STRUCTURE-PRESERVING MODEL REDUCTION FOR NONLINEAR PORT-HAMILTONIAN SYSTEMS*

S. CHATURANTABUT†, C. BEATTIE‡, AND S. GUGERCIN‡

Abstract. This paper presents a structure-preserving model reduction approach applicable to large-scale, nonlinear port-Hamiltonian systems. Structure preservation in the reduction step ensures the retention of port-Hamiltonian structure which, in turn, assures the stability and passivity of the reduced model. Our analysis provides *a priori* error bounds for both state variables and outputs. Three techniques are considered for constructing bases needed for the reduction: one that utilizes proper orthogonal decompositions; one that utilizes $\mathcal{H}_2/\mathcal{H}_\infty$-derived optimized bases; and one that is a mixture of the two. The complexity of evaluating the reduced nonlinear term is managed efficiently using a modification of the discrete empirical interpolation method (DEIM) that also preserves port-Hamiltonian structure. The efficiency and accuracy of this model reduction framework are illustrated with two examples: a nonlinear ladder network and a tethered Toda lattice.

Key words. nonlinear model reduction, proper orthogonal decomposition, port-Hamiltonian, $\mathcal{H}_2$ approximation, structure preservation

AMS subject classifications. 37M05, 65P10, 93A15

1. Introduction and Background. The modeling of complex physical systems often involves systems of coupled partial differential equations, which upon spatial discretization, lead to dynamical models and systems of *ordinary* differential equations with very large state-space dimension. This motivates *model reduction* methods that produce low dimensional surrogate models capable of mimicking the input/output behavior of the original system model. Such reduced-order models could then be used as proxies, replacing the original system model in various computationally intensive contexts that are sensitive to system order, for example as a component in a larger simulation. Dynamical systems frequently have structural features that reflect underlying physics and conservation laws characteristic of the phenomena modeled. Reduced models that do not share such key structural features with the original system may produce response artifacts that are "unphysical" and as a result, such reduced models may be unsuitable for use as dependable surrogates for the original system, even if they otherwise yield high response fidelity. The key system feature that we wish to retain in our reduced models will be port-Hamiltonian structure. In a certain sense, this will be an expression of system passivity.

1.1. Port-Hamiltonian systems. Models of dynamic phenomena may be constructed within a system-theoretic network modeling paradigm that formalizes the interconnection of naturally specified subsystems. If the core dynamics of subsystem components are described by variational principles (e.g., least-action or virtual work), the aggregate system model typically has structural features that characterize it as a *port-Hamiltonian system*. While greater generality is both possible and useful (see, in particular, the review article [28], and the monographs [8] and [30]), it will suffice to consider realizations of finite-dimensional nonlinear port-Hamiltonian (NLPH) systems

---

†Department of Mathematics and Statistics,Thammasat University, Pathumthani, 12120, Thailand. `saifon@mathstat.sci.tu.ac.th`

‡Department of Mathematics, Virginia Polytechnic Institute and State University, 460 McBryde, Virginia Tech, Blacksburg, VA 24061-0123. `{beattie,gugercin}@math.vt.edu`

⁰DRAFT: January 5, 2016

that appear as:

$$\dot{\mathbf{x}} = (\mathbf{J} - \mathbf{R}) \nabla_{\mathbf{x}} H(\mathbf{x}) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y} = \mathbf{B}^T \nabla_{\mathbf{x}} H(\mathbf{x}), \tag{1.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the n-dimensional state vector; $H : \mathbb{R}^n \to [0, \infty)$ is a continuously differentiable scalar-valued vector function - the *Hamiltonian*, describing the internal energy of the system as a function of state; $\mathbf{J} = -\mathbf{J}^T \in \mathbb{R}^{n \times n}$ is the *structure matrix* describing the interconnection of energy storage elements in the system; $\mathbf{R} = \mathbf{R}^T \geq \mathbf{0}$ is the $n \times n$ *dissipation matrix* describing energy loss in the system; and, $\mathbf{B} \in \mathbb{R}^{n \times m}$ is the *port matrix* describing how energy enters and exits the system. We will always assume that the Hamiltonian is bounded below and so without loss of generality, strictly positive, $H(\mathbf{x}) > 0$ for all $\mathbf{x}$.

The family of systems characterized by (1.1) generalizes the classical notion of *Hamiltonian systems* which would be expressed in our notation as $\dot{\mathbf{x}} = \mathbf{J}\nabla_{\mathbf{x}}H(\mathbf{x})$. The analog of conservation of energy for Hamiltonian systems becomes for (1.1):

$$H(\mathbf{x}(t_1)) - H(\mathbf{x}(t_0)) \leq \int_{t_0}^{t_1} \mathbf{y}(t)^T \mathbf{u}(t) \; dt, \tag{1.2}$$

which is to say, the change in the internal energy of the system, as measured by H, is bounded by the total work done on the system. The presumed positivity of H conforms with its use as a "supply function" in the sense of Willems [29], and so NLPH systems are always *stable* and *passive*. Furthermore, the class of NLPH systems defined as in (1.1) is closed under power conserving interconnection - connecting port-Hamiltonian systems together produces an aggregate system that must also be port-Hamiltonian, and hence *a fortiori*, must be both stable and passive. This last fact provides compelling motivation to preserve port-Hamiltonian structure when producing low-order surrogate models intended to be used as proxies for systems of the sort defined by (1.1).

We assume in all that follows that the matrices $\mathbf{J}$, $\mathbf{R}$, and $\mathbf{B}$ are *constant*, however this assumption is adopted here largely for convenience. $\mathbf{J}$, $\mathbf{R}$, and $\mathbf{B}$ may each depend on the state vector, $\mathbf{x}$, input vector, $\mathbf{u}$, and may also carry an explicit time dependence, all without introducing any complications to port-Hamiltonian aspects of system structure; the dissipation inequality (1.2) will still hold and the system remains passive. For this reason, port-Hamiltonian systems can accommodate a very rich variety of nonlinear interactions. Moreover, the structure-preserving strategies for model reduction that we describe below may be adapted with negligible modification in this more complex setting.

1.2. Petrov-Galerkin reduced models. Most model reduction approaches involve some variation of a Petrov-Galerkin projective approximation to the equations describing the system dynamics. This proceeds by choosing two subspaces of $\mathbb{R}^n$: an r-dimensional trial subspace, $\mathcal{V}_r \subset \mathbb{R}^n$, and an r-dimensional test subspace, $\mathcal{W}_r \subset \mathbb{R}^n$. It is convenient and nonrestrictive in practice to assume additionally that $\mathcal{V}_r$ and $\mathcal{W}_r$ have a "generic orientation" with respect to one another so that neither subspace contains any nontrivial vectors that are orthogonal to all vectors in the other subspace. The evolution of an associated reduced-order model may be described in the following

(initially indirect) way:

> Find a trajectory, $\mathbf{v}_r(t)$, contained in $\mathcal{V}_r$ such that
> $$\dot{\mathbf{v}}_r(t) - (\mathbf{J} - \mathbf{R}) \nabla_{\mathbf{x}} H(\mathbf{v}_r) - \mathbf{B}\mathbf{u}(t) \quad \perp \quad \mathcal{W}_r; \qquad (1.3)$$
> the associated output is $\mathbf{y}_r(t) = \mathbf{B}^T \nabla_{\mathbf{x}} H(\mathbf{v}_r)$.

The dynamics described by (1.3) can be represented directly as a dynamical system evolving in a state-space of reduced dimension r once bases are chosen for the two subspaces $\mathcal{V}_r$ and $\mathcal{W}_r$. Let $\mathsf{Ran}(\mathbf{M})$ denote the range of a matrix $\mathbf{M}$. Define matrices $\mathbf{V}_r, \mathbf{W}_r \in \mathbb{R}^{n \times r}$ so that $\mathcal{V}_r = \mathsf{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \mathsf{Ran}(\mathbf{W}_r)$. We can represent reduced system trajectories as $\mathbf{v}_r(t) = \mathbf{V}_r \mathbf{x}_r(t)$ with $\mathbf{x}_r(t) \in \mathbb{R}^r$ for each t; the Petrov-Galerkin approximation (1.3) can be rewritten as

$$\mathbf{W}_r^T \left[\mathbf{V}_r \dot{\mathbf{x}}_r(t) - (\mathbf{J} - \mathbf{R}) \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r) - \mathbf{B}\mathbf{u}(t) \right] = \mathbf{0}$$
$$\text{and} \quad \mathbf{y}_r(t) = \mathbf{B}^T \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r).$$

Since $\mathcal{V}_r$ and $\mathcal{W}_r$ are assumed to have a generic orientation with respect to one another, $\mathbf{W}_r^T \mathbf{V}_r$ is invertible and we may choose bases for $\mathcal{V}_r$ and $\mathcal{W}_r$ such that $\mathbf{W}_r^T \mathbf{V}_r = \mathbf{I}$. This leads to a state-space representation of a reduced-order nonlinear dynamical system approximating (1.1):

$$\dot{\mathbf{x}}_r = \mathbf{W}_r^T (\mathbf{J} - \mathbf{R}) \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r) + \mathbf{W}_r^T \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}_r = \mathbf{B}^T \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r), \qquad (1.4)$$

Typically $r \ll n$ and (1.4) describes a reduced-order model for the original system (1.1). There are two shortcomings that may be anticipated. First, (1.4) will not have the form of (1.1) unless special subspaces are chosen, and so, (1.4) will not typically be a NLPH system and passivity may be lost. Secondly, if the Hamiltonian function, H, is non-quadratic, each evaluation of $\nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r)$ in (1.4) occurring in the course of a simulation will likely require a lifting of $\mathbf{x}_r$ to $\mathbb{R}^n$ (implicit in the formation of $\mathbf{V}_r \mathbf{x}_r$), and so direct simulation of (1.4) is still likely to have complexity proportional to $n \gg r$; little or no savings may be realized from reducing the system order.

We consider each of these issues in subsequent sections. In §2, a structure-preserving model reduction approach for large-scale NLPH systems will be introduced. This approach is built upon Petrov-Galerkin projections that are modified to assure that the resulting reduced system retains port-Hamiltonian structure; thus stability and passivity. Three types of reduced-order bases used to define these projections will be considered: (i) one based on the Proper Orthogonal Decomposition (POD), (ii) one derived from $\mathcal{H}_2$-optimal approaches for a related linear problem (which we refer to as "$\mathcal{H}_2^\varepsilon$-bases"), (iii) hybrid $\mathcal{H}_2^\varepsilon$- POD bases that combine both types. The bases (i) and (ii) were originally considered in [3]. Numerical experiments in §2.4 illustrate that the hybrid $\mathcal{H}_2^\varepsilon$- POD bases significantly outperform the other two. In §2.5, we develop corresponding error analyses and bounds for reduced states and outputs. In order to resolve the "lifting bottleneck" described above, we develop, in §3, a variant of the Discrete Empirical Interpolation Method (DEIM) [6] that incorporates the structure-preserving model reduction approach of §2. In §3.4, the efficiency and accuracy of our approach are illustrated with two examples: a nonlinear ladder network and a tethered Toda lattice. Corresponding *a priori* error bounds for states and outputs are derived in §3.5.

2. Preserving port-Hamiltonian Structure in Reduced Models. The process of obtaining a reduced model from an original full-order model can be viewed as one of identifying and preserving high-value portions of the state space, i.e., portions of the state space that contribute substantively to the system response. We proceed with the following heuristics: Suppose we have identified two r-dimensional subspaces, $\mathcal{V}_r$ and $\mathcal{W}_r$, that are "high-value" in the sense that for "most" input signal profiles, $\mathbf{u}(t)$, in (1.1), we have

$$
\text{HIGH VALUE PH-SPACES:} \quad
\begin{cases}
\bullet \ \text{The associated trajectory, } \mathbf{x}(t), \text{ stays "close" to } \mathcal{V}_r, \text{ so that} \\
\quad \mathbf{x}(t) \approx \mathbf{v}_r(t) \text{ for some } \mathbf{v}_r(t) \in \mathcal{V}_r, \text{ and} \\[4pt]
\bullet \ \text{the internal force, } \nabla_{\mathbf{x}} H(\mathbf{x}), \text{ stays "close" to } \mathcal{W}_r, \text{ so that} \\
\quad \nabla_{\mathbf{x}} H(\mathbf{x}(t)) \approx \mathbf{w}_r(t) \text{ for some } \mathbf{w}_r(t) \in \mathcal{W}_r.
\end{cases}
\tag{2.1}
$$

Evidently if $\mathcal{V}_r = \mathsf{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \mathsf{Ran}(\mathbf{W}_r)$, then $\mathbf{v}_r(t) = \mathbf{V}_r \mathbf{x}_r(t)$ for some trajectory $\mathbf{x}_r \in \mathbb{R}^r$ and $\mathbf{w}_r(t) = \mathbf{W}_r \mathbf{f}_r(t)$, for some choice of $\mathbf{f}_r \in \mathbb{R}^r$. Exactly what comprises "most" input signal profiles and how one measures "closeness" in (2.1) will vary depending on context; we consider different possibilities later in this section. For the time being, note that if $\mathbf{x}(t) \approx \mathbf{V}_r \mathbf{x}_r(t)$ then plausibly,

$$
\nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r(t)) \approx \nabla_{\mathbf{x}} H(\mathbf{x}(t)) \approx \mathbf{W}_r \mathbf{f}_r(t).
$$

Significantly, these statements amount to assertions about the *subspaces*, $\mathcal{V}_r$ and $\mathcal{W}_r$, and do not constrain the choice of bases for the subspaces. As long as $\mathcal{V}_r$ and $\mathcal{W}_r$ have a generic orientation with respect to one another, bases may be chosen so that $\mathbf{W}_r^T \mathbf{V}_r = \mathbf{I}$. With this in mind, note further that

$$
\mathbf{f}_r(t) = \mathbf{V}_r^T \mathbf{W}_r \mathbf{f}_r(t) \approx \mathbf{V}_r^T \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r(t)) = \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r(t)),
\tag{2.2}
$$

where we have introduced a *reduced Hamiltonian*, $H_r(\mathbf{x}_r) = H(\mathbf{V}_r \mathbf{x}_r)$. Thus,

$$
\nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r(t)) \approx \mathbf{W}_r \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r(t)).
\tag{2.3}
$$

Substituting $\mathbf{V}_r \mathbf{x}_r(t)$ for $\mathbf{x}(t)$ and $\mathbf{W}_r \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r(t))$ for $\nabla_{\mathbf{x}} H(\mathbf{x}(t))$ in (1.1) and then multiplying by $\mathbf{W}_r^T$, leads to a state-space representation of a reduced port-Hamiltonian approximation:

$$
\begin{aligned}
\dot{\mathbf{x}}_r &= (\mathbf{J}_r - \mathbf{R}_r) \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r) + \mathbf{B}_r \mathbf{u}(t), \\
\mathbf{y}_r(t) &= \mathbf{B}_r^T \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r)
\end{aligned}
\tag{2.4}
$$

with $H_r(\mathbf{x}_r) = H(\mathbf{V}_r \mathbf{x}_r)$, $\mathbf{J}_r = \mathbf{W}_r^T \mathbf{J} \mathbf{W}_r$, $\mathbf{R}_r = \mathbf{W}_r^T \mathbf{R} \mathbf{W}_r$, and $\mathbf{B}_r = \mathbf{W}_r^T \mathbf{B}$. Note that $H_r : \mathbb{R}^r \to [0, \infty)$ is continuously differentiable; $\mathbf{J}_r = -\mathbf{J}_r^T$; and $\mathbf{R}_r = \mathbf{R}_r^T \geq \mathbf{0}$, so (2.4) retains the structure of (1.1) and is a port-Hamiltonian system.

The earlier works, [10] and [24], also offer structure-preserving model reduction methods for NLPH systems. These papers exploit the Kalman decomposition and balanced truncation in deriving reduced models of NLPH systems. Obtaining the Kalman decomposition of the full-order original system or balancing it, is computationally demanding for nonlinear systems of even modest order; see e.g. [11, 24] and references therein. Such approaches are infeasible for the problem class we consider, which may have thousands of state-variables. In what follows, we develop approaches that remain feasible for this problem class; they depend on the construction of low-dimensional projecting subspaces motivated by the heuristics in (2.1) .

Algorithm 1 : Structure-preserving POD reduction of NLPH systems (POD-PH)

1: Generate a trajectory $\mathbf{x}(t)$, and collect snapshots:
$\mathbb{X} = [\mathbf{x}(t_0), \mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_N)]$

2: Simultaneously collect associated *force* snapshots:
$\mathbb{F} = [\nabla_{\mathbf{x}}H(\mathbf{x}(t_0)), \nabla_{\mathbf{x}}H(\mathbf{x}(t_1)), \ldots, \nabla_{\mathbf{x}}H(\mathbf{x}(t_N))]$.

3: Truncate an SVD of the snapshot matrix, $\mathbb{X}$, to get a POD basis, $\widetilde{\mathbf{V}}_r$, for a "high-value" subspace of the state space. $(\mathbf{x}(t) \approx \widetilde{\mathbf{V}}_r \tilde{\mathbf{x}}_r(t))$

4: Truncate an SVD of $\mathbb{F}$ to get a second POD basis, $\widetilde{\mathbf{W}}_r$, spanning a second "high-value" subspace approximating the range of $\nabla_{\mathbf{x}}H(\mathbf{x}(t)) \approx \widetilde{\mathbf{W}}_r \tilde{\mathbf{f}}_r(t)$.

