21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

On model reduction of physical network systems

Arjan van der Schaft

Abstract— **A method for model order reduction of a large class of physical network systems is discussed. The method is based on clustering of the vertices of the underlying graph, and yields a reduced order model within the same class. This is illustrated on a number of physical examples, including mass-spring-damper systems, chemical reaction networks and hydraulic networks.**

I. Introduction

This paper is concerned with structure-preserving model order reduction of large-scale physical network systems. More specifically, we want to obtain reduced-order models which fall into the same class as the original system. The first class of systems which is considered consists of systems given by a graph, with state variables associated to the vertices and couplings corresponding to the edges of the graph. This includes mass-damper systems, hydraulic networks and isothermal chemical reaction networks.

The basic technique for performing structure-preserving model reduction is *clustering*; the vertices of the graph are merged into a number of cells, corresponding to a *partition* of the set of all vertices. The reduced system will be a system of the same form, defined on the graph whose vertex set is defined by the cells of the original graph.

In the final section we extend our approach to physical networks with second-order dynamics, including mass-spring-damper systems. In state space form this amounts to systems with state variables associated to the vertices, as well as to the edges.

The idea of using clustering for model reduction has been explored before in a number of papers, see e.g. [7], [8], [13], [11] and the references quoted therein. The current paper is particularly influenced by the recent paper [11], as well as by [1], [17] and [9].

Arjan van der Schaft is with the Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, The Netherlands, a.j.van.der.schaft@rug.nl

A. Background on graphs

We recall from e.g. [2], [6] a few standard definitions and facts. A graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, is defined by a set $\mathcal{V}$ of vertices (nodes) and a set $\mathcal{E}$ of edges (links, branches), where $\mathcal{E}$ is identified with a set of unordered pairs $\{i, j\}$ of vertices $i, j \in \mathcal{V}$. We allow for *multiple* edges between vertices, but not for *self-loops* $\{i, i\}$. By endowing the graph with an orientation we obtain a *directed graph*.

A directed graph with n vertices and k edges is specified by its *incidence matrix*. This matrix, denoted by B, is an $n \times k$ matrix where every column corresponds to an edge of the graph, and contains exactly one -1 at the row corresponding to its tail vertex and one $+1$ at the row corresponding to its head vertex, while the other elements are 0. For any diagonal positive semi-definite $k \times k$ matrix R we define the *weighted Laplacian matrix* of the graph as $L := BRB^T$, where the nonnegative diagonal elements $r_1, \cdots, r_k$ of the matrix R are the weights of the edges. It is well-known [2] that L is *independent* of the orientation of the graph, and thus is associated with the undirected[1] graph $\mathcal{G}$.

II. Physical network systems

The first class of systems we will consider in this paper are described as follows. Consider a graph $\mathcal{G}$ as above, and a linear system with state vector x and input vector u of the form

$$\dot{x} = -BRB^T M^{-1} x + Eu, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, \tag{1}$$

[1] Alternatively [2], [6] L is defined as the difference of the diagonal matrix of degrees of the vertices and the adjacency matrix of the undirected graph.

ISBN: 978-90-367-6321-9

where R is a positive semi-definite and M a positive diagonal matrix. Furthermore, E is an $n \times m$ matrix, with every column containing exactly one $+1$ or -1 element at the row corresponding to the vertex (terminal) where control takes place. This means that the i-th component x_i of the vector x is the state variable associated to the i-th vertex of the graph.

The set-up can be easily extended (i.e., by using Kronecker products) to the situation that the scalar variable x_i is replaced by a vector in some higher-dimensional physical space, e.g., $\mathbb{R}^3$; see the remarks later on.

The system (1) is an example of a *port-Hamiltonian system* on a graph, as discussed in [17]. In particular, its Hamiltonian (total energy) is given by $H(x) = \frac{1}{2}x^T M^{-1} x$, while the Laplacian matrix BRB^T determines the resistive structure. See e.g. [16], [4], [14], [15] for further information regarding port-Hamiltonian systems.

A variety of physical network systems is of this form as illustrated by the following examples.

