibit refractive indices slightly smaller than 1; this leads to total external reflection of X-rays when the incident angle is below the critical angle $\theta_c = \cos^{-1}[n_1/n_0]$. A nanostructured thin film will lead to X-ray scattering in the small-angle regime. Scattering within the film at an angle of α_f (relative to the plane of the film) will be refracted to θ_f as it exits the film surface (Fig. 2 inset). We note that since the critical angle can be conveniently measured in GISAXS or X-ray reflectivity (XRR) experiments, it is useful to consider the refraction equations [such as equation (1)] in terms of θ_c,

$$\cos \alpha_i = \frac{\cos \theta_i}{\cos \theta_c},$$
$$\cos \alpha_f = \frac{\cos \theta_f}{\cos \theta_c}. \qquad (2)$$

As can be seen (Fig. 2), the sample scattering of $2\alpha_s$ is measured experimentally at a larger angle of $2\theta_s$. In terms of reciprocal space, scattering from $Q = 2k \sin \alpha_s$ is observed at $q = 2k \sin \theta_s$ (c.f. Fig. 2). The shift in the position of scattering features can be quantified by (Lu et al., 2013),

$$\Delta q_z = q_z - Q_z \qquad (3)$$

$$\Delta q_z = q_z - 2k \sin\left[\frac{1}{2}\cos^{-1}\left(\frac{\cos\theta_i}{\cos\theta_c}\right) + \frac{1}{2}\cos^{-1}\left(\frac{\cos\theta_f}{\cos\theta_c}\right)\right] \qquad (4)$$

$$\Delta q_z \simeq q_z - k\left[\left(\sin^2\theta_i - \sin^2\theta_c\right)^{1/2} + \left(\sin^2\theta_f - \sin^2\theta_c\right)^{1/2}\right]. \qquad (5)$$

The magnitude of the refraction distortion is plotted in Fig. 3(b). In addition to a constant offset throughout q_z (which can be attributed to the refractive shift of the transmitted beam within the film), there are non-linear distortions that are most extreme as one approaches the thin-film critical angle (which can be attributed to refraction of the outgoing scattered ray). There is a gap in GISAXS data between the horizon and the film critical angle, since scattering within the plane of the film will be refracted out at angle θ_c relative to the horizon.

When the exit angle for scattering matches the film critical angle (θ_c), one can infer that the scattering arose from a beam propagating nominally within the film plane. Such scattering events effectively probe a greater scattering volume, leading to a substantial increase in the intensity of the measured scattering. Experimentally, one thus observes a stripe of intense scattering at the detector for angle $\theta_i + \theta_c$ (Fig. 3a and Fig. S1 in the Supporting information). This 'Yoneda scattering' (Yoneda, 1963; Vineyard, 1982) is useful for measuring the critical angle and thus density of the material being probed, and for enhancing otherwise weak scattering signals.

The incident beam is refracted at the ambient/film interface, and further refracted as it crosses the film/substrate interface. Thus, one measures a transmitted beam on the detector at a position slightly shifted (positive q_z) relative to the direct beam (Fig. 3a). In grazing-incidence geometry, the incident beam may also be partially or fully reflected from the substrate interface (red arrows in Fig. 2). Indeed, below the substrate critical angle, the beam will be completely reflected. The specular beam is reflected from the substrate at the incident angle within the film (α_i) and refracted so that it exits the film at the external incident angle (θ_i). In other words, the specular reflection is measured at the expected position on the detector of $2\theta_i$ with respect to the direct beam (Fig. 3a). The reflected beam, while traveling within the sample, can give rise to scattering events as well. This scattering pattern will be shifted on the detector by $2\theta_i$ owing to the direction of the
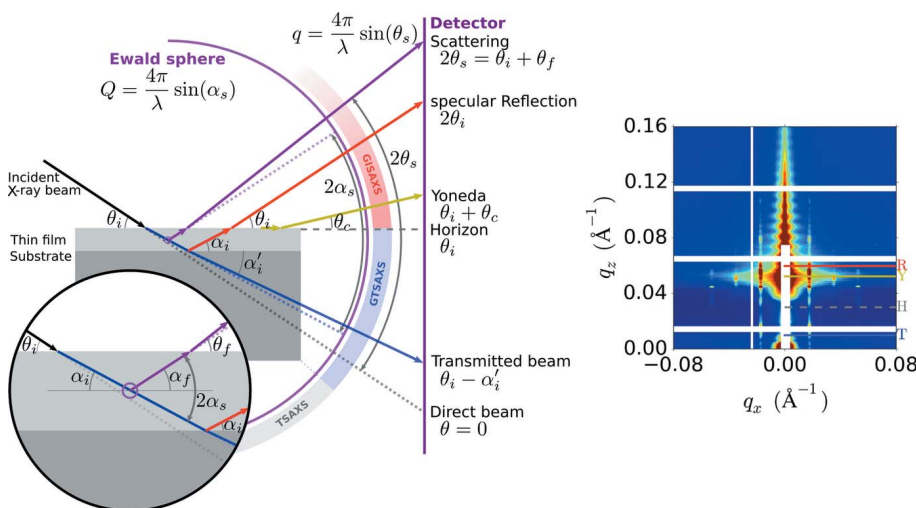

Figure 2
Experimental geometry for grazing-incidence scattering experiments. (Left) An incident X-ray beam (black) is refracted upon entering a thin film, and further refracted upon entering the substrate. Thus, the transmitted beam (blue) is shifted relative to the direct-beam position. Scattered rays (purple) are also refracted as they exit the film. As a result, a scattering event (purple ring) occurring with angle $2\alpha_s = \alpha_i + \alpha_f$ in the sample's reciprocal space (Q) is observed at a different apparent scattering angle ($2\theta_s = \theta_i + \theta_f$) and therefore experimentally measured on the detector at a different reciprocal-space coordinate (q). (Right) An example GISAXS image, with the q_z positions noted for the transmitted beam (T), the horizon (H), the Yoneda (Y) and the specular reflected beam (R).
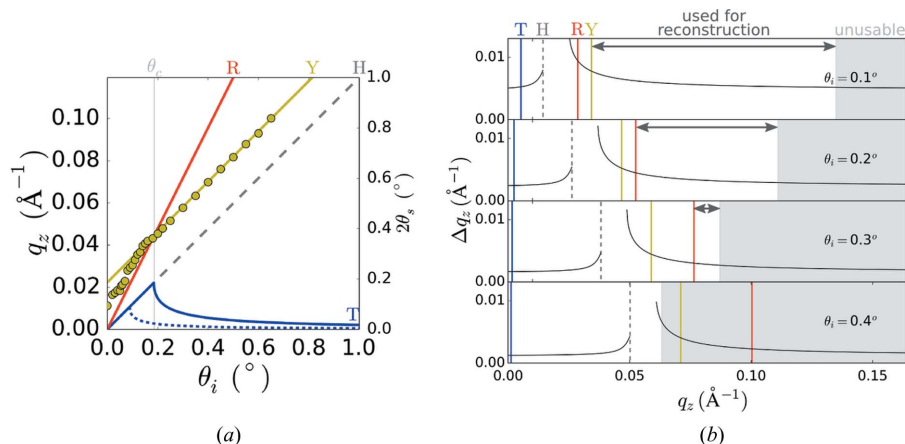
Figure 3
(a) Position along the vertical direction detector (q_z) for various experimental features, as a function of grazing-incidence angle (θ_i), for 13.5 keV X-rays. The horizon (H, dashed gray line) is the plane of the sample itself; sub-horizon scattering must transit through the substrate (and suffer from absorption) while scattering above the horizon exits the film and travels through a vacuum. The specular reflected beam (R, red) appears at $2\theta_i$. When the exit angle matches the critical angle (θ_c) of the film or substrate, one observes enhanced intensity on the detector. This Yoneda (Y, yellow) scattering is offset from the horizon by θ_c. The circles are experimental data, where the position of maximum low-q scattering was identified as the Yoneda, for a block-copolymer thin film on a Ge-coated silicon substrate. The solid yellow line is the theoretical expectation for the critical angle for this substrate. When the incident angle is below the critical angle for the polymer film, the Yoneda shifts to this lower critical angle instead. The transmitted beam (T, blue) is shifted from the direct-beam position ($q_z = 0$) because of beam refraction (the dashed line denotes the position of the transmitted beam within the polymer layer). (b) The refraction distortion (Δq_z) as a function of position along the detector is strongly nonlinear, and depends on both q_z and the incident angle (θ_i). The positions of the various features are denoted by vertical lines: transmitted beam (blue), horizon (dashed gray), substrate Yoneda (yellow), and specular reflection (red). There is a gap in the data between the horizon and the film critical angle. Data close to this gap are generally unusable for reconstruction owing to the large and strongly non-linear distortion. The gray shaded region denotes the area unusable for reconstruction because of the finite range of q_z sampling (refer to *Results and discussion*); the horizontal arrow denotes the span of q_z used for reconstruction.

reflected beam, and by an additional amount because of the aforementioned refraction effects.

These refraction and reflection effects greatly complicate the analysis of GISAXS data, requiring more complex calculations during data fitting, and some knowledge of material properties (such as θ_c). The distortions decrease rapidly (nonlinearly) with increasing incident angle. For GTSAXS at sufficiently large θ_i, the detector image becomes a slightly shifted, linear-intensity representation of reciprocal space; *i.e.* $I_d(q) \simeq I_R(Q = q)$. In practice, only a modest incident angle ($\theta_i \gtrsim 0.6°$) is necessary to greatly reduce the distortions. Thus, in this work we use careful GTSAXS measurements to probe the 'true' scattering, and use this as a means of evaluating the quality of our reconstruction method.

