

Steady state stability preserving nonlinear model reduction using sequential convex optimization

Martin Löhning, Jan Hasenauer, and Frank Allgöwer

Abstract—Models of dynamical systems become increasingly complex. While this allows a more accurate description of the underlying process, it often renders the application of model-based control algorithms infeasible. In this paper, we propose a model reduction procedure for systems described by nonlinear ordinary differential equations. The reduced model used to approximate the input-output map of the system is parameterized via the observability normal form. To preserve the steady states of the system and their stability properties, the set of feasible parameters of the reduced model has to be constrained. Therefore, we derive necessary and sufficient conditions for simultaneous exponential stability of a set of steady states of the nonlinear reduced model. The local approximation of these constraints results in a sequential convex program for computing the optimal parameters. The proposed approach is evaluated using the Fermi-Pasta-Ulam model.

I. INTRODUCTION

Models of electrical, mechanical, (bio-)chemical, and biological processes become more and more detailed to achieve a more accurate description of the underlying mechanisms and process dynamics. Unfortunately, this complicates controller design, as most of the model-based control algorithms, e.g., flatness based approaches, only allow for the consideration of small and medium size systems. Therefore, model reduction for linear and nonlinear systems has become more and more important over the last decade. While reducing the model size and complexity, it is often crucial to preserve certain properties, such as stability.

Most of the research in the field of model reduction focuses on linear systems [2]. Common approaches are balanced truncation [3] and Krylov methods [4]. Balanced truncation has also been extended to nonlinear systems [5, 6] employing the concept of energy functions, respectively empirical gramians. To preserve or guarantee the stability during model reduction, approaches exploiting, e.g., proper orthogonal decomposition [7] and piecewise-linear approximation [8] have been developed.

Another approach is the trajectory-based model reduction procedure proposed in [9] and [10]. In [10], the reduced model is given in observability normal form, thereby implicitly defining a nonlinear mapping from the detailed to

The authors thank Christian Ebenbauer for drawing their attention to [1]. Furthermore, the authors thank Marcus Reble and Christian Breindl for helpful comments.

The authors acknowledge financial support from the German Federal Ministry of Education and Research (BMBF) within the FORSYS-Partner program (grant nr. 0315-280A), and from the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

The authors are with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, Stuttgart, Germany, {martin.loehning, jan.hasenauer, frank.allgower}@ist.uni-stuttgart.de.

the reduced model states. Then, by using system identification to determine the parameters of the observability normal form, a convex formulation for the model reduction problem is derived. Unfortunately, the last method does not admit the explicit consideration of location and stability of steady states. However, preserving these qualitative aspects is important to ensure the correct asymptotic behavior. In this work, we extend the approach of [10] to allow for the integration of knowledge about location and stability of steady states, by introducing a necessary and sufficient condition for the simultaneous stability of a predefined set of steady states. As the resulting problem is nonconvex, it is relaxed to a sequential convex optimization problem. This ensures efficient computations which facilitate the reduction of high-dimensional systems.

The paper is structured as follows: In Section II we describe the problem setup. Subsequently, the steady state stability preserving model reduction procedure is presented in Section III. In Section IV we apply the procedure to the Fermi-Pasta-Ulam lattice. The paper is concluded in Section V.

Mathematical notation: The space of real symmetric $n \times n$ matrices is denoted as \mathbb{S}^n . The nonnegative real numbers are given by \mathbb{R}_+ . The notation I and 0 represents the identity and zero matrix, and $\text{tr}(X)$ is the trace of $X \in \mathbb{R}^{n \times n}$. The positive definiteness and semidefiniteness of a symmetric matrix $X \in \mathbb{S}^n$ is denoted $X \succ 0$ and $X \succeq 0$, respectively. When writing $u(\cdot)$ we refer to the trajectory of u . The first and n -th derivative of u with respect to time is represented by \dot{u} and $\overset{(n)}{u}$, respectively. The notation $u^{(i)}$ is used to number the elements u within a set $\{u^{(i)}\}_{i=1}^N$.

II. PROBLEM STATEMENT

In this work, we consider the problem of model reduction for a system of nonlinear ODEs. The system to be reduced will be called *detailed model* in the following and abbreviated using the subscript D . For notational simplicity the detailed model is assumed to be a SISO system,

$$\Sigma_D : \begin{cases} \dot{x} = f_D(x, u), & x(0) = x_0 \\ y_D = h_D(x, u), \end{cases}$$

in which $x \in \mathbb{R}^n$ is the state of the system, $u \in \mathbb{R}$ is the input, and $y_D \in \mathbb{R}$ is the output. The initial condition is denoted by $x_0 \in \mathbb{R}^n$. To ensure existence and uniqueness of solutions the vector field $f_D : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ is assumed to be globally Lipschitz continuous. Additionally, the mapping $h_D : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be sufficiently smooth.

