Nonspecular x-ray scattering in a multilayer-coated imaging system

D. G. Stearns,[a] D. P. Gaines,[b] and D. W. Sweeney
Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, California 94550

E. M. Gullikson
Center for X-Ray Optics, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, California 94720

(Received 6 January 1998; accepted for publication 11 April 1998)

We present a rigorous theoretical treatment of nonspecular x-ray scattering in a distributed imaging system consisting of multilayer-coated reflective optics. The scattering from each optical surface is obtained using a vector scattering theory that incorporates a thin film growth model to provide a realistic description of the interfacial roughness of the multilayer coatings. The theory is validated by comparing calculations based on measured roughness to experimental measurements of nonspecular scattering from a Mo–Si multilayer coating. The propagation of the scattered radiation through the optical system is described in the context of transfer function theory. We find that the effect of nonspecular scattering is to convolve the image with a point spread function that is independent of the coherence of the object illumination. For a typical soft x-ray imaging system, the scattering within the image field from the multilayer coatings is expected to be slightly greater than for single surfaces (as normalized to the reflectivity). This is because the roughness of the coatings includes both replication of the substrate roughness and the intrinsic roughness of the multilayer growth process. Our analysis indicates that the current multilayer coating technology is capable of producing soft x-ray imaging systems that have acceptably low levels of scattering, provided that the optical substrates are sufficiently smooth. [S0021-8979(98)02214-2]

I. INTRODUCTION

Advancement in areas such as extreme ultraviolet (EUV) lithography[1] and x-ray astronomy[2] are spurring dramatic improvement in the performance of optical imaging systems for the soft x-ray regime $(1<\lambda<100\ \text{nm})$. The ultimate goal is to achieve high throughput with resolution near the diffraction limit. This requires the use of all reflective, distributed imaging systems working near normal incidence. However, the normal incidence reflectivity of all materials is very low at soft x-ray wavelengths. The problem is overcome by coating the optical surfaces with multilayer films, which increases the reflectivity by several orders of magnitude in comparison to a single surface. There are yet many potential problems that can degrade the image formation and limit the resolution. One of the most important of these is the nonspecular scattering from the multilayer coatings.

The ultimate resolution of a soft x-ray imaging system depends in detail on the nonideal nature of the optical substrate and the interfaces in the multilayer coatings. These structures are imperfect at all spatial frequencies. The errors at frequencies less than $\sim 10\ \text{cm}^{-1}$ are called figure error and are treated deterministically. The figure errors produce an aberration of the image that can be calculated using ray tracing techniques. The errors at higher frequency are called roughness (or surface finish) and are usually treated statistically. The effect of the roughness is to remove intensity from the image (the specular field) and scatter it throughout the image field. This nonspecular scattering is problematic for two reasons: (1) it decreases the useful throughput of the optical system and, (2) it produces a background halo which reduces the contrast of the image.

Soft x-ray imaging systems are particularly susceptible to nonspecular scattering. The first reason is that the optics are by necessity all reflective. The second reason is the well-known λ^{-4} dependence of the cross section for dipole scattering. For a given interfacial roughness, the nonspecular scattering increases rapidly with decreasing wavelength. The implications for high resolution imaging are clear: the optical substrates and multilayer coatings must be very smooth to avoid significant scattering. But how smooth is smooth enough? The answer to this question is of both fundamental and practical interest. Fundamentally, there has been until now a general lack of understanding of the effects of scattering in a multilayer-coated imaging system. From a practical perspective, the time and cost of producing figured optical substrates increases dramatically with the smoothness of the surface. An accurate model of scattering is needed to derive realistic specifications for surface finish that can be used as a guideline for manufacturing precision optical components.

The problem of nonspecular scattering of x rays from multilayer films has been addressed by many authors in recent years.[3–7] It is now understood that the scattering from a multilayer film is fundamentally different than the scattering from a single rough surface. The multilayer scattering is characterized by strong interference effects, due to the cor-

[a]Current address: OS Associates, 1174 Castro St., Suite 250, Mountain View, CA 94040; electronic mail: healthst@ricochet.net
[b]Current address: Ultratech Stepper, 3050 Zanker Rd., San Jose, CA 95134.

relation of the roughness of the different interfaces. Resonances in the scattering were theoretically predicted[3,4] and have been experimentally observed.[4,7–10] Dynamical effects arising from the multiple specular reflection and extinction of the scattered radiation can also be important,[5] particularly in the vicinity of a specular Bragg peak. We will show that these unique characteristics of the scattering from multilayer coatings have important consequences in an imaging system.

There have been previous attempts to address the problem of nonspecular x-ray scattering in an imaging system. Church and Takacs[11] have considered the scattering in a simple imaging system consisting of a single reflecting surface operating near normal incidence. Here the multilayer coating is treated as a single rough surface. Harvey[12] has discussed scattering in a distributed imaging system, and has proposed an *ad hoc* method for incorporating the effects of multilayer coatings. More recently, Singh *et al.*[13] have applied a Monte Carlo ray tracing technique to simulate nonspecular scattering in a Schwarzschild imaging system. In spite of this previous work, we believe that the current understanding of scattering in a soft x-ray imaging system is incomplete. In particular, these important issues have yet to be adequately addressed:

(1) Owing to the high spatial frequencies characteristic of roughness there is significant diffraction of the scattered field as it propagates between the optical surfaces that produce the scattering and the exit pupil. How does this diffraction affect the pupil function and the image formation process in the context of the transfer function theory of imaging?

(2) It is well known that the coherence state of the object illumination is an important parameter in the imaging process. What is the relationship between the coherence state and the effects of scattering on the image process?

(3) The scattering from a multilayer coating occurs over a large number of interfaces throughout the volume of the film. The magnitude and distribution of scattering depends on the detailed structure of the roughness of the interfaces. What is a correct description of the roughness of a multilayer coating and how is it related to the roughness of the substrate?

(4) What are the characteristics of the scattering from a multilayer coating, particularly in the vicinity of the Bragg peak, and how do they differ from that of a single rough surface? In particular, the description of the scattering process must include interference effects (from the correlation of the interfacial roughness), dynamical effects (multiple reflection of the incident and scattered fields), and extinction of the scattered radiation in the film.

To investigate these issues we present a comprehensive theoretical description of scattering in a soft x-ray imaging system, which includes rigorous treatments of the image formation process, the roughness of the multilayer coatings, and the scattering process. In the following section we derive, following Born and Wolf,[14] a very general expression for the transfer function that relates the mutual intensity at the object and image planes. In Sec. III we consider the scattering in a distributed imaging system having single reflecting surfaces. We explicitly include the diffraction of the scattered field and obtain an effective pupil function for the distributed imaging system. The intensity distribution at the image plane is derived in Sec. IV. We show that the effect of scattering is to convolve the image with a point spread function independent of the coherence state of the object illumination. In Sec. V we consider the case of multilayer coatings, and we relate the point spread function to the angular scattering distributions from the multilayer coatings. The calculation of these scattering distributions is the topic of the next two sections. In Sec. VI we review a thin film growth model that has been previously developed[15] to describe the interfacial roughness in multilayer coatings, and we apply this model to measurements of roughness in a Mo–Si multilayer film. The multilayer scattering problem is treated in Sec. VII using an existing scattering theory extended to include dynamical effects in the scattered field. To validate the scattering theory, we compare calculations to experimental measurements of nonspecular scattering from a high-performance Mo–Si multilayer coating. In Sec. VIII we evaluate the scattering in a soft x-ray imaging system in terms of conventional performance parameters, namely the point spread function and the optical transfer function. We close our discussion by modeling the effects of scattering in a hypothetical imaging system designed for EUV lithography.

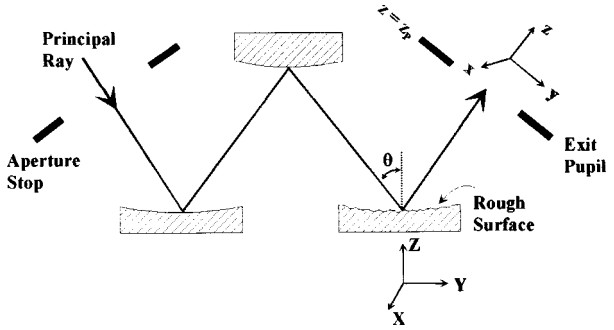


FIG. 1. Schematic diagram of a distributed imaging system showing the trajectory of the principal ray. The (X,Y,Z) coordinate system of the rough surface is related to the (x,y,z) coordinate system of the exit pupil through a rotation of angle θ. The exit pupil is located a distance $z=z_P$ from the rough surface.

Throughout the course of the theoretical development we make several critical assumptions and approximations. It is important to observe the restrictions and constraints imposed by these approximations whenever the theory is applied to specific imaging systems. The approximations are enumerated as they appear in the text and are summarized in Appendix A.

II. DERIVATION OF THE TRANSFER FUNCTION

Consider a distributed imaging system consisting of N reflecting surfaces and having a real aperture stop as shown in Fig. 1. The Cartesian coordinates of the object plane and the image plane are $\mathbf{S}_0$ and $\mathbf{s}_1$, respectively. We introduce the scale normalized coordinates for the object plane, $\mathbf{s}_0 = M\mathbf{S}_0$, where M is the lateral magnification of the optical system. This allows us to describe an object point and its Gaussian image point by the same coordinate values. The aperture stop of the imaging system limits the angular dis-

persion of rays through the system. For a point object, the ray that intersects the center of the aperture stop is called the principal ray. The image of the aperture stop by the part of the optical system which follows it is called the exit pupil. The exit pupil is located a distance R from the image plane. The amplitude of the radiation field at the plane of the exit pupil is called the complex pupil function, $G(\mathbf{s})$.

Let the object be illuminated by quasimonochromatic radiation of wavelength λ. The mutual intensity functions in the object and image planes are $J_0(\mathbf{s}_0,\mathbf{s}_0')$ and $J_1(\mathbf{s}_1,\mathbf{s}_1')$, respectively. Following Born and Wolf[14] the Fourier transforms of the mutual intensity functions, $J_0(\mathbf{f},\mathbf{f}')$ and $J_1(\mathbf{f},\mathbf{f}')$, are related by

$$J_1(\mathbf{f},\mathbf{f}')=G(\lambda R\mathbf{f})G^*(-\lambda R\mathbf{f}')J_0(\mathbf{f},\mathbf{f}'), \tag{1}$$

where $\mathbf{s}=\lambda R\mathbf{f}$ are the coordinates of a point in the plane of the exit pupil. This description of the transfer of the mutual intensity function through the imaging system is only valid within two important approximations. These are:

(a) The angle between the principal ray and any other ray that propagates through the imaging system is small. Specifically, if we denote the angle as φ, then the approximation is

$$\sin^2\varphi \ll 1. \tag{2}$$

We call this the ''small angle approximation.''

(b) For a point object, the pupil function is independent of the location of the point in the object field. In this case the point spread function is independent of the position of the Gaussian image point, and the system is called ''isoplanatic.'' In practice, the assumption of isoplanacity restricts the applicability of the transfer function formalism to objects of small spatial extent.

The effect of scattering from roughness at the optical surfaces is to modify the pupil function in a simple way. Let $G_0(\mathbf{s})$ be the pupil function for the optical system without rough surfaces. Here G_0 contains all of the standard deterministic information about the imaging system such as aberrations. We will show in Sec. II that the effect of the roughness can be represented as a pure phase modulation:

$$G(\mathbf{s})=G_0(\mathbf{s})\exp[i\Phi(\mathbf{s})] \tag{3}$$

where $\Phi(\mathbf{s})$ is a function directly related to the structure of the rough surfaces. Then substituting (3) into (1) we obtain

$$\langle J_1(\mathbf{f},\mathbf{f}')\rangle = G_0(\lambda R\mathbf{f})G_0^*(-\lambda R\mathbf{f}') \times\langle\exp[i(\Phi(\lambda R\mathbf{f})-\Phi^*(-\lambda R\mathbf{f}'))]\rangle J_0(\mathbf{f},\mathbf{f}'). \tag{4}$$

The angular brackets denote taking an ensemble average over many configurations of the rough surface. In practice this is realized by the breaking up of the coherence of the illumination of the surface due to the finite size of the object. For a coherently illuminated object of size L and a distance D between the object plane and the optical surface, the size of the patch on the surface over which the illumination is coherent is approximately $\lambda D/2L$. For typical sources of finite size there will be many of these patches within the area illuminated on the optical surface. If the surface is sufficiently ergodic then each patch represents a different configuration of roughness, and the ensemble average can be replaced by an average over the illuminated surface area.

III. EFFECTIVE PUPIL FUNCTION

To proceed further it is necessary to derive an expression for the pupil function $G(\mathbf{s})$ that includes scattering from the optical surfaces. For the moment let the optical surfaces be single surfaces (not coated with multilayer films), and assume that all but one of the surfaces are perfectly smooth, as indicated in Fig. 1. Since our goal is to determine the effect of scattering on the pupil function, we can neglect the curvature of the incident wavefront and the optical surface (these will be included later). Then let the specular field be a plane wave, $\hat{e}e^{ik\hat{n}\cdot\mathbf{x}}$, of unit amplitude and polarization $\hat{e}$ incident onto the rough surface with an angle θ (measured with respect to the normal). The field is reflected by the surface and propagates to the plane of the exit pupil. Choosing the plane of the exit pupil to be perpendicular to the principal ray, we find that θ is also the angle between the normal to the exit pupil and the normal to the rough surface.

An expression for the field scattered by the rough surface has been derived in a previous paper.[16] The results are valid under the following approximation:

(c) The scattering is weak so that multiple scattering and shadowing effects can be neglected. This is called the ''Born approximation.'' This approximation is generally valid for x-ray wavelengths at angles of incidence away from the critical angle for total external reflection.

The component of the scattered reflected field $\mathbf{E}^R(\mathbf{x})$ having polarization $\hat{a}$ (S or P type) can be written as:

$$\hat{a}\cdot\mathbf{E}^R(\mathbf{x})=\frac{\Delta}{2}(\hat{a}\cdot\hat{e})\int\int\left(\int\int \exp(iq_XX)\exp(iq_YY) \times\frac{\exp[iq_ZH(X,Y)]}{q_Z(q_Z+kn_Z)}\,dXdY\right) \times\exp(ik\hat{m}\cdot\mathbf{x})dm_Xdm_Y. \tag{5}$$

Here $\Delta=1-\epsilon$, where ϵ is the dielectric constant of the surface material, $\mathbf{q}=k(\hat{m}-\hat{n})$ is the change of momentum of the x-ray photon and $H(X,Y)$ is the surface height function describing the roughness of the surface.

The expression (5) for the scattered field has a straightforward physical interpretation: it is simply an expansion of the scattered field using the plane waves $\exp(ik\hat{m}\cdot\mathbf{x})$ as a basis set. The quantity in brackets is the scattering amplitude of the plane wave mode $\hat{m}$. Note that there are two different spatial coordinate systems in Eq. (5). The relationship between these coordinate systems is illustrated in Fig. 1. The coordinate system $(\hat{X},\hat{Y},\hat{Z})$ is defined such that $\hat{Z}$ is normal to the plane of the rough optical surface and the Y-Z plane is the plane of incidence. The coordinate system $(\hat{x},\hat{y},\hat{z})$ has $\hat{z}$ normal to the plane of the exit pupil. The transformation between the two coordinate systems is a rotation through an

angle θ about the $\hat{x}=\hat{x}'$ axis. The momentum transfer vector **q** in the two coordinate systems is related according to

$$\begin{aligned} q_X &= q_x \\ q_Y &= \cos\ \theta q_y + \sin\ \theta q_z \cong \cos\ \theta q_y + 2k\ \cos^2\ \theta\ \sin\ \theta \\ q_Z &= -\sin\ \theta q_y + \cos\ \theta q_z \cong 2k\ \cos\ \theta. \end{aligned} \tag{6}$$

These expressions are valid within the small angle approximation (a), which requires that q_X and q_Y be small. Noting that

$$\hat{n} = \sin\ \theta\hat{Y} - \cos\ \theta\hat{Z} = \sin 2\theta\hat{y} - \cos 2\theta\hat{z} \tag{7}$$

we rearrange Eq. (5) to get

$$\begin{aligned} \hat{a}\cdot\mathbf{E}^R(\mathbf{x}) = (1/4\pi^2) r^{\mathrm{SP}}\ \cos\ \theta\ \exp(iky\ \sin 2\theta)\exp(-ikz\ \cos 2\theta) \int\int\int\int\ \exp(i\mathbf{q}\cdot\mathbf{x})\exp(-iq_xX) \\ \times\exp(-i\ \cos\ \theta q_yY)\exp(-2ik\ \cos^2\ \theta\ \sin\ \theta Y)\exp[-2ik\ \cos\ \theta H(X,Y)]dXdYdq_xdq_y \end{aligned} \tag{8}$$

where $r^{\mathrm{SP}}=\Delta(\hat{a}\cdot\hat{e})/4\cos^2\theta$ is the specular reflectance amplitude from an ideally smooth surface.

Next we propagate the reflected field to the plane of the exit pupil at position $z=z_P$. The field in this plane is the complex pupil function $G(\mathbf{s})$ where $\mathbf{s}=s_x\hat{x}+s_y\hat{y}$. Then using Eq. (8) the pupil function becomes

$$\begin{aligned} G(\mathbf{s}) = (1/4\pi^2) r^{\mathrm{SP}}\ \cos\ \theta\ \exp(ik\ \sin 2\theta s_y)\exp(-ik\ \cos 2\theta z_P) \\ \times\int\int\int\int\ \exp(iq_xs_x)\exp(iq_ys_y)\exp(iq_zz_P) \\ \times\exp(-iq_xX)\exp(-i\ \cos\ \theta q_yY)\exp(-2ik\ \cos^2\ \theta\ \sin\ \theta Y)\exp[-2ik\ \cos\ \theta H(X,Y)]dXdYdq_xdq_y. \end{aligned} \tag{9}$$

This is the correct expression for the pupil function that includes the (significant) diffraction of the scattered field on its way to the exit pupil. However, the propagation of the mutual intensity, as described in Eq. (1), requires the determination of the transfer function $\langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle$. Substituting from Eq. (9) we have

$$\begin{aligned} \langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle = (1/16\pi^4) R_{\mathrm{SP}}\ \cos^2\ \theta\ \exp[ik\ \sin 2\theta(s_y+s_y')] \\ \times\int\int\int\int\ \exp(iq_xs_x)\exp(iq_ys_y)\exp(iq_zz_P)\exp(iq_x's_x')\exp(iq_y's_y')\exp(-iq_z'z_P) \\ \times\int\int\int\int\ \langle\exp\{-2ik\ \cos\ \theta[H(X,Y)-H(X',Y')]\}\rangle\exp(-iq_xX)\exp(iq_x'X') \\ \times\exp[-i\ \cos\ \theta(q_yY-q_y'Y')]\exp[-2ik\ \cos^2\ \theta\ \sin\ \theta(Y-Y')]dXdYdX'dY'dq_xdq_ydq_x'dq_y'. \end{aligned} \tag{10}$$

We make the following assumptions about the statistical properties of the roughness:

(d) The surface height $H(X,Y)$ is a Gaussian random variable, is stationary and is ergodic in the sense that the ensemble average can be replaced by an average over the illuminated surface area.