Example 2.1 (Mass-damper systems): Consider a mass-damper system with n masses with mass parameters $m_1, \cdots, m_n$ and k linear dampers with damping coefficients $r_1, \cdots, r_k$. Define the diagonal matrices $M = \mathrm{diag}(m_1, \cdots, m_n), R = \mathrm{diag}(r_1, \cdots, r_k)$ and the state x to be the vector of momenta of the masses. Furthermore, the non-zero rows of E correspond to masses which are affected by external forces u.

Example 2.2 (Single-species reaction networks): Consider a chemical reaction network governed by mass action kinetics with *single-species reactions* (that is, reactions that involve only one chemical species as substrate and one chemical species as product). In case the network is *detailed-balanced* the dynamics of the vector x of concentrations in the positive orthant takes the form [18]

$$\dot{x} = -BRB^T \frac{x}{x^*}, \quad x^* \in \mathbb{R}^n_{>0},$$

where B is the incidence matrix of the network, and $x^* \in \mathbb{R}^n_{>0}$ is the assumed thermodynamic equilibrium, and where $\frac{x}{x^*}$ denotes elementwise division. The j-th element of the diagonal matrix R is determined by the forward and reverse reaction constants of the j-th reaction and x^* (see [18] for details). Defining the diagonal matrix $M := \mathrm{diag}(x_1^*, \cdots, x_n^*)$ this can be written in the form (1), where Eu denotes the inflow or outflow of the chemical species corresponding to non-zero rows of E.

Example 2.3 (Hydraulic networks): Consider a hydraulic network between n fluid reservoirs whose storage is described by the elements of a vector x. Mass balance corresponds to $\dot{x} = Bf$ where $f \in \mathbb{R}^k$ is the flow through the k pipes linking the reservoirs. Let each storage variable x_i determine a pressure $P_i = \frac{x_i}{m_i}$ for a certain constant m_i. Assuming that the flow f_k is proportional to the difference between the pressure of the head reservoir and the pressure of the tail reservoir this leads to the equations (1), where again Eu denotes the direct inflow or outflow at the reservoirs corresponding to non-zero rows of E. Other transportation networks have a similar form.

Finally we mention that another example of (1) with $M = I_n$ (the $n \times n$ identity matrix) is formed by the standard symmetric consensus algorithm for multi-agent systems in continuous time, with Eu corresponding to leaders of the system (see e.g. [10]).

Before moving on to model reduction we will first introduce some abstractions which will come in handy later on, and are important for the interpretation of our approach. The state space of (1) given by $\mathbb{R}^n$ can be more abstractly defined as follows; cf. [17] for further information. It is given by the linear space Λ_0 of all functions from the vertex set $\mathcal{V}$ to $\mathbb{R}$. Obviously Λ_0 can be identified with $\mathbb{R}^n$. The matrix M^{-1} defines an inner product on Λ_0. As a consequence, any vector $M^{-1}x, x \in \Lambda_0$, can be considered to be an element of the dual space of Λ_0, which is denoted by Λ^0. For a mass-damper system, $v := M^{-1}x$ is the vector of velocities of the n masses. It follows that the system (1) can be represented in the state vector $v := M^{-1}x \in \Lambda^0$ as

$$\dot{v} = -M^{-1}BRB^T v + M^{-1}Eu, \quad v \in \Lambda^0 = \mathbb{R}^n, \tag{2}$$

or equivalently in the *gradient system* representation (with M defining an inner product on the space Λ^0)

$$M\dot{v} = -BRB^T v + Eu, \quad v \in \Lambda^0 = \mathbb{R}^n, u \in \mathbb{R}^m, \tag{3}$$

Furthermore, also the edge space $\mathbb{R}^k$ can be defined more abstractly as the linear space Λ_1 of functions from the edge set $\mathcal{E}$ to $\mathbb{R}$, with dual space denoted by Λ^1. It follows that the incidence matrix B defines a linear map (denoted by the same symbol) $B : \Lambda_1 \to \Lambda_0$ with adjoint map $B^T : \Lambda^0 \to \Lambda^1$. Furthermore, R can be considered to define an inner product on Λ^1, or equivalently, as a map $R : \Lambda^1 \to \Lambda_1$.

