

Stochastic modeling of stationary scalar Gaussian processes in continuous time from autocorrelation data

Martin Hanke^{1*}

^{1*}Institut für Mathematik, Johannes Gutenberg-Universität,
55099 Mainz, Germany.

Corresponding author(s). E-mail(s): hanke@math.uni-mainz.de;

Abstract

We consider the problem of constructing a vector-valued linear Markov process in continuous time, such that its first coordinate is in good agreement with given samples of the scalar autocorrelation function of an otherwise unknown stationary Gaussian process. This problem has intimate connections to the computation of a passive reduced model of a deterministic time-invariant linear system from given output data in the time domain. We construct the stochastic model in two steps. First we employ the AAA algorithm to determine a rational function which interpolates the z -transform of the discrete data on the unit circle and use this function to assign the poles of the transfer function of the reduced model. Second we choose the associated residues as the minimizers of a linear inequality constrained least squares problem which ensures positivity of the transfer function's real part for large frequencies. We apply this method to compute extended Markov models for stochastic processes obtained from generalized Langevin dynamics in statistical physics. Numerical examples demonstrate that the algorithm succeeds in determining passive reduced models, and that the associated Markov processes provide an excellent match of the given data.

Keywords: Stochastic realization, Markov model, model reduction, passive model, positive real function, AAA algorithm

MSC Classification: 65L09 , 94A12 , 93C05 , 65D15 , 41A20

Dedicated to the memory of Claus Schneider

1 Introduction

The stochastic realization problem concerns the question whether a given stationary Gaussian process in continuous time has a realization in terms of a Markov process of higher dimension, cf. Lindquist and Picci [21]. In this paper we refer to stochastic modeling, when an approximating Markov process is determined numerically from given time samples of the autocorrelation function of the stationary process in question, independent of whether or not a stochastic realization exists. This is also known as the inverse problem of stationary covariance generation, cf. Kalman [20] and Anderson [2].

Our interest in stochastic modeling is driven by applications from statistical physics. Consider, e.g., the classical example of the equilibrium dynamics of a particle in a heat bath (Zwanzig [29]). Assume that the velocity of the particle is being measured, and an effective stochastic model for its dynamics is sought. The full Langevin dynamics of the particle together with the constituents of the heat bath provides a stochastic realization of this process, but this is not tractable analytically and much too large and costly to simulate, anyway. Rather, one is interested in a small extended Markov system with a few handful of auxiliary variables, the simulation of which reproduces the main features of the particle's dynamics.

The stochastic realization problem has strong connections to the theory of deterministic time-invariant linear dynamical systems, and the stochastic modelling problem is related to the corresponding model reduction aspect – without any given information about the underlying system, be it finite dimensional or not. The major difference to the common deterministic problem setting is the quality and amount of data: for applications like the one described above the number of time samples may be less than a hundred, and they are noisy.

As will be explained in Section 2 the stochastic problem comes with the additional catch that it is indispensable that the reduced models be passive in the systems theory terminology. Passivity is a constraint that is easiest to formulate in the frequency domain; it is difficult to achieve this property in the time domain when the use of the Laplace transform is prohibited by the quality of the data. Of course, the need for passive models is also relevant in many deterministic applications, but so far there is no available off-the-shelf algorithm which applies to our setting; see Section 2 for a brief review of the pertinent literature.

In this paper we restrict ourselves to scalar problems and present a method which accounts for passivity, but nevertheless operates in the time domain; this way we are able to control the quality of our approximation of the original process. The method is fairly simple to implement and requires only a small number of linear systems to be solved. It employs in a first step the AAA algorithm for rational interpolation by Nakatsukasa, Sète, and Trefethen [22] to determine the eigenvalues of a reduced system matrix, followed by a constrained linear approximation to settle the remaining free parameters in a second step. The (in)equality constraints make sure that passivity is achieved for high frequencies, which is the troublesome frequency regime in our applications. As we demonstrate by numerical examples with data from a statistical physics application, our algorithm is feasible for the treatment of significantly larger data sets than could have been handled previously.

The outline of the paper is as follows. In Section 2 we specify the stochastic modeling problem that we consider and draw the connections to linear systems theory. In Section 3 we review the AAA algorithm and its potential use for model reduction purposes. Subsequently, in Section 4, we develop the algorithm that we propose to compute approximations of the given autocorrelation function, which have all the requirements that are needed to construct a corresponding Markov model. Numerical examples are presented in Section 5. To be self contained we show in Sections 6 and 7 how to set up the extended Markov system: Although the material in these two sections is mostly general folklore, some of the details are difficult to extract from the literature and the presentation is adapted to our needs. Finally, in an appendix, we discuss the consistency of the constrained approximation problem which is considered in Section 4.

2 The stochastic modeling problem

For $t \in \mathbb{R}$ and $s \geq \max\{0, -t\}$ let

$$y(t) = \mathbb{E}[Y(s+t)Y(s)], \quad t \in \mathbb{R},$$

be the autocorrelation function of a real-valued centered stationary Gaussian process $Y = Y(t)$, $t \geq 0$, with positive variance, and assume that we are given $2n$ (possibly noisy) samples

$$y_\nu \approx y(\nu\tau), \quad \nu = 0, 1, \dots, 2n-1,$$

of y on some equidistant time grid with spacing $\tau > 0$. The stochastic model of Y that we want to construct is formulated in terms of an Ornstein-Uhlenbeck equation

$$d \begin{bmatrix} X \\ Z \end{bmatrix} = A \begin{bmatrix} X \\ Z \end{bmatrix} dt + g dW \quad (2.1)$$

with a stable matrix $A \in \mathbb{R}^{m \times m}$ and a scalar Brownian motion W acting in the direction $g \in \mathbb{R}^m$. The goal is to choose A and g in such a way that the scalar component X of the stationary solution of (2.1) approximates the original process Y in the sense that

$$\mathbb{E}[X(s + \nu\tau)X(s)] \approx y_\nu, \quad \nu = 0, \dots, 2n-1. \quad (2.2)$$

The $m-1$ components of $Z \in \mathbb{R}^{m-1}$ are auxiliary (dummy) variables for this purpose and of little interest in practice. We remark that there is no loss of generality in restricting the attention to a scalar Brownian motion in (2.1). This is a consequence of the Positive Real Lemma (see Section 7), as the process Y is itself a scalar one.

We refer to Pavliotis [23] for a general reference concerning the stochastic differential equation (2.1), its stationary solution, and stationary processes in general. We will make use of the fact that the covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ of the stationary solution of (2.1) is determined by the Lyapunov equation

$$A\Sigma + \Sigma A^T = -gg^T, \quad (2.3)$$

and that the autocorrelation function of X is given by

$$\mathbb{E}[X(s+t)X(s)] = e_1^T e^{|t|A} \Sigma e_1, \quad t \in \mathbb{R}, \quad (2.4)$$

cf. [23, Section 3.7]. Here and throughout we denote by e_1 the first Cartesian basis vector in the (real or complex) Euclidean space, the dimension of which is determined by the context. Likewise we write I for the associated identity matrix.

In view of our goal (2.2) we exclude systems (2.1), for which the autocorrelation function of X is identically zero. We may further assume without loss of generality that X and Z are stochastically independent, for this can always be achieved by a coordinate transformation, i.e., by representing the auxiliary variables in a different basis. Then

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & \Sigma_0 \end{bmatrix} \quad (2.5)$$

for some $\sigma^2 > 0$ and some Hermitian positive semidefinite matrix $\Sigma_0 \in \mathbb{R}^{(m-1) \times (m-1)}$, and (2.4) simplifies to

$$\mathbb{E}[X(s+t)X(s)] = \sigma^2 e_1^T e^{|t|A} e_1 =: \varphi(t), \quad t \in \mathbb{R}. \quad (2.6)$$

The autocorrelation function of any second order stationary stochastic process is a function of positive type, cf., e.g., [23, Section 1.2], and hence, the Fourier transform

$$\widehat{\varphi}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} \varphi(t) dt = 2\sigma^2 \operatorname{Re} e_1^T (i\xi I - A)^{-1} e_1 \quad (2.7)$$

of φ is nonnegative for every $\xi \in \mathbb{R}$ by Bochner's theorem. Vice versa, if $\widehat{\varphi}$ is a nonnegative function then there is a vector $g \in \mathbb{R}^m$, such that the stochastic differential equation (2.1) has a stationary solution which satisfies (2.6); see Section 7 for details.

If the approximation $X \approx Y$ were exact, i.e., if $\varphi = y$, then y is the output of the deterministic dynamical system

$$\begin{aligned} \dot{x}(t) &= Ax + \sigma^2 e_1 u, \\ y(t) &= e_1^T x, \end{aligned} \quad (2.8)$$

with homogeneous initial data, where $u = u(t)$ is a delta distribution at $t = 0$. In control theory, u represents the control (or the input) of the system. Associated with (2.8) is the so-called transfer function

$$\kappa(\zeta) = \sigma^2 e_1^T (\zeta I - A)^{-1} e_1, \quad \zeta \in \mathbb{C}. \quad (2.9)$$

In view of (2.7) and Bochner's theorem we observe that κ has a nonnegative real part on the imaginary axis, and since A is stable, κ is analytic in \mathbb{C}^+ , the open right complex half plane. Such functions are termed positive real, and when κ is positive real, then the system (2.8) is called passive. We can thus rephrase our stochastic modelling problem as the task to construct a passive deterministic system such that the output is in good agreement with the given samples y_ν .

