Approximate Reduction of Heterogenous Nonlinear Models with Differential Hulls

Max Tschaikowski and Mirco Tribastone

***Abstract*—We present a model reduction technique for a class of nonlinear ordinary differential equation (ODE) models of heterogeneous systems, where heterogeneity is expressed in terms of classes of state variables having the same dynamics structurally, but which are characterized by distinct parameters. To this end, we first build a system of differential inequalities that provides lower and upper bounds for each original state variable, but such that it is homogeneous in its parameters. Then, we use two methods for exact aggregation of ODEs to exploit this homogeneity, yielding a smaller model of size independent of the number of heterogeneous classes. We apply this technique to two case studies: a multiclass queuing network and a model of epidemics spread.**

***Keywords*-Model reduction, nonlinear ordinary differential equations, differential inequalities, lumpability, a-priori bounds.**

I. Introduction

Many natural and engineered systems feature different classes of entities that are structurally similar, yet behave in a class-specific manner (e.g., [1]). When mathematical models of such systems are considered, heterogeneity between classes is reflected in their dynamics being characterized by different parameters. Here we consider models defined as nonlinear ordinary differential equations (ODEs), whose size will grow with the number of classes. Our main contribution is a new protocol for approximate reduction of such models.

Combining ideas from the theories of differential inequalities [2] and lumpability of ODEs (e.g., [3], [4]), we provide a method that can give bounds on the solution of a heterogenous model by means of an ODE system of smaller size which preserves the structure but is characterized by parameters that represent the extreme values found in the original model. We start from a nonlinear autonomous ODE system, of size n, with solution living in the positive orthant. The method consists of two steps. The first step requires building a *differential hull*: this is a system of differential inequalities of size $2n$; roughly speaking, with each variable x_i in the original system it associates two variables, $\underline{x}_i$ and $\overline{x}_i$, representing a lower and an upper bound for x_i, respectively. The differential hull is carefully chosen such that it is amenable to an *exact* model order reduction that collapses variables associated with analogous states of different classes (e.g., all jobs of any class in a given queuing station) into a single variable. Therefore, the reduced system corresponds to a homogeneous, single-class model providing lower and upper bounds for every state of every distinct class in the original system.

We employ two different model-order reduction techniques. The first, *exact lumpability*, is a well known transformation of the original state space through a linear mapping [4]. The second considers a partition of the state space such that variables in the same partition block have equal solutions if they have equal initial conditions. We call this *uniform lumpability*, which we here generalize with respect to [5], where it was confined to ODE models induced by a formal language [6]. We show the complementarity of these two aggregations by analyzing two cases studies: a multiclass model of epidemics spread [7] and a multiclass queuing network [8]. Finally, we illustrate the applicability of our technique by means of a numerical assessment on these case studies.

In practice, differential hulls may be useful in parameter design, i.e., when a model is to be evaluated over a large parameter space in order to find values of parameters that provide some required dynamical behaviour. The guaranteed a-priori bounds provided by our technique allow, for instance, to plan for systems which enjoy certain reachability properties (e.g., the queue lengths never exceeding a threshold, or diseases never causing a severe outbreak).

Related work: Deriving bounds by means of differential inequalities can be traced back to the seminal work of Müller [9]. In particular, the notion of differential hull can be seen as a special case of Theorem 1 in [2]. Similar results have been established in [10] and [11]. As regarding model reduction, Antoulas offers a substantial account of approximate techniques [12]. For nonlinear systems, methods based on Proper Orthogonal Decomposition require empirical observations of the original system, unlike our approach. Girard and Pappas perform model reduction with *approximate bisimulation*, but it requires the computation of a Lyapunov-like function [13]. Our approach is structure-preserving in that the reduced model shows the dynamics of a representative class in a heterogeneous system (up to certain parameter values). This is different from recent work on approximate model reduction of (linear) dynamical systems evolving over a graph [14], [15], which are structure-preserving because the reduced model preserves topological information. The *averaging principle* presented in [1] allows one, under certain assumptions, to relate a quantity $f(p)$ of interest in a heterogenous model described by parameters $p = (p_1, \ldots, p_n)$ with the quantity $f(p_{\text{avg}})$ obtained by considering the same model where all parameters are replaced with their average $p_{\text{avg}} = \frac{1}{n}\sum_{i=1}^{n} p_i$. A bound in the form $\|f(p_{\text{avg}}) - f(p)\| \leq C \cdot \|p_{\text{avg}} - p\|^2$ is established. Since this result relies on Taylor's formula, the estimation of C requires the homogenous model to have a closed form solution. Unfortunately, this is difficult to obtain in our setting because closed form solutions for nonlinear ODE systems are scarce. Our results are closer to [16], which also studies aggregations of heterogeneous nonlinear ODE system. In contrast to the current work, however, [16] provides looser a-priori bounds based on Gronwall's inequality.

