On the design of wide-field x-ray telescopes

Ronald F. Elsner[a], Stephen L. O'Dell[a], Brian D. Ramsey[a], and Martin C. Weisskopf[a]

[a]NASA Marshall Space Flight Center, VP62, Huntsville, AL 35812

ABSTRACT

X-ray telescopes having a relatively wide field-of-view and spatial resolution vs. polar off-axis angle curves much flatter than the parabolic dependence characteristic of Wolter I designs are of great interest for surveys of the X-ray sky and potentially for study of the Sun's X-ray emission. We discuss the various considerations affecting the design of such telescopes, including the possible use of polynomial mirror surface prescriptions, a method of optimizing the polynomial coefficients, scaling laws for mirror segment length vs. intersection radius, the loss of on-axis spatial resolution, and the positioning of focal plane detectors.

Keywords: X-ray astronomy, X-ray optics, ray trace, wide field-of-view optimization

1. INTRODUCTION

X-ray optics involving polynomial extensions of Wolter I mirror prescriptions are of astronomical interest in the context of wide-field X-ray surveys, the goal of the proposed Wide-Field X-ray Telescope (WFXT) mission, and potentially also for study of the Sun's X-ray disk. In this paper, we explore methods for designing, optimizing, and evaluating such optics.

Mirror presciptions and ray trace methods are reviewed in §2—§7. Expressions for evaluating spatial resolution and effective area using Monte-Carlo simulations are presented in §8 and §9. The approximate analytic expressions for Wolter I optics given in VanSpeybroeck and Chase[1] are reviewed and extended to a broader domain in §10. Effects due to figure errors and surface roughness are briefly discussed in §11. Our methods for designing and optimizing telescopes are described in §12, including the topics of figures of merit §12.1, design considerations §12.2, single shell optimization §12.3, and nested configuations §12.4. A few closing remarks and plans for the future are presented in §13.

2. MIRROR PRESCRIPTIONS

We assume the lengths and nominal grazing angles for the primary (P) and secondary (H) mirror segments are the same. We place the origin of the optical (z) axis at the intersection plane between the two mirror segments, with z increasing toward the front of the primary and decreasing toward the back of the secondary. (This points the z-axis in the direction opposite to that in Burrows, Burg, and Giacconi,[2] henceforth BBG).) The mirror surface prescriptions for the conical approximation are:

$$
\begin{aligned}
r_{c,P}^2(z_P) &= r_0^2 \left[1 + \tan(\alpha_0)(z_P/r_0) \right]^2 \\
r_{c,H}^2(z_H) &= r_0^2 \left[1 + \tan(3\alpha_0)(z_H/r_0) \right]^2
\end{aligned}
\tag{1}
$$

The mirror surface prescriptions for Wolter I and polynomial optics are:

Further author information: (Send correspondence to R.F.E)
R.F.E.: E-mail: ron.elsner@nasa.gov, Telephone: 256 961 7765
S.L.O.: E-mail: steve.o'dell@nasa.gov, Telephone: 256 961 7776
B.D.R: E-mail: brian.ramsey@nasa.gov, Telephone: 256 961 7784
M.C.W.: Email: martin@smoker.msfc.nasa.gov, Telephone: 256 961 7798

$$r_{p,P}^2(z_P) \;=\; r_0^2 \left[\, 1 \,+ 2\,\tan\left(\alpha_0\right)\,(z_P/r_0) \;+\; \delta a_2\,(z_P/r_0)^2 \;+\; \delta a_3\,(z_P/r_0)^3 \,\right]$$

$$r_{p,H}^2(z_P) \;=\; r_0^2 \left[\, 1 \,+ 2\,\tan\left(3\alpha_0\right)\,(z_P/r_0) \;+\; \tan^2\left(3\alpha_0\right)\,h(\alpha_0)\,(z_P/r_0)^2 \right. \tag{2}$$
$$\left. \;+\; \delta b_2\,(z_P/r_0)^2 \;+\; \delta b_3\,(z_P/r_0)^3 \,\right]$$

$$h(\alpha_0) \;=\; 1 - 1/(1 + 2\cos\left(2\alpha_0\right))^2,$$

where

$$\alpha_0 \;=\; \left(\frac{1}{4}\right)\,\tan^{-1}\left(\frac{r_0}{f}\right), \tag{3}$$

r_0 is the radius of the segments at their intersection plane ($z_P = z_H = 0$), and f is the the distance from the intersection plane to the on-axis focus for Wolter I optics. Thus the on-axis focus for Wolter I optics is at $z = -f$. We denote the mirror segment length by ℓ. The mirror shell length is 2ℓ. Wolter I optics require $\delta a_2 = \delta a_3 = \delta b_2 = \delta b_3 = 0$.

3. APERTURE FOR INCIDENT RAYS

For a given shell s, the maximum radius of the aperture, $r_{max,s}$, is independent of off-axis polar angle, θ, and position azimuthal angle, ϕ. The value of $r_{max,s}$ is determined from the equation for the surface of the primary with z set equal to ℓ. The minimum radius of the aperture, $r_{min,s}(\theta, \phi)$, is a function of θ and ϕ. Straightforward geometry leads to the expression

$$r_{min,s}(\theta, \phi) \;=\; -\,\ell\,\tan\theta\,\cos\phi \;+\; \sqrt{r_{0,s}^2 \;-\; \ell^2\,\tan^2\theta\,\sin^2\phi}, \tag{4}$$

where $r_{0,s}$ is the intersection radius for shell s. Amusingly, although the shape and centroid of the incident aperture varies with θ, the total geometric area of the incident aperture is independent of θ, and is given by

$$A_{geom,s} \;=\; \pi\,(\, r_{max,s}^2 \;-\; r_{0,s}^2 \,). \tag{5}$$

These expressions neglect possible shadowing by the next inner shell, a pre-collimator, or any support structures.

