Analysis and Modeling of Anomalous Scattering in the AXAF HETGS

John E. Davis Herman L. Marshall Daniel Dewey
Mark Schattenburg

Center for Space Research
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Tests of the High-Energy Transmission Grating Spectrometer (HETGS) for the Advanced X-ray Astrophysics Facility (AXAF) showed anomalous scattering of monochromatic radiation. The grating resolving power ($E/\mathrm{d}E$) is of order 1000, but tests designed to search for small-angle scattering by the gratings showed events with significant deviations from the dispersed grating orders and concentrated along the dispersion direction.

In this paper, we present a general overview of grating scatter as a result of fluctuations in the grating bar geometry. The grating scatter observed at the AXAF X-Ray Calibration Facility (XRCF) is shown to be consistent with what one expects from scatter due to deviations in the grating bar geometry from perfect bars. For the purposes of modeling, a rectangular bar model is adopted and bar parameters are deduced via fitting the model to the scattering data. The correlations deduced from this model lead to a simple physical picture of grating bar fluctuations where a small fraction of the bars, e.g., 1 in 200 are leaning.

Keywords: AXAF, X-ray, scatter, grating, diffraction, probability, correlations

1. INTRODUCTION

Tests of the HETGS gratings at the AXAF X-Ray Calibration Facility (XRCF) indicated that there was detectable scattering along the High Energy Grating (HEG) dispersion direction.[1] A test of the monochromator scattering indicated that there was no detectable spectral leakage. In this work, we show that the anomalous scattering is consistent with scattering from an imperfect grating.

After a presentation of the scattering measurement data, we treat the problem of scattering from a disordered diffraction grating, and from first principles arrive at an equation for the probability distribution of scatter from such a grating. Without making any reference to a specific geometric model for the grating bar shapes, we show that the anomalous scattering observed at XRCF (e.g., see Figure 1) shares many of the features that one would expect to be produced as a result of grating disorder. In particular, the strong peaks that appear symmetrically about the various diffraction orders can be explained as a natural consequence of one or more correlations between neighboring bars with a correlation length that is characterized by the width of the peaks. The correlation length is estimated to be roughly three grating periods.

A more detailed description of the correlations based on a rectangular model for the bar shapes is presented in the section 4. Although the rectangular model has been shown to be inadequate for modeling the grating efficiency to the one percent calibration goal,[3] it performs well at the ten to twenty percent level. Its most redeeming feature is its simplicity, which allows many of the calculations to be performed analytically. Within the context of this model, expressions for the correlations and the scattering probability are derived and then fitted to the experimental data. The resulting fit, while not perfect, does reflect many of the salient features of the data, although the resulting effective rectangular parameters are somewhat larger than expected from modeling[2] of the grating efficiencies.

In the final section we present a discussion of the correlations as deduced from the modeling. Based upon the form of the correlations, a physical picture of the gratings emerge in which a small fraction (e.g., 1 in every 200) of the grating bars are leaning. Finally, the major results of this work are summarized future modeling improvements are suggested.

Other author information: Send correspondence to J.E.D.: Email: davis@space.mit.edu; Telephone: 617-258-8119; Fax: 617-253-8084

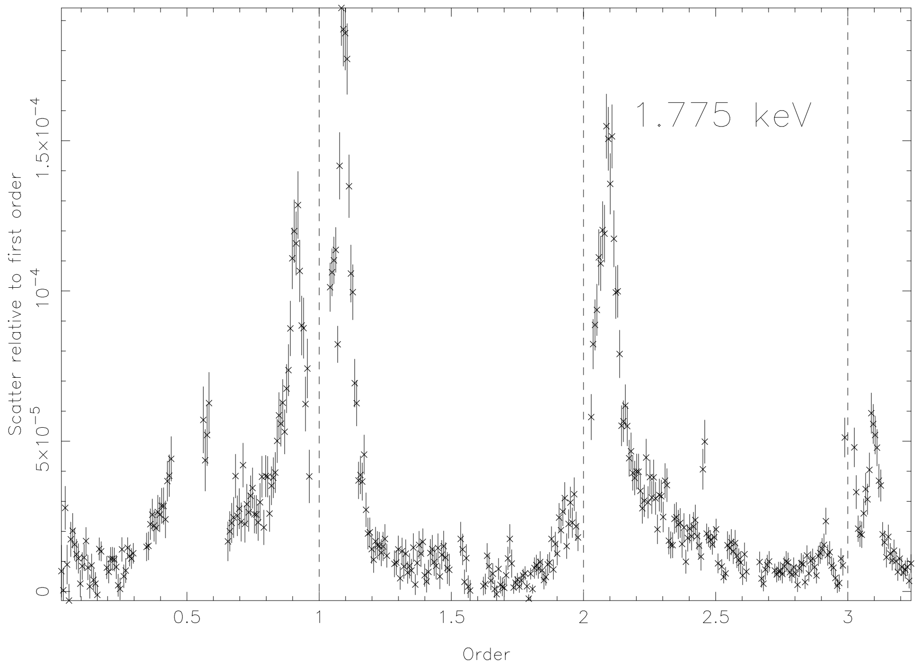

Figure 1. This figure shows the scattering-per-bin as fraction of the counts expected in the first order with a bin size of 0.00716 orders. Note the strong scattering peaks immediately to the left and right of first order, as well as the peaks near second and third order.

Test ID	Detector	Energy(ies) (keV)
G-H2C-SC-88.00{1-5}	ACIS-2C0	1.7754
H-HAS-SC-7.002	ACIS-S	1.3835
H-HAS-SC-7.004	ACIS-S	2.035
H-HAS-SC-17.006	ACIS-S	5.00, 6.00, 7.00, 8.00, 9.00

Table 1. This table shows the set of tests used in the data analysis. For an overview of AXAF calibration, see Weisskopf et al.[4]

2. OBSERVATIONS AND DATA ANALYSIS

Several calibration tests were designed to probe the pattern of the scattering and its dependence on energy. The tests were performed at the AXAF X-ray Calibration Facility (XRCF).[4] The double crystal monochromator (DCM) source was used to reduce spectral contributions far away from the central energy. For three tests, the DCM was tuned to an W-M emission line from the source. For the remaining tests, the DCM was scanned from 5 keV to 9 keV in 1 keV steps for 5 different data sets. The list of tests is shown in Table 1.

Figure 1 shows the scattering features observed at 1.775 keV, while the fractional scattered light integrated over each of many order regions is shown graphically in Figure 2.

2.1. Tests G-H2C-SC-88.00{1-5}

The W-Mα line at 1.7754 keV was the brightest line to which the DCM could be tuned. The shape of the scattering was investigated at this energy using a long integration and the highest DCM current. The test was performed using the backside illuminated CCD in the ACIS-2C* (chip 0), which is a non-flight AXAF detector developed specifically for ground testing and calibration. Five separate exposures of about 900 s each were obtained, shifting by 22 mm between exposures to span a dispersion distance similar to that of the flight detector. The CCD is about 24 mm across, so there was a 2 mm overlap between adjacent exposures. The pixel spacing of the detector was 0.024 mm,

*See `http://space.mit.edu/HETG/acis2c.html` for more details of the ACIS-2C detector