5: Change bases $\widetilde{\mathbf{W}}_r \mapsto \mathbf{W}_r$ and $\widetilde{\mathbf{V}}_r \mapsto \mathbf{V}_r$ such that $\mathbf{W}_r^T \mathbf{V}_r = \mathbf{I}$.

6: With $\mathbf{V}_r$ and $\mathbf{W}_r$ determined in this way, the POD-PH reduced port-Hamiltonian system is then specified by (2.4).

2.1. POD subspaces. The Proper Orthogonal Decomposition (POD) is a natural approach to producing high-value modeling spaces as described in (2.1). POD is a popular approach to (unstructured) model reduction ([16, 25]) which we adapt to our setting as follows: Fix a square integrable input signal, $\mathbf{u}(t)$, for the system (1.1). The corresponding trajectory, $\mathbf{x}(t)$, will then also be square integrable. Denoting orthogonal projections, $\mathbf{P}$ and $\mathbf{Q}$, we consider the minimization problem:

$$\mathbf{P}_\star = \underset{\mathsf{rank}(\mathbf{P}) = r}{\operatorname{argmin}} \int_0^\infty \| (\mathbf{I} - \mathbf{P}) \, \mathbf{x}(t) \|^2 \, dt$$

and

$$\mathbf{Q}_\star = \underset{\mathsf{rank}(\mathbf{Q}) = r}{\operatorname{argmin}} \int_0^\infty \| (\mathbf{I} - \mathbf{Q}) \, \nabla_{\mathbf{x}}H(\mathbf{x}(t)) \|^2 \, dt.$$

We would like to take $\mathcal{V}_r = \mathsf{Ran}(\mathbf{P}_\star)$ and $\mathcal{W}_r = \mathsf{Ran}(\mathbf{Q}_\star)$, but this is not a computationally tractable approach as it stands. If the integrals are truncated and then approximated with a Trapezoid Rule, one arrives at a characterization of POD subspaces which is incorporated into our first method, summarized as *Algorithm 1*. As our numerical results show (and consistent with common experience), the use of POD in *Algorithm 1* provides subspaces $\mathcal{V}_r$ and $\mathcal{W}_r$ that can be very effective in capturing dynamic features that are present in the original sampled system response. However, as an empirical method, it is incapable of providing information about dynamic response features that are absent in the sampled system response, but that *could* have been present had a different choice of input profile been made. One can mitigate this difficulty somewhat by extending POD by including a representative sampling of input profiles, but this leaves open the question of what constitutes a representative sampling of input profiles. A different approach leads to our next class of subspaces.

2.2. $\mathcal{H}_2^\varepsilon$-optimal subspaces. We next consider a choice of subspaces, $\mathcal{V}_r$ and $\mathcal{W}_r$, that are optimal (in a sense we will describe) for *all possible* input profiles that are sufficiently small ("ε-optimal"). It is often the case that this choice will be effective for larger input profiles as well. In contrast to the previous POD approach, no choice of inputs is necessary and no simulations need to be performed in order to derive the approximating subspaces. The proviso that (virtual) inputs be sufficiently small, but

otherwise arbitrary, allows us to tailor the subspaces $\mathcal{V}_r$ and $\mathcal{W}_r$, so as to provide near-optimal reduction for the corresponding *linearized* port-Hamiltonian model. Note that linearization is used here only as a tool to obtain useful information that will be encoded into the projection subspaces, $\mathcal{V}_r$ and $\mathcal{W}_r$, which are then used for reduction of the *nonlinear* system (1.1).

Any input profile, $\mathbf{u}(t)$, may be scaled to have sufficiently small magnitude so that the resulting trajectory, $\mathbf{x}(t)$, is small as well. Then linear terms in the internal forcing dominate, and $\nabla_{\mathbf{x}} H(\mathbf{x}) \approx \mathbf{Q}\mathbf{x}$ for some symmetric positive definite matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$. Indeed, $\mathbf{Q} = \nabla^2 H(\mathbf{0})$, the Hessian matrix for $H(\mathbf{x})$ evaluated at $\mathbf{x} = \mathbf{0}$. This approximation leads to an ancillary *linear port-Hamiltonian system*:

$$\dot{\mathbf{x}} = (\mathbf{J} - \mathbf{R})\,\mathbf{Q}\,\mathbf{x} + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y} = \mathbf{B}^T \mathbf{Q}\,\mathbf{x}. \tag{2.5}$$

We proceed to construct projecting subspaces, $\mathcal{V}_r$ and $\mathcal{W}_r$, that would produce, via the reduction described in (2.4), an effective reduced port-Hamiltonian model approximating (2.5):

$$\dot{\mathbf{x}}_r = (\mathbf{J}_r - \mathbf{R}_r)\mathbf{Q}_r\mathbf{x}_r + \mathbf{B}_r\mathbf{u}(t),$$
$$\mathbf{y}_r(t) = \mathbf{B}_r^T \mathbf{Q}_r\mathbf{x}_r, \tag{2.6}$$

where $\mathbf{J}_r = \mathbf{W}_r^T \mathbf{J} \mathbf{W}_r$, $\mathbf{R}_r = \mathbf{W}_r^T \mathbf{R} \mathbf{W}_r$, $\mathbf{Q}_r = \mathbf{V}_r^T \mathbf{Q} \mathbf{V}_r$ and $\mathbf{B}_r = \mathbf{W}_r^T \mathbf{B}$. Observe that $\mathbf{J}_r = -\mathbf{J}_r^T$, $\mathbf{R}_r = \mathbf{R}_r^T \geq 0$ and $\mathbf{Q}_r = \mathbf{Q}_r^T > 0$. The quality of the subspaces is interpreted now as how well (2.6) approximates (2.5). This, in turn, may be interpreted as a rational approximation problem: We associate the linear dynamical systems (2.5) and (2.6) with their transfer functions,

$$\mathcal{G}(s) = \mathbf{B}^T(s\mathbf{Q}^{-1} - (\mathbf{J} - \mathbf{R}))^{-1}\mathbf{B} \quad \text{and}$$
$$\mathcal{G}_r(s) = \mathbf{B}_r^T(s\mathbf{Q}_r^{-1} - (\mathbf{J}_r - \mathbf{R}_r))^{-1}\mathbf{B}_r, \tag{2.7}$$

respectively. If $\mathcal{G}_r(s)$ approximates $\mathcal{G}(s)$ well with respect to some (appropriately chosen) norm, then the reduced system outputs will approximate the full order system outputs uniformly well over all inputs with bounded energy (square integrable); our model reduction problem has been reduced to a rational approximation problem. We wish to find a degree-r rational function, $\mathcal{G}_r(s)$ having the structure given in (2.7) that also approximates $\mathcal{G}(s)$ well. Tangential rational interpolation provides a useful tool for this: Given r interpolation points $\sigma_1, \dots, \sigma_r$ in the complex plane with corresponding tangent directions $\{\mathbf{b}_1, \dots, \mathbf{b}_r\} \in \mathbb{C}^m$, construct an rth order system, $\mathcal{G}_r$, so that $\mathcal{G}_r$ is port-Hamiltonian, and

$$\mathcal{G}_r(\sigma_i)\mathbf{b}_i = \mathcal{G}(\sigma_i)\mathbf{b}_i \quad \text{for} \quad i = 1, \dots, r. \tag{2.8}$$

A solution to this problem was given in [14].

THEOREM 2.1. *Given interpolation points $\sigma_1, \dots, \sigma_r$ and tangent directions $\mathbf{b}_1, \dots, \mathbf{b}_r$, construct*

$$\widetilde{\mathbf{V}}_r = [(\sigma_1\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_1, \dots, (\sigma_r\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_r].$$

Define the Cholesky factorization of $\widetilde{\mathbf{V}}_r^T \mathbf{Q} \widetilde{\mathbf{V}}_r = \mathbf{R}^T \mathbf{R}$, assign $\mathbf{V}_r = \widetilde{\mathbf{V}}_r \mathbf{R}^{-1}$, and then construct $\mathbf{W}_r = \mathbf{Q}\mathbf{V}_r$. Set

$$\mathbf{J}_r = \mathbf{W}_r^T \mathbf{J} \mathbf{W}_r, \; \mathbf{Q}_r = \mathbf{V}_r^T \mathbf{Q} \mathbf{V}_r = \mathbf{I}_r, \; \mathbf{R}_r = \mathbf{W}_r^T \mathbf{R} \mathbf{W}_r, \quad \text{and} \quad \mathbf{B}_r = \mathbf{W}_r^T \mathbf{B}. \tag{2.9}$$

Then the reduced model, $\mathcal{G}_r$:
$$\begin{aligned} \dot{\mathbf{x}}_r &= (\mathbf{J}_r - \mathbf{R}_r)\mathbf{Q}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u} \\ \mathbf{y}_r &= \mathbf{B}_r^T\mathbf{Q}_r\mathbf{x}_r \end{aligned} \tag{2.10}$$

is port-Hamiltonian (hence stable and passive) and also satisfies the interpolation conditions (2.8).

For information on transfer function interpolation in the special case of single-input/single-output port-Hamiltonian systems, see [21, 13, 22]. For an overview of model reduction methods for linear port-Hamiltonian systems, see [20].

2.2.1. $\mathcal{H}_2^\varepsilon$ port-Hamiltonian approximation. port-Hamiltonian approximations of reduced order may be constructed using Theorem 2.1 once shifts, $\{\sigma_i\}$, and tangent directions, $\{\mathbf{b}_i\}$, are chosen, but this does not give any information on how best to choose $\{\sigma_i\}$ and $\{\mathbf{b}_i\}$. We discuss issues related to finding effective approximations with respect to the $\mathcal{H}_2$ norm: The $\mathcal{H}_2$ norm of $\mathcal{G}$ is defined as

$$\|\mathcal{G}\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathcal{G}(\imath\omega)\|_{\mathrm{F}}^2 \, d\omega \right)^{1/2}.$$

Let $\mathcal{G}_r(s)$ minimize the $\mathcal{H}_2$ error $\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_2}$ over all possible degree-r rational functions and suppose that $\mathcal{G}_r(s)$ has a partial fraction expansion $\mathcal{G}_r(s) = \sum_{k=1}^{r} \frac{\mathsf{c}_k\mathsf{b}_k^T}{s-\widehat{\lambda}_k}$. Then, as shown in [12], the interpolation conditions

$$\mathcal{G}(-\widehat{\lambda}_k)\mathsf{b}_k = \mathcal{G}_r(-\widehat{\lambda}_k)\mathsf{b}_k, \quad \text{for } k = 1, \ldots, r \tag{2.11}$$

are necessary conditions for $\mathcal{H}_2$ optimality (there are additional conditions that must also be satisfied in general). In other words, the optimal $\mathcal{H}_2$ approximant $\mathcal{G}_r$ is a tangential interpolant to $\mathcal{G}$ at the mirror images of the reduced-order poles. For the full set of necessary conditions required for optimality, see [12].

A method was introduced in [14] that produces an interpolatory reduced-order port-Hamiltonian system satisfying the conditions given in (2.11). Since the interpolation points $-\widehat{\lambda}_k$ and the tangential directions b_k depend on the reduced-model to be computed, an iterative process is used to correct the interpolation points and tangential directions until the desired conditions in (2.11) are obtained. These constitute only a subset of the necessary conditions required for $\mathcal{H}_2$-optimality. The remaining degrees of freedom are used in maintaining port-Hamiltonian structure. An algorithm that accomplishes was introduced in [14]. Algorithm 2 below uses this methodology to construct the model reduction bases for the second structure-preserving model reduction of NLPH systems.

Note that we use the linearized port-Hamiltonian model only to obtain the $\mathcal{H}_2^\varepsilon$-PH model reduction subspaces. Once $\mathbf{V}_r$ and $\mathbf{W}_r$ are obtained using the $\mathcal{H}_2^\varepsilon$-approach outlined in Algorithm 2, we use these subspaces to reduce the original nonlinear system as shown in (2.4). See [19, 23] for other approaches that derive useful information from linearized systems in order to reduce nonlinear systems.

2.3. A hybrid POD-$\mathcal{H}_2^\varepsilon$ approach. The model reduction subspaces that POD provides are effective in capturing the dynamics that are represented in the original snapshot data; but naturally will miss features that are absent in this data. To resolve this issue in part, we have proposed to use ε-optimal subspaces that were accurate for input profiles that are sufficiently small ("ε-optimal"). These subspaces are generated from a near-optimal reduction of a linearized port-Hamiltonian model using an approach proposed in . As the trajectory moves away from the linearization

---

Algorithm 2 : Structure-preserving $\mathcal{H}_2^\varepsilon$ reduction of NLPH systems ($\mathcal{H}_2^\varepsilon$-PH)

1: Linearize the NLPH system (1.1) to obtain a linear port-Hamiltonian system of the form in (2.5).

2: Make an initial selection of shifts $\{\sigma_i\}_1^r$, and tangent directions $\{\mathbf{b}_i\}_1^r$.

3: **while** (not converged) **do**

4:   $\widehat{\mathbf{V}}_r = [(\sigma_1\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_1, \ldots, (\sigma_r\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_r]$

5:   Set $\mathbf{V}_r = \widehat{\mathbf{V}}_r\mathbf{L}^{-1}$ with $\widehat{\mathbf{V}}_r^T\mathbf{Q}\widehat{\mathbf{V}}_r = \mathbf{L}^T\mathbf{L}$ (so $\mathbf{Q}_r = \mathbf{V}_r^T\mathbf{Q}\mathbf{V}_r = \mathbf{I}_r$).

6:   Set $\mathbf{W}_r = \mathbf{Q}\mathbf{V}_r$. (so $\mathbf{V}_r^T\mathbf{W}_r = \mathbf{I}_r$).

7:   Set $\mathbf{J}_r = \mathbf{W}_r^T\mathbf{J}\mathbf{W}_r$, $\mathbf{R}_r = \mathbf{W}_r^T\mathbf{R}\mathbf{W}_r$, and $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$.

8:   Calculate left eigenvectors: $\mathbf{z}_i^T(\mathbf{J}_r - \mathbf{R}_r) = \lambda_i\mathbf{z}_i^T$.

9:   Set $\sigma_i \longleftarrow -\lambda_i$ and $\mathbf{b}_i \longleftarrow \mathbf{B}_r^T\mathbf{z}_i$ for $i = 1, \ldots, r$

10: **end while**

   { Calculate final $\mathcal{H}_2^\varepsilon$-PH bases: }

11: Find $\widehat{\mathbf{V}}_r = [(\sigma_1\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_1, \ldots, (\sigma_r\mathbf{I} - (\mathbf{J} - \mathbf{R})\mathbf{Q})^{-1}\mathbf{B}\mathbf{b}_r]$

12: Set $\mathbf{V}_r = \widehat{\mathbf{V}}_r\mathbf{L}^{-1}$ with $\widehat{\mathbf{V}}_r^T\mathbf{Q}\widehat{\mathbf{V}}_r = \mathbf{L}^T\mathbf{L}$.

13: Set $\mathbf{W}_r = \mathbf{Q}\mathbf{V}_r$.

14: With $\mathbf{V}_r$ and $\mathbf{W}_r$ determined in this way, construct the reduced nonlinear PH system using (2.4).

---

point, the efficiency of these subspaces in capturing the true dynamics might degrade. Therefore, we propose to combine the subspaces resulting from POD (i.e., directly obtained from a simulation of the NLPH system) together with the ε-optimal subspaces resulting from the linearized model. The structure-preserving reduction is applied as in (2.4) with the aggregate subspace.

For a given reduced dimension r, let $\widehat{\mathbf{V}}_{\widehat{r}} = [\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \ldots, \widehat{\mathbf{v}}_{\widehat{r}}]$ be the POD-PH basis of dimension $\widehat{r}$ obtained from Algorithm 1 and let $\bar{\mathbf{V}}_{\bar{r}} = [\bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2, \ldots, \bar{\mathbf{v}}_{\bar{r}}]$ be the $\mathcal{H}_2^\varepsilon$-PH basis of dimension $\bar{r}$ obtained from Algorithm 2, where $\widehat{r}$ and $\bar{r}$ are positive integers with $\widehat{r} + \bar{r} = r$. Then, the hybrid basis $\mathbf{V}_r$ of dimension r can be obtained from the orthonormal basis of concatenated matrix $[\widehat{\mathbf{V}}_{\widehat{r}} \quad \bar{\mathbf{V}}_{\bar{r}}] \in \mathbb{R}^{n \times r}$ The other projection basis $\mathbf{W}_r$ can be constructed similarly by combining POD-PH and $\mathcal{H}_2^\varepsilon$-PH basis vectors. This leads to the third algorithm for structure-preserving model reduction of NLPH systems.

2.4. An illustrative example. To illustrate the structure-preserving model reduction techniques described in Section 2, we consider an N-stage nonlinear ladder network (see Figure 2.1) producing reduced models for three cases of bases: (i) POD bases (Algorithm POD-PH), (ii) $\mathcal{H}_2^\varepsilon$ bases (Algorithm $\mathcal{H}_2^\varepsilon$-PH), and (iii) hybrid POD-$\mathcal{H}_2^\varepsilon$ bases (Algorithm POD-$\mathcal{H}_2^\varepsilon$-PH). The system has two inputs and two outputs. The two inputs are a voltage signal applied to the left-hand terminal pair and a current injection across the right-hand terminal pair. The symmetrically paired outputs are the induced current across the left-hand terminal pair and the induced voltage signal across the right-hand terminal pair. For simplicity, we assume that each stage of

---

Algorithm 3 : Structure-preserving POD-$\mathcal{H}_2^\varepsilon$ reduction of NLPH (POD-$\mathcal{H}_2^\varepsilon$-PH)

1: Pick $\widehat{r}$ and $\bar{r} = r$ so that $\widehat{r} + \bar{r} = r$.