II. Differential Hull

Preliminaries: We study the autonomous ODE systems with solutions living in the positive orthant in the form

$$\dot{x}(t) = f(x(t)), \tag{1}$$

where we use Newton's *dot notation* to indicate the derivative with respect to the variable t, here always intended as time.

The authors are at IMT — Institute for Advanced Studies, Lucca, Italy. Email: {max.tschaikowski,mirco.tribastone}@imtlucca.it

Whenever explicit dependence on t is not needed, we shall use the notation $\dot{x} = f(x)$, with $x = (x_1, \ldots, x_n) \in \mathbb{R}^n_{>0}$. In all cases, $f : \mathbb{R}^n_{>0} \to \mathbb{R}^n$ is assumed to be locally Lipschitz continuous. This yields the existence of a unique solution. We denote its components by $f_1, f_2, \ldots, f_n$. For vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, $x \leq y$ is true if and only if $x_i \leq y_i$ for all $1 \leq i \leq n$. Strict inequality, $x < y$, is defined similarly.

We now introduce the notion of a differential hull, which is a vector field of order $2n$ that bounds our dynamics of interest f from above and from below.

Definition 1 (Differential Hull)**.** *A locally Lipschitz function* $(g_{\underline{1}}, \ldots, g_{\underline{n}}, g_{\overline{1}}, \ldots, g_{\overline{n}}) : \mathbb{R}^{2n}_{>0} \to \mathbb{R}^{2n}$ *is a* differential hull *of* $(f_1, \ldots, f_n) : \mathbb{R}^n_{>0} \to \mathbb{R}^n$ *when, for all* $1 \leq i \leq n$ *and* $\underline{x} \leq x \leq \overline{x}$*,*

$$\underline{x}_i = x_i \;\Rightarrow\; g_{\underline{i}}(\underline{x}, \overline{x}) \leq f_i(x)$$

and

$$x_i = \overline{x}_i \;\Rightarrow\; f_i(x) \leq g_{\overline{i}}(\underline{x}, \overline{x})$$

The following restates a special case of Theorem 1 from [2] in convenient way for the present paper.

Theorem 1. *Let g be a differential hull of f. Then, if the solution of the ODE system $(\dot{\underline{x}}, \dot{\overline{x}}) = g(\underline{x}, \overline{x})$ subject to $0 < \underline{x}(0) \leq x(0) \leq \overline{x}(0)$ exists and is positive on $[0; T]$, where $T > 0$, then the solution of $\dot{x} = f(x)$ exists on $[0; T]$ as well and satisfies $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ for all $0 \leq t \leq T$.*

III. Exact Aggregation

As discussed, finding a differential hull of a large ODE system is not useful *per se* in general, because the ODE system doubles in size. For the class of heterogeneous ODE systems of interest here, however, it is possible to find a *homogeneous* differential hull, i.e., one whose equations satisfy certain exact symmetries. We consider two such symmetries. First, we review the notion of exact lumpability, from [4], to which we refer the reader for details.

Exact lumpability: Let $\hat{n} \leq n$ be an integer and let M be a $\hat{n} \times n$ real constant matrix with rank $\hat{n}$.

Definition 2. *If for all solutions of (1), $y(t) := Mx(t)$ obeys an ODE system*

$$\dot{y}(t) = \hat{f}(y(t)) \;, \tag{2}$$

with some function $\hat{f} : E \to \mathbb{R}^{\hat{n}}$, with $E \subseteq \mathbb{R}^{\hat{n}}_{>0}$, then the system (1) *is said to be* exactly lumpable *to* (2) *by M.*

We confine our analysis over a finite interval of time where both (1) and (2) admit a unique solution. The existence of such an interval follows from the fact that Lipschitz continuity is is preserved under linear mappings (see Theorem 2.1, [4]).

Theorem 2 (see Theorem 1.2, [4])**.** *The system* (1) *is exactly lumpable to* (2) *by the matrix M if and only if it holds that $M \circ f = \hat{f} \circ M$ where "$\circ$" denotes function composition.*

Let us remark that in the remainder we will consider linear mappings M induced by a partition of the original state space.

Uniform lumpability: We now present uniform lumpability. In contrast to exact lumpability, it considers symmetries due to different state variables having the same trajectories when starting with the same initial conditions. This is a generalization of the reduction technique presented in [5] for ODE systems induced by formal languages. It is here extended to arbitrary vector fields, which requires a more general definition and a different proof (for Theorem 3).