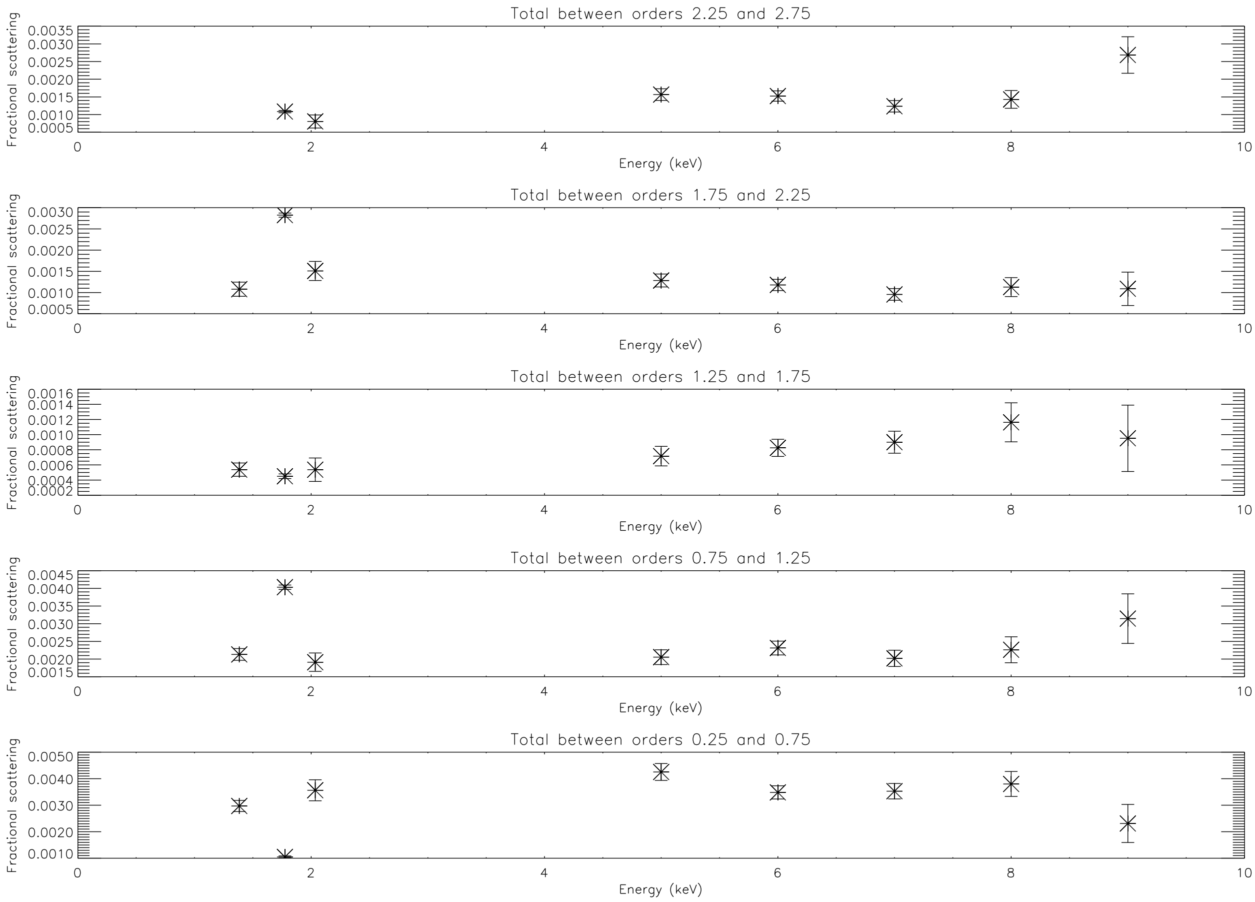

Figure 2. This figure shows a summary of the results of the data analysis. The values are normalized to the total of the +1 and -1 order count rates. Due to the systematic errors in background subtraction, bad columns, etc., the fractional scattering at 0.5 order at 1.775 keV has been systematically underestimated – by as much as a factor of 2 or 3. Error estimates are due to statistical uncertainties only. Note that most scattering fractions are relatively independent of energy above 2.0 keV.

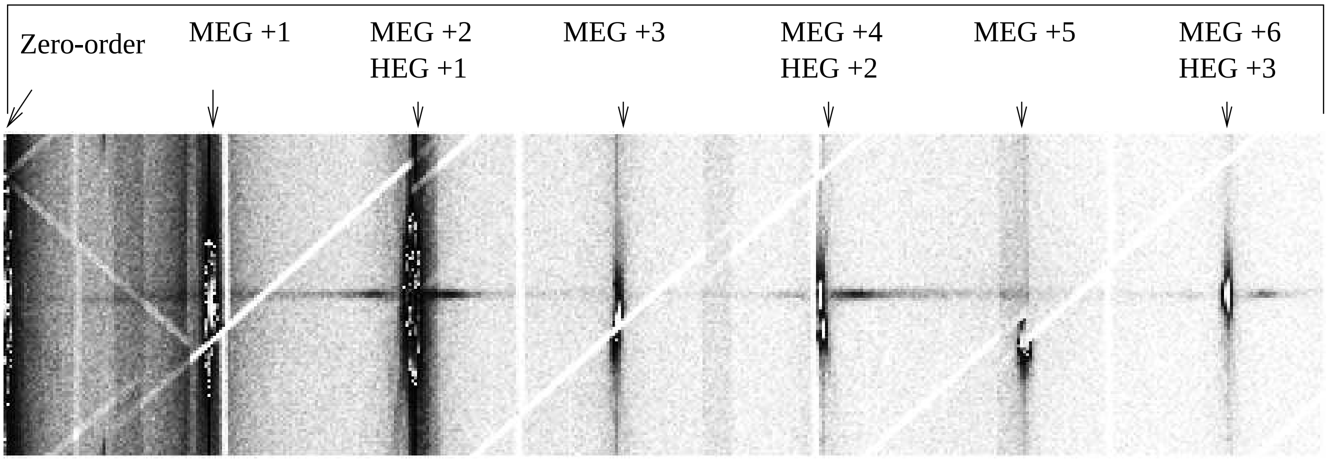

Figure 3. This image is the result of binning data from HETGS scattering tests G-H2C-SC-88.00{1-5} along the HEG dispersion direction. The anomalous scattering is apparent as the broad, dark horizontal streaks in the image. Wavelength increases to the right from the zeroth order at far left. Three HEG orders are observed as well as MEG orders in between. Note the near coincidence of MEG 4th order and HEG 2nd order about 2/3 of the way from the left end. The diagonal features are due to 2 pixel gaps in the event finding algorithm while vertical features are due to bad pixels masked out or to changes in exposure due to regions of overlap between shifted exposures. Scattering is brightest near the HEG +1 and HEG +2 orders.

which was precisely measured during fabrication (M. Bautz, private communication). The CCD was operated in a "fast" mode where 114 rows of data were obtained every 0.66 s, in order to reduce pile-up near the dispersed orders. The frames are shifted downward by 114 rows after each exposure, so the dispersion lines are "wrapped". Event recognition was disabled for the final two rows (wrapped rows were not considered), so there are no events in these two rows (producing a gap that occasionally crosses the dispersion line).

The procedure for reducing these observations started with event lists which included the event location in the 114x1024 window, the event time, and pulse heights for the 5x5 event island. The event positions were randomized within a pixel in order to prevent digitization effects in subsequent analysis, which involved fractional pixel shifts and rotations. The locations were shifted to place the zero point at the 0th order location in facility coordinates and rotated so that the MEG or HEG dispersion direction would be horizontal. Events for each exposure were processed separately due to detector rotations and shifts. A detector-based energy was computed using the 25 pulse heights and the detector gain, 0.322 eV per DN. In order to reduce background, data were screened to include only events with inferred energies in the 1.3 to 4.4 keV range, which gave events with one or two photons each.

The wavelengths for each event, assuming it was dispersed into first order, were computed using the observed dispersion distance for each event, nominal grating periods, and the adopted Rowland distance: 4001.13 Å (MEG), 2000.81 Å (HEG), and 8782.8 mm. The events were "unwrapped" modulo the 114 row period in order to follow the dispersion and then combined and binned in 0.05 Å (HEG) or 0.1 Å intervals (MEG). The images are shown in figures 3 and 4. Background was estimated using a 30 pixel region well away from the dispersed light. The dominant source of background was due to mirror scattering, which is azimuthally symmetric about the 0th order image and all dispersed orders. Many artifacts of the detector are visible but the anomalous scattering is obvious along the HEG direction and is absent along the MEG direction.

The exposure time (see Figure 5) was computed for each wavelength based on the wavelengths of the detector edges for each exposure, and was used to compute count rates as a function of wavelength. The result was normalized to the estimated count rate of the first dispersed order, which was not directly observed due to significant event pile-up. The zeroth order count rate was estimated by modeling the profile of the trail of the zeroth order image that occurs when the CCD is parallel shifted. The trail receives an exposure of 40 μ sec per row for each CCD frame and for each window per frame. There were 9 windows (at 114 rows per 1024 row frame) per frame, 100 frames of data (only the first portion was examined) and 25 rows of the 114 row window were combined, giving a total exposure in the trail of 0.9 sec. The one-dimensional profile of the trail region is shown in Figure 6. Fitting two Gaussians to the trail gave about 1200 $\pm$ 100 counts in each component while a detailed model based on the beam normalization