Using these abstractions it is straightforward to extend the dynamics (1) or (2) to other spatial domains than just the one-dimensional domain $\mathbb{R}$. Indeed, for any linear space $\mathcal{R}$ we can define Λ_0 as the set of functions from $\mathcal{V}$ to $\mathcal{R}$, and Λ_1 as the set of functions from $\mathcal{E}$ to $\mathcal{R}$. In this case we can identify Λ_0 with the tensor product $\mathbb{R}^n \otimes \mathcal{R}$ and Λ_1 with the tensor product $\mathbb{R}^k \otimes \mathcal{R}$. Furthermore, the incidence matrix B defines a linear map $B \otimes I : \Lambda_1 \to \Lambda_0$, where I is the identity map on $\mathcal{R}$. In matrix notation $B \otimes I$ equals the Kronecker product of the incidence matrix B and the identity matrix I. It is straightforward to extend the formulation of the dynamics (1) to the case of a general linear space $\mathcal{R}$ instead of $\mathbb{R}$. For $\mathcal{R} = \mathbb{R}^3$ this will describe the motion of (point)mass-damper systems in $\mathbb{R}^3$. An especially interesting generalization are multibody-systems in $\mathbb{R}^3$, in which case $\mathcal{R} = se(3)$ (the Lie algebra corresponding to the special Euclidean group in $\mathbb{R}^3$); see [17] for further information.

A. Dynamical analysis

We collect the following facts regarding the dynamics of systems of the form (1); see [17]. First we recall (see e.g. [2]) that $\mathbb{1} \in \ker B^T$, and that the graph $\mathcal{G}$ is *connected* if and only if

$$\ker B^T = \operatorname{span} \mathbb{1}$$

It follows that the quantity $\mathbb{1}^T x = x_1 + \cdots + x_n$ is a conserved quantity of the system.

Proposition 2.4: Consider (1) for $u = 0$. Assume that the graph is connected[2]. Then for every initial state x_0 there exists a unique equilibrium x_∞ of (1) satisfying $\mathbb{1}^T x_\infty = \mathbb{1}^T x_0$, and the solution $x(t)$ of (1) with initial condition $x(0) = x_0$ will exponentially converge to x_∞ for $t \to \infty$. Furthermore, x_∞ is uniquely determined by the conditions

$$\mathbb{1}^T x_\infty = \mathbb{1}^T x_0, \quad M^{-1} x_\infty \in \operatorname{span} \mathbb{1}.$$

Remark 2.5: Furthermore, for x_0 in the positive orthant $\mathbb{R}^n_{>0}$ it follows that the whole solution $x(t)$ will remain in the positive orthant; see e.g. the argument for a more general case in [18].

Example 2.6 (Examples continued): In the mass-damper system conservation of $\mathbb{1}^T x$ corresponds to conservation of total momentum, while $M^{-1} x_\infty \in \operatorname{span} \mathbb{1}$ means that the velocities $v_i = \frac{x_i}{m_i}$ will converge to the same value. For the hydraulic network the total fluid is a conserved quantity, while the system will converge to reservoirs with equal pressure.

In the chemical reaction network conservation of $\mathbb{1}^T x$ amounts to conservation, while $M^{-1} x_\infty \in \operatorname{span} \mathbb{1}$ means that the chemical potentials (up to a constant) $\mu_i = \frac{x_i}{x_i^*}$ will converge to the same value.

For constant but non-zero input $\bar{u}$ the state $\bar{x}$ is a steady-state of (1) if

$$0 = -BRB^T M^{-1} \bar{x} + E\bar{u}$$

Thus in order to have a steady-state $\bar{x}$ for $\bar{u}$ we need $E\bar{u}$ to be contained in $\operatorname{im} B$.

