Tip–tilt disturbance model identification for Kalman-based control scheme: application to XAO and ELT systems

Serge Meimon,[1,*] Cyril Petit,[1] Thierry Fusco,[1] and Caroline Kulcsar[2]

[1]*Office National d'Études et de Recherches Aérospatiales, Département d'Optique Théorique et Appliquée, BP 72, F-92322 Châtillon cedex, France*
[2]*Université Paris 13, Institut Galilée, L2TI, 93430 Villetaneuse, France*
**Corresponding author: meimon@onera.fr*

Received April 6, 2010; revised August 13, 2010; accepted August 16, 2010;
posted August 17, 2010 (Doc. ID 126600); published September 17, 2010

Adaptive optics (AO) systems have to correct tip–tilt (TT) disturbances down to a fraction of the diffraction-limited spot. This becomes a key issue for very or extremely large telescopes affected by mechanical vibration peaks or wind shake effects. Linear quadratic Gaussian (LQG) control achieves optimal TT correction when provided with the temporal model of the disturbance. We propose a nonsupervised identification procedure that does not require any auxiliary system or loop opening and validate it on synthetic profile as well as on experimental data. © 2010 Optical Society of America

OCIS codes: 010.1080, 120.7280, 100.3190, 010.1330.

1. INTRODUCTION

Adaptive optics (AO) allows one to perform real-time correction of wave-front distortion effects on the image. In particular, tip–tilt (TT) disturbances have to be corrected down to a fraction of the diffraction-limited spot. For decametric-class telescopes, or Extremely Large Telescopes (ELTs), this corresponds to a few milliarcseconds (mas) RMS. The ATLAS (advanced tomography with lasers for AO systems) module dedicated to the European ELT (E-ELT) considers a 2 mas RMS TT error budget [1]. Due to limitations inherent in the laser guide star concept, this performance has to be reached using natural guide stars (NGSs), preferably faint ones to preserve sky coverage.

Concerning 8-m-class telescopes, future coronographic imaging instruments aim at performing direct detection of extrasolar planets, which leads to very stringent specifications for the residual TT. For instance, the SPHERE project (Spectro-Polarimetric High-contrast Exoplanet REsearch) requires a field stabilization better than 2 mas RMS [2]. Conversely, the Gemini Planet Imager (GPI) requires less than 5 mas RMS (tip+tilt), with a goal of 3 mas RMS [3].

TT disturbances originate from various sources: atmospheric turbulence, of which TT is the main component [4]; telescope wind shake, estimated to approximately 300 mas RMS in the E-ELT case [5]; and mechanical vibrations of the telescope and instruments. For instance, a Strehl ratio (SR) loss up to 25% (in a worst-case scenario) on NAOS [6] was attributed to vibrations, while on Altair [7] the vibration strength was estimated to 10–20 mas RMS. These considerations beg for an accurate TT control law.

It is therefore natural to consider linear quadratic Gaussian (LQG) control laws [8], which allow AO to perform an optimal correction in terms of residual phase variance [9,10]. Degraded solutions are often preferred, due to the complexity of LQG when dealing with a high number of degrees of freedom. However, TT (and possibly focus) can be dealt with separately from high-order modes, as shown in [11]. In this case, complexity issues are less stringent for low orders, and an LQG strategy with a reasonable state-space dimension should definitely be considered. LQG control law performs state estimation, including disturbance estimation/prediction, via a Kalman filter. The key issue is to build the filter with correct parameters, i.e., with a correct temporal model of the disturbance.

Wind shake and vibrations are particularly difficult to characterize, because they vary from one system to another as well as evolve in time along observation. In the case of large telescopes, it is impractical to use artificial excitation. The disturbance input remains unknown, and the system identification algorithms have to deal with output-only measurements. There exists a whole literature on output-only identification methods (see for instance [12] for a review) on a wide variety of application fields [13], making use of raw time data, covariances, or spectra. As regards methods making use of spectrum estimates data, maximum likelihood techniques estimate the parameters of a model by minimizing an *ad hoc* error norm. The identification method presented here belongs to this category.

In this paper, we propose an on-line identification method that makes use of wave-front sensor (WFS) measurement temporal sequences. It can be run periodically as new batches of data are collected, in order to retune the filter parameters along observation and thus adapt to changes in disturbance statistics. As this identification procedure can be run in idle time—the only interface with

the real-time system is an update of the filter parameters—it can be implemented in existing AO systems with only a minor modification of the overall control system architecture. It is an unsupervised method, with only two user-defined parameters: the minimum and maximum frequencies between which the vibration peaks are to be found. Default settings are proposed, which performed successfully on a variety of practical cases. The originality of the method is to make use of the disturbance spectrum to identify sequentially the vibration peaks in decreasing order of energy.

The outputs of the method are (i) a model of the "smooth" component of the disturbance, e.g., wind shake and turbulence; (ii) a model of the "sharp" components, i.e., the vibration peaks, to be used in a state-space framework with an LQG control scheme (Section 2). We try to provide insight for readers not familiar with LQG control. In particular, we show that even in the absence of vibration peaks (e.g., with wind shake alone), knowledge of the smooth component model allows one to reach TT mitigation performance that is much higher than with a classical integrator.

We present a robust model identification method (Section 3), which does not rely on any auxiliary system or require opening the loop to acquire information on the disturbance characteristics. To our knowledge, no similar spectrum-based approach has been proposed in the field of AO. We characterize the performance of our vibration identification and filtering scheme on synthetic profiles corresponding to the SPHERE case (Section 4) as well as on NAOS experimental data (Section 5).

2. STATE-SPACE DESCRIPTION FOR A MONO MODAL DISTURBANCE

We consider that TT modes are decoupled and can be considered separately from high orders and from each other. In the following, the variables correspond to a single mode; i.e., WFS measurements, mirror commands, and disturbance phases are scalar.

A. Temporal Model of the AO System

We first establish the temporal model of the AO system, considering the following assumptions [14]:

- The deformable mirror (DM) provides a linear and instantaneous response (or much faster than the loop rate T), constant over a frame period T (zero-order hold). The correction phase during $[(n-2)T,(n-1)T]$ is then Nu_{n-2}.
- CCD read-out, slope computation, and DM control use one frame period, so that the WFS measurement y_n results from the integration of the residual phase signal over $[(n-2)T,(n-1)T]$:

$$y_n = D(\varphi_{n-1} - Nu_{n-2}) + w_n, \tag{1}$$

where D is a matrix characterizing the WFS, φ_{n-1} is the average disturbance signal over time period $[(n-2)T,(n-1)T]$, and w_n is the measurement noise. w is assumed to be zero-mean white Gaussian noise with a standard deviation σ_w.

We consider mono variable signals, so here D and N are scalar.

