

Adaptive Sampling for Structure-Preserving Model Order Reduction of Port-Hamiltonian Systems[★]

Paul Schwerdtner^{*} Matthias Voigt^{**}

^{*} Technische Universität Berlin, Institut für Mathematik, Straße des
17. Juni 136, 10623 Berlin, Germany (e-mail:
schwerdt@math.tu-berlin.de, corresponding author)

^{**} UniDistance Suisse, Überlandstrasse 12, 3900 Brig, Switzerland
(e-mail: matthias.voigt@fernuni.ch)

Abstract: We present an adaptive sampling strategy for the optimization-based structure-preserving model order reduction (MOR) algorithm developed in [Schwerdtner, P. and Voigt, M. (2020). Structure-preserving model order reduction by parameter optimization, Preprint arXiv:2011.07567]. This strategy reduces the computational demand and the required a priori knowledge about the given full-order model, while at the same time retaining a high accuracy compared to other structure-preserving but also unstructured MOR algorithms. A numerical study with a port-Hamiltonian benchmark system demonstrates the effectiveness of our method when combined with this new adaptive sampling strategy. We also investigate the distribution of the sample points.

Copyright © 2021 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: model reduction, H-infinity optimization, structured systems, port-Hamiltonian systems, structure-preserving methods

1. INTRODUCTION

We consider structure-preserving model order reduction (MOR) for linear time-invariant port-Hamiltonian (pH) systems of the form

$$\Sigma_{\text{pH}} : \begin{cases} \dot{x}(t) = (J - R)Qx(t) + Bu(t), \\ y(t) = B^T Qx(t), \end{cases} \quad (1)$$

where $J, R, Q \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ with $J = -J^T$, $R \geq 0$, and $Q \geq 0$. We call $x : \mathbb{R} \rightarrow \mathbb{R}^n$, $u : \mathbb{R} \rightarrow \mathbb{R}^m$, and $y : \mathbb{R} \rightarrow \mathbb{R}^m$ the *state*, *input*, and *output* of the system, respectively. The *Hamiltonian* $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}^+ := \{a \in \mathbb{R} \mid a \geq 0\}$ of Σ_{pH} , which describes the total internal energy of the system, is given by

$$\mathcal{H}(x(t)) = \frac{1}{2} x(t)^T Q x(t).$$

The *transfer function* of Σ_{pH} is given by

$$H_{\text{pH}}(s) := B^T Q (sI_n - (J - R)Q)^{-1} B.$$

It is a real-rational matrix-valued function that describes the input-to-output mapping in the frequency domain.

The use of pH models in systems and control applications comes with several benefits. First, pH models provide a systematic way to model the interaction of different subsystems, particularly across different physical domains or

time scales (van der Schaft and Jeltsema (2014)). Moreover, pH systems are automatically *passive* (Beattie et al. (2019)). This further simplifies the use of pH subsystems in a networked system since the power-conserving interconnection of passive systems is again passive and the passivity of a system implies its stability (if it is minimal). Furthermore, the pH paradigm allows for structure-preserving spatial semi-discretizations of partial differential equations (Serhani et al. (2019)) such that systems as defined in (1) can be obtained from first principle modeling.

However, the increased demand for highly accurate models leads to systems with a large state-space dimension n in most practical applications. This often renders the analysis, simulation, or control of such systems computationally demanding if not prohibitively expensive. For that reason, MOR can be employed to compute a surrogate *reduced-order model* (ROM) with state-space dimension r such that $r \ll n$ and an input-to-output mapping similar to the *full-order model* (FOM). For a comprehensive introduction to MOR we refer the reader to Antoulas (2005).

Well-established MOR methods such as *balanced truncation* (BT) or the *iterative rational Krylov algorithm* (IRKA) focus on computing ROMs with a high accuracy in terms of the input-to-output mapping but neglect the preservation of additional structural properties of the FOM. In particular, the ROM computed for a given pH model is not necessarily pH. However, when the pH property of that model is used for the automatic construction of a networked system, then the preservation of the pH property is essential. For this reason, the preservation

[★] This work is supported by the German Research Foundation (DFG) within the project VO2243/2-1: “Interpolationsbasierte numerische Algorithmen in der robusten Regelung” and the DFG Cluster of Excellence MATH+ within the project AA4-5: “Energy-based modeling, simulation, and optimization of power systems under uncertainty”.

of additional structural properties during MOR is also considered in the literature.

The method in Gugercin et al. (2012) adapts IRKA to also preserve the pH property of a dynamical system, while the method in Polyuga and van der Schaft (2008) (pH-BT) uses a balancing based approach. In our previous study (Schwerdtner and Voigt (2020)), we observe that for both methods the accuracy is drastically reduced compared to non-structure-preserving IRKA or BT (this can also be observed in Fig. 4), and develop a new optimization-based MOR framework that also preserves the pH structure but has a similar accuracy as the non-structure-preserving MOR methods. We also report that recently Breiten and Unger (2021) have developed a passivity-preserving MOR method built on the *spectral factorization* of a pH system's transfer function that achieves an accuracy similar to the original IRKA. Here the pH structure can be recovered after the reduction using a state-space transformation outlined, e.g., in Beattie et al. (2015). Another passivity-preserving approach is *positive real balanced truncation* (Desai and Pal (1984)).