In contrast to control theory, however, we have no control on u ; the input is necessarily given by a delta distribution. Since this particular input is not localized in frequency space, it prohibits, for example, the use of an algorithm suggested by Cherifi, Goyal, and Benner [10] for finding a passive model that is compatible with given time-domain data. Other works which deal with the construction of passive reduced models, e.g., [9, 16, 25], have proposed fixes of preliminary non-passive models by modifying their coefficients to shift negative values of $\hat{\varphi}$ which are present in finite frequency bands; these algorithms, however, are neither guaranteed to succeed, nor can they treat negative values of $\hat{\varphi}$ in unbounded frequency bands. Unfortunately, the latter is the main cause of difficulties in our application, because the Fourier transforms of autocorrelation functions which are of interest to us, vanish at infinity. This also rules out a method by Fazzi, Gugliemi, and Lubich [13], which requires such a fix in a preliminary step.

Other papers, e.g., [11, 14, 15], have suggested nonlinear optimization to modify the system coefficients in order to find nearby systems which satisfy the so-called Lur'e equations (see (7.3) below) that can be used to characterize passivity. Those methods have subtle convergence issues, are quite complicated to formulate, and their impact on the data fit in the time domain is not transparent. The algorithm which we propose here is much easier to implement, and a mean square approximation (2.2) of the given data is its primary objective.

3 The exponential approximation problem

We will stipulate the assumption that the matrix A in (2.1) is diagonalizable (we will briefly elaborate on this assumption in Remark 3.1). Then the right-hand side of (2.6) is a so-called Prony series, i.e., a real-valued linear combination*

$$\varphi(t) = \sum_{j=1}^m \alpha_j e^{t\lambda_j} \quad \text{for } t \geq 0 \quad (3.1)$$

of complex exponentials with some appropriate weights $\alpha_j \in \mathbb{C}$, $j = 1, \dots, m$. This function decays exponentially because the λ_j are the eigenvalues of A , and since A is stable, those belong to \mathbb{C}^- , the open left complex half plane.

To achieve the goal (2.2) we want to determine a Prony series φ which satisfies

$$\varphi(\nu\tau) \approx y_\nu, \quad \nu = 0, \dots, 2n - 1, \quad (3.2)$$

in a mean square sense. To this end we need to select an appropriate number m of terms in (3.1) and choose corresponding exponents λ_j and weights α_j , $j = 1, \dots, m$. This is a long-standing delicate numerical problem, cf., e.g., Varah [27], because (i) it is nonlinear in the parameters λ_j , and (ii) the outcome is very sensitive to perturbations in the data. In particular, as the exponents have negative real parts, the given data y_ν carry less and less additional information, the higher the value of ν . One can therefore

*Here we adopt the terminology ‘‘Prony series’’, which is commonly used in applications, despite the fact that the sum (3.1) consists of only finitely many terms.

expect the necessary number m of terms in (3.1) to drive the error in (3.2) below a given tolerance to be considerably smaller than n ; note, however, that for $m \geq n$ it is generically possible to realize (3.2) with equality, as there are $2n$ equations to satisfy and $2m$ parameters to choose.

We emphasize that we do not claim or attempt to resolve a presumably valid Prony series representation of the true autocorrelation function, which is a highly ill-conditioned problem. For our purposes it is sufficient to approximate the data by *some* Prony series representation, which is less ambitious and much better conditioned.

A common way to approach the exponential approximation problem (3.2) is via the generating function (or z -transform)

$$f(z) = \sum_{\nu=0}^{\infty} y_{\nu} z^{-\nu-1} \quad (3.3)$$

of the data. Stipulating the very mild restriction that the samples $(y_{\nu})_{\nu \in \mathbb{N}}$ are absolutely summable, this function f is analytic in the exterior of the unit disk and extends continuously to the unit circle. Imagine that (3.2) holds with equality. Then the series can be rewritten as

$$f(z) = \sum_{\nu=0}^{\infty} \sum_{j=1}^m \alpha_j e^{\nu \tau \lambda_j} z^{-\nu-1} = \sum_{j=1}^m \frac{\alpha_j}{z} \sum_{\nu=0}^{\infty} (e^{\tau \lambda_j} / z)^{\nu} = \sum_{j=1}^m \frac{\alpha_j}{z - z_j}$$

with

$$z_j = e^{\tau \lambda_j}, \quad j = 1, \dots, m.$$

These z_j lie inside the unit disk (and are nonzero), and hence the domain of convergence of the generating function extends to the exterior of some disk of radius $\rho < 1$ around the origin. It coincides with a rational function in the complex plane with a zero at infinity and m poles in the points z_j . And vice versa: If f is a rational function with m distinct first order poles $z_j \neq 0$ in the unit disk and corresponding residues α_j , and if f is vanishing at infinity then the associated Prony series φ of (3.1) with

$$\lambda_j = \frac{1}{\tau} \log z_j \quad (3.4)$$

interpolates the coefficients y_{ν} of the Laurent series (3.3) at all grid points $t_{\nu} = \nu \tau$, $\nu \in \mathbb{N}_0$.

We mention that the choice of the complex logarithm in (3.4) has no impact on the values of the Prony series at the grid points, as is obvious from (3.1). However, depending on which branch of the logarithm has been selected, the function φ may oscillate in between the grid points. Assuming that the sampling rate τ has been chosen small enough to capture all relevant frequencies of the target function y , the appropriate branch to choose in (3.4) is the one with imaginary parts in $(-\pi, \pi)$. Further, if some pole z_j happens to be negative, then this pole should be mapped onto two complex conjugate exponents $\lambda_j^{\pm} = (\log |z_j| \pm i\pi) / \tau$, with $\alpha_j^{\pm} = \alpha_j / 2$ as corresponding weights, to obtain a real-valued approximation (3.1); we emphasize that

the special case of negative poles within the unit disk calls for further attention at a later stage of the algorithm; see Remark 4.2 below.

The task (3.2) can therefore be recast as a problem of approximating the generating function f by a rational one, which vanishes at infinity, and then choosing φ to be the Prony series associated with this rational approximation. For example, one can compute the Padé approximation of f of order $(m-1, m)$ at infinity; the associated Prony series then interpolates the given data y_ν , $\nu = 0, \dots, 2n-1$. The resulting algorithm is the classical Prony method for exponential interpolation, cf., e.g., Weiss and McDonough [28], or Plonka et al [24]. This is the route that we have followed in [8]. The resulting algorithm works reasonably well for twenty or so data points, but for larger or noisy data sets the resulting Prony series often fails to be of positive type.[†] Then there is no associated Markov system (2.1) which satisfies (2.6). We briefly mention that Bai and Freund [5] have suggested an ad hoc modification of their Padé approximation scheme to obtain a Prony series of positive type, but they apply the Padé scheme in the frequency domain and not in the domain of the z -transform.

An alternative way of computing rational approximations of the generating function can be based on the AAA algorithm[‡]. To this end one can adapt a suggestion by Derevianko, Plonka, and Petz [12] and first evaluate the available $2n$ leading terms of (3.3) on an equidistant angular grid on the unit circle, e.g., by using a fast Fourier transform of order $2n$ of the samples $\{y_\nu\}$. This gives data

$$f_\nu \approx f(\omega^\nu), \quad \nu = 0, \dots, 2n-1, \quad \text{where } \omega = e^{i\pi/n}. \quad (3.5)$$

The AAA algorithm now proceeds by choosing an appropriate index set $\mathcal{N}_k \subset \{0, \dots, 2n-1\}$ associated with m_k of these grid points on the unit circle, defines a corresponding parameterized rational interpolant r_k of order (m_k-1, m_k-1) in barycentric form,

$$r_k(z) = \sum_{\nu \in \mathcal{N}_k} \frac{w_\nu f_\nu}{z - \omega^\nu} \bigg/ \sum_{\nu \in \mathcal{N}_k} \frac{w_\nu}{z - \omega^\nu}, \quad (3.6)$$

and optimizes the m_k free complex parameters w_ν to approximate the remaining data pairs (3.5) by solving the linear least squares problem

$$\text{minimize } \sum_{\mu \notin \mathcal{N}_k} \left| \sum_{\nu \in \mathcal{N}_k} \frac{w_\nu (f_\mu - f_\nu)}{\omega^\mu - \omega^\nu} \right|^2 \quad (3.7)$$

among all coefficients w_ν with $\sum_{\nu \in \mathcal{N}_k} |w_\nu|^2 = 1$. It follows from (3.7) that any vector $w \in \mathbb{C}^{m_k}$ with minimizing weights is a singular vector associated with the smallest singular value of the so-called (rectangular) Loewner matrix

$$L = \left[\frac{f_\mu - f_\nu}{\omega^\mu - \omega^\nu} \right]_{\mu, \nu} \in \mathbb{C}^{2n-m_k, m_k}.$$

[†]By some abuse of denomination we refer to a Prony series of positive type, if the Fourier transform of its even extension to all of \mathbb{R} is of positive type.