While performing model reduction the main goal is to preserve the key features of the detailed model. Two of these key properties are the location and the stability of steady states. Therefore, we assume that the set of locally exponentially stable steady states of Σ_D , $\mathcal{S} = \{(\bar{u}^{(s)}, \bar{y}^{(s)})\}_{s=1}^{n_s}$, is available. Note that, as I/O-based model reduction is considered, we only specify steady states via a constant output $\bar{y}^{(s)}$. Thus, we allow for internal dynamics, e.g., due to limit cycle oscillations, which are not observable at the output.

Given the detailed model Σ_D , a reduced model Σ_R shall be computed which (i) reflects the location and exponential stability of the steady states and (ii) provides a good approximation of the I/O-behavior of Σ_D on a finite time interval $[0, T_{\text{end}}]$. To compute a reduced model achieving (ii), we use the model reduction procedure presented in [10] and extend it to enforce (i). In the remainder of this section we provide a short summary of the methods presented in [10].

To allow for the development of problem-specific and thus smaller reduced models, we allow for the integration of prior knowledge about relevant initial conditions x_0 and input trajectories $u(\cdot)$. Therefore, importance weights for x_0 and $u(\cdot)$ are introduced and denoted by p_x and p_φ , respectively. For ease of presentation, the input trajectories are assumed to be parameterized by $\varphi \in \mathbb{R}^{n_\varphi}$, e.g., a finite Fourier series. This allows us to express the importance weight p_φ as a function of the parameters φ . In the following the parametrized input is denoted by u_φ . For the weighting functions $p_x: \mathbb{R}^n \rightarrow \mathbb{R}_+$ and $p_\varphi: \mathbb{R}^{n_\varphi} \rightarrow \mathbb{R}_+$ we assume, that the integrals $\int_{\mathbb{R}^n} p_x(x_0) dx_0$ and $\int_{\mathbb{R}^{n_\varphi}} p_\varphi(\varphi) d\varphi$ exist and are equal to one.

To facilitate an efficient model reduction procedure, in the first step the reduced model is parameterized. The reduced model is chosen to have observability normal form [11],

$$\Sigma_R: \begin{cases} \dot{\xi}_i = \xi_{i+1}, & i = 1, \dots, n_R - 1 \\ \dot{\xi}_{n_R} = \theta^T m(\xi, \mu), & \xi_k(0) = y_D^{(k-1)}(0; x_0, u_\varphi), \\ & k = 1, \dots, n_R \\ y_R = \xi_1, \end{cases} \quad (1)$$

with state $\xi \in \mathbb{R}^{n_R}$ and state dimension n_R . The augmented input vector is denoted by $\mu = [u_\varphi, \dot{u}_\varphi, \dots, u_\varphi]^{n_R T}$. Therefore, we allow only n_R times differentiable inputs $u_\varphi(\cdot)$. The right hand side of $\dot{\xi}_{n_R}$ is a weighted sum of ansatz functions, $\theta^T m(\xi, \mu)$. We denote the weighting parameters by $\theta \in \mathbb{R}^{n_\theta}$ and the ansatz functions by $m(\xi, \mu) = [m_1(\xi, \mu), \dots, m_{n_\theta}(\xi, \mu)]^T$. The ansatz functions $m_i(\xi, \mu)$, $i = 1, \dots, n_\theta$, are assumed to be Lipschitz continuous. Common ansatz functions are polynomials or radial basis functions.

The main advantage of the observability normal form is that the states of the reduced model are the time derivatives of the output, $\xi = [y_R, \dot{y}_R, \dots, y_R^{(n_R-1)}]^T$. By using the observability normal form, we exploit the fact that model reduction – and not classical parameter estimation from measurement data – is performed, i.e., the time derivatives of the output are available. Furthermore, this is beneficial as

a direct link between the states ξ of the reduced model and the output of the detailed model y_D is established, which, e.g., makes $\xi_k(0) = y_D^{(k-1)}(0; x_0, u_\varphi)$, $k = 1, \dots, n_R$, a reasonable choice for the mapping of the initial states. The time derivatives of the output $y_D^{(k-1)}(0; x_0, u_\varphi)$ can be computed using the observability map given in [11].

Given the model structure (1), in the second step the parameters θ have to be estimated. For a given x_0 and $u_\varphi(\cdot)$, we thereby use the integrated squared equation error

$$E(x_0, u_\varphi, \theta) = \frac{1}{T_{\text{end}}} \int_0^{T_{\text{end}}} \left(y_D^{(n_R)}(t; x_0, u_\varphi) - \theta^T m(\xi_D, \mu) \right)^2 dt,$$

as error criterion, with $\xi_D = [y_D, \dot{y}_D, \dots, y_D^{(n_R-1)}]^T$. This error criterion is comparable to the 1-step-ahead prediction error in discrete-time systems identification. It penalizes differences in the change of the output of the reduced model y_R and of the detailed model y_D . As the reduced system is given in observability normal form, only the error in the n_R -th time derivative of the output has to be considered.