The purpose of this work is to overcome one of the major issues with our direct optimization approach as introduced in Schwerdtner and Voigt (2020). This approach is based on approximating the transfer function of the given FOM at a set of $n_S \in \mathbb{N}$ sample points $S := \{s_i\}_{i=1 \dots n_S} \subset i\mathbb{R}$, where $i\mathbb{R}$ denotes the imaginary axis. The major issue is that these sample points are fixed and required as user input. However, the approximation quality of the resulting ROM heavily depends on these sample points. Furthermore, the selection of these sample points is not a trivial task. On the contrary, it requires knowledge about the FOM, which is typically computationally expensive to obtain, and an understanding of the methodology that is used to approximate the FOM. Therefore, in this article, we introduce an adaptive sampling strategy, which automatically chooses new sample points as our method progresses. This strategy is based on an argument introduced in Apkarian and Noll (2018).

2. PRELIMINARIES

In this work we consider the space $\mathcal{RH}_\infty^{m \times m}$ of all real-rational and proper $m \times m$ transfer functions that have no poles in the set $\overline{\mathbb{C}^+} := \{\lambda \in \mathbb{C} \mid \operatorname{Re}(\lambda) \geq 0\}$. For $H \in \mathcal{RH}_\infty^{m \times m}$, the $\mathcal{H}_\infty$ norm is given by

$$\|H\|_{\mathcal{H}_\infty} := \sup_{\lambda \in \mathbb{C}^+} \|H(\lambda)\|_2 = \sup_{\omega \in \mathbb{R}} \sigma_1(H(i\omega)),$$

where $\sigma_1(\cdot)$ denotes the largest singular value of its matrix argument. The $\mathcal{H}_\infty$ norm is often used to determine the approximation quality of a ROM due to the well-known error bound

$$\|y - \tilde{y}\|_{L^2([0, \infty), \mathbb{R}^m)} \leq \|H - \tilde{H}\|_{\mathcal{H}_\infty} \|u\|_{L^2([0, \infty), \mathbb{R}^m)},$$

in which y and H as well as $\tilde{y}$ and $\tilde{H}$ denote the output and transfer function of a FOM and its ROM, respectively. Note that, alternatively, the $\mathcal{H}_2$ norm can be used as accuracy measure as it provides a similar error bound. Therefore, most MOR methods for linear systems aim at computing ROMs such that the error of the transfer functions with respect to these norms is small. In the following, we denote the Euclidian norm of the evaluated

error transfer function by $E(s) := \|H(s) - \tilde{H}(s)\|_2$. We focus our investigation on the $\mathcal{H}_\infty$ norm but the results for the $\mathcal{H}_2$ norm are similar, which can also be observed in Schwerdtner and Voigt (2020).

While most MOR methods compute the ROM matrices via subspace projection, we construct a parameterized low-order realization and treat MOR as a parameter optimization problem. A parameterized pH system is given as follows.

Lemma 1. (Schwerdtner and Voigt (2020)). Let $\theta \in \mathbb{R}^{n_\theta}$ be a vector with $n_\theta := n \left(\frac{3n+1}{2} + m \right)$. Furthermore, let θ be partitioned as $\theta := [\theta_J^\top, \theta_R^\top, \theta_Q^\top, \theta_B^\top]^\top$ with $\theta_J \in \mathbb{R}^{n(n-1)/2}$, $\theta_R \in \mathbb{R}^{n(n+1)/2}$, $\theta_Q \in \mathbb{R}^{n(n+1)/2}$, and $\theta_B \in \mathbb{R}^{nm}$. Further define the matrices

$$J(\theta) = \text{vtsu}(\theta_J)^\top - \text{vtsu}(\theta_J), \quad (2a)$$

$$R(\theta) = \text{vtu}(\theta_R)^\top \text{vtu}(\theta_R), \quad (2b)$$

$$Q(\theta) = \text{vtu}(\theta_Q)^\top \text{vtu}(\theta_Q), \quad (2c)$$

$$B(\theta) = \text{vtf}_{n,m}(\theta_B), \quad (2d)$$

where the function $\text{vtu} : \mathbb{R}^{n(n+1)/2} \rightarrow \mathbb{R}^{n \times n}$ maps a vector of length $n(n+1)/2$ to an upper triangular matrix, the function $\text{vtsu} : \mathbb{R}^{n(n-1)/2} \rightarrow \mathbb{R}^{n \times n}$ maps a vector of length $n(n-1)/2$ to a strictly upper triangular matrix, and the function $\text{vtf}_{n,m} : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{n \times m}$ reshapes a vector of length nm to an $n \times m$ matrix. Then, to each $\theta \in \mathbb{R}^{n_\theta}$ one can assign the pH system

$$\Sigma_{\text{pH}}(\theta) : \begin{cases} \dot{x}(t) = (J(\theta) - R(\theta))Q(\theta)x(t) + B(\theta)u(t), \\ y(t) = B(\theta)^\top Q(\theta)x(t). \end{cases} \quad (3)$$