[‡]Spoken “triple-A-algorithm”; the acronym stands for adaptive Antoulas–Anderson (algorithm).

We like to mention in passing that Claus Schneider has contributed in an early paper [26] to the use of the barycentric representation for rational interpolation, and refer to [22] for a detailed historical account of the ideas behind the AAA algorithm.

The AAA algorithm is a greedy iterative scheme, i.e., in each iteration $k = 1, 2, \dots$, the set of interpolation points is extended until the overall match (3.7) is below a given tolerance. Presuming that the final rational function has only first order poles $z_j \neq 0$ within the unit disk, these poles and their residues yield as before an associated Prony series φ which satisfies (3.2).

In order to ensure the rational approximation to be real, i.e., that $r_k(z) \in \mathbb{R}$ for $z \in \mathbb{R}$, we initialize $\mathcal{N}_1 = \{0, n\}$ with the indices ν in (3.5) corresponding to $z = \pm 1$, and in each iteration we extend the index set \mathcal{N}_k by the two indices ν and $2n - \nu$ of one complex conjugate pair of grid points, hence $m_k = 2k$. Then it can be shown that in each iteration we obtain a real rational interpolant (3.6) which minimizes (3.7), provided we choose the weight vector $w \in \mathbb{C}^{m_k}$ appropriately. To be specific, assume that $\{\omega_\nu\}_{\nu \in \mathcal{N}_k}$ is *any* set of minimizing parameters for (3.7) computed by a singular value decomposition of the corresponding Loewner matrix. Assume further that $\nu, \nu' \in \mathcal{N}_k$ correspond to two complex conjugate grid points of the grid (3.5). Then it is not too difficult to see by reordering the summation in (3.7) that the minimal value of (3.7) is also attained if the weight w_ν is replaced by $\bar{w}_{\nu'}$, and $w_{\nu'}$ by \bar{w}_ν , respectively. Accordingly, the vector w' , which contains all these exchanged weights for all complex conjugate pairs of grid points in \mathcal{N}_k , must also be a singular vector of the same Loewner matrix associated with the same singular value, and this implies that $(w + w')/\|w + w'\|_2$ is yet another vector in this singular subspace; the entries of this latter vector, however, are complex conjugate, when the corresponding data points are, and hence the associated rational function (3.6) is real.

Remark 3.1. We have made the assumption that the system matrix A of (2.1) be a diagonalizable matrix. If this fails to be true, then some of the coefficients of the associated autocorrelation function φ may be polynomials $\alpha_j(t)$ instead of being constant. This in turn implies that the corresponding generating function f has poles of higher order. And vice versa: If the rational approximation (computed by whatever means) has poles of degree two or more, then this calls for a system matrix A with nontrivial Jordan blocks. \diamond

Although the AAA algorithm performs well in approximating the generating function, there is no clear interpretation of the sense in which the associated Prony series will approximate the data in (3.2), nor is there any guarantee that this Prony series will be a function of positive type. As we have emphasized before, the latter is mandatory to obtain the desired Markov model (2.1).

4 The new algorithm for the stochastic modeling problem

Recall that it is our objective to approximate the given data by a Prony series in a mean square sense, subject to the constraint that this function be of positive type. The major difficulty in achieving this goal stems from the fact that there is no characterization

(nor understanding) of the nonlinear manifold of data which correspond to Prony series of positive type. We therefore relax our aim in two ways:

- (i) we preassign the number m and the exponents λ_j , $j = 1, \dots, m$, of the exponentials to be used in (3.1);
- (ii) we optimize the remaining free coefficients α_j , $j = 1, \dots, m$, so as to minimize the mean square fit (3.2), subject to the weaker constraint that the Fourier transform of the Prony series is positive near infinity.

These goals are realized in two independent steps of our algorithm. Concerning item (i) we apply in a first step the variant of the AAA method described in the previous section. We take the poles z_j of the resulting rational approximation (which are symmetric with respect to the real axis), ignore the associated residuals – because the final weights of the Prony series φ are subject to item (ii) – and choose the exponents λ_j of φ via (3.4). As has been mentioned in Remark 3.1 higher order poles z_j lead to non-diagonalizable coefficient matrices A in (2.1); although it is very unlikely that this case will ever show up in numerical computations, we simply ignore the order of the poles for item (i). We further remark that some poles may fail to lie in the open unit disk. Those are called spurious poles, because they typically come with very tiny residuals and have small impact on the quality of the approximation; so we can simply eliminate them. The same is true for a possible pole at $z = 0$. In fact, in our numerical implementation we eliminate all poles whose residues are negligible relative to $y_0 = f(\infty)$. For the moment we also make the assumption that none of the remaining poles is negative (see Remark 4.2 below).

Item (ii) is motivated by the fact that in the applications that we have in mind the true autocorrelation function y belongs to $L^1(\mathbb{R})$, and hence, $\widehat{y}(\xi)$ converges to zero for $|\xi| \rightarrow \infty$. Accordingly, violations of the positivity of $\widehat{\varphi}$ are most likely to happen near infinity, in particular, as tiny spurious oscillations in the data due to noise will mostly affect high frequency components of φ . Further, the positivity of $\widehat{\varphi}$ near infinity can be checked with a finite number of linear constraints on the coefficients α_j , which simplifies their optimization. In fact, a straightforward computation shows that

$$\widehat{\varphi}(\xi) = -2 \sum_{j=1}^m \frac{\alpha_j \lambda_j}{\lambda_j^2 + \xi^2}, \quad (4.1)$$

and expanding the right-hand side for large frequencies ξ we obtain

$$\widehat{\varphi}(\xi) = 2 \sum_{k=0}^{\infty} \beta_k \xi^{-2k-2} \quad (4.2)$$

for $|\xi| > \max |\lambda_j|$, with

$$\beta_k = (-1)^{k+1} \sum_{j=1}^m \alpha_j \lambda_j^{2k+1}, \quad k \in \mathbb{N}_0. \quad (4.3)$$

It follows that $\widehat{\varphi}(\xi) \geq 0$ for all sufficiently large frequencies ξ , if and only if the leading term of (4.2) is positive near infinity, i.e., if condition

$$\beta_{k_*} > 0 \quad \text{and} \quad \beta_k = 0 \quad \text{for all } 0 \leq k < k_* \quad (4.4)$$

holds true for some $k_* \in \mathbb{N}_0$. As mentioned before, these are finitely many linear constraints on the searched for coefficients α_j .

Thus, to realize item (ii) we specify the weights α_j via the linear least squares problem

$$\left. \begin{aligned} & \text{minimize } \sum_{\nu=0}^{2n-1} \left| y_\nu - \sum_{j=1}^m \alpha_j e^{\nu\tau\lambda_j} \right|^2 \\ & \text{subject to the (in)equality constraints (4.4).} \end{aligned} \right\} \quad (4.5)$$

The solution of this problem constitutes the second step of our algorithm. Note that the quadratic objective function

$$\psi(a) = \sum_{\nu=0}^{2n-1} \left| y_\nu - \sum_{j=1}^m \alpha_j e^{\nu\tau\lambda_j} \right|^2, \quad a = [\alpha_1, \dots, \alpha_m]^T \in \mathbb{C}^m, \quad (4.6)$$

of (4.5) is strictly convex as long as we restrict ourselves to $m \leq 2n$; compare, for example, [24]. Since the constraints also define a convex set \mathcal{Q} of admissible parameters, problem (4.5) has a unique minimizer, provided that the constraints are consistent. The same argument as in the previous section can be used to show that the minimizing weights come in complex conjugate pairs, so that the corresponding Prony series is real-valued.

The solution of (4.5) can be found by implementing a loop over the number $k_* = 0, 1, 2, \dots$ of active constraints in (4.4). We initialize $k_* = 0$ and consider first the unconstrained least squares problem

$$\text{minimize } \psi(a) \quad (4.7)$$

for the given exponents $\lambda_j \in \mathbb{C}^-$, $j = 1, \dots, m$. If the solution $a^\dagger \in \mathbb{C}^m$ of (4.7) satisfies $\beta_0 = \beta_0(a^\dagger) > 0$ then we have found the unique solution of problem (4.5). Otherwise the solution vector a of (4.5) must satisfy the additional equality constraint

$$\beta_0(a) = - \sum_{j=1}^m \alpha_j \lambda_j = 0. \quad (4.8)$$

We thus let $k_* = 1$ in this case and minimize $\psi(a)$ subject to the equality constraint (4.8). This problem has a unique solution, which we denote again by $a^\dagger \in \mathbb{C}^m$ for simplicity; it can be determined by using the constraint to eliminate one of the free variables and then solving a single linear system of (normal) equations for the

remaining ones; compare, e.g., Björck [7, Section 5.1] for stable numerical implementations. If the corresponding value of $\beta_1 = \beta_1(a^\dagger) > 0$ then the parameter vector a^\dagger is the unique solution of (4.5); if not, we again increase k_* by one, i.e., add another active constraint via (4.4), repeat the optimization process anew, and so on.