B. Temporal Model of the Disturbance

We model the disturbance on one mode at frame n as a composite signal, comprising a turbulent signal φ_n^{tur} and several vibration signals $\varphi_n^{vib,i}$:

$$\varphi_n = \varphi_n^{tur} + \varphi_n^{vib,1} + \varphi_n^{vib,2} + \dots . \tag{2}$$

Each vibration component $\varphi^{vib,i}$ is a discretized version of a continuous-time mechanical oscillatory signal, characterized by its damping coefficient k (related to the vibration bandwidth) and its natural frequency f^{vib}. The discretized version of this continuous-time signal leads to a second-order auto-regressive (AR2) model,

$$\varphi_{n+1}^{vib,i} = a_1^{vib,i}\varphi_n^{vib,i} + a_2^{vib,i}\varphi_{n-1}^{vib,i} + v_n^{vib,i}, \tag{3}$$

with $v^{vib,i}$ a white noise. A link can be found between the continuous-time parameters (k,f^{vib}) and the AR2 parameters (a_1,a_2) (see Appendix B). As shown in Appendix B, second-order models can also be used to describe nonresonating signals such as turbulence by setting the damping coefficient k to values larger than 1. We therefore model the turbulent component also by an AR2:

$$\varphi_{n+1}^{tur} = a_1^{tur}\varphi_n^{tur} + a_2^{tur}\varphi_{n-1}^{tur} + v_n^{tur}. \tag{4}$$

C. Closed-Loop Transfer Function

In this paper, all signals are scalar sequences and all transfer functions are single-input–single-output (SISO). Let G be the controller transfer function. The closed-loop block diagram is shown in Fig. 1. We denote by SF the sensitivity function (transfer between the disturbance φ and the residual phase φ^{res}), and NTF the noise transfer function (transfer between the noise w and the residual phase φ^{res}), given by

$$SF = \frac{1}{1 + DNz^{-2}G}, \tag{5}$$

$$NTF = \frac{-Nz^{-2}G}{1 + DNz^{-2}G}. \tag{6}$$

The aim of the control law is to minimize the energy of the residual phase, which contains a signal and a noise component:

$$\mathrm{Var}\{\varphi^{res}\} = \int_f |SF(f)|^2 \cdot S^{\varphi}(f) + |NTF(f)|^2 \cdot S^{w}(f) \cdot \mathrm{d}f, \tag{7}$$

where S^{φ} and S^{w} denote, respectively, the power spectral densities (PSDs) of the disturbance and the measurement noise. The transfer function of an integrator with gain g is $G(z)=gz/(z-1)$, with an optimized g setting (w.r.t φ^{res} variance minimization) for an optimized gain integrator (OMGI) [15].

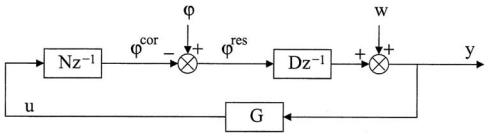

Fig. 1. Closed-loop block diagram.

However, the optimal control law (w.r.t φ^{res} variance minimization) is not the OMGI: the optimal controller has been shown to be a LQGaussian controller consisting in a Kalman filter and a reconstructed state feedback [10]. The precise shape of this controller for the present AO system and disturbance model description is given here below.

D. State-Space Description and LQG Control

Here we briefly provide the mathematical material needed to form the transfer function in the LQG case. The interested reader will find in [16] a concise justification of the state-space description presented hereafter. The whole process described by Eqs. (1), (3), and (4) can be recast as a linear time-invariant state-space model:

$$\boldsymbol{x}_{n+1} = \boldsymbol{\mathcal{A}}\boldsymbol{x}_n + \boldsymbol{v}_n, \tag{8}$$

$$y_n = \boldsymbol{\mathcal{C}}\boldsymbol{x}_n - Du_{n-2} + w_n, \tag{9}$$

with the following state matrix expressions (for sake of clarity, we considered here two vibration components):

$$\underbrace{\begin{bmatrix} \varphi^{tur}_{n+1} \\ \varphi^{tur}_{n} \\ \hline \varphi^{vib,1}_{n+1} \\ \varphi^{vib,1}_{n} \\ \hline \varphi^{vib,2}_{n+1} \\ \varphi^{vib,2}_{n} \end{bmatrix}}_{\boldsymbol{x}_{n+1}} = \underbrace{\left(\begin{array}{cc|cc|cc} a^{tur}_1 & a^{tur}_2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & a^{vib,1}_1 & a^{vib,1}_2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & a^{vib,2}_1 & a^{vib,2}_2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right)}_{\boldsymbol{\mathcal{A}}} \underbrace{\begin{bmatrix} \varphi^{tur}_{n} \\ \varphi^{tur}_{n-1} \\ \hline \varphi^{vib,1}_{n} \\ \varphi^{vib,1}_{n-1} \\ \hline \varphi^{vib,2}_{n} \\ \varphi^{vib,2}_{n-1} \end{bmatrix}}_{\boldsymbol{x}_n} + \underbrace{\begin{bmatrix} v^{tur}_{n} \\ 0 \\ \hline v^{vib,1}_{n} \\ 0 \\ \hline v^{vib,2}_{n} \\ 0 \end{bmatrix}}_{\boldsymbol{v}_n},$$

$$y_n = \underbrace{(0 \quad D \quad |0 \quad D| \quad 0 \quad D)}_{\boldsymbol{\mathcal{C}}}\boldsymbol{x}_n - DNu_{n-2} + w_n. \tag{10}$$

The LQG controller (optimal in terms of residual phase variance) is a Kalman filter observer followed by a reconstructed state feedback, corresponding to the following equations [14]:

$$\hat{\boldsymbol{x}}_{n/n} = \hat{\boldsymbol{x}}_{n/n-1} + \boldsymbol{L}(y_n - \boldsymbol{\mathcal{C}}\hat{\boldsymbol{x}}_{n/n-1} + DNu_{n-2}), \tag{11}$$

$$\hat{\boldsymbol{x}}_{n+1/n} = \boldsymbol{\mathcal{A}}\hat{\boldsymbol{x}}_{n/n}, \tag{12}$$

$$u_n = -\boldsymbol{\mathcal{K}}\hat{\boldsymbol{x}}_{n+1/n}, \tag{13}$$

with $\boldsymbol{\mathcal{K}} = -N^{-1}(1\ 0|1\ 0|1\ 0)$, and $\mathbf{L} = \Sigma_\infty C^T(\boldsymbol{\mathcal{C}}\Sigma_\infty\boldsymbol{\mathcal{C}}^T + \sigma_w^2)^{-1}$, where Σ_∞ is the asymptotic solution of a Riccati equation [14]. Using the z-transform of these equations, we compute the LQG controller transfer function:

$$G(z) = -\boldsymbol{\mathcal{K}}[\boldsymbol{I}\mathrm{d} - (\boldsymbol{I}\mathrm{d} - \boldsymbol{L}\boldsymbol{\mathcal{C}})\boldsymbol{\mathcal{A}}z^{-1} + \boldsymbol{L}DN\boldsymbol{\mathcal{K}}z^{-2}]^{-1}\boldsymbol{L}. \tag{14}$$

These equations show that the LQG control behavior depends on the disturbance model (via the model matrix $\boldsymbol{\mathcal{A}}$ and the noise variances). In the same way the information on the noise/signal balance helps setting the integrator gain g, a model of the noise and signal is needed for the Kalman filter.