Proposition 2.7: Consider (1) for constant $\bar{u}$, satisfying[3] $E\bar{u} \in \operatorname{im} B$. Assume that the graph is connected. Then for every initial state x_0 there exists a unique steady-state $\bar{x}$ of (1) satisfying $\mathbb{1}^T \bar{x} = \mathbb{1}^T x_0$, and the solution $x(t)$ of (1) with initial condition $x(0) = x_0$ and constant input $\bar{u}$ will exponentially converge to $\bar{x}$ for $t \to \infty$.

III. Model reduction by clustering

This paper is concerned with *structure-preserving* model reduction of (1). That is, we

[2]Otherwise the same holds for every connected component of the graph.

[3]For a connected graph this is equivalent to $\mathbb{1}^T E\bar{u} = 0$.

want to reduce (1) to another system of the form (1) of lesser complexity, and thus obtain a suitable approximation of the system. The main idea of the proposed approach will be to *cluster* vertices, so as to obtain a graph with fewer vertices.

Consider a *partition* of the vertex set $\mathcal{V}$ of the graph $\mathcal{G}$ into $\hat{n}$ disjoint cells $C_1, C_2, \cdots, C_{\hat{n}}$, together with a corresponding $n \times \hat{n}$ *characteristic matrix* P. The columns of P equal the characteristic vectors of the cells; the characteristic vector of a cell C_i being defined as the vector with 1-s at the place of every vertex contained in the cell C_i, and 0 elsewhere. With some abuse of notation we will denote the partition simply by its characteristic matrix P.

Based on a partition P we reduce the system (1) to

$$\dot{\hat{x}} = -(P^T B R B^T P)(P^T M P)^{-1}\hat{x} + P^T E u \quad (4)$$

where $\hat{x} := P^T x \in \mathbb{R}^{\hat{n}}$ is the clustered state vector.

We observe that this is again a system of the form (1). In fact, the matrix $P^T B$ consists of column vectors containing exactly one $+1$ and one -1 together with some zero vectors (corresponding to edges which link vertices within a same cell). By leaving out the zero column vectors from $P^T B$ we thus obtain an $\hat{n} \times \hat{k}$ matrix $\hat{B}$, which is the incidence matrix of the *reduced graph* $\hat{\mathcal{G}}$ with vertices being the cells of the original graph, and with edges the union of all the edges between vertices in different cells (leaving out edges *within* cells). Correspondingly we define $\hat{R}$ as the $\hat{k} \times \hat{k}$ diagonal matrix obtained from R by leaving out the rows and columns corresponding to edges between vertices in a same cell. Finally we define the $\hat{n} \times \hat{n}$ diagonal matrix $\hat{M} = P^T M P$, and $\hat{E} = P^T E$. It follows that the reduced system (4) is also given as

$$\dot{\hat{x}} = -\hat{B}\hat{R}\hat{B}^T\hat{M}^{-1}\hat{x} + \hat{E}u, \quad \hat{x} \in \mathbb{R}^{\hat{n}} \quad (5)$$

Remark 3.1: Note that the reduced system (5) will easily have multiple edges between its vertices, that is, between the cells of the original system (1). Of course, multiple edges between two vertices can be combined into a single edge with weight r given by the *sum* of the weights of the multiple edges. However the analysis in the sequel is independent of this.

Remark 3.2: The main difference with the reduced model proposed in [11] is that in the latter paper the case $M = I$ ($n \times n$ identity matrix) is considered[4], while moreover the reduced system is also required to have $\hat{M} = I$ ($\hat{n} \times \hat{n}$ identity matrix). This can be done, but at the expense of having a weighted Laplacian matrix which is *not* anymore symmetric, and therefore not anymore of the form $\hat{B}\hat{R}\hat{B}^T$. (The obtained non-symmetric weighted Laplacian matrix corresponds to a reduced graph which is *directed*.)

Note that in case $u = 0$ by Proposition 2.4 there exists for every initial state $\hat{x}_0$ of the reduced system a unique equilibrium $\hat{x}_\infty$ determined by the conditions $\mathbb{1}^T\hat{x}_\infty = \mathbb{1}^T\hat{x}_0, \hat{M}^{-1}\hat{x}_\infty \in \operatorname{span}\mathbb{1}$ to which the system will exponentially converge.