Based on this setup, the considered problem is:

Problem 1 *Given a detailed model Σ_D and the relevant I/O-behavior defined by the weightings $p_x(x_0)$ and $p_\varphi(\varphi)$, compute a reduced model Σ_R in observability normal form of fixed order n_R and with ansatz functions m , such that*

- 1) *the objective functional*

$$J(\theta) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n_\varphi}} E(x_0, u_\varphi(\cdot), \theta) p_x(x_0) p_\varphi(\varphi) d\varphi dx_0$$

is minimized and

- 2) *Σ_R has locally exponentially stable steady states \mathcal{S} .*

III. MODEL REDUCTION PROCEDURE

In this section we introduce a novel procedure to estimate the parameters of the reduced model in Problem 1, i.e., minimizing the equation error while guaranteeing the location and exponential stability of steady states. Thereby, we employ Monte-Carlo integration, methods from control theory and the cone complementary algorithm.

A. Evaluation of the objective function

For the evaluation of the objection function $J(\theta)$ no analytical methods are available. Fortunately, as discussed in [10], the values of $J(\theta)$ can be approximated by combining the ideas of trajectory-based model reduction, for assessing the performance of the reduced model, and Monte-Carlo integration [12].

The high-dimensional $(1 + n + n_\varphi)$ integral defining $J(\theta)$ is approximated using sampling. At first a sample of initial conditions $\{x_0^{(i)}\}_{i=1}^N$ and input trajectories $\{u_\varphi^{(i)}(\cdot)\}_{i=1}^N$ is drawn, with $x_0^{(i)} \sim p_x(x_0)$, and $\varphi^{(i)} \sim p_\varphi(\varphi)$. Secondly, a sample of time points $\{t^{(i)}\}_{i=1}^N$ is drawn from a uniform distribution over $[0, T_{\text{end}}]$. The number of sample members is denoted by N . Given $\{x_0^{(i)}\}_{i=1}^N$, $\{u_\varphi^{(i)}\}_{i=1}^N$, and $\{t^{(i)}\}_{i=1}^N$, a sample of points from I/O-trajectories is given by

$$\mathcal{T} = \left\{ y_D^{(i)}(t^{(i)}; x_0^{(i)}, u_\varphi^{(i)}) \right\}_{i=1}^N.$$

From classical Monte-Carlo integration it is now known that for $N \gg 1$ a good approximation of $J(\theta)$ employing \mathcal{T} is

$$\hat{J}(\theta) = \frac{1}{N} \sum_{i=1}^N \left(\binom{n_R}{y_D}{}^{(i)}(t^{(i)}; x_0^{(i)}, u_\varphi^{(i)}) - \theta^T m(\xi_D^{(i)}, \mu^{(i)}) \right)^2,$$

in which

$$\xi_D^{(i)} = [y_D^{(i)}(t^{(i)}; x_0^{(i)}, u_\varphi^{(i)}), \dots, y_D^{(n_R-1)(i)}(t^{(i)}; x_0^{(i)}, u_\varphi^{(i)})]^T$$

and $\mu^{(i)} = [u_\varphi^{(i)}(t^{(i)}), \dots, u_\varphi^{(n_R)(i)}(t^{(i)})]^T$. By defining

$$M = [m(\xi_D^{(1)}, \mu^{(1)}), \dots, m(\xi_D^{(N)}, \mu^{(N)})],$$

$$\psi = [y_D^{(n_R)(1)}(t^{(1)}; x_0^{(1)}, u_\varphi^{(1)}), \dots, y_D^{(n_R)(N)}(t^{(N)}; x_0^{(N)}, u_\varphi^{(N)})]^T,$$

the approximated objective function can be written in a quadratic form,

$$\hat{J}(\theta) = \frac{1}{N} \|\theta^T M - \psi^T\|_2^2.$$

Thus, the minimization of $J(\theta)$ can be approximated by a convex, quadratic optimization problem, if the constraints, which ensure the locally exponentially stable steady states, are not considered. In the next step, a formulation for the steady state constraints is derived.

B. Location and stability of steady states

In this section we derive necessary and sufficient conditions for $(\bar{u}^{(s)}, \bar{y}^{(s)})$ being a locally exponentially stable steady state of the reduced system.

1) *Location of steady states:* In a first step it has to be ensured that $(\bar{u}^{(s)}, \bar{y}^{(s)})$ is a steady state. Therefore, the right hand side of Σ_R has to vanish for $\xi = [\bar{y}^{(s)}, 0, \dots, 0]^T$ and $\mu = [\bar{u}^{(s)}, 0, \dots, 0]^T$. For $\xi_1, \dots, \xi_{n_R-1}$, this is fulfilled implicitly. Thus, the reduced model has a steady state at $(\bar{u}^{(s)}, \bar{y}^{(s)})$ if and only if

$$0 = \theta^T m([\bar{y}^{(s)}, 0, \dots, 0]^T, [\bar{u}^{(s)}, 0, \dots, 0]^T) := \theta^T m^{(s)}.$$