The assumption that H is stationary leads to an important simplification: the quantity in the brackets $\langle\ \rangle$ depends only on the separation of the points $U=X-X'$ and $V=Y-Y'$. In particular, if we define

$$F(U,V)\equiv\langle\exp\{-2ik\ \cos\ \theta[H(X,Y)-H(X',Y')]\}\rangle \tag{11}$$

then the inner integral in Eq. (10) becomes

$$\begin{aligned} \int\int\int\int F(U,V)\exp(-iq_xU)\exp(-i\ \cos\ \theta q_yV)\exp(-2ik\ \cos^2\ \theta\ \sin\ \theta V)\exp[-i(q_x-q_x')X'] \\ \times\exp[-i\ \cos\ \theta(q_y-q_y')Y']dX'dY'dUdV \\ = \cos^{-1}\ \theta\delta(q_x-q_x')\delta(q_y-q_y')\int\int F(U,V)\exp(-iq_xU)\exp(-i\ \cos\ \theta q_yV)\exp(-2ik\ \cos^2\ \theta\ \sin\ \theta V)dUdV. \end{aligned} \tag{12}$$

Substituting (12) into (10) we obtain:

$$\langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle = (1/4\pi^2)R^{\mathrm{SP}}\cos\theta\exp[ik\sin 2\theta(s_y+s_y')]\int\int\int\int F(U,V)\exp(-2ik\cos^2\theta\sin\theta V)$$
$$\times\exp[iq_x(s_x+s_x'-U)]\exp[iq_y(s_y+s_y'-\cos\theta V)]dq_x dq_y dU dV$$
$$= R^{\mathrm{SP}}\cos\theta\int\int F(U,V)\,\delta(s_x+s_x'-U)\,\delta(s_y+s_y'-\cos\theta V)dU dV = R^{\mathrm{SP}}F\left(s_x+s_x',\frac{s_y+s_y'}{\cos\theta}\right). \quad (13)$$

We find that the transfer function $\langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle$ does not depend on z_P, the distance between the optical surface and the exit pupil. In other words, the diffraction of the scattered field between the optical surface and the exit pupil does not contribute to the image formation. This surprising conclusion is due to the process of averaging over an ensemble of configurations corresponding to a stationary distribution. Diffraction effects must cancel out in the ensemble average because all points on the optical surface are equivalent. Consequently, for the purpose of evaluating $\langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle$, we can map the rough surface directly onto the exit pupil as if there was no separation ($z_P=0$). The effect of the roughness on the pupil function can be represented as a simple phase modulation

$$G(\mathbf{s}) = G_0(\mathbf{s})\exp\left[-2ik\cos\theta H\left(\frac{s_x}{\alpha_x},\frac{s_y}{\alpha_y\cos\theta}\right)\right]. \quad (14)$$

Here we have reintroduced the effects of the wavefront curvature and surface figure in the factor $G_0(\mathbf{s})$, which is the pupil function in the absence of scattering. The factors $\alpha_x = s_x/X$ and $\alpha_y = s_y/(Y\cos\theta)$ account for the change of scale between the optical surface and the exit pupil, as determined, for example, by the change in separation of the extrema rays. The scaling relationship for s_y includes a factor of $\cos\theta$ to account for the angle of inclination of the optical surface with respect to the plane of the exit pupil. We emphasize that Eq. (14) is not the correct pupil function for any particular configuration of surface roughness. It can be used, however, as an *effective* pupil function, in the sense that it produces a correct result in the calculation of image formation when the quantity $\langle G(\mathbf{s})G^*(-\mathbf{s}')\rangle$ is averaged over a statistically random and *stationary* distribution of configurations.

At first glance it might appear that Eq. (14) cannot correctly account for all of the radiation scattered within the imaging system. For example, radiation scattered at large angles by the first optical surface in the distributed system will not pass through the subsequent optics and reach the exit pupil. Yet within the context of the transfer function theory, the roughness of each optical surface is mapped onto the exit pupil and all of the scattering occurs at the exit pupil. The resolution of this apparent inconsistency is to understand that any scattered radiation that in reality does not reach the exit pupil, is scattered outside of the image field in the transfer function description; the transfer function theory correctly accounts for all scattering that intersects the image field. This can be illustrated using the following argument. Consider the ideal imaging system that images a point object at $\mathbf{s}_0$ to a point $\mathbf{s}_1$ in the image plane. The scattering from a rough optical surface into a particular nonspecular direction $\hat{m}$ will be imaged by the subsequent optics to a different point $\mathbf{s}_1'$ in the image plane. Tracing rays back through the imaging system from the image point $\mathbf{s}_1'$, the propagation through the imaging system of the radiation scattered into direction $\hat{m}$ is found to be equivalent to the imaging of a point source at the conjugate position $\mathbf{s}_0'$ in the object field. Consequently, the propagation of the scattered radiation that intersects the image field is equivalent to the propagation of radiation from an extended object in the absence of scattering, a process that is correctly described by transfer function theory.

Thus far we have considered the scattering from a single rough surface. We next consider an imaging system where each of the N reflecting surfaces has a roughness described by a unique height function $H_n(X_n,Y_n)$. We assume that the roughness of the different surfaces each satisfies the conditions (d) for stationary and ergodic distributions, and are mutually statistically independent. The effective pupil function for the entire imaging system can be derived by imagining that we ''turn on'' the roughness of each surface sequentially and apply Eq. (14) iteratively. In particular, the pupil function for n rough surfaces becomes G_0 for the case of $n+1$ rough surfaces, etc. Then the effective pupil function for the entire distributed system is

$$G(\mathbf{s}) = G_0(\mathbf{s})\exp[i\Phi(\mathbf{s})]$$
$$= G_0(\mathbf{s})\exp\left[-2k\sum_{n=1}^{N}\cos\theta_n H_n\left(\frac{s_x}{\alpha_{xn}},\frac{s_y}{\alpha_{yn}\cos\theta_n}\right)\right]. \quad (15)$$

IV. DESCRIPTION OF THE IMAGE

Having derived the effective pupil function, we now apply the transfer function formalism of Eq. (4) to obtain a description of the image. We begin by evaluating the quantity

$$\langle\exp[i(\Phi(\mathbf{s})-\Phi^*(-\mathbf{s}')]\rangle$$
$$= \left\langle\exp\left\{-2ik\sum_{n=1}^{N}\cos\theta_n\left[H_n\left(\frac{s_x}{\alpha_{xn}},\frac{s_y}{\alpha_{yn}\cos\theta_n}\right)\right.\right.\right.$$
$$\left.\left.\left. - H_n\left(-\frac{s_x'}{\alpha_{xn}},-\frac{s_y'}{\alpha_{yn}\cos\theta_n}\right)\right]\right\}\right\rangle. \quad (16)$$

It can be shown generally[17] that for Gaussian random variables B_m the ensemble average reduces to

$$\left\langle\exp\left(\sum_n B_n\right)\right\rangle = \exp\left[\frac{1}{2}\left\langle\left(\sum_n B_n\right)^2\right\rangle\right]. \quad (17)$$

We have assumed that the roughness of the different optical surfaces is completely uncorrelated, requiring that $\langle H_n H_m \rangle = 0$ for $n \neq m$. Then we obtain

$$\langle \exp[i(\Phi(\mathbf{s}) - \Phi^*(-\mathbf{s}')]\rangle = \exp\left(-4k^2 \sum_{n=1}^{N} \cos^2\theta_n \langle H_n^2(X_n, Y_n)\rangle + 4k^2 \sum_{n=1}^{N} \cos^2\theta_n \langle H_n(X_n, Y_n) H_n(-X_n', -Y_n')\rangle \right), \tag{18}$$

where $(X_n, Y_n) = [(s_x/\alpha_{xn})(s_y/\alpha_{yn} \cos\theta_n)]$ are the spatial coordinates on the nth optical surface. The assumption that H is stationary requires that the quantity $\langle H^2(X,Y)\rangle$ be independent of (X,Y) and the quantity $\langle H(X,Y)H(-X', -Y')\rangle$ depends only on the separation of the points $U = X + X'$ and $V = Y + Y'$. These statistical quantities are conventionally defined in terms of the surface height variance σ^2, where σ is called the root-mean-square (rms) roughness, and the height–height autocorrelation function C, given by

$$\sigma^2 \equiv \frac{1}{A} \int H^2(X,Y)\, dX\, dY \tag{19}$$

and

$$C(U,V) = \frac{1}{A} \int H(X,Y) H(U+X, V+Y)\, dX\, dY. \tag{20}$$

Here we have replaced the ensemble average with an average over the area, A, of the illuminated surface. We write

$$\langle \exp[i(\Phi(\mathbf{s}) - \Phi^*(-\mathbf{s}')]\rangle = \exp\left(-4k^2 \sum_{n=1}^{N} \cos^2\theta_n \sigma_n^2 + 4k^2 \sum_{n=1}^{N} \cos^2\theta_n C_n(\mathbf{u}_n) \right), \tag{21}$$

where

$$\mathbf{u}_n \equiv \frac{s_x + s_x'}{\alpha_{xn}}\,\hat{x} + \frac{s_y + s_y'}{\alpha_{yn} \cos\theta_n}\,\hat{y}.$$

The final step for determining the intensity distribution in the image plane is to take the inverse Fourier transform of $\langle J_1(\mathbf{f}, \mathbf{f}')\rangle$. Applying the convolution theorem to Eq. (4) we obtain

$$\langle J_1(\mathbf{s}_1, \mathbf{s}_1')\rangle = J_1^0(\mathbf{s}_1, \mathbf{s}_1') * \exp\left(-4k^2 \sum_n \cos^2\theta_n \sigma_n^2 \right) \times \int\int \exp\left[4k^2 \sum_n \cos^2\theta_n C_n(\mathbf{u}_n) \right] \times \exp(2\pi i \mathbf{s}_1 \cdot \mathbf{f}) \exp(2\pi i \mathbf{s}_1' \cdot \mathbf{f}')\, d\mathbf{f}\, d\mathbf{f}', \tag{22}$$

Here

$$J_1^0(\mathbf{s}_1, \mathbf{s}_1') = \int\int G_0(\lambda R \mathbf{f}) G_0^*(-\lambda R \mathbf{f}') J_0(\mathbf{f}, \mathbf{f}') \times \exp(2\pi i \mathbf{s}_1 \cdot \mathbf{f}) \exp(2\pi i \mathbf{s}_1' \cdot \mathbf{f}')\, d\mathbf{f}\, d\mathbf{f}', \tag{23}$$

is the mutual intensity at the image plane in the absence of roughness. Since C_n is dependent only on the difference vector $\mathbf{u}_n$, the double integral in (22) collapses to yield,

$$\langle J_1(\mathbf{s}_1, \mathbf{s}_1')\rangle = J_1^0(\mathbf{s}_1, \mathbf{s}_1') * \delta(\mathbf{s}_1 - \mathbf{s}_1') \times \exp\left[-4k^2 \sum_n \cos^2\theta_n \sigma_n^2 \right] \times \int\int \exp\left[4k^2 \sum_n \cos^2\theta_n C_n(\mathbf{u}_n) \right] \times \exp(2\pi i \mathbf{s}_1 \cdot \mathbf{v})\, d\mathbf{v}, \tag{24}$$

where $\mathbf{v} \equiv \mathbf{f} + \mathbf{f}'$. The intensity distribution is obtained by setting $\mathbf{s}_1 = \mathbf{s}_1'$. Then,

$$\langle I_1(\mathbf{s}_1)\rangle = I_1^0(\mathbf{s}_1) * \exp\left[-4k^2 \sum_n \cos^2\theta_n \sigma_n^2 \right] \times \int\int \exp\left[4k^2 \sum_n \cos^2\theta_n C_n(\mathbf{u}_n) \right] \times \exp(2\pi i \mathbf{s}_1 \cdot \mathbf{v})\, d\mathbf{v}. \tag{25}$$

Further simplification is possible if each surface is sufficiently smooth to satisfy the following condition:

(e) The deviations of the surface height $H_n(X_n, Y_n)$ from the ideally smooth surface are small compared to the radiation wavelength such that $2k \cos\theta_n H_n(X_n, Y_n) \ll 1$ for all X_n, Y_n. We call this the ''small roughness approximation.'' A necessary consequence of the small roughness approximation is that the power scattered into the nonspecular field is small compared to the specularly reflected power, a condition that is implicitly satisfied by high-performance optics.

For a randomly rough surface the autocorrelation function C_n will have a maximum value of σ_n^2. Then invoking the small roughness approximation, we can expand the exponential in the integrand in Eq. (25) to obtain

$$\langle I_1(\mathbf{s}_1)\rangle = I_1^0(\mathbf{s}_1) * \exp\left(-4k^2 \sum_n \cos^2\theta_n \sigma_n^2 \right) \times \left[\delta(\mathbf{s}_1) + \frac{4k^2}{\lambda^2 R^2} \sum_n \alpha_{xn}\alpha_{yn} \cos^3\theta_n \times \mathrm{PSD}_n\left(\frac{\alpha_{xn} s_{1x}}{\lambda R}, \frac{\alpha_{yn} \cos\theta_n s_{1y}}{\lambda R} \right) \right]. \tag{26}$$

Here PSD_n is the (power spectral density) of the nth surface, which is the Fourier transform of the autocorrelation function.

The result shown in Eq. (26) has a simple physical interpretation. The effect of the surface roughness on the formation of an image is to convolve the image that would exist in the absence of scattering, I_1^0, with a point spread function due to scattering

$$\langle I_1(\mathbf{s}_1)\rangle = I_1^0(\mathbf{s}_1) * \kappa \mathrm{PSF}^{\mathrm{sc}}(\mathbf{s}_1). \tag{27}$$

The point spread function, PSF^{sc}, consists of two parts:

$$\mathrm{PSF}^{sc}(\mathbf{s}_1)=\frac{1}{\kappa}\left[S\delta(\mathbf{s}_1)+\sum_n \frac{dP_n(\mathbf{s}_1)}{d\mathbf{s}_1}\right]. \tag{28}$$

The first term is the contribution from specular reflection. It is simply a delta function reduced by the Strehl factor, S, where:

$$S=\prod_n S_n=\prod_n \exp(-4k^2\cos^2\theta_n\sigma_n^2). \tag{29}$$

The Strehl factor represents the power removed from the specular field due to nonspecular scattering. The second term in the PSF corresponds to the power per unit area scattered into the image plane given a point source in the object plane, where

$$\frac{dP_n(\mathbf{s}_1)}{d\mathbf{s}_1}=\frac{4k^2S}{\lambda^2R^2}\alpha_{xn}\alpha_{yn}\cos^3\theta_n \times \mathrm{PSD}_n\left(\frac{\alpha_{xn}s_{1x}}{\lambda R},\frac{\alpha_{yn}\cos\theta_n s_{1y}}{\lambda R}\right). \tag{30}$$

The factor κ accounts for any loss of integrated image intensity due to, for instance, scattering outside of the image field or absorption in the multilayer coating. It is defined by

$$\kappa\equiv\int\int\left[S\delta(\mathbf{s}_1)+\sum_n \frac{dP_n(\mathbf{s}_1)}{d\mathbf{s}_1}\right]d\mathbf{s}_1 \tag{31}$$

which ensures that the PSF integrates to unity.

We note that the effect of scattering on the image formation *is independent* of the coherence of the illumination of the object. All of the coherence effects in Eq. (29) are contained in the image $I_1^0(\mathbf{s}_1)$ formed in the absence of scattering. The effect of scattering is to convolve this image with a PSF^{sc} that is independent of the coherence state. This interesting result is due to the process of taking an ensemble average over the statistical distribution of configurations of surface roughness. The coherence state of the source determines the specific illumination of the rough surface. However, taking the ensemble average makes each surface point equivalent in terms of its contribution to scattering, which means that the specific illumination pattern cannot effect the scattering distribution.

V. EXTENSION TO MULTILAYER COATINGS

Thus far we have assumed that the optical surfaces are single reflecting surfaces. In fact, these surfaces are coated with multilayer films to produce efficient reflectivity at soft x-ray wavelengths. Although the reflectance from a single surface is small at these wavelengths, the reflections from each interface in the multilayer coating add constructively to produce a large total specular reflectance. For example, for a single Mo surface the normal incidence reflectance of soft x rays at a wavelength of 13 nm is $\sim$0.15%; the reflectance from a Mo–Si multilayer coating consisting of 40 bilayers of period 6.8 nm is $\sim$70% (see Sec. VII).

Multilayer coatings typically have roughness at the layer boundaries, due to both replication of the substrate roughness and roughness introduced during the film growth process. The incident specular field scatters at each of the multilayer interfaces, and these contributions add coherently to produce a total scattered field. Hence, in principal, the calculation of the scattered field from a multilayer coating requires knowledge of the PSD of each of the interfaces, as well as the correlation of the roughness between every pair of interfaces. It is intractable to measure the roughness of each interface in each coating of an imaging system. We have found that an effective approach to this very complicated problem is to describe the multilayer interface structure in terms of a simple multilayer growth model.[15] The model provides a straightforward method for describing all of the detailed structural information required by the scattering theory in terms of a small set of fundamental parameters. A description of the multilayer growth model is presented in the next section.

The nonspecular scattering from a multilayer film is fundamentally different than the scattering from a single surface. This is because the scattered field is the coherent superposition of the fields scattered by each of the interfaces. Just as the specular reflectance is a resonance property of the coating, the nonspecular scattering exhibits resonance behavior that is absent in the case of a single surface. The phenomenon of resonant nonspecular scattering has been discussed previously.[3,5] The scattering is enhanced whenever the momentum change normal to the film is equal to a reciprocal lattice vector *and* when the structure of the interfaces is correlated from layer to layer (i.e., is at least partially conformal). This has important implications for the scattering in an imaging system. Since the specular field is near the center of the Bragg peak, the scattering at small angles will necessarily satisfy the conditions for resonant scattering. To the extent that the interfacial roughness is purely conformal, the multilayer film will behave like a single surface having the roughness of the substrate. However, we will see that realistic multilayer films have intrinsic interfacial roughness that modifies the substrate roughness and reduces the conformality of the interface structure. These ''multilayer effects'' tend to enhance the resonant scattering at small angles and suppress the scattering at large angles. To accurately describe the scattering from realistic multilayer film we must account for the variation of the interface structure through the film, and the interaction of the radiation field with each interface.

Once the scattered field is outside of the multilayer coating, its propagation through the distributed optical system to the image plane is described by the transfer function formalism derived in the previous section. In particular, the effect of scattering is to convolve the image with a PSF^{sc} as described by Eq. (28), where now the quantities S_n and $dP_n/d\mathbf{s}_1$ correspond to the contribution from the nth multilayer-coated optical surface. The roughness of the nth coating reduces the specular field by a factor of S_n; the power per unit area scattered by nth coating into the image plane (for a point source) is $dP_n/d\mathbf{s}_1$.

