Modeling of surface metrology of state-of-the-art x-ray mirrors as a result of stochastic polishing process

Valeriy V. Yashchuk
Yury N. Tyurin
Anastasia Y. Tyurina
Modeling of surface metrology of state-of-the-art x-ray mirrors as a result of stochastic polishing process

Valeriy V. Yashchuk, a, * Yury N. Tyurin, b, c and Anastasia Y. Tyurina a

aLawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, United States
bMoscow State University, 1 Leninskiye Gory Street, Moscow 119991, Russia
cSecond Star Algonumerics, 19 West Street, Needham, Massachusetts 02494, United States

Abstract. The design and evaluation of the expected performance of optical systems requires sophisticated and reliable information about the surface topography of planned optical elements before they are fabricated. The problem is especially severe in the case of x-ray optics for modern diffraction-limited-electron-ring and free-electron-laser x-ray facilities, as well as x-ray astrophysics missions, such as the X-ray Surveyor under development. Modern x-ray source facilities are reliant upon the availability of optics of unprecedented quality, with surface slope accuracy < 0.1 μrad. The unprecedented high angular resolution and throughput of future x-ray space observatories require high-quality optics of 100 m ² in total area. The uniqueness of the optics and limited number of proficient vendors make the fabrication extremely time-consuming and expensive, mostly due to the limitations in accuracy and measurement rate of metrology used in fabrication. We continue investigating the possibility of improving metrology efficiency via comprehensive statistical treatment of a compact volume of metrology of surface topography, which is considered the result of a stochastic polishing process. We suggest, verify, and discuss an analytical algorithm for identification of an optimal symmetric time-invariant linear filter model with a minimum number of parameters and smallest residual error. If successful, the modeling could provide feedback to deterministic polishing processes, avoiding time-consuming, whole-scale metrology measurements over the entire optical surface with the resolution required to cover the entire desired spatial frequency range. The modeling also allows forecasting of metrology data for optics made by the same vendor and technology. The forecast data are vital for reliable specification for optical fabrication, evaluated from numerical simulation to be exactly adequate for the required system performance, avoiding both over- and underspecification.

Keywords: surface metrology; time-invariant linear filter; autoregressive moving average; power spectral density; fabrication tolerances; x-ray optics; surface slope profilometry.

Paper 160527P received Apr. 8, 2016; accepted for publication Jun. 28, 2016; published online Jul. 15, 2016.

1 Introduction

The design and evaluation of the expected performance of optical systems requires sophisticated and reliable information about the surface topography of planned optical elements before they are fabricated. The problem is especially severe in the case of x-ray optics for modern diffraction-limited-electron-ring and free-electron-laser x-ray source facilities. Modern x-ray source facilities are reliant upon the availability of x-ray optics of unprecedented quality, with surface slope accuracy better than 0.1 μrad and surface height error of < 1 nm.1–5 The uniqueness of the optics and limited number of proficient vendors make the fabrication extremely time-consuming and expensive, mostly due to the limitations in accuracy and measurement rate of the available metrology. Similar problems arise in fabrication of optics for x-ray astrophysics missions under development. In this case, the unprecedented high angular resolution and throughput of future x-ray space observatories, such as the X-Ray Surveyor mission,6 require high-quality optics (with tolerances on the level of 1 μrad) of 100 m ² in total area.

Recently, a possibility of improving metrology efficiency via comprehensive statistical treatment of a compact volume of metrology data has been suggested (see Refs. 7–9 and references therein). It has been demonstrated8,9 that one-dimensional (1-D) slope metrology with super-polished x-ray mirrors can be treated as a result of a stochastic polishing process. In this case, the measured surface slope variation is first detrended to remove the overall shape (trend) and periodic variation (oscillation) that are not stochastic. Second, the residual slope variation is treated as a result of a stochastic polishing by fitting with an autoregressive moving average (ARMA) and an extension of ARMA to time-invariant linear filter (TILF) modeling.10,11 The modeling allows a high degree of confidence in description of the surface topography data (and the polishing process) with a limited number of parameters.

With the parameters of the determined model, the surface slope profiles of the prospective (before fabrication) optics made by the same vendor and technology can be forecast. The forecast data are vital for reliable specification for optical fabrication, evaluated from numerical simulation to be necessary and sufficient for the required system performance, avoiding both over and underspecification.12,13

Considering surface slope topography the result of a stationary stochastic polishing process and using a compact volume of metrology data on the topography, modeling can be utilized to provide feedback for deterministic optical polishing. This can avoid time-consuming whole-scale metrology

*Address all correspondence to: Valeriy V. Yashchuk, E-mail: VVYashchuk@lbl.gov
measurements over the entire optical surface with the resolution required to cover the entire spatial frequency range, important for the optical system performance.