2) *Stability of steady states:* A necessary and sufficient condition for local exponential stability of $(\bar{u}^{(s)}, \bar{y}^{(s)})$ is that all poles of the Jacobi linearization of Σ_R at $(\bar{u}^{(s)}, \bar{y}^{(s)})$ lie in the open left half plane. Linearizing the reduced model Σ_R at $(\bar{u}^{(s)}, \bar{y}^{(s)})$ provides for the autonomous part

$$\Delta \dot{\xi} = \mathcal{A}^{(s)}(\theta) \Delta \xi \quad (2)$$

with $\mathcal{A}^{(s)}(\theta) = A + b\theta^T C^{(s)}$ depending affinely on θ , and

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n_R \times n_R}, \quad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{n_R},$$

$$C^{(s)} = \begin{bmatrix} c_1^{(s)} & \cdots & c_{n_R}^{(s)} \end{bmatrix} \in \mathbb{R}^{n_\theta \times n_R},$$

$$c_k^{(s)} = \frac{\partial m}{\partial \xi_k} \Big|_{\xi = [\bar{y}^{(s)}, 0, \dots, 0]^T, \mu = [\bar{u}^{(s)}, 0, \dots, 0]^T}, \quad k = 1, \dots, n_R.$$

It is well-known that for a given θ the linear system (2) is exponentially stable if and only if there exists a symmetric matrix $P^{(s)} \in \mathbb{S}^{n_R}$ satisfying the Lyapunov inequality [13],

$$P^{(s)} \mathcal{A}^{(s)}(\theta) + (\mathcal{A}^{(s)}(\theta))^T P^{(s)} \prec 0, \quad P^{(s)} \succ 0. \quad (3)$$

To eliminate the multiplication of the two unknowns $P^{(s)}$ and θ , we use the following lemma which is based on [1].

Lemma 1 *Given a constant matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ and symmetric matrix $P \in \mathbb{S}^n$ with $P \succ 0$, the following statements are equivalent:*

- 1) $P\mathcal{A} + \mathcal{A}^T P \prec 0$
- 2) $\exists \rho \in \mathbb{R} : \begin{bmatrix} \rho P & -\mathcal{A}^T - \rho I \\ -\mathcal{A} - \rho I & \rho P^{-1} \end{bmatrix} \succ 0.$

Proof: As $P \succ 0$, there exists a unique $Z = Z^T \succ 0$ such that $Z^2 = P$, denoted by $P^{1/2} := Z$. Therefore,

$$\begin{aligned} P\mathcal{A} + \mathcal{A}^T P \prec 0 \\ \Leftrightarrow P^{-1/2}(-P\mathcal{A} - \mathcal{A}^T P)P^{-1/2} \succ 0 \\ \Leftrightarrow \begin{bmatrix} P^{-1/2} & P^{1/2} \end{bmatrix} \begin{bmatrix} 0 & -\mathcal{A}^T \\ -\mathcal{A} & 0 \end{bmatrix} \begin{bmatrix} P^{-1/2} \\ P^{1/2} \end{bmatrix} \succ 0. \end{aligned} \quad (4)$$

As the rows of $U := \begin{bmatrix} P^{1/2} & -P^{-1/2} \end{bmatrix}$ form a basis for the null space of $\begin{bmatrix} P^{-1/2} & P^{1/2} \end{bmatrix}$, the Lemma [13, page 33] provides that (4) is equivalent to the existence of $\rho \in \mathbb{R}$ such that

$$\begin{bmatrix} 0 & -\mathcal{A}^T \\ -\mathcal{A} & 0 \end{bmatrix} + \rho U^T U \succ 0 \quad (5)$$

which concludes the proof. \blacksquare

Applying Lemma 1 to (3) yields the necessary and sufficient conditions for local exponential stability of $(\bar{u}^{(s)}, \bar{y}^{(s)})$:

$$\begin{bmatrix} \rho^{(s)} P^{(s)} & -(\mathcal{A}^{(s)}(\theta))^T - \rho^{(s)} I \\ * & \rho^{(s)} (P^{(s)})^{-1} \end{bmatrix} \succ 0, \quad P^{(s)} \succ 0. \quad (6)$$

C. Model reduction ensuring steady state properties

From the above derivations follows:

Theorem 1 *Suppose the objective $\hat{J}(\theta)$ and the linearizations of Σ_R at the steady states $\mathcal{A}^{(s)}(\theta)$ are given. Then a parameter vector θ ensures location and local exponential stability of the steady states in \mathcal{S} , if and only if it is a feasible solution of*

$$\begin{aligned} & \underset{\theta, P^{(s)}, \rho}{\text{minimize}} && \hat{J}(\theta) \\ & \text{subject to} && \theta^T m^{(s)} = 0 \\ & && P^{(s)} \succ 0 \\ & && \left. \begin{bmatrix} \rho P^{(s)} & -(\mathcal{A}^{(s)}(\theta))^T - \rho I \\ * & \rho (P^{(s)})^{-1} \end{bmatrix} \succ 0 \right\} \end{aligned} \quad (C1) \quad (C2)$$

(7)

The optimization of $\hat{J}(\theta)$ results in a minimization of the prediction error:

In (C2), the same ρ is used for all steady states. This is possible, as from the positive definiteness of $U^T U$ in (5) it follows that (6) is also fulfilled when $\rho^{(s)}$ is replaced by any value which is larger. In consequence and to avoid

the nonlinear terms $\rho P^{(s)}$ and $\rho (P^{(s)})^{-1}$, we set ρ to a predefined large value.