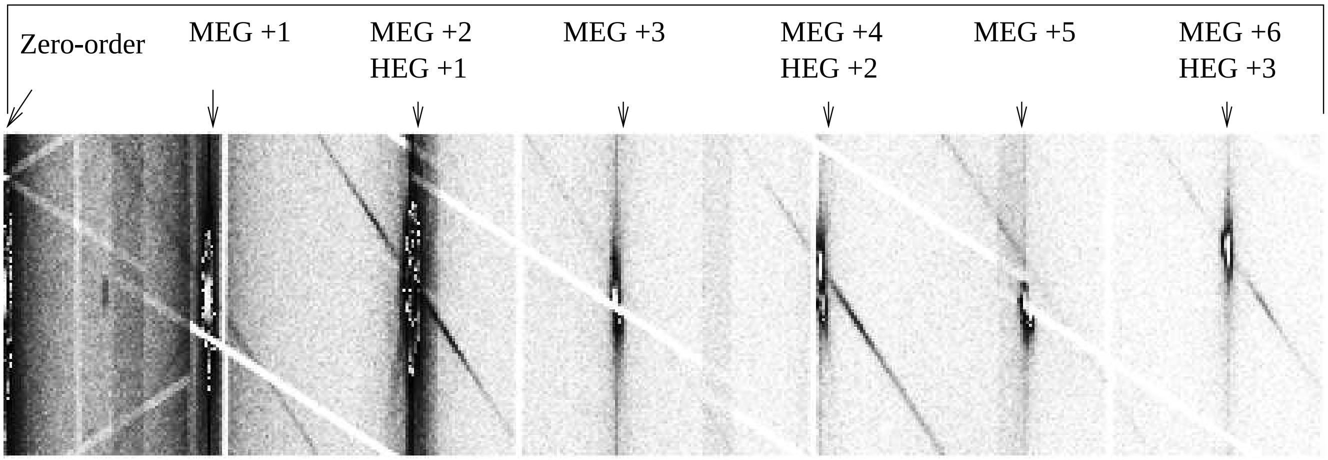

Figure 4. An image derived from tests G-H2C-SC-88.00{1-5} as in Fig. 3 but binned along the MEG dispersion direction and binned at 0.1 Å bins (so the image size is about the same as the HEG image). The same detector features are apparent, such as stripes where there are no events. The HEG scattering shows up as steep, wrapped, dark, diagonal lines. No MEG scattering is observed, so stringent upper limits can be set. The faint elliptical image between zeroth and first order is the first order dispersed image of the DCM output at twice the tuned energy.

detector (BND) data only predicted 1170 in the broad, flat component and 1280 in the Gaussian component. The broad component was modeled as an intensity gradient that varied by a factor of two (as given by the BNDs) across the HRMA aperture and then adjusting the defocus and centroid until a good match was obtained to the observed profile. The normalization was not adjusted but was fixed by the north and south BND count rates.† Fitting a Gaussian after subtracting the model of the flat component gave 1180 ± 80 count, so the model agrees well with the data. The first order count rate, $R_{1,HEG}$, was then estimated as $R_{0,HETGS}A_{1,HEG}/(A_{0,HEG}+A_{0,MEG})$, where A_m is the predicted effective area at 1.7754 keV for order m and the grating subset (either HEG or MEG). The model predicts a count rate for HEG 1st order (both sides) of 2980 count/s and 5420 count/s in the MEG 1st order.

Detector and grating order artifacts were eliminated after close examination of the raw data. MEG orders 1, 3, and 5 caused systematic errors at 3.5, 10.5, and 17.5 Å while HEG orders 1, 2, and 3 required excising data at 7, 14, and 21 Å. The excised regions were generally 0.3 Å wide and no scattering due to the grating was apparent in these regions. Bad exposure corrections due to the 2 row gap were eliminated near 4.3, 11.2, 2.2, and 18.4 Å while bad columns in the detector required eliminating data near 18.95, 1.30, and 8.85 Å. The result is shown in Figure 1. The total scattered light from 0th order through 3rd order is 1.00±0.01 % of HEG first orders based on this data set.

A similar analysis of the MEG data shows no significant detections of scattered light. There are systematic background subtraction errors as the HEG scattered light contributed to the background. Nevertheless, the MEG scattered light density never appears to exceed a value of 5×10^{-6} per 0.014 order. The total scattering is difficult to estimate given the systematic errors but appears to be less than 5×10^{-5} between orders 1 and 3 while the comparable HEG scattering is 6×10^{-3}.

2.2. Tests H-HAS-SC-7.002, 7.004, and 17.006

The ACIS-S was used in phase H for these scattering tests, so there were no problems associated with windows or shifting of the detector. In the first test, the DCM was tuned to the W M_ζ 1.384 keV line. In the second test, the DCM was tuned to the W M_γ 2.035 keV line. For the third test, the DCM was scanned from 5.00 keV to 9.00 keV in 1.00 keV steps. The HETGS dispersion relation constants derived from phase C data were verified to a part in 2400 using the first data set. The data reduction steps were essentially identical for these tests except

†See the AXAF Project Science web page `http://wwwastro.msfc.nasa.gov/xray/xraycal/bu/bucat.html` and the figure linked from "TAP2", tap2.eps. RunID 112250 shows a good example of the profile of the DCM beam when tuned to a bright line. Additional data taken during the XRCF rehearsal phase with a BND centered in the XRCF vacuum pipe were used to chart the 1.775 keV line and the nearby continuum. Note that the north and south BNDs are placed about 70 cm to either side of the HRMA aperture center while the top and bottom BNDs are centered in XRCF Y coordinates.

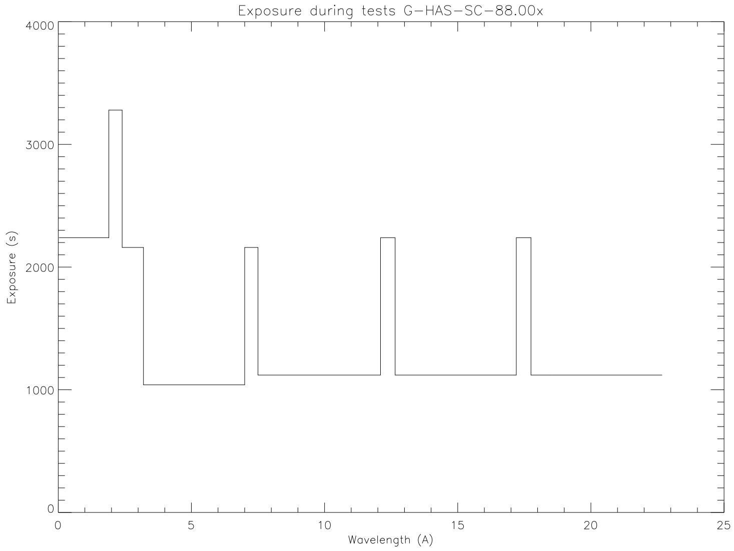

Figure 5. This figure shows the exposure function derived for tests G-H2C-SC-88.00{1-5}. Regions where two adjacent observations overlapped show twice the average exposure and there is double exposure near zeroth order as data from the negative dispersion direction is included.

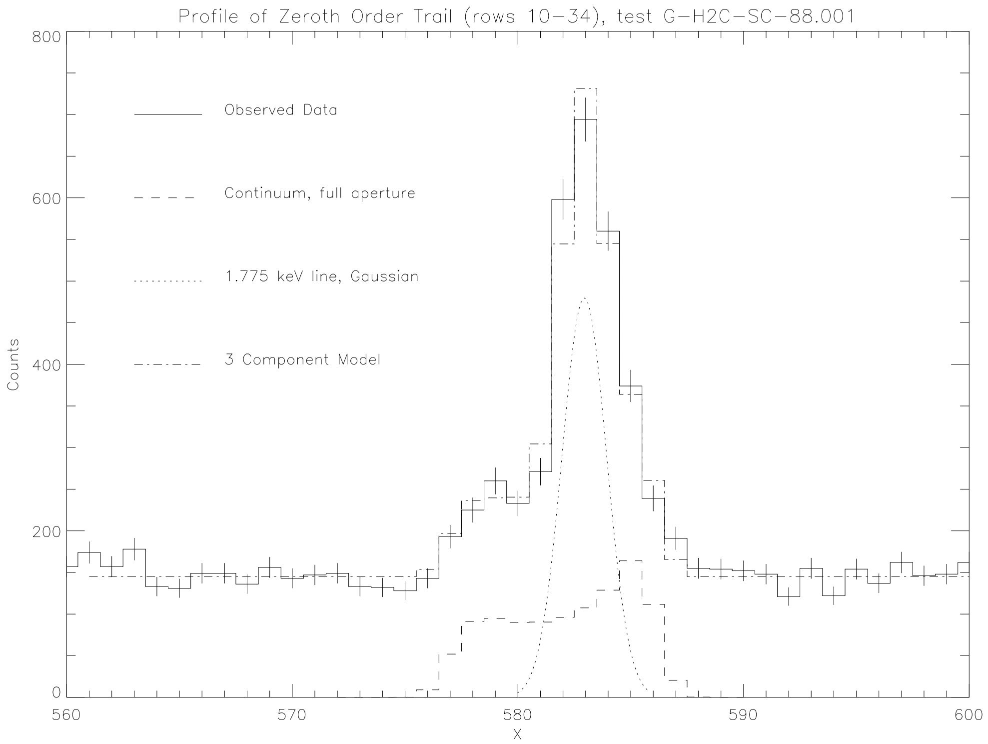

Figure 6. This figure shows the profile of the zeroth order image in test G-H2C-SC-88.001. The profile was obtained by summing 25 rows far from the image centroid so that pile-up would not be a concern. The profile was fitted with three components: 1) *dashed line* the pedestal of the profile is dominated by off-line continuum as the beam pattern does not fill the mirror aperture and the detector was (unintentionally) placed slightly out of focus; 2) *dotted line* a Gaussian which approximates the pattern produced by the DCM when tuned to a source emission line; and 3) the uniform background is primarily due to mirror scattering.

that for H-HAS-SC-17.006 time selections were needed to separate the periods when the DCM energy was constant. In addition, for the last test only, the five axis mounts (FAM) that moves the ACIS-S relative to the HRMA was positioned at Y=0.25 mm in XRCF coordinates.