E. Bode's Theorem

The LQG control scheme makes use of the disturbance temporal model. A qualitative interpretation is that its sensitivity function is "tailored" to reject as much power as possible. Although the exact shape of the LQG sensitivity function SF is computed using Eqs. (5) and (14), it can be predicted with simple arguments in a high-signal-to-noise ratio (SNR) regime. This approximation has no practical interest but provides insight into the comparative filtering capabilities between integrator and LQG.

Bode's theorem, expressed in the discrete-time system case by Mohtadi [17], states that for any stabilizing controller,

$$\int_{f \leqslant f_s/2} \ln|SF(f)|\mathrm{d}f = \beta, \tag{15}$$

where β is a constant depending only on the system we want to control (independent of the controller). It means that the balance between the frequencies for which the disturbance is mitigated ($\ln|SF(f)| < 0$) and the frequencies for which the disturbance is amplified ($\ln|SF(f)| > 0$) is fixed. The optimal filter corresponds to an optimal choice of where to amplify and where to mitigate, according to the disturbance and noise PSD shapes. The LQG control law is optimal in the sense that it minimizes the residual phase variance, which decreases in a high-SNR regime to [see Eq. (7)]:

$$\mathrm{Var}\{\varphi^{res}\} \sim \int_{f \leqslant f_s/2} |SF|^2 \cdot S \cdot \mathrm{d}f.$$

The minimization of this expression under Eq. (15) constraint can be solved with Lagrange multipliers, i.e., by minimizing

$$\mathcal{L}(|SF|,\lambda) = \int_{f\leqslant f_s/2} |SF|^2 \cdot S \cdot \mathrm{d}f + \lambda\left[\beta - \int_{f\leqslant f_s/2} \ln|SF(f)| \cdot \mathrm{d}f\right]. \quad (16)$$

We approximate the integrals by a sum on discrete values f_k of f. The condition $\partial\mathcal{L}/[\partial|SF(f_k)|]=0$ yields

$$|SF|^2 \propto 1/S. \quad (17)$$

In a high-SNR regime, the LQG squared sensitivity function in close loop is proportional to the inverse of the disturbance PSD. This classical result was witnessed by Dessenne [18] although explained otherwise. It is illustrated in Fig. 2.

We can see that the LQG sensitivity function (SF) is better suited to the disturbance temporal behavior, resulting in an almost white residual phase PSD, because

- it has notches at the location of the vibration peaks;
- it is bent in the low-frequency regime, resulting in a better rejection of the disturbance plateau.

The latter argument led to consideration of an LQG control law in the framework of ATLAS [1], the laser tomography module for the E-ELT. The goal was to mitigate the wind-shake-induced tip–tilt variance from 84.10^3 mas^2 in open loop [5] down to less than 2 mas^2. We compared the performance obtained with the LQG control scheme and for an integrator with a 0.53 gain (leading to 45° phase margin and 3 dB gain margin). For both control laws, we computed the frame rate yielding the best results and the associated residual TT variance. The results (see Table 1) show a significant gain of LQG over integrator.

Fitting a single AR2 model on a time sequence does not produce any particular difficulty. Problems arise when the signal PSD contains spiky vibrational components in addition to the "smooth" turbulence/wind shake contributor. Moreover, the performance is much more sensitive to errors in the vibration model, since a small frequency error ("small" depending on the damping in the vibration model) will result in notches in the SF being at the wrong frequency and having no effect. In what follows, we propose a method to identify both smooth and spiky components of the disturbance temporal model, their respective energies, and the measurement noise variance.

Table 1. Comparison of Wind-Shake-Mitigation Performance for LQG and Integrator Control Laws for ATLAS [19] on a Magnitude m_H=19 star[a]

Integrator		LQG	
f_s	var(φ^{res})	f_s	var(φ^{res})
500 Hz	7.0 mas^2	400 Hz	1.8 mas^2

[a]Frame rates are optimized for each control law.

3. ON-LINE IDENTIFICATION METHOD

A. Motivations for an Idle-Time Spectral Fitting Method

1. Data

A straightforward solution would be to require from the telescope an auxiliary system dedicated to measuring the parameters of the disturbance model. However, we are in favor of utilizing directly the AO loop data and use the auxiliary system, if existing, as a complementary source of data. It makes the system more autonomous and avoids differential disturbances between the WFS and the disturbance model monitor as well as cross-calibration errors. Identification should therefore be performed on the WFs measurements of the disturbance. Still, as the AO system works in closed loop, the quantity cannot be accessed directly. Consequently, we reconstruct from closed-loop measurements what would have been open-loop measurements, or pseudo-open-loop (POL), knowing the DM correction applied during the acquisition. Doing so, we implicitly assume that the measurement noise is the

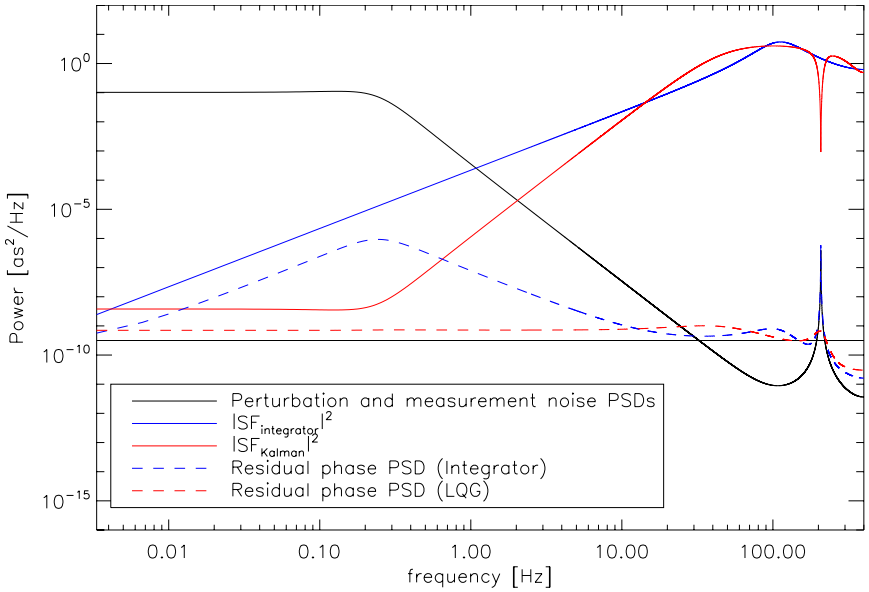

Fig. 2. (Color online) Sensitivity functions for an integrator law and for an LQG control law, and residual phase PSDs for both control laws.

same in open and closed loop—an acceptable hypothesis. In the worst case, it would lead to an underestimation of the measurement noise.

2. Idle Time Method

Adaptive methods, i.e., which update the estimated model in real time, have been successfully used in laser beam jitter control (see for instance [20] and references therein), or in the field of interferometry [21]. The present identification method is designed for extreme AO (XAO) systems and ELTs. These systems deal with high degrees of freedom, and the computation burden is of primary concern. Although this constraint is less stringent on TT (provided that these modes are treated separately from higher orders), for which an LQG strategy with a reasonable state-space dimension is considered, the complexity of the filter-parameter (i.e., disturbance model) update algorithm must not slow the loop. Considering that the disturbance model evolves much more slowly than the AO loop frame rate—with a few-thousand-frames timescale, typically—we opted for an batch identification process. It makes use of a time sequence of POL data and is run periodically (at a typical 1 Hz rate) as new batches of data are collected in order to retune the filter parameters along observation and thus adapt to disturbance statistics changes. Hinnen *et al.* made the same choices in the framework of multimodal turbulence model identification [22].