Example 3.3 (Examples continued): In the mass-damper example the i-th component of $\hat{x}$ denotes the total momentum of the masses contained in cell C_i, while the i-th diagonal element of $\hat{M}$ is the total mass of the mass contained in cell C_i. The solution $\hat{x}(t)$ for $\hat{x}(0) = \hat{x}_0$ converges to the unique $\hat{x}_\infty$ determined by $\mathbb{1}^T\hat{x}_\infty = \mathbb{1}^T\hat{x}_0, \hat{M}^{-1}\hat{x}_\infty \in \operatorname{span}\mathbb{1}$. This means that the velocities of the clustered masses of cells $C_1, \cdots, C_{\hat{n}}$ (all masses within a cell rigidly interconnected, and leaving out intermediate dampers) will converge to a common value. Note that the velocity v_i of cell C_i is defined as $v_i = \sum x_j / \sum m_j$, with the summation ranging over the indices of the vertices in cell C_i.

Similar interpretations hold for the other two examples.

A *Petrov-Galerkin* interpretation interpretation of the reduced model (4) can be given as follows. Recall that a Petrov-Galerkin reduction of a linear set of differential equations $\dot{x} = Ax$ is given as $\dot{\hat{x}} = W^T A V \hat{x}$, where V and W are matrices of dimension $n \times \hat{n}$ such that $W^T V = I_{\hat{n}}$. Now the reduced system (4) is a Petrov-Galerkin reduction

[4]Corresponding to the standard continuous-time consensus algorithm.

of (1) with $W := P$, and

$$V := MP\hat{M}^{-1} = MP(P^TMP)^{-1} \qquad (6)$$

It follows that $W^TV = I_{\hat{n}}$. Note furthermore that the matrix $V = MP\hat{M}^{-1}$ defined in (6) equals the *Moore-Penrose pseudo-inverse* of the clustering map $\hat{x} = P^Tx$, with respect to the inner products M^{-1} on the space of full-order state vectors $x \in \mathbb{R}^n$ and $\hat{M}^{-1}$ on the space of reduced state vectors $\hat{x} \in \mathbb{R}^{\hat{n}}$. Furthermore, we note that $V^TV = P^TP$ is a diagonal matrix, and hence V is orthogonal, while also

$$W^TM^{-1}W = \hat{M}^{-1},$$

implying orthogonality of W with respect to the inner product defined by M^{-1}. Using the abstractions introduced earlier on we note that the linear maps V and W are actually defined as maps

$$V : \hat{\Lambda}_0 \to \Lambda_0, \quad W : \hat{\Lambda}^0 \to \Lambda^0$$

where $\hat{\Lambda}_0$ is the vertex space of the reduced graph, with dual space $\hat{\Lambda}^0$. Moreover, note that $M : \Lambda^0 \to \Lambda_0$ and $\hat{M} : \hat{\Lambda}^0 \to \hat{\Lambda}_0$.

A. The nonlinear case

The model reduction method can be straightforwardly extended from linear systems (1) to nonlinear systems of the form

$$\dot{x} = -B\mathcal{R}\left(B^T\frac{\partial H}{\partial x}(x)\right) + Eu, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, \qquad (7)$$

where $H : \mathbb{R}^n \to \mathbb{R}$ is a Hamiltonian function (total energy), and

$$\mathcal{R} : \mathbb{R}^k \to \mathbb{R}^k$$

is a nonlinear damping characteristic. Note that (1) is a special case of (7) with $H(x) = \frac{1}{2}x^TM^{-1}x$, and $\mathcal{R}$ given by the linear map R.