Unfortunately, due to the appearance of $(P^{(s)})^{-1}$ in (C2), problem (7) is nonlinear and nonconvex. Therefore, in the next step a relaxation is performed to allow for the development of an efficient computational scheme.

D. Formulation as sequential convex optimization

To overcome the nonconvexity, we propose a sequential convex programming approach employing the cone complementarity linearization [14]. Therefore, $Q^{(s)} = (P^{(s)})^{-1}$ is introduced into the optimization problem. Using $Q^{(s)}$, the constraint (C2) for the local exponential stability of a single steady state $(\bar{u}^{(s)}, \bar{y}^{(s)})$ can be written as

$$P^{(s)} \succ 0, \quad Q^{(s)} \succ 0 \quad (8)$$

$$P^{(s)} Q^{(s)} = I \quad (9)$$

$$\begin{bmatrix} \rho P^{(s)} & -(\mathcal{A}^{(s)}(\theta))^T - \rho I \\ * & \rho Q^{(s)} \end{bmatrix} \succ 0 \quad (10)$$

Due to the equality constraint $P^{(s)} Q^{(s)} = I$, this problem is still nonconvex. Nevertheless, an approximation of (9) can be derived by employing the linear matrix inequality

$$\begin{bmatrix} P^{(s)} & I \\ I & Q^{(s)} \end{bmatrix} \succcurlyeq 0. \quad (11)$$

Therefore, note that (11) is feasible with $\text{tr}(P^{(s)} Q^{(s)}) = n_R$, $P^{(s)} \succ 0$, and $Q^{(s)} \succ 0$ if and only if $P^{(s)} Q^{(s)} = I$ [14]. Otherwise $\text{tr}(P^{(s)} Q^{(s)}) > n_R$. Given this, a feasible solution for (8)–(10) may be found by solving the minimization problem

$$\begin{aligned} & \underset{\theta, P^{(s)}, Q^{(s)}}{\text{minimize}} && \text{tr}(P^{(s)} Q^{(s)}) \\ & \text{subject to} && (8), (10) \text{ and } (11) \end{aligned}$$

As for this optimization problem the objective function is nonlinear, a linear approximation is derived at a feasible point $(P_0^{(s)}, Q_0^{(s)})$. This approximation is given by $\frac{1}{2} \text{tr}(P_0^{(s)} Q^{(s)} + P^{(s)} Q_0^{(s)})$. This yields the linearized problem

$$\begin{aligned} & \underset{\theta, P^{(s)}, Q^{(s)}}{\text{minimize}} && \frac{1}{2} \text{tr}(P_0^{(s)} Q^{(s)} + P^{(s)} Q_0^{(s)}) \\ & \text{subject to} && (8), (10) \text{ and } (11), \end{aligned}$$

which is solved sequentially and shown to converge within a few iterations for similar problems [14].

Applying this procedure to all stability constraints in (7) and defining a mixed objective, we obtain

$$\begin{aligned} & \underset{\theta, P^{(s)}, Q^{(s)}}{\text{minimize}} && \hat{J}(\theta) + \alpha \sum_{s=1}^{n_S} \text{tr}(P_0^{(s)} Q^{(s)} + P^{(s)} Q_0^{(s)}) \\ & \text{subject to} && \left. \begin{aligned} & \theta^T m^{(s)} = 0 \\ & P^{(s)} \succ 0, \quad Q^{(s)} \succ 0 \\ & \begin{bmatrix} P^{(s)} & I \\ I & Q^{(s)} \end{bmatrix} \succcurlyeq 0 \\ & \begin{bmatrix} \rho P^{(s)} & -(\mathcal{A}^{(s)}(\theta))^T - \rho I \\ * & \rho Q^{(s)} \end{bmatrix} \succ 0 \end{aligned} \right\} \text{(C1)} \\ & && \left. \begin{aligned} & \text{(RC2)} \end{aligned} \right\} \\ & && s = 1, \dots, n_S \end{aligned}$$

where $\alpha \in \mathbb{R}_+$ determines the weighting for the multi-objective optimization. Instead of introducing an additional term into the objective function, one may also add constraints

$$\frac{1}{2} \text{tr}(P_0^{(s)} Q^{(s)} + P^{(s)} Q_0^{(s)}) < (1 + \epsilon) n_R \quad (12)$$

and solve the optimization problem:

Problem 2

$$\begin{aligned} \mathcal{OP}(\{(P_0^{(s)}, Q_0^{(s)})\}_{s=1}^{n_S}) : & \underset{\theta, P^{(s)}, Q^{(s)}}{\text{minimize}} && \hat{J}(\theta) \\ & \text{subject to} && (12), \text{(C1)}, \text{(RC2)} \\ & && s = 1, \dots, n_S \end{aligned}$$

From our experience it is easier to determine a suitable constant ϵ than a weighting α .