3. Spectral Fit Versus Subspace Identification

Many methods can be considered to extract the disturbance model from a sequence of POL data. Based on state-space models, it is natural to consider, for example, data-driven sub-space identification [22] (SSI), or extended Kalman filter [8] (EKF). But one could also consider maximum likelihood and prediction error methods, which rely on the optimization of a suitably chosen cost function with respect to model parameters.

In this paper, we are interested in estimating the physical parameters. In that sense, data-driven SSI methods have not been selected, as we clearly impose the particular modal structure [Eqs. (8)–(10)] for the state model. Without this constraint and for turbulence-only state-model identification, Hinnen *et al.* have successfully applied this identification scheme [22]. Also, we face here the fact that all AR2 parameters (one AR2 for each disturbance component) together with input-noise variances and disturbance need to be estimated from only a single output data sequence corresponding to the sum of all processes. The problem then becomes structurally unidentifiable [23]. For an EKF based on model (8)–(10), approximate prior variances could be used, but the EKF in this particular case is consequently very sensitive to initialization values, leading to poor performance or instability of the closed loop.

The way we circumvent this issue is to exploit the spectral separation of the components. We first clip the peaks of the disturbance PSD to identify the smooth turbulence wind shake component. Then we sequentially identify the vibration components, one at a time, in decreasing order of energy. Both operations are done by minimizing an *ad hoc* cost function. As previously mentioned, it is an unsupervised method, with only two user-defined parameters: the minimum and maximum frequencies between which the vibration peaks are to be found (default settings are proposed).

B. Direct Model

We adopt the disturbance model described in Subsection 2.B:

$$\varphi = \varphi^{tur} + \varphi^{vib,1} + \varphi^{vib,2} + \dots,$$

$$\varphi^{tur}_{n+1} = a^{tur}_1 \varphi^{tur}_n + a^{tur}_2 \varphi^{tur}_{n-1} + v^{tur}_n,$$

$$\varphi^{vib,i}_{n+1} = a^{vib,i}_1 \varphi^{vib,i}_n + a^{vib,i}_2 \varphi^{vib,i}_{n-1} + v^{vib,i}_n.$$

POL data γ correspond to WFS measurements when the DM is off, i.e., with $u_n = 0$ [cf. Eq. (1)]. Without loss of generality, we consider that the WFS gain D is equal to 1:

$$\gamma_n = \varphi_{n-1} + w_n = \varphi^{tur}_{n-1} + \varphi^{vib,1}_{n-1} + \varphi^{vib,2}_{n-1} + \dots + w_n. \quad (18)$$

We consider that the components $w, \varphi^{tur}, \varphi^{vib,i}$ of γ are statistically independent, as they may correspond to independent physical processes. Therefore, the PSD of POL data γ is the sum of the PSDs of each component:

$$S^\gamma = S^{tur} + S^{vib,1} + S^{vib,2} + \dots + S^w. \quad (19)$$

We do not have access directly to the PSD of γ but only to a periodogram P, which is asymptotically a noisy version of the PSD. Let η be the periodogram convergence noise:

$$P^\gamma = S^\gamma + \eta = S^{tur} + S^{vib,1} + S^{vib,2} + \dots + S^w + \eta. \quad (20)$$

Without averaging the periodogram, i.e., by considering directly the squared Fourier transform of γ, the SNR of the periodogram is 1:

$$\sigma_\eta(f) = S^\gamma(f).$$

Note that η dimension is the same as S^γ (for instance, mas^2/Hz), as it is an estimation error on S^γ. This estimation noise error could be reduced by using, for instance, averaged periodograms (i.e., modified Welch's periodograms). However, these methods lead to a spectrum estimate on a coarser grid, which affects the precision of the peak detection. Moreover, the identification method presented here makes use of several thousand periodogram points to estimate a few parameters by maximum likelihood, so the convergence noise $\sigma_\eta(f)$ is *de facto* filtered. The selection of the spectrum estimation method best suited to our identification procedure is still an open issue.

C. Source Separation

Let f_s be the data sampling frequency. The principle of the method is to segment the periodogram frequency range $[0, f_s/2]$ in regions where only one component is present.

1. Noise Isolation

We first isolate the measurement noise PSD S^w by assuming that for frequencies higher than f_w (see Fig. 3), the measurement noise is much stronger than the disturbance:

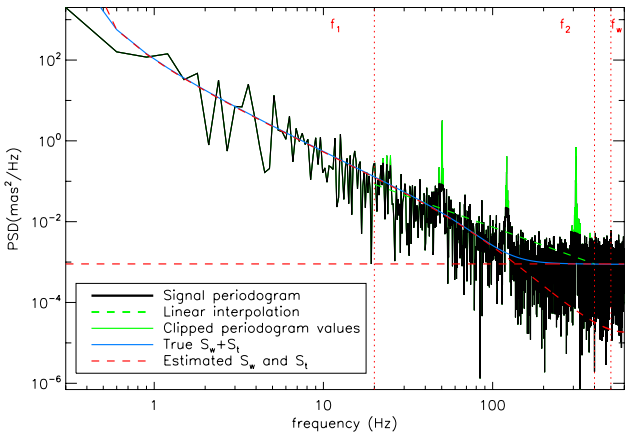

Fig. 3. (Color online) Source separation by segmenting the frequency range.

$$P^{\gamma} \simeq S^{w} + \eta \text{ for } f > f_w.$$

The average of P^{γ} in this frequency range provides an estimate $\hat{S}^{w}$ of the measurement noise spectrum level (S^{w} is constant, as the measurement noise is assumed to be white). The measurement noise variance is obtained by $\hat{\sigma}_w^2 = f_s \hat{S}^{w}$, with f_s the sampling frequency of the time sequence.

2. Turbulence/Wind Shake Isolation

We then assume that no vibration peak is present outside a $[f_1, f_2]$ frequency range. We build a linear (in log-log space) interpolation of the periodogram P in the $[f_1, f_2]$ range. In this range, the periodogram values higher than the linear interpolation plus $3\sigma_{\eta}$ belong to the vibration peaks. We clip these values and obtain the clipped periodogram $\bar{P}$ of Fig. 3. Considering the previously obtained measurement noise PSD estimate $\hat{S}^{w}$, we form the following data model for $\bar{P}$:

$$\bar{P} \simeq \hat{S}^{w} + S^{tur} + \eta \text{ for } f \notin [f_1, f_2], \qquad f \leqslant f_w. \tag{21}$$

3. Default Settings for f_1, f_2, f_w

The vibration frequency range $[f_1, f_2]$, within which are all the vibration peaks, and the noise limit frequency f_w are user-defined parameters. In a variety of cases we found that the following setting is satisfactory:

$$f_1 \triangleq f_s/60, \qquad f_2 = f_w \triangleq f_s/3,$$

but these parameters can be adjusted to the considered application. No other parameter setting is required in the identification method

D. Disturbance Model Identification

1. Parameters to Identify

As previously mentioned, turbulence/wind shake and vibration components can be described as AR2 processes, defined by parameters a_1, a_2 and the variance of the generating noise (which conveys the energy of the considered component). As shown in Appendix B, we can use an equivalent description with physical parameters: the damping coefficient k, the central frequency f, and still the variance of the generating noise. In the following, we show how to estimate the turbulence/wind shake model parameters $[k^{tur}, f^{tur}, \sigma^2(v^{tur})]$ and then to sequentially estimate the vibration parameters $[k^{vib,i}, f^{vib,i}, \sigma^2(v^{vib,i})]$.