To derive an expression for $dP_n/d\mathbf{s}_1$, we relate this quantity to the angular distribution of power scattered from the nth coating. The power incident at the point $\mathbf{s}_1$ in the image plane corresponds to radiation that is scattered by the nth coating into the direction $\hat{m}$, given by

$$m_X = \frac{\alpha_{nx} s_{1x}}{R}$$

$$m_Y \cong \frac{\alpha_{ny} s_{1y} \cos\ \theta_n}{R} + \sin\ \theta_n \tag{32}$$

$$m_Z \cong -\frac{\alpha_{ny} \sin\ \theta_n}{R}\, s_{1y} + \cos\ \theta_n .$$

Note that $\hat{m}$ is projected onto the (X,Y,Z) coordinate system of the rough coating. The power distribution in the image plane is related to the power scattered per unit solid angle by the nth coating according to

$$\frac{dP_n(\mathbf{s}_1)}{d\mathbf{s}_1} = \frac{\alpha_{nx}\alpha_{ny}}{R_n^{\mathrm{SP}} R^2} \sum_{\hat{a}} \frac{dP_n(\hat{m},\hat{a};\hat{n},\hat{e})}{d\Omega}, \tag{33}$$

where R_n^{SP} is the specular reflectivity of the coating, $\hat{n}$ and $\hat{e}$ are the direction of propagation and the polarization of the incident field, respectively, and we must sum over both polarization states $\hat{a}$ of the scattered field. In essence, Eqs. (32) and (33) represent two mappings: first the nth rough coating is mapped onto the exit pupil and second the angular distribution of scattering from the rough coating is mapped to the spatial coordinates of the image plane. Our task then reduces to finding the angular distribution of scattering, $dP_n/d\Omega$, and the Strehl factor S_n for a multilayer-coated surface given an incident plane wave propagating in direction $\hat{n}$ with polarization $\hat{e}$. This problem is addressed in Sec. VII. However, before we can calculate the scattering from a multilayer coating, we must obtain a realistic description of the roughness of the multilayer interfaces, which is the subject of the next section.

VI. GROWTH MODEL FOR MULTILAYER FILMS

Much attention has been given to the problem of the roughening of the surface of a thin film by growth and erosion. Stochastic theories of the evolution of the surface have been developed using two generally different approaches. In the first approach,[18] the phenomenological observation that randomly rough surfaces are self-affine is used to derive simple scaling laws describing the width of the interface as a function of film thickness and the amount of surface area sampled. One consequence of the scaling theory is that the autocorrelation function for the rough surface is approximately described by:[19]

$$C(r) = \begin{cases} \sigma^2\left[1 - \dfrac{\alpha+1}{2}\left(\dfrac{r}{\xi}\right)^{2\alpha}\right], & \text{for } r \leqslant \xi \\ 0, & \text{for } r > \xi \end{cases} \tag{34}$$

where ξ is the correlation length and α is an independent scaling parameter called the ''roughness exponent.'' The corresponding PSD of the rough surface is given by

$$\mathrm{PSD}(f) = \begin{cases} \dfrac{\alpha}{\pi}\,\sigma^2\xi^2, & \text{for } q < 1/\xi \\ \dfrac{\alpha}{\pi}\,\dfrac{\sigma^2}{\xi^{2\alpha}}\, f^{-2(\alpha+1)}, & \text{for } q \geqslant 1/\xi \end{cases} . \tag{35}$$

The rms roughness σ and correlation length ξ of the surface change with the thickness of the growing (or eroding) film. The scaling laws predict that $\sigma \sim \tau^{\beta}$ and $\xi \sim \tau^{\beta/\alpha}$, where τ is the thickness of the film and β is a second independent scaling parameter.

The second approach[19] of describing the roughening of a surface is through the use of a kinetic continuum equation for the evolution of the surface height $H(\mathbf{r})$. The linear version of such a continuum equation (to lowest order in H) has the form:

$$\frac{\partial H(\mathbf{r})}{\partial \tau} = -\nu\left|\nabla^n H(\mathbf{r})\right| + \frac{\partial \eta}{\partial \tau}. \tag{36}$$

In this approach the evolution of the surface roughness is viewed to be a competition between relaxation of the surface, where ν is an independent growth parameter that characterizes the relaxation process, and the stochastic roughening due to the random shot noise η of the deposition (or removal) process. When ν is positive, the first term in Eq. (36) tends to dampen the surface roughness while the second term increases the roughness with film thickness. The exponent n in the relaxation term varies according to the kinetic mechanism that dominates the smoothing process. Edwards and Wilkinson[20] first applied Eq. (36) with $n=2$ to describe the settling of a granular layer under the influence of the gravitational potential. Herring[21] has identified several relaxation mechanisms relevant to film growth corresponding to viscous flow $(n=1)$, evaporation and condensation $(n=2)$, bulk diffusion $(n=3)$ and surface diffusion $(n=4)$. It has been pointed out by Salditt *et al.*[22] that, for high-energy deposition processes such as sputtering at low pressures, the case of $n=2$ will more likely correspond to the sputter and redeposition of adatoms via atomic bombardment of the surface. Tong and Williams[19] have suggested that by using a negative value of ν the first term in Eq. (36) can also describe roughening of the surface due to three-dimensional island growth. In this case $n=1$ corresponds to island growth via deposition onto the surfaces of existing islands, and $n=3$ represents the growth of islands via the diffusion of atoms on the surface.

Taking the Fourier transform of Eq. (36) readily yields a solution for the PSD of the growing surface:[15]

$$\mathrm{PSD}(\mathbf{f}) = \Omega\,\frac{1-\exp[-2\nu|2\pi f|^n \tau]}{2\nu|2\pi f|^n}. \tag{37}$$

Here Ω is the volume of a constituent element of the film (e.g., atom, molecule, cluster). It is surprising to find that the scaling and kinetic continuum models predict essentially the same form for the PSD of the surface! A comparison of Eqs. (35) and (37) shows that the scaling parameters are related to the exponent n according to

$$\alpha = (n-2)/2, \quad \beta = (n-2)/2n. \tag{38}$$

The kinetic model (36) predicts a characteristic shape for the PSD of a single layer grown on a smooth substrate given by Eq. (37). An example is shown in Fig. 2 for several different film thicknesses and reasonable parameter values of $\Omega = 0.02\ \mathrm{nm}^3$, $\nu = 2.5\ \mathrm{nm}^3$, and $n=4$. The PSD is flat at low frequencies and rolls over to asymptotically approach a

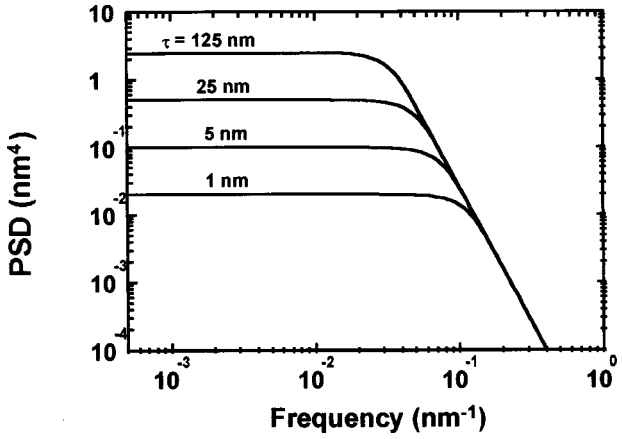


FIG. 2. The theoretical PSD of the top surface of a film grown on a perfectly smooth substrate at different values of the film thickness. The growth parameters are $\Omega = 0.02\ \mathrm{nm}^3$, $\nu = 2.5\ \mathrm{nm}^3$, and $n = 4$.

power law dependence of $\sim f^{-n}$ at high frequencies. The transition point moves to lower frequency as the film thickness increases. In the limit of infinite film thickness the PSD of the surface becomes a pure power law, which is the signature of a self-affine (fractal) surface, and explains in part the similarity between the kinetic continuum and scaling models. The behavior of the PSDs shown in Fig. 2 has a straightforward physical explanation. The low frequency (large wavelength) components of roughness have a flat response characteristic of white noise, which is simply the shot noise of the random deposition process. At high frequency (small wavelength) the PSD rolls off due to local relaxation of the growing surface. In particular, surface features having a size less than $(\nu t)^{1/n}$ are unstable and are damped out.

It should be emphasized that Eq. (36) is the simplest possible kinetic model for roughening. It is a linear and local description of the roughening process, and is expected to be valid only when the surface heights and slopes are small. The first nonlinear correction, corresponding to a term $\sim(\nabla H)^2$, has been considered by Kardar *et al.*[23] Physically, this term represents growth along the local normal to the film surface as might be expected under the conditions of isotropic deposition characteristic of, for example, chemical vapor deposition. The assumption that the roughening is a local process breaks down when the distribution of deposition angles is large and the surface slopes are large. In this case the deposition at a point on the surface depends on the topology of the surrounding surface due to shadowing effects. Karunasiri *et al.*[24] and later Tang *et al.*[25] have proposed growth models that explicitly include a nonlocal growth mechanism (shadowing). It is found that when the nonlinear and nonlocal effects dominate the roughening process, the film surface rapidly develops discontinuities in the form of cusps and columns. These features have often been observed in thin film morphology, particularly for films grown using low-energy deposition processes.[26–28] In contrast, the high-performance multilayer optical coatings which we are considering in this paper have small roughness by design. This is achieved by using a high-energy growth process, such as sputtering at low pressure, that incorporates a significant relaxation mechanism (large ν) to compensate the natural roughening due to the stochastic nature of the deposition. By balancing the roughening and smoothing mechanisms, the roughness is never allowed to become large enough to trigger the nonlinear and nonlocal growth modes. We believe that under these conditions the continuum model of Eq. (36) is an appropriate description of the thin film growth. This view is supported by recent experimental studies[7] of roughness in multilayer films.

Thus far we have considered the growth of a single layer. We next extend the kinetic model to describe the evolution of interfacial roughness in a multilayer film. This is achieved by considering the growth of a multilayer film to be a sequence of single layers, each growing upon a ''substrate'' corresponding to the underlying layer. Then the roughness of an interface naturally separates into two components: (1) the ''intrinsic'' roughness due to the growth of the ith layer, as would occur if the underlying layer was perfectly smooth, and (2) the ''extrinsic'' roughness due to the replication of the roughness of the underlying layer. To represent the growth of such a sequence of layers, we recast Eq. (36) as a finite difference equation[15]

$$h_i(\mathbf{f}) = \gamma_i(\mathbf{f}) + a_i(\mathbf{f}) h_{i-1}(\mathbf{f}), \tag{39}$$

where $h_i(\mathbf{f})$ is the frequency spectrum of the roughness of the ith interface, $H_i(\mathbf{r})$, and

$$a_i(\mathbf{f}) = \exp[-\nu_i |2\pi f|^n \tau_i], \tag{40}$$

is the replication factor that describes the fraction of the frequency component $\mathbf{f}$ in the $(i-1)$th interface that is replicated in the ith interface. The first and second terms on the right-hand side of Eq. (39) correspond to the intrinsic and extrinsic roughness components of the ith interface, respectively.

The growth theory is typically applied to measurements of surface or interfacial roughness using a statistical description of the roughness in terms of the power spectral density

$$\mathrm{PSD}_i(\mathbf{f}) = \frac{1}{A} \langle h_i(\mathbf{f}) h_i^*(\mathbf{f}) \rangle, \tag{41}$$

or the autocorrelation function

$$C_i(\mathbf{r}) = \langle H_i(\mathbf{x}) H_i(\mathbf{x}+\mathbf{r}) \rangle, \tag{42}$$

where the expectation value denotes an average over an ensemble of interface structures having statistically equivalent random roughness. These quantities are related by a simple Fourier transform, and in principal are equivalent descriptions of the structure of the rough surface. In practice, however, all measurements of surface roughness are limited to a finite instrumental bandwidth, and the PSD has the distinct advantage of being accurately measurable within the instrumental bandwidth.[29] We consider the case where the multilayer film is grown by alternately depositing N pairs of high-index (H) and low-index (L) layers onto a substrate (S) having an isotropic surface roughness described by a power spectral density $\mathrm{PSD}_{\mathrm{sub}}$. The PSD of the top surface of the multilayer film is found by successive iteration of Eq. (39) to be

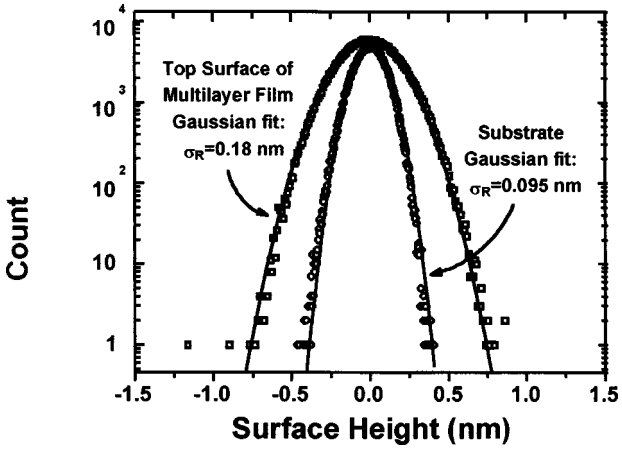


FIG. 3. Surface height distributions of a fused silica substrate and a Mo–Si multilayer film grown on the substrate, measured using atomic force microscopy. The structure of the multilayer film is [Mo(2.1 nm)/Si(4.75 nm)] $\times 40$. The solid lines are Gaussian fits to the data.

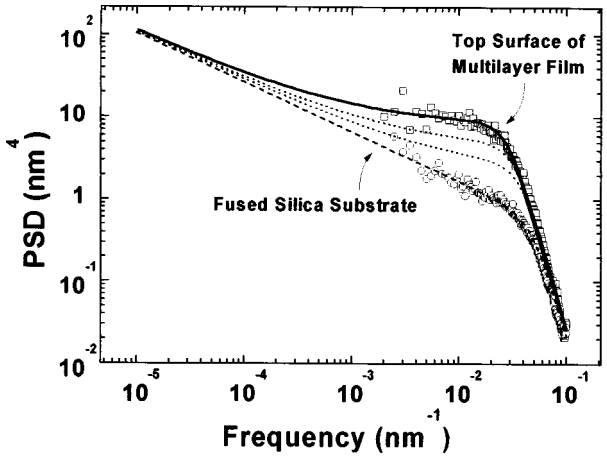


FIG. 4. The PSDs of a fused silica substrate and the top surface of a Mo–Si multilayer film grown on the substrate, measured using atomic force microscopy. The PSD of the substrate is empirically described by $\mathrm{PSD}_{\mathrm{SUB}}(f) = (1.4\times 10^{-6}/f^{4.2})[1-\exp(-7.5\times 10^{4} f^{3.6})]$ in nm^4 (dashed line). The solid line is a best fit to the PSD of the top surface of the multilayer film using the growth model discussed. The dotted lines are the calculated PSDs for the interfaces at the 10th and 20th bilayer period as measured from the substrate.

$$\mathrm{PSD}_N = \frac{1-(a_H^2 a_L^2)^N}{1-a_H^2 a_L^2}\,(\mathrm{PSD}_{\mathrm{int}}^L + a_L^2 \mathrm{PSD}_{\mathrm{int}}^H) + (a_H^2 a_L^2)^N \mathrm{PSD}_{\mathrm{sub}}, \tag{43}$$

where $a_{H,L}$ is the replication factor of the high- or low-index layer and

$$\mathrm{PSD}_{\mathrm{int}}^i(f) = \frac{1}{A}\langle \gamma_i \gamma_i^* \rangle = \Omega_i \frac{1-\exp[-2\nu_i |2\pi f|^n \tau_i]}{2\nu_i |2\pi f|^n}, \tag{44}$$

is the PSD of the intrinsic roughness of the layer. Here we have used the fact that the intrinsic roughness of each interface is statistically independent so that $\langle \gamma_i \gamma_j^* \rangle = 0$ when $i \neq j$.

Our measurements of the roughness of high-performance multilayer optical coatings are generally in good agreement with the predictions of Eqs. (43) and (44). As an example, we show in Figs. 3 and 4 results of surface metrology measurements on a superpolished fused silica substrate and the top surface of a Mo–Si multilayer film grown on the substrate, which we will refer to as our ''canonical'' multilayer sample throughout this paper. The multilayer film was deposited using magnetron sputter deposition in an Ar plasma of 1.75 mTorr pressure as described in detail elsewhere.[30] The film consisted of 40.5 layer pairs with individual layer thicknesses of 2.1 nm for Mo and 4.75 nm for Si. The first and last layers deposited were Si. Images of the surface height were measured using a Digital Instruments Dimension 5000 atomic force microscope operating in the tapping mode. The lateral resolution was ~ 10 nm, due to the width of the tip of the single-crystal Si probe, resulting in a bandwidth limit at high frequency of $\sim 0.1\ \mathrm{nm}^{-1}$. The height resolution of the microscope was ~ 0.01 nm and the surface area sampled was a square region of width 5 μm. To obtain images of these ultrasmooth surfaces it was necessary to operate the microscope inside an environmental chamber that significantly reduced noise from vibrations and air currents. The surface height distributions for the two surfaces are shown in Fig. 3. The data are seen to be well described by Gaussian distributions. The PSDs of the surfaces were obtained directly from the surface images through Fourier transform. The two-dimensional PSDs were found to be isotropic and were averaged over all directions to generate the radial PSDs shown in Fig. 4.

A complete description of the interfacial roughness in the multilayer film can be inferred by fitting the data shown in Fig. 4 to the model of Eqs. (43) and (44). To limit the degrees of freedom of the fitting process, we place the following constraints on the values of the growth parameters:

(1) We set the growth unit volume Ω for Si to the atomic volume of $0.020\ \mathrm{nm}^3$. The Si layers in the multilayer film are deposited in an amorphous phase. The choice of the atomic volume assumes that the final position of the adatoms on the amorphous growth surface are random and uncorrelated.

(2) We set the relaxation parameters for the Mo and Si layers to be equal, $\nu = \nu_{\mathrm{Si}} = \nu_{\mathrm{Mo}}$. This is an arbitrary and unrealistic constraint, which is likely to produce a result that is a weighted average of the true values.

The solid line in Fig. 4 represents the best fit using values of $\nu = 2.5\ \mathrm{nm}^3$, $\Omega_{\mathrm{Mo}} = 0.050\ \mathrm{nm}^3$, and $n = 4$ for the remaining free parameters. We note that the growth unit volume for Mo is approximately three times the atomic volume. This suggests that the final position of the atoms on the growth surface are partially correlated. Indeed, the Mo layers in this film have a polycrystalline bcc phase with a strong $\langle 110 \rangle$ texture,[31] and the ordering due to crystal growth is expected to increase the size of the growth unit. Within the context of the growth theory, the growth parameters Ω, ν, and n, along with the PSD of the substrate, provide a comprehensive description of the roughness of the multilayer film. From these parameters we can determine the PSD of any interface in the multilayer. For example, the dotted lines in Fig. 4 show the PSDs calculated for the interfaces corresponding to the tenth and twentieth bilayer periods as measured from the substrate.