E. Model reduction algorithm

Based on the above results, we propose Algorithm 1 below for model reduction ensuring steady state properties. For ease of presentation we omit the range of $s = 1, \dots, n_S$ in Algorithm 1. The sequential convex optimization is stopped if the reduced model is exponentially stable in all steady states and if the objective criterion is less than γ .

Algorithm 1 Model reduction ensuring steady state properties

Require: $\{(P_0^{(s)}, Q_0^{(s)})\}$ and θ_0 satisfying (C1), (RC2), and (12).

$k = 0$.

while $\mathcal{A}^{(s)}(\theta_k)$ not exponentially stable or $\hat{J}(\theta_k) \geq \gamma$ **do**

Find $\{(P_{k+1}^{(s)}, Q_{k+1}^{(s)})\}$ and θ_{k+1} by solving the convex optimization problem $\mathcal{OP}(\{(P_k^{(s)}, Q_k^{(s)})\})$.

Update $k = k + 1$.

end while

return θ_k .

Finding an initially feasible solution satisfying (12), (C1), and (RC2) is a topic for future research. We set $P_0^{(s)} = I$ and $Q_0^{(s)} = I$ for $s = 1, \dots, n_S$ to initialize Algorithm 1, which was no restriction for the examples considered.

Remark 1 *Computing the solution of Problem 2 is independent of the detailed model, assuming the sample of points T has been computed beforehand. This emphasizes that only the complexity of the detailed models I/O-behavior is relevant for a model reduction procedure based on Theorem 1. Thus, the model reduction procedure can be applied to large-scale detailed models. Unfortunately, the order of the reduced model is limited due to numerical difficulties originating from the model structure with the derivatives of the output as states.*

IV. EXAMPLE: FERMI-PASTA-ULAM LATTICE

In this section the proposed model reduction approach is applied to the Fermi-Pasta-Ulam (FPU) lattice.

A. Introduction and FPU model

Prior to 1955, E. Fermi, J. Pasta, and S. Ulam studied dynamical systems consisting of a chain of particles with forces between neighbors that can be modeled by nonlinear springs. Through numerical experiments they analyzed the time evolution of the energy of each normal mode discovering the well known FPU paradox [15], which is studied till recently.

The FPU lattice has been chosen due to its importance in nonlinear science and the existence of nonlinear oscillations. Due to the nonlinear oscillations, the poles of the Jacobi linearization around a steady state are close to, respectively on the imaginary axis. Therefore, the reduced model is prone to be not locally exponentially stable around the steady states.

For this study, the FPU model as in [15, Figure 4] consisting of 16 particles with fixed end points and cubic forces is used. Denoting the displacement and velocity of the i -th particle with x_i and v_i , respectively, the cubic force acting on particle i , $i = 2, \dots, 15$, is given by

$$F_i(x) = (x_{i-1} - x_i) + 8(x_{i-1} - x_i)^3 + (x_{i+1} - x_i) + 8(x_{i+1} - x_i)^3.$$

Due to the fixed end points the displacement and velocity of the first and last particle is 0. Therefore, the detailed model contains 28 states. As we require exponentially stable steady states we extended the model with a viscous damping

$$D_i(v) = 0.2(v_{i-1} - v_i) + 0.2(v_{i+1} - v_i), \quad i = 2, \dots, 15.$$

Furthermore, as we consider I/O based model reduction, we define the input as an external force acting on particle 8 and the position of this particle as output. Thus, the differential equations of this model are given by

$$\begin{aligned} \dot{x}_i &= \begin{cases} 0 & \text{for } i = 1, 16 \\ v_i & \text{for } i = 2, \dots, 15 \end{cases} \\ \dot{v}_i &= \begin{cases} 0 & \text{for } i = 1, 16 \\ D_i(v) + F_i(x) & \text{for } i = 2, \dots, 15, i \neq 8 \\ D_i(v) + F_i(x) + u_\varphi & \text{for } i = 8 \end{cases} \end{aligned}$$

$$y_D = x_8.$$

B. Model reduction

For the model reduction of the FPU model, I/O-data is generated by simulation of the detailed model. Denote the steady state for $\bar{u} = 1.5$ by \bar{x} and $\Phi(x|\alpha, \beta)$ the probability density of the uniform distribution in the interval $x \in [\alpha, \beta]$. The weighting of the initial condition p_x is chosen as

$$p_x(x_0) = \prod_{i=1}^n \Phi(x_{0,i}|0.99\bar{x}_i, 1.01\bar{x}_i).$$

With the Heaviside step function denoted $\sigma(t)$, the input trajectories are chosen to be multi-step functions

$$u_\varphi(t) = 1.5 + \sum_{i=1}^4 (h_i - h_{i-1})\sigma(t - \tau_i)$$

smoothed by the transfer function $G(s) = (s + 1)^{-n_R}$. The parameters of the input trajectories are

$$\varphi := [h_1, h_2, h_3, h_4, \tau_1, \tau_2, \tau_3, \tau_4]^T.$$

The weighting function for the input is

$$p_\varphi(\varphi) = \left(\prod_{i=1}^4 \Phi(h_i|0.5, 2.5) \right) \left(\prod_{i=1}^4 \Phi(\tau_i|0, 100) \right),$$

in which the order $\tau_1 \leq \tau_2 \leq \tau_3 \leq \tau_4$ is assumed. Within the interval of possible input values, the two values $\bar{u}_1 = 0.75$ and $\bar{u}_2 = 2$ define the set of steady states $\mathcal{S} = \{(0.75, 1.86), (2, 3.11)\}$.