4. RAY INTERSECTION WITH OPTIC

The equations of ray tracing X-ray optics are simple in principle. Given a starting position $(x_0,\,y_0,\,z_0)$ and normalized wavevector $\hat{k}_0 = (k_{x,0},\,k_{y,0},\,k_{z,0})$, the intersection $(x_1,\,y_1,\,z_1)$ with an optic is determined from:

$$x_1 \;=\; x_0 \;+\; k_{x,0}\,t$$

$$y_1 \;=\; y_0 \;+\; k_{y,0}\,t \tag{6}$$

$$z_1 \;=\; z_0 \;+\; k_{z,0}\,t.$$

The parametric variable t is determined from

$$(\, x_0 \;+\; k_{x,0}\,t\,)^2 \;+\; (\, y_0 \;+\; k_{y,0}\,t\,)^2 \;=\; r_O^2(z_0 \;+\; k_{z,0}\,t\,). \tag{7}$$

Here $r_O(z)$ is the radius of the optic as a function of distance along the optical axis. For simple ideal surfaces with no figure errors, $r_O(z)$, Eq. (7) can be solved analytically for t, and hence for (x_1, y_1, z_1). For more complicated surfaces or in the presence of surface errors, we proceed by first finding the intersection (x_{10}, y_{10}, z_{10}) with a simple surface, $r_{O,0}(z)$, approximating the actual one, $r(z) = r_{O,0}(z) + \Delta r(z)$, and then correcting the solution using a perturbation method. The 1st order corrections to the intersection point are given by:

$$
\begin{aligned}
x_1 &= x_{10} + k_{x,0}\,\delta t \\[2mm]
y_1 &= y_{10} + k_{y,0}\,\delta t \\[2mm]
z_1 &= z_{10} + k_{z,0}\,\delta t,
\end{aligned}
\tag{8}
$$

where δt is given by

$$
\delta t = \frac{r_{O,0}(z_{10})\,\Delta r(z_{10})}{k_{x,0}\,x_{10} + k_{y,0}\,y_{10} - k_{z,0}\,r_{O,0}(z_{10})\,\frac{\partial r_{O,0}}{\partial z}\,|_{z=z_{10}}}.
\tag{9}
$$

In Eqs. (8) & (9), we've neglected radial perturbations dependent on x and y. This is justified because contributions from axial slope perturbations dominate those from (x, y, z) perturbations in intersection points or from azimuthal slope errors to the PSF. This method can also be used to deal with low-frequency figure errors.

5. SPECULAR REFLECTION FROM A SURFACE

Neglecting azimuthal variations as is appropriate for ideal surfaces, unit normals $\hat{n}_1 = (n_{x,1}, n_{y,1}, n_{z,1})$ to the surface O at the intersection point (x_1, y_1, z_1) are given by

$$
\begin{aligned}
n_{x,1} &= \frac{1}{n_n}\left(-\frac{x_1}{\sqrt{x_1^2 + y_1^2}}\right) \\[3mm]
n_{y,1} &= \frac{1}{n_n}\left(-\frac{y_1}{\sqrt{x_1^2 + y_1^2}}\right) \\[3mm]
n_{z,1} &= \frac{1}{n_n}\left(\frac{\partial r_O}{\partial z}\,|_{z=z_1}\right),
\end{aligned}
\tag{10}
$$

where

$$
n_n \equiv \sqrt{1 + \left(\frac{\partial r_O}{\partial z}\,|_{z=z_1}\right)^2}.
\tag{11}
$$

For example, for a polynomial primary surface:

$$
\frac{\partial r_{p,P}}{\partial z} = \frac{\tan(\alpha_0) + \delta a_2(z/r_0) + (3/2)\delta a_3(z/r_0)^2}{\sqrt{1 + 2(z/r_0)\tan(\alpha_0) + \delta a_2(z/r_0)^2 + \delta a_3(z/r_0)^3}}.
\tag{12}
$$

The direction vector $\hat{k}_1 = (k_{x,1}, k_{y,1}, k_{z,1})$ of the ray specularly reflected from the surface is given by

$$\hat{k}_1 \;=\; (\hat{n}_1 \times \hat{k}_0) \times \hat{n}_1 \;-\; \hat{n}_1\,(\hat{k}_0 \cdot \hat{n}_1) \;=\; \hat{k}_0\,(\hat{n}_1 \cdot \hat{n}_1)\;-\;2\,\hat{n}_1\,(\hat{k}_0 \cdot \hat{n}_1)$$

$$k_{x,1} \;=\; k_{x,0}\,(n_{y,1}^2 + n_{z,1}^2 - n_{x,1}^2) - 2\,n_{x,1}\,(k_{y,0}\,n_{y,1} + k_{z,0}\,n_{z,1})$$

$$k_{y,1} \;=\; k_{y,0}\,(n_{x,1}^2 + n_{z,1}^2 - n_{y,1}^2) - 2\,n_{y,1}\,(k_{x,0}\,n_{x,1} + k_{z,0}\,n_{z,1}) \tag{13}$$

$$k_{z,1} \;=\; k_{z,0}\,(n_{x,1}^2 + n_{y,1}^2 - n_{z,1}^2) - 2\,n_{z,1}\,(k_{x,0}\,n_{z,1} + k_{y,0}\,n_{y,1}).$$