2. Turbulence/Wind Shake Parameter Estimation

We determine $[k^{tur}, f^{tur}, \sigma^2(v^{tur})]$ by a maximum-data likelihood method based on Eq. (21). To avoid local minima issues, we search for the three parameters' optimal values by an exhaustive search on a 3D grid. As illustrated in Fig. 3, this fitted model is very close to the actual turbulence PSD in the frequency range where the measurement noise is negligible. The turbulence PSD obtained with these identified parameters is denoted by $\hat{S}^{tur}$.

3. Vibration Parameter Estimation

The identification of the vibration parameters is performed sequentially. At step i, i vibration components have been identified, and the corresponding estimated PSDs $\hat{S}^{vib,i}$ have been computed. The total identified PSD at this stage is

$$\hat{S} \triangleq \hat{S}^{w} + \hat{S}^{tur} + \sum_{j=0}^{i} \hat{S}^{vib,j}.$$

We assume that the periodogram contains *one more* vibration peak $S^{vib,i+1}$ and consider the following data model [cf. Eqs. (19) and (20)]:

$$P \simeq \hat{S} + S^{vib,i+1} + \eta, \quad \sigma_{\eta} \simeq \hat{S}. \tag{22}$$

As we did for the turbulence parameter estimation, we determine $[k^{vib,i+1}, f^{vib,i+1t}, \sigma^2(v^{vib,i})]$ by a maximum-data likelihood exhaustive optimization on a 3D grid. Figure 4 illustrates the first three steps of the identification pattern.

This procedure estimates vibration peaks with decreasing energy. It stops when reaching a specified number N_{max} of vibration components or when the energy of the last identified peak is below α_{min} times the total energy of the signal. N_{max} shall be set according to the maximum state-space dimension allowed in the AO loop architecture. In the applications of this identification method presented below, we chose $N_{max} = 10$ and $\alpha_{min} = 10^{-}6$.

4. VIBRATION FILTERING PERFORMANCE ASSESSMENT ON SYNTHETIC DATA

In this section, we aim at characterizing the performance of our vibration identification method in the SPHERE case [2].

A. Principle of the Study

1. Simulation Tool

The simulation tool we use has been developed for SPHERE AO (SAXO) preliminary and final design [24]. This is an end-to-end numerical simulator of the full SAXO system, based on the following modules: a turbulence simulator, a filtered Shack–Hartmann WFS module, a DM module, a control module, and an imaging module with coronographic imaging. The control architecture is based on a decoupling of tip and tilt modes, those modes

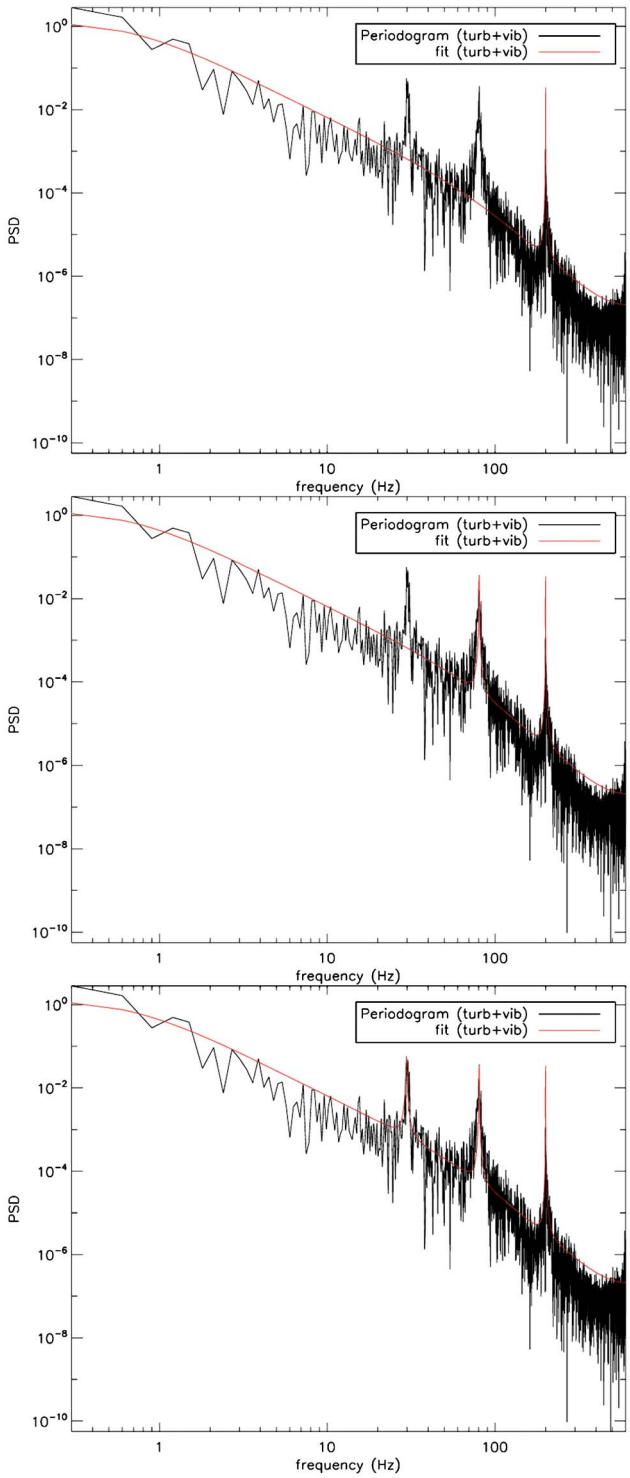

Fig. 4. (Color online) Total estimated PSD $\hat{S}$ at each step.

being controlled through LQG, while higher-order modes correction is ensured by OMGI: this is consistent with the assumptions made in Section 2. The sampling frequency of the loop is 1200 Hz, and we consider time sequence data up to 4000 points for model identification as well as for residual phase variance computation.

2. Vibration Filtering Efficiency in Closed Loop p

We assess the quality of the identification by computing the empirical residual phase variance via a closed-loop end-to-end simulation. We compute performance as the percentage of the vibration energy that is filtered. To do so, we compute

- the residual phase variance σ_a^2 when there are no vibration peaks in the disturbance (it corresponds to the ideal case of 100% vibration energy filtered);
- the residual phase variance σ_b^2 obtained when the vibration peaks are present in the disturbance but when the vibration model in the LQG control law is void. This is the worst case, corresponding to 0% vibration energy filtered.