This loop over k_* terminates whenever the solution of (4.5) has been found, or when the set of constraints becomes inconsistent, i.e., when $\beta_0 = \beta_1 = \dots = \beta_{k_*-1} = 0$ implies that $\beta_k = 0$ for all $k \geq k_*$. Due to the Vandermonde type structure of the constraints this happens exactly when $k_* = m$, i.e., when the number of active constraints has reached the number of free parameters. If this happens then the algorithm has failed.

Such a failure indicates that the Prony series approximation (3.1) does not have enough degrees of freedom or that the data are too noisy. To account for the former, one can enforce a smaller tolerance for (3.7) in the AAA algorithm in view of item (i) above, because this will typically give more poles, i.e., more terms for the Ansatz (3.1). To cope with too much noise in the data, one option is to use a larger grid spacing τ , or to come up with some individual sophisticated treatment to reduce the noise in the data.

We also emphasize that even when (4.5) has a solution, then this is no guarantee that the resulting Prony series is of positive type, for the Fourier transform may still be negative in some bounded frequency band. It remains open how to best proceed in such a situation.

Remark 4.1. Depending on the application it may be appropriate to append additional constraints to (4.5). For example, in many situations the variance $y(0)$ of the underlying stationary process Y can be measured more accurately than the other autocorrelations, or may be known beforehand from theoretical considerations. In this case it may be reasonable to enforce the approximation $\varphi \approx y$ to be exact at the origin, i.e., by adding the constraint

$$\varphi(0) = \sum_{j=1}^m \alpha_j = y_0. \quad (4.9)$$

In other applications, for example the one in Section 5, it may be known a priori that the true autocorrelation function is differentiable, in which case $y'(0)$ must vanish due to the symmetry of y . In this case one may wish to add the constraint

$$\varphi'(0) = \sum_{j=1}^m \alpha_j \lambda_j = 0, \quad (4.10)$$

which means that $\beta_0 = 0$, compare (4.3). Accordingly, (4.4) can only hold for some $k_* \geq 1$, and in this case we initialize $k_* = 1$ when executing the loop of our algorithm for the constrained least squares problem for the first time.

We will show in the appendix that the inhomogeneous constraint (4.9) and the equality constraints in (4.5) are consistent as long as $k_* < m$. \diamond

Remark 4.2. So far we have excluded the case that the rational approximation determined by the AAA algorithm has negative poles. Now let us assume that some

pole z_j happens to be negative. Then the corresponding two exponents

$$\lambda_j^\pm = \mu_j \pm i \frac{\pi}{\tau}, \quad \mu_j = \frac{1}{\tau} \log |z_j|,$$

differ by $2i\pi/\tau$, and the exponentials $e^{\nu\tau\lambda_j^\pm}$ in (4.5) are indistinguishable. Accordingly, the objective function ψ of (4.6) is convex, but not strictly convex. If (4.9) is the only equality constraint in (4.5) then there is an infinite family of Prony series which solves (4.5): If φ is one of them, then

$$\varphi(t) + \gamma e^{\mu_j t} \sin(\pi t/\tau), \quad t \geq 0,$$

with $\gamma \in \mathbb{R}$, $|\gamma|$ small, are further ones. By tuning γ , one can change $\varphi'(0)$ to any value one might like – provided the positivity of $\widehat{\varphi}$ is not being violated. Which of them is “best” (and in what sense), however, is not clear.

In our implementation we enforce this derivative to be zero in this case, i.e., we add the constraint (4.10). If the rational function has only one negative pole then this heals the problem: the objective ψ is strictly convex over the set \mathcal{Q} of admissible parameters defined by (4.4) with $k_* \geq 1$. If the rational function has $\ell \in \mathbb{N}$ negative poles, then a similar fix is possibly by constraining $k_* \geq \ell$ in (4.4); see Theorem A.4 in the appendix. \diamond

5 Numerical example: Motion of a colloid within a fluid

For our numerical results we focus on applications where the underlying stationary process $Y : \mathbb{R} \rightarrow \mathbb{R}$ satisfies a generalized Langevin equation

$$m \dot{Y}(t) = -m \int_{-\infty}^t \gamma(t - \tau) Y(\tau) d\tau + F(t), \quad (5.1)$$

with an even integral kernel $\gamma \in L^1(\mathbb{R})$ of positive type. This sort of equation is used, for example, in statistical physics as a coarse-grained model for the velocity Y of a macroparticle, a colloid say, dispersed in a fluid, where the so-called memory friction γ represents the interactions of the macroparticle with the constituents of the fluid. In (5.1) m is the mass of the macroparticle and the external forcing F is a random process, whose autocorrelation function is given by

$$\mathbb{E}[F(s+t)F(s)] = \frac{m}{\beta} \gamma(t), \quad t \in \mathbb{R},$$

for a system in thermodynamical equilibrium, where $\beta > 0$ is the inverse temperature, cf., e.g., Jung and Schmid [19]. Moreover, Y is a Gaussian process whenever the random forcing is.

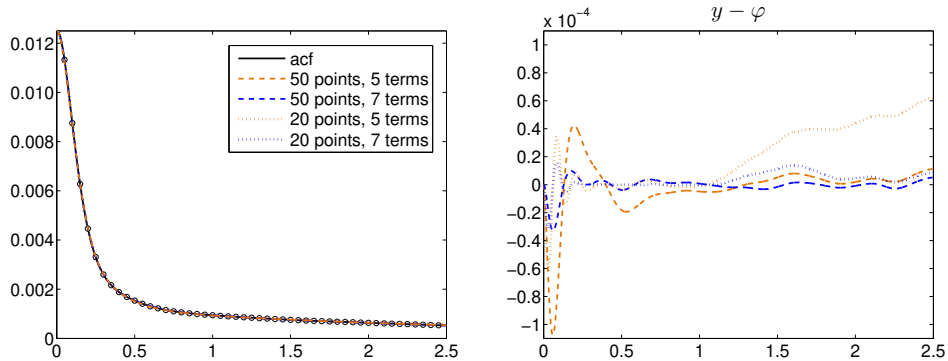


Fig. 1 Autocorrelation function (acf) of Examples 5.1 and 5.2 and its Prony series approximations (left panel) with five and seven terms for $2n = 50$ ($\tau = 0.05$) and $2n = 20$ ($\tau = 0.06$) data points, respectively. The panel on the right shows the corresponding errors, i.e., the graphs of $y - \varphi$.

The numerical implementation of (5.1) is involved and expensive, and therefore approximate Markov models (2.1) are often sought for simulations of the coarse-grained system. Typically autocorrelation data of Y are used to find such Markov approximations. Different techniques have been proposed for this in the literature, cf., e.g., [8], and the references given there. Most of these work well for a handful of auxiliary variables only and a corresponding limited number of data points. As will be demonstrated below, the new algorithm of Section 4 can handle much larger data sets, while the size of the resulting Markov models remains manageable.

For the generalized Langevin equation (5.1) it is not difficult to see that the autocorrelation function y of the stationary process Y satisfies the delay differential equation

$$\dot{y}(t) = - \int_0^t \gamma(t - \tau)y(\tau) d\tau, \quad y(0) = \frac{1}{\beta m}, \quad (5.2)$$

cf., e.g., [17]. Note that the variance $1/(\beta m)$ of the process corresponds to the average kinetic energy of a particle of mass m in thermodynamical equilibrium, and for physical consistency it suggests itself to prescribe $\varphi(0)$ to be exactly this value. Further, it follows readily from (5.2) that $\dot{y}(t) \rightarrow 0$ as $t \rightarrow 0$, and hence, the even autocorrelation function y is differentiable with $\dot{y}(0) = 0$. We therefore include the two additional constraints (4.9) and (4.10) of Remark 4.1 in our algorithm for the determination of the weights α_j of the Prony series (3.1).

The data for the numerical examples have been borrowed from our previous work [8, 18], where more details about the physical setting and the particular application can be found, as well as the technical details concerning the generation of these data. For the detection of spurious poles of the rational approximations in the first step of our algorithm we have used a tolerance of 10^{-3} for the absolute values of the residues of the poles, relative to the value of y_0 .

Example 5.1. As a first test case we consider the autocorrelation function from [18] shown as black solid line in the left panel of Figure 1. We assume that we are given $2n = 50$ samples of this autocorrelation function with grid spacing $\tau = 0.05$. These data points correspond to the black circles in the plot on the left hand side.

For this example the Padé approximation which is computed in the Prony method mentioned in Section 3 would have 25 poles, and if the associated Prony series is of positive type then the corresponding Markov system would have 24 auxiliary variables. Of course, some of these poles may have negligible residues. With a relative tolerance of 10^{-3} for the data fit (3.7) the AAA algorithm determines a rational approximation of the generating function with only five poles. The optimization of the corresponding weights α_j of the Prony series succeeds in the very first execution of the loop (with counter $k_* = 1$) to find the global solution of (4.5). One can check that the Fourier transform of the associated Prony series has no real zeros, so this Prony series approximation of the autocorrelation function is of positive type.