The right-hand-side of the top relation in Eqs. (13) follows from the vector identity $(\hat{A} \times \hat{B}) \times \hat{C} = \hat{B}(A \cdot C) - \hat{A}(\hat{B} \cdot \hat{C})$. The vector $\hat{k}_1$ will be normalized to unity if $\hat{k}_0$ and $\hat{n}_1$ are, as the reader can demonstrate for himself.

If included directly in the ray trace, scattering due to dust or surface roughness introduces perturbations to the specularly reflected direction vector.

6. PROPAGATION TO THE FOCAL PLANE

Given the intersection position (x_2, y_2, z_2) with the secondary and the normalized wavevector ($k_{x,2}$, $k_{y,2}$, and $k_{z,2}$) reflected from the secondary, the ray is propagated to the focal surface plane $z = -f + \delta z$ by solving $-f + \delta z = z_2 + k_{z,2}\,t$ for t and substituting in $x_f = x_2 + k_{x,2}\,t$ and $y_f = y_2 + k_{y,2}\,t$ to find x_f and y_f:

$$x_f \;=\; x_2 + \frac{k_{x,2}}{k_{z,2}}\,(-f + \delta z - z_2)$$

$$\tag{14}$$

$$y_f \;=\; y_2 + \frac{k_{y,2}}{k_{z,2}}\,(-f + \delta z - z_2).$$

From Eqs. (14), and the quantities x_2, y_2, z_2, $k_{x,2}/k_{z,2}$ and $k_{y,2}/k_{z,2}$ we can construct the rms blur radius, $\sigma_{rms}(\theta, \delta z)$, as described in § 8.1.

7. GRAZE ANGLES

The graze angles, α_1 and α_2, for the ray at the primary and secondary, respectively, are:

$$\alpha_1 \;=\; \left(\frac{1}{2}\right)\,\cos^{-1}(\hat{k}_1 \cdot \hat{k}_0),$$

$$\tag{15}$$

$$\alpha_2 \;=\; \left(\frac{1}{2}\right)\,\cos^{-1}(\hat{k}_2 \cdot \hat{k}_1).$$

Values of the graze angles are needed for applications requiring the calculation of reflectivities for the weighting of individual rays.

8. MEASURES OF SPATIAL RESOLUTION

8.1 Calculation of σ_{rms} from a ray bundle

RMS dispersion radii for PSFs derived from a Monte-Carlo simulation are determined as follows:

$$< Q > \equiv \frac{1}{N} \sum_1^N Q_i \tag{16}$$

$$\sigma_{rms}^2 = \sigma_{x,rms}^2 + \sigma_{x,rms}^2$$

$$\sigma_{x,rms}^2 = \frac{N}{N-1}(< x^2 > - < x >^2) \tag{17}$$

$$\sigma_{y,rms}^2 = \frac{N}{N-1}(< y^2 > - < y >^2),$$

where N is the number of doubly reflected rays reaching the focal plane. We define $x_i = x_{f,i} + \delta z(k_{x,i}/k_{z,i})$ and $y_i = y_{f,i} + \delta z(k_{y,i}/k_{z,i})$, where $x_{f,i}$ is the ray x-position in the flat focal plane through the on-axis focus, $k_{x,i}$ the x-component of the doubly reflected ray direction vector, x_i the ray x-component at a height δz above that plane, and similar definitions for the y-components. Then Eqs. (17) become:

$$\sigma_{rms}^2(\delta z) = \sigma_{x,rms}^2(\delta z) + \sigma_{y,rms}^2(\delta z), \tag{18}$$

$$\sigma_{x,rms}^2(\delta z) = \left(\frac{N}{N-1}\right)\left(c_{x,rms} + 2\,b_{x,rms}\,\delta z + a_{x,rms}\,\delta z^2\right)$$

$$c_{x,rms} = < x_f^2 > - < x_f >^2$$

$$b_{x,rms} = \left[< x_f \left(\frac{k_x}{k_z}\right) > - < x_f >< \left(\frac{k_x}{k_z}\right) >\right] \tag{19}$$

$$a_{x,rms} = \left[< \left(\frac{k_x}{k_z}\right)^2 > - < \left(\frac{k_x}{k_z}\right) >^2\right]$$

$$\sigma_{y,rms}^2(\delta z) = \left(\frac{N}{N-1}\right)\left(c_{y,rms} + 2\,b_{y,rms}\,\delta z + a_{y,rms}\,\delta z^2\right)$$

$$c_{y,rms} = < y_f^2 > - < y_f >^2$$

$$b_{y,rms} = \left[< y_f \left(\frac{k_y}{k_z}\right) > - < y_f >< \left(\frac{k_y}{k_z}\right) >\right] \tag{20}$$

$$a_{y,rms} = \left[< \left(\frac{k_y}{k_z}\right)^2 > - < \left(\frac{k_y}{k_z}\right) >^2\right].$$