2.2. Distorted-wave Born approximation

For conventional SAXS data, one can typically model the data in the Born approximation, wherein one assumes that the photon field throughout the sample is uniform and simply given by the incident plane wave. In other words, the scattering of objects within the sample is assumed to be sufficiently weak (compared to the high-flux incident beam) that it does not substantially perturb the overall field. This ignores the possibility of multiple scattering, wherein the scattering from one nano-object in the sample is scattered again by another nano-object. In the case of GISAXS, this approximation breaks down. In particular, the shallow incident angles give rise to non-negligible reflection at interfaces. Thus, scattering can occur because of the incident beam, as well as the beam reflected off the substrate. The scattering itself can also be

reflected. In principle, arbitrary sequences of scattering and reflection events contribute to the overall image measured on the detector. When the incident angle is close to the film critical angle, the beam can couple into waveguide modes, where the beam travels essentially within the film plane, and generates a pattern of standing waves within the film (Narayanan *et al.*, 2005; Lee *et al.*, 2006; Jiang *et al.*, 2011, 2012). This greatly enhances the effective scattering volume probed by the beam and thus the measured scattering intensity. These complex multiple-scattering effects can be accounted for by computing the total field within the sample through the interference of all possible transmission/reflection/scattering events. This computationally complex problem is typically simplified, with only the most intense scattering contributions considered. The most popular approach is the DWBA, wherein one considers the scattering from the incident field (Born approximation) as well as the first three multiple-scattering terms (Sinha *et al.*, 1988; Lazzari *et al.*, 2007; Renaud *et al.*, 2009).

In the DWBA, the measured scattering intensity $I_d(q_z)$ is computed using

$$I_d(q_z) = \big| T(\alpha_i)T(\alpha_f)F(+Q_{z1}) + T(\alpha_i)R(\alpha_f)F(-Q_{z2})$$
$$+ R(\alpha_i)T(\alpha_f)F(+Q_{z2}) + R(\alpha_i)R(\alpha_f)F(-Q_{z1}) \big|^2, \quad (6)$$

where Q_{z1} and Q_{z1} are vectors in reciprocal space that take into account the beam direction within the film,

$$Q_{z1} = k\sin\alpha_i + k\sin\alpha_f$$
$$Q_{z2} = k\sin\alpha_i - k\sin\alpha_f, \quad (7)$$

and $k = 2\pi/\lambda$. While Q_{z1} corresponds to the typical definition of reciprocal space, Q_{z2} represents a shift caused by a (single) reflection event. If two reflection events occur, the reciprocal vector instead becomes $-Q_{z1}$. As noted previously, the Q_z values must be computed after taking into account the refraction effect appropriate for that (θ_i, θ_f). The scattering of the sample itself is captured by $F(Q)$, which may be thought of as the form factor and/or structure factor of the sample. The four terms in the DWBA can be interpreted physically as four possible configurations for a scattering event (Fig. 4). The first term represents scattering that occurs directly from the incident refracted beam (Born approximation). We denote this term TT since the incident and scattered rays are transmitted (as opposed to reflected). The second term represents the case where scattering from the incident beam is reflected from the substrate interface (TR). The third term represents scattering from the reflected beam (RT). The final term covers the case where the incident beam is reflected, scattering occurs, and this scattering is itself reflected from the substrate (RR). The factors of $T(\alpha)$ and $R(\alpha)$ represent the amplitudes of the transmission and reflectivity factors, respectively, evaluated for the angle of the beam within the film (α_i or α_f). These can be computed in the usual way by applying the multilayer matrix formalism appropriate for the system being considered (Rauscher *et al.*, 1999; Renaud *et al.*, 2009) (also refer to Fig. S2 in the Supporting information).

Although the DWBA treatment of GISAXS data is necessarily an approximation, it is very successful in reproducing experimentally observed data (Lee, Yoon *et al.*, 2005; Tate *et al.*, 2006; Busch *et al.*, 2006) since the four terms it considers are the dominant sources of scattering, with additional terms involving progressively more $R(\alpha)$ prefactors, which correspondingly decreases their intensity.

It is important to note that both the refraction effects and multiple-scattering effects of GISAXS occur in the q_z direction. Beam refraction shifts scattering along the q_z direction, and the transmission/reflection coefficients [$T(\alpha)$ and $R(\alpha)$] depend only on the angles within the plane of incidence (and are invariant across q_x at constant q_z). This allows us to directly equate the x coordinate of the detector space and the true reciprocal space ($Q_x = q_x$), and allows different q_x columns within the GISAXS reconstruction to be considered independently.

2.3. Simplified DWBA

We investigate a series of simplifications to the DWBA formalism for two reasons. Firstly, any simplification which reduces computational complexity will allow the iterative reconstruction to converge more rapidly. Secondly, we wish to use a form of the DWBA equations that emphasizes physically measurable quantities (θ_c, measurable aspects of reflectivity, and intensities rather than complex amplitudes) to minimize the number of additional assumptions about the sample's structure or scattering required during reconstruction. The DWBA equation (6) can be expanded *via* multiplication to yield 16 terms, which can be regrouped as (the notation $T_i = T(\alpha_i)$ and $F_{+1} = F(+Q_{z1})$ is used for compactness),

$$
\begin{aligned}
I_d(q_z) = & |T_i T_f|^2 |F_{+1}|^2 + |T_i R_f|^2 |F_{-2}|^2 + |R_i T_f|^2 |F_{+2}|^2 \\
& + |R_i R_f|^2 |F_{-1}|^2 + 2|T_i|^2 \mathrm{Re}[T_f R_f^* F_{+1} F_{-2}^*] \\
& + 2|T_f|^2 \mathrm{Re}[T_i R_i^* F_{+1} F_{+2}^*] + 2|R_i|^2 \mathrm{Re}[T_f R_f^* F_{-1}^* F_{+2}] \\
& + 2|R_f|^2 \mathrm{Re}[T_i R_i^* F_{-1}^* F_{-2}] + 2\mathrm{Re}[T_i R_i^* T_f R_f^* F_{+1} F_{-1}^*] \\
& + 2\mathrm{Re}[T_i R_i^* T_f^* R_f F_{+2}^* F_{-2}]. \quad (8)
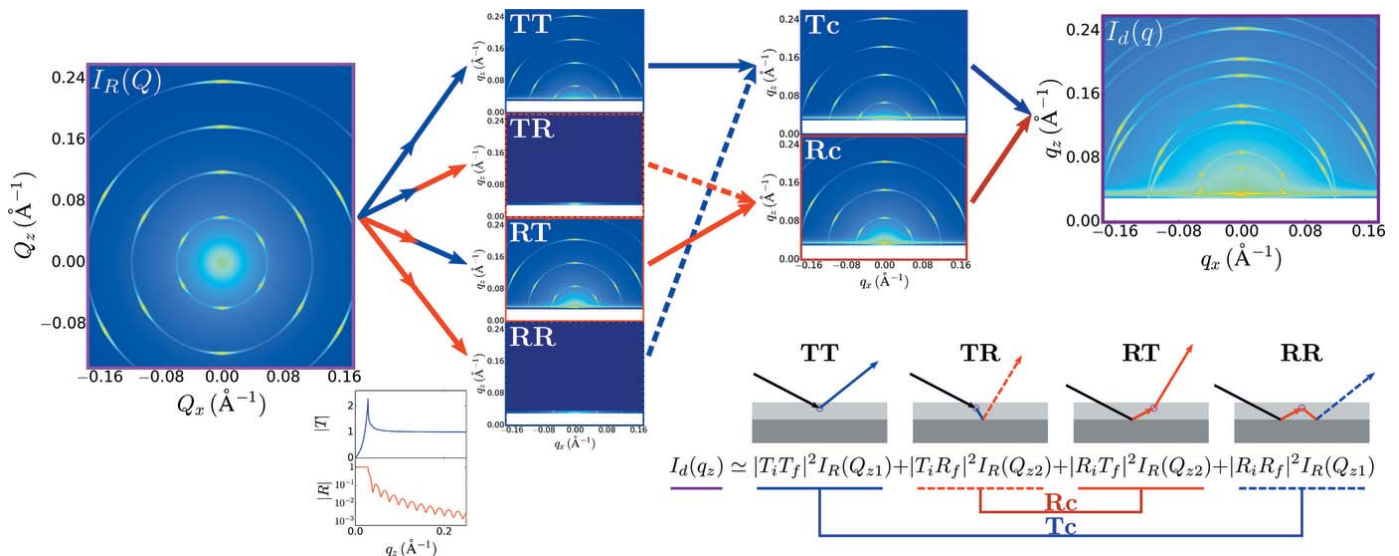\end{aligned}
$$

Figure 4
Graphical representation of the distorted-wave Born approximation (DWBA) used to calculate the total scattering intensity in a GISAXS experiment. The DWBA has four terms (TT, TR, RT and RR) which can be understood in terms of the four most probable configurations for a scattering event. The scattering amplitude from these four terms interfere coherently; to a good approximation their intensity contributions are simply additive. The four scattering terms can be grouped into two scattering channels: one where the scattering pattern is centered about the transmitted beam (Tc channel) and one where the scattering is centered about the reflected beam (Rc channel). The summation of these two channels yields the measured detector image [$I_d(q)$], accounting for the doubling, shift and distortion of the scattering patterns observed in GISAXS.

The first four terms in equation (8) can be thought of as the intensity (amplitude squared) of the four scattering configurations considered in the DWBA (the independent scattering), while the remaining terms are cross terms representing interference effects between these primary components. We note that many of these terms decrease significantly as we probe to larger q_z (especially any that contain a factor of R_f). As we shall see later, our method focuses on data in this large-q_z regime, and thus many of these terms can be ignored. It is also worth emphasizing several differences between the independent terms and the cross terms. While the independent terms are strictly real and positive, the cross terms are complex and may be positive or negative; in the general case, the cross terms will not be in-phase and will thus partially or fully cancel out one another. Furthermore, in a real GISAXS experiment, the beam probes an ensemble of nanostructures that give rise to an average scattering signal. As noted by Lee, Park, Yoon *et al.* (2005), the ensemble average over scatterers with uncorrelated randomness causes the average of the cross terms to be small. Thus, in many experimental scenarios, these cross terms give rise to only subtle high-frequency modulations in the data (especially near the critical angles of the film or substrate) and can be neglected in the data analysis (Omote *et al.*, 2003; Lee, Park, Hwang *et al.*, 2005) (refer to Fig. S3 in the Supporting information for an example). On the other hand, in the limit of monodisperse structures being probed by a highly coherent beam, these cross terms may substantially influence the scattering image.