The value of $n=4$ indicates that the dominant relaxation mechanism in these Mo–Si multilayer films grown by sputter deposition is surface diffusion. The importance of surface diffusion in Mo–Si multilayer growth has been previously noted[32] in a study where films were grown using electron beam evaporation onto substrates at elevated temperatures. It was found that the roughness of the interfaces decreased dramatically as the temperature of the growing film was increased, up to a point where interdiffusion caused broadening of the interfaces. The decrease in roughness was accompanied by an increase in the size and texture of the crystallites in the polycrystalline Mo layers. Smooth and abrupt interfaces, comparable to the best sputtered multilayer films, were obtained in a rather narrow temperature range around 525 K. This result was interpreted to indicate that the formation of smooth interfaces required sufficient energy at the growth surface to allow adequate surface mobility without activating bulk diffusion. A similar result of $n=4$ has been found for sputtered $Ni_{0.81}Fe_{0.19}$–Au multilayer films.[9] In contrast, Salditt *et al.*[22] have observed $n=2$ behavior in sputtered W–Si multilayer films, indicating that sputtering and redeposition might be another important smoothing mechanism for specific material systems. It is interesting to note that in these experiments both the W and Si layers were amorphous, whereas in the case of the Mo–Si and $Ni_{0.81}Fe_{0.19}$–Au multilayer films at least one of the layers was polycrystalline. Hence it is possible that surface diffusion is the dominant smoothing mechanism when the major source of interfacial roughness is polycrystalline faceting.

For the case of relaxation via surface diffusion, it is shown in Appendix B that the parameter ν can be related to other standard growth parameters according to

$$\nu=\frac{\xi D V_0^{4/3}}{r_D kT}. \tag{45}$$

Here ξ is the surface energy, D is the surface diffusion coefficient, V_0 is the atomic volume, r_D is the deposition rate, k is Boltzmann's constant, and T is the local temperature of the growth surface. Based on our measurement of the relaxation parameter ν we can infer from Eq. (45) an estimate of the surface diffusion coefficient for the Mo–Si multilayer film growth. Using values[33] for Mo of $\xi=2250\ \mathrm{erg/cm^2}$, $V_0=0.016\ \mathrm{nm^3}$, $r_D=0.2$ nm/s, and $T=525$ K we obtain a surface diffusion coefficient of $D=4\times10^{-15}\ \mathrm{cm^2/s}$. Then the range of an adatom on the surface before it is ''frozen'' by the deposition of the next monolayer would be approximately given by $\sqrt{Dt}=0.9$ nm. This suggests that the adatom has the opportunity to relax to a energetically favorable position within a radius of several atomic sites. Monte Carlo simulations[34] of thin film growth have indicated that such a relaxation mechanism is sufficient to produce low defect films of near bulk density.

It is evident in Figs. 3 and 4 that the roughness of the top surface of the multilayer film is significantly greater than the substrate. The kinetic growth theory predicts three distinct regimes:

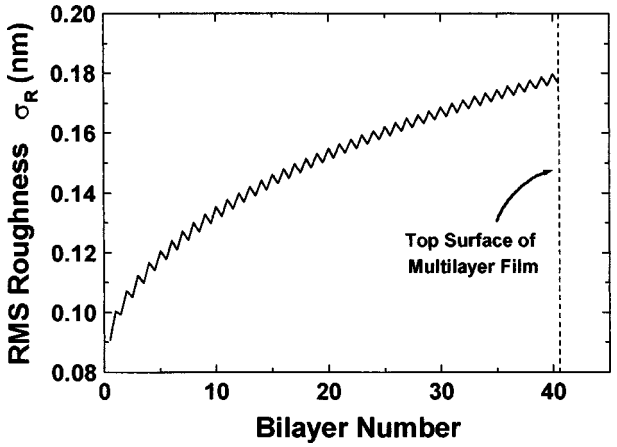


FIG. 5. The variation of the rms roughness of the interfaces within the Mo–Si multilayer film, obtained by integrating the PSDs over the frequency range of 10^{-4}–$10^{-1}\ \mathrm{nm^{-1}}$.

(1) At low spatial frequencies (less than $\sim10^{-4}\ \mathrm{nm^{-1}}$ in Fig. 4) the interfacial roughness replicates the substrate roughness and is purely conformal.
(2) There is an intermediate frequency range (10^{-4}–$10^{-1}\ \mathrm{nm^{-1}}$ in Fig. 4) in which the top surface of the multilayer film is significantly rougher than the substrate due to the additional intrinsic roughness associated with the growth of the film.
(3) At high frequencies (greater than $\sim10^{-1}\ \mathrm{nm^{-1}}$ in Fig. 4) the multilayer film growth process has a smoothing effect and can actually result in a damping of the roughness of the substrate.

In general, the transition points between these frequency regimes are not unique; they can vary with both the multilayer growth parameters and the PSD of the substrate.

The increase in the multilayer film roughness in the frequency range of 10^{-4}–$10^{-1}\ \mathrm{nm^{-1}}$ is a particular concern for soft x-ray imaging, as it represents in some sense the limit of smoothness that can be obtained for a multilayer-coated surface. For 13 nm radiation near normal incidence these frequency components will produce scattering at angles ranging from ~0.1 to 90 degrees. The integrated power scattered over this angular range is proportional to the variance of the surface height, σ^2, which is the second moment of the surface height distribution shown in Fig. 3, and can also be obtained from the PSD according to

$$\sigma^2=2\pi\int_{f\ \min}^{f\ \max}\mathrm{PSD}(f)f\,df. \tag{46}$$

The rms roughness, σ, determined by integrating the PSD over the frequency range of 10^{-4}–$10^{-1}\ \mathrm{nm^{-1}}$, is plotted in Fig. 5 for all of the interfaces in the multilayer film. The rms roughness is observed to double throughout the thickness of the film, increasing from a value of $\sigma=0.09$ nm at the substrate to $\sigma=0.18$ nm at the top surface. From this example it is evident that the intrinsic roughness associated with the growth of the multilayer film can be a significant part of the

total roughness, and should be included to obtain an accurate description of nonspecular scattering from multilayer-coated optical surfaces.

VII. THEORY OF NONSPECULAR SCATTERING FROM A MULTILAYER FILM

The last decade has seen considerable progress towards developing a rigorous theory of the scattering of x rays from multilayer films. The existing theoretical framework exploits the fact that the interaction between x rays and matter is typically weak, and treats the nonspecular scattering from interfacial roughness using first- or at most second-order perturbation theory. Consequently the theory is limited to the case where the scattered power is small compared to the incident power. Within the context of perturbation theory there are basically two different formulations, each valid in a different regime. Stearns[3] has presented a theory of scattering from multilayer structures based upon previous work[16] describing the scattering from a single interface within the Born approximation. In this treatment, the incident field at each interface consists of plane waves incoming from both sides, corresponding to the exact eigenstate of the ideal multilayer structure (no roughness). The rough interface is considered to be the perturbation, and a solution of Maxwell's equations is found for the scattered vector field, which includes the polarization dependence. The total scattered field, consisting of outgoing plane waves from each of the interfaces, is treated kinematically. The Born approximation neglects the refraction of the incident field, which becomes important near the critical angle for total internal reflection, and hence this scattering theory is only valid for angles greater than the critical angle (as measured from the surface).

In an alternate approach, Holy *et al.*[5] have developed a scattering theory using the distorted-wave Born approximation (DWBA), based upon the description of scattering from a single rough surface presented by Sinha *et al.*[35] The theory has been extended to second order by de Boer.[36] In this approach the incident field is an eigenstate of the ideal multilayer structure (in the absence of roughness, although the interfaces need not be abrupt[36]) having a wave vector $\mathbf{k}_0$, including both the incoming and outgoing waves, and the scattered field is a time-reversed version of an incident field having a different wave vector $\mathbf{k}$. The perturbation is just the change of the multilayer structure upon introducing the roughness. At first glance the DWBA formulation appears to be unphysical since the scattered field includes incoming plane waves. This paradox is resolved by realizing that the final state only needs to be a reasonable approximation of the scattered field within the interaction region, that is, the regions where the roughness exists. When the reflectivity is large, such as near the critical angle, the ''outgoing'' scattered plane wave will experience multiple reflections within the interaction region, thereby creating a strong ''incoming'' plane wave component, and this state is well-represented by the DBWA final state. Hence this theory is valid at angles near the critical angle, but is generally not applicable at larger angles. Another limitation of the DBWA approach is that it is based on the Helmholtz equation for the scalar field and neglects polarization effects, which is strictly only valid at grazing incidence.

An important practical issue in the formulation of the scattering theory is the way in which the roughness of the interfaces is incorporated. It is easily shown[35] that the scattering from a single surface is proportional to the Fourier transform of the quantity $\exp[q_z^2 C(\mathbf{r})/2]$. Consequently, there has been a tendency to describe and model surface roughness in terms of the autocorrelation function. However, in the limit where $q_z^2\sigma^2 \ll 1$, which we call the ''small roughness approximation,'' the exponential can be expanded to obtain the well-known result that the scattering cross section is proportional to the PSD of the surface roughness. The formulation of the scattering theory in terms of the PSD has several distinct advantages:

(1) The PSD is directly measurable by instruments having finite bandwidths, as mentioned previously.
(2) Knowledge of the PSD within a limited bandwidth is sufficient to model scattering for a given angular range. In contrast, the complete autocorrelation function is required to describe the scattering within any angular range.
(3) The scattering problem can be inverted to determine the PSD of a surface from the angular distribution of the scattering.

In the case of scattering from multilayer films the situation gets more complicated; the scattering is proportional to a sum of Fourier transforms of exponential terms containing cross-correlation functions between every pair of interfaces. Specific models of the correlation functions are typically introduced in an *ad hoc* fashion. In contrast, applying the small roughness approximation (when valid) provides the important simplification of linearizing the dependence of the scattering amplitude on the interface roughness $h(\mathbf{f})$. It then becomes possible to directly integrate into the scattering theory the linear growth model described in the previous section.

In this paper we are interested in modeling the scattering from high-performance multilayer optical coatings in configurations near normal incidence. Under these conditions it is appropriate to apply the scattering theory of Stearns based on the Born approximation and linearized using the small roughness approximation. It is necessary, however, to extend the previous theory to account for two effects which are important in modeling the performance of realistic imaging systems. These new developments are:

(1) The scattered field is treated dynamically, that is, we take into account the multiple specular reflection of the scattered field within the multilayer structure. The dynamical treatment of the scattered field is important when the scattering angle is within a Bragg resonance condition, as will generally be the case for small angle scattering in an imaging system.

(2) The description of the unperturbed multilayer structure (without roughness) is allowed to included interfacial diffuseness. The diffuseness, corresponding to a broadened composition profile across the interface, arises from interdiffusion and reaction at the layer boundaries. For example, high performance Mo–Si multilayer coatings exhibit inter-

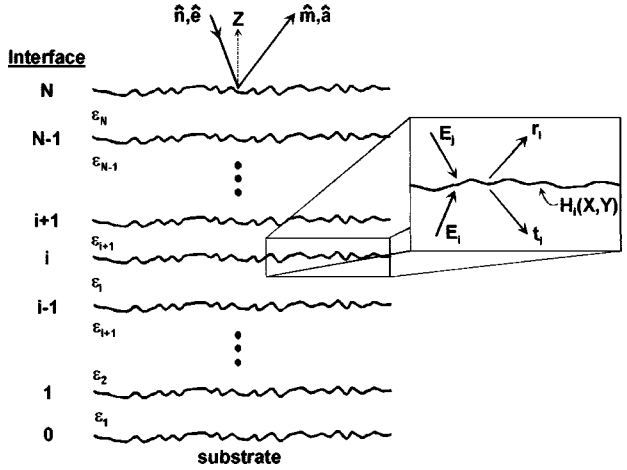


FIG. 6. Schematic diagram of a multilayer film having rough and diffuse interfaces. The inset shows the scattering process at the ith interface. The specular fields E_i and E_j are incident on either side of the interface. The nonspecular scattering into mode $(\hat{m},\hat{a})$ consists of two parts, the field r_i that is scattered towards the top of the film and the field t_i that is scattered towards the substrate.

diffusion zones of ~ 1 nm at the Mo-on-Si interface and ~ 0.5 nm at the Si-on-Mo interface.[37] The relative effects of roughness and diffuseness on x-ray scattering are discussed in Ref. 16. In general, diffuseness of the interfaces reduces the specular reflectivity of each interface, and thereby redistributes the specular field within the multilayer structure. In this way diffuseness can have an important effect on the nonspecular scattering without producing any scattering *per se*.

We begin by considering a multilayer film having a series of rough and diffuse interfaces as shown schematically in Fig. 6. Let the incident field be a plane wave with wavevector $k\hat{n}$ and polarization $\hat{e}$, corresponding to either S or P type. Multiple specular reflection within the multilayer film produces counterpropagating plane waves in each layer and refraction modifies the wave vectors and polarization vectors. The wave vectors in the ith layer, $k_i\hat{n}_i^{\pm}$, for the fields propagating towards the top of the film $(+)$ and towards the substrate $(-)$ are related to the incident field according to

$$\begin{aligned} k_i n_{iX}^{\pm} &= k n_X \\ k_i n_{iY}^{\pm} &= k n_Y \\ k_i n_{iZ}^{\pm} &= \pm k\sqrt{\epsilon_i - n_X^2 - n_Y^2}. \end{aligned} \tag{47}$$

Here the wavenumber $k_i = \epsilon_i^{1/2} k$ becomes a complex quantity for the case of an absorbing medium. The polarization vectors within the ith layer are given by

$$S \text{ type:} \quad e_{iX}^{\pm} = \frac{n_{iY}}{\sqrt{(n_{iX})^2 + (n_{iY})^2}},$$

$$e_{iY}^{\pm} = -\frac{n_{iX}}{\sqrt{(n_{iX})^2 + (n_{iY})^2}}, \quad e_{iZ}^{\pm} = 0$$

$$P \text{ type:} \quad e_{iX}^{\pm} = \frac{n_{iX} n_{iZ}^{+}}{\sqrt{(n_{iX})^2 + (n_{iY})^2}},$$

$$e_{iY}^{\pm} = \frac{n_{iY} n_{iZ}^{+}}{\sqrt{(n_{iX})^2 + (n_{iY})^2}}, \quad e_{iZ}^{\pm} = \mp\sqrt{(n_{iX})^2 + (n_{iY})^2}\,. \tag{48}$$

We define interface i to be located between layers i and $j = i+1$. Then there are two plane waves, $E_j^- \hat{e}_j^-\exp(ik_j\hat{n}_j^- \cdot \mathbf{x})$ and $E_i^+ \hat{e}_i^+ \exp(ik_i\hat{n}_i^+ \cdot \mathbf{x})$, incident on the ith interface from above and below, respectively. Within the spirit of perturbation theory, the incident plane waves correspond to the specular field in the multilayer film in the absence of roughness.

Let us consider the field scattered by the multilayer film into direction $\hat{m}$ with polarization $\hat{a}$. As before, the scattered field undergoes multiple specular reflections within the film which produces counterpropagating waves in each layer having wave vectors $k_i\hat{m}_i^{\pm}$ and polarizations $\hat{a}_i^{\pm}$. Following the formalism of Ref. 3, the amplitude of the field scattered by the ith interface towards the top of the multilayer film is given by

$$r_i(\hat{m}_j^+, \hat{a}_j^+) = \frac{\Delta_{ji} k_j^3}{8\pi^2 \epsilon_j m_{jZ}^+}\left(E_j^-(\hat{a}_j^+ \cdot \hat{e}_j^-)\,\frac{g_i(\mathbf{q}_r)}{q_{rZ}} + E_i^+(\hat{a}_j^+ \cdot \hat{e}_i^+)\,\frac{\tilde{g}_i(\mathbf{q}_r')}{q_{rZ}'}\right). \tag{49}$$

Here $\mathbf{q}_r = k_j\hat{m}_j^+ - k_j\hat{n}_j^-$ and $\mathbf{q}_r' = k_j\hat{m}_j^+ - k_i\hat{n}_i^+$ are the momentum transfer vectors, $\Delta_{ji} = \epsilon_j - \epsilon_i$ is the change in the dielectric function across the interface, $g_i(\mathbf{q})$ is the Fourier transform of the normalized gradient of the dielectric function given by

$$g_i(\mathbf{X}) \equiv \frac{1}{\Delta}\frac{\partial \epsilon(\mathbf{X})}{\partial Z}, \tag{50}$$

and $\tilde{g}_i(q_X, q_Y, q_Z) = g_i(q_X, q_Y, -q_Z)$. Since we are treating the scattered field dynamically, we must also include the scattering that is initially directed towards the substrate, as this radiation can be reflected back out of the film by underlying interfaces. The field scattered by the ith interface towards the substrate is given by

$$t_i(\hat{m}_i^-, \hat{a}_i^-) = \frac{\Delta_{ij} k_i^3}{8\pi^2 \epsilon_i m_{iZ}^-}\left(E_i^+(\hat{a}_i^- \cdot \hat{e}_i^+)\,\frac{\tilde{g}_i(\mathbf{q}_t)}{q_{tZ}} - E_j^-(\hat{a}_i^- \cdot \hat{e}_j^-)\,\frac{g_i(\mathbf{q}_t')}{q_{tZ}'}\right), \tag{51}$$

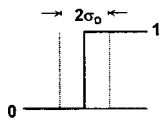
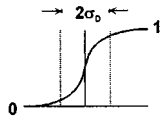
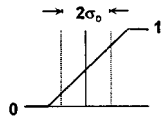
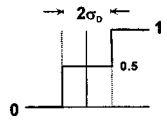
where $\mathbf{q}_t = k_i\hat{m}_i^- - k_i\hat{n}_i^+$ and $\mathbf{q}_t' = k_i\hat{m}_i^- - k_j\hat{n}_j^-$.

All of the information about the structure of the interface is contained in the quantity g, which can be thought of as a structure factor for the interface in the terms of x-ray diffraction theory. We choose to describe the interface using a model

$$g(\mathbf{X}) = W^D[Z - H(X,Y)]. \tag{52}$$

Here the function $W^D(Z)$ represents the gradient of the dielectric function across the interface due to the diffuseness; the position of this diffuse interface is modulated by the

TABLE I. Several examples of the gradient function $W^D(Z)$ for a diffuse interface and its Fourier transform $w^D(q_Z)$.