For a flat detector with face perpendicular to the optical axis, the optimal focal height, δz_{min}, determined by minimizing the RMS blur σ_{rms}, then is:

$$\delta z_{min} = -\frac{b_{x,rms} + b_{y,rms}}{(a_{x,rms} + a_{y,rms})}. \tag{21}$$

Substituting Eq. (21) into Eqs. (19)—(20), and assuming $N/(N-1) \simeq 1$, leads to:

$$\sigma_{rms}^2(\delta z = 0) = c_{x,rms} + c_{y,rms}, \tag{22}$$

$$\sigma_{rms}^2(\delta z_{min}) = \sigma_{rms}^2(0) - \frac{(b_{x,rms} + b_{y,rms})^2}{(a_{x,rms} + a_{y,rms})} \tag{23}$$

These equations apply independent of the mirror prescription.

For Wolter I optics, δz_{min} is nearly a parabola in the polar off-axis angle θ, and the best focal surface is usually not a flat plane. This leads to schemes for arranging focal plane detectors to follow the best focal surface as closely as physically possible.

8.2 Measures of encircled energy

Encircled energy $(EE(r))$ vs. r is just the fraction of rays falling inside a radius r, and obviously varies from 0 at $r = 0$ to 1 as $r \to \infty$. Common measures of spatial resolution are $r_{50} = HPD/2$, the radius inside which 50% of the rays fall, and $r_{80} = D_{80}/2$, the radius inside which 80% of the rays fall. Here HPD is the PSF half-power diameter. For a two-dimensional gaussian PSF with a $1/r$ cusp, one finds $HPD = 1.35\ \sigma_{rms}$ and $D_{80} = 2.56\ \sigma_{rms}$.

9. EFFECTIVE AREA $A_{EFF}(\theta)$ AT $E = 0$

In the zero energy limit, where reflectivities are all unity, the effective area is defined by

$$A_{eff}(E = 0, \theta) = A_{geom}(\theta) \times n_2(\theta)/n_{inc}(\theta), \tag{24}$$

where $A_{geom}(\theta)$ is the geometrical area of the incident aperture (whether for a single shell or a set of nested shells) at polar off-axis angle θ (see § 3), $n_2(\theta)$ the number of doubly reflected rays through the optic, and $n_{inc}(\theta)$ the number incident on the aperture.

10. APPROXIMATE ANALYTIC EXPRESSIONS FOR WOLTER I OPTICS

We now consider Wolter I optics, for which $\delta a_2 = \delta a_3 = \delta b_2 = \delta b_3 = 0$ in Eqs. (2).

10.1 Single shell

VanSpeybroeck & Chase[1] provide a number of useful approximate analytic expressions characterizing the performance of Wolter 1 optics. For the RMS blur radius, $\sigma_{rms,0}(\theta)$, as a function of off-axis angle, θ, in a flat plane perpendicular to the optical axis at the on-axis focus ($\delta z = 0$), their expression is

$$\sigma_{rms,0}(\theta, \alpha_0, \mu_0) \simeq \mu_0 (2/4)(\ell/f)(tan^2\theta / tan\alpha_0) + 4\ tan\theta\ tan^2\alpha_0. \tag{25}$$

In Eq. (25), $\alpha_0 = (1/4)\ tan^{-1}(r_0/f)$, the graze angle at the optic radius of the primary-secondary intersection, f the focal length, ℓ the mirror segment length (so that $2\ \ell$ is the shell length), and $\mu_0 \sim 1$ an adjustment factor. The second term on the RHS of this equation is due to coma. Since coma grows linearly with θ, while the spherical aberration term grows as θ^2, coma is only important at the smallest values of θ. In the flat on-axis focal plane, the cross-over occurs at $tan\theta \simeq (8\ f\ /\ \ell)\ tan^3\alpha_0$.

The VanSpeybroeck & Chase[1] expression for the best focal position, $\delta z_{best}(\theta)$, above the on-axis focal plane as a function of off-axis angle, θ, is

$$\delta z_{best}(\theta, \alpha_0, \mu_{\delta z}) \simeq \mu_{\delta z} (1/8) \ell (\tan\theta / \tan\alpha_0)^2, \tag{26}$$

where $\mu_{\delta z} \sim 1$ is another adjustment factor. On the best focal surface given by Eq. (26), VanSpeyBroeck & Chase[1] find that the RMS blur radius, $\sigma_{rms}(\theta)$, neglecting coma, is

$$\sigma_{rms,best}(\theta, \alpha_0, \mu_1) \simeq \mu_1 (1/4) (\ell/f) (tan^2\theta / tan\alpha_0). \tag{27}$$

Comparing Eqs. (25) and (27), we see that on the best focal surface the first term on the RHS, due to spherical abberations, is half its value in the on-axis focal plane.

In the low energy limit $E \to 0$, VanSpeybroeck & Chase (1972) find that the effective area as a function of θ is

$$A_{eff}(\theta, \alpha_0, \mu_{A,1}, \mu_{A,2}) \simeq \mu_{A,1} A_{geom} [1 - \mu_{A,2} (2/3) (\theta/\alpha_0)], \tag{28}$$

where A_{geom} is the on-axis geometric area, and $\mu_{A,1} \sim 1$ and $\mu_{A,2} \sim 1$ are adjustment factors.