The ACIS-S was run in the timed event mode so there was no ambiguity between MEG and HEG. The DCM current was set at the highest possible value, which caused significant pile-up in the dispersed orders and made zeroth order readout trail bright enough to measure in order to estimate the beam flux. The trail was somewhat fainter than in the tests taken at 1.775 keV, so total counts in a 3 column region starting 200 rows on either side of the zeroth order image were compared to a similar region rotated 15° from the vertical. The exposure time in the zeroth order trail was 40μs/row/frame $\times(1024 - 2 \cdot 200)$ rows, or 25ms per frame. Events were selected so that the energy derived from the ACIS pulse height was within 100 eV of the expected value. The event positions were transformed using the grating constants into wavelengths and a simple extraction was performed using a window 0.2 mm wide around the dispersed events while background was measured in a region 0.2 to 0.8 mm from the dispersion line. An additional rotation of -0.13° was required to make more horizontal dispersion lines; this rotation could be either in the ACIS to XRCF coordinate frame or could be a residual error in measuring the HETG to XRCF coordinate frame. Data near 0th order were eliminated from the final results as background measurements failed to account for the azimuthal symmetry of the 0th order (mirror) scattered light. Data near orders were also eliminated in general. The results are shown in figures 9 and 10.

3. THEORETICAL OVERVIEW

3.1. General Formulation

Consider a diffraction grating composed of N lines, or bars, with period d, and a photon with wave number $k = 2\pi/\lambda$ incident upon the grating. It follows from Huygen's principle that the probability for scattering into an angle between θ and $\theta + \mathrm{d}\theta$ with respect to the normal of the grating is given by

$$p(k,s)\mathrm{d}s = \frac{kd}{2\pi N}\left|\sum_{j=0}^{N-1} e^{ijksd} F_j(k,s)\right|^2 \mathrm{d}s\ , \tag{1}$$

where $s = \sin\theta$ and the *structure factor* $F_j(k,s)$ is defined by

$$F_j(k,s) = \frac{1}{d}\int_0^d \mathrm{d}x\ e^{iksx+i\phi_j(k,x)}. \tag{2}$$

The shape and composition of an individual grating bar determines the complex phase shift $\phi_j(k,x)$ that, in general, depends upon energy and, possibly, the polarization of the photon. Assuming the validity of scalar diffraction theory, and neglecting reflection and refraction, the phase shift can be written in the form

$$\phi_j(k,x) = -k(\delta - i\beta)z_j(x), \tag{3}$$

where δ and β are energy dependent functions related to the dielectric constant of the grating bars. The function $z_j(x)$ represents the path length of the photon as it passes though the jth grating bar; it is sometimes called, rather loosely, the "grating bar shape", and more rigorously, the "path-length function".

In general, the structure factor $F_j(k,s)$ will vary with the grating bars because the "bar shape" will vary from bar to bar. Since an exact computation of the scattering probability would require complete knowledge of all the bar shapes, and since there are millions of bars that contribute to the scattering, we must resort to a statistical means of computing Eq. 1. This is completely analogous to the situation in statistical physics where one deals with systems that consist of large numbers of degrees of freedom, and whose complete characterization requires a knowledge of the initial conditions for each degree of freedom. These problems are treated statistically where one introduces the notion of an ensemble of macroscopically identical systems and makes an ergodic hypothesis that at any given instant the microscopic state of the system is represented by one of the systems in the ensemble. Hence the assumption of ergodicity allows one to compute the time average of a macroscopic quantity by averaging that quantity over the ensemble. In a similar vein we shall introduce the idea of an ensemble of gratings and make the assertion that the observed scattering does not depend upon the exact "microscopic" state of the grating, but that it is an indicator of the "macroscopic" state. In other words, we shall assume that the scattering is a macroscopic property of the ensemble and may be computed via

$$p(k,s)\mathrm{d}s \; = \frac{kd}{2\pi N}\left\langle \left| \sum_{j=0}^{N-1} e^{ijksd} F_j(k,s) \right|^2 \right\rangle \mathrm{d}s \; , \tag{4}$$

where $\langle\cdot\rangle$ denotes averaging over the ensemble. The specification of the ensemble involves knowing the probability for a particular configuration of grating bars, denoted by the ordered set of bar shapes $\{z_j\}$. Let this probability be represented by the probability density $\mathcal{P}(z_0, z_1, \ldots, z_{N-1})$ such that the ensemble average of some function $f(z_0, \ldots z_{N-1})$ is given by ‡

$$\langle f(z_0,\ldots,z_{N-1})\rangle = \int \mathcal{D}z_0(x)\cdots\mathcal{D}z_{N-1}(x) f(z_0,\ldots,z_{N-1})\mathcal{P}(z_0,\ldots,z_{N-1}). \tag{5}$$

Furthermore, we shall assume that the ensemble is such that $\langle f(z_j)\rangle$ is independent of j, and that $\langle f(z_j, z_l)\rangle$ depends only upon the difference $j - l$. For obvious reasons, a probability distribution with these properties is said to be *homogeneous* or *stationary*.[6]

With the assumption of a homogeneous probability distribution for bar shapes, denote the ensemble averaged structure factor by

$$\bar{F}(k,s) = \langle F_j(k,s)\rangle \tag{6}$$

and define $f_j(k,s)$ to be the deviation of the jth bar from this value via

$$F_j(k,s) = \bar{F}(k,s) + f_j(k,s). \tag{7}$$

Substituting the above definitions into Eq. 4 yields for the probability density

$$p(k,s) = \frac{kd\sin^2(\frac{1}{2}Nksd))}{2\pi N\sin^2(\frac{1}{2}ksd)}|\bar{F}(k,s)|^2 + \frac{kd}{2\pi N}\sum_{j,l} e^{iksd(j-l)}\langle f_j f_l^*\rangle. \tag{8}$$

For very large N, the first term goes to zero like $1/N$ except when ksd is a multiple of 2π where it becomes proportional to N. In fact, one can show that the factor involving N in first term is a representation of a series of delta functions:

$$\frac{kd\sin^2(\frac{1}{2}Nksd)}{2\pi N\sin^2(\frac{1}{2}ksd)} = \sum_n \delta_N(s - \frac{2\pi n}{kd}). \tag{9}$$

This equivalence allows the probability density to be written

$$p(k,s) = |\bar{F}(k,s)|^2 \sum_n \delta_N(s - \frac{2\pi n}{kd}) + \frac{kd}{2\pi N}\sum_{j,l} e^{iksd(j-l)}\langle f_j f_l^*\rangle. \tag{10}$$

Thus, the first term represents the familiar interference pattern that consists of a series of sharp peaks, weighted by $|\bar{F}(k,s)|^2$, whose height varies with N. The peaks occur at values of ksd that are multiples of 2π, or equivalently when

$$n\lambda = d\sin\theta, \tag{11}$$

which is the familiar grating equation. The second term is the scattering term and is more complicated. Superficially, it appears to vanish in the limit of a large number of bars; however, owing to the double sum (hence N^2 terms), this is not necessarily true.

‡Since the bar shape $z_j(x)$ is a function, the integrals in this equation are functional integrals and the symbol $\mathcal{D}z(x)$ is called the *Weiner measure*. See the book by Feynman and Hibbs[5] for a simple introduction to functional integration. Alternatively, one can view such an integral as a multiple integral over the parameters defining a shape, e.g., the bar width and height for a rectangular bar.