Description of interface	$W^D(Z)$	$w^D(q_Z)$
(a) Ideal 	$\delta(Z)$	1
(b) Error Function 	$\frac{1}{\sqrt{2\pi\sigma_D^2}}\exp(-Z^2/2\sigma_D^2)$	$\exp(-\sigma_D^2 q_Z^2/2)$
(c) Linear 	$\begin{cases} 0, & \lvert Z\rvert > \sqrt{3}\sigma_D \\ \frac{1}{2\sqrt{3}\sigma_D}, & \lvert Z\rvert < 3\sigma_D \end{cases}$	$\frac{\sin(\sqrt{3}\sigma_D q_Z)}{\sqrt{3}\sigma_D q_Z}$
(d) Step 	$\frac{1}{2}[\delta(Z+\sigma_D)+\delta(Z-\sigma_D)]$	$\cos(\sigma_D q_Z)$

roughness function $H(X,Y)$. The key underlying assumption of this model is that the diffuseness is constant over the interface. We would expect this model to be valid when the physical mechanisms causing the roughness and diffuseness are essentially independent. This would be the case, for example, when the interface roughness is predominantly due to the replication of roughness of the underlying layer and the diffuseness is due to the local interdiffusion or reaction of the layers. One example where our model would be inappropriate is the case where the interface roughness is produced by nonuniform interdiffusion or reaction at the layer boundary.

Several important examples of a diffuse interface are listed in Table I, along with the gradient function $W^D(Z)$ and its Fourier transform $w^D(q_Z)$. In case (a) we show an ideal interface, where the dielectric function changes abruptly between layers. In this case the gradient is a delta function with a Fourier transform of unity. Classical interdiffusion is represented by case (b), where the dielectric function across the interface is described by an error function, and both the gradient and its Fourier transform are Gaussian. When a compound is formed at the interface, and the growth of the interlayer is rate limited by diffusion through the interlayer, then the dielectric function should have a linear profile as shown in case (c). However, if the growth of the interlayer is limited by the reaction rate at the interlayer boundary, then the dielectric function has a step profile as shown in case (d). In Table I we normalize the width of the different interface models to the second moment of the gradient function, σ_D, defined as

$$\sigma_D^2 \equiv \int Z^2 W^D(Z)\,dZ. \tag{53}$$

Choosing the appropriate model requires a detailed knowledge of the microstructure of the interface. This information is accessible from high-resolution imaging techniques such as transmission electron microscopy of cross-section specimens.[31] Advances in new scanning techniques such as high-angle annular dark-field microscopy[38] make it possible to map the composition gradient at interfaces with nanometer resolution.

Taking the Fourier transform of Eq. (52) we obtain for the nonspecular case $(q_X, q_Y \neq 0)$:

$$\begin{aligned} g(\mathbf{q}) &= w^D(q_Z)\int \exp(-iq_X X)\exp(-iq_Y Y) \\ &\quad \times \exp[-iq_Z H(X,Y)]\,dX\,dY \\ &\cong -iq_Z h(q_X,q_Y) w^D(q_Z) \end{aligned} \tag{54}$$

where we have used the small roughness approximation (e) to expand the exponential. Substituting into Eqs. (49) and (51) yields

$$\begin{aligned} r_i(\hat{m}_j^+, \hat{a}_j^+) &= -\frac{i\Delta_{ji} k_j^3 h_i(q_X,q_Y)}{8\pi^2 \epsilon_j m_{jZ}^+}\,[E_j^-(\hat{a}_j^+\cdot\hat{e}_j^-) w^D(q_{rZ}) \\ &\quad + E_i^+(\hat{a}_j^+\cdot\hat{e}_i^+) w^D(-q'_{rZ})] \\ t_i(\hat{m}_i^-, \hat{a}_i^-) &= \frac{i\Delta_{ij} k_i^3 h_i(q_X,q_Y)}{8\pi^2 \epsilon_i m_{iZ}^-}\,[E_i^+(\hat{a}_i^-\cdot\hat{e}_i^+) w^D(-q_{tZ}) \\ &\quad + E_j^-(\hat{a}_i^-\cdot\hat{e}_j^-) w^D(q'_{tZ})]. \end{aligned} \tag{55}$$

The power per unit solid angle scattered by all of the interfaces in the multilayer film is

$$\frac{dP(\hat{m},\hat{a})}{d\Omega} = \frac{4\pi^2 m_Z^2}{k^2 A |n_Z|}\left|\sum_i (\phi_i^r r_i + \phi_i^t t_i)\right|^2, \tag{56}$$

where A is the area of the film illuminated by the incident field. The propagation factors ϕ_i^r and ϕ_i^t account for the phase shift and attenuation of the field scattered from the ith interface as it propagates to the top surface of the multilayer film. We rearrange Eq. (56) to show explicitly the dependence on the interface structure:

$$\frac{dP(\hat{m},\hat{a})}{d\Omega} = \frac{m_Z^2}{16\pi^2 k^2 A |n_Z|}\sum_{ik}(\Gamma_i \Gamma_k^* h_i h_k^*), \tag{57}$$

where

$$\begin{aligned} \Gamma_i &= \frac{\Delta_{ji} k_j^3 \phi_i^r}{\epsilon_j m_{jZ}^+}\,[E_j^-(\hat{a}_j^+\cdot\hat{e}_j^-) w^D(q_{rZ}) \\ &\quad + E_i^+(\hat{a}_j^+\cdot\hat{e}_i^+) w^D(-q'_{rZ})] \\ &\quad - \frac{\Delta_{ji} k_i^3 \phi_i^t}{\epsilon_i m_{iZ}^+}\,[E_i^+(\hat{a}_i^-\cdot\hat{e}_i^+) w^D(-q_{tZ}) \\ &\quad + E_j^-(\hat{a}_i^-\cdot\hat{e}_j^-) w^D(q'_{tZ})]. \end{aligned} \tag{58}$$

The next step is to incorporate the multilayer growth model to describe the roughness of the interfaces, including the correlation of roughness between interfaces. In particular, iteration of Eq. (39) shows that the roughness h_i of the ith interface is a superposition of the intrinsic roughness γ_i of each of the underlying layers (and the substrate). The roughness h_i can be written as

$$h_i = \sum_{n=0}^{i} c_{\mathrm{in}} \gamma_n . \tag{59}$$

The factor c_{in} represents the amount of intrinsic roughness of the nth layer that propagates to the ith interface. It is explicitly related to the replication factors, a_m, of the intervening interfaces according to

$$c_{\mathrm{in}} = \frac{\Pi_{m=0}^{i} a_m}{\Pi_{m=0}^{n} a_m}. \tag{60}$$

Assuming that the intrinsic roughness γ_i of each interface is statistically independent, we have

$$h_i h_k^* \xrightarrow{\text{for } k<i} \sum_{n=0}^{k} c_{\mathrm{in}} c_{kn} \gamma_n \gamma_n^* = A \sum_{n=0}^{k} c_{\mathrm{in}} c_{kn} \mathrm{PSD}_{\mathrm{int}}^{n}. \tag{61}$$

Substituting into Eq. (57) yields

$$\frac{dP(\hat{m},\hat{a})}{d\Omega} = \frac{m_Z^2}{16\pi^2 k^2 |n_Z|} \times \sum_{i=0}^{N} \left[\left(\sum_{n=0}^{i} c_{\mathrm{in}}^2 \mathrm{PSD}_{\mathrm{int}}^{n} \right) \Gamma_i \Gamma_i^* + \sum_{k=0}^{i-1} \left(\sum_{n=0}^{k} c_{\mathrm{in}} c_{kn} \mathrm{PSD}_{\mathrm{int}}^{n} \right) (\Gamma_i \Gamma_k^* + \Gamma_k \Gamma_i^*) \right]. \tag{62}$$

This expression is the central result of the multilayer scattering theory. The angular distribution of the scattering power is directly related to the detailed structure of the interfaces through the $\mathrm{PSD}_{\mathrm{int}}$ of the intrinsic roughness (including the substrate) and the factors c_{in} which describe the replication of roughness between interfaces. These quantities are in turn defined, within the context of our growth model, by the growth parameters Ω, ν, and n, which are characteristic of the film media and deposition conditions. It is evident in Eq. (62) that the scattering separates naturally into two terms. The first term corresponds to the uncorrelated scattering, and is simply the sum of the intensities scattered by each interface independently. The second term corresponds to the correlated scattering. This contribution represents the interference of the radiation fields scattered by interfaces that are correlated due to the replication of roughness from layer to layer.

We note that Eq. (62) is only applicable under the conditions for which the multilayer growth model is expected to be valid, that is, when the roughness is sufficiently small so that the growth kinetics are local and linear. This limit is consistent with the ''small roughness approximation'' (e) that was invoked previously, and is expected to be satisfied by high performance multilayer optical coatings, where the roughness is minimized by design. Equation (62) should not be valid when the roughness is large and the film growth is dominated by nonlinear and nonlocal effects such as shadowing and columnar growth.

Although the formulation of the scattering theory is complete, its implementation requires a method of calculating the incident field amplitudes, E_i^+ and E_j^-, and the propagation factors ϕ_i^r and ϕ_i^t. These quantities are to be determined in the absence of roughness, since we are considering the scattering process as a first-order perturbation (the Born approximation). This is accomplished using a well-known matrix approach to analyze the propagation of the specular fields within the multilayer film. The matrix method is described in detail in Appendix C.

The description of the image formation in the presence of scattering (Eq. [29]) also requires the calculation of the Strehl factor, S. This factor accounts for the total decrease in the specular intensity due to scattering and, for the case of a multilayer film, includes losses due to the absorption of the scattered field within the film. In practice, the accurate calculation of the Strehl factor is problematic when there is a strong reflected field. This is because the scattering from the roughness significantly alters the configuration of the incident specular field. In particular, the roughness redistributes the power between the reflected and transmitted specular fields, which generates loss through increased absorption. This is a purely dynamical effect, requiring second-order perturbation theory to correctly describe the lowest-order change in the specular field due to the interfacial roughness. Unfortunately, a general second-order theory of scattering from rough multilayer structures has not yet been developed. An alternative approach for estimating the Strehl factor, based on an *ad hoc* treatment of these interference effects, is presented at the end of this section.

As an example of the application of the multilayer scattering theory, we have modeled the specular reflectivity and nonspecular scattering from the canonical Mo–Si multilayer film described in Sec. VI, and we compare these results to experimental measurements. The x-ray scattering measurements were performed using synchrotron radiation (S polarization) provided by Beamline 6.3.2 of the Advanced Light Source at Lawrence Berkeley Laboratory. A detailed description of the beamline and reflectometer is presented elsewhere.[39] A unique feature of this experimental facility, which is particularly important for performing scattering measurements on high-quality multilayer films, is the combination of high photon flux ($\sim 10^{12}$ photons/s in 0.1% bandwidth) and excellent collimation of the incident beam. An example of the profile of the incident beam is shown in Fig. 7. At angles greater than four degrees the wings of the incident beam are suppressed by over nine orders of magnitude, making it possible to measure very low levels of scattering.

The measured specular reflectivity of the Mo–Si multilayer as a function of x-ray wavelength is shown in Fig. 8 for angles of incidence ranging from 5 to 20 degrees as measured from the film normal. The reflectivity was modeled using the matrix method described in Appendix C, with optical constants provided by the CXRO World Wide Web

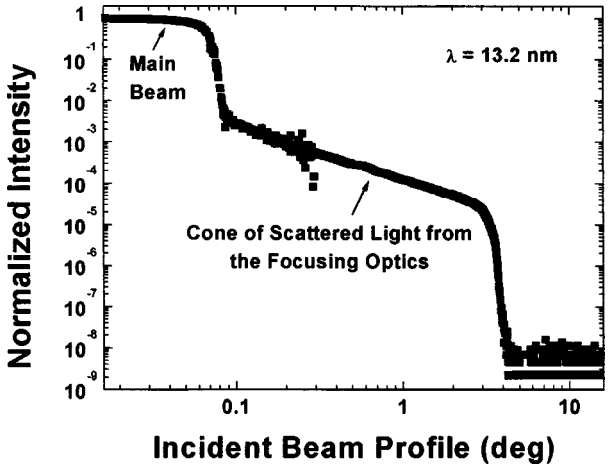


FIG. 7. The measured angular profile of the incident beam for the experimental configuration used to obtain the scattering data. The intensity of the incident beam is reduced by nine orders of magnitude within four degrees of the center of the beam. The high spatial purity is achieved using a combination of focusing optics and apertures.

site.[40] Modeling the position and width of the Bragg peaks provides an accurate and unique determination of the individual layer thicknesses. The interfaces have been modeled as asymmetric zones of intermixing having linear composition profiles, based upon previous detailed studies of the microstructure of similar Mo–Si multilayer films.[41] The width, σ_D, of these zones is 0.3 nm for the Mo-on-Si interfaces and 0.15 nm for the Si-on-Mo interfaces. The best fits are shown as the solid lines in Fig. 8, and correspond to a layer structure of $[\mathrm{Mo}(2.1\ \mathrm{nm})/\mathrm{Si}(4.75\ \mathrm{nm})]\times 40$. The amplitude of the reflectivity (i.e., the Strehl factor) is not accurately modeled by the matrix method, as discussed above, and has been treated as a free fitting parameter.

The nonspecular scattering measured at normal incidence ($\theta = 0$ deg) and at wavelengths of 12.8 and 13.2 nm is shown in Fig. 9 as a function of the scattering angle. We also show the calculated *specular* reflectivity for comparison

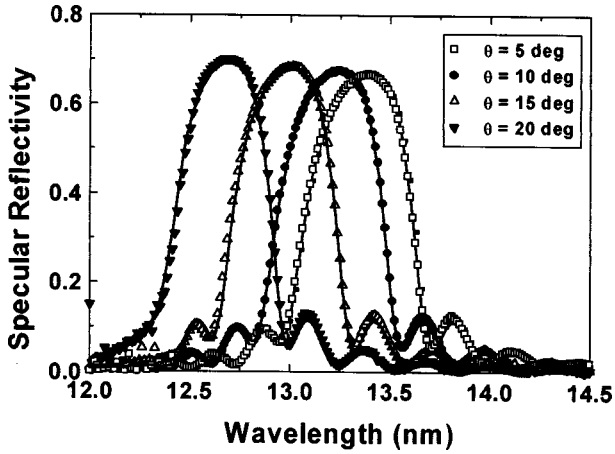


FIG. 8. Measured specular reflectivity of the canonical Mo–Si multilayer film as a function of soft x-ray wavelength for several different angles of incidence. The peak reflectivity varies from 66% at five degrees and 13.4 nm to 70% at 20 degrees and 12.6 nm (just above the Si L edge). The solid lines are best fits using a multilayer structure of $[\mathrm{Mo}(2.1\ \mathrm{nm})/\mathrm{Si}(4.75\ \mathrm{nm})]\times 40$.

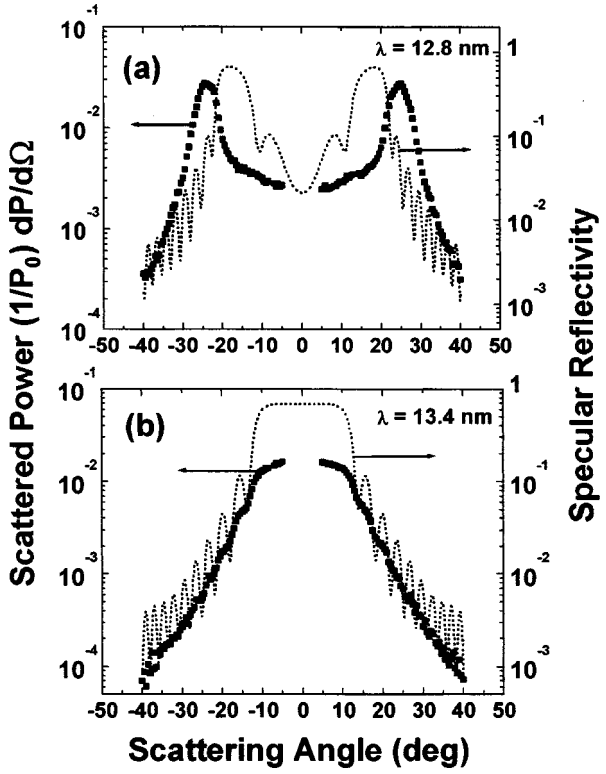


FIG. 9. Nonspecular scattering measured from the canonical Mo–Si multilayer film for normal incidence radiation at a wavelength of (a) 12.8 nm and (b) 13.4 nm. The dotted lines are the calculated specular reflectivity of the film as a function of incident angle.

(dotted lines). The measurements correspond to the scattered power per unit solid angle, normalized to the incident power. These data were obtained using a channeltron detector operating in pulse counting mode. The solid angle subtended by the detector was defined by a 2.0-mm-diameter pinhole positioned 225 mm from the multilayer sample. The scattering was measured by scanning the detector in the plane of incidence while keeping the incident beam fixed. The scattering could be measured to within approximately 4 degrees of the angle of specular reflection, at which point the background level overwhelmed the scattering signal.

The angular dependence of the nonspecular scattering exhibits several characteristic features. In both cases there is a broad peak that mimics the specular Bragg peak. At 12.8 nm [Fig. 9(a)] the scattering peak is shifted to a larger angle than the specular Bragg peak. At 13.4 nm [Fig. 9(b)] the scattering and specular peaks appear to coincide. Furthermore, the nonspecular scattering exhibits a small oscillation, more easily observed in the 13.4 nm data [Fig. 9(b)], which dies out at larger scattering angles.