10.2 Single shell generalization

The results of §8.1 show that we can write the RMS blur radius, in linear dimensions (such as cm), as a function of off-axis angle θ and height δz above the flat on-axis focal plane as

$$\sigma^2_{rms}(\theta, \delta z) \simeq c_s + 2 b_s \delta z + a_s \delta z^2. \tag{29}$$

Here

$$
\begin{aligned}
c_s &= <x_f^2 + y_f^2> - <x_f>^2 - <y_f>^2 \\
b_s &= \left[<x_f\left(\frac{k_x}{k_z}\right) + y_f\left(\frac{k_y}{k_z}\right)> - <x_f><\left(\frac{k_x}{k_z}\right)> - <y_f><\left(\frac{k_y}{k_z}\right)> \right] \\
a_s &= \left[<\left(\frac{k_x}{k_z}\right)^2 + \left(\frac{k_y}{k_z}\right)^2> - <\left(\frac{k_x}{k_z}\right)>^2 - <\left(\frac{k_y}{k_z}\right)>^2 \right].
\end{aligned}
\tag{30}
$$

According to the results from §10.1, we can write

$$c_s(\theta, \alpha_0, \mu_c) \simeq \left[\mu_c (2/4) \ell (tan^2\theta / tan\alpha_0) + 4 f \, tan\theta \, tan^2\alpha_0 \right]^2, \tag{31}$$

where $\mu_c \sim 1$.

Monte-Carlo simulations show that a very good approximation for a_s is given by

$$a_s(\alpha_0, \mu_a) \simeq (4 \mu_a \tan\alpha_0)^2, \tag{32}$$

where $\mu_a \sim 1$.

The condition that the first term in Eq. (25) is cut in half on the best focal surface now gives us a way to determine b_s. Neglecting the coma term in the expression for c_s, the result is

$$b_s(\theta, \mu_c, \mu_a) \simeq - \mu_a \mu_c \sqrt{3} \ell \tan^2\theta. \tag{33}$$

So far this expression has been followed by the results of the Monte-Carlo simulations for Wolter I optics that we have done. Then we have

$$\sigma_{rms}^2(\delta z, \theta, \alpha_{0,s}, \mu_c, \mu_a) \; \simeq \; c_s(\theta, \alpha_0, \mu_c) \; + \; 2\,b_s(\theta, \mu_c, \mu_a)\,\delta z \; + \; a_s(\alpha_0, \mu_a)\,\delta z^2 \tag{34}$$

$$\delta z_{min} \; = \; -\,\frac{b_s(\theta, \mu_c, \mu_a)}{a_s(\alpha_0, \mu_a)}.$$

Substitution of $a_s(\alpha_0, \mu_a)$ and $b_s(\theta, \mu_c, \mu_a)$ in the last expression above reproduces the dependences in the expression for the best focal surface, Eq. (26).

10.3 Nested shells

Eqs. (29)—(34) are for a single Wolter I shell. For a telescope consisting of nested shells, we might consider summing expressions for each shell, weighted by effective area:

$$A_{eff,nested}(\theta, \mu_4, \mu_5) \; = \; \sum_{s=1}^{N_S} A_{eff}(\theta, \alpha_{0,s}, \mu_{A,1}, \mu_{A,2}), \tag{35}$$

$$
\begin{aligned}
\sigma_{rms,0,nested}(\theta, \mu_c, \mu_{A,1}, \mu_{A,2}) \; &= \; \left(\frac{1}{A_{eff,nested}(\theta, \mu_{A,1}, \mu_{A,2})} \right) \times \\
&\quad \sum_{s=1}^{N_S} A_{eff}(\theta, \alpha_{0,s}, \mu_{A,1}, \mu_{A,2})\,\sigma_{rms}(0, \theta, \alpha_{0,s}, \mu_c, \mu_a) \\
\delta z_{best,nested}(\theta, \mu_c, \nu_a, \mu_{A,1}, \mu_{A,2}) \; &= \; \left(\frac{1}{A_{eff,nested}(\theta, \mu_{A,1}, \mu_{A,2})} \right) \times \\
&\quad \sum_{s=1}^{N_S} A_{eff}(\theta, \alpha_{0,s}, \mu_{A,1}, \mu_{A,2})\,\delta z_{best}(\theta, \alpha_{0,s}, \mu_c, \mu_a) \\
\sigma_{rms,nested}(\theta, \mu_c, \mu_a, \mu_{A,1}, \mu_{A,2}) \; &= \; \left(\frac{1}{A_{eff,nested}(\theta, \mu_{A,1}, \mu_{A,2})} \right) \times \\
&\quad \sum_{s=1}^{N_S} A_{eff}(\theta, \alpha_{0,s}, \mu_{A,1}, \mu_{A,2})\,\sigma_{rms}(\delta z, \theta, \alpha_{0,s}, \mu_c, \mu_a)
\end{aligned}
\tag{36}
$$

11. EFFECT OF FIGURE ERRORS AND SURFACE ROUGHNESS

All equations have been for ideal figures. There will of course be figure errors and surface roughness resulting from the manufacturing process. Once the optics are manufactured and their figures measured, it is possible to evaluate the effects of sufficiently low amplitude and low frequency figure errors using the perturbation equation given in §4. It is also possible to evaluate the effects of surface roughness by convolving the predicted scattering wings with the predicted geometric core (see, *e.g.*, O'Dell *et al.*(1993)[3]). A rough estimate of the resulting dispersion follows from the root-summed-square of the contribution expected from manufacturing errors and the contribution expected from ideal figures. We have already mentioned that positioning focal plane detectors to follow as closely as physically possible the best focal surface leads to the best performance.