Defining the scattered intensity in reciprocal space as $I_R(Q) = |F(Q)|^2$ (and ignoring cross terms), equation (8) can be written as

$$I_d(q_z) \simeq |T(\alpha_i)T(\alpha_f)|^2 I_R(+Q_{z1}) + |T(\alpha_i)R(\alpha_f)|^2 I_R(-Q_{z2})$$
$$+ |R(\alpha_i)T(\alpha_f)|^2 I_R(+Q_{z2})$$
$$+ |R(\alpha_i)R(\alpha_f)|^2 I_R(-Q_{z1}). \qquad (9)$$

Since reciprocal space is centrosymmetric [$I(+Q_z) = I(-Q_z)$], scattering patterns will exhibit twofold symmetry in the small-angle limit (although this is violated in cases of highly oriented materials). Taking advantage of this simplification, and noting that $Q_{z2} = Q_{z1} - 2k \sin(\alpha_i)$, we can rearrange equation (9) into

$$I_d(q_z) \simeq [|T(\alpha_i)T(\alpha_f)|^2 + |R(\alpha_i)R(\alpha_f)|^2] I_R(Q_{z1})$$
$$+ [|T(\alpha_i)R(\alpha_f)|^2 + |R(\alpha_i)T(\alpha_f)|^2]$$
$$\times I_R[Q_{z1} - 2k \sin(\alpha_i)]. \qquad (10)$$

The first term of equation (10) ($TT + RR$) can be identified as the scattering that is centered about the transmitted beam on the detector image; we refer to this as the transmitted channel (Tc). The second term ($TR + RT$) can be identified as the scattering that is shifted on the detector image (by $\sim 2\theta_i$) and is centered about the specular reflected beam; we refer to this as the reflection channel (Rc). The summation of these two channels gives rise to the overall detector image (Fig. 4).

2.4. Iterative reconstruction

We deploy a straightforward iterative algorithm to reconstruct the sample's reciprocal space. We generate a candidate reconstruction [$I_{R,\text{test}}(Q)$], convert this to a detector image [$I_{d,\text{test}}(q_z)$] using the above-described refraction and multiple-scattering corrections {which can be thought of in terms of a transformation function $\mathbb{D}[I_R(Q_z), \theta_i] = I_d(q_z, \theta_i)$}, and compare this with the experimental data. We simultaneously fit multiple GISAXS images taken at different incident angles. Thus, we compute the mismatch (χ_d^2) between the candidate detector image and the experimental GISAXS data [$I_{d,\text{true}}(q_z)$] by summing the residuals across the N_{q_z} available pixels of the N_{θ_i} GISAXS images,

$$\chi_d^2(q_x) = \sum^{N_{\theta_i}} \sum^{N_{q_z}} |\log[I_{d,\text{test}}(q_x, q_z, \theta_i)] - \log[I_{d,\text{true}}(q_x, q_z, \theta_i)]|^2$$
$$= \sum^{N_{\theta_i}} \sum^{N_{q_z}} |\log\{\mathbb{D}[I_{R,\text{test}}(q_x, Q_z), \theta_i]\}$$
$$- \log[I_{d,\text{true}}(q_x, q_z, \theta_i)]|^2. \qquad (11)$$

The residuals are computed on the logarithm of the data, which helps to fairly balance the contribution from intense scattering features (low-q diffuse scattering, sharp peaks, *etc.*) and weaker features (high-q scattering, higher order peaks, *etc.*). As will be described later, selecting an appropriate sub-range for q_z is crucial to obtaining a reliable reconstruction. Conversely, including a greater number of GISAXS images will generally improve the quality of the reconstruction. We minimize χ_d^2 using a multi-variable constrained optimizer, which iteratively converges towards smaller values of the target function using successive local refinements. In particular, we select a bound pseudo-Newton method (Byrd *et al.*, 1995; Zhu *et al.*, 1997; Morales & Nocedal, 2011), finding it to converge more rapidly than unconstrained Newton or trust-region methods (Conn *et al.*, 2000). We terminate the iteration when χ_d^2 no longer improves with continued searching. This optimization procedure is repeated for each column of constant q_x across the GISAXS image. In principle, these columns are mathematically independent and can be solved in parallel. In practice, there are advantages to exploiting the similarity in neighboring columns to improve the convergence speed.

Our method is implemented in the Python programming language (Millman & Aivazis, 2011), taking advantage of the *scipy* library for fast optimization (Oliphant, 2007), the *numpy* library for numeric handling (Walt *et al.*, 2011) and the *matplotlib* library for plotting (Hunter, 2007). The minimization is performed using the `scipy.optimize.minimize` function, and invoking the `L-BFGS-B` method.

2.5. Experimental data

To validate our methods, we conducted a series of X-ray scattering measurements on nanostructured thin films using synchrotron beamlines. Data presented here were collected at the X9 beamline of NSLS (at X-ray energy of 13.5 keV), at the Soft Matter Interfaces (SMI, 12-ID) beamline at NSLS-II (11.0 keV) and at the Complex Materials Scattering (CMS, 11-

BM) beamline at NSLS-II (13.5 keV). Samples were carefully aligned with respect to the incident beam using both the direct-beam scattering intensity and the position of the reflected beam on the detector. GISAXS data were collected across a range of incident angles. Corresponding GTSAXS data were collected by increasing the incident angle and realigning the sample height such that the beam was impinging on the downstream edge of the sample, thereby maximizing the sub-horizon scattering intensity (Lu *et al.*, 2013). We selected a range of samples to study, in order to test the performance of our method on different types of scattering images, taking advantage of both lithographic methods, which can generate extremely precise nanostructures (Luttge, 2009), and block-copolymer self-assembly, which can be used to generate a variety of two-dimensional and three-dimensional morphologies over wide areas (Doerk & Yager, 2017). Lithographic line-gratings fabricated using electron-beam

lithography were measured with the X-ray beam along the grating groove length (Johnston *et al.*, 2014). Three-dimensional polymer nanostructures were fabricated using a commercial two-photon absorption laser lithography tool (Nanoscribe Photonic Professional GT). Thin films of a block-copolymer-cylinder phase, with cylinder axes oriented perpendicular to the film normal were shear-aligned using photo-thermal methods (Majewski & Yager, 2015a,b; Majewski, Rahman *et al.*, 2015) and measured with the X-ray beam along the shear direction (*cf.* Figs. 1, 2, 3 and 5a). Multilayered nanostructures were fabricated by iterative assembly of block-copolymer phases combined with sequential infiltration synthesis to convert the self-assembled organic phases into inorganic replicas (Rahman *et al.*, 2016) (Figs. 5b and 6). An ordered and aligned array of hexagonally packed nano-dots, fabricated using block-copolymer self-assembly was also measured (Fig. 5c) (Choo *et al.*, 2017). To compute
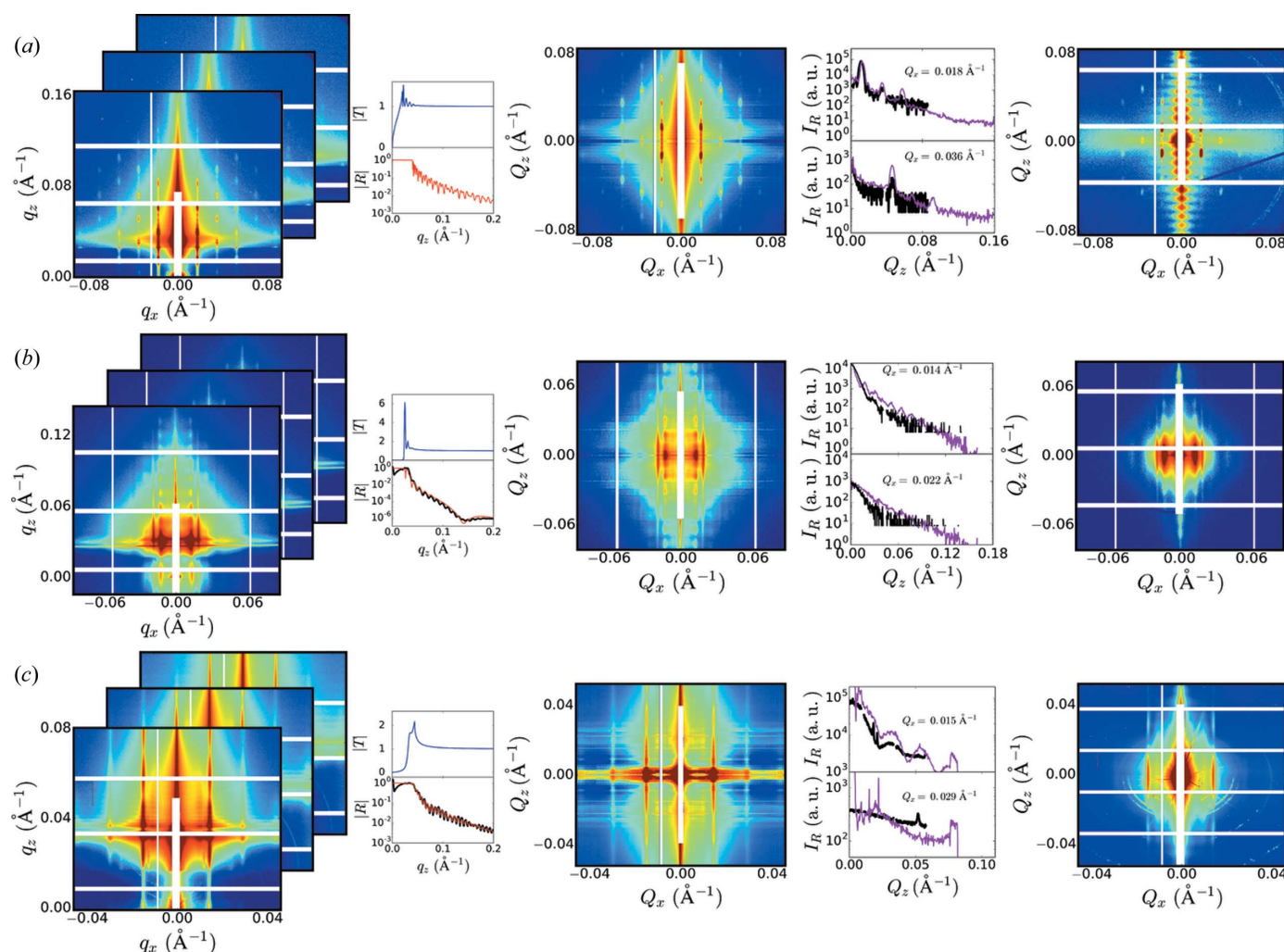

Figure 5
Example reconstructions based on experimental GISAXS data. Each row shows (from left to right) the set of GISAXS images, the transmission and reflectivity curves (model curves in blue/red; experimentally measured reflectivity in black), the reciprocal-space reconstruction [$I_R(Q_z)$], select linecuts through the reconstruction (reconstruction in purple, GTSAXS data in black) and the experimental GTSAXS image. The presented data are (*a*) a self-assembled phase of horizontally oriented cylinders that form a hexagonal lattice, (*b*) multilayered block-copolymer nanostructures and (*c*) an in-plane array of inorganic nano-dots.

the transmission and reflection coefficients (T and R) for these samples, we either estimated the layer profile from knowledge of sample fabrication, or experimentally measured the X-ray reflectivity (XRR) curve (Rigaku Ultima III), and used the best fit to compute the coefficients. For instance, for the sample shown in Fig. 5(a) (data collected for $\lambda = 0.9184$ Å) we use a model with a 125 nm polymer layer ($\theta_c = 0.09°$), on a 100 nm Ge layer ($\theta_c = 0.17°$), on a SiO$_2$ substrate ($\theta_c = 0.14°$). For some tests, we generated synthetic data (Yager *et al.*, 2017), allowing us to systematically vary properties of reciprocal space and the film structure. For the synthetic data used throughout the paper (such as Fig. 8), we assume a 30 nm polymer layer ($\theta_c = 0.11°$), on a 100 nm Si$_2$N$_3$ layer ($\theta_c = 0.09°$), on a Ge substrate ($\theta_c = 0.20°$).