For the model reduction a sample with 100 members is drawn from $p_x(x_0)$ and $p_\varphi(u_\varphi)$, respectively. For the resulting pairs the FPU model is simulated for the time interval $[0, 100]$. In order to use most of the information contained in every I/O-trajectory, the trajectories are sampled with a sampling time of 0.1, yielding the sample of I/O-trajectories \mathcal{T} .

Given the sample \mathcal{T} , the order n_R of the reduced model has to be chosen. For n_R several values have been tested, yielding that a reasonable approximation of the I/O-behavior is achieved with $n_R = 2$. Due to the cubic force existing in the FPU model, the ansatz functions $m(\xi, \mu)$ consist of all monomials with degree less or equal three. Thus, the reduced model with 56 parameters and monomials is given by

$$\Sigma_R : \begin{cases} \dot{\xi}_1 = \xi_2 \\ \dot{\xi}_2 = \theta^T m(\xi, \mu) \\ y_R = \xi_1. \end{cases}$$

Based on this structure of the reduced model and the sample \mathcal{T} , two parameterizations of the reduced model are determined using the MATLAB toolbox Yalmip [16] with the solver SeDuMi [17]. The first parameterization or reduced model is determined by minimizing $\hat{J}(\theta)$ subject to (C1). The second reduced models is computed using Algorithm 1 and thus guarantees in addition to the location also the local exponential stability of the steady states. To enforce sparsity of θ the ℓ_1 -norm formulation of [10] is adopted. The optimization results in the reduced model 1 with 8 parameters determined by

$$\begin{aligned} \ddot{y}_R &= -0.038 y_R & + 0.14 u_\varphi & + 0.057 \ddot{u}_\varphi \\ &+ 0.0095 u_\varphi y_R & - 0.0074 y_R^3 & - 0.076 \dot{y}_R^3 \\ &- 0.02 u_\varphi^2 \dot{u}_\varphi & + 0.96 \dot{u}_\varphi^3 \end{aligned}$$

and the reduced model 2 with 13 parameters given by

$$\begin{aligned} \ddot{y}_R &= -0.035 y_R & - 0.04 \dot{y}_R & + 0.14 u_\varphi \\ &+ 0.0019 y_R^2 & - 0.0097 y_R \dot{y}_R & - 0.0001 u_\varphi y_R \\ &+ 0.011 u_\varphi \dot{y}_R & - 0.011 y_R^3 & + 0.017 y_R^2 \dot{y}_R \\ &+ 0.011 u_\varphi y_R^2 & - 0.035 u_\varphi y_R \dot{y}_R & - 0.0065 u_\varphi^2 y_R \\ &+ 0.02 u_\varphi^2 \dot{y}_R. \end{aligned}$$

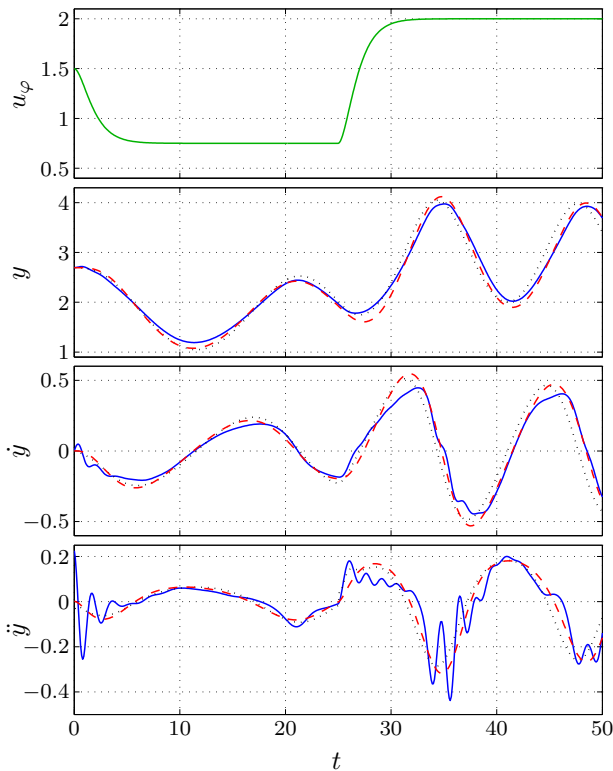


Fig. 1. Input trajectory (—) and the corresponding trajectories of the detailed model (—), reduced model 1 (---) and reduced model 2 (---).