12. WIDE-FIELD X-RAY OPTICS WITH POLYNOMIAL COEFFICIENTS

In 1992 BBG[2] showed how by adding higher order polynomial terms to the Wolter I prescriptions, and hence giving up some on-axis spatial resolution, one can obtain prescriptions for reflecting surfaces that provide improved spatial resolution over a wide field-of-view (say $\sim 30 - 45$ arcmin). Such so-called polynomial optics, Eqs. (2) with non-zero δa_2, δa_3, δb_2, and δb_3, would be particularly useful for moderately deep to deep surveys, to be carried out by observatories such as for the proposed Wide-Field X-ray Telescope (WFXT) mission, and for solar x-ray observations.

12.1 Figures of merit

Many of the methods outlined above for Wolter I optics carry over to the case of polynomial optics. The most important new ingredient is how to choose the polynomial coefficients δa_2, δa_3, δb_2, and δb_3 so as to optimize the telescope performance for the chosen task. **In the case of a survey mission, for example, in principle one needs to create many simulated fields-of-view, provide accurate background, optical and detector performance models, and chose the polynomial coefficients and other mission parameters, subject to spacecraft and launch constraints, that unambiguously find the most sources.** In practise, one often designs the optics to optimize some suitably chosen merit function,[2, 4–8] and then test its performance using simulated fields-of-view. Methods for opimizing parameters can get complex, one example being the Markov-Chain-Monte-Carlo method.[7, 8]

One simple merit function stressing the minimization of the spatial resolution over some specified field-of-view, θ_{FOV}, is

$$M(\delta z, \theta_{FOV}, \delta a_2, \delta a_3, \delta b_2, \delta b_3) \equiv \frac{\int_0^{\theta_{FOV}} d\theta \sin\theta \; w(\theta) \; \sigma_{rms}^2(\delta z(\theta), \theta, \delta a_2, \delta a_3, \delta b_2, \delta b_3)}{\int_0^{\theta_{FOV}} d\theta \sin\theta \; w(\theta)} \tag{37}$$

The weighting factor, $w(\theta)$, is often set unity, but may also be set to the effective area, $A_{eff}(\theta)$. Since maximizing $A_{eff}(\theta)$ is also important for greatest sensitivity, one may want to consider the weighting factor $w(\theta) = 1/A_{eff}(\theta)$. However, normally the effective area is constrained mostly by mass considerations and spacecraft dimensions. Eq. (37) provides a spatial resolution figure of merit for the full field-of-view. **It is important to recognize that polynomial optics with ideal surfaces necessarily results in the loss of on-axis spatial resolution. Of course, manufacturing errors may make this point irrelevant.**

12.2 Design considerations

Since X-ray optics necessarily involve grazing incidence, effective area is built up by nesting many shells in each telescope module. This complicates the optimization procedure because:

1. The polynomial coefficients for each shell will be different, depending on the shell intersection radius, $r_{0,s}$, and segment length, ℓ_s. Here s denotes the shell number; our convention is to number the outermost shell as $s = 1$ with s increasing inward. Most telescope designs for wide-field surveys contain many shells (as many as 100), so this fact greatly complicates the optimization procedure. In practise one often optimizes the polynomial coefficients for each shell separately rather than all at once.

2. The best focal surfaces for the different shells will be different depending on the shell segment length, ℓ_s. Thus the optimization procedure should include the possibility that the segment lengths of the various shells are not all the same. The optimization procedure should also include placement of the focal plane detectors.

3. Constraints due to mass limits, spacecraft dimensions, and optic manufacturing limits normally lead to an envelope for the telescope optics prescribing values for the maxima and minima intersection radii and shell segment lengths $r_{0,max}$ and $r_{0,min}$, and ℓ_{max}, and ℓ_{min}.

For Wolter I optics, we see from Eq. (26) and the fact $\alpha_0 = (1/4)\ \tan^{-1}(r_0/f) \simeq r_0/4f$ (see Eq. (3)) that

$$\delta z_{best}(\theta, \alpha_0, \mu_{\delta z}) \simeq \mu_{\delta z}\ (1/8)\ \ell\ (\tan\theta\ /\ \tan\alpha_0)^2 \simeq \mu_{\delta z}\ 2f\ (\frac{\ell}{r_0^2})\ \theta^2 \tag{38}$$

Eq. (38) shows that Wolter I mirror shells satisfying the scaling relation

$$\ell_{max}\ \geq\ \ell_s = \ell_{max}\ (\frac{r_{0,s}}{r_{0,max}})^2\ \geq\ \ell_{min} \tag{39}$$

are expected to have approximately the same best focal surfaces. The difficulty with this scaling relationship is that for closely packed shells, necessary to build up effective area and to reduce or eliminate ghost rays due to single reflections, the limit on shell length is normally reached well before the intersection radius $r_{0,s}$ approaches the value $r_{0,min}$, thus limiting the effective area. In addition, of course, introduction of the higher order polynomial terms in the surface equations alter the shape of the best focal surface.