Conversely, to each pH system Σ_{pH} with n states and m inputs and outputs one can assign a vector $\theta \in \mathbb{R}^{n_\theta}$ such that $\Sigma_{\text{pH}} = \Sigma_{\text{pH}}(\theta)$ with $\Sigma_{\text{pH}}(\theta)$ as in (3).

For details of the construction, we refer the reader to Schwerdtner and Voigt (2020). In the following, we denote the transfer function of $\Sigma_{\text{pH}}(\theta)$ by $H_{\text{pH}}(\cdot, \theta)$.

The MOR method in Schwerdtner and Voigt (2020) consists of two stages. First, an initialization is performed to obtain the initial parameter vector θ_0 . Then the parameterization is optimized with respect to a sequence of objective functions designed to minimize the $\mathcal{H}_\infty$ norm of the difference between the transfer function of the given FOM and the transfer function of the ROM parameterization.

The initialization step is based on a greedy interpolation strategy presented in Beddig et al. (2019). In particular, the initial ROM is constructed by iteratively adding imaginary interpolation points at which the $\mathcal{H}_\infty$ norm of the current error transfer function is attained. This construction uses the projection introduced in Gugercin et al. (2012). The $\mathcal{H}_\infty$ norms of the respective error transfer functions, which are large-scale, are computed using the method proposed in Schwerdtner and Voigt (2018). Interpolation points are added, until the ROM has a prescribed state-space dimension. Then the parameter θ_0 is extracted from this initial model to create the initial parameterized pH system and the optimization stage is started.

A direct gradient-based minimization of the $\mathcal{H}_\infty$ error is avoided, since

- (i) the repeated $\mathcal{H}_\infty$ norm computation for large-scale systems is computationally demanding despite efficient methods that were proposed recently,
- (ii) the $\mathcal{H}_\infty$ norm depends nonsmoothly on the model parameters, which obstructs the optimization, and
- (iii) the gradient of the $\mathcal{H}_\infty$ norm (if it exists) contains little information for a good descent direction because it only considers the error at a single point.

Instead we propose to minimize the objective function

$$\mathbf{L}(\gamma, H, H_{\text{pH}}(\cdot, \theta), S) := \frac{1}{\gamma} \sum_{s_i \in S} \left((\|H(s_i) - H_{\text{pH}}(s_i, \theta)\|_2 - \gamma)_+ \right)^2, \quad (4)$$

where

$$(\cdot)_+ : \mathbb{R} \rightarrow \overline{\mathbb{R}}^+, \quad x \mapsto \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0 \end{cases}$$

with respect to θ for decreasing values of $\gamma > 0$ to fit $H_{\text{pH}}(\cdot, \theta)$ to a given transfer function H .

Remark 2. A few remarks on $\mathbf{L}$ are in order.

- (i) The value of $\mathbf{L}$ is zero if the norm of the error between the given and the reduced transfer function is less than γ at all sample points. In this way, a connection between $\mathbf{L}$ and the $\mathcal{H}_\infty$ error minimization problem is established.
- (ii) Under mild assumptions it can be shown that $\mathbf{L}$ is differentiable with respect to θ at some point $\theta_0 \in \mathbb{R}^{n_\theta}$, which facilitates the optimization.
- (iii) The gradient of $\mathbf{L}$ with respect to θ contains a descent direction that considers all sample points, at which the error is larger than the current level γ .

The optimization step consists of first finding a suitable value for γ and then minimizing $\mathbf{L}$ using nonlinear optimization techniques (we use the implementation of BFGS from Mogensen and Riseth (2018)). In this work, we select γ using bisection, i.e., when the optimization determines a minimum at which $\mathbf{L}$ is zero, we reduce γ for the construction of the next optimization problem, otherwise it is increased. The method is similar to the method described in Alg. 2. However, Alg. 2 includes the adaptive sampling strategy, which is explained in the following section.

3. OUR METHOD WITH ADAPTIVE SAMPLING

Since a zero value of $\mathbf{L}$ only guarantees that the error is less than γ at the given sample points, the distribution of these points plays an essential role in connecting the minimization of $\mathbf{L}$ to the minimization of the overall $\mathcal{H}_\infty$ error. If the sample points are distributed such that some peaks in the error transfer function are missed entirely, then a minimization of $\mathbf{L}$ does not lead to a good $\mathcal{H}_\infty$ fit. This consideration may cause the user to prescribe a large number of sample points. However, the cost of evaluating $\mathbf{L}$ and its gradient grows with the number of sample points, so a large number of unnecessary sample points should also be avoided. For these reasons, an automatic selection of the sample points is desirable.