The quantities $\langle f_j f_l^* \rangle$ represent correlations between the various grating elements. If the shape of each grating bar element is independent of the others, then $\langle f_j f_l^* \rangle$ vanishes except when $j = l$. More explicitly, if no correlations exist between the fluctuations of the grating bars then

$$\langle f_j f_l^* \rangle = \langle |f_j|^2 \rangle \delta_{jl}. \tag{12}$$

If there are correlations between the various grating bars, then Eq. 12 is not correct. We make the assumption that if any correlations are present, then they have a short range. This is really an assumption about the fabrication method, but one could imagine the presence of long range correlations from e.g., a holographic fabrication process in which a non-monochromatic light source would create a beating effect. Similarly, for a "ruled" grating, the mechanical system could introduce periodic long range correlations. Nevertheless, we shall assume that $\langle f_j f_l^* \rangle$ is zero for $|j - l|$ greater than some number M. For example, if $M = 1$, then only nearest neighbor correlations may be present. If $M \ll N$, then one may neglect edge effects in the sum and write the scattering term in Eq. 10 as

$$\frac{kd}{2\pi N} \sum_{j=0}^{N-1} \sum_{m=-M}^{M} e^{imksd} \langle f_j f_{j+m}^* \rangle. \tag{13}$$

Since we have assumed that the ensemble governing the probability distributions is homogeneous, by definition we know that the correlations are translation invariant, i.e., $\langle f_j f_{j+m}^* \rangle$ does not depend upon j. Hence, the scattering term reduces to

$$\frac{kd}{2\pi} \sum_{m=-M}^{M} e^{imksd} c_m(k, s), \tag{14}$$

where c_m has been defined by

$$c_m(k, s) = \langle f_j(k, s) f_{j+m}^*(k, s) \rangle. \tag{15}$$

Using the property that $c_{-m} = c_m^*$, which follows from the definition of c_m, and separating out real and imaginary parts, it follows that Eq. 10 may be written

$$\begin{aligned} p(k, s) =& |\bar{F}(k, s)|^2 \sum_n \delta_N(s - \frac{2\pi n}{kd}) + \frac{kd}{2\pi} c_0(k, s) \\ &+ \frac{kd}{\pi} \sum_{m=1}^{M} \left\{ \mathrm{Re}[c_m(k, s)] \cos(mksd) - \mathrm{Im}[c_m(k, s)] \sin(mksd) \right\}. \end{aligned} \tag{16}$$

This equation shows that the intensity pattern will consist of sharp diffraction peaks on top of a weak but diffuse scattering background. The actual angular dependence of the diffuse scattering will depend upon the precise nature of the correlations and upon the correlation length given by M. Nevertheless, this equation permits us to make several observations about the scattering without reference to a detailed model of the grating bar shapes.

First of all, note that if the correlations $c_m(k, s)$ are slowly varying functions of s, then the scattering will be quasi-periodic with a period of $2\pi/kd$. That is, the scattering will roughly repeat itself between diffraction orders. The scattering data shown in Figure 1 appears to share this feature. For example, strong scattering peaks appear just to the right of the first, second, and third diffraction orders. In addition, there is a peak just to the left of first order and there appears to be some evidence for peaks just to the left of the second and third orders, although the existence of the peak at the left of second order is questionable. On the other hand, the fact that the scattering is not strictly periodic is just an indication that $c_m(k, s)$ depends upon s.

3.2. Displacement Variations

One might think that since the scattering is relatively weak near zeroth order with no strong peaks in that region, the apparent quasi-periodicity reflected in the data around first, second, and third orders is coincidental. However, an important prediction of Eq. 16 is that for fluctuations involving simple displacements of the grating bars, there can be no scattering near zeroth order. In fact, this is a well known result from the theory of X-ray diffraction from crystals.[7] To prove this in the context of a diffraction grating, consider a model of N grating bar shapes that differ from one another only by a shift, or displacement from their ideal lattice positions. That is, the path-length function for the jth bar is assumed to be given by $z_j(x) = z(x + a_j)$ where a_j denotes the shift of the bar from its ideal position in the unit cell. Furthermore, assume that $z(x)$ is non-zero only for values of x such that $|a_j| \le x \le d - |a_j|$ for all j. Physically, this means that the center of the grating bar whose shape is described by $z(x)$ lies near middle of the cell. Then, the structure factor F_j for the jth unit cell can be written

$$\begin{aligned} F_j(k,s) &= \frac{1}{d}\int_0^d \mathrm{d}x\ e^{iksx+i\phi(x+a_j)} \\ &= \frac{1}{d}e^{-iksa_j}\int_{a_j}^{d+a_j} \mathrm{d}x\ e^{iksx+i\phi(x)} \\ &= \frac{1}{d}e^{-iksa_j}\left(\int_0^d + \int_d^{d+a_j} - \int_0^{a_j}\right)\mathrm{d}x\ e^{iksx+i\phi(x)}. \end{aligned} \tag{17}$$

With the assumption that $z(x)$, and hence $\phi(x)$, vanishes for $x < |a_j|$ or for $x \ge d - |a_j|$, the latter two integrations can be performed with the result

$$F_j(k,s) = e^{-iksa_j}F(k,s) + \frac{1}{iksd}(1 - e^{-iksa_j})(e^{iksd} - 1). \tag{18}$$

Here, $F(k,s)$ has been defined

$$F(k,s) = \frac{1}{d}\int_0^d \mathrm{d}x\ e^{iksx+i\phi(x)}. \tag{19}$$

From this, it trivially follows from Eq. 6, that

$$\bar{F}(k,s) = \langle e^{-iksa_j}\rangle F(k,s) + \frac{1}{iksd}(e^{iksd} - 1)(1 - \langle e^{-iksa_j}\rangle) \tag{20}$$

and we also deduce from Eq. 7 that

$$f_j(k,s) = \left[e^{-iksa_j} - \langle e^{-iksa_j}\rangle\right]\left[F(k,s) - \frac{1}{iksd}(e^{iksd} - 1)\right]. \tag{21}$$

Thus, from this equation it is immediately clear that a model of grating bar *fluctuations consisting of pure displacements produces no scattering in the immediate vicinity of zeroth order.* This result is quite general since no further assumptions were made about about the actual shape of the grating bars.

To see the impact of these fluctuations on the diffraction orders, assume that the shift parameters a_j are independently Gaussian distributed with a mean of zero and a variance of σ^2. For such a distribution, one can show that

$$\langle e^{-i2\pi n a_j/d}\rangle = e^{-\frac{1}{2}(2\pi n\sigma/d)^2} \tag{22}$$

from which it follows

$$|\bar{F}(k,s_n)|^2 = |F(k,s_n)|^2 e^{-(2\pi n\sigma/d)^2}, \tag{23}$$

where $s_n = 2n\pi/kd$ is the value of s at the nth diffraction order. This equation explicitly shows that the fluctuations reduce the power in the diffraction orders, as one would expect from considerations of flux conservation.

3.3. Bar Geometry Variations

The geometry of the jth grating bar may be specified by a number of geometric quantities. Let the μth such quantity for the jth bar be denoted by $\xi_\mu^{(j)}$. For example, a rectangular grating bar may be specified by three quantities: the width (ξ_1), height (ξ_2), and displacement (ξ_3). The structure factor F_j will be a function of these quantities and indicated symbolically by $F_j(k,s,\xi)$. Let $\Delta\xi_\mu^{(j)}$ be the deviation of the μth parameter from its mean, i.e.,

$$\xi_\mu^{(j)} = \langle\xi_\mu^{(j)}\rangle + \Delta\xi_\mu^{(j)}. \tag{24}$$

Assuming that the fluctuations $\Delta\xi_\mu^{(j)}$ are small, we can approximate the structure factor by

$$F_j(k,s,\xi) = F(k,s,\langle\xi\rangle) + \sum_\mu \frac{\partial F}{\partial \xi_\mu}\Delta\xi_\mu^{(j)}, \tag{25}$$

where $F(k,s,\langle\xi\rangle)$ is independent of j by virtue of the assumed homogeneity of the ensemble. It then follows from Eq. 7

$$f_j(k,s,\xi) = \sum_\mu \frac{\partial F}{\partial \xi_\mu}\Delta\xi_\mu^{(j)} \tag{26}$$

and from Eq. 15

$$\begin{aligned} c_m(k,s) &= \langle f_j f_{l+m}^* \rangle \\ &= \sum_{\mu\nu} \frac{\partial F}{\partial \xi_\mu}\frac{\partial F^*}{\partial \xi_\nu}\langle \Delta\xi_\mu^{(j)}\Delta\xi_\nu^{(j+m)}\rangle. \end{aligned} \tag{27}$$

Since the ensemble is assumed to be homogeneous, the correlations $\langle \Delta\xi_\mu^{(j)}\Delta\xi_\nu^{(j+m)}\rangle$ are independent of j, and will be denoted by $G_{\mu\nu}(m)$, i.e.,

$$G_{\mu\nu}(m) = \langle \Delta\xi_\mu^{(j)}\Delta\xi_\nu^{(j+m)}\rangle. \tag{28}$$