Definition 3. *Let $\mathcal{H}$ be a partition of $\{1, \ldots, n\}$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field. Then*

- *The vector $x \in \mathbb{R}^n$ is said to be* uniform *on $\mathcal{H}$, if $x_i = x_j$ for all $H \in \mathcal{H}$ and $i, j \in H$.*
- *$\mathcal{H}$ is called* uniformly lumpable *(UL) with respect to f, if $f(x)$ is uniform on $\mathcal{H}$ whenever x is uniform on $\mathcal{H}$.*

Example 1. *Let us fix the partition $\mathcal{H} = \{\{1, 2\}, \{3\}\}$. Then, the vector (x_1, x_2, x_3) is uniform on $\mathcal{H}$ only if $x_1 = x_2$. Let us now consider the ODE system:*

$$\dot{x}_1 = x_1 x_3 \qquad \dot{x}_2 = x_2 x_3 \qquad \dot{x}_3 = -(x_1 + x_2) x_3$$

We have that $\mathcal{H}$ is UL with respect to this vector field because $x_1 x_3 = x_2 x_3$ whenever $x_1 = x_2$.

Thus, an UL partition identifies equal derivatives for components belonging to the same partition block. The next theorem establishes a result of equivalence for the ODE solution.

Theorem 3. *Let $\mathcal{H}$ be an UL partition of $\{1, \ldots, n\}$. Then, under the assumption that the initial condition $x(0)$ is uniform on $\mathcal{H}$, the solution of the ODE system $\dot{x} = f(x)$ is such that $x(t)$ is uniform on $\mathcal{H}$, provided that t is in the domain of x.*

Proof: Let $\mathcal{H} = \{H_1, \ldots, H_m\}$ for some $1 \leq m \leq n$. By renaming the variables $x_1, \ldots, x_n$, we may assume without loss of generality that $j \in H_j$ for all $1 \leq j \leq m$. For all $1 \leq i \leq n$, set $[i] = j$ where j is uniquely determined by $i \in H_j$. Further, for any $y \in \mathbb{R}^m_{>0}$, let $y^{\mathcal{H}} \in \mathbb{R}^n_{>0}$ be given by $y^{\mathcal{H}}_i := y_{[i]}$, where $1 \leq i \leq n$. By setting $g_j(y) := f_j(y^{\mathcal{H}})$ for all $1 \leq j \leq m$ and $y \in \mathbb{R}^m_{>0}$, let y denote the unique solution of the locally Lipschitz ODE system $(\dot{y}_1, \ldots, \dot{y}_m) = (g_1(y), \ldots, g_m(y))$ subject to some $y(0) \in \mathbb{R}^m_{>0}$. Then, for all $1 \leq i \leq n$, it holds that $(\dot{y}(t))^{\mathcal{H}}_i = (\dot{y}(t))_{[i]} = g_{[i]}(y(t)) = f_{[i]}(y(t)^{\mathcal{H}}) = f_i(y(t)^{\mathcal{H}})$, where the last identity follows form the fact that $\mathcal{H}$ is a UL partition and $y(t)^{\mathcal{H}}$ is uniform on $\mathcal{H}$. Thus, $t \mapsto y(t)^{\mathcal{H}}$ is the unique solution of $\dot{x} = f(x)$ with $x(0) = y(0)^{\mathcal{H}}$ and the claim holds true. ■

Thus, given a uniform partition $\mathcal{H}$ with respect to f and an initial condition $x(0)$ that is uniform on $\mathcal{H}$, Theorem 3 ensures that the corresponding solution of $\dot{x} = f(x)$ is such that $x(t)$ is uniform on $\mathcal{H}$ for all t that are in the domain of x. This readily yields a reduced model, obtained by choosing one equation for each representative and replacing every variable in the vector field with the representative of its block. In Example 1, choosing x_1 and x_3 as representatives gives

$$\dot{x}_1 = x_1 x_3 \qquad\qquad \dot{x}_3 = -2 x_1 x_3 \tag{3}$$

Remark 1. *In the forthcoming examples we shall identify a UL partition in terms of the ODE variables. For instance, the partition $\{\{1, 2\}, \{3\}\}$ will be denoted by $\{\{x_1, x_2\}, \{x_3\}\}$.*

IV. Case Studies

In this section we apply our technique to two case studies: an ODE model of a multi-class queueing network and a multi-type version (introduced in [7]) of the well-known susceptible-infected-recovered (SIR) epidemiological model (e.g., [17]). They show the complementarity of the two types of aggregation—the former requiring uniform lumpability while the latter being also amenable to exact lumpability. Furthermore, we present numerical results to evaluate the quality of the bounds produced by the differential hulls.

For each case study, we identify one model parameter that introduces *heterogeneity* into the system (for instance, in the SIR model we associated different infection rates to each population class). The differences between these values depend on a parameter, Δ, which thus indicates the degree of heterogeneity. Each model perturbed in this way was then analyzed with increasing values of Δ, and with varying parameter configurations so as to be exercised under different dynamics, e.g., to attain different steady states.