In the present work, we continue the investigations started in Refs. 8–13. First, we briefly review the mathematical fundamentals of 1-D ARMA modeling of topography of random rough surfaces (Sec. 2). In Sec. 3, we analyze a generalization of ARMA modeling with the TILF approach. We analytically show that the suggested symmetric TILF approximation has all the advantages of one-sided autoregressive (AR) and ARMA modeling, but it additionally has improved fitting accuracy. It is free of the causality problem, which can be thought of as a limitation of ARMA modeling of surface metrology data. An algorithm for identification of an optimal symmetric TILF model with a minimum number of parameters and smallest residual error is derived in Sec. 4. Finally, in Sec. 5, we verify the efficiency of the developed algorithm in application for modeling of a series of stochastic processes, which are generated with the known ARMA model, determined for surface slope data for a state-of-the-art x-ray mirror. The paper concludes (Sec. 6) by summarizing the main concepts discussed throughout the paper and stating a plan for extending the suggested approach to parameterize the results of two-dimensional (2-D) surface metrology data.

2 One-Dimensional Statistical Modeling and Forecasting of Random Rough Surfaces

2.1 Autoregressive Moving Average Modeling

Let us consider the surface slope metrology of high-quality x-ray optics. For the 1-D case, the result of the metrology is a distribution (trace) of residual (after subtraction of the best-fit figure and trends) slopes X[n] measured over discrete points xₙ = n · ΔX [n = 1, . . . , N], where N is the total number of observations and (N − 1)ΔX is the total length of the trace, uniformly, with an increment ΔX distributed along the trace.

ARMA modeling describes the discrete surface slope distribution α[n] as the result of a uniform stochastic process,

\[ X[n] = \sum_{l=1}^{p} a_l X[n-l] + \sum_{l=0}^{q} b_l \nu[n-l], \]

where \( \nu[n] \) is zero-mean variance white Gaussian noise (referred to as “white Gaussian noise”), i.e., the driving noise of the model. The parameters \( p \) and \( q \) are the orders of the AR and moving average (MA) processes, respectively. At \( q = 0 \) and \( b_0 = 1 \), the ARMA process [Eq. (1)] reduces to an AR stochastic process. In addition to the linearity, the ARMA transformation is time-invariant since its coefficients depend on the relative lags \( l \) rather than on \( n \). The goal of the modeling is to determine the ARMA orders and estimate the corresponding AR and MA coefficients \( a_l \) and \( b_l \).

Due to the availability of sophisticated statistical software capable of ARMA modeling of experimental data, ARMA fitting becomes a rather routine task for finding the ARMA model parameters and verifying the statistical reliability of the model. We use a commercially available software package, EViews 8.9. In particular, the software provides easy-to-use ARMA modeling tools oriented to econometric analysis, forecasting, and simulation.

ARMA fitting allows for replacement of the spectral estimation problem with a problem of parameter estimation. In principle, the parameters of a successful ARMA model for a rough surface should relate to the polishing process. The analytical derivation of such a relation is a separate difficult task; there are just a few works that try to solve this problem. Instead, most of the existing work provides an empirical ARMA description for the results from polishing processes. When an ARMA model is identified, the corresponding power spectral density (PSD) distribution can be analytically derived

\[ P_X(f) = \frac{\sigma^2 |B[e^{i2\pi f}]|^2}{|A[e^{i2\pi f}]|^2} \]

where the frequency \( f \in [-0.5, 0.5] \)

\[ A[e^{i2\pi f}] = 1 + a_1 e^{i2\pi f} + \cdots + a_p e^{i2\pi pf}, \]

\[ B[e^{i2\pi f}] = b_0 + b_1 e^{i2\pi f} + \cdots + b_q e^{i2\pi qf}. \]

Equation (2) can be expressed as

\[ P_X(f) = \sigma^2 \frac{(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})(1 - a_1 z^{-1} - \cdots - a_p z^{-p})}{(1 - a_1 z^{-1} - \cdots - a_p z^{-p})}, \]

where \( z = e^{i2\pi f} \) and \( \sigma^2 \) are the variance of the driving noise \( \nu[n] \) and \( \sigma \) is also called the standard error of regression.

Therefore, a low-order ARMA fit, if successful, allows parametrization of both the PSD and the autocovariance function (ACF) of a random rough surface. As a result, the PSD distributions appear as highly smoothed versions of the corresponding estimates via a direct digital Fourier transform. Description of a rough surface as the result of an ARMA stochastic process provides a model-based mechanism for extrapolating the spectra outside the measured bandwidth.

Trustworthy ARMA modeling and forecasting based on a limited number of observations assume statistical stability of the data used. The data are statistically stable if they are the result of a so-called wide sense stationary (WSS) random process (see, e.g., Ref. 14). The process \( X[n] \), where \( n = 1, \ldots , N \) and \( N \) is the number of observations, is a WSS process if its ACF

\[ r_X[l] = E[X[n]X[n-l]], \]

depends only on the lag \( l \) and does not depend on the value of \( n \). In Eq. (1), \( E \) is the expectation operator. Note that the PSD of the WSS random process \( X[n] \) can be found from the ACF [compare to Eqs. (2) and (5)]

\[ P_X(f) = \sum_{l=-\infty}^{\infty} r_X[l] e^{-i2\pi fl} = \sum_{l=-\infty}^{\infty} r_X[l] z^{-l}. \]

According to Eq. (5), \( r_X[l] \) is a nonlinear function of the ARMA coefficients, \( a_l \) for \( l = 1, \ldots , p \) and \( b_l \) for \( l = 1, \ldots , q \).
Recent publications\textsuperscript{8,9} describe a successful application of ARMA modeling to the experimental surface slope data for a 1280-m spherical reference mirror.\textsuperscript{23,24} The data were obtained with the Advanced Light Source (ALS) developmental long trace profiler (DLTP)\textsuperscript{23} and verified in cross-comparison with measurements performed with the HZB/BESSY-II nanometer optical component measuring machine,\textsuperscript{26-29} one of the world’s best slope measuring instruments.