Moreover, they are independent of k and s, and assuming left-right invariance of the grating, they satisfy the symmetry relations

$$G_{\mu\nu}(m) = G_{\mu\nu}(-m) = G_{\nu\mu}(m). \tag{29}$$

From this symmetry it follows that $c_m(k,s)$ is real. Since $G_{\mu\nu}(m)$ is an even function of m, it may be represented as Fourier cosine series. In addition, from physical considerations the correlations must fall off with increasing $|m|$. For simplicity, assume that the correlations $G_{\mu\nu}(m)$ may be represented in the form

$$G_{\mu\nu}(m) = \sigma_{\mu\nu}e^{-\alpha_{\mu\nu}|m|}\cos(2\pi m\Omega_{\mu\nu}). \tag{30}$$

In terms of these correlations, the probability density for scattering from Eq. 14 may be expressed as

$$\frac{kd}{2\pi}\sum_{\mu\nu}\sigma_{\mu\nu}\frac{\partial F}{\partial \xi_\mu}\frac{\partial F^*}{\partial \xi_\nu}\left[\sum_{m=-\infty}^{\infty} e^{imksd}e^{-\alpha_{\mu\nu}|m|}\cos(2\pi m\Omega_{\mu\nu})\right]. \tag{31}$$

In deriving this equation, the exponential fall off of the correlations have allowed the sum to extended to all m. The sum over m may be performed by writing the cosine in its complex representation and recognizing that the resulting sum is geometric. It is left as an exercise for the reader to show that

$$\begin{aligned} &\sum_{m=-\infty}^{\infty} e^{imksd}e^{-\alpha|m|}\cos(2\pi m\Omega) \\ &\qquad = \frac{1+e^{-\alpha}}{2}\left[\frac{1-e^{-\alpha}}{(1-e^{-\alpha})^2+2e^{-\alpha}(1-\cos(ksd+2\pi\Omega))}\right. \\ &\qquad\qquad \left. + \frac{1-e^{-\alpha}}{(1-e^{-\alpha})^2+2e^{-\alpha}(1-\cos(ksd-2\pi\Omega))}\right] \end{aligned} \tag{32}$$

The right hand side of this equation shows that the correlations will produce peaks whenever $\cos(ksd \pm 2\pi\Omega)$ is unity, i.e., at values of s when $ksd/2\pi = n \pm \Omega$. In other words, *the scattering will consist of peaks that are symmetrically located about the diffraction orders*, although the strength of the peaks will depend upon the exact form of $c_m(k,s)$. A quick glance at Figure 1 shows that the peaks immediately to the left and right of first order have this property. In addition, the data has a peak just to the right of third order and it appears to have a small peak symmetrically

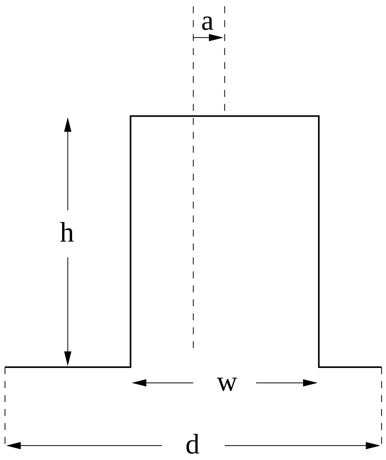

Figure 7. Figure showing the geometric parameters for a rectangular grating bar of width w and height h in a unit cell with period d. The displacement of the bar from the center of the unit cell is given by a.

placed at the left of third order. It is harder to make the argument about second order because if there is a peak to the left of second order then it is suppressed.

The location of the peaks and their widths provide valuable information about the parameters Ω and α. In particular the location of the peaks give information about Ω, and their widths are governed by α. To see this, suppose that α and $\theta = ksd + 2\pi\Omega$ are small. Then the first term term on the right hand side of Eq. 32 may be approximated by

$$\frac{\alpha}{\alpha^2 + \theta^2}. \tag{33}$$

This function has a maximum at $\theta = 0$ and the half-width at half-maximum is α. Hence, the width of the scattering peak provides information about α, which in turn dictates the range of the correlation responsible for the peak. Armed with this knowledge one can make a rough estimate of Ω and α for the correlation associated with the peaks closest to the diffraction orders of Figure 1. We estimate Ω to be around 0.1 and α to be roughly $0.05 \times 2\pi \approx 0.3$. This implies that the correlations have a length of $1/\alpha$ or about 3 grating bars.

4. CORRELATIONS IN THE RECTANGULAR MODEL

In this section, we assume that the grating bar shapes may be represented by a simple rectangle with height h and width w. The center of the bar is assumed to offset a distance a from the center of the lattice node (see Figure 7).

The rectangular geometry readily permits the integral in Eq. 2 to computed with the result

$$F_j(k,s) = \frac{\sin(\frac{1}{2}ksd)}{\frac{1}{2}ksd} + e^{iksa_j}(e^{-kh_j\beta - ikh_j\delta} - 1)\frac{\sin(\frac{1}{2}ksw_j)}{\frac{1}{2}ksd}. \tag{34}$$

As in the previous section, we assume that the correlations are small so that Eq. 27 is valid. Expressing Eq. 27 in terms of h, w, and a, results in

$$\begin{aligned} c_m(k,s) =& d^2\left|\frac{\partial F}{\partial a}\right|^2 G_{aa}(m) + h^2\left|\frac{\partial F}{\partial h}\right|^2 G_{hh}(m) + d^2\left|\frac{\partial F}{\partial w}\right|^2 G_{ww}(m) \\ &+ 2hd\,\mathrm{Re}\left[\frac{\partial F}{\partial a}\frac{\partial F^*}{\partial h}\right]G_{ah}(m) + 2hd\,\mathrm{Re}\left[\frac{\partial F}{\partial w}\frac{\partial F^*}{\partial h}\right]G_{wh}(m) \\ &+ 2d^2\,\mathrm{Re}\left[\frac{\partial F}{\partial a}\frac{\partial F^*}{\partial w}\right]G_{aw}(m), \end{aligned} \tag{35}$$

h	$=$	0.64μm	w/d	$=$	0.84			
σ_{ww}	$=$	6.5×10^{-5}	α_{ww}	$=$	0.27	Ω_{ww}	$=$	0.092
σ_{hh}	$=$	5.0×10^{-5}	α_{hh}	$=$	0.60	Ω_{hh}	$=$	0.50
σ_{wh}	$=$	4.2×10^{-5}	α_{wh}	$=$	0.30	Ω_{wh}	$=$	0.91

Table 2. This table shows the best fit parameter values as deduced from fitting the scattering data at 1.384, 1.775, and 2.035 keV. Displacement variations were not considered. Figure 9 shows the resulting fit.

where factors of d and h have been introduced to make the correlation functions $G_{\mu\nu}(m)$ unitless, e.g.,

$$\begin{aligned} G_{wh}(m) &= \langle \frac{\Delta w}{d} \frac{\Delta h}{h} \rangle \\ &= \sigma_{wh} e^{-\alpha_{wh}|m|} \cos(2\pi m \Omega_{wh}). \end{aligned} \tag{36}$$

The partial derivatives of the structure factor $F(k,s)$ are rather straightforward to carry out with the result:

$$\begin{aligned} \left|\frac{\partial F}{\partial a}\right|^2 &= \frac{4}{d^2}[1 - 2e^{-kh\beta}\cos(kh\delta) + e^{-2kh\beta}]\sin^2(\frac{1}{2}ksw) \\ \left|\frac{\partial F}{\partial w}\right|^2 &= \frac{1}{d^2}[1 - 2e^{-kh\beta}\cos(kh\delta) + e^{-2kh\beta}]\cos^2(\frac{1}{2}ksw) \\ \left|\frac{\partial F}{\partial h}\right|^2 &= k^2(\delta^2+\beta^2)e^{-2kh\beta}\left[\frac{\sin(\frac{1}{2}ksw)}{\frac{1}{2}ksd}\right]^2 \\ \mathrm{Re}\left[\frac{\partial F}{\partial w}\frac{\partial F^*}{\partial h}\right] &= \frac{k}{d}e^{-kh\beta}\left[\beta\cos(kh\delta) + \delta\sin(kh\delta) - \beta e^{-kh\beta}\right]\frac{\sin(ksw)}{ksd} \\ \mathrm{Re}\left[\frac{\partial F}{\partial a}\frac{\partial F^*}{\partial h}\right] &= \frac{4k}{d}e^{-kh\beta}\left[-\beta\sin(kh\delta) + \delta\cos(kh\delta) - \delta e^{-kh\beta}\right]\frac{\sin^2(\frac{1}{2}ksw)}{ksd} \\ \mathrm{Re}\left[\frac{\partial F}{\partial a}\frac{\partial F^*}{\partial w}\right] &= 0, \end{aligned} \tag{37}$$

This model contains 17 independent parameters that include the bar width w and height h as well as the 15 parameters that define the correlation functions $G_{\mu\nu}(m)$. A χ^2 minimization procedure based upon the Marquardt algorithm[8] was used in deducing the parameters. Since the scattering contains energy-dependent terms, a global fit to the scattering data at 1.384, 1.775, and 2.035 keV was performed. These particular data were used because we feel that they have the the least experimental uncertainties and would provide the best estimates for the parameters.