For each test, we compute the residual phase variance σ^2 and calculate the vibration filtering percentage p by

$$p = 100\frac{\sigma_b^2 - \sigma^2}{\sigma_b^2 - \sigma_a^2}.$$

It is equal to 100% when $\sigma^2 = \sigma_a^2$ and 0% when $\sigma^2 = \sigma_b^2$.

3. Turbulence and Vibration Profiles

The turbulence profile used in the simulation is a three-layer profile, whose parameters are summarized in Table 2. We consider two vibration profiles. V1 is made of three vibration peaks at 50, 120.5, and 310.5 Hz (see Table 3). All peaks are sharp (damping coefficient k below 10^{-3}), so that there is no overlap between peaks. V2 is a more complex vibration set comprising 10 vibration components of various damping coefficient values. Some peaks are very close one to the other, and the damping coefficient k ranges from 10^{-2} to 10^{-5} (see Table 4).

The PSDs of these profiles are shown in Fig. 5.

Table 2. Three-Layer Turbulence Profile T1[a]

Altitude (m)	100	2200	8000
Relative C_n^2	50%	20%	30%

[a] L_0=25 m, wind speed=12.4 m/s, seeing=0.85″@0.5 μm.

B. Results

We have considered three measurement noise levels, denoted by Worst (W), Typical (T), and Best (B), corresponding to a flux per subpupil worth respectively 50, 100, and 250 photons. The results presented in Table 5 show that the closed-loop vibration filtering performance in the nominal case is very close to what would be obtained with perfect knowledge of the turbulence and vibration models. With a more energetic vibration profile such as "V2," using our identification and filtering is even more beneficial, even with a limited SNR (corresponding to the "W" noise level).

All the aforementioned tests were conducted on 4000-point time sequences. Table 6 provides the performance obtained with 1000 and 2000 points.

Table 3. Three-Peak Vibration Profile V1

f [Hz]	50.0	120.5	310.5
k	5e−4	0.001	0.002
$\sigma^2(v^{vib})$ [mas^2]	3.3e−6	2.5e−5	1.5e−4

Table 4. Ten-Peak Vibration Profile V2

f [Hz]	40.0	50.0	75.0	90.0	95.0	110.0	180.0	200.0	300.0	350.0
k	1e−5	1e−5	1e−3	1e−5	1e−5	1e−2	1e−4	1e−2	1e−3	1e−5
$\sigma^2(v^{vib})$ [mas^2]	5e−5	5e−4	5e−5	5e−5	1.5e−4	2.5e−5	5e−5	4e−4	5e−5	1e−4

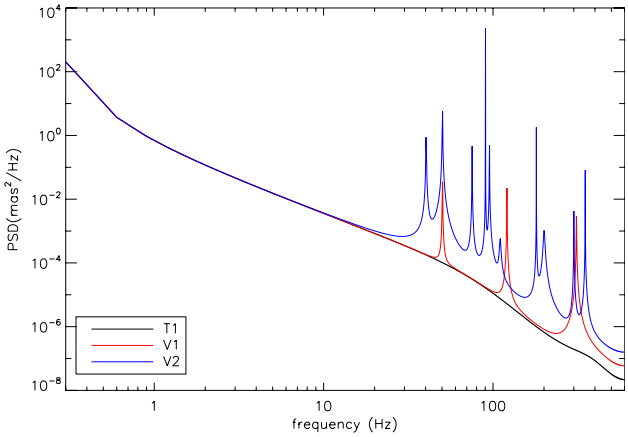

Fig. 5. (Color online) Turbulence and vibration profiles used in the simulations.

We can see that the performance does not degrade when considering only 2000 data points for the identification, whereas it drops dramatically when using 1000 data points. We have therefore specified a baseline number of data points worth 2000, with 4000 as goal value. In the SAXO case, the baseline value corresponds to an actualization of the models *typically every second*. This value can certainly be refined with respect to the typical evolution of the disturbance model parameters.

C. Tracking Capabilities

The identification procedure can be run periodically and can identify a whole new vibration set each 2000 frames, i.e., typically every second. As noted in Subsection 3.A, the disturbance statistics evolve at a much slower rate.

Table 5. Vibration Identification and Filtering Performance[a]

Test Case	Profile	Noise Level (W/T/B)		
Perfect model	T1/V1	81%	82%	83%
Identified	T1/V1	68%	79%	79%
Perfect model	T1/V2	92%	92%	92%
Identified	T1/V2	90%	91%	90%

[a]Up: with vibration profile "V1." Down: with vibration profile "V2."

Table 6. Vibration Identification and Filtering Performance as a Function of the Time Sequence Length

Test Case	Profile	Noise Level (W/T/B)		
Identified (4000 points)	T1/V1	68%	79%	79%
2000 points	T1/V1	68%	79%	80%
1000 points	T1/V1	44%	46%	62%

On specific cases with fast-varying disturbance statistics, one might run the identification process more often, by considering the 2000 last data points. Another possibility would be to use a tracking system such as in [21], with an initialization model provided by our method and refreshed every 2000 frames. However, such cases are outside the scope of the present paper.

We also studied what happens if the profile does not contain any vibrations. The identification method will still identify peaks with low energy. These "false alarm" identified peaks will affect the correction of the turbulence. We estimated the turbulence filtering degradation to be lower than 3% of the residual phase variance obtained with a perfect model. Actually, even when the vibration model V1 is used where there are no vibrations, the overall performance loss does not exceed 6%. Qualitatively, our interpretation is that the LQG control law is robust to a mis-estimation of the vibration peak magnitude but needs to be fed with precise estimates of the peak frequencies.

D. Limitations of the Method

The identification method we developed is an iterative algorithm, identifying components of decreasing energy. Nothing is done in the process to minimize the number of identified peaks, so that it is possible that a lighter model—i.e., with fewer components—would yield equivalent performance. This flaw is common to methods estimating sequentially elementary components to describe a complex signal, such as the image reconstruction algorithm CLEAN [25] widely used in radioastronomy. This could be circumvented by adding an Akaike-like penalty [13] to the optimization or by running a peak-picking procedure.

Also, we have considered an instantaneous-response DM. This approximation is true as long as the typical response time of the DM is much smaller than the loop rate. We are currently working on the evaluation of the DM dynamic effects on the TT filtering performance in existing AO systems. Including the DM dynamics in the model, as in [20,26,27] for instance, would constitute an upgrade of the method.

5. VALIDATION ON EXPERIMENTAL DATA

We have used the identification procedure on NAOS experimental reconstructed open-loop tip–tilt measurements. Data are courtesy of Yann Clenet. Figure 6 shows the data periodogram and the estimated profile.

Both vibration and turbulence components are successfully fitted. It was not possible to experimentally check how a Kalman filter would behave in this case because such a control law is not implemented on NAOS. However, we estimated the performance obtained with and without vibrations for an OMGI and an LQG control law,

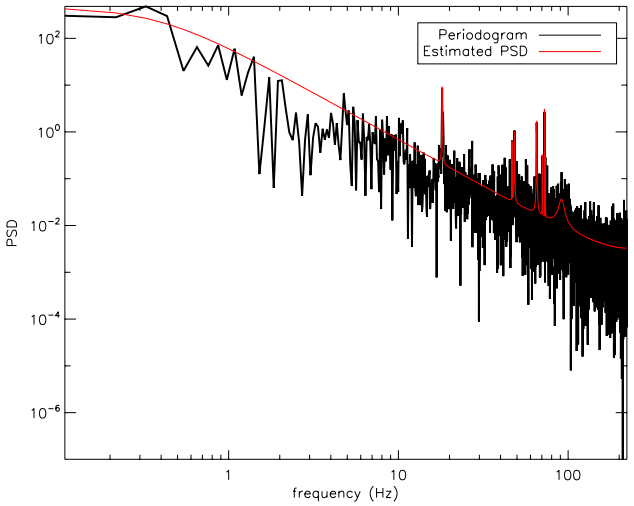

Fig. 6. (Color online) Estimated PSD on NAOS experimental tilt slopes.