### 2.2 Two-Sided Symmetrical Autoregressive Moving Average Modeling

With the obvious success and perspective of the application of 1-D ARMA modeling to 1-D surface slope metrology, the inherent causality of the modeling is thought of as a limiting factor that also complicates extension of the method for modeling 2-D surface metrology available, e.g., with high-precision interferometers and microscopes.

Indeed, ARMA modeling is inherently causal, assuming that the current value of the process depends only upon the past, as expressed with Eq. (1). While in the case of time series, the property of causality is natural, in the case of modeling of surface metrology data, the causality can be thought of as a limitation. A valid model should describe the reversed surface metrology data corresponding to the measurements with the optic rotated (flipped) by 180 deg with respect to the scanning direction of the profiler. The direct and reversed residual slope traces are related through a straightforward transformation of the coordinate system and change to the scanning direction of the profiler. The causality limitation is solved by a straightforward merging of the causal stochastic processes [Eqs. (1) and (8)] into another (filtered) process [Eq. (9)]

\[
X[t] = \sum_{i=-\infty}^{\infty} c_i X[t - i] \equiv C * X[t].
\]  

Similar to the ARMA transformation, the TILF C is linear and time-invariant. The filter C possesses the property of causality if

\[
c_i = 0, \quad \text{ for } i < 0.
\]  

The requirement of stability of the transformation implies that the filter is absolutely summable.

\[
\sum_{i=-\infty}^{\infty} |c_i| < \infty.
\]  

Also similar to ARMA modeling, when an optimal TILF is identified, the corresponding PSD distribution can be analytically derived [compare with Eq. (5)]

\[
P_Y(f) = \left| \int_{-\infty}^{\infty} c_i e^{2\pi if} \right|^2 P_X(f).
\]

Almost any ARMA process \(X[t]\) with the parameters \(p\) and \(q\) can be obtained from white Gaussian noise \(\nu[t]\) by application of the corresponding causal TILF\textsuperscript{29}

\[
X[t] = \sum_{i=0}^{\infty} c_i \nu[t - i].
\]

The weights \(c_i\) in Eq. (14) are determined by the relation

\[
\sum_{i=0}^{\infty} c_i z^i = \frac{b(z)}{a(z)}, \quad |z| \leq 1.
\]

where the AR and MA polynomials on the right-hand side are, respectively

\[
a(z) = 1 - a_1 z^{-1} - \cdots - a_p z^{-p} \quad \text{and}
\]

\[
b(z) = 1 + b_1 z^1 + \cdots + b_q z^q.
\]

Consequently, the two-sided ARMA process given in Eq. (9) can be expressed via a TILF in the form of Eq. (9), which is free of the causality limitation

\[
X[t] = \frac{1}{2} \left\{ \sum_{i=0}^{\infty} c_i \nu[t - i] + \sum_{i=0}^{\infty} c_i \nu[t + i] \right\}
\]

\[
= \sum_{i=-\infty}^{\infty} c_i \nu[t - i].
\]
Therefore, in the case of 1-D metrology data, if ARMA modeling is successful, there is a corresponding TILF operator that describes the metrology result as filtered white Gaussian noise. The identified TILF can be used for forecasting of a new slope distribution possessing the same statistical properties as the measured one, but with different parameters, such as the distribution length and the rms variation. A straightforward generalization of the 1-D expressions [Eqs. (10)–(17)] to the 2-D case opens a way for parametrization and forecasting of 2-D metrology data by applying 2-D TILF modeling.

Note that there is a simple relation between the coefficients of the AR terms of Eq. (9) and the weights of a TILF that transform the two-sided AR process into the noise process \( \nu[t] \). In some sense, such a TILF is the inverse operator to the one in Eq. (14). In this case, the AR part of Eq. (9) can be written as

\[
X[t] = \frac{1}{2} \sum_{l=-p}^{p} a_l X[t-l] - \frac{1}{2} a_0 X[t] + \nu^*[t],
\]

with the coefficients \( a_l, l = \pm 1, \ldots, \pm p \) determined by AR modeling the direct and reversed traces of the same slope measurement. Assigning \( a_0 = 0 \), Eq. (18) is rewritten in a form of a TILF transformation

\[
\nu[t] = \sum_{l=-p}^{p} c_l X[t-l] - X[t] \equiv (C-I) \ast X[t],
\]

where white Gaussian noise \( \nu[t] = -\nu^*[t] \), \( I \) is the identity operator, and \( C \) is a finite TILF of order \( p \) with the weights \( c_l = a_l/2 \), for \( l = \pm 1, \ldots, \pm p \) and \( c_0 = 0 \), for \( l = 0 \).