3. Results and discussion

Although the transformation from SAXS data to GISAXS data is well understood (see Section 2), inverting this transformation is not readily possible. The inversion is a mathematically ill-posed problem, owing to experimental noise, the one-to-many mapping between the Q_z and q_z spaces, and the limited range of the data. A scattering feature at a particular Q_z will contribute intensity to two different q_z locations (Tc and Rc channels). Similarly, in the inverse problem, the intensity at a particular q_z has contributions from two different Q_z. Reconstructing $I_R(Q_z)$ thus requires solving a coupled set of equations; however, the limited range of experimentally available q_z makes the problem formally underdefined (with references outside the available data range). On the other hand, the inversion is only weakly underdefined, and an approximate solution can be found using modest, physically reasonable assumptions. Our method for reconstructing the true reciprocal-space scattering image consists of first constructing a reasonable guess for $I_R(Q_z)$, and then iteratively adjusting this estimate by fitting available GISAXS data [$I_d(q_z)$], minimizing the fit-error [χ^2_d, equation (11)]. Reconstruction involves repeatedly applying the DWBA transformation (Fig. 4), and thus requires knowledge of θ_i for the GISAXS data, and reasonable estimates for θ_c, $|T|$, and $|R|$. The latter can be estimated from knowledge of the sample makeup, or measured using X-ray reflectivity.

Only a subset of a GISAXS image contains data usable for the reconstruction. While the finite size of the X-ray detector provides a hard limit on the q_z range, other effects impose additional constraints. At high q_z, the signal becomes weak and unreliable. At small q_z (approaching the film or substrate critical angle), the refraction distortion becomes large and highly nonlinear, making the mapping to Q_z error prone. In this regime, intensity corrections also become more error prone (sensitive to the assumed form of $|T|$ and $|R|$). In practice, we select only the GISAXS data that are strictly above both the substrate critical angle and the reflected-beam position (yellow and red lines in Fig. 3). The Rc channel is symmetric about the reflected-beam position, and thus data below the reflected beam are largely redundant. The reconstruction method is implicitly identifying a form of $I_R(Q_z)$ that

yields Tc and Rc channel predictions mutually consistent with the data. This procedure is only well defined over the q_z range where the two channels overlap. Since the Rc channel is shifted by $+2\theta_i$ relative to the Tc channel, any data within $2\theta_i$ of the upper edge of the detector (gray shaded region in Fig. 3) are not usable (overlap condition is not satisfied).

The GISAXS incident angle influences the reconstruction quality. For large θ_i, the overlap region becomes very small and the corresponding Q_z range of the reconstruction is limited. In principle, large θ_i should simplify the reconstruction since the reflection components become weak, and one need only consider the Born-approximation term of the DWBA. However, the concomitant decrease in overall scattering intensity makes the data noisier. A key feature of our approach is to simultaneously fit a set of GISAXS images at different θ_i (rather than a single image). This provides some redundancy in the data, which helps to average out noise in the data and artifacts from the fitting procedure, and to converge to a unique minimum during reconstruction. This is especially important since scattering features are shifted because of refraction, and convolved with the transmission and reflectivity curves. At a particular incident angle, two distinct scattering features may overlap, or a scattering feature may be suppressed (*e.g.* because of a minimum in $|R|$). By combining multiple incident angles, each feature in $I_R(Q_z)$ is cleanly represented somewhere in the overall GISAXS data set. The selection of incident angles of course influences the quality of the reconstruction. Selecting similar θ_i is suboptimal since the GISAXS curves are similar and thus do not offer meaningfully independent constraints. On the other hand, selecting very large θ_i is not helpful since such data have a very limited usable q_z range. It is also important to note that the computation-time scales linearly with the number of q_z data points, and thus with the number of θ_i one considers. We empirically find that selecting three GISAXS images, with incident angles spaced by $\Delta_{\theta_i} \simeq 0.02°$ yields robust reconstructions (*e.g.* 0.14, 0.16, and 0.18°). An additional benefit to using multiple incident angles is that this implicitly fills any gaps in the data caused by, for example, inter-module gaps in the detector. Because each θ_i probes a slightly different part of Q_z space, the gaps are not aligned and thus a set of GISAXS images can be used to generate an uninterrupted span of Q_z. Any gaps remaining in Q_x can also be filled using previously reported image-healing methods (Liu *et al.*, 2017).

Fig. 6 shows an example of fitting an experimental set of three GISAXS images. The fitting generates a set of GISAXS images that closely match the experimental data. The underlying $I_R(Q_z)$ reconstruction is a close match to the true undistorted scattering (which can be estimated from the experimentally measured GTSAXS pattern). A direct comparison between the reconstruction and the GTSAXS data (Fig. 6, lower-left) highlights that the reconstruction reproduces the correct positions and scaling of scattering features (*e.g.* peaks). Moreover, the SNR of the reconstruction is much better than that available for GTSAXS (*e.g.* higher order peaks are visible), since it takes advantage of the

GISAXS intensity-enhancement effects (especially beam projection over the sample surface).

A crucial aspect for rapid convergence of the reconstruction method is the selection of the initial guess for $I_R(Q_z)$. We find that the iterative reconstruction is well behaved in the sense that any initial guess will converge to the same (correct) reconstruction. For instance, initializing $I_R(Q)$ to noise, or starting with $I_R(Q) = 0$ will eventually converge to a valid reconstruction (see Fig. S4 in the Supporting information). It is also worth noting that the fit-error to the experimental data (χ_d^2) is highly correlated to the mismatch of the reconstruction (χ_R^2); in other words, minimizing the fit-error does converge to the true reciprocal-space scattering (see Figs. S5 and S6 in the Supporting information). Although the final reconstruction quality is robustly insensitive to the starting guess for $I_R(Q_z)$, the number of iterations required for a high-quality reconstruction depends strongly on the initial guess. Thus, there is a huge computational benefit to selecting an appropriate initialization. A straightforward initialization is to simply copy the GISAXS curve into Q_z space. While this is necessarily incorrect in many ways (e.g. exhibiting double peaks), it at least sets a roughly correct scale for the scattering intensity (capturing the experimentally observed decay with q_z). Since each Q_x column can be reconstructed separately, another option is to select a neighboring reconstructed column as the initial guess for a new column. As might be expected, this greatly improves convergence if one selects a nearby neighbor, and is somewhat less successful if the column is far away or the scattering intensity varies dramatically along q_x (see Fig. S4 in the Supporting information).

We can generate a much better initial guess by taking advantage of the fact that a GISAXS curve conceptually contains two copies of the underlying scattering pattern because of the contributions of the Tc and Rc channels. We reformulate the problem of estimating the individual channel intensities $[I_{d,Tc}(q_z)$ and $I_{d,Rc}(q_z)]$ into a problem of estimating
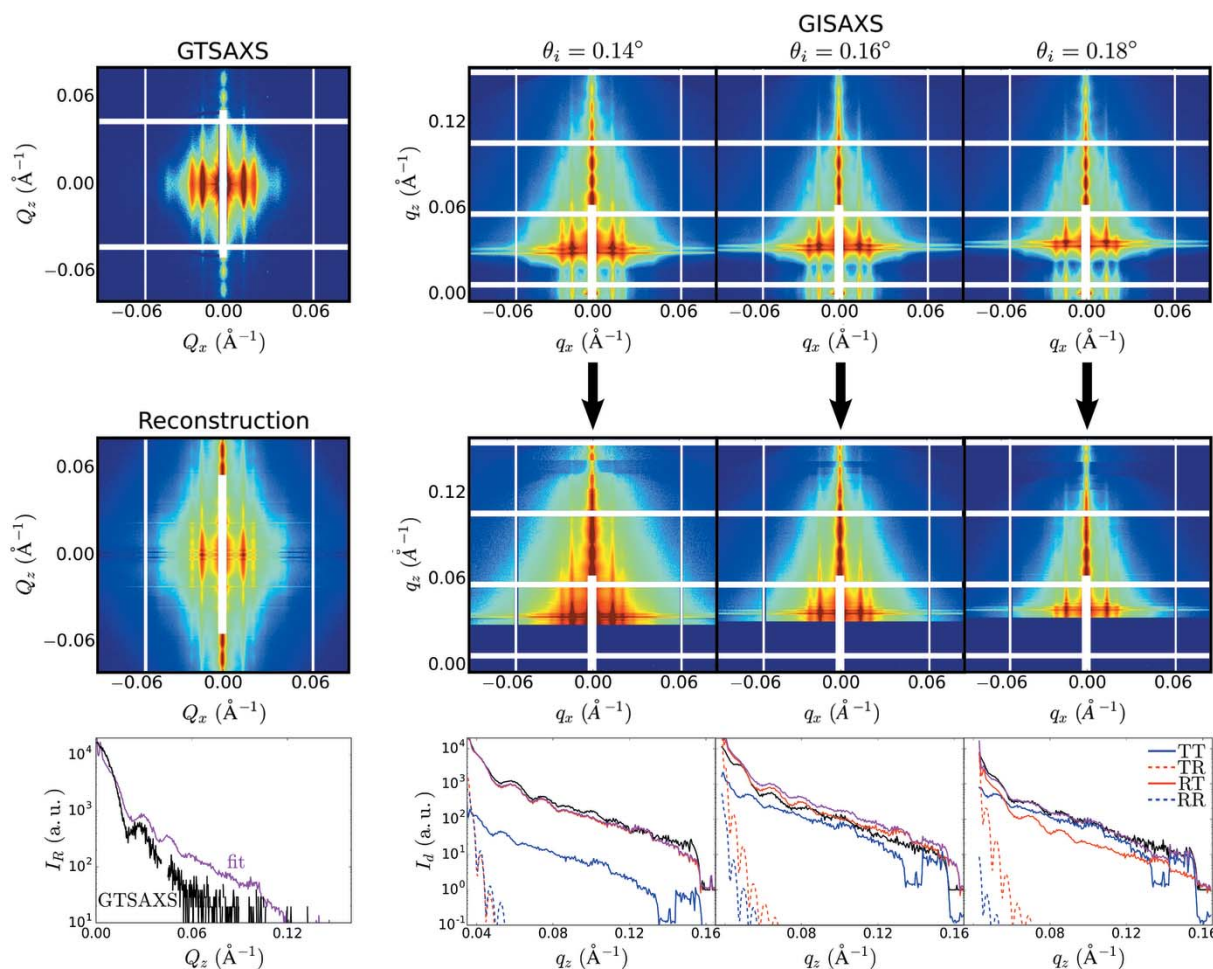