Given a partition P of the underlying graph $\mathcal{G}$ we reduce the system (7) to

$$\dot{\hat{x}} = -P^TB\mathcal{R}\left(B^TP\frac{\partial \hat{H}}{\partial \hat{x}}(\hat{x})\right) + P^TEu \qquad (8)$$

where $\hat{x} := P^Tx \in \mathbb{R}^{\hat{n}}$ is the clustered state vector, and where the reduced Hamiltonian function

$$\hat{H} : \mathbb{R}^{\hat{n}} \to \mathbb{R}$$

is defined as follows. Given the Hamiltonian $H : \Lambda_0 = \mathbb{R}^n \to \mathbb{R}$ define its Legendre transform $H^* : \Lambda^0 = \mathbb{R}^n \to \mathbb{R}$. (In particular, if H is *convex* then H^* is defined as its convex conjugate $H^*(x^*) := \sup_x \left[x^Tx^* - H(x)\right]$.) Then define the reduced function $\hat{H}^* : \hat{\Lambda}^0 = \mathbb{R}^{\hat{n}} \to \mathbb{R}$ as

$$\hat{H}^*(\hat{x}^*) = H^*(P\hat{x}^*)$$

Finally, define $\hat{H} : \hat{\Lambda}_0 = \mathbb{R}^{\hat{n}} \to \mathbb{R}$ as the Legendre transform of $\hat{H}^*$.

It is easily checked that for $H(x) = \frac{1}{2}x^TM^{-1}x$ the function $\hat{H}$ defined above is given as $\hat{H}(\hat{x}) = \frac{1}{2}\hat{x}^T(P^TMP)^{-1}\hat{x}$, thus generalizing the linear case considered before.

Example 3.4: Examples of nonlinear physical network systems are mass-damper systems with nonlinear dampers, nonlinear hydraulic networks (with nonlinear pressure functions), as well as (under certain conditions) general detailed-balanced mass-action kinetics chemical reaction networks; see [18].

IV. Extension to second-order systems

Consider the port-Hamiltonian system on a directed graph [17] with incidence matrix B given by

$$\begin{bmatrix}\dot{p}\\ \dot{q}\end{bmatrix} = \begin{bmatrix}-BRB^T & -B\\ B^T & 0\end{bmatrix}\begin{bmatrix}M^{-1}p\\ Kq\end{bmatrix} \qquad (9)$$

Here $p \in \mathbb{R}^n$ is a vector of state variables associated to the vertices of the graph (similar to the vector denoted before by x), and $q \in \mathbb{R}^k$ is a vector of state variables associated to the k edges of the graph. As before, the matrix M is a positive diagonal matrix and R is a positive semi-definite diagonal matrix. Similarly K is a positive semi-definite diagonal matrix. Furthermore, the total energy is given as

$$H(q,p) = \frac{1}{2}p^TM^{-1}p + \frac{1}{2}q^TKq$$

Equivalently, this can be written out as a second-order system in the vector of state variables p associated to the vertices as

$$\begin{aligned}\ddot{p} &= -BRB^TM^{-1}\dot{p} - BK\dot{q} = \\ &\quad -BRB^TM^{-1}\dot{p} - BKB^TM^{-1}p\end{aligned}$$

or equivalently

$$\ddot{p} + BRB^T M^{-1}\dot{p} + BKB^T M^{-1}p = 0 \quad (10)$$

Note that since the weighted Laplacian matrices BRB^T and BKB^T are actually independent of the orientation of the graph the second-order system (10) is in fact defined on the *undirected* graph $\mathcal{G}$.

Example 4.1: A typical example (9) and (10) is a mass-spring-damper system, where the masses are associated to the vertices, and the springs and dampers to the edges of the graph. The vector p denotes the momenta of the masses, with mass parameters $m_1, \cdots, m_n$ appearing at the diagonal of M, and q denotes the vector of elongations of the springs at the edges. Furthermore, the j-th diagonal element of R, respectively K, denote the damping constant r_j and spring constant k_j of the damper and spring associated to the j-th edge (possibly r_j or k_j are zero). The Hamiltonian $H(q,p) = \frac{1}{2}p^T M^{-1}p + \frac{1}{2}q^T Kq$ is the sum of the kinetic and potential energy.