Filter \( C \) in Eq. (19), when applied to the process \( X[t] \), gives a new stationary random process \( Y[t] \) that differs from the process \( X[t] \) by the noise process \( \nu[t] \). If the difference is small (e.g., the variance of the noise is much smaller than that of the processes \( X[t] \) and \( Y[t] \)), the TILF \( C \) can be thought of as a good model of the stochastic process \( X[t] \), representing its structure with the weight coefficients given by Eq. (20).

Practically, to determine a TILF filter \( C \) that best models the observed stochastic process \( X[t] \), one has to find a set of the weight coefficients \( c_l \) that minimizes the deviation

\[
E([X[t] - Y[t]]^2) \equiv E(\nu^2[t]),
\]

of the modeled process from the observed one.

\section{Symmetry of Time-Invariant Linear Filters for Modeling of Surface Slope Metrology}

Generally, the values of the TILF weights with the same positive and negative lags are not necessarily equal, i.e.,

\[
c_l \neq c_{-l}.
\]

However, as we mathematically prove in this section, among all TILFs (including AR and ARMA models) of the same order, the symmetrical filter with

\[
c_l = c_{-l}.
\]

provides the smallest variance of the residual noise, which is equal to the difference between the measured trace and the best-fitted model. In the case of causal TILFs (such as AR and ARMA models), it can be intuitively understood as a result of averaging of the residual noises of the fits with the corresponding causal filters of the direct and reversed processes. Assuming that the residual noises are not mutually correlated, one should expect a suppression of the variance of the averaged residual noise by a factor of 2 with respect to the corresponding causal filter.

For mathematical proof of the statement given in Eq. (23), we will show that replacement of a given TILF with its symmetric form reduces the variance of the difference between the observed and modeling stochastic processes.

Let us define

\[
c_l = \tilde{c}_l + \delta c_l \quad \text{and} \quad c_{-l} = \tilde{c}_l - \delta c_l,
\]

where \( \tilde{c}_l = (c_l + c_{-l})/2 \), \( \delta c_l = (c_l - c_{-l})/2 \), and \( |l| \leq p \), and \( p \) is the order of the TILF model \( C \). In these notations

\[
Y[t] = \sum_{k=-1}^{p} \tilde{c}_k [X[t-k] + X[t+k]] + \sum_{l=1}^{p} \delta c_l [X[t-l] - X[t+l]] \equiv u[t] + v[t].
\]

Therefore, the variance of the difference between \( X[t] \) and \( Y[t] \) is

\[
E([X[t] - Y[t]]^2) = E([X[t] - u[t]]^2) + E(\nu^2[t]) + 2 \sum_{l=1}^{p} \delta c_l E\left( X[t] - \sum_{k=1}^{p} \tilde{c}_k [X[t-k] + X[t+k]] \right. \\
\left. \times [X[t-l] - X[t+l]] \right).
\]

Let us show that the last sum in Eq. (26) equals zero, as each of its terms is zero. The elements in the sum are of two types (up to multipliers) that can be reduced to the ACF of the process \( X[t] \)

\[
E[X[t][X[t-k] - X[t+k]]] = E[X[t]X[t-k]] - E[X[t]X[t+k]] = r_X[k] - r_X[-k] = 0,
\]

and

\[
E([X[t-k] + X[t+k]][X[t-k] - X[t+k]]) = r_X[0] + r_X[2k] - r_X[2k] = 0.
\]

Therefore, the variance of the difference between \( X[t] \) and \( Y[t] \) equals (compare with Eq. (26))
the variance equation [Eq. (30)] can be written in the matrix form
\[
E([X[t] - Y[t]]^2) = r_X[0] - 4\bar{c}\bar{r}_X + 2\bar{c}Q\bar{c}^T.
\] (33)

The weight coefficients of the optimal symmetric TILF correspond to the minimum value of the variance in Eq. (33). We derive an analytical expression that allows determination of the optimal weight coefficients for the case where the inverse of matrix \(Q\) exists.

Let us add the finite-difference derivative \(\delta c\) to vector \(\bar{c}\), \(\bar{c} + \delta c, \delta c \ll \bar{c}\), and insert the result in Eq. (33)
\[
\delta E([X[t] - Y[t]]^2) = -4(\bar{c} + \delta c)\bar{r}_X^2 + 2(\bar{c} + \delta c)Q(\bar{c} + \delta c)^T + 4\bar{c}\bar{r}_X^2 - 2\bar{c}Q\bar{c}^T.
\] (34)

By performing straightforward algebraic transformations and leaving in the right term only the part linear with respect to \(\delta c\), Eq. (34) can be transformed to
\[
\delta E([X[t] - Y[t]]^2) = -4\delta c\bar{r}_X^2 + 2\delta cQ\bar{c}^T + 2\bar{c}Q\delta c^T
\] (35)

\[
= 4\delta c(-\bar{r}_X^2 + Q\bar{c}^T).
\]