Figure 6
Example reconstruction of true (SAXS) scattering data based on combined fitting of multiple experimental GISAXS images. The corresponding experimental GTSAXS data are presented as an alternative estimate for the true SAXS scattering pattern (upper left). The upper row shows the experimental GISAXS images (at multiple incident angles), while the middle row shows the corresponding GISAXS fits obtained from reconstruction. The bottom row shows example one-dimensional linecuts through the data (at $q_x = 0.0138$ Å^{-1}). The experimental data are shown in black, the reconstruction/fit is shown in purple and the individual DWBA components are shown in blue or red. The direct comparison between the reconstruction and the GTSAXS data (bottom left) demonstrates that the GISAXS reconstruction is robust and has a higher SNR (GTSAXS curve, black, was offset vertically to aid in comparison).

the ratio of these contributions, $w(q_z)$. This simplifies the problem since one can estimate $w(q_z)$ by evaluating the mismatch between predictions for $I_R(Q_z)$ obtained from $I_{d,Tc}(q_z)$ and $I_{d,Rc}(q_z)$ separately. The experimental data $I_d(q_z)$ result from the summation of two contributions,

$$I_d(q_z) = (|T_i T_f|^2 + |R_i R_f|^2)I_{d,Tc}(q_z) + (|T_i R_f|^2 + |R_i T_f|^2)I_{d,Rc}(q_z)$$
$$= |Tc|^2 I_{d,Tc}(q_z) + |Rc|^2 I_{d,Rc}(q_z). \quad (12)$$

We define the ratio between the channels to be

$$w(q_z) = \frac{I_{d,Tc}(q_z)}{I_{d,Tc}(q_z) + I_{d,Rc}(q_z)}, \quad (13)$$

such that one can compute the two components from

$$I_{d,Tc}(q_z) = \frac{I_d(q_z)}{|Tc|^2 + (|Rc|^2/w) - |Rc|^2} \quad (14)$$

and

$$I_{d,Rc}(q_z) = \frac{I_d(q_z)}{|Rc|^2} - \frac{|Tc|^2}{|Rc|^2}I_{d,Tc}(q_z)$$
$$= \frac{I_d(q_z)}{|Tc|^2 [(w)/(1-w)] + |Rc|^2}. \quad (15)$$

Focusing on solving for $w(q_z)$ conceptually simplifies the problem. In particular, if $w(q_z)$ is known, one can easily compute the contributions from the two channels and can thus reconstruct the correct $I_R(Q_z)$ from either of these; conversely, one can estimate $w(q_z)$ by comparing the mismatch between predictions obtained from $I_{d,Tc}(q_z)$ and $I_{d,Rc}(q_z)$, separately. An algorithm for initializing $w(q_z)$ to a reasonable estimate is shown in Fig. 7. The steps are as follows: (i) we begin by arbitrarily selecting a guess for $w(q_z)$. The simplest starting point is to assume equal contributions from both channels *i.e.* $w(q_z) = 0.5$. We find empirically that using $w(q_z) \simeq 0.44$ is slightly better, since in most cases $I_{d,Rc} > I_{d,Tc}$, since the *Rc* channel is shifted to higher q_z than the *Tc* channel. (ii) Based on the assumed $w(q_z)$, we divide the experimental $I_d(q_z)$ into contributions from the $I_{d,Tc}$ and $I_{d,Rc}$ channels [equations (14) and (15)]. (iii) Based on knowledge of the refraction distortion, the two channels can be shifted from the detector space (q_z) into reciprocal space (Q_z). These act as two related predictions for the undistorted scattering $I_R(Q_z)$. These two curves should be identical; any mismatch between them thus represents an error in our current guess of $w(q_z)$. (iv) We compute the difference between our two predictions,

$$\delta(Q_z) = I_{R,Rc}(Q_z) - I_{R,Tc}(Q_z). \quad (16)$$

(v) From the difference between the curves, we compute a local mixing ratio $m(Q_z)$,

$$m = \frac{1}{1 + \exp(\delta/\sigma_m)}, \quad (17)$$

where σ_m controls the character of the mixing. (vi) The purpose of $m(Q_z)$ is to regulate how we mix together the two channel predictions into an improved estimate for $I_R(Q_z)$. In particular,

$$I_R(Q_z) = m \times I_{R,Tc}(Q_z) + (1 - m) \times I_{R,Rc}(Q_z). \quad (18)$$

For $\sigma_m \to \infty$, $m = 0.5$, and we are thus averaging together the two channel predictions. For $\sigma_m \to 0$, m becomes a step function, and we are thus thresholding, and selecting whichever curve has the lowest value (at each Q_z). Intermediate values of σ_m represent an intermediate strategy, where we include contributions from both curves but emphasize the curve with lower intensity. The underlying rationale is that any mismatch between the two predictions represents a mistake in our guess of w. In particular, if a feature (*e.g.* peak) appears only in one of the channel predictions, then it is erroneous (and can be suppressed *via* m); whereas features that appear in both channels are considered correct and will be preserved by m. (vii) The current estimate for $I_R(Q_z)$ is distorted back into the detector space (q_z). (viii) A new estimate for $w(q_z)$ is calculated [equation (13)]. This is used as a new input [step (ii)] for another iteration of this strategy.

On each iteration, this algorithm produces an improved estimate for $w(q_z)$. Even a single iteration of this method already yields a reasonable guess for the structure of reciprocal space (Fig. 7). The quality of this guess improves rapidly with additional iterations (see Figs. S7 and S8 in the supporting information), since the algorithm specifically focuses on changing the estimate in any regions where the two channels disagree. It is worth noting that the algorithm does not, at first, correctly predict the overall intensity of scattering features (*e.g.* peaks). However, because the method converts from intensity space [$I_R(Q_z)$] to a ratio representation [$w(q_z)$], it immediately shifts back to experimentally validated intensities. This method can be applied (in parallel) to the columns of a two-dimensional data set, allowing a GISAXS image to be converted into a corresponding SAXS image in a very small number of iterations (see Fig. S9 in the supporting information).

Through a series of empirical tests, we find that the w-initialization method produces a reasonable reconstruction of reciprocal space for essentially any form of $I_R(Q_z)$. We investigated well separated sharp peaks, scattering signals with many overlapping peaks, overlapping broad peaks, diffuse scattering, power-law backgrounds, and combinations thereof; we found acceptable reconstructions in all cases. We note that, in general, this method introduces small artifacts (intensity jumps) at locations that are influenced by the overall q range of the data. Minor artifacts also persist at q values where sharp features (*e.g.* peaks) initially appeared in the 'wrong' channel. These artifacts cannot be completely eliminated using the w-initialization method, but are removed once the curve is iteratively improved by directly fitting $I_d(q_z)$. We empirically find that five to ten iterations is sufficient to obtain a reasonable w estimate, with additional iterations not yielding improvements in χ_d^2. The exact rate of convergence is influenced by the m mixing function. We find that larger values of σ_m are somewhat better for scattering data with broad/diffuse features (*i.e.* averaging channel estimates is preferable) while smaller σ_m is preferable for sharp features (*i.e.* thresholding channel estimates is preferable). Moreover, there is an
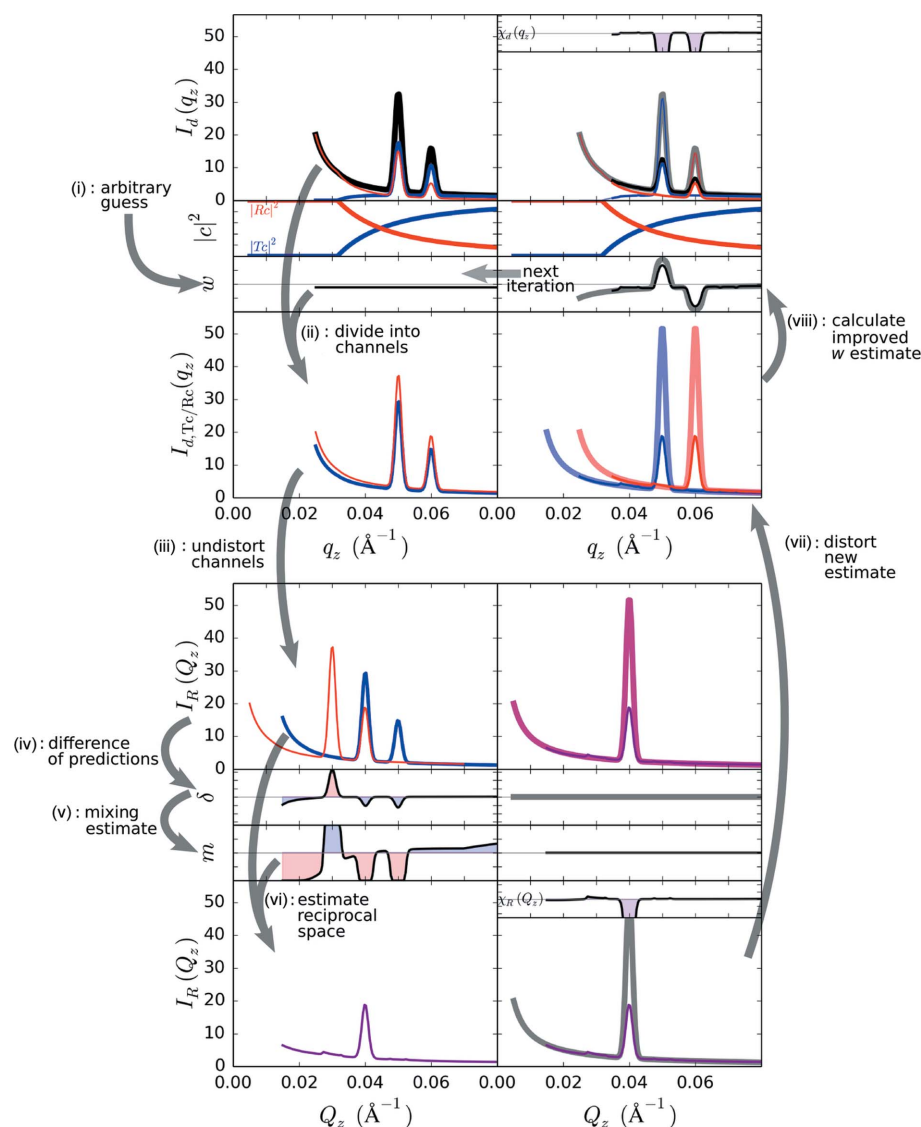
Figure 7