We propose a sample selection algorithm based on the following result from Apkarian and Noll (2018). Therein, for $\omega_1 \leq \omega_2$, $L[\omega_1, \omega_2]$ is called a first-order upper bound

for a continuously differentiable $\phi : \mathbb{R} \rightarrow \overline{\mathbb{R}}^+$, if $L[\omega_1, \omega_2] \geq |\phi'(\omega)|$ for all $\omega \in [\omega_1, \omega_2]$.

Lemma 3. (Apkarian and Noll (2018)). Let $\phi : \mathbb{R} \rightarrow \overline{\mathbb{R}}^+$ be continuously differentiable and let ω_i and ω_{i+1} be two consecutive nodes of a piecewise linear interpolation P_ϕ of ϕ such that $\gamma^* \geq \max\{\phi(\omega_i), \phi(\omega_{i+1})\}$. Furthermore, let $L[\omega_i, \omega_{i+1}]$ be a first-order upper bound for ϕ on the interval $[\omega_i, \omega_{i+1}]$. Assume that for a given tolerance $\tau > 0$ we have that

$$L[\omega_i, \omega_{i+1}](\omega_{i+1} - \omega_i) < 2\gamma^* + 2\tau - \phi(\omega_i) - \phi(\omega_{i+1}). \quad (5)$$

Then, $\phi(\omega) < \gamma^* + \tau$ for every $\omega \in [\omega_i, \omega_{i+1}]$.

In Apkarian and Noll (2018) this result is used to construct an accuracy certificate after their main optimization algorithm (which computes $\mathcal{H}_\infty$ controllers) has terminated to restart it on a finer grid, if necessary. Since our method, on the other hand, is based on multiple optimizations of $\mathbf{L}$ for varying values of γ , and, moreover, a meaningful tolerance is also given by the target value for γ , we can use Lemma 3 already within our algorithm to update the sample points as γ is decreased before the nonlinear optimization step is started.

We use Alg. 1 to update the sample points that are used in the objective in Alg. 2 to achieve a good $\mathcal{H}_\infty$ fit. The sampling update in Alg. 1 works as follows. Between each pair of adjacent sample points, a new sample $i\omega_{\text{test}}$ is chosen and the difference quotients between this new sample and the two respective adjacent points are computed. The maximum of these difference quotients serves as estimate for the first-order upper bound in the respective interval. The tolerance (τ in Lemma 3) is set to the current level γ . If now the criterion given in the inequality (5) which guarantees that the error between the two sample points is less than $\gamma^* + \tau$ is not met, the extra sample point $i\omega_{\text{test}}$ is added to the sample point set. Due to the fact that multiple additional points may be required

Algorithm 1 Logarithmic Sampling Adaptation

Input: The error $E : \mathbb{C}^+ \rightarrow \overline{\mathbb{R}}^+$, imaginary parts of current sample points $S = \{\omega_1, \dots, \omega_{n_S}\}$, current level $\gamma > 0$.

Output: Updated set of sample points S .

```

1: Set  $n_{\text{new}} := 1$ .
2: while  $n_{\text{new}} > 0$  do
3:   Sort  $\{\omega_1, \dots, \omega_{n_S}\}$  such that  $\omega_i < \omega_{i+1}$  for  $i = 1, \dots, n_S - 1$ .
4:   Set  $n_{\text{new}} := 0$ .
5:   for  $i = 1, \dots, n_S - 1$  do
6:     Set  $\omega_{\text{test}} := 10^{(\log_{10}(\omega_i) + \log_{10}(\omega_{i+1}))/2}$ .
7:     Set  $d_1 := (E(i\omega_{\text{test}}) - E(i\omega_i))/(\omega_{\text{test}} - \omega_i)$ .
8:     Set  $d_2 := (E(i\omega_{i+1}) - E(i\omega_{\text{test}}))/(\omega_{i+1} - \omega_{\text{test}})$ .
9:     Set  $\gamma^* := \max\{E(i\omega_i), E(i\omega_{i+1})\}$ .
10:    Set  $d^* := \max\{d_1, d_2\}$ .
11:    if  $d^*(\omega_{i+1} - \omega_i) \geq 2(\gamma^* + \gamma) - (E(i\omega_i) + E(i\omega_{i+1}))$  then
12:      Set  $n_{\text{new}} := n_{\text{new}} + 1$ .
13:      Set  $n_S := n_S + 1$  and set  $S := S \cup \{\omega_{\text{test}}\}$ .
14:    end if
15:  end for
16: end while