Table 7. Estimated Residual Phase Variance Based on an Experimental NAOS Periodogram, with and Without Vibrations for an OMGI and an LQG

Without Vibrations		With Vibrations	
OMGI	LQG	OMGI	LQG
15.4 mas^2	9.5 mas^2	61.0 mas^2	24.8 mas^2

based on the identified profiles. The results, summarized in Table 7, were obtained by computing the residual phase variance with Eq. (7) on the identified profile. They show a significant vibration filtering improvement using an LQG control law associated with our identification method.

Of course, these results are preliminary and have to be confirmed on a variety of data sets, and eventually on sky. The corresponding experimental tests will be performed during the assembly integration and tests of the SAXO system.

6. CONCLUSION AND FUTURE WORKS

We have proposed an identification method, suited to AO systems affected by turbulence or wind shake only, as well as in the presence of vibration peaks. It does not rely on a prior information about the number of peaks nor the central frequency of each vibration peak. This process can be done on-line, i.e., without opening the loop, and cycled in order to periodically retune the filter in the case of an evolution of the disturbance statistics. Its performance has been computed on synthetic profiles and estimated on experimental data, leading to a considerable gain over classical integrator control schemes. A formal robustness study remains to be done, yet our method has been applied successfully to a variety of profiles.

Although designed for tip tilt, our method can be extended to a multi-modal context (as soon as a basis of independent turbulent modes is available, the procedure can be run on each mode). It could also be profitably extended to a joint estimation of turbulence parameters (outer scale, wind speed) as well as multilayer characterization. A peak-picking procedure could be run on the data to automatically set the number of vibration components to identify. Additionally, we are currently working on a model refinement procedure using the closed-loop data to further improve vibration filtering. Several alternative trails may be explored among the numerous methods in parametric identification, in particular if the model structure constraint is relaxed, such as real-time identification with an EKF on long AR models, batch identification such as SSI methods [22], or pencil matrix methods [28], to cite a few.

The next step will be the experimental validation of the procedure on SAXO, first in lab in 2010, then on sky. Indeed, the present TT vibration and filtering strategy is the one that has been specified for SAXO's RTC [11]. We thus hope to demonstrate the feasibility of an unsupervised vibration filtering scheme in an extreme AO system.

APPENDIX A: DEFINITION AND PROPERTIES OF AN AR PROCESS

A random discrete time signal x_n is said to be ARp if there exists a set of real coefficients $\{a_q\}_{1\leqslant q\leqslant p}$ and a sampled white noise v of variance $\sigma^2(v)$ such that

$$x_n = \sum_{q=1}^{p} a_q x_{n-q} + v_n. \tag{A1}$$

The z transform of this equation gives

$$H(z) = Tz(x)/Tz(v) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}}.$$

The PSD S^x of x is given by

$$S^x(f) = \sigma^2 \cdot \left| 1 - \sum_{q=1}^{p} a_q \, \mathrm{e}^{-2\mathrm{i}\pi q f} \right|^{-2}. \tag{A2}$$

APPENDIX B: AUTO-REGRESSIVE PROCESS OF THE SECOND ORDER

Let us suppose that the incoming vibration is an oscillatory signal, characterized by its damping coefficient k (related to the vibration bandwidth) and its natural frequency f^{vib}. The sampled version of the vibration, obtained by averaging the signal over one time period $[(n-1)T,(n)T]$, can suitably be approximated by a scalar second-order auto-regressive (AR2) model [29]: Thus, we can link the AR2 model coefficients to more meaningful ones, namely, the oscillatory continuous-time parameters k^{vib}, f^{vib}:

$$a_1 = 2e^{-2k^{vib}\pi f^{vib}T} \cos(2\pi f^{vib} T \sqrt{1 - k^{vib^2}}), \tag{B1}$$

$$a_2 = -e^{-4k^{vib}\pi f^{vib}T}. \tag{B2}$$

Variance σ_ξ^2 behaves as a scaling factor. For values of k higher than 0.66, the PSD is strictly decreasing. For values of k higher than 1, the a_1 expression remains valid us-

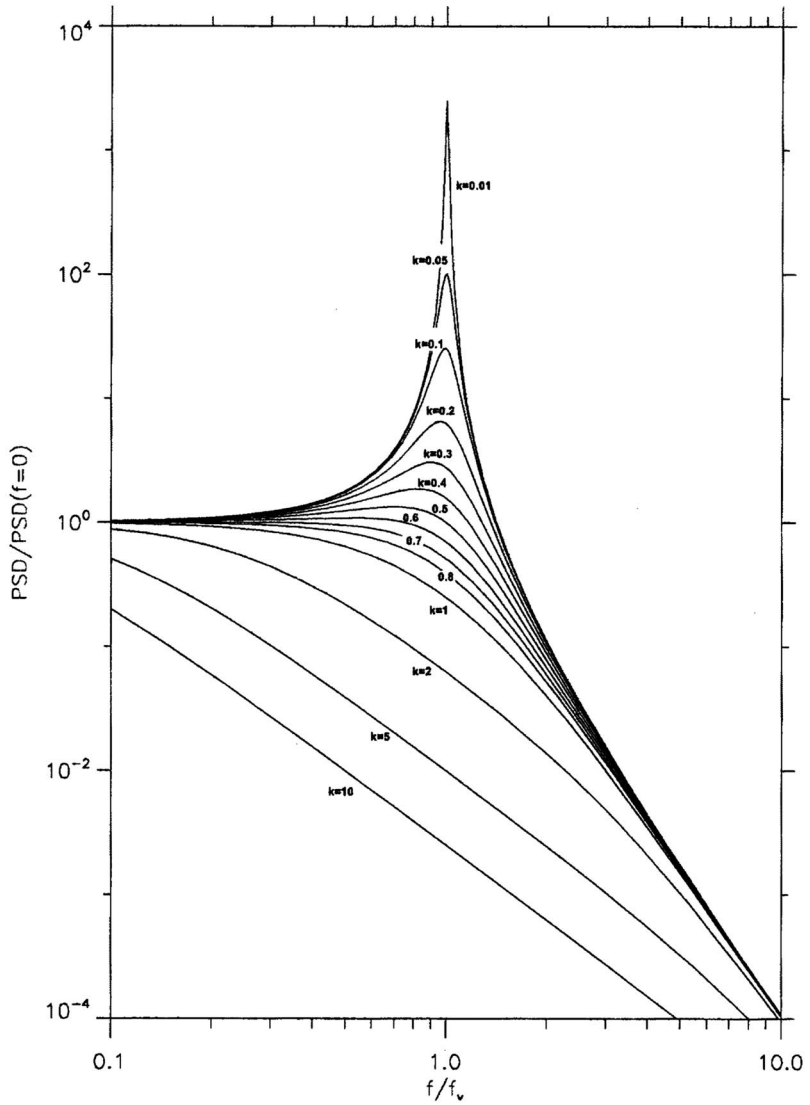

Fig. 7. PSD trends of an AR2 obtained for various values of k.

ing $\sqrt{1-k^2}=i\sqrt{k^2-1}$. The PSD trends of an AR2 model obtained for various values of k are shown in Fig. 7.