The origin of the broad peak in the nonspecular scattering is the same as the Bragg peak in specular reflection: the interference of the fields scattered from the different interfaces. This phenomenon has been called ''quasi-Bragg scattering'' [3] or ''resonant diffuse scattering'' [5] in the literature. We will refer to it as resonant nonspecular scattering (RNS) in this paper. Two conditions are required to observe a peak in RNS. First there must be correlation between the roughness of the interfaces in the multilayer. Second, the

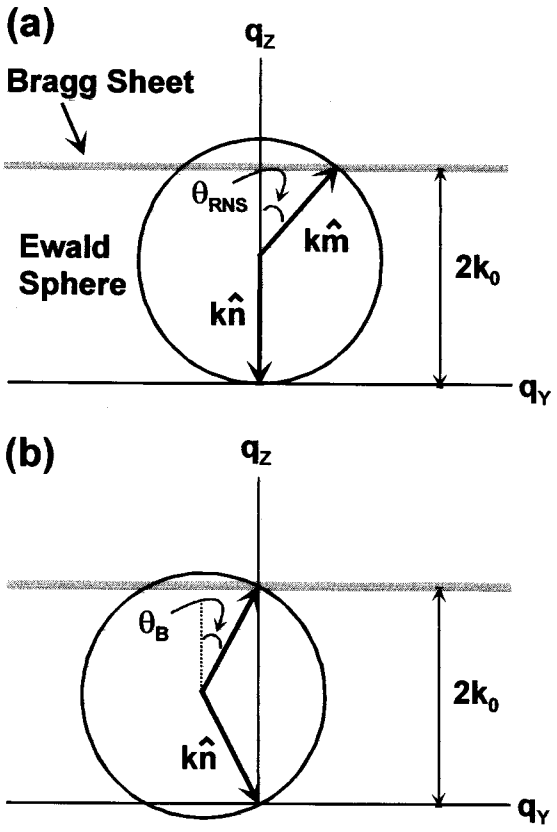


FIG. 10. (a) An Ewald construction showing the nonspecular scattering process in reciprocal space. The Bragg sheet is located at $q_z=2k_0$ and the incident radiation, of momentum $k\hat{n}$, is normal to the film. The nonspecular scattered field, of momentum $k\hat{m}$, is constrained to the surface of the Ewald sphere. A peak in the resonant nonspecular scattering (RNS) occurs when the Ewald sphere intersects the Bragg sheet. (b) An Ewald construction illustrating the conditions for specular reflectivity.

fields scattered from the different interfaces must add in phase constructively. These requirements are best illustrated using an Ewald construction in reciprocal space, as shown in Fig. 10. Here the Bragg reflection at $q_Z=2k_0$ is spread out into a sheet parallel to the q_X-q_Y plane due to the correlated roughness of the multilayer interfaces, which produces coherent scattering with finite momentum transfer in the X-Y plane. Figure 10(a) shows the configuration for scattering at normal incidence with an x-ray momentum $k>k_0$ [as is the case in Fig. 9(a)]. The allowed (elastic) values of momentum transfer are constrained to be on the surface of the Ewald sphere. The peak in the RNS, θ_{RNS}, is given by the angle at which the scattering vector $k\hat{m}$ intersects the Bragg sheet. The RNS peak generally occurs at a larger angle than the specular Bragg peak. This can be seen in Fig. 10(b), where we show an Ewald construction for specular reflection from the multilayer at the same value of x-ray momentum k. Here the angle of incidence and reflection is θ_B. Inspection of the two diagrams in Fig. 10 shows that the relationship between the angular positions of the RNS and Bragg peaks is

$$2k\cos\theta_B=k+k\cos\theta_{\mathrm{RNS}}, \tag{63}$$

which reduces to $\theta_{\mathrm{RNS}}\cong\sqrt{2}\,\theta_B$ in the limit of small angles. This predicted behavior is consistent with the data of Fig. 9(a), where the RNS and Bragg peaks are at $\sim$25 and 18 degrees, respectively. The two peaks only coincide when the wavelength and angle of the incident field satisfy the condition for Bragg reflection ($k=k_0$), as is the case in Fig. 9(b).

An important consequence of the relationship between the position of the RNS and Bragg peaks is that the total integrated scattering increases significantly as k exceeds k_0. The total integrated scattering (TIS) corresponds to the total power nonspecularly scattered into all angles for a given incident angle and wavelength. The TIS obtained by integrating the scattering shown in Fig. 9 is found to be 0.98% and 0.30% at the wavelengths of 12.8 and 13.4 nm, respectively. The increase in TIS at the shorter wavelength (larger k) is due to a significant increase in the transmission of the multilayer film as the RNS peak shifts away from the specular Bragg peak. Within the Bragg peak, scattered radiation is trapped inside the film in a standing wave similar to the specular field, which reflects much of the radiation back into the film. However, when the RNS is outside of the Bragg peak, as occurs at shorter wavelengths [Fig. 9(a)]. The radiation scattered at interfaces within the film propagates to the top surface with little loss. This purely dynamical effect causes the extinction of the RNS to vary dramatically with scattering angle.

In Fig. 9 the RNS exhibits a small oscillation having the same period as the high frequency oscillation of the specular reflectivity (sometimes called ''Kiessig fringes''). The oscillation in the RNS has the same origin as that in the specular reflectivity, namely the interference of the radiation scattered from the front and back surfaces of the multilayer film. This can occur in scattering whenever there is a correlation between the roughness of the substrate and the roughness of the top surface. Indeed, the existence of a finite thickness oscillation in the scattering intensity is an unequivocal indicator of conformality in the roughness of the multilayer film. The amplitude of the oscillation indicates the degree to which the roughness of the substrate is replicated at the top surface. The growth theory of Sec. VI asserts that the degree of replication is a strong function of frequency, decreasing at higher frequency. Hence we expect that the RNS at small scattering angles, corresponding to lower frequency roughness, should exhibit larger finite thickness oscillations than the RNS at large scattering angles. This is consistent with the data shown in Fig. 9, where the amplitude of the oscillation is observed to dampen with increasing scattering angle, and is essentially absent at angles greater than $\sim$25 degrees.

In Fig. 11 we present additional measurements of nonspecular scattering from the canonical Mo–Si multilayer film for normal incidence and wavelengths of 12.8, 13.0, 13.2, and 13.4 nm. The solid lines are calculations of the scattering intensity using the parameters summarized in Table II. The growth parameters that characterize the multilayer roughness are based on the measured PSDs of the substrate and top surface of the multilayer film, shown in Fig. 4. The layer thicknesses are derived from measurements of the specular reflectivity, shown in Fig. 8. We emphasize that all of the input parameters are obtained from independent measurements and the calculations of the nonspecular scattering have no adjustable parameters. The good agreement between the measured scattering and the calculations based on the mea-

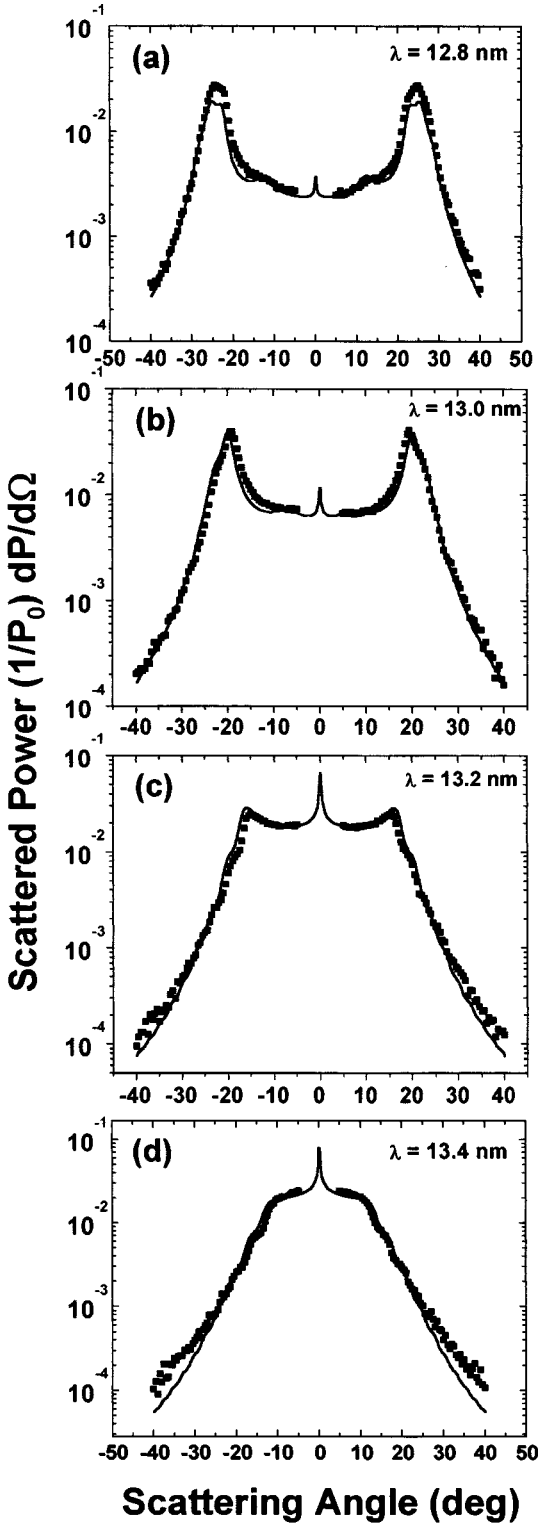


FIG. 11. Nonspecular scattering measured from the canonical Mo–Si multilayer film at normal incidence and for several different wavelengths. Data within four degrees of the specular direction is obscured by the wings of the specularly reflected beam. The solid lines are the scattering distributions predicted by the theory, based on the measured roughness of the multilayer film.

sured roughness provides an important validation of the multilayer growth model and scattering theory.

Finally we address an issue that is central to the application of multilayer coatings in imaging systems: how does the scattering from a multilayer coating compare to the scattering from a single reflecting surface? The unique characteristics of scattering from a multilayer film are illustrated in Fig. 12(a) where we show calculations of the nonspecular scattering *normalized to the specular reflectivity* for three configurations of the surface. In each case the incidence field is unpolarized, has a wavelength of 13.4 nm and is incident normal to the surface. The dotted line corresponds to scattering from a single surface having the roughness of the fused silica substrate shown in Fig. 4. The scattering is featureless and decreases relatively slowly with increasing scattering angle. The rolloff is mostly due to the frequency dependence of the PSD. In contrast, the dashed line represents the scattering from our canonical Mo–Si multilayer film having completely conformal interfaces, that is, the roughness at each interface is identical to the substrate. Here the scattering is characterized by strong interference effects (RNS). The RNS from the conformal multilayer is comparable to the scattering from the single surface at angles less than $\sim$12 degrees. (The scattering from the multilayer is slightly reduced due to an increase of the x-ray wavelength within the film.) Beyond 12 degrees the scattering intensity drops precipitously, as the radiation fields scattered by the different interface interfere destructively. Hence the conformal multilayer scatters like a single surface within the RNS peak, and strongly suppresses scattering at larger angles.

The solid line in Fig. 12(a) shows the scattering calculated for the canonical Mo–Si multilayer film (Fig. 9), where the interfacial roughness is due to both the replication of the substrate and the intrinsic roughness of the film growth process. The scattering exhibits an angular dependence similar to the case of the purely conformal multilayer, but has a nearly sixfold increase in scattering at all angles greater than $\sim$1 degree. The scattering is increased because the multilayer interfaces are rougher than the substrate, particularly towards the top of the film where most of the scattering originates (see Fig. 5). It is also apparent that the finite thickness oscillation is smaller and the decrease in scattering at large angles is less rapid than for the case of the purely conformal multilayer film, behavior which is consistent with the partial correlation of the interfacial roughness. Comparing the realistic multilayer film to the single surface, we find that the scattering is equivalent only at very small angles of $<$1 degree, where the interfacial roughness in the multilayer film is purely conformal. The scattering from the multilayer is enhanced in the region of 1–20 degrees due to the intrinsic roughness of the film, and is suppressed at angles greater than 20 degrees due the interference effects characteristic of RNS.

The total integrated scatter (TIS) within a cone of half-width θ centered about the film normal is plotted as a function of θ in Fig. 12(b). The TIS is normalized to the specular reflectivity, R^{SP}. We observe that, compared to the single surface, the scattering from the multilayer coatings is concentrated in the relatively small annular region within $\sim$14 degrees from the normal, corresponding to the peak in RNS. This suggests a simple *ad hoc* method for estimating the Strehl factor for the multilayer film in the condition

TABLE II. A list of the parameters used to model the nonspecular scattering from the canonical Mo–Si multilayer film. The multilayer growth parameters Ω, ν, and n are described in Sec. VI. These together with the substrate PSD (shown in Fig. 4) define the roughness of the multilayer interfaces. The diffuseness of the interfaces is described by a linear profile of width σ_D (case [c] in Table I). The thickness of the individual layers is τ. The atomic scattering factors are f_1 and f_2, the ρ is the mass density, and W is the atomic weight.

	Growth parameters			Structural parameters		Optical parameters				
	Ω (nm^3)	ν (nm^3)	n	τ (nm)	σ_D (nm)	λ (nm)	f_1	f_2	ρ (g/cm^3)	W (g/mole)
Mo	0.05	2.5	4	2.1	0.3	12.8	14.34	1.270	10.2	95.94
					(Mo-on-Si)	13.0	14.52	1.320		
						13.2	14.67	1.372		
						13.4	14.82	1.424		
Si	0.02	2.5	4	4.75	0.15	12.8	−1.397	0.487	2.33	28.086
					(Si-on-Mo)	13.0	−0.763	0.475		
						13.2	−0.321	0.464		
						13.4	0.023	0.452		

where the RNS and Bragg peaks effectively coincide. Our main assertion is that the effect of roughness can be divided into two regimes, corresponding to the frequencies that scatter within and without the RNS peak. Defining θ_C as the scattering angle at the edge of the RNS peak, we divide the rms roughness into two parts: the low frequency roughness, σ_L, obtained by integrating the PSD over the spatial frequencies less than the value $f_C = \sin\theta_C/\lambda$, and the high frequency roughness σ_H obtained by integrating the PSD over the spatial frequencies greater than f_C. The low-frequency roughness produces scattering within the RNS peak. Since the roughness is conformal at these frequencies, the scattering from the different interfaces is coherent and interferes constructively. Then the reduction in the specular reflectivity due to these low frequencies can be estimated by a simple factor of the Debye–Waller type, $\exp(-16\pi^2\sigma_L^2\cos^2\theta/\lambda^2)$. For the frequencies of the roughness that scatter outside of the RNS peak, the situation is very different. Here the loss due to scattering is significantly reduced for two reasons: (1) the interference of the fields scattered by the different interfaces becomes destructive, and (2) the coherence of the scattering is reduced due to the decreased correlation of the interfacial roughness. Thus in the high-frequency range the primary effect of roughness is not to produce scattering, but instead to increase the transmission of the interfaces, thereby resulting in a larger penetration depth for the incident specular field and correspondingly greater absorption. This is in fact the same loss mechanism as the case of a diffuse interface broadened by intermixing or chemical reaction, and can be treated in a similar way. Specifically, we estimate the reduction of the specular reflectivity due to the high-frequency roughness by including a contribution from the high-frequency roughness in the interface width $\sigma_D \rightarrow \sqrt{\sigma_D^2+\sigma_H^2}$. This modified value of σ can be applied in the matrix method described in Appendix C to calculate the specular reflectivity $R^{\mathrm{SP}}(\sigma)$ of the multilayer coating reduced by both diffuseness and high-frequency roughness. Then the Strehl factor describing the reduction in the specular reflectivity due to interfacial roughness is estimated as

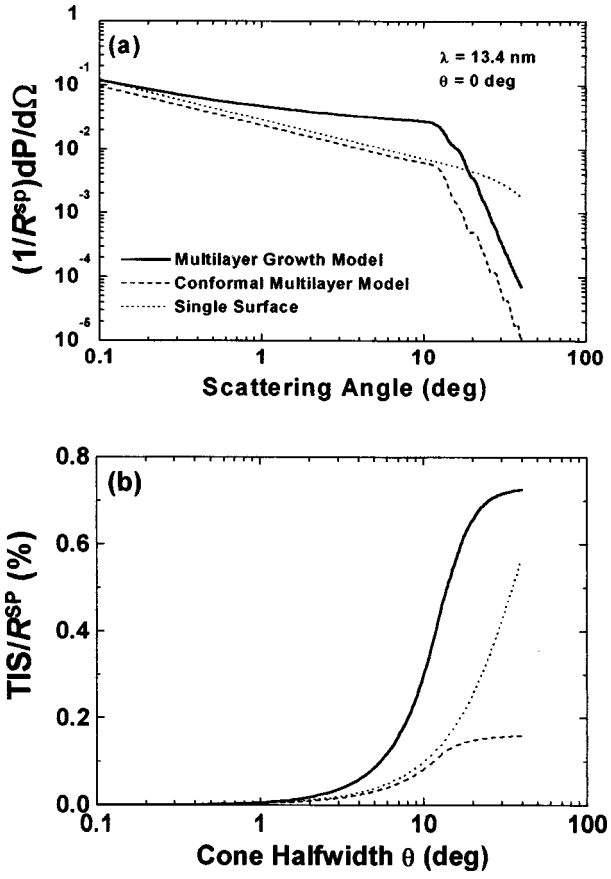


FIG. 12. (a) Calculations of nonspecular scattering from three different surfaces, normalized to the specular reflectivity. The dotted line is a single surface having the roughness of the fused silica substrate shown in Fig. 4. The dashed line corresponds to an ideally conformal Mo–Si multilayer film where the roughness of each interface is identical to the roughness of the single surface. The solid line is a realistic Mo–Si multilayer film having interfacial roughness due to both replication of the substrate roughness and the intrinsic roughness of the growth process. (b) The total integrated scattering (TIS) within a cone of halfwidth θ, calculated for the three surfaces and normalized to the specular reflectivity. Nearly all of the scattering from the multilayer films is within 20 degrees of the specular direction.

$$S=\frac{R^{\mathrm{SP}}(\sigma=\sqrt{\sigma_D^2+\sigma_H^2})}{R^{\mathrm{SP}}(\sigma=\sigma_D)}\exp\left(-\frac{16\pi^2\sigma_L^2\cos^2\theta}{\lambda^2}\right). \quad (64)$$

As an example, consider the scattering from the canonical multilayer coating calculated from the PSDs of Fig. 4 and shown as the solid line in Fig. 12(a). The separation of the low- and high-frequency ranges is found by inspection to be at a scattering angle of $\theta_C \cong 14$ deg, corresponding to a frequency of $f_C = 0.02\ \mathrm{nm}^{-1}$. We integrate the PSD of the top surface of the multilayer coating as prescribed in Eq. (46) over the low- and high-frequency ranges to obtain $\sigma_L = 0.09$ nm and $\sigma_H = 0.15$ nm, respectively. Using these values in Eq. (64) in conjunction with the structural parameters of Table II, we obtain a Strehl factor of 0.988 for normal incidence and $\lambda = 13.4$ nm.

We emphasize that the method described by Eq. (64) for estimating the Strehl factor is only necessary when there is a strong reflected field, such that dynamical effects associated with the interference of the fields scattered by the different interfaces are important. When the incident field is not near the Bragg peak, or if the specular reflectivity is small (<0.1) then the scattering process is well approximated by kinematical theory, that is, the scattering does not significantly alter the configuration of the incident field. In this case, the Strehl factor can be estimated from either the Debye–Waller factor of Eq. (29) or the matrix method of Appendix C with modified Fresnel coefficients, using a rms roughness σ obtained by integrating the PSD over all frequencies.