A. Queuing Network Model

Specification: Let us consider a Markov chain model of a queueing network (QN) with n classes of users, labeled $1, 2, \ldots, n$, where each user requires service at a resource with capacity α. Let X_i and U be the total number of class-i users ($1 \leq i \leq n$) requiring service and available resources, respectively. Then, we use a capacity-sharing argument (e.g., [18]) whereby one class-i user receives service at an exponential rate equal to $\alpha X_i / \left(\sum_{j=1}^n X_j\right) \min(\sum_{j=1}^n X_j, U)$, if $X_i \geq 0$. Furthermore, after each request we assume that each user goes idle for some time, distributed with rate β_i; after the wait, the user goes into service again. Similarly, after each service the resource is made temporarily unavailable for some time (an example of service *with vacations*, e.g., [8]), distributed with rate γ, after which it is ready to serve again. Using standard arguments (see, e.g., [19]), it can be shown that in the case of many users and many resources the Markov chain converges in probability to the following ODE system:

$$\begin{aligned}
\dot{x}_i &= -\alpha \frac{x_i}{\sum_{j=1}^n x_j} \min\left(\sum_{j=1}^n x_j, u\right) + \beta_i(1 - x_i) \\
\dot{u} &= -\alpha \min\left(\sum_{j=1}^n x_j, u\right) + \gamma(1 - u)
\end{aligned}$$

Here, for all $1 \leq i \leq n$, x_i represents the fraction of class-i users in service and u is the fraction of servers that are available; hence $1 - x_i$ is the fraction of idle clients and $1 - u$ is the fraction of servers on vacation. We consider the model with initial conditions $0 < x_i(0), u(0) < 1$, from which we have that every solution satisfies $0 < x_i(t), u(t) < 1$ for all t. In particular, the solutions $x_i(t)$ are directly related to the load of the system, with $\sum_{j=1}^n x_j(t)$ representing an estimate of the total queue length at the resource at time t.

Differential hull: The following is a differential hull for the QN model.

$$\begin{aligned}
\dot{\underline{x}}_i &= -\alpha \frac{\underline{x}_i}{\sum_{j=1}^n \underline{x}_j} \min\left(\underline{x}_i + \sum_{j \neq i} \overline{x}_j, \overline{u}\right) + \underline{\beta} - \overline{\beta}\underline{x}_i \\
\dot{\overline{x}}_i &= -\alpha \frac{\overline{x}_i}{\sum_{j=1}^n \overline{x}_j} \min\left(\overline{x}_i + \sum_{j \neq i} \underline{x}_j, \underline{u}\right) + \overline{\beta} - \underline{\beta}\overline{x}_i \\
\dot{\underline{u}} &= -\alpha \min\left(\sum_{j=1}^n \overline{x}_j, \underline{u}\right) + \gamma - \gamma\underline{u} \\
\dot{\overline{u}} &= -\alpha \min\left(\sum_{j=1}^n \underline{x}_j, \overline{u}\right) + \gamma - \gamma\overline{u}
\end{aligned} \tag{4}$$

where we set $\underline{\beta} := \min_j \beta_j$ and $\overline{\beta} := \max_j \beta_j$. The initial conditions are given by $\overline{x}_i(0) := \max_j x_j(0)$, $\underline{x}_i(0) := \min_j x_j(0)$ and $\overline{u}(0) := \underline{u}(0) := u(0)$ for all $1 \leq i \leq n$. By using the extrema of the parameters values of the original model, we can now apply model reduction. Indeed, it holds that $\left\{\{\underline{x}_1, \ldots, \underline{x}_n\}, \{\overline{x}_1, \ldots, \overline{x}_n\}, \{\underline{u}\}, \{\overline{u}\}\right\}$ is a UL partition for (4). We can write one equation for each partition block, thus obtaining the reduced differential hull:

$$\begin{aligned}
\dot{\underline{x}} &= -\alpha \min\left(\underline{x} + (n-1)\overline{x}, \overline{u}\right) + \underline{\beta} - \overline{\beta}\underline{x} \\
\dot{\overline{x}} &= -\alpha \min\left(\overline{x} + (n-1)\underline{x}, \underline{u}\right) + \overline{\beta} - \underline{\beta}\overline{x} \\
\dot{\underline{u}} &= -\alpha \min\left(n\overline{x}, \underline{u}\right) + \gamma - \gamma\underline{u} \\
\dot{\overline{u}} &= -\alpha \min\left(n\underline{x}, \overline{u}\right) + \gamma - \gamma\overline{u}
\end{aligned}$$

Therefore, we get that the solution satisfies $\underline{x}(t) \leq x_i(t) \leq \overline{x}(t)$ and $\underline{u}(t) \leq u \leq \overline{u}(t)$ for all i and all t. Let us remark that (4) cannot be reduced by exact lumpabililty.

Numerical evaluation: From the construction of (4), it emerges that the dynamical behavior of the differential hull depends on: (i) the number of heterogeneous components, n; (ii) on the actual values of β_i; and (iii) on the initial conditions $x_i(0)$, for all $1 \leq i \leq n$. The last point, in particular, is because the uniform lumpability criterion requires setting the initial conditions equal within each partition block. Thus, clearly, the best estimations are achieved when $\max_j x_j(0) = \min_j x_j(0)$.