2: Obtain the POD-PH bases $\widehat{\mathbf{V}}_{\widehat{r}}$ and $\widehat{\mathbf{W}}_{\widehat{r}}$ of dimension $\widehat{r}$ using Algorithm 1.

3: Obtain the $\mathcal{H}_2^\varepsilon$-PH bases $\bar{\mathbf{V}}_{\bar{r}}$ and $\bar{\mathbf{W}}_{\bar{r}}$ of dimension $\bar{r}$ using Algorithm 2.

4: Construct $\widetilde{\mathbf{V}}_r$ of dimension $n \times r$ as the orthonormal basis of the concatenated matrix $[\widehat{\mathbf{V}}_{\widehat{r}} \ \ \bar{\mathbf{V}}_{\bar{r}}] \in \mathbb{R}^{n \times r}$

5: Construct $\widetilde{\mathbf{W}}_r$ of dimension $n \times r$ as the orthonormal basis of the concatenated matrix $[\widehat{\mathbf{W}}_{\widehat{r}} \ \ \bar{\mathbf{W}}_{\bar{r}}] \in \mathbb{R}^{n \times r}$

6: Change bases $\widetilde{\mathbf{W}}_r \mapsto \mathbf{W}_r$ and $\widetilde{\mathbf{V}}_r \mapsto \mathbf{V}_r$ such that $\mathbf{W}_r^T \mathbf{V}_r = \mathbf{I}$.

7: With $\mathbf{V}_r$ and $\mathbf{W}_r$ determined in this way, the POD-$\mathcal{H}_2^\varepsilon$ reduced port-Hamiltonian system is then specified by (2.4).

---

the ladder network is built from identical components and that the current injection from the right is zero. Resistors and inductors are assumed to behave linearly. The capacitors have a nonlinear C-V characteristic of the form $C_k(V) = \frac{C_0 V_0}{V_0 + V}$.

FIG. 2.1. *Ladder network circuit topology (capacitors are nonlinear)*
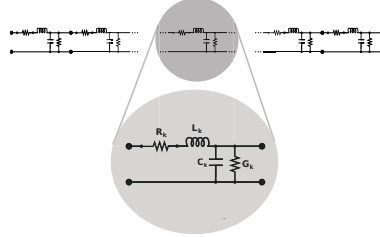
Inductors and capacitors are evidently the energy storage elements of the circuit, so we take as state variables the magnetic fluxes in the inductors, $\{\phi_k(t)\}_{k=1}^N$, and the charges on the capacitors, $\{Q_k\}_{k=1}^N$, with labels referring to the Stages $k = 1, \ldots, N$, respectively, where they occur. The energy stored in the Stage k (linear) inductor may be expressed in terms of its magnetic flux as $\frac{1}{2L_0}\phi_k^2$. To determine the energy stored in the nonlinear capacitors, note first that the charge on a capacitor may be expressed as a function of the voltage, V, held across the capacitor:

$$Q_k(V) = \int_0^V C(v)\,dv = C_0 V_0 \log\left(1 + \frac{V}{V_0}\right),$$

which may be inverted to find

$$V_k(Q_k) = V_0 \left[\exp\left(\frac{Q_k}{C_0 V_0}\right) - 1\right].$$

The energy stored in the capacitor at Stage k of the circuit, is then given by

$$\int_0^{Q_k} V_k(q)\,dq = C_0 V_0^2 \left[\exp\left(\frac{Q_k}{C_0 V_0}\right) - 1\right] - Q_k V_0,$$

and the total energy stored in Stage k is then

$$H^{[k]}(\phi_k, Q_k) = C_0 V_0^2 \left[\exp\left(\frac{Q_k}{C_0 V_0} \right) - 1 \right] - Q_k V_0 + \frac{1}{2 L_0} \phi_k^2.$$

The Hamiltonian for this system is

$$H(Q_1, \ldots, Q_N, \phi_1, \ldots, \phi_N) = \sum_{k=1}^{N} H^{[k]}(\phi_k, Q_k).$$

We order the state variables so that $\mathbf{x} = [Q_1, \ldots, Q_N, \phi_1, \ldots, \phi_N]^T$ and therefore $\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathsf{S} \\ -\mathsf{S}^T & \mathbf{0} \end{bmatrix}$ where S is an upper bidiagonal matrix with 1 on the diagonal and -1 on the superdiagonal; $\mathbf{R} = \begin{bmatrix} G_0 \mathbf{I} & 0 \\ 0 & R_0 \mathbf{I} \end{bmatrix}$; and $\mathbf{B} = [\mathbf{e}_{N+1}, \mathbf{e}_N]$ where $\mathbf{e}_k$ denotes the kth column of the identity. Consider the particular case of a 50-stage ($N = 50$) circuit with parameters: $L_0 = 2\mu\mathsf{H}$, $V_0 = 1\mathsf{V}$, $R_0 = 1\Omega$, $G_0 = 10\mu\mho$.

We applied a voltage pulse to the left port of the network (Gaussian pulse windowed to 3μsec with a magnitude of 3V, σ of 0.5) and observed the output voltage at the right port. The output is displayed as a solid green trace in Figure 2.2. The induced response of the *linearized* full-order network is also displayed (green dashed line) for comparison. Notice that nonlinearity sharpens the peak of the response and significantly reduces dispersion. The POD basis sets are generated from uniformly sampled snapshots $\mathbf{x}(t)$ and $\nabla_{\mathbf{x}} H(\mathbf{x}(t))$ of this full-order system with Gaussian impulse training input. These basis sets are then used to construct the POD-PH reduced system. In the cases of $\mathcal{H}_2^\varepsilon$-PHbases, the procedure described in Algorithm 2 is used. We also use Algorithm 3 to generate the POD-$\mathcal{H}_2^\varepsilon$-PHbases. All three reduced models are then simulated for the same Gaussian impulse training input, which was used for generating the POD snapshot as well as a different one, a sinusoidal input. First we investigate the accuracy of POD-PH and $\mathcal{H}_2^\varepsilon$-PH for $r = 6$. Figure 2.2 illustrates the two structure-preserving nonlinear reduced models capture the output of the original nonlinear system very accurately for both types of excitations.

FIG. 2.2. *Ladder network: Time responses of the reduced-order systems from Gaussian pulse (left) and from sinusoidal input(right).*
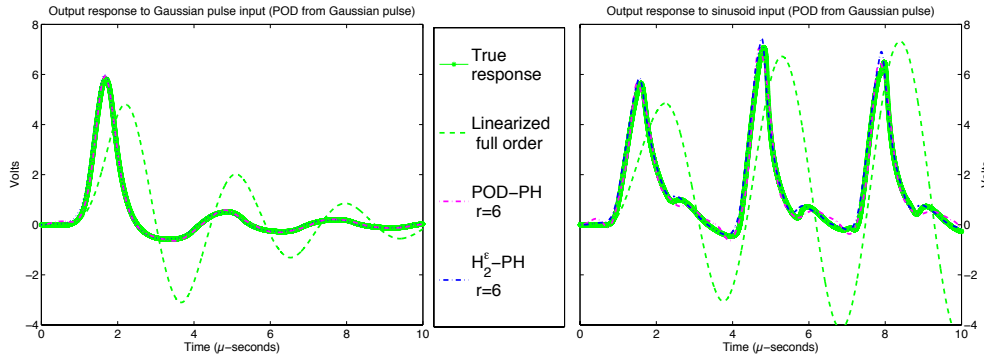
Next, we consider the effect on the accuracy of the state space solutions of using different proportions of POD-PH and $\mathcal{H}_2^\varepsilon$-PH basis vectors in the hybrid POD-$\mathcal{H}_2^\varepsilon$-PH basis.

FIG. 2.3. *Ladder Network: Average relative errors of outputs and state variables of structure-preserving reduced systems (2.4) using hybrid bases with different numbers of POD and $\mathcal{H}_2^\varepsilon$ vectors.*
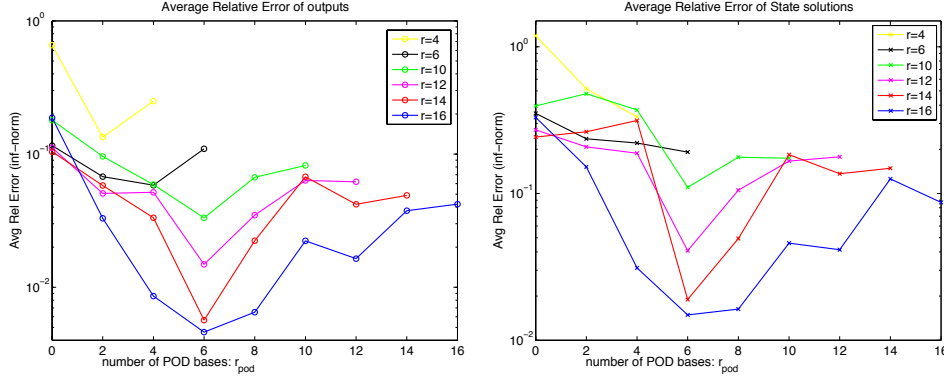

FIG. 2.4. *Ladder Network: Average relative errors of outputs and state variables of reduced systems using bases from (i) $\mathcal{H}_2^\varepsilon$, (ii) POD (iii) combination of POD and $\mathcal{H}_2^\varepsilon$ (with same number of basis vectors for both POD and $\mathcal{H}_2^\varepsilon$). The POD-PH reduced systems are constructed using Gaussian impulse training input. The top and the bottom plots, respectively, use the impulse input (same as the training input for POD) and with sinusoidal input (different from the training input for POD).*
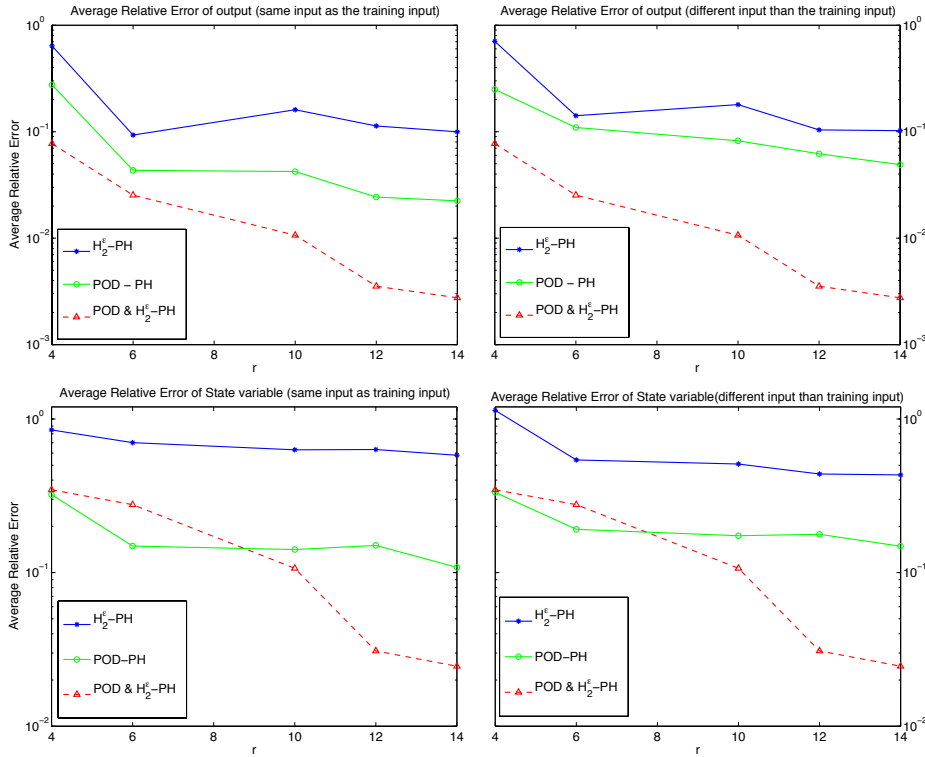
Results for a wide range of r values are illustrated in Figure 2.3, depicting the relative error in the output. In these figures, the colors correspond to a fixed order r. The x-axis is the number of POD-PH basis vectors for that given order, so e.g., the blue line corresponds to relative error for $r = 16$. For that line, the value corresponding to

$x = 12$ means that for the $r = 16$ model, $\hat{r} = 12$ POD-PH basis vectors are combined with $\bar{r} = r - \hat{r} = 4$ $\mathcal{H}_2^\varepsilon$-PH basis vectors. The most accurate approximation resulted from reduced systems with bases that combined roughly equal numbers of POD-PH and $\mathcal{H}_2^\varepsilon$-PH basis vectors. Figure 2.4 shows that reduced systems constructed from these hybrid bases (using equal numbers of POD-PH and $\mathcal{H}_2^\varepsilon$-PH vectors) can give much more accurate approximations than those constructed with POD-PH alone or $\mathcal{H}_2^\varepsilon$-PH alone. For $r = 12$, say, a reduced system with a hybrid basis using $\hat{r} = \bar{r} = 6$ (red dashed line) produced roughly 10 times smaller error for both state variables and outputs than those using only POD-PH (green solid lines) or $\mathcal{H}_2^\varepsilon$-PH bases (blue solid lines).

2.5. An *a priori* error bound for NLPH-reduced models. Fix a reduction order, r, and let $\mathcal{V}_r = \mathsf{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \mathsf{Ran}(\mathbf{W}_r)$ denote r-dimensional reduction subspaces used in creating reduced NLPH systems as in (2.4). We provide an error analysis here that bounds the deviation between the true state trajectory of an NLPH system and that provided by a reduced NLPH system of the form given in (2.4). This leads in turn to a bound on the error between the true system output and the reduced NLPH system output. Typically, the reduction subspaces $\mathcal{V}_r$ and $\mathcal{W}_r$ will be chosen consistently with the heuristics of (2.1) but the bounds we derive apply more generally.

Let $\mathbf{Q} \in \mathbb{R}^{n \times n}$ be a symmetric positive-definite matrix and define a weighted inner product on $\mathbb{R}^n$ as $\langle \mathbf{x}, \mathbf{z} \rangle_\mathbf{Q} = \mathbf{x}^T \mathbf{Q} \mathbf{z}$, with a related norm, $\|\mathbf{x}\|_\mathbf{Q} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_\mathbf{Q}}$. We leave the choice of $\mathbf{Q}$ open for the time being, however the choices $\mathbf{Q} = \mathbf{I}$ and $\mathbf{Q} = \nabla^2 H(\mathbf{x}_0)$ at a locally stable equilibrium point $\mathbf{x}_0$ will have particular merit.

For a mapping $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, we define the associated *Lipschitz constant* and *logarithmic Lipschitz constant* of $\mathbf{F}$ relative to $\mathbf{Q}$ (see [26]) as

$$L_\mathbf{Q}[\mathbf{F}] = \sup_{\mathbf{u} \neq \mathbf{v}} \frac{\|\mathbf{F}(\mathbf{u}) - \mathbf{F}(\mathbf{v})\|_\mathbf{Q}}{\|\mathbf{u} - \mathbf{v}\|_\mathbf{Q}}, \quad \text{and} \quad \mathcal{L}_\mathbf{Q}[\mathbf{F}] = \sup_{\mathbf{u} \neq \mathbf{v}} \frac{\langle \mathbf{u} - \mathbf{v}, \mathbf{F}(\mathbf{u}) - \mathbf{F}(\mathbf{v}) \rangle_\mathbf{Q}}{\|\mathbf{u} - \mathbf{v}\|_\mathbf{Q}^2} \quad (2.12)$$

respectively. Note that $\mathcal{L}_\mathbf{Q}[\mathbf{F}]$ could be negative and $-L_\mathbf{Q}[\mathbf{F}] \leq \mathcal{L}_\mathbf{Q}[\mathbf{F}] \leq L_\mathbf{Q}[\mathbf{F}]$.