This suggests one should investigate scaling laws of the form

$$\ell_{max}\ \geq\ \ell_s = \ell_{max}\ (\frac{r_{0,s}}{r_{0,max}})^\nu\ \geq\ \ell_{min}. \tag{40}$$

If the available space is to be filled up with mirror shells, one has

$$\nu = \frac{\log(\ell_{max}/\ell_{min})}{\log(r_{0,max}/r_{0,min})}. \tag{41}$$

12.3 Symbolic optimization of a single shell using *Mathematica*©

We use the symbolic mathematics system *Mathematica*©[9] to get an expression for $M(\delta z, \theta_{FOV}, \delta a_2, \delta a_3, \delta b_2, \delta b_3)$ for given numerical values of r_0, f (and therefore α_0) and ℓ. This is done at each stage of the ray trace (intersection with mirror surface, computing normals and reflected direction vectors, etc.) by expanding the appropriate quantity to second order in δa_2, δa_3, δb_2, and δb_3, including cross-terms. For example, the x-component of the normal vector from the secondary (H) surface is computed in the form:

$$
\begin{aligned}
n_{x,H}\ =\ & n_{xH,0000}\ +\ \delta a_2\ n_{xH,1000}\ +\ \delta a_3\ n_{xH,0100}\ +\ \delta b_2\ n_{xH,0010}\ +\ \delta b_3\ n_{xH,0001} \\[2mm]
& +\ \delta a_2^2\ n_{xH,2000}\ +\ \delta a_3^2\ n_{xH,0200}\ +\ \delta b_2^2\ n_{xH,0020}\ +\ \delta b_3^2\ n_{xH,0002} \\[2mm]
& +\ \delta a_2\ \delta a_3\ n_{xH,1100}\ +\ \delta a_2\ \delta b_2\ n_{xH,1010}\ +\ \delta a_2\ \delta b_3\ n_{xH,1001} \\[2mm]
& +\ \delta a_3\ \delta b_2\ n_{xH,0110}\ +\ \delta a_3\ \delta b_3\ n_{xH,0101}\ +\ \delta b_2\ \delta b_3\ n_{xH,0011},
\end{aligned}
\tag{42}
$$

where each of the 15 $n_{xH,ijkl}$ is a number. The normal vectors and direction vectors are kept normalized to 1 to second order. Using *Mathematica*©, ray traces over a set Θ of θ values produces expressions for $a_{x,rms}, b_{x,rms}, c_{x,rms}, a_{y,rms}, b_{y,rms}, c_{y,rms}$ (see Eqs. (19) and (20)) in the form of Eq. (42). Then we construct an expression for $\sigma_{rms}^2(\delta z, \theta, \delta a_2, \delta a_3, \delta b_2, \delta b_3)$, which can be integrated over θ to produce an expression for $M(\delta z, \theta_{FOV}, \delta a_2, \delta a_3, \delta b_2, \delta b_3)$. Finally the merit function can be minimized by setting derivatives of $M(\delta z, \theta_{FOV}, \delta a_2, \delta a_3, \delta b_2, \delta b_3)$ to zero and solving for $\delta z, \delta a_2, \delta a_3, \delta b_2$, and δb_3.

Maintaining symbolic expressions as described above is time and memory consuming. Because of these limitations, we proceed in three steps. The first step is not a Monte-Carlo simulation at all. Instead, we choose rays spaced uniformly, but not randomly, over the full shell aperture. For each value for $\theta \leq \theta_{FOV}$, the subset of rays incident on the shell are traced through the optic.

The second step uses the results from the first in order to determine trial values for δz, δa_2, δa_3, δb_2, and δb_3 for the value of θ_{FOV}.

The third step is a full Monte-Carlo ray trace using these trial values with large numbers of rays per θ value. The only approximation in the full Monte-Carlo ray trace involves calculating the intersections with the mirror segments as if they were Wolter I optics, and then correcting the location of the intersections using a pertubation method. The effects of this single approximation are insignificant compared to those of, say, axial slope errors. This third step provides a full evaluation of the mirror shell performance for the given values of the polynomial coefficients.

12.4 Nested Mirror Shells

As previously remarked, X-ray telescopes usually involve nested mirror shells. For example the Chandra X-ray Observatory has 4 shells and ESA's XMM-Newton 58. Since the plate scale and best focal surface for each shell are different, it is not sufficient to optimize each shell separately. Assuming two coefficients per mirror segment, one would need to simultaneously optimize 16 and 232 coefficients for the two telescopes mentioned above. Thus far we have succeeded in optimizing polynomial coefficients for 7 mirror shells simultaneously. For telescopes with significantly larger number of shells, as currently proposed for WFXT, we have not so far found it possible to simultaneously optimize all the polynomial coefficients.

Instead, for a large number of mirror shells, we proceed as follows:

1. Given a focal length f, a desired field-of-view θ_{FOV}, values for $r_{0,max}$, $r_{0,min}$, ℓ_{max}, and ℓ_{min}, shell thickness t, the dependence of ℓ_s on $r_{0,s}$, and an algorithm for nesting the shells, we produce a Wolter I telescope design.