We present numerical results where these values are changed in a controlled way by means of the parameter Δ. Specifically, for a given Δ we set $\beta_1, \ldots, \beta_n$ uniformly spaced in $[1.00 - \Delta; 1.00 + \Delta]$, and initial conditions $x_1(0), \ldots, x_n(0)$, uniformly spaced between in $[0.50 - 0.10\Delta; 0.50 + 0.10\Delta]$. We set $u(0) = 0.50$. These conditions were chosen arbitrarily, however we experimented with different values of α in order to exercise a wide spectrum of the system's dynamics. Intuitively, small values of α yield large service times, increasing the proportion of jobs (of every class) in queue; conversely, larger values lead to a network with less load on the resource.

As a measure of the accuracy of the bound, we fixed a maximum tolerance, W, for the differential hull, normalized with respect to the maximum value attained by the solution of the original system across the whole time horizon. This was set to 50 time units, which ensured convergence to equilibrium in all cases considered. Specifically, call T^* the smallest t such that $\max_{z \in \{x_1, \ldots, x_n, u\}} \{(\overline{z}(t) - \underline{z}(t)) / \max_{\tau \leq 50.0} z(\tau)\} \geq W$.

Then $[0; T^*]$ defines the time interval within which the bound on every solution, normalized with respect to the maximum value of the solution, is less than or equal to W. Let us remark that we use this normalization in order to relate the width of the bound to the range of values attained by the solution. This is analogous to error metrics widely used for approximation techniques in queuing networks, e.g., [20]. Other notions of relative error are less robust. For instance, normalizing with respect to $z(t)$ would amplify small differences for small values, although this would have a negligible impact on the actual accuracy. Having fixed T^* we similarly define the *tolerance error* as $\epsilon = \max_{z\in\{x_1,\ldots,x_n,u\}}\{(z(T^*) - z(\infty))/\max_{\tau\leq 50.0} z(\tau)\}$. This estimates how far away the solution at time T^* is from the original steady state solution.[1]

Table I shows the tolerance error ϵ with varying α and Δ, for $n = 30$. The second column lists the difference between the maximum and minimum value of the steady-state solution $x_i(\infty)$ across all values of Δ, as an indication of the different steady-state conditions attained as a function of α (the values of α were chosen to obtain queue-length proportions at every decimal point in $[0.00; 1.00]$). In general, the results indicate that the quality of the differential hull worsens with increasing Δ; this is expected because more heterogeneity must be accounted for within the bounds. The tolerance also depends on α, with a tendency to improved accuracy in the central region of the parameter space.

In order to show the dependency of the error on the system's size, we repeated the same analysis for $n = 60$. However, for these tests we used different values of α than in Table I. This is motivated by the fact that, in this queuing model, the larger n the higher the load on the resource, all other parameters being fixed. The new values of α were chosen to produce models that are comparable, in terms of steady-state queue-length proportions, with those of Table I. The results are reported in Table II, showing a similar pattern of the error behavior. To numerically confirm this trend, we collected the error statistics—minimum, average, maximum values and percentiles—for both $n = 30$ and $n = 60$ with varying tolerance levels W. There results are reported in Table III, quantifying the trade-off between the degree of precision tolerated and the "coverage" of the transient dynamics of the original system. In all cases, doubling the original system size while maintaining the same reduced order model did not cause significant degradation.

B. Multiclass SIR Model

Specification: The well-known SIR model describes the dynamics of spread of an infection from a population of infected individuals (I) to a population of susceptible individuals (S), accounting for the possibility of recovering (R) from infection after some time [21]. In its deterministic version the model is described by the following three ODEs:

$$\dot{S} = -\beta SI \qquad \dot{I} = -\gamma I + \beta SI \qquad \dot{R} = \gamma I$$

[1]This is well posed for the tests presented here. We verified that the numerical ODE solutions monotically attained convergence to a steady state.

W	n	*Min*	*5%*	*50%*	*Mean*	*95%*	*Max*
0.15	30	0.000	0.000	0.164	0.207	0.670	0.776
0.15	60	0.000	0.000	0.172	0.222	0.766	0.871
0.17	30	0.000	0.000	0.136	0.189	0.621	0.733
0.17	60	0.000	0.000	0.146	0.205	0.725	0.847
0.20	30	0.000	0.000	0.125	0.165	0.565	0.689
0.20	60	0.000	0.000	0.118	0.181	0.696	0.815
0.22	30	0.000	0.000	0.108	0.151	0.530	0.673
0.22	60	0.000	0.000	0.098	0.166	0.659	0.791

TABLE III
QN MODEL. ERROR STATISTICS FOR DIFFERENT VALUES OF W AND n.