Suppose $\mathbf{Q}$-orthogonal bases for $\mathcal{V}_r$ and $\mathcal{W}_r$ are chosen: $\widetilde{\mathbf{V}}_r, \widetilde{\mathbf{W}}_r \in \mathbb{R}^{n \times r}$ such that $\mathcal{V}_r = \mathsf{Ran}(\widetilde{\mathbf{V}}_r)$ and $\mathcal{W}_r = \mathsf{Ran}(\widetilde{\mathbf{W}}_r)$, with $\widetilde{\mathbf{V}}_r^T \mathbf{Q} \widetilde{\mathbf{V}}_r = \mathbf{I}$ and $\widetilde{\mathbf{W}}_r^T \mathbf{Q} \widetilde{\mathbf{W}}_r = \mathbf{I}$. Consider $\mathbf{Q}$-orthogonal projectors $\mathbf{\Pi}_\mathcal{V} : \mathbb{R}^n \to \mathcal{V}_r$, $\mathbf{\Pi}_\mathcal{W} : \mathbb{R}^n \to \mathcal{W}_r$ defined by $\mathbf{\Pi}_\mathcal{V} = \widetilde{\mathbf{V}}_r \widetilde{\mathbf{V}}_r^T \mathbf{Q}$ and $\mathbf{\Pi}_\mathcal{W} = \widetilde{\mathbf{W}}_r \widetilde{\mathbf{W}}_r^T \mathbf{Q}$. Define an ancillary state space projection as $\mathcal{P}_r = \mathbf{V}_r \mathbf{W}_r^T$. For a given (true) system trajectory, $\mathbf{x}(t)$, the best approximation (relative to the $\mathbf{Q}$-norm) that is available by a path in $\mathcal{V}_r$ is given by $\mathbf{\Pi}_\mathcal{V} \mathbf{x}(t)$, and so the state space error $\mathbf{x}(t) - \mathbf{V}_r \mathbf{x}_r(t)$ is bounded pointwise below as

$$\|(\mathbf{I} - \mathbf{\Pi}_\mathcal{V}) \mathbf{x}(t)\|_\mathbf{Q} \leq \|\mathbf{x}(t) - \mathbf{V}_r \mathbf{x}_r(t)\|_\mathbf{Q}$$

for all $t \geq 0$. The error associated with the reduced internal force, $\nabla_{\mathbf{x}_r} H_r$, is bounded similarly

$$\|(\mathbf{I} - \mathbf{\Pi}_\mathcal{W}) \nabla_\mathbf{x} H(\mathbf{x}(t))\|_\mathbf{Q} \leq \|\nabla_\mathbf{x} H(\mathbf{x}(t)) - \mathbf{W}_r \nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r(t))\|_\mathbf{Q}.$$

Let the optimal state-space and internal force residual vectors be defined as

$$\varepsilon_\mathbf{x}(t) = (\mathbf{I} - \mathbf{\Pi}_\mathcal{V}) \mathbf{x}(t) \quad \text{and} \quad \varepsilon_\mathbf{F}(t) = (\mathbf{I} - \mathbf{\Pi}_\mathcal{W}) \nabla_\mathbf{x} H(\mathbf{x}(t)).$$

The squared residual state-space and internal force errors integrated over $[0, T]$ will be denoted as

$$\mathcal{E}_\mathbf{x} = \int_0^T \|\varepsilon_\mathbf{x}(t)\|_\mathbf{Q}^2 \, dt = \int_0^T \|(\mathbf{I} - \mathbf{\Pi}_\mathcal{V}) \mathbf{x}(t)\|_\mathbf{Q}^2 \, dt$$

and

$$\mathcal{E}_{\mathbf{F}} = \int_0^T \|\boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}}^2 \, dt = \int_0^T \|(\mathbf{I} - \boldsymbol{\Pi}_{\mathcal{W}})\nabla_{\mathbf{x}} H(\mathbf{x}(t))\|_{\mathbf{Q}}^2 \, dt.$$

We seek to bound the state space error, $\mathbf{x} - \mathbf{V}_r \mathbf{x}_r$, and output error, $\mathbf{y} - \mathbf{y}_r$, in terms of $\mathcal{E}_{\mathbf{x}}, \mathcal{E}_{\mathbf{F}}$, and the deviation between the projected and reduced initial condition. Note that when the reduction spaces, $\mathcal{V}_r$ and $\mathcal{W}_r$, are generated via a POD approach (e.g., Algorithm 1) then the aggregate errors $\mathcal{E}_{\mathbf{x}}, \mathcal{E}_{\mathbf{F}}$ are approximately minimized and can be expressed directly as the sum of neglected singular values of the snapshot matrix (either of $\mathbf{x}(t)$ or of $\nabla_{\mathbf{x}} H(\mathbf{x}(t))$). As we have seen in §2.4, the actual state space error or output error might not be minimized with this choice.

THEOREM 2.2. *Suppose* $\mathbf{Q} \in \mathbb{R}^{n \times n}$ *is symmetric positive definite and that a reduced port-Hamiltonian system as in (2.4) is constructed to approximate the full order* NLPH *system (1.1) using reduction bases* $\mathbf{V}_r, \mathbf{W}_r \in \mathbb{R}^{n \times r}$ *that are defined so that* $\mathbf{V}_r^T \mathbf{Q} \mathbf{V}_r = \mathbf{I}_r$ *and* $\mathbf{V}_r^T \mathbf{W}_r = \mathbf{I}_r$. *(Note that it may or may not be the case that* $\mathbf{W}_r = \mathbf{Q} \mathbf{V}_r$.*) Suppose further that* $\mathbf{F}(\mathbf{x}) = \nabla_{\mathbf{x}} H(\mathbf{x})$ *in (1.1) is Lipschitz continuous. Denote* $\mathcal{A} = \mathbf{J} - \mathbf{R}$ *and* $\mathcal{P}_r = \mathbf{V}_r \mathbf{W}_r^T$. *Then*

$$\int_0^T \|\mathbf{x}(t) - \mathbf{V}_r \mathbf{x}_r(t)\|_{\mathbf{Q}}^2 \, dt \leq C_{\mathbf{x}} \, \mathcal{E}_{\mathbf{x}} + C_{\mathbf{F}} \, \mathcal{E}_{\mathbf{F}} + C_0 \|\mathbf{W}_r^T \mathbf{x}(0) - \mathbf{x}_r(0)\|^2 \quad (2.13)$$

$$\int_0^T \|\mathbf{y}(t) - \mathbf{y}_r(t)\|^2 \, dt \leq \widehat{C}_{\mathbf{x}} \mathcal{E}_{\mathbf{x}} + \widehat{C}_{\mathbf{F}} \mathcal{E}_{\mathbf{F}} + \widehat{C}_0 \|\mathbf{W}_r^T \mathbf{x}(0) - \mathbf{x}_r(0)\|^2 \quad (2.14)$$

where $\alpha = \mathcal{L}_{\mathbf{Q}}[\mathcal{P}_r \mathcal{A} \mathcal{P}_r^T \mathbf{F}], \quad \beta = \|\mathcal{P}_r \mathcal{A}\|_{\mathbf{Q}} \|\mathcal{P}_r^T\|_{\mathbf{Q}}, \quad \gamma = L_{\mathbf{Q}}[\mathbf{F}] \, \|\mathcal{P}_r\|_{\mathbf{Q}},$

$$\delta = 2\|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\| \, \|\mathcal{P}_r^T\|_{\mathbf{Q}}^2, \quad c_\alpha(t) = \int_0^t e^{2\alpha\tau} \, d\tau, \quad C_\alpha(t) = \int_0^t c_\alpha(\tau) \, d\tau,$$

$$C_{\mathbf{x}} = (2\beta\gamma)^2 \, C_\alpha(T) + 2 \, \|\mathcal{P}_r\|_{\mathbf{Q}}^2, \quad C_{\mathbf{F}} = (2\beta)^2 \, C_\alpha(T), \quad C_0 = 2 \, c_\alpha(T), \text{ and}$$

$$\widehat{C}_{\mathbf{x}} = \delta \cdot L_{\mathbf{Q}}[\mathbf{F}]^2 C_{\mathbf{x}}, \quad \widehat{C}_{\mathbf{F}} = \delta \cdot (1 + L_{\mathbf{Q}}[\mathbf{F}]^2 C_{\mathbf{F}}), \quad \widehat{C}_0 = \delta \cdot L_{\mathbf{Q}}[\mathbf{F}]^2 C_0.$$

Proof. First note that

$$(\mathbf{I} - \mathcal{P}_r)\mathbf{x}(t) = (\mathbf{I} - \mathcal{P}_r)(\mathbf{I} - \boldsymbol{\Pi}_{\mathcal{V}})\mathbf{x}(t) = (\mathbf{I} - \mathcal{P}_r)\boldsymbol{\varepsilon}_{\mathbf{x}}(t) \quad \text{and} \quad (2.15)$$

$$(\mathbf{I} - \mathcal{P}_r^T)\mathbf{F}(t) = (\mathbf{I} - \mathcal{P}_r^T)(\mathbf{I} - \boldsymbol{\Pi}_{\mathcal{W}})\mathbf{F}(t) = (\mathbf{I} - \mathcal{P}_r^T)\boldsymbol{\varepsilon}_{\mathbf{F}}(t). \quad (2.16)$$

The state space error can be separated into the sum of a component in $\mathsf{Ker}(\mathcal{P}_r)$ (i.e., orthogonal to $\mathcal{W}_r$) and a component contained in $\mathsf{Ran}(\mathcal{P}_r) = \mathcal{V}_r$:

$$\mathbf{x}(t) - \mathbf{V}_r \mathbf{x}_r(t) = \boldsymbol{\rho}(t) + \mathbf{V}_r \boldsymbol{\theta}(t),$$

where $\boldsymbol{\rho}(t) = \mathbf{x}(t) - \mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t) = (\mathbf{I} - \mathcal{P}_r)\boldsymbol{\varepsilon}_{\mathbf{x}}(t) \in \mathcal{W}_r^\perp$ and $\boldsymbol{\theta}(t) = \mathbf{W}_r^T \mathbf{x}(t) - \mathbf{x}_r(t)$. Then

$$\int_0^T \|\mathbf{x}(t) - \mathbf{V}_r \mathbf{x}_r(t)\|_{\mathbf{Q}}^2 \, dt \leq 2 \int_0^T \|\boldsymbol{\rho}(t)\|_{\mathbf{Q}}^2 \, dt + 2 \int_0^T \|\boldsymbol{\theta}(t)\|^2 \, dt. \quad (2.17)$$

For the first term, we may estimate immediately

$$\int_0^T \|\boldsymbol{\rho}(t)\|_{\mathbf{Q}}^2 \, dt \leq \int_0^T \|(\mathbf{I} - \mathcal{P}_r)\boldsymbol{\varepsilon}_{\mathbf{x}}(t)\|_{\mathbf{Q}}^2 \, dt \leq \|\mathcal{P}_r\|_{\mathbf{Q}}^2 \, \mathcal{E}_{\mathbf{x}}$$

where we have made use of the identity, $\|\mathbf{I} - \boldsymbol{\mathcal{P}}_r\|_{\mathbf{Q}} = \|\boldsymbol{\mathcal{P}}_r\|_{\mathbf{Q}}$ (see [27]).

To bound the second term of (2.17), note that

$$\dot{\boldsymbol{\theta}}(t) = \mathbf{W}_r^T \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}_r(t) = \mathbf{A}_r \mathbf{V}_r^T \left[\mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r(t)) \right] + \boldsymbol{\eta}(t),$$

where $\mathbf{A}_r = \mathbf{J}_r - \mathbf{R}_r$ (see (2.9)) and $\boldsymbol{\eta}(t) = \mathbf{W}_r^T (\mathbf{J} - \mathbf{R}) \left[\mathbf{F}(\mathbf{x}(t)) - \mathbf{W}_r \mathbf{V}_r^T \mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) \right]$.
Note that $\|\boldsymbol{\theta}(t)\| \frac{d}{dt} \|\boldsymbol{\theta}(t)\| = \frac{1}{2} \frac{d}{dt} \|\boldsymbol{\theta}(t)\|^2 = \left\langle \boldsymbol{\theta}(t), \dot{\boldsymbol{\theta}}(t) \right\rangle$, so we have

$$\frac{d}{dt} \|\boldsymbol{\theta}(t)\| = \left\langle \frac{\boldsymbol{\theta}(t)}{\|\boldsymbol{\theta}(t)\|}, \dot{\boldsymbol{\theta}}(t) \right\rangle = \left\langle \frac{\boldsymbol{\theta}(t)}{\|\boldsymbol{\theta}(t)\|}, \mathbf{A}_r \mathbf{V}_r^T \left[\mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r(t)) \right] + \boldsymbol{\eta}(t) \right\rangle.$$

Observe that

$$| \langle \boldsymbol{\theta}(t), \boldsymbol{\eta}(t) \rangle | = \left| \left\langle \mathbf{V}_r \boldsymbol{\theta}(t), \mathbf{V}_r \mathbf{W}_r^T (\mathbf{J} - \mathbf{R}) \left[\mathbf{F}(\mathbf{x}(t)) - \mathbf{W}_r \mathbf{V}_r^T \mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) \right] \right\rangle_{\mathbf{Q}} \right|$$

$$\leq \|\mathbf{V}_r \boldsymbol{\theta}(t)\|_{\mathbf{Q}} \ \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \left[\mathbf{F}(\mathbf{x}(t)) - \boldsymbol{\mathcal{P}}_r^T \mathbf{F}(\boldsymbol{\mathcal{P}}_r \mathbf{x}(t)) \right] \|_{\mathbf{Q}}$$

$$\leq \|\boldsymbol{\theta}(t)\| \ \|\boldsymbol{\mathcal{P}}_r \mathcal{A}\|_{\mathbf{Q}} \|(\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T) \mathbf{F}(\mathbf{x}(t)) + \boldsymbol{\mathcal{P}}_r^T (\mathbf{F}(\mathbf{x}(t)) - \mathbf{F}(\boldsymbol{\mathcal{P}}_r \mathbf{x}(t)) \|_{\mathbf{Q}}$$

$$\leq \|\boldsymbol{\theta}(t)\| \ \|\boldsymbol{\mathcal{P}}_r \mathcal{A}\|_{\mathbf{Q}} \cdot \left(\|(\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T) \boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}} + \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \|\mathbf{F}(\mathbf{x}(t)) - \mathbf{F}(\boldsymbol{\mathcal{P}}_r \mathbf{x}(t))\|_{\mathbf{Q}} \right)$$

$$\leq \|\boldsymbol{\theta}(t)\| \ \|\boldsymbol{\mathcal{P}}_r \mathcal{A}\|_{\mathbf{Q}} \cdot \left(\|\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \|\boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}} + \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \ L_{\mathbf{Q}}[\mathbf{F}] \|(\mathbf{I} - \boldsymbol{\mathcal{P}}_r) \mathbf{x}(t)\|_{\mathbf{Q}} \right)$$

$$\leq \|\boldsymbol{\theta}(t)\| \left(\|\boldsymbol{\mathcal{P}}_r \mathcal{A}\|_{\mathbf{Q}} \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \right) \cdot (\|\boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}} + L_{\mathbf{Q}}[\mathbf{F}] \ \|\boldsymbol{\mathcal{P}}_r\|_{\mathbf{Q}} \|\boldsymbol{\varepsilon}_{\mathbf{x}}(t)\|_{\mathbf{Q}})$$

$$\leq \|\boldsymbol{\theta}(t)\| \ \cdot \ \beta \left(\|\boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}} + \gamma \|\boldsymbol{\varepsilon}_{\mathbf{x}}(t)\|_{\mathbf{Q}} \right)$$

and

$$\left\langle \boldsymbol{\theta}(t), \mathbf{A}_r \mathbf{V}_r^T \left[\mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r(t)) \right] \right\rangle$$

$$= \left\langle \mathbf{V}_r \boldsymbol{\theta}(t), \boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \left[\mathbf{F}(\mathbf{V}_r \mathbf{W}_r^T \mathbf{x}(t)) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r(t)) \right] \right\rangle_{\mathbf{Q}}$$

$$\leq \mathcal{L}_{\mathbf{Q}}[\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \mathbf{F}] \ \cdot \ \|\mathbf{V}_r \boldsymbol{\theta}(t)\|_{\mathbf{Q}}^2 = \alpha \ \|\boldsymbol{\theta}(t)\|^2 ,$$

where we make use of the fact that $\|\mathbf{I} - \boldsymbol{\mathcal{P}}_r\|_{\mathbf{Q}} = \|\boldsymbol{\mathcal{P}}_r\|_{\mathbf{Q}}$ and $\|\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} = \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}$.