VIII. MODELING THE PERFORMANCE OF A SOFT X-RAY IMAGING SYSTEM

We now turn our attention back to the problem of modeling the nonspecular scattering in a distributed optical system consisting of multilayer coated optics. The conventional method for characterizing the performance of an imaging system is to measure the optical transfer function (OTF). The OTF is the ratio of the image intensity to the object intensity at a particular spatial frequency, and is also the Fourier transform of the point spread function. It is only defined for the case of incoherent illumination, where there is a purely linear relationship between the Fourier transforms of the intensities in the image and object planes. However, we have shown that the effect of scattering is to produce a point spread function, $\mathrm{PSF}^{\mathrm{sc}}$, that is independent of the coherence of the illumination. Then the Fourier transform of $\mathrm{PSF}^{\mathrm{sc}}$ yields an $\mathrm{OTF}^{\mathrm{sc}}$ that describes the modulation of the image intensity due to scattering under any illumination conditions. In particular, collecting the results of Secs. II and III, the intensity at the image plane is given by

$$\langle I_1(\mathbf{s}_1)\rangle = I_1^0(\mathbf{s}_1) * \kappa \mathrm{PSF}^{\mathrm{sc}}(\mathbf{s}_1), \quad (65)$$

where I_1^0 is the image produced by the optical system in the absence of scattering and

$$\mathrm{PSF}^{\mathrm{sc}}(\mathbf{s}_1) = \frac{1}{\kappa}\left(\prod_n S_n\delta(\mathbf{s}_1) + \sum_n \frac{\alpha_{nx}\alpha_{ny}}{R_n^{\mathrm{SP}}R^2}\sum_{\hat{a}}\frac{dP_n(\hat{m},\hat{a};\hat{n},\hat{e})}{d\Omega}\right), \quad (66)$$

is the point spread function due to scattering. The quantity in brackets is calculated using the scattering theory presented in the previous section; the angular scattering distributions $dP_m/d\Omega$ are obtained from Eq. (62), where the scattering vector is related to the point $\mathbf{s}_1$ in the image field through Eq. (32). The Strehl factor S_n, corresponding to the ratio between the specular reflectivity of the nth surface with and without roughness, can be estimated from the PSD of the top surface of the multilayer coating using Eq. (64). The normalizing factor κ is

$$\kappa = \int\int\left(\prod_n S_n\delta(\mathbf{s}_1) + \sum_n \frac{\alpha_{nx}\alpha_{ny}}{R_n^{\mathrm{SP}}R^2}\sum_{\hat{a}}\frac{dP_n(\hat{m},\hat{a};\hat{n},\hat{e})}{d\Omega}\right)d\mathbf{s}_1, \quad (67)$$

where the integration is over the image field. This factor accounts for the loss of image intensity due to the scattering outside of the image field and the increased absorption arising from the high-frequency interfacial roughness.

Taking the Fourier transform of (65) we obtain

$$I_1(\mathbf{f})/I_1^0(\mathbf{f}) = \mathrm{OTF}^{\mathrm{sc}}(\mathbf{f}), \quad (68)$$

where $I_1(\mathbf{f})$ is the Fourier transform of the image intensity, and

$$\mathrm{OTF}^{\mathrm{sc}}(\mathbf{f}) = \frac{1}{\kappa}\prod_n S_n + \frac{1}{\kappa}\int\int\sum_n \frac{\alpha_{nx}\alpha_{ny}}{R_n^{\mathrm{SP}}R^2} \times \sum_{\hat{a}}\frac{dP_n(\hat{m},\hat{a};\hat{n},\hat{e})}{d\Omega}\exp(-2\pi i\mathbf{s}_1\cdot\mathbf{f})d\mathbf{s}_1. \quad (69)$$

Equations (68) and (69) provide the basis for relating the optical performance in the presence of scattering, as characterized by the $\mathrm{OTF}^{\mathrm{sc}}$, to the key structural parameters of the optical surfaces, including the surface finish of the substrates and the roughness of the multilayer coatings. This allows us to determine specifications for the roughness of the substrates and coatings, given certain performance requirements. The operational procedure for deriving such specifications consists of first defining a minimally acceptable $\mathrm{OTF}^{\mathrm{sc}}(\mathbf{f})$ that will allow the production of useful images. Then, taking into account the intrinsic roughness of the multilayer coating, we determine the limits of substrate roughness required to satisfy the specified value of $\mathrm{OTF}^{\mathrm{sc}}$ at each frequency.

A significant simplification is possible in the limit of very smooth substrates, where the scattering is dominated by the intrinsic roughness of the multilayer coatings. In this case the scattering is uniform out to fairly large angles ($\sim$10 degrees) because the PSD of the intrinsic roughness is flat for frequencies less than $\sim 10^{-1}\ \mathrm{nm}^{-1}$ (see Fig. 2). When the scattered light is uniformly distributed throughout

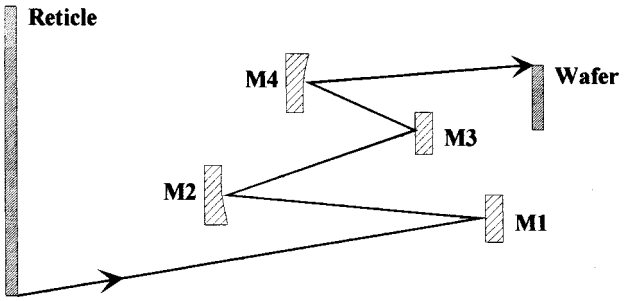


FIG. 13. Schematic diagram of a hypothetical soft x-ray imaging system for EUV lithography. The optical system, consisting of four multilayer-coated mirrors, projects an image of the reticle onto the wafer.

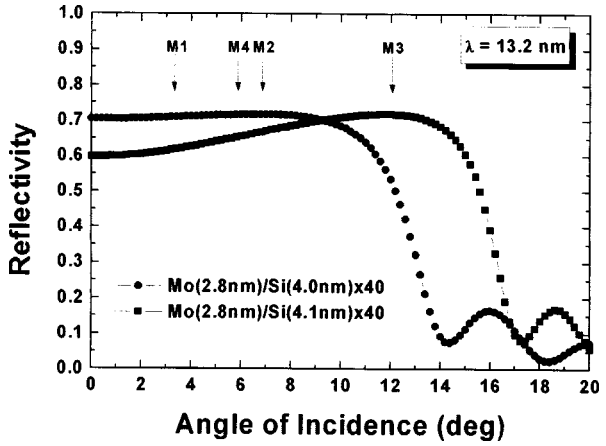


FIG. 14. Calculated reflectivity of the two Mo–Si multilayer coatings used in the hypothetical imaging system. The arrows indicate the angles of incidence of the principal ray on each of the optical surfaces.

the image field it is called ''veiling glare.'' Under these conditions the second term in Eq. (69) is negligible and the OTF^{sc} can be approximated as having a constant value of S/κ, that is, the image contrast is reduced by a constant amount at all but the lowest frequencies.

As an example, we model the nonspecular scattering in the distributed optical system shown schematically in Fig. 13. The imaging system is purely hypothetical, but is based on generic designs being considered for applications in EUV lithography.[42,43] It consists of four reflecting surfaces and, although designed to be used as a ring field, we will consider for the purpose of modeling scattering that the image field is a square of width 2.5 cm. A requirement for lithography applications is that the imaging system be telecentric at the wafer, that is, the principal ray is parallel to the optical axis for all points in the image field. A consequence of telecentricity is that the exit pupil is infinitely large and is located an infinite distance from the image plane. However, we showed in Sec. III that the effects of scattering on the transfer function are independent of the position of the actual exit pupil. We can choose to evaluate the transfer function at any position on the image side of the last optical surface, provided that we apply the correct scaling factors. In this case we evaluate the transfer function at the position of the last mirror (M4), located 23 cm from the image plane. The scaling factors and angles of incidence of the principal ray, as determined by ray tracing calculations, are listed in Table III. We assume that the optical surfaces are coated with Mo–Si multilayer films designed for an operating wavelength of 13.2 nm. Since the dispersion of the angles of incidence on any given optical surface is small, the multilayer coatings will have uniform bilayer spacing. The multilayer structure that optimizes the reflectivity for the first, second, and fourth surfaces is $[\mathrm{Mo}(2.80\ \mathrm{nm})/\mathrm{Si}(4.00\ \mathrm{nm})]\times 40$. The third surface requires a slightly different design of $[\mathrm{Mo}(2.80\ \mathrm{nm})/\mathrm{Si}(4.10\ \mathrm{nm})]\times 40$, due to the larger angle of incidence. The calculated reflectivity of these designs at a wavelength of 13.2 nm is shown in Fig. 14. These calculations include interdiffusion at the interfaces of $\sigma_D = 0.3$ nm for the Mo-on-Si interface and $\sigma_D = 0.15$ nm for the Si-on-Mo interface. All roughness and oxidation is neglected, and hence the reflectivity values are slightly overestimated.

TABLE III. Design parameters for a hypothetical soft x-ray imaging system consisting of four mirrors. The scaling factors, α_X and α_Y, are set equal and θ is the angle of incidence of the principal ray as measured from the normal to the optical surface. Also listed are the Strehl factors S_n for the multilayer coatings calculated at an x-ray wavelength of 13.2 nm.

Mirror	$\alpha_X = \alpha_Y$	θ (deg)	Multilayer structure	S_n
M1	1.7	3.0	$[\mathrm{Mo}(2.8\ \mathrm{nm})/\mathrm{Si}(4.0\ \mathrm{nm})]\times 40$	0.985
M2	0.9	7.0	$[\mathrm{Mo}(2.8\ \mathrm{nm})/\mathrm{Si}(4.0\ \mathrm{nm})]\times 40$	0.985
M3	1.8	12.0	$[\mathrm{Mo}(2.8\ \mathrm{nm})/\mathrm{Si}(4.1\ \mathrm{nm})]\times 40$	0.985
M4	1.0	6.0	$[\mathrm{Mo}(2.8\ \mathrm{nm})/\mathrm{Si}(4.0\ \mathrm{nm})]\times 40$	0.986

The PSF^{sc} is calculated using the methodology described above. The plane of incidence of the principal ray is the $y-z$ plane for all optical surfaces, and the radiation incident on each surface is assumed to be unpolarized. We assume that the roughness of the substrates is the same as the superpolished fused silica flat having the PSD shown in Fig. 4. Furthermore, we assume that the roughness of the Mo–Si multilayer coatings is equivalent to the ''canonical'' film discussed in the previous sections, and that the roughness is described by the thin film growth model and the corresponding growth parameters listed in Table II. Consequently, the multilayer coating contributes roughness to the optical surfaces, causing the rms roughness to increase from the substrate to a final value of $\sigma = 0.18$ nm at the top surface. The Strehl factors S_n for the coatings are determined from Eq. (64), where the rms roughness is divided into low- and high-frequency components of $\sigma_L = 0.09$ nm and $\sigma_H = 0.15$ nm, corresponding to those frequencies that scatter within and out of the RNS peak, respectively. The calculated values of the Strehl factors are listed in Table III.

The PSF^{sc} calculated over a 2.5×2.5-cm image field is shown in Fig. 15. The delta function corresponding to the specular field [the first term in Eq. (66)] has been omitted from the plot. The PSF^{sc} exhibits a peak at the optical axis which drops off to a relatively constant level of scattering at distances beyond ~ 0.5 cm from the optical axis. The edge of the image field corresponds to a scattering angle of only ~ 3 degrees. Since the strong resonant scattering (RNS) from the multilayer coatings exists out to an angle of ~ 14 degrees, it is evident that the image field intercepts only

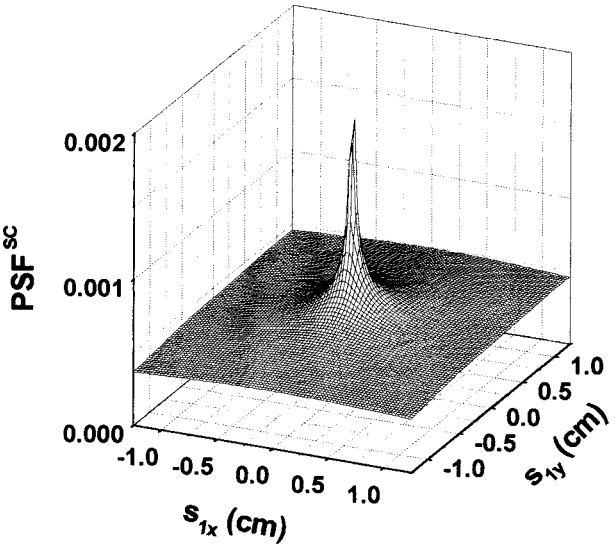


FIG. 15. The point spread function due to scattering (PSF^{sc}) of the hypothetical imaging system, calculated assuming that the roughness of the Mo–Si multilayer coatings is equivalent to the canonical film shown in Fig. 4. The delta-function component of PSF^{sc} is not shown.

a small fraction of the total scattered radiation. In this case the Strehl factor S for the imaging system is 0.94, corresponding to a 6% decrease in the specular image intensity. However the fraction of the specular intensity scattered into the image field is only 0.3% ($S/\kappa=0.997$). The OTF^{sc} obtained by taking the Fourier transform of the PSF^{sc} is shown in Fig. 16. The OTF^{sc} drops rapidly to a constant value of $S/\kappa=0.997$ for frequencies greater than $\sim 2.5\ \mathrm{cm}^{-1}$. This behavior is characteristic of veiling glare and illustrates the dominant effect of the intrinsic roughness of the multilayer coatings in this example; the scattering is fairly uniformly distributed throughout the image field and hence reduces the

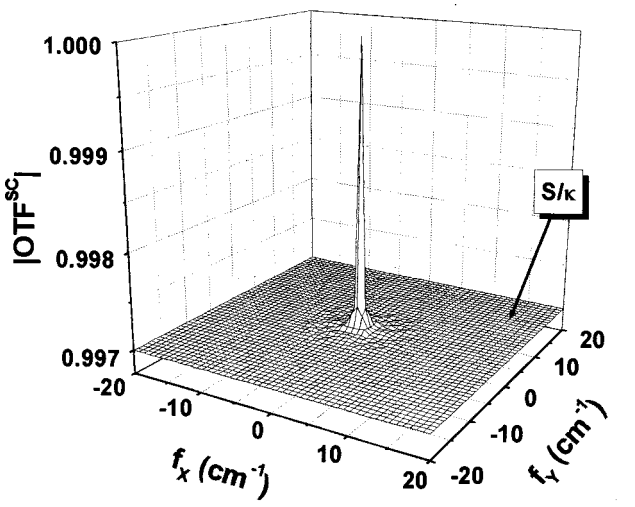


FIG. 16. The optical transfer function due to scattering (OTF^{sc}) of the hypothetical imaging system. The nonspecular scattering reduces the OTF of the imaging system by a factor $S/\kappa=0.997$ at all but the lowest frequencies.

image contrast by a constant amount at all but the lowest frequencies.

At the risk of oversimplification, we can use our example to make some general comments regarding the effect of scattering in the performance of multilayer-coated imaging systems. High performance imaging systems for soft x rays require a numerical aperture of ~ 0.1 or less in order to have a reasonable depth of focus ($\geqslant 1\ \mu\mathrm{m}$). Consequently the range of scattering angles subtended by the image field will be limited to a few degrees, and all of this scattering will be within the RNS peak of the multilayer coatings. The components of roughness that scatter into the image field are in the midspatial frequency range of $\sim 10^{-5}/\lambda - 10^{-1}/\lambda\ \mathrm{nm}^{-1}$; it is this range of frequencies that is responsible for the decrease in image resolution and contrast. In this range the roughness of the multilayer coatings is dominated by the replication of the substrate roughness and includes (at the higher end of the frequency range) some intrinsic roughness of the multilayer film. Consequently, the scattering within the image field for a multilayer-coated imaging system is expected to be comparable to, and slightly greater than, an equivalent system having single reflecting surfaces. Most of the intrinsic roughness of the multilayer-coatings occurs at higher frequencies ($10^{-3}-10^{-1}\ \mathrm{nm}^{-1}$) that will scatter outside of the image field. Hence the main detrimental effect of the intrinsic roughness of the coatings is to reduce the throughput of the imaging system.

IX. CONCLUSION

In summary, we have presented a theoretical framework for modeling nonspecular scattering in a soft x-ray imaging system consisting of multilayer-coated reflecting optics. The theory directly relates the image degradation due to scattering to the statistical properties of the interfacial roughness of the multilayer coatings. Consequently, the theory can be a versatile tool for tasks such as modeling the performance of optical components of known roughness, deriving specifications for the roughness of optical substrates and coatings, and comparing the performance of different optical designs. When applying the theory in practice, it is important to recall the numerous approximations that were invoked, and to observe the restrictions imposed by these approximations. We summarize the key approximations of the theory in Appendix A.

Throughout the course of this presentation we have attempted to illustrate the theoretical formalism with realistic examples and analysis of experimental data whenever possible. In particular, the good agreement between the measured scattering from Mo–Si multilayer films and the calculations based on the measured roughness of the films serves to validate our treatment of the scattering problem. The ultimate test of the theory will require the complete characterization of a distributed imaging system, corresponding to independent measurements of the PSF and the roughness of the multilayer-coated optics. This will be the goal of future work.

In addition to its practical importance, our theoretical investigation has broadened the fundamental understanding

of scattering in a distributed imaging system and, in particular, the effects of multilayer coatings. The most important results are summarized below:

(1) Image formation in a distributed optical system can be described as a convolution of the image formed in the absence of scattering with a PSF^{sc} due to scattering.

(2) The PSF^{sc} is independent of the coherence state of the object.

(3) The roughness of a multilayer coating originates from both the intrinsic roughness of the growth process and the replication of the roughness of the substrate. At the lowest spatial frequencies the multilayer film exactly replicates the substrate roughness. In the range $\sim 10^{-3}-10^{-1}$ nm^{-1} the roughness increases from the substrate to the top surface of the multilayer, and the roughness of the different interfaces is partially correlated. At higher spatial frequencies the multilayer film tends to smooth the substrate roughness.

(4) The nonspecular scattering from a multilayer coating exhibits strong interference effects due to the partial correlation of the roughness of the interfaces. This produces a resonance in the nonspecular scattering (RNS) in the vicinity of the specular Bragg peak. In particular, the RNS peak and the Bragg peak are coincident when the incident field satisfies the Bragg condition for specular reflection.

(5) In a soft x-ray imaging system, the roughness in the midspatial frequency range of $\sim 10^{-5}/\lambda - 10^{-1}/\lambda$ nm^{-1} produces the scattering that reaches the image field. In this frequency range the roughness of the multilayer coatings is due to replication of the substrate roughness and, to a small extent, the intrinsic roughness of the multilayer film. Hence the scattering within the image field for a multilayer-coated imaging system is expected to be slightly greater than an equivalent system having single reflecting surfaces, and depends predominantly on the roughness of the optical substrates.

(6) The main detrimental effect of the intrinsic roughness of the multilayer coatings is to scatter radiation outside the image field, thereby reducing the throughput of the optical system.

We conclude with a comment regarding the impact of scattering on the performance of soft x-ray imaging systems. The example that we have presented using a hypothetical imaging system designed for EUV lithography shows a very minor degradation of performance due to scattering; the OTF is reduced by only 0.3% and the throughput is decreased by 6%. One might be tempted to infer that scattering is not a significant problem in such an imaging system. In fact, our example demonstrates that scattering can be limited to acceptably low levels in a soft x-ray imaging system *if the optical substrates can be fabricated with roughness equivalent to the best superpolished flats measured to date*. Our analysis indicates that, once these ultrasmooth substrates are available, the existing multilayer-coating technology is capable of producing soft x-ray imaging systems that have acceptably low levels of scattering. This conclusion is supported by experiments performed[44] on a soft x-ray telescope designed for normal incidence operation at $\lambda = 6.35$ nm. The optical components were coated with Co–C multilayer films, and were measured to have roughness similar to the PSDs shown in Fig. 4. Analysis of the images obtained with the telescope showed no measurable scattering. Although these results are encouraging, they must be considered a best case scenario. We anticipate that fabricating sufficiently smooth figured optics will be a significant technical challenge, particularly as the size of the optical components increases and the figures become aspherical. In practice the effect of scattering on the performance of the soft x-ray imaging system will define the acceptable limits of roughness for these optics.