With the smaller tolerance $0.5 \cdot 10^{-3}$ for (3.7) the resulting AAA approximation has seven poles. This time the optimization of the corresponding weights requires more than one execution of the loop, because β_1 turns out to be negative initially. The loop terminates for $k_* = 2$ with the global solution of (4.5). Again the resulting Prony series is of positive type.

The graphs of the two approximations are displayed as dashed lines in Figure 1. In the panel on the left they can hardly be distinguished from the given autocorrelation function. Note that they correspond to autocorrelation functions of the first component X of a Markov system (2.1) with four, respectively six, auxiliary variables only. The panel on the right of Figure 1 reveals that the autocorrelation function of the approximation with four auxiliary variables is somewhat less accurate for small times.

We mention that with even smaller tolerances for (3.7) the AAA approximations develop spurious poles and the final mean square fit does not improve any further. \diamond

Example 5.2. To compare the outcome of the new method with the numerical results in [8, Section 3.2], we also consider the corresponding smaller data set with $2n = 20$ samples and slightly larger grid spacing $\tau = 0.06$ of the same autocorrelation function. Again, for a relative tolerance of 10^{-3} the rational function computed with the AAA algorithm has $m = 5$ poles, but this time one of these poles is negative and leads to two associated terms of the Prony series, see Remark 4.2. Since we restrict ourselves to $k_* \geq 1$ anyway, the constrained least squares problem (4.5) has a unique solution, which is, indeed, obtained for $k_* = 1$ as in Example 5.1. When the tolerance for (3.7) is 10^{-4} then the associated rational function has seven poles. Again, one of these poles is negative and the optimal weights are obtained for $k_* = 1$.

As in Example 5.1 both Prony series can be checked to be of positive type; their graphs are included in Figure 1 as dotted lines. Note that they use other interpolation data than the approximations in Example 5.1, with samples restricted to $t \leq 1.2$. As can be seen from the plot on the right the quality of the approximations of the corresponding Prony series deteriorate for larger times.

Due to the fact that one of the poles of the rational function has been negative, the corresponding two Markov systems have one auxiliary variable more than the ones of Example 5.1, namely five and seven, respectively. To compare the results with the ones of [8] we point out that the autocorrelation function constructed in [8] with the Prony type algorithm is of similar quality, but requires nine auxiliary variables. \diamond

Example 5.3. For a third example we consider an autocorrelation function from a different generalized Langevin equation for a similar physical problem. This example

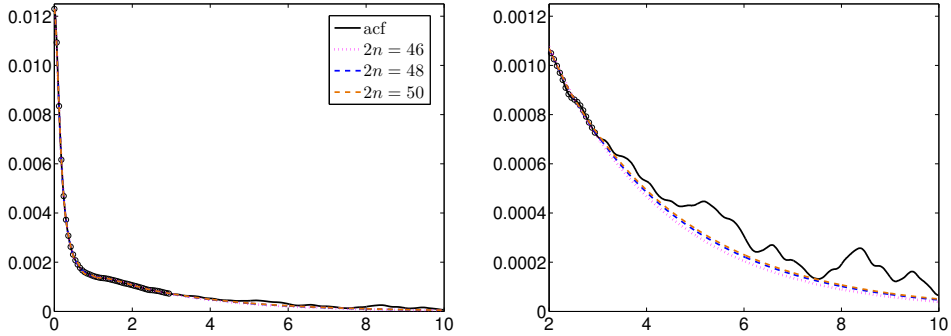


Fig. 2 Autocorrelation function (acf) of Example 5.3 and corresponding Prony series; right-hand panel shows a detail.

is taken from [8, Section 3.3], where it is termed “medium noise setting”; as the name suggests the corresponding data are somewhat noisy. With the same grid spacing $\tau = 0.06$ as in Example 5.2 the method of [8] failed to produce an approximation of positive type when the first $2n = 50$ data points were used, and also for $2n = 48$ data points. It was successful for $2n = 46$ data points, though, but despite the $m = 13$ terms of the approximating Prony series, the fit of the autocorrelation function wasn’t too good in that case, either; compare the middle panel of Figure 6 in [8].

With our new method and each of the three data sets the default tolerance of 10^{-3} for the fit (3.7) of the AAA algorithm is sufficient to obtain Markov models for which the autocorrelation functions in question are in good agreement with the given data. The graphs of two of them, namely the ones for $2n = 46$ and $2n = 50$ data points are displayed in Figure 2, together with the underlying autocorrelation function (black line) and the data samples (black circles). The associated Prony series have five terms ($2n = 46$) and nine terms ($2n = 50$), respectively; in both cases the rational approximations had two further poles, which were decided to be spurious. The two Prony series approximate the data equally well, both in terms of the mean square and absolute error. The fact that more poles are necessary for 50 data points may reflect the aforementioned difficulties of the Prony method for that situation.

The approximation shown in Figure 2 for $2n = 48$ grid points has been obtained for the slightly smaller tolerance $0.5 \cdot 10^{-3}$ for (3.7). Its data fit is about a factor of two better than for the other two models. In this case the AAA rational approximation had 13 poles, four of which were considered spurious, while two poles in the exterior of the unit disk just barely failed the criterion of being spurious poles. One may suspect that poles outside the unit circle may be more common when the data are noisy. Of course these two poles had to be eliminated to make the algorithm work, much to the success of the corresponding approximation, which thus had seven terms, eventually.

The plot on the right-hand side of Figure 2 zooms into a detail of the left-hand plot, in which the noise in the autocorrelation data is striking. Take note that the given data samples correspond to times $t < 3$. As can be seen from this plot the Prony series approximations provide a smoother and more reliable extrapolation of these data to times $t > 3$ as the measured values of the autocorrelation function do. This

In particular, for $t = 0$ this gives

$$v^T w = \varphi(0) = \sigma^2$$

by virtue of (2.6).

Now one can construct a nonsingular matrix $V \in \mathbb{R}^{m \times m}$ which satisfies

$$Vw = \sigma^2 e_1 \quad \text{and} \quad V^{-T}v = e_1. \quad (6.3)$$

(See, for example, Bai and Freund [4, Lemma 1] for an explicit construction.) With the help of (6.2) and (6.3) we thus arrive at the representation (2.6), namely

$$\varphi(t) = \sigma^2 v^T V^{-1} V e^{|t|J} V^{-1} e_1 = \sigma^2 (V^{-T}v)^T e^{|t|VJV^{-1}} e_1 = \sigma^2 e_1^T e^{|t|A} e_1$$

with the stable matrix

$$A = VJV^{-1}.$$

7 Determination of the driving noise direction

The final step of our construction of the Markov system (2.1) consists in finding the direction $g \in \mathbb{R}^m$ of the driving Brownian motion, so that the autocorrelation function of the stationary solution of (2.1) satisfies (2.6). For the ease of presentation we will make the stronger assumption that the computed Prony series φ is of strict positive type, i.e., that

$$\widehat{\varphi}(\xi) > 0 \quad \text{for all } \xi \in \mathbb{R}; \quad (7.1)$$

we refer to Anderson and Vongpanitlerd [3] for a treatment of the general case. We further assume that the coefficients α_j of (3.1) are all different from zero, and that φ can be written in the form (2.6) for some (stable) matrix $A \in \mathbb{R}^{m \times m}$, e.g., by choosing A as in Section 6. Then, in the systems theory terminology, the representation (2.9) of κ is a minimal realization of this transfer function. From (7.1) and (2.7) it follows that κ is analytic in a neighborhood of the closed right half complex plane with

$$\operatorname{Re} \kappa(i\xi) > 0 \quad \text{for all } \xi \in \mathbb{R}. \quad (7.2)$$

Given these properties the Positive Real Lemma states that the singular Lur'e equations

$$A\Sigma + \Sigma A^T = -gg^T, \quad \Sigma e_1 = \sigma^2 e_1, \quad (7.3)$$

have a solution – consisting of a symmetric positive definite matrix $\Sigma \in \mathbb{R}^{m \times m}$ and a vector $g \in \mathbb{R}^m$; cf. [3]. If the driving force of the Ornstein-Uhlenbeck equation (2.1) operates in the direction g then Σ is the covariance matrix of its stationary solution, compare (2.3); the second equation in (7.3) implies that Σ can be written in the form (2.5), and hence, the autocorrelation function of the stationary solution satisfies (2.6) as desired.

Note that we not only need g if we want to simulate the process X , but also Σ for the initial value

$$\begin{bmatrix} X(0) \\ Z(0) \end{bmatrix} \sim \mathcal{N}(0, \Sigma) \quad (7.4)$$

of the stationary solution of (2.1). To compute Σ and g , i.e., to solve the singular Lur'e equations numerically, we reduce them to a set of Lur'e equations of lower dimension, cf. Anderson [1]. Since this is an essential step of our algorithm we provide some more details, as those are rarely found in the pertinent literature. We represent A in the block form

$$A = \begin{bmatrix} -a_0 & c_0^T \\ -b_0 & A_0 \end{bmatrix} \quad (7.5)$$

with appropriate $A_0 \in \mathbb{R}^{(m-1) \times (m-1)}$, $b_0, c_0 \in \mathbb{R}^{m-1}$, and $a_0 \in \mathbb{R}$. It follows from (2.6) that

$$a_0 = -e_1^T A e_1 = -\frac{1}{\sigma^2} \lim_{t \rightarrow 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0,$$

because functions of positive type attain their maximum at the origin.