2. We select a subset of these shells, separated for example by a constant amount in $r_{0,s}$, for which we individually determine the polynomial coefficients δa_2, δa_3, δb_2, and δb_3 as described in §12.3.

3. The coefficients determined in this way form smooth functions of $r_{0,s}$, which we use to interpolate values for the polynomial coefficients of the intervening shells.

4. Now having polynomial coefficients for every shell, we run large numbers of rays in *Mathematica*© for each individual shell.

5. We use the results of these ray traces to evaluate the performance of the wide-field X-ray telescope we have constructed.

6. One way to improve spatial resolution performance is to shift the individual shells relative to each other by a small amount in the axial direction.[2,4,5]

7. A particularly effective way to improve spatial resolution performance is to tilt focal plane detectors so as to conform more closely to the best focal surface. This was done with the ACIS CCD detectors on the Chandra X-ray Observatory.

In this way, we are able to construct and fully evaluate a wide-field X-ray telescope consisting of nested polynomial optics, as are needed for the WFXT mission.

13. CONCLUDING REMARKS AND FUTURE WORK

We have outlined our methods for designing and evaluating wide-field X-ray telescopes for use on X-ray astronomy survey missions, with the proposed WFXT mission particularly in mind. Along the way we have extended to a broader domain the analytic approximations for Wolter I optics given in VanSpeybroeck and Chase.[1]

Our method for determining the polynomial coefficients for wide-field grazing incidence optics uses the symbolic mathematics system *Mathematica*©.[9] This method does not require a Monte-Carlo search of parameter space to find values for these coefficients. However, Monte-Carlo simulations are still required to fully evaluate

performance. Up to the present, we have been able to simultaneously optimize coefficients for up to seven mirror shells. Most designs require many more shells and for those cases we still must optimize each shell individually. The resulting differences in plate scale and best focal surface can be partially compensated for by varying the mirror segment length, ℓ_s, with intersection radius, $r_{0,s}$, adjusting the relative axial positions of the individual shells with respect to each other,[2,4,5] and tilting the focal plane detectors to most closely conform to the telescope's best focal surface.

In the future, we intend to:

1. Explore and evaluate wide-field designs for the WFXT mission.

2. Investigate designs satisfying the scaling relation Eq. (40) in order to reach the most favorable compromise between best spatial resolution over the field-of-view and largest effective area for the available physical volume.

3. Explore incorporating detector tilt into the optimization of polynomial coefficients.

4. Provide point-spread functions for individual designs in order to enable their evaluation using simulated fields-of-view.

ACKNOWLEDGMENTS

REFERENCES

[1] VanSpeybroeck, L. P. and Chase, R. C., "Design parameters of paraboloid-hyperboloid telescopes for x-ray astronomy," *Appl. Optics* **11**, 440–445 (1972).

[2] Burrows, C. J., Burg, R., and Giacconi, R., "Optimal grazing incidence optics and its application to wide-field x-ray imaging," *Ap, J.* **392**, 760–765 (1992).

[3] O'Dell, S. L., Elsner, R. F., Kolodziejczak, J. J., Weisskopf, M. C., Hughes, J. P., and VanSpeybroeck, L. P., "X-ray evidence for particulate contamination on the axaf veta-1 mirrors," in [*Multilayer and grazing incidence X-ray/EUV optics for astronomy and projection lithography*], Hoover, R. B. and II, A. B. C. W., eds., *Proc. SPIE* **1742**, 171–182 (1993).

[4] Conconi, P. and Campana, S., "Optimization of grazing incidence mirrors and its application to surveying x-ray telescopes," *Astron. Ap.* **372**, 1088–1094 (2001).

[5] Conconi, P., Pareschi, G., Campana, S., Chincarini, G., and Tagliaferri, G., "Wide-field x-ray imaging for future missions, including xeus," in [*Optics for EUV, X-ray, and Gamma-Ray Astronomy*], Citterio, O. and O'Dell, S. L., eds., *Proc. SPIE* **5168**, 334–345 (2004).

[6] Roming, P. W. A., Burrows, D. N., Garmire, G. P., Shoemaker, J. R., and Roush, W. B., "Grazing incidence optics for wide-field x-ray survey imaging: a comparison of optimization techniques," in [*X-Ray Optics, Instruments, and Missions III*], Truemper, J. E. and Aschenbach, B., eds., *Proc. SPIE* **4012**, 359–369 (2000).

[7] Roming, P. W. A., Liechty, J. C., Sohn, D. H., Shoemaker, J. R., Burrows, D. N., and Garmire, G. P., "Markov chain monte carlo algorithms for optimizing grazing incidence optics for wide-field x-ray survey imaging," in [*X-Ray Optics for Astronomy: Telescopes, Multilayers, Spectrometers, and Missions*], Gorenstein, P. and Hoover, R. B., eds., *Proc. SPIE* **4496**, 146–153 (2002).

[8] Roming, P. W. A., Liechty, J. C., Shoemaker, J. R., Sohn, D. H., Roush, W. B., Burrows, D. N., and Garmire, G. P., "Grazing incidence optics for wide-field x-ray survey imaging: A comparison of optimization techniques," *eprint arXiv:astro-ph/0406097* (2004).

[9] Wolfram, S., [*The Mathematica Book*], Wolfram Media, Inc., and Cambridge University Press, Champaign, Ill. and Cambridge, United Kingdom (2003).