For (2.14), we find

$$\|\mathbf{y} - \mathbf{y}_r\| = \|\mathbf{B}^T \mathbf{F}(\mathbf{x}) - \mathbf{B}^T \mathbf{W}_r \mathbf{V}_r^T \mathbf{F}(\mathbf{V}_r \mathbf{x}_r)\|$$

$$= \|\mathbf{B}^T \mathbf{Q}^{-1/2} \mathbf{Q}^{1/2} \left(\mathbf{F}(\mathbf{x}) - \boldsymbol{\mathcal{P}}_r^T \mathbf{F}(\mathbf{V}_r \mathbf{x}_r) \right) \|$$

$$\leq \|\mathbf{B}^T \mathbf{Q}^{-1/2}\| \ \|(\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T) \boldsymbol{\varepsilon}_{\mathbf{F}} + \boldsymbol{\mathcal{P}}_r^T \left(\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r) \right) \|_{\mathbf{Q}}$$

$$\leq \|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\|^{1/2} \left(\|(\mathbf{I} - \boldsymbol{\mathcal{P}}_r^T) \boldsymbol{\varepsilon}_{\mathbf{F}}\|_{\mathbf{Q}} + \|\boldsymbol{\mathcal{P}}_r^T \left(\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r) \right) \|_{\mathbf{Q}} \right)$$

$$\leq \|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\|^{1/2} \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} (\|\boldsymbol{\varepsilon}_{\mathbf{F}}\|_{\mathbf{Q}} + \| \left(\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{V}_r \mathbf{x}_r) \right) \|_{\mathbf{Q}})$$

$$\leq \|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\|^{1/2} \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} (\|\boldsymbol{\varepsilon}_{\mathbf{F}}\|_{\mathbf{Q}} + L_{\mathbf{Q}}[\mathbf{F}] \|\mathbf{x} - \mathbf{V}_r \mathbf{x}_r\|_{\mathbf{Q}})$$

Thus,

$$\int_0^T \|\mathbf{y}(t) - \mathbf{y}_r(t)\|^2 \, dt \leq 2\|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\| \ \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}^2 \int_0^T \left(\|\boldsymbol{\varepsilon}_{\mathbf{F}}(t)\|_{\mathbf{Q}}^2 + L_{\mathbf{Q}}[\mathbf{F}]^2 \|\mathbf{x} - \mathbf{V}_r \mathbf{x}_r\|_{\mathbf{Q}}^2 \right) \, dt$$

$$\leq 2\|\mathbf{B}^T \mathbf{Q}^{-1} \mathbf{B}\| \ \|\boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}^2 \ \cdot$$

$$\left(\mathcal{E}_{\mathbf{F}} + L_{\mathbf{Q}}[\mathbf{F}]^2 \left(C_{\mathbf{x}} \mathcal{E}_{\mathbf{x}} + C_{\mathbf{F}} \mathcal{E}_{\mathbf{F}} + C_0 \|\mathbf{W}_r^T \mathbf{x}(0) - \mathbf{x}_r(0)\|^2 \right) \right)$$

and (2.14) follows. $\square$

3. Structure-preserving model reduction with DEIM. The performance of projection-based reduels of nonlinear systems can be degraded by the need to lift the reduced state to the full state dimension in order to evaluate the nonlinear term. For example, consider the evaluation of $\mathbf{V}_r^T \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r)$ in (2.2). If $\nabla_{\mathbf{x}} H(\mathbf{x})$ is nonlinear in $\mathbf{x}$, it is likely that $\mathbf{V}_r^T \nabla_{\mathbf{x}} H(\mathbf{V}_r \mathbf{x}_r)$ cannot be precomputed explicitly as a map from $\mathbb{R}^r$ to $\mathbb{R}^r$ without an intermediate lifting to $\mathbb{R}^n$. The order of complexity required to evolve the reduced system will then remain at least $\mathcal{O}(n)$ and there may be little benefit seen in the use of a reduced model. Several approaches have been proposed to address this difficulty; see, e.g., [1, 2, 4, 6, 9]. We resolve the complexity issue by developing a variant of the *discrete empirical interpolation method* (DEIM) [6], which is itself a discrete variant of the Empirical Interpolation Method introduced in [2]. Since DEIM (as presented in [6]) does not typically preserve port-Hamiltonian structure, we develop here a structure-preserving variant of DEIM together with associated error estimates. There have been other structure-preserving methods developed recently that employ methods resolving the lifting bottleneck, e.g., see [5] for an approach that preserves Lagrangian structure in structural dynamics using gappy POD.

3.1. The Discrete Empirical Interpolation Method. The 'lifting bottleneck' in the reduction of large scale nonlinear models as described above is resolved by DEIM through the approximation of the nonlinear system function via interpolation. This approximation is done in such a way as to not require a prolongation of the reduced state variables ('lifting') back to the original high dimensional state space. Only a few selected entries of the original nonlinear term need be evaluated at each time step.

In particular, let $\mathbf{f} : \mathscr{D} \mapsto \mathbb{R}^n$ be a nonlinear vector-valued function defined on a domain $\mathscr{D} \subseteq \mathbb{R}^n$. Let $\mathbf{U}_m = [\mathbf{u}_1, \ldots, \mathbf{u}_m] \in \mathbb{R}^{n \times m}$ have rank m and define $\mathbf{E}_m = [\mathbf{e}_{\wp_1}, \ldots, \mathbf{e}_{\wp_m}] \in \mathbb{R}^{n \times m}$ with the index set $\{\wp_1, \ldots, \wp_m\}$ output from Algorithm 4 using the input basis $\{\mathbf{u}_i\}_{i=1}^m$. $\mathbf{e}_{\wp_j} \in \mathbb{R}^n$ denotes the $\wp_j$th column of the n-by-n identity matrix.

The DEIM approximation of order $m \leq n$ for $\mathbf{f}$ in span$\{\mathbf{U}_m\}$ is given by

$$\widehat{\mathbf{f}}(\tau) := \mathbb{P} \mathbf{f}(\tau), \quad \text{where} \quad \mathbb{P} = \mathbf{U}_m (\mathbf{E}_m^T \mathbf{U}_m)^{-1} \mathbf{E}_m^T. \tag{3.1}$$

Note that in the original work, [6], $\mathbf{U}_m$ is assumed to have orthonormal columns, yet linear independence suffices for $\mathbb{P}$ to be well defined and that is all that we require here.

We adapt an error bound for the DEIM function approximation derived in [6] for the case that $\mathbf{U}_m$ has $\mathbf{Q}$-orthonormal columns.

LEMMA 3.1. *If $\mathbf{U}_m^T \mathbf{Q} \mathbf{U}_m = \mathbf{I}_m$, then $\|\mathbf{f}(\tau) - \widehat{\mathbf{f}}(\tau)\|_{\mathbf{Q}} \leq \|\mathbb{P}\|_{\mathbf{Q}} \; \mathscr{E}_{\mathbf{Q}}(\mathbf{f}(\tau), \mathbf{U}_m)$, where $\mathscr{E}_{\mathbf{Q}}(\mathbf{f}(\tau), \mathbf{U}_m)$ is the best $\mathbf{Q}$-norm approximation error for $\mathbf{f}(\tau)$ from $\mathsf{Ran}(\mathbf{U}_m)$.*

The invertibility of $\mathbf{E}^T \mathbf{U}$ at the end of each cycle of the DEIM procedure is verified in [6] where it is shown that each DEIM interpolation index is selected in order to limit the stepwise growth of the factor $\|(\mathbf{E}_m^T \mathbf{U}_m)^{-1}\| = \|\mathbb{P}\|_{\mathbf{Q}}$ in the error bound. This will be used in the next section to assess the accuracy of the state variables in the DEIM reduced system. DEIM shown in Algorithm 4 uses LU with partial pivoting for interpolation indices. A new selection operator for DEIM based on the pivoted QR has been recently introduced by Drmač and Gugercin [7]. Our numerical results in this paper are based the original implementation in Algorithm 4. We also refer the reader to [17] for a recently introduced localized version of DEIM and to [18] for online adaptivity approach to DEIM that adjusts the DEIM subspace and the interpolation indices with online low-rank updates.

Algorithm 4 : DEIM [6]

INPUT: $\{\mathbf{u}_\ell\}_{\ell=1}^m \subset \mathbb{R}^n$ linearly independent
OUTPUT: $\vec{\wp} = [\wp_1, \ldots, \wp_m]^T \in \mathbb{R}^m$

1: $[|\rho|, \ \wp_1] = \mathtt{max}\{|\mathbf{u}_1|\}$

2: $\mathbf{U} = [\mathbf{u}_1]$, $\mathbf{E} = [\mathbf{e}_{\wp_1}]$, $\vec{\wp} = [\wp_1]$

3: **for** $\ell = 2$ to m **do**

4:　　Solve $(\mathbf{E}^T\mathbf{U})\mathbf{c} = \mathbf{E}^T\mathbf{u}_\ell$ for $\mathbf{c}$

5:　　$\mathbf{r} = \mathbf{u}_\ell - \mathbf{U}\mathbf{c}$

6:　　$[|\rho|, \ \wp_\ell] = \mathtt{max}\{|\mathbf{r}|\}$　　$\left(\begin{array}{l} |\rho| = |v_\wp| = \max_{i=1,\ldots,n}\{|v_i|\}, \text{ with the} \\ \text{smallest index taken for } \wp \text{ in case of ties.} \end{array} \right)$

7:　　$\mathbf{U} \leftarrow [\mathbf{U} \ \mathbf{u}_\ell]$, $\mathbf{E} \leftarrow [\mathbf{E} \ \mathbf{e}_{\wp_\ell}]$, $\vec{\wp} \leftarrow \left[\begin{array}{c} \vec{\wp} \\ \wp_\ell \end{array} \right]$

8: **end for**

Consider the nonlinear PH system in (1.1) where $\mathbf{J}$, $\mathbf{R}$, and $\mathbf{B}$ are constant, and $\mathbf{F}(\mathbf{x}) = \nabla_\mathbf{x} H(\mathbf{x})$ is nonlinear. In the reduced system (2.4), we use the approximation (2.3): $\nabla_\mathbf{x} H(\mathbf{x}) \approx \mathbf{W}_r \mathbf{V}_r^T \nabla_\mathbf{x} H(\mathbf{x})$, for $\mathbf{x} \in \mathsf{Ran}(\mathbf{V}_r)$. DEIM can be applied directly to $\mathbf{F}$ (using $\mathbf{W}_r$ instead of $\mathbf{U}$) to obtain an approximation in the form of

$$\mathbf{F}(\mathbf{V}_r\mathbf{x}_r) \approx \mathbf{W}_r(\mathbf{E}^T\mathbf{W}_r)^{-1}\mathbf{E}^T\mathbf{F}(\mathbf{V}_r\mathbf{x}_r),$$

allowing us to evaluate the nonlinear term with low complexity. However, this approach will not preserve the underlying PH structure; it generally will not produce a passive system; and indeed, the reduced system might no longer be stable. We modify DEIM to overcome these shortcomings.

3.2. The DEIM Hamiltonian. We continue to assume that the source of nonlinearity in the system (1.1) lies in the Hamiltonian gradient: $\nabla_\mathbf{x} H(\mathbf{x})$, and we focus on approximating this nonlinear term in a way that is consistent with the PH structure, so that the complexity does not depend on the original full-order dimension. This restriction comes largely without loss of generality since additional state-space dependence of $\mathbf{J}$, $\mathbf{R}$, and $\mathbf{B}$ can be accommodated with usual DEIM-based approaches with no threat to the underlying port-Hamiltonian structure.

We first identify a linear component of $\nabla_\mathbf{x} H(\mathbf{x})$, or equivalently, a quadratic component of $H(\mathbf{x})$:

$$H(\mathbf{x}) = \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + h(\mathbf{x}) \tag{3.2}$$

where $\mathbf{Q}$ is an $n \times n$ positive-definite, constant matrix. A typical choice may be $\mathbf{Q} = \nabla_\mathbf{x}^2 H(\mathbf{x}_0)$ at an equilibrium point, $\nabla_\mathbf{x} H(\mathbf{x}_0) = \mathbf{0}$. Similar strategies are considered in [15] to maintain high accuracy. Once $\mathbf{Q}$ is selected, (3.2) determines $h(\mathbf{x})$ and $\nabla_\mathbf{x} h(\mathbf{x})$ then captures the remaining nonlinear portion of $\nabla_\mathbf{x} H(\mathbf{x})$.

We select a new modeling basis, $\mathbf{U}_m$, orthogonalized with respect to $\mathbf{Q}$, so that $\nabla_\mathbf{x} h(\mathbf{x}(t)) \approx \mathbf{U}_m \mathbf{g}_m(t)$ and $\mathbf{U}_m^T \mathbf{Q} \mathbf{U}_m = \mathbf{I}$. There will be a variety of choices for $\mathbf{U}_m$; we consider a couple of them below. Using $\mathbf{U}_m$, we calculate DEIM indices, $\wp_1, \ldots, \wp_m$ with Algorithm 4, define the associated DEIM projection, $\mathbb{P} = \mathbf{U}_m(\mathbf{E}_m^T\mathbf{U}_m)^{-1}\mathbf{E}_m^T$, and

finally, introduce a "DEIM Hamiltonian":

$$\widehat{H}(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^T\,\mathbf{Q}\,\mathbf{x} \,+\, h(\mathbb{P}^T\mathbf{x}). \tag{3.3}$$

Observe that $\nabla_{\mathbf{x}}\widehat{H}(\mathbf{x}) = \mathbf{Q}\mathbf{x} + \mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{x})$, so the error induced by $\nabla_{\mathbf{x}}\widehat{H}(\mathbf{x})$ is

$$\nabla_{\mathbf{x}}H(\mathbf{x}) - \nabla_{\mathbf{x}}\widehat{H}(\mathbf{x}) = \nabla_{\mathbf{x}}h(\mathbf{x}) - \mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{x}).$$

If $\mathbf{Q}$ is chosen well then $\nabla_{\mathbf{x}}\widehat{H}(\mathbf{x})$ can exactly recover the "linear part" of $\nabla_{\mathbf{x}}H(\mathbf{x})$. Observe that the evaluation of the remaining nonlinear term, $\mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{x})$, only involves the evaluation of m elements of $\nabla_{\mathbf{x}}h$ on only m nonzero arguments - the remaining arguments having only m nonzero values.

Observe that trivially $\nabla_{\mathbf{x}}H(\mathbf{x}) = \nabla_{\mathbf{x}}\widehat{H}(\mathbf{x})$ when $m = n$. However, the approximation $\nabla_{\mathbf{x}}H(\mathbf{x}) \approx \nabla_{\mathbf{x}}\widehat{H}(\mathbf{x})$ can be effective even for significantly smaller $m \ll n$. Nonetheless, the enforced symmetry in this approximation appears to require some additional considerations.

Let $\Omega \subset \mathbb{R}^n$, be a compact convex set containing each of the trajectories $\mathbf{x}(t)$ and $\mathbb{P}^T\mathbf{x}(t)$, for $0 \le t \le T$ on its interior. Define

$$\mathbf{K}^{\circ} = \int_{\Omega} \nabla_{\mathbf{x}}h(\mathbf{x})\nabla_{\mathbf{x}}h(\mathbf{x})^T\mathbf{Q}\ d\nu(\mathbf{x}).$$

$\mathbf{K}^{\circ}$ is $\mathbf{Q}$-selfadjoint and positive-definite with associated eigenpairs $\mathbf{K}^{\circ}\mathbf{u}_k^{\circ} = \sigma_k^{\circ}\mathbf{u}_k^{\circ}$ for $\sigma_1^{\circ} \ge \sigma_2^{\circ} \ge \ldots \ge \sigma_n^{\circ} \ge 0$ and $\mathbf{Q}$-orthonormal eigenvectors, $\mathbf{u}_i^{\circ T}\mathbf{Q}\mathbf{u}_j^{\circ} = \delta_{ij}$. For any choice of $0 < m \le n$, designate the dominant modes as $\mathbf{U}_m^{\circ} = [\mathbf{u}_1^{\circ}, \ldots, \mathbf{u}_m^{\circ}]$ and the complementary subdominant modes as $\widetilde{\mathbf{U}}_m^{\circ} = [\mathbf{u}_{m+1}^{\circ}, \ldots, \mathbf{u}_n^{\circ}]$. Observe that if Ω shrank to $\mathbf{x}(t)$, $\mathbf{U}_m^{\circ}$ could be considered as the limiting case of a usual POD basis of dimension m.

We have in that case,

$$\mathbf{K}^{\circ}\,\mathbf{U}_m^{\circ} = \mathbf{U}_m^{\circ}\mathsf{diag}(\sigma_1^{\circ}, \sigma_2^{\circ}, \ldots, \sigma_m^{\circ}) \quad \text{and} \quad \mathbf{K}^{\circ}\,\widetilde{\mathbf{U}}_m^{\circ} = \widetilde{\mathbf{U}}_m^{\circ}\mathsf{diag}(\sigma_{m+1}^{\circ}, \ldots, \sigma_n^{\circ}).$$

If $\mathscr{E}_{\mathbf{Q}}(\nabla_{\mathbf{x}}h(\mathbf{x}), \mathbf{U}_m^{\circ})$ denotes the (local) best $\mathbf{Q}$-norm approximation error to $\nabla_{\mathbf{x}}h(\mathbf{x})$ out of $\mathbf{U}_m^{\circ}$ (as defined in Lemma 3.1), then

$$\left(\varepsilon_{\mathbf{Q}}^{\circ(m)}\right)^2 = \int_{\Omega} \mathscr{E}_{\mathbf{Q}}(\nabla_{\mathbf{x}}h(\mathbf{x}), \mathbf{U}_m^{\circ})^2\, d\nu(\mathbf{x}) = \int_{\Omega} \|(\mathbf{I} - \mathbf{U}_m^{\circ}\mathbf{U}_m^{\circ T}\mathbf{Q})\nabla_{\mathbf{x}}h(\mathbf{x})\|_{\mathbf{Q}}^2\ d\nu(\mathbf{x})$$

$$= \int_{\Omega} \mathsf{trace}\left((\mathbf{I} - \mathbf{U}_m^{\circ}\mathbf{U}_m^{\circ T}\mathbf{Q})\nabla_{\mathbf{x}}h(\mathbf{x})\nabla_{\mathbf{x}}h(\mathbf{x})^T(\mathbf{I} - \mathbf{Q}\mathbf{U}_m^{\circ}\mathbf{U}_m^{\circ T})\mathbf{Q}\right)\ d\nu(\mathbf{x})$$

$$= \int_{\Omega} \mathsf{trace}\left(\nabla_{\mathbf{x}}h(\mathbf{x})\nabla_{\mathbf{x}}h(\mathbf{x})^T\mathbf{Q}(\mathbf{I} - \mathbf{U}_m^{\circ}\mathbf{U}_m^{\circ T}\mathbf{Q})\right)\ d\nu(\mathbf{x})$$

$$= \mathsf{trace}\left(\mathbf{K}^{\circ}(\mathbf{I} - \mathbf{U}_m^{\circ}\mathbf{U}_m^{\circ T}\mathbf{Q})\right) = \sum_{k=m+1}^{n} \sigma_k^{\circ}.$$