Inserting (7.5) and (2.5) into (7.3), we see that the driving force vector g has to have the form

$$g = \sigma \begin{bmatrix} \sqrt{2a_0} \\ g_0 \end{bmatrix}, \quad (7.6)$$

where $g_0 \in \mathbb{R}^{m-1}$ and $\Sigma_0 \in \mathbb{R}^{(m-1) \times (m-1)}$ satisfy the system of equations

$$A_0 \Sigma_0 + \Sigma_0 A_0^T = -g_0 g_0^T, \quad b_0 - \Sigma_0 c_0 = \sqrt{2a_0} g_0. \quad (7.7)$$

This is the general form of the Lur'e equations, with the singular case corresponding to $a_0 = 0$. Again, the Positive Real Lemma states that there is a vector g_0 and a symmetric positive definite matrix Σ_0 , which solve (7.7), if A_0 is stable and the associated transfer function

$$\kappa_0(\zeta) = a_0 + c_0^T (\zeta I - A_0)^{-1} b_0, \quad \zeta \in \mathbb{C}, \quad (7.8)$$

is positive real.

Note that $\zeta + \kappa_0(\zeta)$ is the Schur complement of $\zeta I - A_0$ in $\zeta I - A$, i.e., the reciprocal of $\zeta + \kappa_0(\zeta)$ is the (1, 1)-element of the inverse of $\zeta I - A$,

$$\kappa_0(\zeta) = \frac{1}{e_1^T (\zeta I - A)^{-1} e_1} - \zeta = \frac{\sigma^2}{\kappa(\zeta)} - \zeta, \quad (7.9)$$

where we have used (2.9) for the final step. Since κ is analytic in $\mathbb{C} \setminus \mathbb{C}^-$, its real part is harmonic, and it follows from (7.2) and the maximum principle that it is positive in the closed right half plane. This implies, that (i) κ_0 is analytic in $\mathbb{C} \setminus \mathbb{C}^-$, and (ii) its real part also satisfies (7.2). Accordingly, κ_0 is, indeed, positive real.

To see that the lower right block A_0 of A in (7.5) is stable, we assume to the contrary that μ is an eigenvalue of A_0 , which does not belong to \mathbb{C}^- . Let x_r and x_l be associated right and left eigenvectors of A_0 , respectively. Then it follows from

(7.5) that neither $c_0^T x_r$ nor $x_l^* b_0$ may vanish, for otherwise $[0; x_r]$ or $[0; x_l]$ would be right/left eigenvectors of A for the same eigenvalue; but this is not possible because A is stable. Therefore, since neither $c_0^T x_r$ nor $x_l^* b_0$ vanish, it follows from (7.8) that κ_0 has a pole at $\zeta = \mu$, but this is not possible as we have discussed above. Accordingly, all eigenvalues of A_0 belong to \mathbb{C}^- . We thus have verified that (7.7) has a solution.

If $a_0 > 0$ then a solution of the Lur'e system (7.7) can be obtained numerically by first solving the Riccati matrix equation

$$B_0 \Sigma_0 + \Sigma_0 B_0^T + \Sigma_0 c_0 c_0^T \Sigma_0 + b_0 b_0^T = 0$$

with

$$B_0 = 2a_0 A_0 - b_0 c_0^T$$

for Σ_0 , and then computing g_0 from the second equation in (7.7), cf. [3]. This is a well-studied numerical problem; compare, e.g., Bini, Iannazzo, and Meini [6].

On the other hand, when $a_0 = 0$, i.e., in the case of a singular Lur'e equation (7.7), then we first note that

$$\sigma_0^2 = c_0^T b_0 > 0. \quad (7.10)$$

In fact, it follows from (7.7) that in the singular case

$$c_0^T b_0 = c_0^T \Sigma_0 c_0 \geq 0,$$

and equality can only occur when $\Sigma_0 c_0 = 0$. But this means that $b_0 = 0$, and then A is singular according to (7.5), which would contradict the stability of A . Therefore, (7.10) is valid, and we can choose (as in Section 6) a similarity transformation $V \in \mathbb{R}^{(m-1) \times (m-1)}$ with

$$V b_0 = \sigma_0^2 e_1 \quad \text{and} \quad V^{-T} c_0 = e_1,$$

so that we can rewrite

$$\kappa_0(\zeta) = \sigma_0^2 e_1^T (\zeta I - V A_0 V^{-1})^{-1} e_1.$$

Replacing A by (the stable matrix) $V A_0 V^{-1}$, we can then proceed as above to reduce the solution of the singular Lur'e system (7.7) of dimension $m - 1$ to a Lur'e system of dimension $m - 2$, and so forth.

Take note that

$$\operatorname{Re} \kappa(i\xi) = -\sigma^2 e_1^T A (A + \xi^2 I)^{-1} e_1$$

is an even function of $\xi \in \mathbb{R}$, and it has an asymptotic expansion of the form

$$\operatorname{Re} \kappa(i\xi) = \sigma^2 (\gamma_k + o(1)) \xi^{-2k}, \quad |\xi| \rightarrow \infty, \quad (7.11)$$

for some $k \in \mathbb{N}$ and some $\gamma_k \neq 0$; as usual, we denote by $o(1)$ a term which converges to zero in the respective limiting process. It thus follows from (7.9) and (2.9) that

$$\operatorname{Re} \kappa_0(i\xi) = \sigma^2 \frac{\operatorname{Re} \kappa(i\xi)}{|\kappa(i\xi)|^2} = \frac{\sigma^4 (\gamma_k + o(1)) \xi^{-2k}}{\sigma^4 (1 + o(1)) \xi^{-2}} = (\gamma_k + o(1)) \xi^{-2(k-1)}.$$

Therefore, if $k > 1$ in (7.11) then we necessarily have $a_0 = 0$, compare (7.8), which means that the reduced Lur'e system (7.7) is also singular and another reduction step is necessary for its solution.

More precisely, (2.7), (2.9), and (4.2) imply that

$$\operatorname{Re} \kappa(i\xi) = \frac{1}{2} \widehat{\varphi}(\xi) = (\beta_{k_*} + o(1)) \xi^{-2(k_*+1)}$$

with $\beta_{k_*} > 0$, where k_* is the parameter occurring in the inequality constraint (4.4). In view of (7.11) we thus conclude that $k_* + 1$ reduction steps are necessary to turn the singular Lur'e equations (7.3) into a regular Lur'e system.

Appendix. Zeros of polynomials, consistency of constraints, and the issue of negative poles

This appendix deals with the consistency of the constrained least squares problems proposed in Section 4 in the two special situations encountered in Remarks 4.1 and 4.2.

As in Section 4 let $a = [\alpha_1, \dots, \alpha_m]^T \in \mathbb{C}^m$ be the vector with the free weight parameters, let $\beta_k = \beta_k(a)$ be defined by (4.3), and let $k_* \in \mathbb{N}$. We will first turn to Remark 4.1 and prove that the linear equality constraints

$$\beta_k(a) = 0, \quad k = 0, 1, \dots, k_* - 1, \quad (\text{A.1})$$

are consistent with the optional inhomogeneous constraint (4.9), as long as $m \geq k_* + 1$, i.e., as long as there are at least as many variables as there are equality constraints.

To this end we need the following auxiliary result on the zeros of certain real polynomials.

Theorem A.1. *Let $k \in \mathbb{N}$ and*

$$p(z) = \pi_0 + \pi_1 z + \pi_3 z^3 + \dots + \pi_{2k-1} z^{2k-1} \quad (\text{A.2})$$

be a real polynomial with $\pi_0 \neq 0$. Then p has at least $k-1$ zeros (counting multiplicities) in \mathbb{C}^- and in \mathbb{C}^+ , respectively.