where $\beta > 0$ is the infection rate and $\gamma > 0$ is the recovery rate. Here we consider a multiclass variant with n class of individuals with class-specific infection rates and recovery rates [7]. The ODEs for this model are, for all $1 \leq i \leq n$:

$$\dot{S}_i = -S_i \sum_{j=1}^{n} \beta_{i,j} I_j \quad \dot{I}_i = -\gamma_i I_i + S_i \sum_{j=1}^{n} \beta_{i,j} I_j \quad \dot{R}_i = \gamma_i I_i$$

Differential hull: Together with $\overline{\beta} := \max_{i,j} \beta_{i,j}$, $\underline{\beta} := \min_{i,j} \beta_{i,j}$, $\overline{\gamma} := \max_i \gamma_i$ and $\underline{\gamma} := \min_i \gamma_i$, the following is a differential hull for the multi-class SIR model.

$$\dot{\underline{S}}_i = -\overline{\beta}\underline{S}_i\Big(\sum_{j=1}^{n} \overline{I}_j\Big) \qquad \dot{\overline{S}}_i = -\underline{\beta}\overline{S}_i\Big(\sum_{j=1}^{n} \underline{I}_j\Big)$$

$$\dot{\underline{I}}_i = -\overline{\gamma}\underline{I}_i + \underline{\beta}\underline{S}_i\Big(\sum_{j=1}^{n} \underline{I}_j\Big) \quad \dot{\overline{I}}_i = -\underline{\gamma}\overline{I}_i + \overline{\beta}\overline{S}_i\Big(\sum_{j=1}^{n} \overline{I}_j\Big)$$

$$\dot{\underline{R}}_i = \underline{\gamma}\underline{I}_i \qquad \dot{\overline{R}}_i = \overline{\gamma}\overline{I}_i$$

Similarly to the QN example,

$$\{\{\underline{S}_1, \ldots \underline{S}_n\}, \{\overline{S}_1, \ldots \overline{S}_n\}, \{\underline{I}_1, \ldots \underline{I}_n\}, \{\overline{I}_1, \ldots \overline{I}_n\}, \{\underline{R}_1, \ldots, \underline{R}_n\}, \{\overline{R}_1, \ldots, \overline{R}_n\}\}$$

is a UL partition, by setting initial conditions $\overline{S}_i(0) = \max_j S_j(0)$, $\overline{I}_i(0) = \max_j I_j(0)$, $\overline{R}_i(0) = \max_j R_j(0)$, $\underline{S}_i(0) = \min_j S_j(0)$, $\underline{I}_i(0) = \min_j I_j(0)$, $\underline{R}_i(0) = \min_j R_j(0)$. This will yield an aggregate differential hull with a tolerance error that also depends on the relative differences in the initial conditions.

However, this model can also be reduced by exact lumpability, by means of an aggregation matrix M which yields the above partition. This leads to ODEs for the upper and lower bounds of the *overall sums* of susceptible, infected, and recovered individuals, denoted by $\mathcal{S}$, $\mathcal{I}$, and $\mathcal{R}$, respectively:

$$\dot{\underline{\mathcal{S}}} = -\overline{\beta}\underline{\mathcal{S}}\overline{\mathcal{I}} \qquad \dot{\overline{\mathcal{S}}} = -\underline{\beta}\overline{\mathcal{S}}\underline{\mathcal{I}}$$

$$\dot{\underline{\mathcal{I}}} = -\overline{\gamma}\underline{\mathcal{I}} + \underline{\beta}\underline{\mathcal{S}}\underline{\mathcal{I}} \qquad \dot{\overline{\mathcal{I}}} = -\underline{\gamma}\overline{\mathcal{I}} + \overline{\beta}\overline{\mathcal{S}}\overline{\mathcal{I}}$$

$$\dot{\underline{\mathcal{R}}} = \underline{\gamma}\underline{\mathcal{I}} \qquad \dot{\overline{\mathcal{R}}} = \overline{\gamma}\overline{\mathcal{I}}$$

This offers a trade off: while with exact lumpability the evolution of single class cannot be recovered, the initial conditions do not have to be perturbed as one can set $\mathcal{S}(0) = \sum_i S_i(0)$, $\mathcal{I}(0) = \sum_i I_i(0)$, and $\mathcal{R}(0) = \sum_i R_i(0)$, instead of working with their extremal values.