ACKNOWLEDGMENTS

This work benefited from fruitful discussions with Henry-François Raynaud, Carlos Correia, Jean-Marc Conan, and Vincent Michau. NAOS experimental data are courtesy of Yann Clenet. This work was partly supported by the Centre National de la Recherche Scientifique (Subcontract No. 1105601/SPHERE_CNRS_ALPES/LAOG/ONERA) and by the European Southern Observatory (Contract No. 22562/ESO/INS/08/19554/JSC).

REFERENCES

1. T. Fusco, V. Michau, S. Meimon, Y. Clenet, M. Cohen, J. Paufique, E. Gendron, D. Gratadour, N. Hubin, J. Montri, C. Petit, C. Robert, and G. Rousset, "Atlas: The ltao module of the E-ELT," presented at the Adaptive Optics for Extremely Large Telescopes (AO4ELT) 2009 Conference, June 22–26, 2009, Paris.
2. J.-L. Beuzit, M. Feldt, D. Mouillet, C. Moutou, K. Dohlen, P. Puget, T. Fusco, P. Baudoz, A. Boccaletti, S. Udry, D. Ségransan, R. Gratton, M. Turatto, H.-M. Schmid, R. Waters, D. Stam, P. Rabou, A.-M. Lagrange, F. Ménard, J.-C. Augereau, M. Langlois, F. Vakili, L. Arnold, T. Henning, D. Rouan, M. Kasper, and N. Hubin, *A Planet Finder Instrument for the VLT*, Proceedings of the International Astronomical Union (International Astronomical Union, 2005), Vol. 1, pp. 317–322.
3. J.-P. Véran and L. Poyneer, "Evaluation of the T/T conditions at GEMINI south using NICI AO telemetry data," in *Proceedings of 1st AO4ELT Conference—Adaptative Optics for Extremely Large Telescopes* (EDP Sciences, 2010), p. 05002.
4. R.-J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
5. B. Sedghi, "E-ELT main axis control analysis," Tech. Rep. 3 (European Southern Observatory, 2007).
6. Y. Clénet, M. E. Kasper, N. Ageorges, C. Lidman, T. Fusco, O. Marco, M. Hartung, D. Mouillet, B. Koehler, G. Rousset, and N. N. Hubin, "NACO performance: Status after 2 years of operation," Proc. SPIE **5490**, 107–117 (2004).
7. J. A. Stoesz, J.-P. Véran, F. J. Rigaut, G. Herriot, L. Jolissaint, D. Frenette, J. Dunn, and M. Smith, "Evaluation of the on-sky performance of ALTAIR," Proc. SPIE **5490**, 67–78 (2004).

8. B. Anderson and J. Moore, eds., *Optimal Filtering* (Dover, 2005).
9. C. Petit, J.-M. Conan, C. Kulcsar, H.-F. Raynaud, T. Fusco, J. Montri, and D. Rabaud, "Optimal control for multiconjugate adaptive optics," C. R. Phys. **6**, 1059–1069 (2005).
10. B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. M. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," J. Opt. Soc. Am. A **21**, 1261–1276 (2004).
11. C. Petit, J.-M. Conan, T. Fusco, E. Fedrigo, C. Kulcsár, and H.-F. Raynaud, "Optimization of the control laws for the SPHERE XAO system," Proc. SPIE **7015**, 70151D (2008).
12. B. Peeters and G. De Roeck, "Reference-based stochastic subspace identification for output-only modal analysis," Mech. Syst. Signal Process. **13**, 855–878 (1999).
13. L. Ljung, *System Identification: Theory for the User*, 2nd ed. (Prentice Hall, 1998).
14. C. Petit, J.-M. Conan, C. Kulcsár, and H.-F. Raynaud, "Linear quadratic Gaussian control for adaptive optics and multiconjugate adaptive optics: experimental and numerical analysis," J. Opt. Soc. Am. A **26**, 1307–1325 (2009).
15. E. Gendron and P. Léna, "Astronomical adaptive optics I. modal control optimization," Astron. Astrophys. **291**, 337 (1994).
16. C. Kulcsár, H.-F. Raynaud, C. Petit, J.-M. Conan, and P. Viaris de Lesegno, "Optimal control, observers and integrators in adaptive optics," Opt. Express **14**, 7464–7476 (2006).
17. C. Mohtadi, "Bode's integral theorem for discrete-time systems," IEE Proc.-D: Control Theory Appl. **137**, 57–66 (1990).
18. C. Dessenne, "Commande modale et prédictive en optique adaptative," Ph.D. thesis (Université de Paris VII, 1998).
19. S. Meimon, T. Fusco, Y. Clenet, J.-M. Conan, and V. Michau, "The hunt for 100% sky coverage," Proc. SPIE **7736**, 77360Y (2010).
20. N. Perez-Arancibia, J. Gibson, and T.-C. Tsao, "Frequency-weighted minimum-variance adaptive control of laser beam jitter," IEEE/ASME Trans. Mechatron. **14**, 337–348 (2009).
21. N. D. Lieto, P. Haguenauer, J. Sahlmann, and G. Vasisht, "Adaptive vibration cancellation on large telescopes for stellar interferometry," Proc. SPIE **7013**, 70130H (2008).
22. K. Hinnen, M. Verhaegen, and N. Doelman, "Exploiting the spatiotemporal correlation in adaptive optics using data-driven h2-optimal control," J. Opt. Soc. Am. A **24**, 1714–1725 (2007).
23. E. Walter and L. Pronzato, *Identification of Parametric Models from Experimental Data* (Springer-Verlag, 1997).
24. T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper, and P. Puget, "High order adaptive optics requirements for direct detection of extra-solar planets. application to the sphere instrument," Opt. Express **14**, 7515–7534 (2006).
25. Hogbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," Astron. Astrophys. Suppl. Ser. **15**, 417–426 (1974).
26. C. Correia, C. Kulcsár, J.-M. Conan, and H.-F. Raynaud, "On the optimal wave-front reconstruction and control in adaptive optics with mirror dynamics," J. Opt. Soc. Am. A **27** (2010) (to be published).
27. D. P. Looze, "Linear-quadratic-gaussian control for adaptive optics systems using a hybrid model," J. Opt. Soc. Am. A **26**, 1–9 (2009).
28. Y. Hua and T. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," IEEE Trans. Acoust., Speech, Signal Process. **38**, 814–824 (1990).
29. C. Petit, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and T. Fusco, "First laboratory validation of vibration filtering with LQG control law for Adaptive Optics," Opt. Express **16**, 87–97 (2008).