ACKNOWLEDGMENTS

We thank F. Weber for fabricating the Mo–Si multilayer films. The atomic force microscopy measurements were performed by S. L. Baker. We are grateful to J. H. Underwood for his assistance in designing and performing the scattering measurements on Beamline 6.3.2 at the Advanced Light Source. This work was supported by the Department of Energy under Contract DE-AC03-76SF00098 and CRADA TC-0191/0192-92.

APPENDIX A: SUMMARY OF APPROXIMATIONS

We summarize below the key assumptions and approximations underlying the theoretical development presented in this paper.

(a) The angle between the principal ray and any other ray that propagates through the imaging system is small. Specifically, if we denote the angle as φ, then the approximation is

$$\sin^2 \varphi \ll 1. \tag{2}$$

We call this the ''small angle approximation.''

(b) For a point object, the pupil function is independent of the location of the point in the object field. In this case the point spread function is independent of the position of the Gaussian image point, and the system is called ''isoplanatic.'' In practice, the assumption of isoplanacity restricts the applicability of the transfer function formalism to objects of small spatial extent.

(c) The scattering is weak so that multiple scattering and shadowing effects can be neglected. This is called the ''Born approximation.'' This approximation is generally valid for x-ray wavelengths at angles of incidence away from the critical angle for total external reflection.

(d) The surface height $H(X,Y)$ is a Gaussian random variable, is stationary and is ergodic in the sense that the ensemble average can be replaced by an average over the illuminated surface area.

(e) The deviations of the surface height $H_n(X_n, Y_n)$ from the ideally smooth surface are small compared to the radiation wavelength such that $2k \cos \theta_n H_n(X_n, Y_n) \ll 1$ for all X_n, Y_n. We call this the ''small roughness approximation.'' A necessary consequence of the small roughness approximation is that the power scattered into the nonspecular field is small compared to the specularly reflected power, a condition that is implicitly satisfied by high-performance optics.

APPENDIX B: RELAXATION OF A ROUGH SURFACE BY SURFACE DIFFUSION

Consider a rough surface $H(X,Y)$ where the growth units (e.g., atoms) have mobility to move between sites on the surface. The chemical potential at each point on the surface depends on the curvature at that point. The gradient of the chemical potential is the driving force for the surface diffusion that causes the smoothing of the surface. To express these concepts quantitatively let us model the local curvature of the surface at a particular point as a sphere of radius R. Then the curvature at that point is

$$\nabla^2 H(X,Y) = -\frac{2}{R}. \tag{B1}$$

The chemical potential of the point $H(X,Y)$ is found by letting the radius of the spherical feature change by an infinitesimal amount ΔR. The chemical potential is given by

$$\mu = \frac{\Delta E_s}{\Delta N}, \tag{B2}$$

where ΔE_s is the change in the surface energy and ΔN is the change in the number of atoms within the sphere. The change in the surface energy is just proportional to the change in the surface area

$$\Delta E_S = 8\pi R \xi \Delta R, \tag{B3}$$

where ξ is the surface energy per unit area. The change in the number of atoms in the sphere is

$$\Delta N = \frac{4\pi R^2 \Delta R}{V_0}, \tag{B4}$$

where V_0 is the atomic volume. Combining Eqs. (B1)–(B4) we obtain

$$\mu = \frac{2\xi V_0}{R} = -\xi V_0 \nabla^2 H(X,Y). \tag{B5}$$

This shows explicitly that the chemical potential is proportional to the local curvature of the surface.

The driving force, F, for surface diffusion is the gradient of the chemical potential

$$F = -\nabla \mu. \tag{B6}$$

From the Nernst–Einstein relation[45] the mean velocity of an atom on the surface is given by

$$\langle v \rangle = \frac{F D_S}{kT} = -\frac{D_S \nabla \mu}{kT} \tag{B7}$$

where D_S is the surface diffusion coefficient, k is Boltzmann's constant, and T is the surface temperature. The flux of atoms, J, on the surface is

$$J = \frac{\langle v \rangle}{V_0^{2/3}}. \tag{B8}$$

The surface flux is the mechanism of mass transport through which the smoothing of the surface takes place. However, the surface height at a given point can change only if there is a divergence of flux. Then the change in the surface height per unit time is given by

$$\frac{dH(X,Y)}{dt} = -V_0 \nabla \cdot J. \tag{B9}$$

Combining Eqs. (B5)–(B9) we obtain

$$\frac{dH(X,Y)}{dt} = -\frac{\xi D_S V_0^{4/3}}{kT} \nabla^4 H(X,Y). \tag{B10}$$

In the thin film growth model the evolution of the surface is measured as a function of film thickness, τ, and not time. However, time and thickness are simply related through the deposition rate $r_D = d\tau/dt$. Then we can rewrite Eq. (B10) as

$$\frac{dH(X,Y)}{d\tau} = -\frac{\xi D_S V_0^{4/3}}{r_D kT} \nabla^4 H(X,Y). \tag{B11}$$

Comparing this result to Eq. (36) we identify the relaxation parameter for the thin film growth model

$$\nu = \frac{\xi D_S V_0^{4/3}}{r_D kT}. \tag{B12}$$

Not surprisingly, the rate at which the smoothing occurs is proportional to the surface diffusion coefficient. The temperature dependence of the relaxation parameter is dominated by the surface diffusion coefficient which is proportional to $\exp(-E_A/kT)$, where E_A is the activation energy. In general the relaxation of the surface will be enhanced at higher temperature and lower deposition rate.

APPENDIX C: MATRIX METHOD FOR PROPAGATING SPECULAR FIELDS IN A MULTILAYER FILM

The multilayer scattering theory requires as input parameters the incident field amplitudes, E_i^+ and E_j^-, and the propagation factors ϕ_i^r and ϕ_i^t for each interface of the multilayer film in the absence of roughness. These are most easily calculated using a well-known matrix approach[46] to analyze the propagation of the specular fields within the multilayer film. First we define a scattering matrix, $\mathbf{T}_i$, that relates the specular fields across the ith interface according to

$$\begin{pmatrix} E_j^- \\ E_j^+ \end{pmatrix} = \mathbf{T}_i \begin{pmatrix} E_i^- \\ E_i^+ \end{pmatrix}. \tag{C1}$$

It is easily shown that

$$\mathbf{T}_i = \frac{1}{t_{ji}} \begin{pmatrix} 1 & r_{ji} \\ r_{ij} & t_{ji}t_{ij} + r_{ji}^2 \end{pmatrix}, \tag{C2}$$

where t_{ji} and r_{ji} are the specular transmission and reflection amplitudes, respectively, for the ith interface. For a compositionally abrupt interface the transmission and reflection amplitudes are given by the Fresnel equations

$$\begin{aligned} &S \text{ polarization: } t_{ji}^0 = \frac{2\chi_j}{\chi_j + \chi_i}, \quad r_{ji}^0 = \frac{\chi_j - \chi_i}{\chi_j + \chi_i} \\ &P \text{ polarization: } t_{ji}^0 = \frac{2\sqrt{\epsilon_j \epsilon_i}\chi_j}{\epsilon_i \chi_j + \epsilon_j \chi_i}, \quad r_{ji}^0 = \frac{\epsilon_i \chi_j - \epsilon_j \chi_i}{\epsilon_i \chi_j + \epsilon_j \chi_i} \end{aligned} \tag{C3}$$

where

$$\chi_i = k\sqrt{\epsilon_i - n_X^2 - n_Y^2}. \tag{C4}$$

Diffuseness at the interface, characterized by a composition gradient $W^D(Z)$, modifies the transmission and reflection amplitudes according to the well-known formulas first derived by Nevot and Croce[47]

$$\begin{aligned} t_{ji} &= t_{ji}^0 \frac{1}{w^D(k_j n_{jZ}^+ - k_i n_{iZ}^+)} \\ r_{ji} &= r_{ji}^0 \frac{w^D(k_j n_{jZ}^+ + k_i n_{iZ}^+)}{w^D(k_j n_{jZ}^+ - k_i n_{iZ}^+)}. \end{aligned} \tag{C5}$$

Specific examples of the function $w^D(q_z)$, corresponding to several simple interface profiles, are listed in Table I.

The scattering matrix describes the propagation of the specular field across an interface. To propagate the field through the ith layer of thickness τ_i we define the propagation matrix $\mathbf{P}_i$

$$\mathbf{P}_i = \begin{pmatrix} \exp(-i\varphi_i) & 0 \\ 0 & \exp(i\varphi_i) \end{pmatrix}, \tag{C6}$$

where

$$\varphi_i = k\tau_i\sqrt{\epsilon_i - n_X^2 - n_Y^2}. \tag{C7}$$

The propagation of the specular field from layer to layer in the multilayer film is represented by a series of matrix multiplications. In particular, the field above interface i is related to the field below the underlying interface m through a matrix $\mathbf{A}^{i,m}$ according to

$$\begin{pmatrix} E_{i+1}^- \\ E_{i+1}^+ \end{pmatrix} = \mathbf{A}^{i,m} \begin{pmatrix} E_m^- \\ E_m^+ \end{pmatrix}, \tag{C8}$$

where

$$\mathbf{A}^{i,m} = \begin{pmatrix} A_{11}^{i,m} & A_{21}^{i,m} \\ A_{12}^{i,m} & A_{22}^{i,m} \end{pmatrix} = \mathbf{T}_i\mathbf{P}_i\mathbf{T}_{i-1}\cdots\mathbf{T}_{m+1}\mathbf{P}_{m+1}\mathbf{T}_m. \tag{C9}$$

The transmission and reflection amplitudes for the entire multilayer film, t_{ML} and r_{ML}, are obtained from $\mathbf{A}^{N,0}$, corresponding to propagation through all N interfaces. In particular,

$$t_{\mathrm{ML}} = \frac{1}{A_{11}^{N,0}}, \quad r_{\mathrm{ML}} = \frac{A_{21}^{N,0}}{A_{11}^{N,0}}, \tag{C10}$$

and the specular transmission and reflectance are correspondingly,

$$T^{\mathrm{SP}} = |t_{\mathrm{ML}}|^2, \quad R^{\mathrm{SP}} = |r_{\mathrm{ML}}|^2. \tag{C11}$$

Knowledge of the transmitted amplitude allows us to determine the specular fields above and below each interface using the relations

$$\begin{pmatrix} E_{i+1}^- \\ E_{i+1}^+ \end{pmatrix} = \mathbf{A}^{i,0} \begin{pmatrix} t_{\mathrm{ML}} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} E_i^- \\ E_i^+ \end{pmatrix} = \mathbf{P}_i\mathbf{A}^{i-1,0} \begin{pmatrix} t_{\mathrm{ML}} \\ 0 \end{pmatrix}.$$

We introduced the factors ϕ_i^r and ϕ_i^t in the scattering theory to account for the propagation of the scattered field from the ith interface to the top surface of the multilayer film. Consider a field of unit amplitude initially scattered at the ith interface towards the top of the film. This produces a field exiting the front surface of the film of amplitude ϕ_i^r and a field exiting the bottom surface into the substrate of amplitude E_s. Then the matrix equation that describes the propagation of this scattered radiation through the system of layers is given by

$$\begin{pmatrix} 0 \\ \phi_i^r \end{pmatrix} = \mathbf{A}^{N,i+1}\mathbf{P}_{i+1}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mathbf{A}^{i,0}\begin{pmatrix} E_s \\ 0 \end{pmatrix}\right], \tag{C12}$$

where the matrices $\mathbf{A}$ and $\mathbf{P}$ now correspond to a scattered plane wave mode. Similarly, for a field of unit amplitude initially scattered towards the bottom of the multilayer film we can write a matrix equation:

$$\begin{pmatrix} 0 \\ \phi_i^t \end{pmatrix} = \mathbf{A}^{N,i}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathbf{P}_i\mathbf{A}^{i-1,0}\begin{pmatrix} E_S \\ 0 \end{pmatrix}\right]. \tag{C13}$$

Solving these matrix equations for ϕ_i^r and ϕ_i^t we obtain

$$\begin{aligned} \phi_i^r &= e^{i\varphi_i}(A_{22}^{N,i+1} - r_{\mathrm{ML}}A_{12}^{N,i+1}) \\ \phi_i^t &= A_{21}^{N,i} - r_{\mathrm{ML}}A_{11}^{N,i}. \end{aligned} \tag{C14}$$

[1] G. D. Kubiak and D. R. Kania, OSA Trends in Optics and Photonics Vol. 4, *Extreme Ultraviolet Lithography* (Optical Society of America, Washington, DC, 1996).
[2] E. Spiller, in *Soft X-Ray Optics* (SPIE, Bellingham, 1994), p. 235.
[3] D. G. Stearns, J. Appl. Phys. **71**, 4286 (1992).
[4] D. E. Savage, J. Kleiner, N. Schimke, Y.-H. Phang, T. Jankowski, J. Jacobs, R. Kariotis, and M. G. Lagally, J. Appl. Phys. **69**, 1411 (1991).
[5] V. Holy and T. Baumbach, Phys. Rev. B **49**, 668 (1994); V. Holy *et al.*, *ibid.* **47**, 896 (1993).
[6] A. P. Payne and B. M. Clemens, Phys. Rev. B **47**, 2289 (1993).
[7] T. Salditt, D. Lott, T. H. Metzger, J. Peisl, G. Vignaud, P. Hoghoj, O. Scharpf, P. Hinze, and R. Lauer, Phys. Rev. B **54**, 5860 (1996); T. Salditt, T. H. Metzger, and J. Peisl, Phys. Rev. Lett. **73**, 2228 (1994).
[8] J. B. Kortright, J. Appl. Phys. **70**, 3620 (1991).
[9] R. Paniago, H. Homma, P. C. Chow, S. C. Moss, Z. Barnea, S. S. P. Parkin, and D. Cookson, Phys. Rev. B **52**, 52 (1995).
[10] D. R. Lee, Y. J. Park, D. Kim, Y. H. Jeong, and K. B. Lee, Phys. Rev. B. (submitted for publication).
[11] E. L. Church and P. Z. Takacs, Opt. Eng. (Bellingham) **34**, 353 (1995); E. L. Church and P. Z. Takacs, Appl. Opt. **32**, 3344 (1993).
[12] J. E. Harvey, Appl. Opt. **34**, 3715 (1995); J. E. Harvey, K. L. Lewotsky, and A. Kotha, Opt. Eng. (Bellingham) **35**, 2423 (1996).
[13] S. Singh, H. Solak, and F. Cerrina, Rev. Sci. Instrum. **67**, 3355 (1996).
[14] M. Born and E. Wolf, in *Principles of Optics*, 6th ed. (Pergamon, New York, 1987), p. 480.
[15] D. G. Stearns, Appl. Phys. Lett. **62**, 1745 (1993).
[16] D. G. Stearns, J. Appl. Phys. **65**, 491 (1989).
[17] B. R. Frieden, in *Probability, Statistical Optics, and Data Testing*, 2nd ed. (Springer, New York, 1991), p. 75.
[18] For a review see, A. L. Barabasi and H. E. Stanley, in *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995), p. 19.
[19] W. M. Tong and R. S. Williams, Annu. Rev. Phys. Chem. **45**, 401 (1994).
[20] S. F. Edwards and D. R. Wilkinson, Proc. R. Soc. London, Ser. A **381**, 17 (1982).
[21] C. Herring, J. Appl. Phys. **21**, 301 (1959).
[22] T. Salditt, D. Lott, T. H. Metzger, J. Peisl, R. Fischer, J. Zweck, P. Hoghoj, O. Scharpf, and G. Vignaud, Europhys. Lett. **36**, 565 (1996).
[23] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).
[24] R. P. U. Karunasiri, R. Bruinsma, and J. Rudnick, Phys. Rev. Lett. **62**, 788 (1989).
[25] C. Tang, S. Alexander, and R. Bruinsma, Phys. Rev. Lett. **64**, 772 (1990).
[26] J. A. Thornton, Thin Solid Films **45**, 387 (1977).
[27] R. Messier and J. E. Yehoda, J. Appl. Phys. **58**, 3739 (1985).
[28] D. J. Miller, K. E. Gray, R. T. Kampwirth, and J. M. Murduck, Europhys. Lett. **19**, 27 (1992).
[29] E. L. Church, Appl. Opt. **27**, 1518 (1988).

[30] D. G. Stearns, R. S. Rosen, and S. P. Vernon, J. Vac. Sci. Technol. A **9**, 2662 (1991).
[31] Y. Cheng, D. J. Smith, M. B. Stearns, and D. G. Stearns, J. Appl. Phys. **72**, 5165 (1992).
[32] M. B. Stearns, C.-H. Chang, and D. G. Stearns, J. Appl. Phys. **71**, 187 (1992).
[33] E. S. Machlin, in *An Introduction to Thermodynamics and Kinetics Relevant to Materials Science* (Giro, Croton-on-Hudson, 1991), p. 115.
[34] K. H. Muller, Surf. Sci. **184**, L375 (1987); Phys. Rev. B **35**, 7906 (1987).
[35] S. K. Sinha, E. B. Sirota, S. Garoff, and H. B. Stanley, Phys. Rev. B **38**, 2297 (1988).
[36] D. K. G. de Boer, Phys. Rev. B **53**, 6048 (1996).
[37] D. G. Stearns, M. B. Stearns, Y. Cheng, J. H. Stith, and N. M. Ceglio, J. Appl. Phys. **67**, 2415 (1990).
[38] Y. Cheng, J. Liu, M. B. Stearns, and D. G. Stearns, Proc. SPIE **1547**, 167 (1992).
[39] J. H. Underwood, E. M. Gullikson, M. Koike, P. J. Batson, P. E. Denham, K. D. Franck, R. E. Tackaberry, and W. F. Steele, Rev. Sci. Instrum. **67**, 3372 (1996), CD-ROM only.
[40] URL: www-cxro.lbl.gov; B. L. Henke, E. M. Gullikson, and J. C. Davis, Atomic Data Nucl. Data Tables **54**, 181 (1993).
[41] R. S. Rosen, D. G. Stearns, and S. P. Vernon, Appl. Opt. **32**, 6975 (1993).
[42] W. C. Sweatt, OSA Trends in Optics and Photonics Vol. 4, *Extreme Ultraviolet Lithography*, edited by G. D. Kubiak and D. Kania (Optical Society of America, Washington, DC, 1996), p. 178.
[43] G. E. Sommargren and L. Seppala, Appl. Opt. **32**, 6938 (1993).
[44] E. Spiller, D. Stearns, and M. Krumrey, J. Appl. Phys. **74**, 107 (1993).
[45] *Physical Metallurgy*, edited by R. W. Cahn and P. Haasen, 3rd ed. (Elsevier, Amsterdam, 1983), p. 398.
[46] M. Born and E. Wolf, in *Principles of Optics*, 6th ed. (Pergamon, New York, 1987), p. 51; O. S. Heavens, *Optical Properties of Thin Films* (Dover, New York, 1966).
[47] L. Nevot and P. Croce, Rev. Phys. Appl. **15**, 761 (1980).