LEMMA 3.2. *Suppose $h(\mathbf{x})$ is continuously differentiable in Ω. Construct the DEIM projection using $\mathbf{U}_m^{\circ}$ with $\mathbb{P}_{\circ} = \mathbf{U}_m^{\circ}(\mathbf{E}_m^T\mathbf{U}_m^{\circ})^{-1}\mathbf{E}_m^T$. Then*

$$|h(\mathbf{x}) - h(\mathbb{P}_{\circ}^T\mathbf{x})| \le \|(\mathbf{I} - \mathbb{P}_{\circ}^T)\mathbf{x}(t)\|_{\mathbf{Q}^{-1}} \cdot \varepsilon_{\mathbf{Q}}^{\circ(m)} \tag{3.4}$$

Proof. Define the columns of $\widetilde{\mathbf{E}}_m$ to be complementary to those of $\mathbf{E}_m$ so that $[\mathbf{E}_m, \widetilde{\mathbf{E}}_m]$ is an $n \times n$ permutation matrix. Observe that $\mathsf{Ran}(\mathbb{P}_\circ^T) = \mathsf{Ran}(\mathbf{E}_m)$ and $\mathsf{Ran}(\mathbf{I} - \mathbb{P}_\circ^T) = \mathsf{Ker}(\mathbb{P}_\circ^T) = \mathsf{Ker}(\mathbf{U}_m^{\circ T}) = \mathsf{Ran}(\mathbf{Q}\widetilde{\mathbf{U}}_m^\circ)$. Hence,

$$\mathbf{x} = \mathbb{P}_\circ^T \mathbf{x} + (\mathbf{I} - \mathbb{P}_\circ^T)\mathbf{x} = \mathbf{E}_m \boldsymbol{\zeta}_1 + \mathbf{Q}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2$$

for continuous functions given as

$$\boldsymbol{\zeta}_1(\mathbf{x}) = (\mathbf{U}_m^{\circ T}\mathbf{E}_m)^{-1}\mathbf{U}_m^{\circ T}\mathbf{x} \quad \text{and} \quad \boldsymbol{\zeta}_2(\mathbf{x}) = (\widetilde{\mathbf{E}}_m^T \mathbf{Q}\widetilde{\mathbf{U}}_m^\circ)^{-1}\widetilde{\mathbf{E}}_m^T \mathbf{x}.$$

Note that, $\mathbf{Q}^{\frac{1}{2}}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2 = \mathbf{Q}^{-\frac{1}{2}}(\mathbf{I} - \mathbb{P}_\circ^T)\mathbf{x}$ and $\|\boldsymbol{\zeta}_2\| = \|(\mathbf{I} - \mathbb{P}_\circ^T)\mathbf{x}\|_{\mathbf{Q}^{-1}}$.

For each fixed $\mathbf{x} \in \Omega$, define the path connecting $\mathbf{x}$ and $\mathbb{P}_\circ^T \mathbf{x}$:

$$\mathcal{C}(\mathbf{x}) = \{\theta\mathbf{x} + (1-\theta)\mathbb{P}_\circ^T \mathbf{x} \mid \theta \in [0,1]\} = \{\mathbf{E}_m \boldsymbol{\zeta}_1 + \theta\,\mathbf{Q}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2 \mid \theta \in [0,1]\}.$$

If we define, $f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) = h(\mathbf{x}) = h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2)$, observe that

$$\nabla_{\boldsymbol{\zeta}_2} f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) = \widetilde{\mathbf{U}}_m^{\circ T}\mathbf{Q}\nabla_{\mathbf{x}}h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \mathbf{Q}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2).$$

Integrating $\nabla_{\boldsymbol{\zeta}_2} f$ along $\mathcal{C}(t)$, we find

$$\begin{aligned}
|h(\mathbf{x}) - h(\mathbb{P}_\circ^T \mathbf{x})| = |f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) - f(\boldsymbol{\zeta}_1, 0)| &= \left| \int_{\mathcal{C}(\mathbf{x})} \widetilde{\mathbf{U}}_m^{\circ T}\mathbf{Q}\nabla_{\mathbf{x}}h(\boldsymbol{\xi}) \cdot d\boldsymbol{\xi} \right| \\
&= \left| \int_0^1 \boldsymbol{\zeta}_2^T \widetilde{\mathbf{U}}_m^{\circ T}\mathbf{Q}\nabla_{\mathbf{x}}h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \theta\,\mathbf{Q}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2)\, d\theta \right| \\
&= \|\boldsymbol{\zeta}_2\| \left(\int_0^1 \|\widetilde{\mathbf{U}}_m^{\circ T}\mathbf{Q}\nabla_{\mathbf{x}}h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \theta\,\mathbf{Q}\widetilde{\mathbf{U}}_m^\circ \boldsymbol{\zeta}_2)\|^2\, d\theta \right)^{\frac{1}{2}} \\
&\leq \|\boldsymbol{\zeta}_2\| \left(\int_\Omega \|\widetilde{\mathbf{U}}_m^{\circ T}\mathbf{Q}\nabla_{\mathbf{x}}h(\boldsymbol{\xi})\|^2\, d\nu(\boldsymbol{\xi}) \right)^{\frac{1}{2}} \\
&\leq \|(\mathbf{I} - \mathbb{P}_\circ^T)\mathbf{x}(t)\|_{\mathbf{Q}^{-1}} \cdot \varepsilon_{\mathbf{Q}}^{\circ(m)}.
\end{aligned}$$

$\square$

Thus, a modest extension of the usual POD basis is sufficient to produce DEIM approximations and associated DEIM Hamiltonians that converge uniformly to the true Hamiltonian with a rate related to the decay rate of σ_k°. Stronger hypotheses on the approximating modes appear to be necessary in order to assure rapid convergence of the corresponding gradients. Towards that end, we assume that h is twice continuously differentiable and define

$$\mathbf{K} = \int_\Omega \nabla_{\mathbf{x}}h(\mathbf{x})\nabla_{\mathbf{x}}h(\mathbf{x})^T \mathbf{Q}\, d\nu(\mathbf{x}) + \int_\Omega \left(\nabla_{\mathbf{x}}^2 h(\mathbf{x})\right)^2 \mathbf{Q}\, d\nu(\mathbf{x}).$$

$\mathbf{K}$ is $\mathbf{Q}$-selfadjoint and positive-definite with associated eigenpairs $\mathbf{K}\,\mathbf{u}_k = \sigma_k\,\mathbf{u}_k$ for $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ and $\mathbf{Q}$-orthonormal eigenvectors, $\mathbf{u}_i^T \mathbf{Q}\mathbf{u}_j = \delta_{ij}$. Analogous to what has gone before, designate dominant modes as $\mathbf{U}_m = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ and subdominant modes as $\widetilde{\mathbf{U}}_m = [\mathbf{u}_{m+1}, \ldots, \mathbf{u}_n]$, with

$$\mathbf{K}\,\mathbf{U}_m = \mathbf{U}_m \mathsf{diag}(\sigma_1, \sigma_2, \ldots, \sigma_m), \qquad \mathbf{K}\,\widetilde{\mathbf{U}}_m = \widetilde{\mathbf{U}}_m \mathsf{diag}(\sigma_{m+1}, \ldots, \sigma_n).$$

and an associated error, $\varepsilon_{\mathbf{Q}}^{(m)}$, defined similarly

$$\left(\varepsilon_{\mathbf{Q}}^{(m)}\right)^2 = \int_\Omega \|(\mathbf{I} - \mathbf{U}_m\mathbf{U}_m^T \mathbf{Q})\nabla_{\mathbf{x}}h(\boldsymbol{\xi})\|_{\mathbf{Q}}^2 + \|(\mathbf{I} - \mathbf{U}_m\mathbf{U}_m^T \mathbf{Q})\nabla_{\mathbf{x}}^2 h(\boldsymbol{\xi})\|_{\mathbf{Q}}^2\, d\nu(\boldsymbol{\xi}) = \sum_{k=m+1}^n \sigma_k.$$

LEMMA 3.3. *Given the modeling basis, $\mathbf{U}_m$, and the basis completion, $\widetilde{\mathbf{U}}_m$, as described above,*

$$\|\nabla_{\mathbf{x}} h - \mathbb{P} \circ (\nabla_{\mathbf{x}} h) \circ \mathbb{P}^T\|_{\mathcal{L}_2(\Omega)} \leq \sqrt{2} \|\mathbb{P}\|_{\mathbf{Q}} \; M \; \varepsilon_{\mathbf{Q}}^{(m)}$$

where $M = \max\{1, \|\mathbb{P}\|_{\mathbf{Q}} \sqrt{\frac{\|\mathbf{E}_m^T \mathbf{Q} \mathbf{E}_m\|}{n-m+1}} \sup_{\mathbf{x} \in \Omega} \|\mathbf{x}\|_{\mathbf{Q}^{-1}}\}$

Proof. Note that

$$\|\nabla_{\mathbf{x}} h - \mathbb{P} \circ (\nabla_{\mathbf{x}} h) \circ \mathbb{P}^T\|_{\mathcal{L}_2(\Omega)}^2 = \int_\Omega \|\nabla_{\mathbf{x}} h(\mathbf{x}) - \mathbb{P}\nabla_{\mathbf{x}} h(\mathbb{P}^T\mathbf{x})\|_{\mathbf{Q}}^2 \, d\nu(\mathbf{x})$$

$$\leq 2\int_\Omega \|\mathbb{P}\left(\nabla_{\mathbf{x}} h(\mathbf{x}) - \nabla_{\mathbf{x}} h(\mathbb{P}^T\mathbf{x})\right)\|_{\mathbf{Q}}^2 \, d\nu(\mathbf{x}) + 2\int_\Omega \|\left(\mathbf{I} - \mathbb{P}\right)\nabla_{\mathbf{x}} h(\mathbf{x})\|_{\mathbf{Q}}^2 \, d\nu(\mathbf{x})$$

$$\leq 2\|\mathbb{P}\|_{\mathbf{Q}}^2 \|\mathbf{E}_m^T \mathbf{Q} \mathbf{E}_m\| \int_\Omega \|\mathbf{E}_m^T\left(\nabla_{\mathbf{x}} h(\mathbf{x}) - \nabla_{\mathbf{x}} h(\mathbb{P}^T\mathbf{x})\right)\|^2 \, d\nu(\mathbf{x})$$

$$+ 2\|\mathbf{I} - \mathbb{P}\|_{\mathbf{Q}}^2 \int_\Omega \|\widetilde{\mathbf{U}}_m^T \mathbf{Q} \nabla_{\mathbf{x}} h(\mathbf{x})\|^2 \, d\nu(\mathbf{x})$$

Consider the function $f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)$ as defined in the proof of Lemma 3.2, but now adapted to the new basis, $\mathbf{U}_m$. Note that

$$\nabla_{\boldsymbol{\zeta}_1} f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) = \mathbf{E}_m^T \nabla_{\mathbf{x}} h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \mathbf{Q}\widetilde{\mathbf{U}}_m \boldsymbol{\zeta}_2).$$

Since $\nabla_{\mathbf{x}} h$ is continuously differentiable in an open neighborhood of $\mathcal{C}(t)$, we have that

$$\mathbf{E}_m^T(\nabla_{\mathbf{x}} h(\mathbf{x}) - \nabla_{\mathbf{x}} h(\mathbb{P}^T\mathbf{x})) = \nabla_{\boldsymbol{\zeta}_1} f(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) - \nabla_{\boldsymbol{\zeta}_1} f(\boldsymbol{\zeta}_1, 0)$$

$$= \nabla_{\boldsymbol{\zeta}_1} \int_0^1 \nabla_{\mathbf{x}} h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \theta \, \mathbf{Q}\widetilde{\mathbf{U}}_m \boldsymbol{\zeta}_2) \cdot \left(\mathbf{Q}\widetilde{\mathbf{U}}_m \boldsymbol{\zeta}_2\right) d\theta$$

$$= \left(\int_0^1 \mathbf{E}_m^T \nabla_{\mathbf{x}}^2 h(\mathbf{E}_m \boldsymbol{\zeta}_1 + \theta \, \mathbf{Q}\widetilde{\mathbf{U}}_m \boldsymbol{\zeta}_2) \, \mathbf{Q}\widetilde{\mathbf{U}}_m \, d\theta\right) \boldsymbol{\zeta}_2$$

and so,

$$\int_\Omega \|\mathbf{E}_m^T\left(\nabla_{\mathbf{x}} h(\mathbf{x}) - \nabla_{\mathbf{x}} h(\mathbb{P}^T\mathbf{x})\right)\|^2 \, d\nu(\mathbf{x})$$

$$\leq \sup_{\mathbf{x} \in \Omega} \|\boldsymbol{\zeta}_2(\mathbf{x})\|^2 \int_0^1 \int_\Omega \|\mathbf{E}_m^T \nabla_{\mathbf{x}}^2 h(\theta \, \mathbf{x} + (1-\theta)\mathbb{P}^T\mathbf{x}) \, \mathbf{Q}\widetilde{\mathbf{U}}_m\|^2 \, d\nu(\mathbf{x}) \, d\theta$$

$$\leq \sup_{\mathbf{x} \in \Omega} \|(\mathbf{I} - \mathbb{P}^T)\mathbf{x}\|_{\mathbf{Q}^{-1}}^2 \int_0^1 \int_\Omega \|\mathbf{E}_m^T \nabla_{\mathbf{x}}^2 h(\boldsymbol{\xi}) \, \mathbf{Q}\widetilde{\mathbf{U}}_m\|^2 \, |\det(\theta \, \mathbf{I} + (1-\theta)\mathbb{P}^T)| \, d\nu(\boldsymbol{\xi}) \, d\theta$$

$$\leq \|\mathbb{P}\|_{\mathbf{Q}}^2 \sup_{\mathbf{x} \in \Omega} \|\mathbf{x}\|_{\mathbf{Q}^{-1}}^2 \left(\int_0^1 |\det(\theta \, \mathbf{I} + (1-\theta)\mathbb{P}^T)| \, d\theta\right) \int_\Omega \|\widetilde{\mathbf{U}}_m^T \mathbf{Q} \nabla_{\mathbf{x}}^2 h(\boldsymbol{\xi})\mathbf{E}_m\|^2 \, d\nu(\boldsymbol{\xi})$$

$$\leq \frac{\|\mathbb{P}\|_{\mathbf{Q}}^2}{n-m+1} \sup_{\mathbf{x} \in \Omega} \|\mathbf{x}\|_{\mathbf{Q}^{-1}}^2 \int_\Omega \|\widetilde{\mathbf{U}}_m^T \mathbf{Q} \nabla_{\mathbf{x}}^2 h(\boldsymbol{\xi})\|^2 \, d\nu(\boldsymbol{\xi})$$

The interchange of order of integration is justified in the first inequality since the integrand is uniformly bounded on the joint domain $[0,1] \times \Omega$. In the second inequality, we have introduced the (linear) change of variable $\boldsymbol{\xi} = [\theta \, \mathbf{I} + (1-\theta)\mathbb{P}^T]\mathbf{x} \in \Omega$ for all $\theta \in [0,1]$. The third inequality incorporates the observation $\|\mathbf{I} - \mathbb{P}^T\|_{\mathbf{Q}^{-1}} = \|\mathbb{P}\|_{\mathbf{Q}}$. The last inequality uses the elementary observation that the Jacobian matrix for the change-of-variables has exactly two eigenvalues: $\lambda = 1$ with multiplicity m and $\lambda = \theta$ with multiplicity $n - m$. $\square$

3.3. Preserving port-Hamiltonian Structure with DEIM. Equipped with the DEIM Hamiltonian, $\widehat{H}(\mathbf{x})$, we apply our previously described structure-preserving approach: we define a *reduced* DEIM Hamiltonian, $\widehat{H}_r(\mathbf{x}_r) = \widehat{H}(\mathbf{V}_r\mathbf{x}_r)$. As before, we expect that $\nabla_{\mathbf{x}}\widehat{H}(\mathbf{V}_r\mathbf{x}_r(t)) \approx \mathbf{W}_r\nabla_{\mathbf{x}_r}\widehat{H}_r(\mathbf{x}_r(t))$, and the DEIM-reduced port-Hamiltonian approximation becomes

$$\dot{\mathbf{x}}_r = (\mathbf{J}_r - \mathbf{R}_r)\nabla_{\mathbf{x}_r}\widehat{H}_r(\mathbf{x}_r) + \mathbf{B}_r\mathbf{u}(t),$$

$$\mathbf{y}_r(t) = \mathbf{B}_r^T\,\nabla_{\mathbf{x}_r}\widehat{H}_r(\mathbf{x}_r) \tag{3.5}$$

with $\mathbf{J}_r = \mathbf{W}_r^T\mathbf{J}\mathbf{W}_r$, $\mathbf{R}_r = \mathbf{W}_r^T\mathbf{R}\mathbf{W}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, and

$$\nabla_{\mathbf{x}_r}\widehat{H}_r(\mathbf{x}_r) = \mathbf{V}_r^T\nabla_{\mathbf{x}}\widehat{H}(\mathbf{V}_r\mathbf{x}_r) = \mathbf{x}_r + \mathbf{V}_r^T\mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{V}_r\mathbf{x}_r).$$

In the evaluation of $\mathbf{V}_r^T\mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{V}_r\mathbf{x}_r)$, the $r \times m$ matrix, $\mathbf{V}_r^T\mathbf{U}_m$, may be precomputed and $\nabla_{\mathbf{x}}h$ is evaluated only on arguments having m nonzero values and only m values need to be evaluated (that is, only at the DEIM indices). Algorithm 5 summarizes the steps for constructing a structure-preserving POD-DEIM reduced system.