Algorithm for generating an initial estimate of the undistorted scattering [$I_R(Q_z)$] from experimental GISAXS data [$I_d(q_z)$]. The experimental intensity is first divided into contributions from the transmitted and reflected channels (Tc and Rc) based on known transmission/reflectivity curves ($|Tc|^2$ and $|Rc|^2$) and an arbitrary guess for the ratio between the channels (w). The two channels are both undistorted into reciprocal space, which provides two predictions for the true scattering (which should agree). The difference between these predictions is used to compute an improved estimate for I_R. This I_R estimate can be distorted to yield new estimates for the contributions from the two channels, which provides an improved estimate for w. This new w estimate can be fed back as an improved initial guess. (Thick faded lines show true scattering contributions, while χ shows the corresponding residuals; these are of course not known during reconstruction of experimental data.) By iterating through this procedure, the algorithm converges towards a self-consistent prediction for $w(q_z)$ and $I_R(Q_z)$.

$I_R(Q_z)$ in a computationally simple manner. This initial guess is then input into the iterative reconstruction, which then refines this estimate so that it closely matches the experimental $I_d(q_z)$.

The w-initialization strategy, combined with iterative fitting of GISAXS data, is able to robustly reconstruct undistorted scattering from a wide range of experimental GISAXS data (see examples in Fig. 5 and Fig. S10 in the Supporting information). The reconstruction resolves many subtle features (such as inter-peak oscillations) that are smeared or indistinct in the GISAXS data. In a data-visualization context (Zhong *et al.*, 2018), it also removes the ambiguity of deciding whether a GISAXS peak position should be interpreted with respect to the transmitted or reflected beam. Seemingly complex GISAXS patterns become much simpler to interpret after unwarping. The reconstruction method leverages the intensity enhancements of GISAXS to yield a high SNR SAXS image. When comparing with corresponding experimental GTSAXS data, it is clear that the unwarped GISAXS data resolve many features (such as weak higher order peaks) that are lost in GTSAXS because of beam attenuation through the substrate. On the other hand, the reconstructions sometimes exhibit intensity artifacts. For instance, the intensity near $Q_z \simeq 0$ is not reconstructed faithfully, since this data comes from the error-prone portion of the GISAXS image (q_z close to substrate critical angle). Where both GISAXS and GTSAXS data are available, the two can be combined (Fig. 1), wherein one can leverage both the high SNR of GISAXS data (for large Q_z) and the cleaner measurement of GTSAXS (for small Q_z).

advantage to shifting gradually towards smaller σ_m as iteration continues, which conceptually means shifting from averaging to thresholding. We select values of σ_m using the scale of the δ values themselves. In particular, we define δ_{max} as the maximum (absolute) δ value, and find that a value of $\sigma_m \simeq 0.1\delta_{max}/(1 + i)$ (where i is the iteration number) works in the general case (with $\sigma_m \simeq 0.015\delta_{max}$ being better for sharp features, and $\sigma_m \simeq 10\delta_{max}$ for diffuse features). Overall, the w-initialization strategy is a simple and robust means of obtaining a reasonable initial estimate for the structure of

The reconstruction method works across a wide range of different scattering patterns, and is relatively insensitive to the shape of the reciprocal-space scattering. We evaluated the performance of our method using synthetic data (Fig. 8), allowing us to systematically compare the reconstruction to the correct scattering. Both sharp and broad features in scattering data are properly reconstructed (Fig. 8a), with sharp features being somewhat easier to fit (see Fig. S6 in the Supporting information). Since the reconstruction method

requires knowledge of several experimental and materials parameters, it is natural to ask whether the method is sensitive to errors in these inputs. Any error in the assumed value of θ_i will naturally shift the reconstructed Q_z by that amount (Fig. 8b), and may introduce additional intensity artifacts. We also find that an erroneous θ_i sharply increases χ_d^2. Thus, the value of θ_i could also be refined during the reconstruction. We similarly observe that errors in θ_c will shift the reconstruction along Q_z (see Fig. S11 in the Supporting information); this angle value itself can be refined since the correct value is a deep and well defined minimum in the χ_d^2 error surface. While θ_i is typically precisely known from the experimental alignment, θ_c is materials and sample dependent, and should ideally be measured from the data rather than assumed. The reconstruction method requires estimates for the transmission

and reflectivity curves ($|T|$ and $|R|$). The X-ray-reflectivity curve can be measured experimentally, or estimated based on the known layering of the sample. Although the reconstruction quality obviously depends upon these curves, we find that the reconstruction is surprisingly robust to errors in these curves. For modest errors in the assumed curves, the reconstruction remains qualitatively correct (Fig. 8c). This suggests that as long as a reasonable estimate for the X-ray reflectivity curve can be provided by the experimenter, the reconstruction method can yield a useful estimate for the structure of reciprocal space. With respect to the materials-layer model that underlies the reflectivity curve, we find that the reconstruction is quite robust. Errors in the film thickness shift the overall $I_R(Q_z)$ intensity (see Fig. S12 in the Supporting information) while correctly capturing the position of features; in extreme
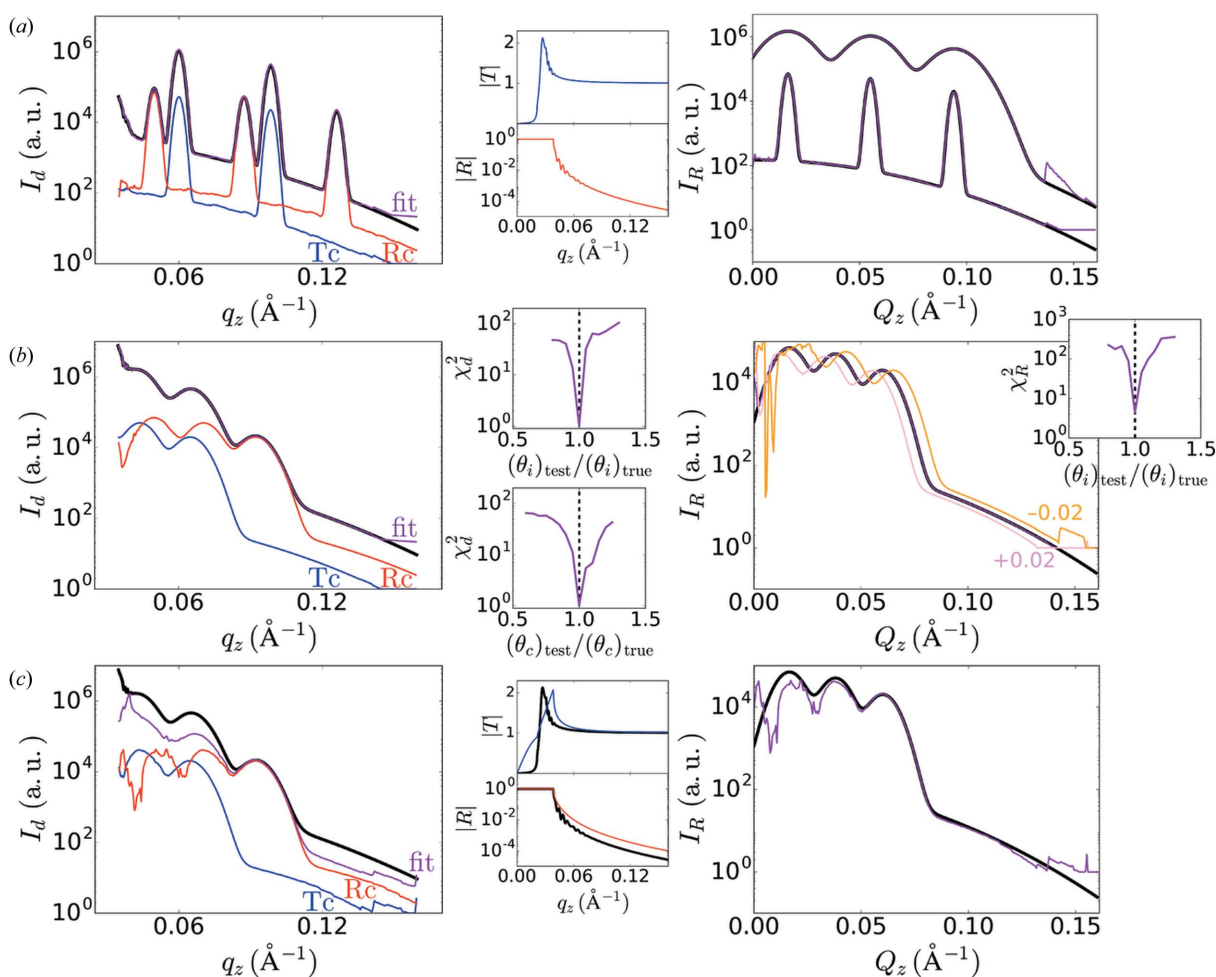