We first considered the case involving no correlations. This reduced the parameter space from 17 dimensions down to 5. Figure 8 shows the best fit of the rectangular model with no correlations to the data 1.775 keV. From the figure, one can immediately see that the model without correlations fails to reproduce any of the rich structure seen in the data. The lack of any prominent scattering features is a strong indication that the grating scatter must be the result of correlations.

Figure 9 shows the result of a simultaneous fit to the 1.384, 1.775, and 2.035 keV scattering data when correlations were included. The parameters deduced from the fit are presented in Table 2. Fluctuations involving the bar displacement a were not included in this fit, reducing the number of free parameters down to 11 (including them did not substantially improve the fit so we felt that it was better to restrict the model to 11 parameters).

We can immediately see that including correlations results in a dramatic improvement of the fit. Although not perfect, this model shows all the main features in the scattering data at these energies. This is especially true for our best data set, the data at 1.775 keV. The main areas of disagreement are near the half order locations, and at the second and third order scattering peaks where it underestimates the scattering. Also the average effective-rectangular bar width resulting from the fit is about 35 percent larger than the accepted average width: w/d is 0.84 compared to

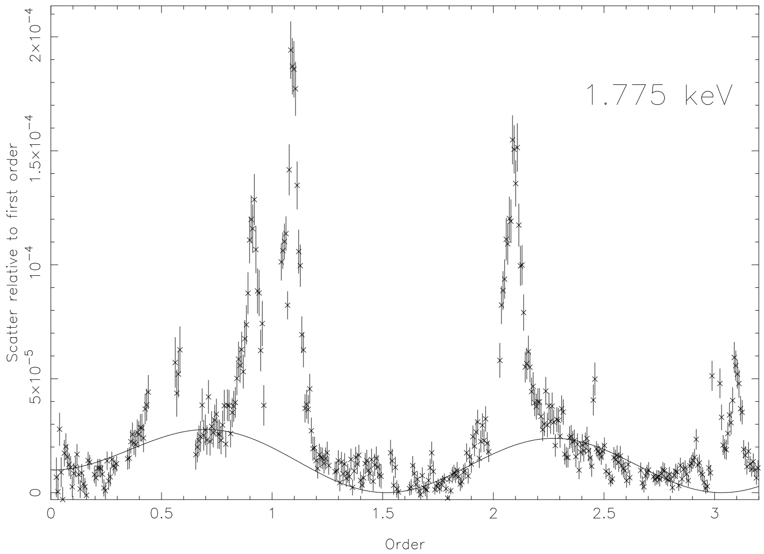

Figure 8. This figure shows the best fit to the 1.775 keV scattering data using a rectangular model with no correlations.

$w/d \approx 0.60$ based on detailed vertex-model fits[3] to the facet data.[2] Nevertheless, we feel that the fit is remarkable considering the simplicity and known shortcomings of the model.

Figure 10 shows the predictions of the model at higher energies ($5 - 8$ keV) overlaid with scattering data at those energies. The model is not too bad at predicting the scatter between the first and third orders in the high energy data, although it performs poorly when extrapolated past third order. This should not be too surprising since the data used to determine the model's parameters did not extend past third order.

5. DISCUSSION

While the fit adequately reproduces many of the features observed in the data, the real physics lies in the form of the correlation functions deduced from the fit. A plot of the three correlation functions $G_{ww}(m)$, $G_{hh}(m)$, and $G_{wh}(m)$ is shown in Figure 11. The dominate correlation is $G_{ww}(m)$, i.e., the width-width correlation. As can been seen from the figure, the shape of this correlation implies that the grating bar widths of the first and second nearest neighbor bars are positively correlated whereas the widths of the fourth and fifth nearest neighbors are anti-correlated. Physically, this correlation could arise by groups of two or three grating bars leaning in the same general direction. It is this correlation that appears to be responsible for the strong peaks closest to the diffraction orders.

The form of the height-height correlations, $G_{hh}(m)$, indicate that heights of nearest neighbors are anti-correlated. This has the effect of doubling the grating period and it is this correlation that contributes to the scattering near the half-orders. The simplest physical picture of this type of correlation is that of the left and right nearest neighbors leaning in the direction of the central bar.

One must not come to the conclusion that these physical pictures of the width-width and height-height correlations are mutually exclusive. Indeed there is nothing that prevents the gratings from having both types of disorder because each type may occur at disjoint parts of the grating. Thus the physical picture that emerges from these correlations is that the grating consists of localized groups of leaning bars and that the groups may be classified according to at least two varieties. The first type consists of several bars that lean in the same direction, possibly three or four together. The second type of fluctuation involves a group of three bars where the left and right bar leans on the central one. The final correlation, $G_{wh}(m)$, can be though of as the result of a combination of these two physical pictures.

If the explanation of the source of the correlations is correct, then the size of the correlations may be used to deduce the frequency of such fluctuations. For example, consider the width-width correlation, which is roughly 10^{-4} and implies that the fluctuation $\Delta w/d$ is 0.01 or about 1 percent of the period. This is consistent with 1 in 200 grating bars having a fluctuation in width by half a period, while the other 199 bars do not fluctuate at all.

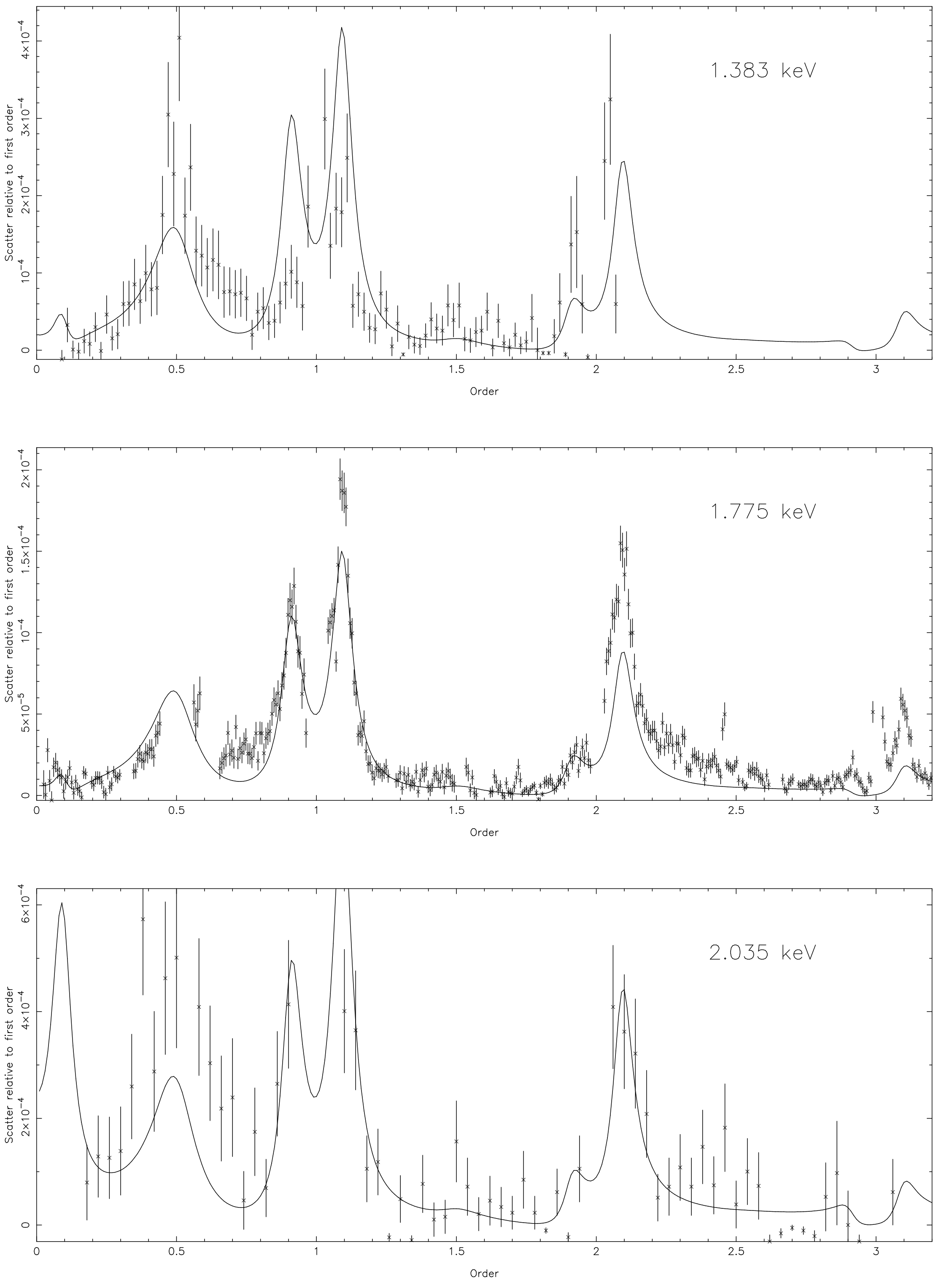

Figure 9. This figure shows results of a global fit to the 1.384, 1.775, and 2.035 keV scattering data. The model included width-width, height-height, and width-height correlations.