Algorithm 5 : POD-DEIM Structure-preserving DEIM reduction of NLPH systems

---

1: Select a positive-definite matrix $\mathbf{Q}$ (nominally approximating $\nabla_{\mathbf{x}}^2 H(\mathbf{x}_0)$) and define $h(\mathbf{x}) = H(\mathbf{x}) - \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}$.

2: Generate the trajectory $\mathbf{x}(t)$, (or ensemble of trajectories) and collect snapshots:
$\mathbb{X} = [\mathbf{x}(t_0), \mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_N)]$.
$\mathbb{F} = [\nabla_{\mathbf{x}}H(\mathbf{x}(t_0)), \nabla_{\mathbf{x}}H(\mathbf{x}(t_1)), \ldots, \nabla_{\mathbf{x}}H(\mathbf{x}(t_N))]$.
$\mathbb{G} = [\nabla_{\mathbf{x}}h(\mathbf{x}(t_0)), \nabla_{\mathbf{x}}h(\mathbf{x}(t_1)), \ldots, \nabla_{\mathbf{x}}h(\mathbf{x}(t_N))]$.

3: Truncate the SVDs of the snapshot matrices, $\mathbb{X}$, $\mathbb{F}$, and $\mathbb{G}$ to obtain, respectively, the POD basis matrices $\widetilde{\mathbf{V}}_r$, $\widetilde{\mathbf{W}}_r$, and $\mathbf{U}_m$ used to approximate
$\mathbf{x}(t) \approx \mathbf{V}_r\mathbf{x}_r(t)$; $\nabla_{\mathbf{x}}H(\mathbf{x}) \approx \mathbf{W}_r\mathbf{f}_r(t)$; and $\nabla_{\mathbf{x}}h(\mathbf{x}(t)) \approx \mathbf{U}_m\mathbf{g}_m(t)$.
Choose bases so that $\mathbf{V}_r^T\mathbf{Q}\mathbf{V}_r = \mathbf{I}$, $\mathbf{V}_r^T\mathbf{W}_r = \mathbf{I}$, and $\mathbf{U}_m^T\mathbf{Q}\mathbf{U}_m = \mathbf{I}$.

4: From Algorithm 4, calculate DEIM indices, $\wp_1, \ldots, \wp_m$ and the DEIM projection, $\mathbb{P} = \mathbf{U}_m(\mathbf{E}^T\mathbf{U}_m)^{-1}\mathbf{E}^T$.

5: The DEIM-reduced port-Hamiltonian approximation becomes

$$\dot{\widehat{\boldsymbol{x}}}_r = (\mathbf{J}_r - \mathbf{R}_r)\nabla_{\widehat{\boldsymbol{x}}_r}\widehat{H}_r(\widehat{\boldsymbol{x}}_r) + \mathbf{B}_r\mathbf{u}(t)$$

$$\widehat{\boldsymbol{y}}_r = \mathbf{B}_r^T\nabla_{\widehat{\boldsymbol{x}}_r}\widehat{H}_r(\widehat{\boldsymbol{x}}_r)$$

with $\mathbf{J}_r = \mathbf{W}_r^T\mathbf{J}\mathbf{W}_r$, $\mathbf{R}_r = \mathbf{W}_r^T\mathbf{R}\mathbf{W}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, and

$$\nabla_{\widehat{\boldsymbol{x}}_r}\widehat{H}_r(\widehat{\boldsymbol{x}}_r) = \mathbf{V}_r^T\nabla_{\mathbf{x}}\widehat{H}(\mathbf{V}_r\widehat{\boldsymbol{x}}_r) = \mathbf{x}_r + \mathbf{V}_r^T\mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\mathbf{V}_r\widehat{\boldsymbol{x}}_r).$$

Note that we can apply this structure-preserving DEIM approach to bases $\mathbf{V}_r$ and $\mathbf{W}_r$ associated with $\mathcal{H}_2^\varepsilon$-PH bases obtained from Algorithm 2, instead of POD-PH bases obtained from Algorithm 1. In the numerical examples of §3.4, we consider both strategies, refering to the variant using Algorithm 2 as $\mathcal{H}_2^\varepsilon$-DEIM-PH.

3.4. Numerical examples. We consider two system models: the nonlinear LC ladder network from Section 2.4 and a Toda lattice model with exponential interactions. We will illustrate the performance of the structure-preserving DEIM-based model reduction approaches introduced above using the POD-DEIM-PH, $\mathcal{H}_2^\varepsilon$-DEIM-PH and the hybrid POD-$\mathcal{H}_2^\varepsilon$-PH bases.

3.4.1. Ladder network. The N-stage nonlinear ladder network studied in Section 2.4 will be considered again. We will investigate here the effect of incorporating the structure-preserving DEIM approximation in reducing the nonlinear term. As in Section 2.4, we use a Gaussian pulse as the training input and test the reduced models on both the Gaussian pulse and a sinusoidal input. Figure 3.1 shows that POD-DEIM-PH and $\mathcal{H}_2^\varepsilon$-DEIM-PH reduced systems capture the behavior of the original output responses accurately for both inputs, repeating the success of the POD-PH and $\mathcal{H}_2^\varepsilon$-PH reduced models. Notice from Figure 3.1 that the linearized system does not give a good approximation to the true output response of the nonlinear system, as pointed out in Section 2. Notice also that, the accuracy of the POD-DEIM-PH reduced system ($r = 6$, $m = 24$) is very close to the accuracy of the POD-PH reduced system ($r = 6$); the POD-PH reduced system is slightly more accurate as expected; however the POD-DEIM-PH reduced system has much lower computational complexity as explained in the previous section; thus the structure-preserving DEIM approximation reduces the computational complexity yet does not degrade the accuracy of the reduced model. As the DEIM dimension increases, the POD-DEIM-PH yields the same accuracy as POD-PH as shown in Figure 3.2 for a fixed POD dimension of $r = 6$. Figure 3.3 shows the convergence of the average relative errors. Analogous to the previous numerical test in Section 2.4, we also consider projection bases constructed using the hybrid POD-$\mathcal{H}_2^\varepsilon$-PH approach combined with DEIM approximation. As one may see in Figure 3.4, this hybrid approach is more accurate than the ones constructed using only POD-bases (green lines) or using only $\mathcal{H}_2^\varepsilon$-bases (blue lines).

FIG. 3.1. *Ladder Network: ROM responses to Gaussian pulse (left) and sinusoidal input (right).*
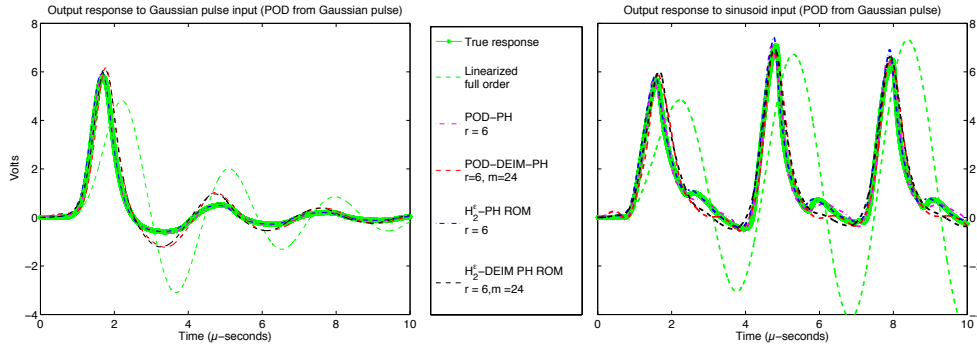
3.4.2. Toda Lattice. A Toda lattice model describes the motion of a chain of particles, each one connected to its nearest neighbors with 'exponential springs'. The equations of motion for the N-particle Toda lattice with such exponential interactions can be written in the form of a nonlinear port-Hamiltonian system as in (1.1) with

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n\times n},\ \mathbf{R} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}(\gamma_1,\ldots,\gamma_N) \end{bmatrix} \in \mathbb{R}^{n\times n},\ \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_1 \end{bmatrix} \in \mathbb{R}^n,$$

FIG. 3.2. *Ladder Network: Outputs from $\mathcal{H}_2^\varepsilon$-PH reduced system (dimension $r = 6$) and $\mathcal{H}_2^\varepsilon$-DEIM-PH reduced systems ($\mathcal{H}_2^\varepsilon$ dimension $r = 6$ with DIEM dimensions $m = 6, 12, 14, 48$)–similarly for the plot of POD-PH and POD-DEIM-PH reduced systems.*
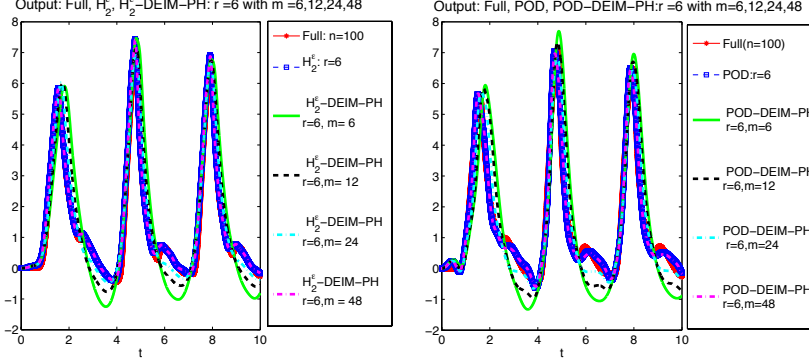

FIG. 3.3. *Ladder Network: Average relative errors of the reduced systems constructed from POD-PH, POD-DEIM-PH, $\mathcal{H}_2^\varepsilon$-PH, and $\mathcal{H}_2^\varepsilon$-DEIM-PH bases (with dimensions $r = 2, 4, ..., 18$ and DEIM dimension $m = 3r$).*
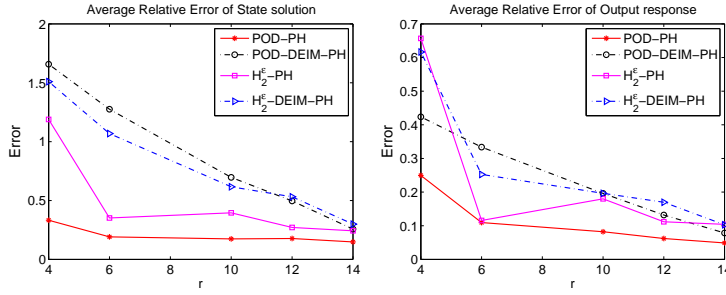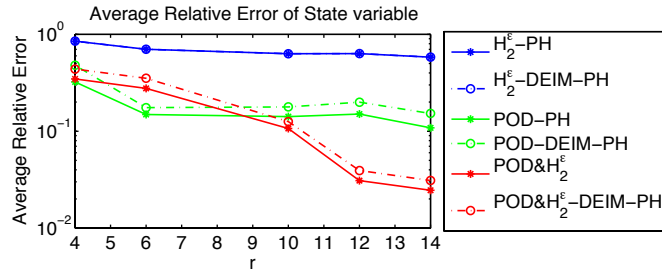

FIG. 3.4. *Ladder Network: Average relative errors using the POD-PH, POD-DEIM-PH, $\mathcal{H}_2^\varepsilon$-PH, $\mathcal{H}_2^\varepsilon$-DEIM-PH, and the hybrid POD-$\mathcal{H}_2^\varepsilon$-PH and POD-$\mathcal{H}_2^\varepsilon$-DEIM-PH bases (with $r = 2, 4, ..., 18$ and DEIM dimension $m = 3r$).*

where $n = 2N$ and

$$\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} \in \mathbb{R}^n, \quad \mathbf{q} = [q_1, \ldots, q_N]^T, \quad \mathbf{p} = [p_1, \ldots, p_N]^T$$

with q_j and p_j being, respectively, the displacement of the j-th particle from its equilibrium position and the momentum of the j-th particle for $j = 1, \ldots, N$. In this example; we use $N = 1000$ (i.e., the system dimension is $n = 2000$), $\mathbf{x}_0 = \mathbf{0}$, and

$\gamma_j = 0.1$, for $j = 1, \ldots, N$. The corresponding nonlinear Hamiltonian is given by

$$H(\mathbf{x}) = H([\mathbf{q}; \mathbf{p}]) = \sum_{k=1}^{N} \frac{1}{2} p_k^2 + \sum_{k=1}^{N-1} \exp(q_k - q_{k+1}) + \exp(q_N) - q_1 - N.$$

The decomposition $H(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + h(\mathbf{x})$ is used with

$$\mathbf{Q} = \left[\begin{array}{cc} \mathbf{Q}_0 & 0 \\ 0 & \mathbf{I} \end{array}\right], \quad \mathbf{Q}_0 = \left[\begin{array}{ccccccc} 1 & -1 & 0 & \ldots & & 0 & 0 \\ -1 & 2 & -1 & & & & 0 \\ 0 & -1 & 2 & \ddots & & & 0 \\ \vdots & & & \ddots & \ddots & & \\ & & & & & -1 & 0 \\ 0 & & & \ldots & -1 & 2 & -1 \\ 0 & 0 & & & 0 & -1 & 2 \end{array}\right],$$

and $h(\mathbf{x}) = \sum_{k=1}^{N-1}(q_k - q_{k+1})^3 \varphi(q_k - q_{k+1})$ where $\varphi(z) = \frac{1}{2}\int_0^1 \theta^2 \, e^{(1-\theta)z} \, d\theta$. The system was excited with two different inputs: $u(t) = 0.1$ and $u(t) = 0.1\sin(t)$. Figure 3.5 shows that the outputs due to both inputs are accurately approximated by the outputs of the POD-PH and POD-DEIM-PH reduced systems. The average relative errors and the simulation times relative to the simulation time of the full model are illustrated in Figures 3.6 and 3.7. Notice that the accuracy of the POD-DEIM-PH reduced model captures that of the POD-PH model as the DEIM dimension m increases. Note also that thePOD-DEIM-PH reduced model cuts the simulation time by nearly 96% while retaining accuracy.

FIG. 3.5. *Toda Lattice: The outputs from inputs $u(t) = 0.1$ and $u(t) = 0.1\sin(t)$.*
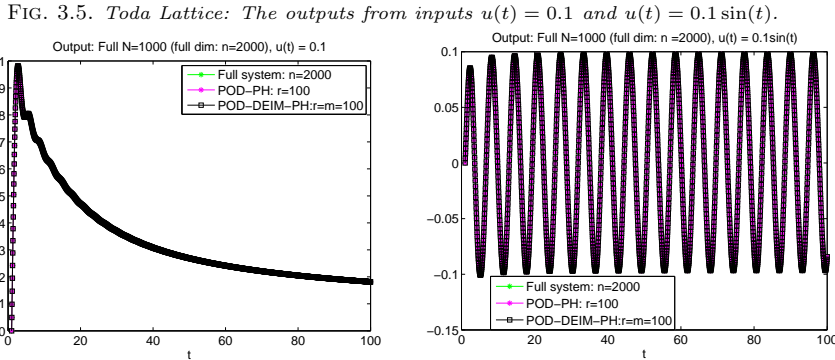
3.5. An *a priori* error bound for PH-preserving DEIM reduction. We derive error bounds for a DEIM-based reduced order model preserving PH-structure by estimating additional errors that occur by introducing the symmetrized-DEIM approximations into the structure-preserving reduction framework of §2. In particular, suppose that reduction bases, $\mathbf{V}_r$ and $\mathbf{W}_r$, have been chosen in some manner (e.g., as described as in §2.1, §2.2, or §2.3, say), and then are used to produce a reduced model that preserves the PH structure of the original system:

$$\begin{aligned} \dot{\mathbf{x}}_r &= (\mathbf{J}_r - \mathbf{R}_r)\nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r &= \mathbf{B}_r^T\nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r). \end{aligned} \tag{3.6}$$

FIG. 3.6. *Toda Lattice with input $u(t) = 0.1$: Relative errors of the outputs and the online CPU times for* POD-PH *and* POD-DEIM-PH *reduced systems with POD basis dimension r and DEIM dimension $m = r, m_1, m_2$ where $m_1 = r+$* `ceil(r/3)`*, $m_2 = r+$*`ceil(2r/3)`*; full-order system: $n = 2000$.*
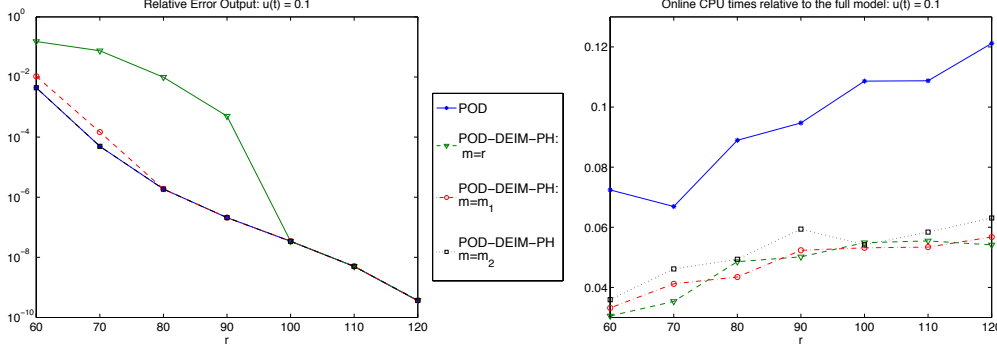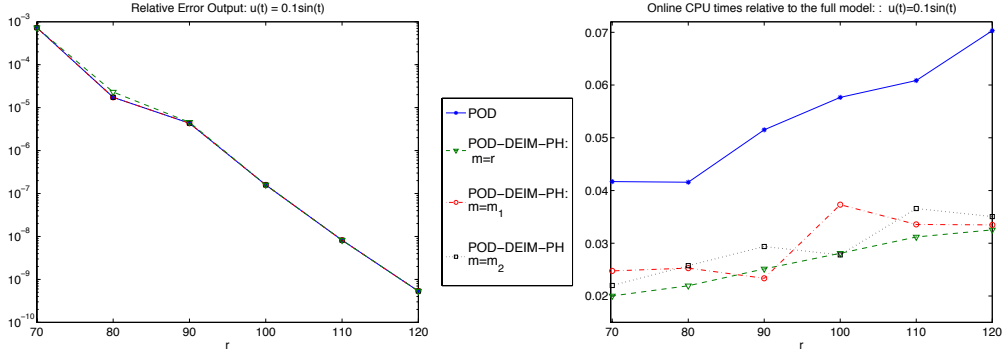

FIG. 3.7. *Toda Lattice with input $u(t) = 0.1 \sin(t)$: Relative errors of the outputs and the online CPU times for* POD-PH *and* POD-DEIM-PH *reduced systems with POD basis dimension r and DEIM dimension $m = r, m_1, m_2$ where $m_1 = r+$* `ceil(r/3)`*, $m_2 = r+$*`ceil(2r/3)`*; full-order system: $n = 2000$.*

We then symmetrically "sparsify" the nonlinear interactions in the Hamiltonian gradient evaluation and introduce a further DEIM-reduction as described in §3.3 and Algorithm 5 in order to produce:

$$\dot{\hat{\boldsymbol{x}}}_r = (\mathbf{J}_r - \mathbf{R}_r)\nabla_{\hat{\boldsymbol{x}}_r}\widehat{H}_r(\hat{\boldsymbol{x}}_r) + \mathbf{B}_r\mathbf{u}(t)$$
$$\hat{\boldsymbol{y}}_r = \mathbf{B}_r^T\nabla_{\hat{\boldsymbol{x}}_r}\widehat{H}_r(\hat{\boldsymbol{x}}_r) \tag{3.7}$$

where $\mathbf{x}(0) = \hat{\boldsymbol{x}}(0)$ and the DEIM Hamiltonian, $\widehat{H}$, has been defined in (3.3).