To get Eq. (35), we use the facts that \(\delta cQ\bar{c}^T\) and \(\bar{c}Q\delta c^T\) are just constants and that the matrix \(Q\) is symmetric; therefore
\[
\delta cQ\bar{c}^T = (\delta cQ\bar{c}^T)^T = (Q\bar{c}^T)^T\delta c^T = \bar{c}Q^T\delta c^T = \bar{c}Q\delta c^T.
\] (36)

From Eq. (35), the variance between the observed process \(X[t]\) and the approximating one \(Y[t]\) reaches its minimum at
\[
\langle -\bar{r}_X^2 + Q\bar{c}^T \rangle = 0,
\]
i.e., \(\bar{r}_X = Q\bar{c}\). (37)

If the inverse of matrix \(Q\) exists, one gets a condition for determining the weight coefficients of the optimal symmetric TILF
\[
\bar{c} = \bar{r}_XQ^{-1}.
\] (38)

To determine the minimum value achieved by the variance [Eq. (33)], we substitute Eq. (38) into Eq. (33)
\[
E([X[t] - Y[t]]^2)_{\text{MIN}} = r_X[0] - 4\bar{r}_XQ^{-1}\bar{r}_X^T
\]
\[
+ 2\bar{r}_XQ^{-1}Q(\bar{r}_XQ^{-1})^T
\]
\[
= r_X[0] - 4\bar{r}_XQ^{-1}\bar{r}_X^T + 2\bar{r}_X(\bar{r}_XQ^{-1})^T\bar{r}_X.
\] (39)

And finally, the expression for calculating the minimum value of the variance between the observed process \(X[t]\) and the approximating one \(Y[t]\) is
\[
E([X[t] - Y[t]]^2)_{\text{MIN}} = r_X[0] - 2\bar{r}_XQ^{-1}\bar{r}_X.
\] (40)

Equations (38) and (40) provide an algorithm for evaluation of the best TILF with given values of the filter order \(p\).
5 Verification of the Developed Algorithm for Identification of Optimal Symmetrical Time-Invariant Linear Filter

In this section, we will verify the developed algorithm for determining weight coefficients of an optimal symmetrical TILF (Sec. 4) applied to a series of stochastic processes generated with the ARMA model\textsuperscript{12,13} built from surface slope data, measured with the ALS DLTP\textsuperscript{25} of the SLAC Linac Coherent Light Source (LCLS) beam split and delay mirror.\textsuperscript{30} The DLTP is capable of slope metrology for plane surfaces with absolute error better than 80 nrad and rms error <50 nrad.\textsuperscript{31,32} The overall error of the used data is estimated to be <60 nrad (rms).

5.1 Autoregressive Moving Average Model for Slope Data Measured with the Linac Coherent Light Source Beam Split and Delay Mirror

Figure 1(a) (blue solid line) shows the residual (after subtraction of the best-fit third-order polynomial) slope variation over the mirror clear aperture of 138 mm. The trace consists of \( N = 691 \) points measured with an increment of \( \Delta x = 0.2 \) mm.

The best-fit slope trace, shown in Fig. 1(a) with the red dashed line, corresponds to the ARMA model specified in Table 1. The table, generated by EViews 8 software\textsuperscript{19} as the regression output, includes only the statistically significant ARMA parameters.

The slope difference trace shown in Fig. 1(b) is the driving noise of the ARMA model, i.e., \( \nu[n] \) in Eq. (1). It should be distinguished from any observation noise or measurement error. The details of the ARMA modeling of the slope metrology with the LCLS beam split and delay mirror can be found in Refs. 12 and 13.

5.2 Autoregressive Moving Average Forecasting of Surface Topography for Mirrors with 500-mm Length, Statistically Identical to the Linac Coherent Light Source Split and Delay Mirror

The ARMA model established for the LCLS beam split and delay mirror and depicted in Table 1 was used to forecast a number of new surface slope distributions that can be thought of as the data for prospective mirrors manufactured with the same polishing process at the fabrication facility of the real mirror modeled with ARMA.

Forecasting a new slope trace with the determined ARMA model is performed in two steps. First, we generate a new sequence of white-noise-like, normally distributed residuals \( \nu[n] \) with a length of 500 mm (with 0.2-mm increment) corresponding to the new desired mirror length. Second, by using Eq. (1) with the ARMA parameters in Table 1 and the extended residuals, a new slope trace is generated and normalized to get the rms slope variation of 0.1 \( \mu \)rad.

Using uncorrelated sets of residuals \( \nu[n] \), a number of statistically independent (inherently noncorrelating) but statistically identical (with the predetermined ARMA parameters) slope traces are generated with the EViews 8 software.\textsuperscript{19}

The blue solid line in Fig. 2(a) presents one of the slope distributions, slope 09, forecast based on the described procedure and the ARMA model in Table 1. The best-fit slope trace corresponding to the ARMA model specified in Table 3 is shown in Fig. 2(a) with the red dashed line. By ARMA modeling the traces with EViews 8 in a similar manner to that described in Refs. 8–13, we verify the statistical identity of the generated slope traces to the used ARMA model.