		Degree of heterogeneity Δ														
α	x_1–x_{30}	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090	0.100	0.110	0.120	0.130	0.140	0.150
1.50	0.94–0.95	0.000	0.000	0.000	0.029	0.106	0.165	0.211	0.247	0.276	0.301	0.320	0.336	0.353	0.364	0.375
4.50	0.83–0.86	0.000	0.000	0.000	0.046	0.115	0.167	0.205	0.235	0.259	0.277	0.293	0.305	0.317	0.326	0.335
7.50	0.73–0.77	0.000	0.000	0.000	0.059	0.115	0.156	0.185	0.208	0.224	0.239	0.250	0.258	0.265	0.271	0.275
10.50	0.63–0.68	0.000	0.000	0.000	0.062	0.102	0.129	0.148	0.162	0.171	0.179	0.184	0.189	0.191	0.192	0.193
13.50	0.53–0.60	0.000	0.000	0.012	0.048	0.069	0.081	0.088	0.092	0.093	0.094	0.092	0.090	0.086	0.081	0.077
16.50	0.44–0.52	0.000	0.000	0.004	0.007	0.004	0.010	0.020	0.030	0.042	0.055	0.068	0.080	0.094	0.108	0.124
19.50	0.36–0.45	0.000	0.000	0.020	0.056	0.082	0.105	0.125	0.145	0.163	0.182	0.200	0.216	0.234	0.252	0.272
24.00	0.24–0.35	0.000	0.000	0.038	0.134	0.192	0.236	0.274	0.306	0.334	0.361	0.385	0.408	0.431	0.455	0.480
30.00	0.10–0.22	0.000	0.000	0.000	0.217	0.326	0.399	0.455	0.501	0.539	0.576	0.613	0.642	0.672	0.703	0.729
60.00	0.01–0.02	0.069	0.077	0.120	0.128	0.131	0.267	0.409	0.504	0.572	0.620	0.670	0.709	0.735	0.763	0.776

TABLE I
QN MODEL. NUMERICAL EVALUATION USING A TOLERANCE $W = 0.15$ AND $n = 30$.

		Degree of heterogeneity Δ														
α	x_1–x_{60}	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090	0.100	0.110	0.120	0.130	0.140	0.150
3	0.94–0.95	0.000	0.000	0.000	0.029	0.106	0.165	0.211	0.247	0.276	0.301	0.320	0.337	0.354	0.365	0.376
9	0.84–0.86	0.000	0.000	0.000	0.045	0.115	0.167	0.205	0.235	0.259	0.277	0.294	0.308	0.320	0.330	0.339
15	0.74–0.78	0.000	0.000	0.000	0.057	0.117	0.159	0.190	0.213	0.232	0.247	0.259	0.268	0.275	0.282	0.287
21	0.66–0.71	0.000	0.000	0.000	0.063	0.109	0.141	0.163	0.179	0.192	0.202	0.208	0.214	0.218	0.222	0.224
30	0.55–0.61	0.000	0.000	0.011	0.052	0.076	0.090	0.099	0.105	0.108	0.109	0.109	0.108	0.105	0.102	0.097
39	0.45–0.53	0.000	0.000	0.006	0.011	0.010	0.007	0.012	0.022	0.034	0.046	0.058	0.071	0.084	0.098	0.113
48	0.36–0.45	0.000	0.000	0.019	0.052	0.077	0.099	0.119	0.138	0.156	0.174	0.192	0.208	0.226	0.244	0.263
60	0.26–0.37	0.000	0.000	0.036	0.118	0.170	0.208	0.243	0.272	0.298	0.323	0.346	0.368	0.390	0.413	0.437
90	0.07–0.20	0.000	0.000	0.000	0.226	0.343	0.419	0.483	0.531	0.572	0.610	0.649	0.679	0.711	0.743	0.770
120	0.01–0.08	0.000	0.000	0.000	0.279	0.456	0.572	0.657	0.719	0.766	0.792	0.813	0.834	0.849	0.864	0.871

TABLE II
QN MODEL. NUMERICAL EVALUATION USING A TOLERANCE $W = 0.15$, $n = 60$.

		Degree of heterogeneity Δ					
γ_i	S_1–S_{30}	$\leq$0.020	0.040	0.060	0.080	0.100	0.120
10	0.03–0.03	0.000	0.000	0.130	0.841	0.927	0.962
15	0.08–0.08	0.000	0.000	0.017	0.143	0.260	0.745
20	0.14–0.14	0.000	0.000	0.000	0.038	0.114	0.210
25	0.19–0.19	0.000	0.000	0.000	0.000	0.038	0.099
30	0.23–0.23	0.000	0.000	0.000	0.000	0.000	0.027
40	0.29–0.30	0.000	0.000	0.000	0.000	0.000	0.000
50	0.34–0.34	0.000	0.000	0.000	0.000	0.000	0.000
60	0.36–0.36	0.000	0.000	0.000	0.000	0.000	0.000
100	0.42–0.42	0.000	0.000	0.000	0.000	0.000	0.000
200	0.46–0.46	0.000	0.000	0.000	0.000	0.000	0.000

TABLE IV
SIR MODEL. NUMERICAL EVALUATION USING A TOLERANCE $W = 0.15$.

W	*Min*	*5%*	*50%*	*Mean*	*95%*	*Max*
0.15	0.000	0.000	0.000	0.076	0.793	0.962
0.17	0.000	0.000	0.000	0.064	0.689	0.962
0.20	0.000	0.000	0.000	0.038	0.174	0.886
0.22	0.000	0.000	0.000	0.024	0.122	0.841

TABLE V
SIR MODEL. ERROR STATISTICS FOR DIFFERENT ERROR TOLERANCES W.