```

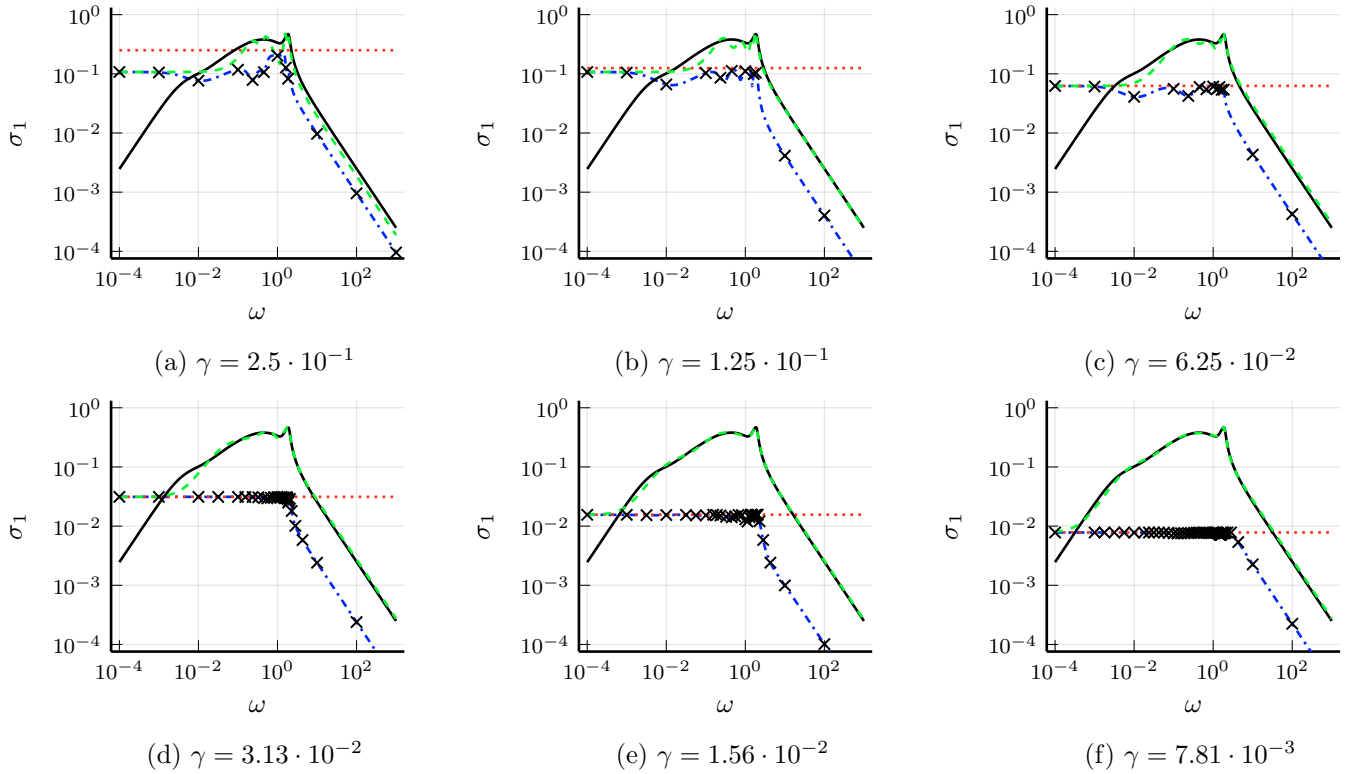


Fig. 1. The progress of Alg. 2 for reduced model order of $r = 8$ is depicted for decreasing levels γ . The FOM transfer function is illustrated as black solid line, the ROM transfer function is depicted as green dashed line, and the error is shown as blue dash-dotted line. The sample points are depicted as black crosses.

Algorithm 2 Iterative Reduction Algorithm

Input: Transfer function $H \in \mathcal{RH}_{\infty}^{m \times m}$, initial parameterized transfer function $H_{\text{pH}}(\cdot, \theta) \in \mathcal{RH}_{\infty}^{m \times m}$, $\gamma_{\text{max}} > 0$, initial sample points S , bisection tolerance τ_b .

Output: Optimized ROM transfer function $H_{\text{pH}}(\cdot, \theta_{\text{opt}})$

```

1: Set  $\gamma_{\text{min}} := 0$  and  $\gamma := \gamma_{\text{max}}$ .
2: while  $\gamma\tau_b < (\gamma_{\text{max}} - \gamma_{\text{min}})$  do
3:   Set  $\gamma := (\gamma_{\text{max}} + \gamma_{\text{min}})/2$ .
4:   Update  $S$  using Alg. 1.
5:   Update  $\theta_{\text{opt}} := \arg\min_{\theta \in \mathbb{R}^{n_{\theta}}} \mathbf{L}(\gamma, H, H_{\text{pH}}(\cdot, \theta), S)$ .
6:   if  $\mathbf{L}(\gamma, H, H_{\text{pH}}(\cdot, \theta_{\text{opt}}), S) > 0$ . then
7:     Set  $\gamma_{\text{min}} := \gamma$ .
8:   else
9:     Set  $\gamma_{\text{max}} := \gamma$ .
10:  end if
11: end while

```

between two original sample points, we iterate through the set of sample points until no more new points are added. Note that in Alg. 1, for simplicity, we only refer to the imaginary parts of the sample points.