Figure 8
Sensitivity of reconstruction method, evaluated using synthetic data. (a) The input synthetic GISAXS data (black curve), $I_d(q_z)$ (example on left) can be iteratively fitted (purple curve), from which we reconstruct the undistorted scattering [$I_R(Q_z)$, right]; the corresponding transmission channel (blue $I_{d,Tc}$) and reflection channel (red $I_{d,Rc}$) components are calculated using the known transmission ($|T|$) and reflectivity ($|R|$) curves (shown in center). Two different examples (right) are shown (reconstruction shown in purple, true scattering in black), demonstrating that the method works for sharp and diffuse features (with artifacts appearing towards the edges of the available q_z range). (b) The method is relativity robust to errors in the assumed incident angle (θ_i) and film critical angle (θ_c). For a typical data set (left) with intentional errors introduced into θ_i, the reconstruction becomes correspondingly shifted (right, numbers indicate intentional error in θ_i). The fit-error (χ_d^2) with respect to errors in the angles has a deep minimum at the true value, allowing this value to also be iteratively refined. (c) Errors in the assumed transmission and/or reflectivity curves necessarily corrupt the reconstruction. However, for modest errors in these curves, the reconstruction remains qualitatively correct. In the example presented here, incorrect transmission and reflectivity curves were used intentionally (center; true curves in black). The resulting reconstruction (right) exhibits artifacts yet nevertheless maintains the correct overall shape and intensity.

cases an erroneous thickness estimate can introduce spurious oscillations into the reconstruction. Errors in the film roughness or film absorbance (the imaginary part of refractive index) lead to only very minor deviations in the final reconstruction (see Figs. S13 and S14 in the Supporting information).

The iterative reconstruction method we employ can be computationally expensive since the method must repeatedly perform the DWBA calculations and must refine the intensity for each point in the (Q_x, Q_z) space one is constructing. The computation time increases linearly with the number of detector pixels and the number of θ_i one includes. However, there are a number of ways to minimize the computational burden. Many intermediate results [distortion shift of equation (3), channel prefactors in equation (10)] can be computed once and used repeatedly. We also greatly minimize the number of iterations required during reconstruction by first generating a high-quality initial guess for reciprocal space; by using either a w-initialization algorithm, or using nearby reconstructed data. The Q_x columns can be refined independently, and it is thus trivial to parallelize in this manner. We find that such a parallelization provides nearly linear speedup (e.g. a 6× speedup when computing using eight CPU cores). The exact computation time depends on the size of the data set and the hardware being used. As a guide, we find that when fitting multiple (three) high-resolution (1000×1000 pixels) GISAXS detector images the reconstruction takes ∼2000 s using a single modern CPU and the near-neighbor initialization strategy. Reconstructing the same image using eight CPU cores and the w-initialization strategy requires <200 s. The reconstruction time could be further reduced in a number of ways. Smaller images will, of course, be even faster to reconstruct; that is, there is an advantage to only reconstructing the desired region of interest. The required set of computations is sufficiently simple that this method could also be greatly accelerated using GPUs (graphics processing units). One could also take advantage of the relative smoothness and continuity of reciprocal space, in both the Q_x and Q_z directions, by first performing a coarse-grained reconstruction, and then using this as an initial guess to a finer-grained reconstruction.

A key limitation of the method we have proposed is that one must have reasonable estimates for the film makeup. While some values (especially θ_c) can be easily measured from the input GISAXS data themselves, other parameters (the reflectivity curve, $|R|$) may not be easy to estimate. In the general case, a material of interest can have an arbitrary density profile in the film normal direction, with a correspondingly complex transmission and reflectivity curve. In such cases, the reconstruction will be valid only if one accurately models the layer profile. On the other hand, a great many experiments are performed on thin films that are relatively uniform. If the scattering contrast from nanostructures within the film is much smaller than the scattering contrast between the film and the ambient medium, then the reflectivity curve will be dominated by the thickness and average density of the film. In other words, for a great many GISAXS experiments, an adequate estimate for the reflectivity curve can be obtained by simply knowing the film thickness and average composition.

Although our present implementation is highly successful in reconstruction of an estimate for the undistorted scattering, there are many improved variants one could consider. As previously noted, a straightforward extension of the present method would be to include other parameters in the iterative reconstruction. Our results demonstrate that θ_i and θ_c can be robustly fitted as part of the reconstruction. Some aspects of the reflectivity curve (especially the film thickness) could also be fitted as part of the χ_d^2 minimization. A more challenging case would be to attempt to fit the entire X-ray reflectivity curve during reconstruction; or, equivalently, to develop a model of the film's density profile. In principle, such information is encoded within the GISAXS image, and iterative reconstruction could simultaneously extract estimates for $I_R(Q_z)$ and $|R(q_z)|$ that match the GISAXS data. In our current implementation of the reconstruction, we take advantage of several approximations, which simplify the formalism and reduce computation time. However, this limits the generality of the method. In particular, we expect the method to fail when materials have layer profiles that deviate strongly from the uniform-film approximation, and when the nanostructures within the film are near-perfectly ordered over large in-plane distances. An avenue for future improvement of our method is to be able to handle these cases. To correctly reconstruct materials with complex layer profiles, one would need to use DWBA variants that include graded or multi-layered profiles for the film (Boer, 1996; Sentenac & Greffet, 1998; Lazzari et al., 2007; Renaud et al., 2009). This allows one to account for the intensity of transmitted and reflected components throughout the entirety of the film depth; however, in such a case one would need to have an explicit model for the real-space structure of the sample in order to account for which features in reciprocal space arise from which depths within the film. Similarly, an improved reconstruction could be achieved if one included the contribution from the cross terms of equation (8). Correct calculation of these terms requires knowledge of the complex amplitude $F(Q_z)$, which could also be carried out using an explicit real-space model.

4. Summary

We have demonstrated a method for unwarping GISAXS data, by iteratively reconstructing the true (undistorted) reciprocal-space scattering (i.e. a SAXS-like scattering pattern) such that it is consistent with the experimental GISAXS. We simultaneously fitted multiple GISAXS images to improve robustness, and filled any regions of reciprocal space not covered by a single image. We exploited a simplified version of the DWBA formalism to compute GISAXS images from reciprocal space. We found that despite several simplifications, and the use of assumed forms of the transmission and reflectivity curves, we are able to reconstruct reasonable estimates of undistorted scattering for a wide variety of sample

structures. Overall, this unwarping method should prove to be a useful and generic method for inspecting GISAXS data, allowing experimenters to view data unencumbered by refraction distortion and multiple scattering, and in some cases yield reconstructed data of sufficient quality for rigorous analysis and fitting.

References

Altarelli, M., Kurta, R. P. & Vartanyants, I. A. (2010). *Phys. Rev. B*, **82**, 104207.

Boer, D. K. G. de (1996). *Phys. Rev. B*, **53**, 6048–6064.

Breiby, D. W., Bunk, O., Andreasen, J. W., Lemke, H. T. & Nielsen, M. M. (2008). *J. Appl. Cryst.* **41**, 262–271.

Busch, P., Rauscher, M., Smilgies, D.-M., Posselt, D. & Papadakis, C. M. (2006). *J. Appl. Cryst.* **39**, 433–442.

Byrd, R., Lu, P., Nocedal, J. & Zhu, C. (1995). *SIAM J. Sci. Comput.* **16**, 1190–1208.

Chen, G., Modestino, M. A., Poon, B. K., Schirotzek, A., Marchesini, S., Segalman, R. A., Hexemer, A. & Zwart, P. H. (2012). *J. Synchrotron Rad.* **19**, 695–700.

Choo, Y., Majewski, P. W., Fukuto, M., Osuji, C. O. & Yager, K. G. (2017). *Nanoscale*, **10**, 416–427.

Conn, A., Gould, N. & Toint, P. (2000). *Trust Region Methods*. MOS SIAM Series on Optimization. Philadelphia: Society for Industrial and Applied Mathematics.

Cristofolini, L. (2014). *Curr. Opin. Colloid Interface Sci.* **19**, 228–241.

Croset, B. (2017). *J. Appl. Cryst.* **50**, 1245–1255.

Doerk, G. S. & Yager, K. G. (2017). *Mol. Syst. Des. Eng.* **2**, 518–538.

Förster, S., Apostol, L. & Bras, W. (2010). *J. Appl. Cryst.* **43**, 639–646.

Förster, S., Timmann, A., Konrad, M., Schellbach, C., Meyer, A., Funari, S. S., Mulvaney, P. & Knott, R. (2005). *J. Phys. Chem. B*, **109**, 1347–1360.

Gommes, C. J. (2016). *J. Appl. Cryst.* **49**, 1162–1176.

Heitsch, A. T., Patel, R. N., Goodfellow, B. W., Smilgies, D.-M. & Korgel, B. A. (2010). *J. Phys. Chem. C*, **114**, 14427–14432.

Hexemer, A. & Müller-Buschbaum, P. (2015). *IUCrJ*, **2**, 106–125.

Hu, T., Jones, R. L., Wu, W., Lin, E. K., Lin, Q., Keane, D., Weigand, S. & Quintana, J. (2004). *J. Appl. Phys.* **96**, 1983–1987.

Hunter, J. D. (2007). *Comput. Sci. Eng.* **9**, 90–95.

Jiang, Z., Lee, D. R., Narayanan, S., Wang, J. & Sinha, S. K. (2011). *Phys. Rev. B*, **84**, 075440.

Jiang, Z., Li, X., Strzalka, J., Sprung, M., Sun, T., Sandy, A. R., Narayanan, S., Lee, D. R. & Wang, J. (2012). *J. Synchrotron Rad.* **19**, 627–636.

Johnston, D. E., Yager, K. G., Hlaing, H., Lu, X., Ocko, B. M. & Black, C. T. (2014). *ACS Nano*, **8**, 243–249.

Jones, R. L., Hu, T., Lin, E. K., Wu, W.-L., Kolb, R., Casa, D. M., Bolton, P. J. & Barclay, G. G. (2003). *Appl. Phys. Lett.* **83**, 4059–4061.