TABLE I

POLES OF THE JACOBI LINEARIZATION AT THE STEADY STATES.

$\bar{u}^{(s)}$	reduced model 1	reduced model 2
0.75	$0 \pm 0.328i$	$-0.0147 \pm 0.336i$
2.00	$0 \pm 0.482i$	$-0.0125 \pm 0.475i$

C. Validation of reduced model

To validate the reduced model, trajectories from simulations with x_0 drawn according to $p_x(x_0)$ are used as further validation. Exemplary, trajectories for the input trajectory with $\varphi = [0.75, 2.0, 1.5, 0.5, 0, 25, 50, 75]^T$ are shown in Figure 1. It is obvious, that the quantitative I/O-behavior for increasing and decreasing steps is reproduced well by both reduced models. Comparing the second time derivative of the outputs, one realizes that the major oscillation, which is apparently nonlinear, is captured by the reduced models whereas the faster oscillations of the detailed model are not reproduced due to the limitation to two states.

To examine the local steady state behavior, the poles of the Jacobi linearization at the steady states of both reduced models is given in Table I, showing that the reduced model 1 exhibits undamped oscillations. For the reduced model 2 the locally exponential stability is guaranteed by Algorithm 1 resulting in lightly damped poles.

V. CONCLUSION

In this paper an extension of a trajectory-based nonlinear model reduction technique with conditions for location and stability of a set of steady states is proposed. The trajectory-based approach parametrizes the reduced models which results in a convex objective function, while the steady

state constraints restrict the space of feasible parameters. To ensure a maximal number of degrees of freedom despite the stability requirement, a novel necessary and sufficient condition for simultaneous local exponential stability is proposed. This formulation can also be utilized when performing simultaneously stabilizing static output feedback design.

As the resulting overall optimization problem has nonlinear constraints, the cone complementarity linearization is employed. This results in a sequential convex optimization problem which has good convergence properties. Hence, the model reduction for nonlinear ODEs could be reformulated to a series of efficient convex programs.

To evaluate the proposed approach the Fermi-Pasta-Ulam's spring-mass system is studied. The system has sixteen masses and nonlinear spring potentials. Employing the proposed algorithm a reduced model with only two states is computed, which reproduces the important I/O-behavior and satisfies the stability criteria.

REFERENCES

- [1] T. P. Nguyen, "Semidefinite programming relaxations of the matrix stability problem," Student thesis, University of Stuttgart Institute for System Theory and Automatic Control, 2004.
- [2] A. C. Antoulas, *Approximation of large-scale dynamical systems*, ser. Advances in design and control. Philadelphia: SIAM, 2005.
- [3] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Autom. Contr.*, vol. 26, no. 1, pp. 17–32, 1981.
- [4] R. Freund, "Model reduction methods based on Krylov subspaces," *Act. Num.*, vol. 12, pp. 267–319, 2003.
- [5] J. M. A. Scherpen, "Balancing for nonlinear systems," *Syst. Control Lett.*, vol. 21, no. 2, pp. 143–153, 1993.
- [6] S. Lall, J. Marsden, and S. Glavaski, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *Int. J. Robust and Nonlinear Control*, vol. 12, no. 6, pp. 519–535, 2002.
- [7] S. Prajna, "POD model reduction with stability guarantee," in *Proc. IEEE Conf. Decision and Control*, 2003, pp. 5254–5258.
- [8] B. Bond and L. Daniel, "Stable reduced models for nonlinear descriptor systems through piecewise-linear approximation and projection," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, no. 10, pp. 1467–1480, 2009.
- [9] P. Wolfrum, A. Vargas, M. Gallivan, and F. Allgöwer, "Complexity reduction of a thin film deposition model using a trajectory based nonlinear model reduction technique," in *Proc. Amer. Contr. Conf.*, 2005, pp. 2566–2571.
- [10] M. Löhring, J. Hasenauer, and F. Allgöwer, "Trajectory-based model reduction of nonlinear biochemical networks employing the observability normal form," in *Proc. 18th IFAC World Congress*, 2011, pp. 10442–10447.
- [11] M. Zeitz, "Observability canonical (phase-variable) form for nonlinear time-variable systems," *Int. J. Syst. Science*, vol. 15, no. 9, pp. 949–958, 1984.
- [12] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2005.
- [13] S. Boyd, *Linear matrix inequalities in system and control theory*. Philadelphia, Pa. : Society for Industrial and Applied Mathematics, 1994.
- [14] L. El Ghaoui, F. Oustry, and M. Ait Rami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE Trans. Autom. Contr.*, vol. 42, no. 8, pp. 1171–1176, 1997.
- [15] E. Fermi, J. Pasta, and S. Ulam, "Studies of nonlinear problems," *Los Alamos report LA-1940*, 1955.
- [16] J. Löfberg, "Dualize it: software for automatic primal and dual conversions of conic programs," *Optim. Meth. and Softw.*, vol. 24, no. 3, pp. 313–325, 2009.
- [17] J. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optim. Meth. and Softw.*, vol. 11, pp. 625–653, 1999.