Similarly as before we can define a reduced model for (10) as the following second-order system with clustered vector $\hat{p} := P^T p$ given as

$$\ddot{\hat{p}} + \hat{B}\hat{R}\hat{B}^T\hat{M}^{-1}\dot{\hat{p}} + \hat{B}\hat{K}\hat{B}^T\hat{M}^{-1}\hat{p} = 0, \quad (11)$$

where $\hat{K}$ is defined similarly as $\hat{R}$ (that is, $\hat{K}$ is the diagonal matrix obtained from K by leaving out the rows and columns corresponding to the edges *within* any of the cells).

A first-order state space formulation of the reduced system (11) is again given as the port-Hamiltonian system on the reduced graph

$$\begin{bmatrix}\dot{\hat{p}}\\ \dot{\hat{q}}\end{bmatrix} = \begin{bmatrix}-\hat{B}\hat{R}\hat{B}^T & -\hat{B}\\ \hat{B}^T & 0\end{bmatrix}\begin{bmatrix}\hat{M}^{-1}p\\ \hat{K}\hat{q}\end{bmatrix}, \quad (12)$$

where the elements of $\hat{q}$ corresponds to the edges which survive in the reduced model.

A *nonlinear* extension of (9) is given by

$$\begin{bmatrix}\dot{p}\\ \dot{q}\end{bmatrix} = \begin{bmatrix}-B\mathcal{R}\left(B^T\frac{\partial H_K}{\partial p}(p)\right) - B\frac{\partial H_U}{\partial q}(q)\\ B^T\frac{\partial H_K}{\partial p}(p)\end{bmatrix}, \quad (13)$$

where $H_K(p)$ is the kinetic energy, and $H_U(q)$ the potential energy part of the Hamiltonian.

For simplicity we will assume that the potential energy H_U associated to the edges is *additive*, i.e., of the form $H_U(q) = H_{U1}(q_1) + \cdots + H_{Uk}(q_k)$. Then the reduced model corresponding to a partition P is given by

$$\begin{bmatrix}\dot{\hat{p}}\\ \dot{\hat{q}}\end{bmatrix} = \begin{bmatrix}-\hat{B}\hat{\mathcal{R}}\left(\hat{B}^T\frac{\partial \hat{H}_K}{\partial \hat{p}}(\hat{p})\right) - \hat{B}\frac{\partial \hat{H}_U}{\partial \hat{q}}(\hat{q})\\ \hat{B}^T\frac{\partial \hat{H}_K}{\partial \hat{p}}(\hat{p})\end{bmatrix}, \quad (14)$$

where $\hat{p}, \hat{q}, \hat{B}$ are as defined before. Furthermore, $\hat{\mathcal{R}}$ is the nonlinear damping characteristic on the surviving edges, $\hat{H}_K(\hat{p})$ is the reduced Hamiltonian as introduced before for the first-order case, while finally $\hat{H}_U(\hat{q}) = H_{U1}(\hat{q}_1) + \cdots + H_{U\hat{k}}(\hat{q}_{\hat{k}})$, where $\hat{k}$ is the number of surviving edges (i.e., the edges *not* within cells of the partition).

V. Conclusions and Outlook

We have described a structure-preserving model order reduction approach for physical systems on graphs, which is based on clustering of the vertices. This is illustrated on a number of example classes including mass-spring-damper systems, single-species chemical reaction networks, and hydraulic networks. Furthermore, we have shown how this can be interpreted as a Petrov-Galerkin approximation.

The proposed approach naturally raises a number of questions. A main question under current consideration is the choice of the partition leading to clustering, and the resulting approximation properties of the reduced system. Recently, it has been shown in [12] how the notion of an *almost equitable partition* as defined in [3], see also [6], can be generalized to graphs with weights both on the vertices and the edges, and how for such partitions we obtain an explicit H_2 error expression between the full-order and the reduced-order model (with the outputs being defined as $y = B^T M^{-1}x$), thus generalizing the results of [11]. For the second-order case we need a partition P which is almost equitable with respect to the subgraph defined by the 'spring edges' (with weights defined by K), as well as with respect to the subgraph defined by the 'damper edges' (with weights given by R).