Proof. Without loss of generality we may assume that $\pi_0 > 0$. We define the family of polynomials

$$p(z; \gamma) = \gamma + iq(z/i), \quad (\text{A.3})$$

where $\gamma \in \mathbb{R}$, and q is a fixed real and odd polynomial of degree $2k - 1$. Obviously, one can rewrite p of (A.2) in this form with $\gamma = \pi_0$ and

$$q(\xi) = \pi_1 \xi - \pi_3 \xi^3 + \dots + (-1)^{k+1} \pi_{2k-1} \xi^{2k-1}.$$

Note that for $\gamma \neq 0$ the polynomial $p(\cdot; \gamma)$ cannot have any zeros on the imaginary axis, because

$$\operatorname{Re} p(i\xi; \gamma) = \operatorname{Re} (\gamma + iq(\xi)) = \gamma \quad \text{for every } \xi \in \mathbb{R}.$$

Consider first the case that q has only simple zeros on the real axis. One of these zeros is $\xi_0 = 0$; then there may be further positive zeros $\xi_1 < \xi_2 < \dots < \xi_r$ for some $0 \leq r < k$, and the corresponding negative ones, $-\xi_1, \dots, -\xi_r$. Accordingly, the odd polynomial $p(\cdot; 0)$ has $2r + 1$ zeros on the imaginary axis, and $k - r - 1$ zeros in each of the two half planes \mathbb{C}^- and \mathbb{C}^+ , respectively. Since the real zeros of q are assumed to be simple, the graph of $\xi \mapsto q(\xi)$, $\xi \in \mathbb{R}$, changes sign at each of them, i.e., the local Taylor expansion of q near ξ_j has the form

$$q(\xi) = (\eta_j + o(1))(\xi - \xi_j), \quad \xi \rightarrow \xi_j, \quad (\text{A.4})$$

with

$$(-1)^j \eta_j \eta_0 > 0, \quad j = -r, \dots, r. \quad (\text{A.5})$$

Now we consider the homotopy that turns $p(\cdot; 0)$ into $p(\cdot; \pi_0)$ by increasing the parameter γ from zero to π_0 . Since the zeros of $p(\cdot; \gamma)$ depend continuously on γ , we conclude that for small enough γ the $k - r - 1$ zeros of $p(\cdot; 0)$ in \mathbb{C}^- stay in \mathbb{C}^- , and the $k - r - 1$ zeros in \mathbb{C}^+ stay in \mathbb{C}^+ . Further, for γ sufficiently small, the remaining zeros $z_j = z_j(\gamma)$, $j = -r, \dots, r$, of $p(\cdot; \gamma)$ can be found near the purely imaginary zeros $i\xi_j$ of $p(\cdot; 0)$: by virtue of (A.3) and (A.4) they satisfy

$$0 = \gamma + (i\eta_j + o(1))\left(\frac{z_j}{i} - \xi_j\right) = \gamma + (\eta_j + o(1))(z_j - i\xi_j),$$

which gives

$$z_j(\gamma) = i\xi_j - \frac{\gamma}{\eta_j + o(1)} = i\xi_j - \frac{\gamma}{\eta_j}(1 + o(1)), \quad \gamma \rightarrow 0. \quad (\text{A.6})$$

From (A.6) and (A.5) we conclude that the sign of

$$\operatorname{Re} z_j(\gamma) = -\frac{\gamma}{\eta_j}(1 + o(1))$$

depends only on the parity of j (and the sign of η_0 , of course), i.e., $z_{\pm 1}(\gamma), z_{\pm 3}(\gamma), \dots$ and $z_0(\gamma), z_{\pm 2}(\gamma), z_{\pm 4}(\gamma), \dots$ move into opposite half planes as γ becomes positive. Accordingly, for small positive values of γ at least $k - 1$ zeros of $p(\cdot; \gamma)$ belong to \mathbb{C}^+ and to \mathbb{C}^- , respectively. Since $p(\cdot; \gamma)$ is nonzero on the imaginary axis for $\gamma \neq 0$, these zeros stay in \mathbb{C}^+ , resp. \mathbb{C}^- , for all positive values of γ . We thus have proved the assertion for any polynomial p of the form (A.3), when the real zeros of q are all simple.

In the case that some real zeros of q happen to have higher multiplicities, we can construct a sequence (q_ν) of real and odd polynomials of degree $2k - 1$, which have only simple real zeros, and whose coefficients converge to those of q as $\nu \rightarrow \infty$. As we have shown in the first part of this proof every polynomial

$$p_\nu(z; \pi_0) = \pi_0 + iq_\nu(z/i)$$

has at least $k - 1$ zeros in \mathbb{C}^+ and those converge to zeros of $p(\cdot; \pi_0)$ as $\nu \rightarrow \infty$. Since none of them can come to lie on the imaginary axis, they all stay in \mathbb{C}^+ . This shows

that $p(\cdot; \pi_0)$ has at least $k - 1$ zeros (counting multiplicities) in \mathbb{C}^+ , and the same argument applies to \mathbb{C}^- . This concludes the proof. \square

Corollary A.2. *Let $m > k_* \geq 1$ and $\lambda_1, \dots, \lambda_m \in \mathbb{C}^-$. Then the constraints (A.1) and (4.9) are consistent.*

Proof. We rewrite the constraints in matrix form

$$Ba = y_0 e_1, \quad (\text{A.7})$$

where $a = [\alpha_1, \dots, \alpha_m]^T \in \mathbb{C}^m$ as before and

$$B = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \lambda_1^3 & \lambda_2^3 & \dots & \lambda_m^3 \\ \vdots & & & \vdots \\ \lambda_1^{2k_*-1} & \lambda_2^{2k_*-1} & \dots & \lambda_m^{2k_*-1} \end{bmatrix} \quad (\text{A.8})$$

is a complex matrix with $k_* + 1 \leq m$ rows and m columns. Accordingly, the constraints are consistent, if B has full row rank, i.e., if the null space of B^* is trivial.

Note that we can add and subtract complex conjugate columns of B to transform B into a real matrix of the same rank. This means that if B should fail to have full row rank then there is a nontrivial *real* vector $\pi = [\pi_0, \pi_1, \pi_3, \dots, \pi_{2k_*-1}]^T \in \mathbb{R}^{k_*+1}$, such that $B^T \pi = 0$. The latter is the case, if and only if the associated real polynomial p of degree $2k_* - 1$ of (A.2) is vanishing for $z = \lambda_1, \dots, \lambda_m$. This means that p has (at least) m roots in \mathbb{C}^- .

Now we distinguish two cases. If $\pi_0 = 0$ then p is odd, and for each zero in \mathbb{C}^- there is a corresponding one in \mathbb{C}^+ . This would imply that p has $2m \geq 2k_* + 2$ zeros (at least). On the other hand, if $\pi_0 \neq 0$ then Theorem A.1 states that p has $k_* - 1$ zeros (at least) in \mathbb{C}^+ ; accordingly, p has at least $k_* - 1 + m \geq 2k_*$ zeros in this case. Since p has degree $2k_* - 1$ this shows that p must be the zero polynomial in either case, and hence, all its coefficients are zero, i.e., $\pi = 0$. This is a contradiction, which shows that the null space of B^* is trivial, and hence, B has full row rank. \square

Remark A.3. Note that we have made the assumption that $y_0 = \mathbb{E}[Y^2] > 0$. In this case the system (A.1) and (4.9) of constraints becomes inconsistent, whenever $k_* \geq m$. To see this it remains to establish the inconsistency of the first $m + 1$ constraints. We use the arguments and the notation from the proof of Corollary A.2. For the system of the first $m + 1$ constraints the matrix B of (A.8) has one more row than columns, and its first m rows have full rank according to Corollary A.2. It follows that the null space of B^T is one-dimensional. Let $\pi \in \mathbb{R}^{m+1}$ be a nontrivial vector from this null space and p of (A.2) be the corresponding polynomial of degree $2m - 1$. Then p has m zeros $\lambda_1, \dots, \lambda_m \in \mathbb{C}^-$. Assume now that $\pi_0 = 0$. Then p is odd, and hence, it has m further zeros in \mathbb{C}^+ . This makes $2m$ zeros, which exceed the degree of p . Hence, if $\pi_0 = 0$ then p must be the zero polynomial which gives a contradiction.

Accordingly, $B^* \pi = \overline{B^T} \pi = 0$, i.e., the null space of B^* contains a vector π , whose first entry is nonzero. But then the right-hand side of (A.7) cannot belong to the range

space of B , because e_1 is not orthogonal to the null space of B^* . This shows that the system (A.7) is inconsistent, whenever the number of constraints exceeds the number of free variables. \diamond

The final result of this appendix deals with Remark 4.2 and considers the case that the rational function determined by the AAA algorithm has $\ell \in \mathbb{N}$ (distinct) negative poles.

Theorem A.4. *Let \mathcal{Q} be the set of parameter vectors $a = [\alpha_1, \dots, \alpha_m]^T \in \mathbb{C}^m$ which satisfy (4.4) with the additional constraint that $k_* \geq \ell$. Then ψ of (4.6) is strictly convex over \mathcal{Q} .*

Proof. Without loss of generality let us assume that

$$\lambda_{2l-1} = \mu_l + i\frac{\pi}{\tau} \quad \text{and} \quad \lambda_{2l} = \mu_l - i\frac{\pi}{\tau}, \quad (\text{A.9})$$

$l = 1, \dots, \ell$, are the exponents of the Prony series associated with the negative poles. Since all poles are assumed to be simple the (negative) μ_l are pairwise different.