Jones, R. L., Hu, T., Soles, C. L., Lin, E. K., Reano, R. M., Pang, S. W. & Casa, D. M. (2006). *Nano Lett.* **6**, 1723–1728.

Jones, R. L., Soles, C. L., Lin, E. K., Hu, W., Reano, R. M., Pang, S. W., Weigand, S. J., Keane, D. T. & Quintana, J. P. (2006). *J. Micro. Nanolithogr. MEMS MOEMS*, **5**, 013001.

Khaira, G., Doxastakis, M., Bowen, A., Ren, J., Suh, H. S., Segal-Peretz, T., Chen, X., Zhou, C., Hannon, A. F., Ferrier, N. J., Vishwanath, V., Sunday, D. F., Gronheid, R., Kline, R. J., de Pablo, J. J. & Nealey, P. F. (2017). *Macromolecules*, **50**, 7783–7793.

Kiapour, M. H., Yager, K., Berg, A. C. & Berg, T. L. (2014). *IEEE Winter Conference on Applications of Computer Vision*. pp. 933–940. Steamboat Springs, CO: IEEE.

Kline, S. R. (2006). *J. Appl. Cryst.* **39**, 895–900.

Lazzari, R. (2002). *J. Appl. Cryst.* **35**, 406–421.

Lazzari, R., Leroy, F. & Renaud, G. (2007). *Phys. Rev. B*, **76**, 125411.

Lee, D. R., Hagman, A., Li, X., Narayanan, S., Wang, J. & Shull, K. R. (2006). *Appl. Phys. Lett.* **88**, 153101.

Lee, B., Park, Y.-H., Hwang, Y.-T., Oh, W., Yoon, J. & Ree, M. (2005). *Nat. Mater.* **4**, 147–151.

Lee, B., Park, I., Yoon, J., Park, S., Kim, J., Kim, K.-W., Chang, T. & Ree, M. (2005). *Macromolecules*, **38**, 4311–4323.

Lee, B., Yoon, J., Oh, W., Hwang, Y.-T., Heo, K., Jin, K. S., Kim, J., Kim, K.-W. & Ree, M. (2005). *Macromolecules*, **38**, 3395–3405.

Lehmkühler, F., Grübel, G. & Gutt, C. (2014). *J. Appl. Cryst.* **47**, 1315–1323.

Lehmkühler, F., Schulz, F., Schroer, M. A., Frenzel, L., Lange, H. & Grübel, G. (2018). *IUCrJ*, **5**, 354–360.

Lhermitte, J. R., Tian, C., Stein, A., Rahman, A., Zhang, Y., Wiegart, L., Fluerasu, A., Gang, O. & Yager, K. G. (2017). *J. Appl. Cryst.* **50**, 805–819.

Li, X., Shew, C.-Y., He, L., Meilleur, F., Myles, D. A. A., Liu, E., Zhang, Y., Smith, G. S., Herwig, K. W., Pynn, R. & Chen, W.-R. (2011). *J. Appl. Cryst.* **44**, 545–557.

Liu, J., Lhermitte, J., Tian, Y., Zhang, Z., Yu, D. & Yager, K. G. (2017). *IUCrJ*, **4**, 455–465.

Lu, X., Yager, K. G., Johnston, D., Black, C. T. & Ocko, B. M. (2013). *J. Appl. Cryst.* **46**, 165–172.

Luttge, R. (2009). *J. Phys. D Appl. Phys.* **42**, 123001.

Mahadevapuram, N., Strzalka, J. & Stein, G. E. (2013). *J. Polym. Sci. B Polym. Phys.* **51**, 602–610.

Majewski, P. W., Rahman, A., Black, C. T. & Yager, K. G. (2015). *Nat. Commun.* **6**, 7448.

Majewski, P. W. & Yager, K. G. (2015a). *Macromolecules*, **48**, 4591–4598.

Majewski, P. W. & Yager, K. G. (2015b). *ACS Nano*, **9**, 3896–3906.

Malmerberg, E., Kerfeld, C. A. & Zwart, P. H. (2015). *IUCrJ*, **2**, 309–316.

Martin, A. V. (2017). *IUCrJ*, **4**, 24–36.

Meister, N., Guan, Z., Wang, J., Lashley, R., Liu, J., Lhermitte, J., Yager, K. G., Qin, H., Sun, B. & Yu, D. (2017). *New York Scientific Data Summit*, pp. 1–6. Steamboat Springs, CO: IEEE.

Millman, K. J. & Aivazis, M. (2011). *Comput. Sci. Eng.* **13**, 9–12.

Morales, J. L. & Nocedal, J. (2011). *ACM Trans. Math. Softw.* **38**, 7:1–7:4.

Müller-Buschbaum, P. (2003). *Anal. Bioanal. Chem.* **376**, 3–10.

Müller-Buschbaum, P. (2016). *Eur. Polym. J.* **81**, 470–493.

Narayanan, S., Lee, D. R., Guico, R. S., Sinha, S. K. & Wang, J. (2005). *Phys. Rev. Lett.* **94**, 145504.

Oliphant, T. E. (2007). *Comput. Sci. Eng.* **9**, 10–20.

Omote, K., Ito, Y. & Kawamura, S. (2003). *Appl. Phys. Lett.* **82**, 544–546.

Pedersen, J. S. (1997). *Adv. Colloid Interface Sci.* **70**, 171–210.

Posselt, D., Smilgies, D.-M., Busch, P. & Papadakis, C. M. (2002). *Synchrotron Rad. News*, **15**, 35–42.

Rahman, A., Majewski, P. W., Doerk, G., Black, C. T. & Yager, K. G. (2016). *Nat. Commun.* **7**, 13988.

Rauscher, M., Paniago, R., Metzger, H., Kovats, Z., Domke, J., Peisl, J., Pfannes, H.-D., Schulze, J. & Eisele, I. (1999). *J. Appl. Phys.* **86**, 6763–6769.

Renaud, G., Lazzari, R. & Leroy, F. (2009). *Surf. Sci. Rep.* **64**, 255–380.

research papers

Senesi, A. J. & Lee, B. (2015). *J. Appl. Cryst.* **48**, 1172–1182.

Sentenac, A. & Greffet, J.-J. (1998). *JOSA A*, **15**, 528–532.

Settens, C. M., Cordes, A., Bunday, B. D., Bello, A. F., Kamineni, V. K., Paul, A., Fronheiser, J. & Matyi, R. J. (2014). *J. Micro. Nanolithogr. MEMS MOEMS*, **13**, 041408.

Sinha, S. K., Sirota, E. B., Garoff, S. & Stanley, H. B. (1988). *Phys. Rev. B*, **38**, 2297–2311.

Smilgies, D.-M. (2009). *J. Appl. Cryst.* **42**, 1030–1034.

Smilgies, D.-M., Heitsch, A. T. & Korgel, B. A. (2012). *J. Phys. Chem. B*, **116**, 6017–6026.

Sunday, D. F., Hammond, M. R., Wang, C., Wu, W.-L., Delongchamp, D. M., Tjio, M., Cheng, J. Y., Pitera, J. W. & Kline, R. J. (2014). *ACS Nano*, **8**, 8426–8437.

Sunday, D. F., Hammond, M. R., Wang, C., Wu, W.-L., Kline, R. J. & Stein, G. E. (2013). *J. Micro. Nanolithogr. MEMS MOEMS*, **12**, 031103.

Sunday, D. F., List, S., Chawla, J. S. & Kline, R. J. (2015). *J. Appl. Cryst.* **48**, 1355–1363.

Székely, P., Ginsburg, A., Ben-Nun, T. & Raviv, U. (2010). *Langmuir*, **26**, 13110–13129.

Tate, M. P., Urade, V. N., Kowalski, J. D., Wei, T.-C., Hamilton, B. D., Eggiman, B. W. & Hillhouse, H. W. (2006). *J. Phys. Chem. B*, **110**, 9882–9892.

Toney, M. F. & Brennan, S. (1989). *Phys. Rev. B*, **39**, 7963–7966.

Vineyard, G. H. (1982). *Phys. Rev. B*, **26**, 4146–4159.

Walt, S. van der, Colbert, S. C. & Varoquaux, G. (2011). *Comput. Sci. Eng.* **13**, 22–30.

Wang, B., Guan, Z., Yao, S., Qin, H., Nguyen, M. H., Yager, K. G. & Yu, D. (2016). *New York Scientific Data Summit*, pp. 1–5. Steamboat Springs, CO: IEEE.

Wang, C., Jones, R. L., Lin, E. K., Wu, W.-L., Rice, B. J., Choi, K.-W., Thompson, G., Weigand, S. J. & Keane, D. T. (2007). *J. Appl. Phys.* **102**, 024901.

Wang, B., Yager, K. G., Yu, D. & Hoai, M. (2017). *IEEE Winter Conference on Applications of Computer Vision*, pp. 697–704. Steamboat Springs, CO: IEEE.

Weidman, M. C., Yager, K. G. & Tisdale, W. A. (2015). *Chem Mater.* **27**, 474–482.

Wochner, P., Gutt, C., Autenrieth, T., Demmer, T., Bugaev, V., Ortiz, A. D., Duri, A., Zontone, F., Grübel, G. & Dosch, H. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 11511–11514.

Wu, W. (1993). *J. Chem. Phys.* **98**, 1687–1694.

Wu, W. (1994). *J. Chem. Phys.* **101**, 4198–4204.

Yager, K. G., Lhermitte, J., Liu, J., Yu, D., Wang, B. & Guan, Z. (2017). *Dataset of synthetic X-ray scattering images for classification using deep learning*, http://dx.doi.org/doi:10.18126/M2Z30Z.

Yager, K. G. & Majewski, P. W. (2014). *J. Appl. Cryst.* **47**, 1855–1865.

Yager, K. G., Zhang, Y., Lu, F. & Gang, O. (2014). *J. Appl. Cryst.* **47**, 118–129.

Yoneda, Y. (1963). *Phys. Rev.* **131**, 2010–2013.

Zhong, W., Xu, W., Yager, K. G., Doerk, G. S., Zhao, J., Tian, Y., Ha, S., Xie, C., Zhong, Y., Mueller, K. & Van Dam, K. K. (2018). *Vis. Inf.* **2**, 14–25.

Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. (1997). *ACM Trans. Math. Softw.* **23**, 550–560.