Within statistical uncertainty, AR parameters of the ARMA models identified for the generated slope traces (see

---

**Table 1** Parameters of the ARMA model [red dashed line in Fig. 1(a)] that best fit the surface slope trace for the LCLS beam split and delay mirror measured with the ALS DLTP. In Eqs. (1)–(5), \( b_2 = 1 \) and \( \alpha_2 \) is equal to the standard error of the regression of 0.053 \( \mu \)rad (rms). The data are the regression outputs generated by EViews\textsuperscript{19} software. In the table, the standard error of the corresponding coefficient determines the statistical significance of the parameter.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1):  ( a_1 )</td>
<td>0.637</td>
<td>0.036</td>
</tr>
<tr>
<td>AR(2):  ( a_2 )</td>
<td>0.352</td>
<td>0.041</td>
</tr>
<tr>
<td>AR(5):  ( a_5 )</td>
<td>-0.147</td>
<td>0.027</td>
</tr>
<tr>
<td>MA(2):  ( b_2 )</td>
<td>-0.0659</td>
<td>0.0035</td>
</tr>
</tbody>
</table>

---

**Fig. 1** (a) Measured slope trace (blue solid line) after subtracting the best-fit third-order polynomial shape to remove the trend, i.e., characteristic for short x-ray mirrors, and the best-fit slope trace (red dashed line), corresponding to the ARMA model specified in Table 1. The rms variation of the measured slope trace is 0.099 \( \mu \)rad. (b) Difference between the measured and fitted traces, i.e., the driving noise of the model in Eq. (1). The rms variation of the slope difference is 0.053 \( \mu \)rad.
Table 2) are equal to those of the ARMA model of the measured slope trace.

However, the variation of the MA(2) parameter values is very large. A comparison with the results of an ARMA fit for an individual trace (e.g., slope 09 in Table 3) suggests that, in general, these ARMA fits are not sensitive to the MA term. In the case of trace slope 09, the value of \( b_2 \) is larger than its error only by a factor of 1.5. Nevertheless, the MA(2) term was kept in the parent ARMA model because it is needed to randomize the residuals of the ARMA fit for the measured slope trace in Fig. 1.

The generated slope traces, such as the one shown in Fig. 2 and with the best-fit ARMA parameters in Table 2, are used to test the developed algorithm for determining the TILF weight parameters.

### 5.3 AR-Time-Invariant Linear Filters with Analytically Derived Weight Coefficients for One-Dimensional Data Generated with the Known Autoregressive Moving Average Model

Here, we investigate the performance of modeling the stochastic polishing process using symmetric TILFs and the developed analytical procedure for determining the weight coefficients of the optimal filter. For this, we apply the calculation algorithm based on Eqs. (38) and (40) to fit nine slope traces generated with the known parent ARMA model as described in Secs. 5.1 and 5.2. Such generated traces are a priori the results of a uniform, stationary stochastic process;

![Generated slope traces](attachment:image.png)

**Fig. 2** (a) Generated slope trace slope 09 (blue solid line) and best-fit slope trace (red dashed line), corresponding to the ARMA model specified in Table 3. The rms variation of the generated slope trace is 0.101 \( \mu \)rad. (b) Difference between the generated and fitted traces. The rms variation of the slope difference is 0.0523 \( \mu \)rad.

<table>
<thead>
<tr>
<th>Trace\ parameter</th>
<th>AR(1): ( a_1 )</th>
<th>AR(2): ( a_2 )</th>
<th>AR(5): ( a_5 )</th>
<th>Mean value ( b_2 )</th>
<th>Standard error of regression (( \mu )rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope 01</td>
<td>0.6705</td>
<td>0.3008</td>
<td>-0.1512</td>
<td>-0.0041</td>
<td>0.0519</td>
</tr>
<tr>
<td>Slope 02</td>
<td>0.6362</td>
<td>0.3251</td>
<td>-0.1325</td>
<td>-0.0175</td>
<td>0.0523</td>
</tr>
<tr>
<td>Slope 03</td>
<td>0.6032</td>
<td>0.3737</td>
<td>-0.1545</td>
<td>-0.0376</td>
<td>0.0541</td>
</tr>
<tr>
<td>Slope 04</td>
<td>0.6517</td>
<td>0.3208</td>
<td>-0.1447</td>
<td>-0.0587</td>
<td>0.0524</td>
</tr>
<tr>
<td>Slope 05</td>
<td>0.6098</td>
<td>0.3805</td>
<td>-0.1556</td>
<td>-0.0950</td>
<td>0.0517</td>
</tr>
<tr>
<td>Slope 06</td>
<td>0.6936</td>
<td>0.1874</td>
<td>-0.0737</td>
<td>0.1624</td>
<td>0.0538</td>
</tr>
<tr>
<td>Slope 07</td>
<td>0.6166</td>
<td>0.4010</td>
<td>-0.1712</td>
<td>-0.1027</td>
<td>0.0531</td>
</tr>
<tr>
<td>Slope 08</td>
<td>0.6563</td>
<td>0.3244</td>
<td>-0.1349</td>
<td>-0.0560</td>
<td>0.0531</td>
</tr>
<tr>
<td>Slope 09</td>
<td>0.6378</td>
<td>0.3462</td>
<td>-0.1306</td>
<td>-0.0419</td>
<td>0.0523</td>
</tr>
<tr>
<td>Mean value</td>
<td>0.6418</td>
<td>0.3289</td>
<td>-0.1388</td>
<td>-0.0279</td>
<td>0.0527</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0296</td>
<td>0.0623</td>
<td>0.0277</td>
<td>0.0783</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