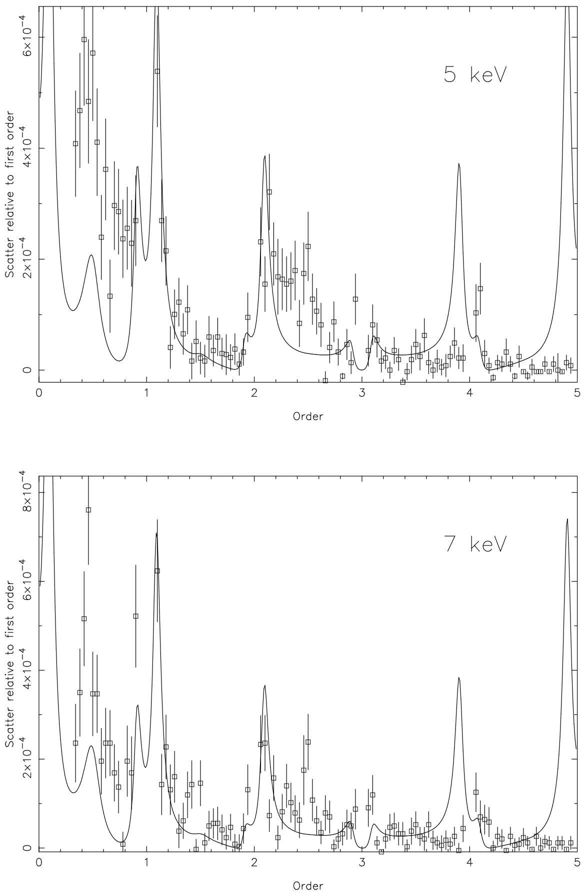

Figure 10. These figures show the predictions of the correlation model of Figurefig:en0-2 evaluated at at 5, 6, 7, and 8 keV overlaid with the scattering data at these energies.

Finally, if the scattering is due to the presence of leaning bars, then the absence of any detectable scattering by the MEG gratings is very easy to understand. The MEG gratings have a bar width of about 0.2 microns, and a bar height of about 0.36 microns. Hence, compared to the HEG gratings, which have a height to width ratio of about 5, the MEG gratings have a height to width ratio that is less than 2, and would be much less likely to lean.

6. CONCLUSION AND NEXT STEPS

Despite the apparent simplicity of both the rectangular model and the assumed form of the correlations, the resulting fits to the data are relatively good. The primary goal of this work was to demonstrate that the observed scatter could be the result of imperfections in the gratings and not to some other more exotic physical explanation. We are quite content that this goal has been achieved and we believe that the rectangular model has served us quite well in this regard.

Nevertheless, the fits are far from perfect and there are several discrepancies that should be accounted for by a more accurate and sophisticated model. Over 144 distinct grating facets contributed to the scatter and it is not inconceivable that a better fit could be achieved by using more than one set of bar parameters. This is perhaps the most trivial extension of the simple rectangular model used here. Also, we assumed that the deviations in the bar parameters from their ideal values were small enough that only first order deviations were important. Although the average fluctuation was found to be very small, any one bar parameter could deviate significantly from its mean, with the implication that higher order terms in the expansion could be significant.

Since the actual bars have path-length functions that are more trapezoidally shaped than rectangular, it makes more sense to consider a model involving trapezoidal bars. Such a model would also permit a more direct study of

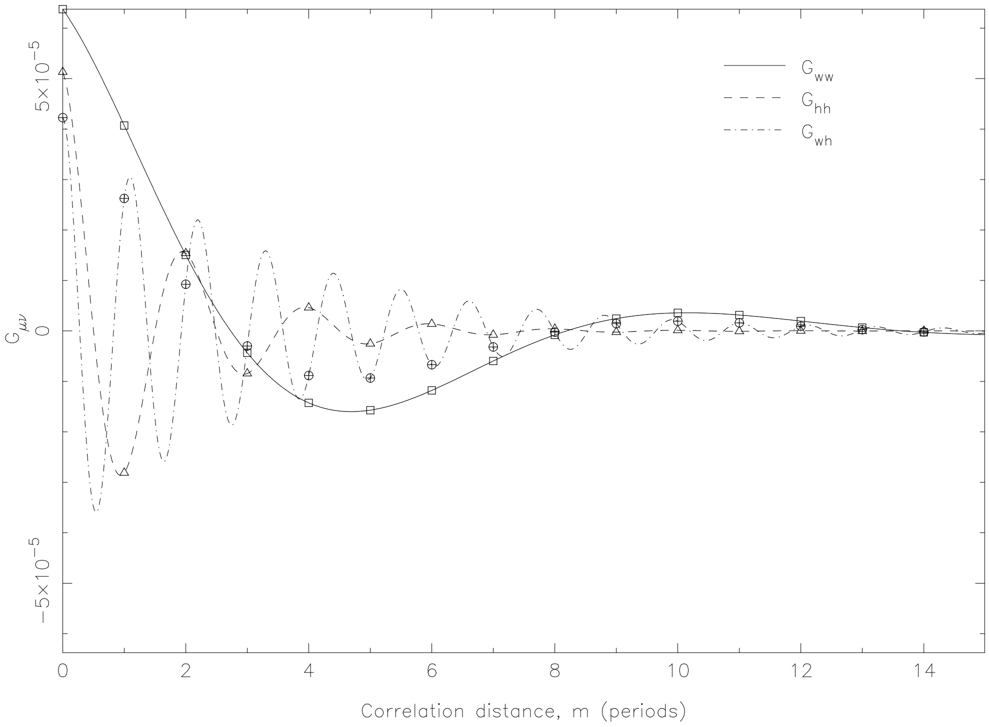

Figure 11. This figure shows a plot of the correlation functions parameterized in Eq. 30 using the parameters from Table 2.

leaning bars— something that is not possible within the confines of the rectangular model. Finally, there are other components of the gratings that have been totally ignored in this work, and the fluctuations in these quantities could contribute to scatter. A more complete treatment would also include these fluctuations, e.g., fluctuations in the plating base as well as in the thickness of the polyimide film.

ACKNOWLEDGMENTS

This work benefited from the encouragement of Claude Canizares and the other members of the MIT/ASC and HETG teams. We also thank the PSU ACIS group for making the ACIS XRCF data sets conveniently available.

The work was supported in part by NASA under the HETG contract NAS-38249 and the AXAF Science Center contract to the Smithsonian Astrophysical Observatory, NAS8-39073.

REFERENCES

1. Marshall, H.L., et al, "Towards the Calibration of the HETGS Line Response Function," in Grazing Incidence and Multilayer X-ray Optical Systems, Proc. SPIE, vol 3113, pp 160-171, 1997.
2. The AXAF Observatory Guide, Rev. 1.0, Figure 7.4, p.175
3. T.H. Markert et al, "Modeling the Diffraction Efficiencies of the AXAF High Energy Transmission Gratings", in EUV, X-Ray, and Gamma-Ray Instrumentation for Astronomy VI, Proc. SPIE, vol 2518, pp 424-437, 1995.
4. Weisskopf, M.C., et al, "Calibration of the AXAF Observatory: Overview," in Grazing Incidence and Multilayer X-ray Optical Systems, Proc. SPIE, vol 3113, pp 2-17, 1997.
5. Feynman, R.P., and Hibbs, A.R., *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965
6. Stratonovich, R.L., *Topics in the Theory of Random Noise*, Volume 1, Gordon and Breach, 1963
7. Warren, B.E., *X-Ray Diffraction*, Dover Publications, 1990
8. Bevington, P.R., and Robinson, D.K, *Data Reduction and Error Analysis for the Physical Sciences*, Second Edition, McGraw-Hill, 1992.