Remark 4. Apkarian and Noll (2018) use the difference quotients between the two adjacent sample points already in the set as estimate for the derivative upper bound. However, since our method tends to cause the error function to attain similar values (in magnitude) at the sample points, the actual derivative is heavily underestimated, when this difference quotient is used. Inserting the extra point between the two samples, on the other hand, provides better estimates for the derivative.

In Alg. 2, our adaptive sampling MOR algorithm based on a bisection over γ is described. In each iteration, first the sample points are updated as described above. Then a minimizer θ_{opt} is computed for the objective function for the current γ level. For that, we use a BFGS optimization solver from Mogensen and Riseth (2018). If the final objective value is zero, i.e., the error at all sample points is below γ , then γ_{max} is updated to the current value of γ such that γ is reduced in the next iteration. Otherwise, the currently required accuracy is not met and γ is increased by setting γ_{min} to the current value of γ .

4. NUMERICAL EXPERIMENTS

We evaluate our new adaptive method on a pH mass-springer-damper chain benchmark example also used in Gugercin et al. (2012) to assess the performance of the pH structure-preserving variant of IRKA. We set the state-space dimension only to 100 to be able to compute the global $\mathcal{H}_{\infty}$ approximation error during the evaluation. However, we demonstrate in our previous work (Schwerdtner and Voigt (2020)) that for larger state-space dimensions, neither the shape of the transfer function of the given pH model nor the performance of our reduction algorithm change significantly. Moreover, the model has $m = p = 2$ inputs and outputs. The purpose of our experiments is (i) to demonstrate that the $\mathcal{H}_{\infty}$ accuracy of the adaptive method is comparable to our previous method despite using significantly fewer sample points; and (ii) to investigate where the adaptive sampling method places the sample points.

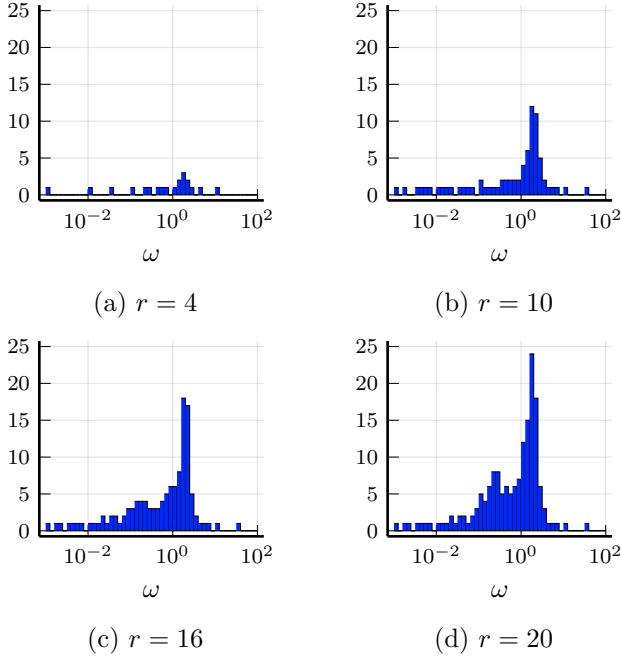


Fig. 2. Histograms showing the distribution of the sample points after Alg. 2 has terminated for different reduced model orders.

Our experimental setup is as follows. We compute ROMs for the pH benchmark system for reduced state-space dimensions 4, 6, $\dots$, 20. For that, we run our greedy interpolation initialization procedure and stop at 2, 3, $\dots$, 10 interpolation points. After that, the initial parameter θ_0 is extracted from the initial reduced model and the initial parameterized model is constructed. As initial set of sample points, we concatenate a set of logarithmically spaced points $\{10^{-8}, 10^{-7}, \dots, 10^5\}$ with the interpolation points computed during initialization. For the bisection parameters we choose $\tau_b = 0.1$ and initially set $\gamma_{\max} = 0.5$. Note that $\gamma_{\min}$ is always set to zero. All experiments are run in *julia* (version 1.5.3) on an Intel® Core™ i9-9900K CPU at 3.60 GHz with 32 GB of RAM.

In Fig. 1, we illustrate the progress of our method over the first 6 iterations, as γ is decreasing. At early stages of the algorithm (Fig. 1 (a–c)), the error at the different sample points varies, while it attains similar values close to γ at later iterations, except beyond a certain frequency at which all depicted transfer functions converge to zero. This rather flat error curve indicates that a good $\mathcal{H}_\infty$ fit is found. Furthermore, most sample points are added around 10^0 , which is where the transfer function of the FOM has two peaks.