**Table 3** Parameters of the ARMA model [the red dashed line in Fig. 2(a)] that best fit the generated surface slope trace slope 09. In Eqs. (1)–(5), \( b_2 = 1 \) and \( \sigma \) is equal to the standard error of the regression of 0.0523 \( \mu \)rad (rms). Note that the ARMA fit is not sensitive to the MA term; the value of \( b_2 \) is larger than its error only by a factor of 1.5. The data are the regression outputs generated by EViews19 software.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1): ( a_1 )</td>
<td>0.6378</td>
<td>0.0197</td>
</tr>
<tr>
<td>AR(2): ( a_2 )</td>
<td>0.3462</td>
<td>0.0284</td>
</tr>
<tr>
<td>AR(5): ( a_5 )</td>
<td>-0.1306</td>
<td>0.0172</td>
</tr>
<tr>
<td>MA(2): ( b_2 )</td>
<td>-0.0419</td>
<td>0.0296</td>
</tr>
</tbody>
</table>
therefore, they are ideal objects for testing the developed TILF-based modeling approach.

Figure 3 illustrates the TILF approximation of the same generated trace (slope 09) as the one shown in Fig. 2 with its best ARMA fit. The blue solid line in Fig. 3 represents the generated slope trace. The optimal approximation with the developed symmetric TILF is shown in Fig. 3(a) with the red dashed line.

For modeling, we use a symmetric TILF of order $p = 5$, corresponding to the AR order parameter $p$ of the parent ARMA model. The question about the optimal number of coefficients is out of the scope of this present work and will be investigated elsewhere.

Table 4 presents the weight coefficients of the optimal symmetric TILF determined by application of the developed fitting procedure for nine slope traces generated with the parent ARMA model. The mean values of the coefficients estimated by averaging of nine coefficients with the same lag, as well as the standard deviations of the coefficients, are shown in the last two rows of Table 4.

Note that if we double the mean values of the TILF weight coefficients in Table 4, they will be close to the values of the corresponding (with the same lag) AR parameters of the parent ARMA model, given in Table 1. The small difference is probably due to the extra two fitting AR-like parameters in the TILF.

One of the major advantages of the developed TILF approximation is that the residual slope, calculated as a difference between the generated slope trace and the corresponding fit, has, as predicted (see Sec. 3.2), a smaller variation than the ARMA approximation. For the slope trace slope 09, the improvement is about a factor of 1.38, close to $\sqrt{2}$ [compare Figs. 2(b) and 3(b)].

The algorithm of approximation with symmetric TILF developed in this paper is based on analytical transformation of the ACF of the stochastic process under treatment [refer to Sec. 4 and Eqs. (38) and (40)]. Therefore, in our case, it is natural to use, as a measure of fidelity of the TILF model, the difference between the ACFs of the generated trace and its TILF approximation. As a typical result, we illustrate the fidelity of the developed TILF modeling with the example of the generated trace slope 09.

Figure 4 shows the ACFs of trace slope 09 and its TILF approximation [Fig. 3(a)], as well as the difference of the
Fig. 4 (a) ACF of the generated trace slope 09 and (b) its TILF approximation. The inset shows the difference of the ACFs in plots (a) and (b). Except for a very tiny central region, the difference has a clear random character, indicating high-accuracy of the determined TILF.

Fig. 5 (a) Central regions of the ACFs of the generated trace slope 09 (blue solid line) and its TILF approximation (red dashed line). (b) Difference between the ACFs of the generated and fitted traces in plot (a). Note that the vertical scale in (b) is increased by a factor of ∼6.

Fig. 6 (a) Central regions of the ACF of the difference between the generated trace and its TILF approximation shown in Fig. 3(b). (b) ACF of white-noise trace with the same value of the rms variation. The inset shows the difference of the ACFs in plots (a) and (b). Note that the vertical scale in inset is increased by a factor of ∼4.
ACFs. Almost over the entire range of lag values, except for a very tiny central region, the difference has a clear random character. The regions of the ACFs in the central vicinity of lag $l = 0$ are shown in Fig. 5 with enlarged scale. Here, also, there is a very close resemblance of the two ACFs, almost over the entire range of represented lags. This means that the two stochastic processes, generated and approximated, are spectrally close; therefore, the determined TILF is highly accurate.

The ACF of the residual trace [Fig. 3(b)], which is the difference between the generated trace and its TILF approximation, is shown in Fig. 6(a). The delta-function-like ACF suggests a white-noise-like character of the residual trace. For comparison, the ACF of a computer-generated white-noise trace with the same value of the rms variation is shown in Fig. 6(b). The inset in Fig. 6 represents the difference of the ACFs in Figs. 6(a) and 6(b), plotted with significantly larger scale.