THEOREM 3.4. *Let $\mathbf{x}_r(t)$ and $\mathbf{y}_r(t)$ be the state trajectory and output associated with the reduced port-Hamiltonian system (3.6). Suppose a* DEIM *basis of order m has been chosen and used to define a* DEIM *projection $\mathbb{P}$ and* DEIM *Hamiltonian $\widehat{H}(\mathbf{x})$ as introduced in (3.3). Recalling the discussion of §3.2, suppose $\Omega \subset \mathcal{V}_r$ contains the trajectories $\mathbf{V}_r\mathbf{x}_r(t)$ and $\mathbb{P}^T\mathbf{V}_r\mathbf{x}_r$ and that $\varepsilon_h > 0$ satisfies*

$$\sup_{\boldsymbol{\xi} \in \Omega} \|\nabla_{\mathbf{x}}h(\boldsymbol{\xi}) - \mathbb{P}\nabla_{\mathbf{x}}h(\mathbb{P}^T\boldsymbol{\xi})\|_{\mathbf{Q}} \leq \varepsilon_h$$

Let $\hat{\boldsymbol{x}}_r(t)$ and $\hat{\boldsymbol{y}}_r(t)$ be the state trajectory and output associated with the DEIM-*reduced port-Hamiltonian system (3.7).*

Then with

$$\alpha = \mathcal{L}_{\mathbf{Q}}[\mathbf{G}] - \rho_{min}, \quad \mathbf{G}(\boldsymbol{\xi}) = \boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \boldsymbol{\xi}), \quad and \quad \rho_{min} = \min \lambda(\mathbf{R}_r)$$

$$\beta = \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}, \qquad \gamma = 1 + L_{\mathbf{Q}}[\mathbb{P}\, \nabla_{\mathbf{x}} h \circ \mathbb{P}^T], \qquad \delta = \|\mathbf{B}\|\, \|\boldsymbol{\mathcal{P}}_r\|$$

we have, for $\alpha \neq 0$,

$$\|\mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)\| \leq \tfrac{\beta}{\alpha}\, (e^{\alpha t} - 1)\, \varepsilon_h$$
$$\|\mathbf{y}_r(t) - \hat{\boldsymbol{y}}_r(t)\| \leq \delta \left(1 + \tfrac{\beta \gamma}{\alpha}\, (e^{\alpha t} - 1) \right) \varepsilon_h, \tag{3.8}$$

whereas for $\alpha = 0$,

$$\|\mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)\| \leq \beta\, t\, \varepsilon_h$$
$$\|\mathbf{y}_r(t) - \hat{\boldsymbol{y}}_r(t)\| \leq \delta\, (1 + \beta\, t)\, \varepsilon_h.$$

Proof. Let $\mathbf{e}_r(t) := \mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)$ be the difference between state trajectories generated by (3.6) and (3.7). Define $\mathbf{f}(\mathbf{x}) = \nabla_{\mathbf{x}} h(\mathbf{x})$ and observe that

$$\frac{d}{dt}\|\mathbf{e}_r(t)\| = \left\langle \frac{\mathbf{e}_r(t)}{\|\mathbf{e}_r(t)\|}, \dot{\mathbf{e}}_r(t) \right\rangle = \left\langle \frac{\mathbf{e}_r(t)}{\|\mathbf{e}_r(t)\|}, \mathbf{A}_r \left[\nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r) - \nabla_{\hat{\boldsymbol{x}}_r} \widehat{H}_r(\hat{\boldsymbol{x}}_r) \right] \right\rangle$$
$$= \left\langle \frac{\mathbf{e}_r(t)}{\|\mathbf{e}_r(t)\|}, \mathbf{A}_r \left[\mathbf{e}_r + \mathbf{V}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \right] \right\rangle$$

We have

$$\left\langle \mathbf{e}_r, \mathbf{A}_r \left[\mathbf{e}_r + \mathbf{V}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \right] \right\rangle$$
$$= -\langle \mathbf{e}_r, \mathbf{R}_r \mathbf{e}_r \rangle + \left\langle \mathbf{V}_r \mathbf{e}_r, \boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \right\rangle_{\mathbf{Q}}$$
$$= -\langle \mathbf{e}_r, \mathbf{R}_r \mathbf{e}_r \rangle + \left\langle \mathbf{V}_r \mathbf{e}_r, \boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \left(\mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \right\rangle_{\mathbf{Q}}$$
$$\quad + \left\langle \mathbf{V}_r \mathbf{e}_r, \boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r) \right) \right\rangle_{\mathbf{Q}}$$
$$\leq -\rho_{min}\|\mathbf{e}_r\|^2 + \mathcal{L}_{\mathbf{Q}}[\mathbf{G}]\, \|\mathbf{e}_r\|^2$$
$$\quad + \|\mathbf{e}_r\|\, \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}\, \|\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r)\|_{\mathbf{Q}}$$

Thus,

$$\frac{d}{dt}\|\mathbf{e}_r(t)\| \leq (\mathcal{L}_{\mathbf{Q}}[\mathbf{G}] - \rho_{min})\, \|\mathbf{e}_r(t)\|$$
$$\quad + \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}}\, \|\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r(t)) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r(t))\|_{\mathbf{Q}}.$$

If $\alpha = \mathcal{L}_{\mathbf{Q}}[\mathbf{G}] - \rho_{min} = 0$ then trivially,

$$\|\mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)\| \leq \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \int_0^t \|\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r(\tau)) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r(\tau))\|_{\mathbf{Q}}\, d\tau \leq \beta\, t\, \varepsilon_h.$$

If $\alpha \neq 0$ then by Gronwell's inequality,

$$\|\mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)\| \leq \|\boldsymbol{\mathcal{P}}_r \mathcal{A} \boldsymbol{\mathcal{P}}_r^T\|_{\mathbf{Q}} \int_0^t e^{\alpha(t-\tau)}\, \|\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r(\tau)) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r(\tau))\|_{\mathbf{Q}}\, d\tau$$
$$\leq \frac{\beta}{\alpha}\, (e^{\alpha t} - 1)\, \varepsilon_h$$

This expression holds with a positive bound regardless of whether $\alpha > 0$ or $\alpha < 0$.

The resulting difference in output maps may be directly bounded:

$$
\begin{aligned}
\|\mathbf{y}_r - \hat{\boldsymbol{y}}_r\| &= \|\mathbf{B}_r^T \left(\nabla_{\mathbf{x}_r} H_r(\mathbf{x}_r) - \nabla_{\hat{\boldsymbol{x}}_r} \widehat{H}_r(\hat{\boldsymbol{x}}_r) \right) \| \\
&= \|\mathbf{B}_r^T \left(\mathbf{x}_r - \hat{\boldsymbol{x}}_r + \mathbf{V}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \right) \| \\
&\leq \|\mathbf{B}_r^T \mathbf{e}_r\| + \|\mathbf{B}_r^T \mathbf{V}_r^T \left(\nabla_{\mathbf{x}} h(\mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r) \right) \| \\
&\qquad + \|\mathbf{B}_r^T \mathbf{V}_r^T \left(\mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \mathbf{x}_r) - \mathbb{P} \nabla_{\mathbf{x}} h(\mathbb{P}^T \mathbf{V}_r \hat{\boldsymbol{x}}_r) \right) \| \\
&\leq \|\mathbf{B}\| \, \|\boldsymbol{\mathcal{P}}_r\| \, \left(1 + L_{\mathbf{Q}}[\mathbb{P} \nabla_{\mathbf{x}} h \circ \mathbb{P}^T] \right) \|\mathbf{x}_r(t) - \hat{\boldsymbol{x}}_r(t)\| + \|\mathbf{B}\| \, \|\boldsymbol{\mathcal{P}}_r\| \, \varepsilon_h,
\end{aligned}
$$

which leads to the second conclusions respectively for $\alpha \neq 0$ and $\alpha = 0$. $\square$

4. Conclusions. We have introduced a structure-preserving projection-based model reduction framework for large-scale multi-input/multi-output nonlinear port-Hamiltonian systems. We constructed projection subspaces using three different approaches: POD, $\mathcal{H}_2$-based, and a hybrid POD-$\mathcal{H}_2$ based. We showed that for the same reduced order, the hybrid basis significantly outperforms the other two. We introduced a modification of DEIM within this framework to approximate the nonlinear part of the Hamiltonian gradient symmetrically. In all cases, the resulting reduced system preserves port-Hamiltonian structure, and thus retains the stability and passivity of the original system. We have derived the corresponding *a priori* error bounds of the state variables and outputs by using an application of generalized logarithmic norms for unbounded nonlinear operators. The effectiveness of the proposed approaches were shown on a nonlinear ladder network and a Toda lattice model.

REFERENCES

[1] P. ASTRID, S. WEILAND, K. WILLCOX, AND T. BACKX, *Missing point estimation in models described by proper orthogonal decomposition*, IEEE Transactions on Automatic Control, (2008), pp. 2237–2251.

[2] M. BARRAULT, Y. MADAY, N. NGUYEN, AND A. PATERA, *An "empirical interpolation" method: Application to efficient reduced-basis discretization of partial differential equations*, Comptes Rendus Mathématique. Académie des Sciences. Paris, I (2004), pp. 339–667.

[3] C. BEATTIE AND S. GUGERCIN, *Structure-preserving model reduction for nonlinear port-hamiltonian systems*, in Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on, Dec., pp. 6564–6569.

[4] K. CARLBERG, C. FARHAT, J. CORTIAL, AND D. AMSALLEM, *The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows*, Journal of Computational Physics, 242 (2013), pp. 623–647.

[5] K. CARLBERG, R. TUMINARO, AND P. BOGGS, *Preserving lagrangian structure in nonlinear model reduction with application to structural dynamics*, SIAM Journal on Scientific Computing, 37 (2015), pp. B153–B184.

[6] S. CHATURANTABUT AND D. C. SORENSEN, *Nonlinear model reduction via discrete empirical interpolation*, SIAM Journal on Scientific Computing, 32 (2010), pp. 2737–2764.

[7] Z. DRMAČ AND S. GUGERCIN, *A new selection operator for the discrete empirical interpolation method–improved a priori error bound and extensions*, SIAM Journal on Scientific Computing. Accepted to appear. Available as http://arxiv.org/abs/1505.00370, (2015).

[8] V. DUINDAM, A. MACCHELLI, S. STRAMIGIOLI, AND H. BRUYNINCKX, *Modeling and control of complex physical systems*, Springer, 2009.

[9] R. EVERSON AND L. SIROVICH, *The Karhunen-Loeve Procedure for Gappy Data*, Journal of the Optical Society of America, 12 (1995), pp. 1657–1664.

[10] K. FUJIMOTO AND H. KAJIURA, *Balanced realization and model reduction of port-Hamiltonian systems*, in American Control Conference, 2007, 2007, pp. 930–934.

[11] K. FUJIMOTO AND J. SCHERPEN, *Balanced realization and model order reduction for nonlinear systems based on singular value analysis*, SIAM Journal on Control and Optimization, 48 (2010), pp. 4591–4623.

[12] S. GUGERCIN, A. C. ANTOULAS, AND C. A. BEATTIE, $\mathcal{H}_2$ *model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638.

[13] S. GUGERCIN, R. POLYUGA, C. BEATTIE, AND A. VAN DER SCHAFT, *Interpolation-based $\mathcal{H}_2$ Model Reduction for port-Hamiltonian Systems*, in Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, PR China, 2009, pp. 5362–5369.

[14] S. GUGERCIN, R. POLYUGA, C. BEATTIE, AND A. VAN DER SCHAFT, *Structure-preserving tangential interpolation for model reduction of port-hamiltonian systems*, Automatica, 48 (2012), pp. 1963–1974.

[15] A. HOCHMAN, B. BOND, AND J. WHITE, *A stabilized discrete empirical interpolation method for model reduction of electrical, thermal, and microelectromechanical systems*, in Design Automation Conference (DAC), 2011 48th ACM/EDAC/IEEE, June, pp. 540–545.

[16] J. LUMLEY, *The Structures of Inhomogeneous Turbulent Flow*, Atmospheric Turbulence and Radio Wave Propagation, (1967), pp. 166–178.

[17] B. PEHERSTORFER, D. BUTNARU, K. WILLCOX, AND H.-J. BUNGARTZ, *Localized discrete empirical interpolation method*, SIAM Journal on Scientific Computing, 36 (2014), pp. A168–A192.

[18] B. PEHERSTORFER AND K. WILLCOX, *Online adaptive model reduction for nonlinear systems via low-rank updates*, SIAM Journal on Scientific Computing, 37 (2015), pp. A2123–A2150.

[19] J. PHILLIPS, *Projection frameworks for model reduction of weakly nonlinear systems*, in Proceedings of the 37th Annual Design Automation Conference, ACM, 2000, pp. 184–189.

[20] R. V. POLYUGA, *Model Reduction of Port-Hamiltonian Systems*, PhD thesis, University of Groningen, 2010.

[21] R. V. POLYUGA AND A. VAN DER SCHAFT, *Structure preserving model reduction of port-Hamiltonian systems by moment matching at infinity*, Automatica, 46 (2010), pp. 665–672.

[22] ——, *Structure preserving moment matching for port-Hamiltonian systems: Arnoldi and Lanczos*, To appear in IEEE Transactions on Automatic Control, (2010).

[23] M. REWIENSKI AND J. WHITE, *A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices*, Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, 22 (2003), pp. 155–170.

[24] J. SCHERPEN AND A. VAN DER SCHAFT, *A structure preserving minimal representation of a nonlinear port-Hamiltonian system*, in Decision and Control, 2008, 47th IEEE Conference on, 2008, pp. 4885–4890.

[25] L. SIROVICH, *Turbulence and the dynamics of coherent structures. Part 1: Coherent structures*, Quarterly of Applied Mathematics, 45 (1987), pp. 561–571.

[26] G. SÖDERLIND, *The logarithmic norm. history and modern theory*, BIT Numerical Mathematics, 46 (2006), pp. 631–652. 10.1007/s10543-006-0069-9.

[27] D. B. SZYLD, *The Many Proofs of an Identity on the Norm of Oblique Projections*, Numerical Algorithms, 42 (2006), pp. 309–323.

[28] A. VAN DER SCHAFT, *Port-Hamiltonian systems: an introductory survey*, in Proceedings of the International Congress of Mathematicians Vol. III, Madrid, M. Sanz-Sole, J. Soria, J. L. Varona, and J. Verdera, eds., Madrid, Spain, 2006, European Mathematical Society Publishing House (EMS Ph), pp. 1339–1365.

[29] J. C. WILLEMS, *Dissipative dynamical systems part i: General theory*, Archive for rational mechanics and analysis, 45 (1972), pp. 321–351.

[30] H. ZWART AND B. JACOB, *Distributed-parameter port-hamiltonian systems*, tech. rep., Technischer Bericht, Lehrstuhl für angewandte Mathematik, Universität Twente, Niederlande, 2009.