In Fig. 2, histograms showing the distribution of the sample points after the completion of Alg. 2 are depicted for different reduced model orders and, consequently, different final accuracies. We can observe that as the reduced model order and at the same time the final accuracy increases, the concentration of sample points around 10^0 intensifies. This is due to the fact that the error has many small peaks in this frequency range, which lead to a larger estimate of the first-order derivative bound. This can also be observed in Fig. 3, in which the spiky behavior of the error transfer function between 10^{-1} and 10^1 is illustrated.

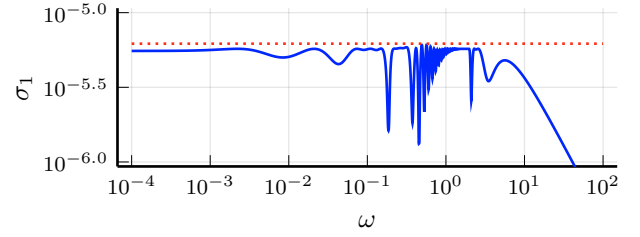


Fig. 3. Final error transfer function for reduced model order $r = 20$ depicted in blue. The corresponding γ level is shown as red dotted line.

Table 1. Number of sample points for different reduced model orders during the execution of Alg. 2 at different intermediate values for γ .

$r \backslash \gamma$	4	6	8	10	12	14	16	18	20
2.50e-01	16	17	18	19	20	21	22	23	24
1.25e-01	16	17	18	19	20	21	22	23	24
6.25e-02	21	21	21	22	22	23	23	23	25
3.13e-02	-	38	36	34	29	30	27	24	26
1.56e-02	-	-	44	46	54	53	53	45	32
7.81e-03	-	-	57	59	63	70	75	65	68
3.91e-03	-	-	63	63	66	74	81	78	106
1.95e-03	-	-	-	66	67	78	87	84	108
9.77e-04	-	-	-	76	76	83	94	88	111
4.88e-04	-	-	-	81	82	90	95	97	118
2.44e-04	-	-	-	-	88	91	99	101	121
1.22e-04	-	-	-	-	94	99	104	103	123
6.10e-05	-	-	-	-	-	103	110	106	128
3.05e-05	-	-	-	-	-	110	115	113	133
1.53e-05	-	-	-	-	-	-	128	124	144
7.63e-06	-	-	-	-	-	-	-	132	152
3.81e-06	-	-	-	-	-	-	-	-	159

In Tab. 1, the number of samples for each model order and value for γ that is attained during optimization is given. Note that we only show the γ values as long as the bisection leads to a decreasing value, since no points are deleted, when γ is increased. The increased number of sample points for larger reduced model orders in the first row is due to the initialization, which can use one more interpolation point as r is increased by two. It is remarkable that despite the fact that γ decreases over several orders of magnitude, the number of sample points stays moderate and typically only a few sample points are added as γ is halved.

This only slight increase in the number of samples is due to our iterative decrease of γ . Since the accuracy of the ROM is gradually increased, the criterion in line 11 of Alg. 1, requires fewer sample points even for a smaller tolerance (which is always set to γ). If we, instead, were to fix γ to the final accuracy and consult Alg. 1 for the sample points based on the initial ROM, the number of required sample points is drastically higher, from 38 samples for $r = 4$, to 8962 for $r = 12$, and to 160536 samples for $r = 20$. Such a large number of sample points would make the evaluation of L and thus its gradient-based minimization prohibitively expensive. This underlines the importance of combining the adaptive sampling strategy with our iterative reduction of γ in our objective function L .

Table 2. Runtime comparison between Alg. 2 and its fixed sampling counterpart.

r	runtime time in seconds		
	fixed sampling	Alg. 2	speed up ratio
4	4.50e+01	2.40e+00	18.7
6	1.61e+02	8.44e+00	19.1
8	1.66e+02	6.06e+01	2.73
10	9.26e+02	1.29e+02	7.16
12	9.46e+02	1.94e+02	4.87
14	1.13e+03	4.76e+02	2.37
16	4.21e+03	3.09e+02	13.6
18	3.64e+03	1.10e+03	3.32
20	5.39e+03	1.58e+03	3.40