The two-peak character of the ACF difference in Fig. 6 can be a signature of residual MA contributions to the stochastic process that are not approximated with the developed symmetric TILF model. However, ARMA modeling of the residual trace in Fig. 3(b) with EViews 8 software does not provide any reasonable model that would noticeably decrease the variance of the residual trace. This suggests an almost perfect white-noise-like distribution of the residuals.

6 Conclusion
In this work, we have continued the investigation started in Refs. 8–11, which will potentially allow us to analytically characterize/parameterize the polishing capabilities of different vendors for x-ray optics. Based on the parametrization, the expected surface profiles of the prospective x-ray optics will be reliably simulated (forecast) prior to purchasing. The simulated surface slope and height distributions of prospective optics (before they are fabricated) can be used for estimations of the expected performance of x-ray optical systems (beamlines and x-ray telescopes).12,13

We have analyzed a generalization of ARMA modeling with the TILF approach. We have analytically shown that the suggested symmetric TILF approximation has all the advantages of one-sided AR and ARMA modeling, along with improved fitting accuracy. It is also free of the causality problem, which can be thought of as a limitation of ARMA modeling of surface metrology data.

An algorithm for the identification of an optimal symmetric TILF with a minimum number of parameters and smallest residual error has been derived. We have verified the efficiency of the developed algorithm applied to modeling of a series of stochastic processes, which were generated with the known ARMA model determined from surface slope data of a state-of-the-art x-ray mirror.

The major application of the performed investigation of stochastic modeling of 1-D optical surface topography is in the field of x-ray reflecting optics, where the requirements for the surface quality in the direction perpendicular to the direction of the light incident at very small angle are significantly (typically by a few orders of magnitude) relaxed (see, e.g., Ref. 1 and references therein).

For more general 2-D applications, the considered TILF-based modeling of surface metrology data provides the possibility of a direct, straightforward generalization of TILF modeling to 2-D random fields. The mathematical foundations of the generalization are well established.29 However, its practical realization requires the development of calculation algorithms and dedicated software for determination of the optimal TILF best-fit of measured 2-D surface slope and height distributions. The optimization can be done in a standard way, consisting of searching for the optimal filter’s weights by using, e.g., a method similar to one developed in this work. For reliable TILF forecasting of new surface topography based on measured and fitted ones, the residual noise of the fit has to have a zero-mean variance white Gaussian distribution. This is similar to the ARMA modeling; therefore, the corresponding methods and criteria could be applied to the statistical analysis of TILF modeling in dedicated software under development.

The forthcoming investigations have to solve the question about the uniqueness of the ARMA and TILF parametrizations for a certain polishing process. This can be performed, e.g., by cross-comparing the ARMA and TILF models for different optics of identical fabrication. This work is also in progress.

We did not discuss here the questions related to the application of the considered stochastic modeling to performance evaluation of particular optic and/or optical systems, or the details of forecasting surface slope topographies suitable for the simulations that also account for the detrended trend and cycles. These topics will be discussed elsewhere. Note that some related questions have been discussed in Refs. 12 and 13.

Acknowledgments
The authors are very grateful to Daniel J. Merthe, Nikolay A. Artemiev, and Daniele Cocco for their help with high-accuracy surface slope measurements of the LCLS beam split and delay mirror and to the Gary Centers and Wayne McKinney for very useful discussions. This work was supported in part by the NASA Small Business Innovation Research grant to Second Star Algonumerics, Project No. 15-1 S2.04-9193. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Material Science Division, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 at the Lawrence Berkeley National Laboratory. This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.
Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
Yashchuk, Tyurin, and Tyurina: Modeling of surface metrology of state-of-the-art x-ray mirrors...

References


Valeriy V. Yashchuk is currently leading the X-Ray Optics Laboratory at the Advanced Light Source, Lawrence Berkeley National Laboratory. He has authored and co-authored more than 150 scientific publications. In 1986, he was awarded the Leningrad Komsomol Prize in physics. In 2007 and 2015, he received R&D Magazine’s R&D 100 Awards for the development of laser-detected magnetic resonance imaging and binary pseudorandom calibration tool, respectively. He is an OSA fellow and a member of SPIE and APS.

Yury N. Tyurin is currently a part-time professor with the Department of Probability Theory, Faculty of Mathematics and Mechanics at Moscow State Lomonosov University, Russia. He is an author of more than 100 publications and 10 books in probability and mathematical statistics. He also enjoys developing practical applications of mathematics and statistics in industry, economics, biology, and so on, some of the developments being transferred to State Standards for Industry.

Anastasia Y. Tyurina graduated from Moscow State University as a mathematician. She is an image processing scientist and software developer. She founded a Second Star Algonumerics company in 2008. She is currently working on commercialization of her patented method of superresolution for detection of crowded point sources beyond the diffraction limit. She is consulting for governmental organizations and commercial companies, and in the course of the work, she developed a number of patented algorithms for her clients.