Another venue for research concerns the connection of model reduction by clustering as discussed in the present paper with model reduction by taking Schur complements of the weighted Laplacian of the graph; see e.g. [18], [5] and the references quoted therein.

ACKNOWLEDGMENTS

The writing of this paper was stimulated by discussions with various people, including Kanat Camlibel, Jacquelien Scherpen, Claudio De Persis, Harry Trentelman, Jorick Klaassens and Nima Monshizadeh; all of the University of Groningen.

REFERENCES

[1] M. Arcak, "Passivity as a design tool for group coordination", *IEEE Transactions on Automatic Control*, 52(8), pp. 1380 –1390, 2007.

[2] B. Bollobas, *Modern Graph Theory*, Graduate Texts in Mathematics 184, Springer, New York, 1998.

[3] D.M. Cardoso, C. Delorme, P. Rama, "Laplacian eigenvectors and eigenvalues and almost equitable partitions", *European Journal of Combinatorics*, 28(3), pp. 665–673, 2007.

[4] M. Dalsmo, A.J. van der Schaft, "On representations and integrability of mathematical structures in energy-conserving physical systems", *SIAM J. Control and Optimization, vol.37*, pp. 54–91, 1999.

[5] F. Doerfler, F. Bullo, "Kron reduction of graphs with applications to electrical networks", *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1), 2013, pp. 150 - 163.

[6] C. Godsil, G. Royle, *Algebraic graph theory*, Springer-Verlag, New York, 2001.

[7] J.-I. Imura, "Clustered model reduction of large-scale complex networks", in Proceedings of the 20th Int. Symp. on Mathematical Theory of Networks and Systems (MTNS), Melbourne, Australia, 2012.

[8] T. Ishizaki, K. Kashida, J.-I. Imura, K. Aihara, " Network clustering for SISO linear dynamical networks via reaction-diffusion transformation", Proc. 18th IFAC World Congress, Milan, Italy, pp. 5639–5644, 2011.

[9] J. Klaassens, Model reduction of networks, Master Thesis, University of Groningen, 2012.

[10] M. Mesbahi, M. Egerstedt, *Graph theoretic methods in multiagent networks*, Princeton University Press, 2010.

[11] N. Monshizadeh, M.K. Camlibel, H.L. Trentelman, "Projection based model reduction of multi-agent systems using graph partitions", to appear in *IEEE Transactions on Control of Network Systems*, 2014.

[12] N. Monshizadeh, A.J. van der Schaft, "Structure-preserving model reduction of physical network systems by clustering", submitted to the CDC2014, available at http://arxiv.org/abs/1403.4789.

[13] H. Sandberg, R.M. Murray, "Model reduction of interconnected linear systems", *Optimal Control Applications and Methods*, 30(3), pp. 225–245, 2009.

[14] A.J. van der Schaft, *L_2-Gain and Passivity Techniques in Nonlinear Control*, Lect. Notes in Control and Information Sciences, Vol. 218, Springer-Verlag, Berlin, 1996, p. 168, 2nd revised and enlarged edition, Springer-Verlag, London, 2000 (Springer Communications and Control Engineering series), p.xvi+249.

[15] A.J. van der Schaft, "Port-Hamiltonian Systems", Chapter 2 in *Modeling and Control of Complex Physical Systems; The Port-Hamiltonian Approach*, eds. V. Duindam, A. Macchelli, S. Stramigioli, H. Bruyninckx, Springer, 2009.

[16] A.J. van der Schaft, B.M. Maschke, "The Hamiltonian formulation of energy conserving physical systems with external ports", *Archiv für Elektronik und Übertragungstechnik*, 49: 362–371, 1995.

[17] A.J. van der Schaft, B.M. Maschke, "Port-Hamiltonian systems on graphs", *SIAM J. Control Optim.*, 51(2), 906-937, 2013.

[18] A.J. van der Schaft, S. Rao, B. Jayawardhana, "On the Mathematical Structure of Balanced Chemical Reaction Networks Governed by Mass Action Kinetics", *SIAM J. Appl. Math.*, 73(2), 953-973, 2013.