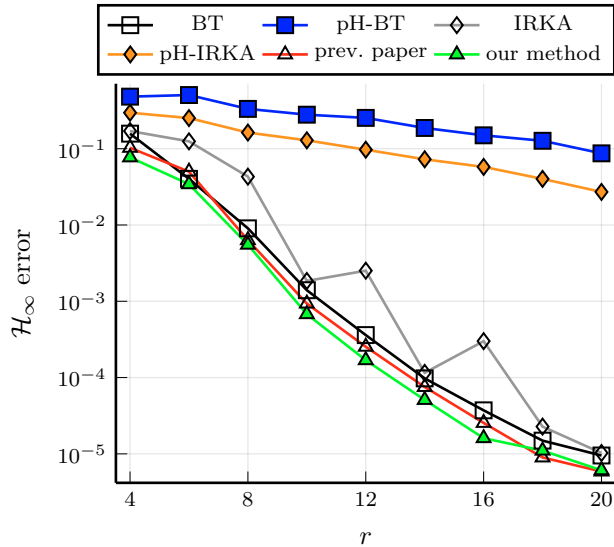


Fig. 4. $\mathcal{H}_\infty$ error comparison

In Tab. 2, the runtimes of Alg. 2 for different reduced model orders are reported and compared to a fixed sampling variant of this algorithm, in which around 800 logarithmically spaced sample points are chosen before the optimization is started. The speedup ranges from around $3\times$ to $19\times$ and we note that for the two smallest reduced model orders the speedup is the largest. For each reduced order except for $r = 18$ the number of optimizations (line 5 in Alg. 2) was the same for both the fixed and the adaptive sampling approach. Also the resulting $\mathcal{H}_\infty$ errors, which are studied in Fig. 4, are approximately the same.

Fig. 4 emphasizes that our new adaptive method leads to a similar accuracy compared to our previous method with a large fixed number of samples. We also mention that our method yields structure-preserving ROMs with an orders-of-magnitude higher accuracy compared to the other structure-preserving MOR algorithms pH-BT and pH-IRKA and even outperforms the non-structure-preserving methods BT and IRKA for all reduced model orders considered in this comparison.

5. CONCLUSION

We have equipped our optimization-based structure-preserving MOR algorithm from Schwerdtner and Voigt (2020) with an adaptive sampling strategy to increase its usability and lower its computational footprint. In our numerical experiments, we have demonstrated that combin-

ing this adaptive sampling with our reduction algorithm, which is based on an objective function that iteratively increases the overall accuracy, is beneficial to keep the number of samples low while retaining a high accuracy.

Future research includes extending the method to parameter-dependent systems as well as to additional model structures. Finally, the improved accuracy (also compared to the unstructured MOR algorithms) motivates studying the accuracy of our method also for unstructured systems.

REFERENCES

- Antoulas, A.C. (2005). *Approximation of Large-Scale Dynamical Systems*, volume 6 of *Adv. Des. Control*. SIAM Publications, Philadelphia, PA.
- Apkarian, P. and Noll, D. (2018). Structured $\mathcal{H}_\infty$ -control of infinite-dimensional systems. *Internat. J. Robust Nonlinear Control*, 28(9), 3212–3238.
- Beattie, C., Mehrmann, V., and Van Dooren, P. (2019). Robust port-Hamiltonian representations of passive systems. *Automatica J. IFAC*, 100, 182–186.
- Beattie, C.A., Mehrmann, V., and Xu, H. (2015). Port-Hamiltonian realizations of linear time invariant systems. Preprint 23-2015, Institute of Mathematics, Technische Universität Berlin.
- Beddig, R.S., Benner, P., Dorschky, I., Reis, T., Schwerdtner, P., Voigt, M., and Werner, S.W. (2019). Model reduction for second-order dynamical systems revisited. *PAMM Proc. Appl. Math. Mech.*, 19(1), e201900224.
- Breiten, T. and Unger, B. (2021). Passivity preserving model reduction via spectral factorization. Preprint arXiv:2103.13194.
- Desai, U.B. and Pal, D. (1984). A transformation approach to stochastic model reduction. *IEEE Trans. Automat. Control*, 29(12), 1097–1100.
- Gugercin, S., Polyuga, R.V., Beattie, C., and van der Schaft, A. (2012). Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems. *Automatica J. IFAC*, 48(9), 1963–1974.
- Mogensen, P.K. and Riseth, A.N. (2018). Optim: A mathematical optimization package for Julia. *J. Open Source Softw.*, 3(24), 615–618.
- Polyuga, R.V. and van der Schaft, A.J. (2008). Structure preserving model reduction of port-Hamiltonian systems. In *Proceedings of the 18th International Symposium on Mathematical Theory of Networks and Systems*. Blacksburg, VA.
- Schwerdtner, P. and Voigt, M. (2018). Computation of the $\mathcal{L}_\infty$ -norm using rational interpolation. *IFAC-PapersOnLine*, 51(25), 84–89. 9th IFAC Symposium on Robust Control Design 2018, Florianópolis, Brazil.
- Schwerdtner, P. and Voigt, M. (2020). Structure preserving model order reduction by parameter optimization. Preprint arXiv:2011.07567.
- Serhani, A., Matignon, D., and Haine, G. (2019). A partitioned finite element method for the structure-preserving discretization of damped infinite-dimensional port-Hamiltonian systems with boundary control. In F. Nielsen and F. Barbaresco (eds.), *Geometric Science of Information*, 549–558. Springer International Publishing, Cham.
- van der Schaft, A. and Jeltsema, D. (2014). Port-Hamiltonian systems theory: An introductory overview. *Found. Trends Systems Control*, 1(2–3), 173–378.