We rewrite

$$\psi(a) = \|b - Ga\|_2^2 \quad \text{with} \quad G = [e^{\nu\tau\lambda_j}]_{\nu,j} \in \mathbb{R}^{2n \times m} \quad \text{and} \quad b = [y_\nu]_\nu \in \mathbb{R}^{2n}.$$

Then G has a nontrivial null space, spanned by the vectors $e_{2l} - e_{2l-1}$, $l = 1, \dots, \ell$, where e_j , $j = 1, \dots, m$, are the Cartesian basis vectors in \mathbb{C}^m . So, if $a, a' \in \mathcal{Q}$ are two solutions of (4.5), then

$$a - a' = \sum_{l=1}^{\ell} \gamma_l (e_{2l} - e_{2l-1}) \quad (\text{A.10})$$

for certain $\gamma_l \in \mathbb{C}$. By assumption, a and a' both satisfy the equality constraints (A.1) for $k = 0, \dots, \ell - 1$, and hence, $[\gamma_1, \dots, \gamma_\ell]^T$ belongs to the null space of

$$\begin{bmatrix} \lambda_1 - \overline{\lambda_1} & \cdots & \lambda_{2l-1} - \overline{\lambda_{2l-1}} \\ \lambda_1^3 - \overline{\lambda_1^3} & \cdots & \lambda_{2l-1}^3 - \overline{\lambda_{2l-1}^3} \\ \vdots & & \vdots \\ \lambda_1^{2\ell-1} - \overline{\lambda_1^{2\ell-1}} & \cdots & \lambda_{2l-1}^{2\ell-1} - \overline{\lambda_{2l-1}^{2\ell-1}} \end{bmatrix} = 2i \begin{bmatrix} \text{Im } \lambda_1 & \cdots & \text{Im } \lambda_{2l-1} \\ \text{Im } \lambda_1^3 & \cdots & \text{Im } \lambda_{2l-1}^3 \\ \vdots & & \vdots \\ \text{Im } (\lambda_1^{2\ell-1}) & \cdots & \text{Im } (\lambda_{2l-1}^{2\ell-1}) \end{bmatrix}.$$

Using (A.9) we obtain

$$\begin{aligned} \text{Im } (\lambda_{2l-1}^{2k+1}) &= \left(\mu_l^2 - \frac{\pi^2}{\tau^2} \right) \text{Im } (\lambda_{2l-1}^{2k-1}) + 2\mu_l \frac{\pi}{\tau} \text{Re } (\lambda_{2l-1}^{2k-1}), \\ \text{Re } (\lambda_{2l-1}^{2k+1}) &= \left(\mu_l^2 - \frac{\pi^2}{\tau^2} \right) \text{Re } (\lambda_{2l-1}^{2k-1}) - 2\mu_l \frac{\pi}{\tau} \text{Im } (\lambda_{2l-1}^{2k-1}), \end{aligned}$$

for every $l = 1, \dots, \ell$ and every $k \geq 1$. It thus follows by induction that there exists a lower triangular matrix L with nonzero entries on the diagonal such that

$$\begin{bmatrix} \operatorname{Im} \lambda_1 & \cdots & \operatorname{Im} \lambda_{2\ell-1} \\ \operatorname{Im} \lambda_1^3 & \cdots & \operatorname{Im} \lambda_{2\ell-1}^3 \\ \vdots & & \vdots \\ \operatorname{Im} (\lambda_1^{2\ell-1}) & \cdots & \operatorname{Im} (\lambda_{2\ell-1}^{2\ell-1}) \end{bmatrix} = L \begin{bmatrix} 1 & \cdots & 1 \\ \mu_1^2 & \cdots & \mu_\ell^2 \\ \vdots & & \vdots \\ \mu_1^{2(\ell-1)} & \cdots & \mu_\ell^{2(\ell-1)} \end{bmatrix}.$$

The matrix on the right-hand side is the square Vandermonde matrix of $\mu_1^2, \dots, \mu_\ell^2$, which is nonsingular because the μ_i^2 are pairwise different, since all μ_i are negative. Therefore the null space of the matrix on the left-hand side is trivial, and hence, the coefficients $\gamma_1, \dots, \gamma_\ell$ are all equal to zero, showing that $a = a'$ in (A.10). Therefore, ψ is strictly convex over \mathcal{Q} . \square

Acknowledgements

The author likes to thank Peter Benner and Volker Mehrmann for valuable pointers to the literature.

Declarations

Code availability. MATLAB code and data for the numerical examples in Section 5 are available at <https://mathematik.uni-mainz.de/hanke-aaa>.

Funding. The research leading to this work has been done within the Collaborative Research Center TRR 146; corresponding funding by the DFG is gratefully acknowledged.

Conflict of interests. The author declares no competing interests.

References

1. Anderson, B.D.O.: An algebraic solution to the spectral factorization problem. *IEEE Trans. Automat. Control* **12**, 410–414 (1967)
2. Anderson, B.D.O.: The inverse problem of stationary covariance generation. *J. Stat. Phys.* **1**, 133–147 (1969)
3. Anderson, B.D.O., Vongpanitlerd, S.: *Network Analysis and Synthesis: A modern Systems Theory Approach*. Prentice-Hall, Englewood Cliffs, NJ (1973)
4. Bai, Z., Freund, R.W.: Eigenvalue-based characterization and test for positive realness of scalar transfer functions. *IEEE Trans. Automat. Control* **45**, 2396–2402 (2000)

5. Bai, Z., Freund, R.W.: A partial Padé-via-Lanczos method for reduced-order modeling. *Linear Algebra Appl.* **332-334**, 139–164 (2001)
6. Bini, D.A., Iannazzo, B., Meini, B.: *Numerical Solution of Algebraic Riccati Equations*. SIAM, Philadelphia (2012)
7. Björck, Å.: *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia (1996)
8. Bockius, N., Shea, J., Jung, G., Schmid, F., Hanke, M.: Model reduction techniques for the computation of extended Markov parameterizations for generalized Langevin equations. *J. Phys.: Condens. Matter* **33**, 214003 (2021)
9. Brüll, T., Schröder, C.: Dissipativity enforcement via perturbation of para-Hermitian pencils. *IEEE Trans. Circuits Syst. I. Regul. Pap.* **60**, 164–177 (2012)
10. Cherifi, K., Goyal, P., Benner, P.: A non-intrusive method to inferring linear port-Hamiltonian realizations using time-domain data. *Electron. Trans. Numer. Anal.* **56**, 102–116 (2022)
11. Coelho, C.P., Phillips, J., Silveira, L.M.: Convex programming approach for generating guaranteed passive approximations to tabulated frequency-data. *IEEE Trans. Computer-Aided Design* **23**, 293–301 (2004)
12. Derevianko, N., Plonka, G., Petz, M.: From ESPRIT to ESPIRA: estimation of signal parameters by iterative rational approximation. *IMA J. Numer. Anal.* (2022), <https://doi.org/10.1093/imanum/drab108>
13. Fazzi, A., Guglielmi, N., Lubich, C.: Finding the nearest passive or nonpassive system via Hamiltonian eigenvalue optimization. *SIAM J. Matrix Anal. Appl.* **42**, 1553–1580 (2021)
14. Freund, R.W., Jarre, F., Vogelbusch, C.H.: Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling. *Math. Program.* **109**, 581–611 (2007)
15. Gillis, N., Sharma, P.: Finding the nearest positive-real system. *SIAM J. Numer. Anal.* **56**, 1022–1047 (2018)
16. Grivet-Talocia, S.: Passivity enforcement via perturbation of Hamiltonian matrices. *IEEE Trans. Circuits Syst. I. Regul. Pap.* **51**, 1755–1749 (2004)
17. Hanke, M.: Mathematical analysis of some iterative methods for the reconstruction of memory kernels. *Electron. Trans. Numer. Anal.* **54**, 483–498 (2021)
18. Jung, G., Hanke, M., Schmid, F.: Iterative reconstruction of memory kernels. *J. Chem. Theory Comput.* **13**, 2481–2488 (2017)

19. Jung, G., Schmid, F.: Fluctuation-dissipation relations far from equilibrium: a case study. *Soft Matter* **17**, 6413–6425 (2021)
20. Kalman, R.E.: Linear stochastic filtering theory–reappraisal and outlook. In: *Proceedings of the Symposium on System Theory*, New York, 1965, Polytechnic Press, Brooklyn, pp. 197–205 (1965)
21. Lindquist, A., Picci, G.: *Linear Stochastic Systems*. Springer, Heidelberg (2015)
22. Nakatsukasa, Y., Sète, O., Trefethen, L.N.: The AAA algorithm for rational approximation. *SIAM J. Sci. Comput.* **40**, A1494–A1522 (2018)
23. Pavliotis, G.A.: *Stochastic Processes and Applications. Diffusion Processes, the Fokker-Planck and Langevin Equations*. Springer, New York (2014)
24. Plonka, G., Potts, D., Steidl, G., Tasche, M.: *Numerical Fourier Analysis*. Springer, Cham (2018)
25. Saraswat, D., Achar, R., Nakhla, M.: Enforcing passivity for rational function based macromodels of tabulated data. In: *Electrical Performance of Electrical Packaging (IEEE Cat. No. 03TH8710)*, IEEE, pp. 295–298 (2003)
26. Schneider, C., Werner, W.: Some new aspects of rational interpolation. *Math. Comp.* **47**, 285–299 (1986)
27. Varah, J.M.: On fitting exponentials by nonlinear least squares. *SIAM J. Sci. Comput.* **6**, 30–44 (1985)
28. Weiss, L., McDonough, R.N.: Prony’s method, z -transforms, and Padé approximation. *SIAM Rev.* **5**, 145–149 (1963)
29. Zwanzig, R.: *Nonequilibrium Statistical Mechanics*. Oxford University Press, Oxford (2001)