Numerical evaluation: The numerical evaluation was designed in a way similar to the QN case study, but here we consider the results for the differential hull aggregated according to exact lumpability. The source of heterogeneity was assigned to the infection parameters $\beta_{i,j}$. For a given degree Δ, $\{\beta_{1,1}, \ldots, \beta_{1,n}, \beta_{2,1}, \ldots, \beta_{2,n}, \ldots, \beta_{n,n}\}$ were set uniformly spaced in $[1.00 - \Delta; 1.00 + \Delta]$. For all $1 \leq i \leq n$, the initial conditions $S_i(0)$, were uniformly spaced between in $[0.50 - 0.1\Delta; 0.50 + 0.1\Delta]$ while we set $I_i(0) = 1 - S_i(0)$ and $R_i(0) = 0$.[2]

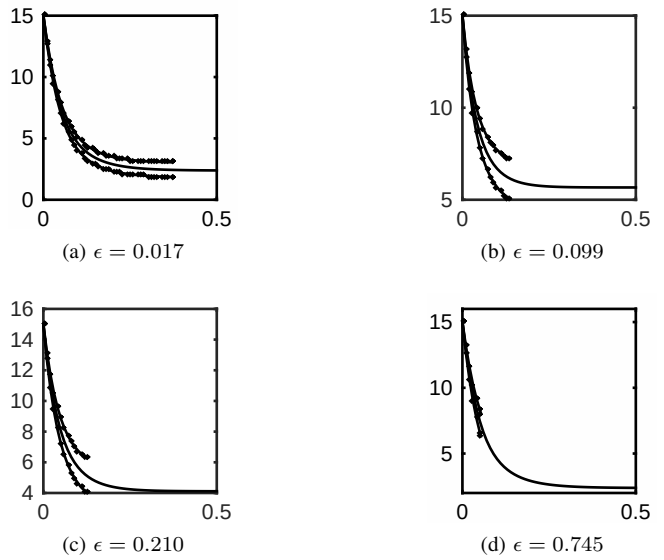
The notion of error is defined similarly to the QN case, taking into account that the aggregated hull now provides bounds for sums of the original trajectories of the original ODE system. More precisely, define T^* as the smallest t such that

$$\max\Big\{(\bar{S}(t) - \underline{S}(t))/\max_{\tau\leq 50.0}\sum_i S_i(\tau),\\ (\bar{I}(t) - \underline{I}(t))/\max_{\tau\leq 50.0}\sum_i I_i(\tau),\\ (\bar{R}(t) - \underline{R}(t))/\max_{\tau\leq 50.0}\sum_i R_i(\tau)\Big\} \geq W,$$

where the bound is normalized with respect to the maximum sums of the respective classes across all time points (here, again, the ODEs were numerically solved until time point 50.0). Then, the *tolerance error* is defined as $\epsilon = \max_{z\in\{S_1,\ldots,S_n,I_1,\ldots,I_n,R_1,\ldots,R_n\}}\{(z(T^*) - z(\infty))/\max_{\tau\leq 50.0} z(\tau)\}$.

The results for $n = 30$ are presented in Table IV, where we varied γ_i, for all $1 \leq i \leq n$, between 10 and 200; the second column indicates the range of susceptible individuals of each class in the steady state, for each value of γ across of all values of Δ. Notice that this covers a wide range behaviors, since the population of susceptible individuals cannot be larger than the largest initial condition, 0.515, across all experiments.

[2]The theorems hold also if f and g are locally Lipschitz at any point in $\mathbb{R}^n_{\geq 0}$ and the differential hull $(\underline{x}, \overline{x})$ is known to be nonnegative on $[0; T]$.



(a) $\epsilon = 0.017$ (b) $\epsilon = 0.099$ (c) $\epsilon = 0.210$ (d) $\epsilon = 0.745$

Fig. 1. Comparison of differential hull (marked line) against the original SIR model (solid line) for the sum of susceptibles in four selected cases. The solution of the differential hull is plotted until its normalized tolerance is less than $W = 0.15$. Parameters: (a) $\Delta = 0.060$, $\gamma_i = 15$; (b) $\Delta = 0.120$, $\gamma_i = 25$; (c) $\Delta = 0.120$, $\gamma_i = 20$; (d) $\Delta = 0.120$, $\gamma_i = 15$, $1 \leq i \leq n$.

The trend is different than in the QN case study. Here, the differential hull performs better for smaller values of Δ. Substantial errors occur for small γ_i and large Δ (0.140 and 0.150) which correspond to a degree of heterogeneity of ca 15% across the values of $\beta_{i,j}$. We attribute this difference to the fact that, for small Δ, a nonnegligible contribution to the error behavior is given by width of the differential hull trajectory at the initial condition. While this is equal 0.2Δ in the QN case, it is equal to zero in the SIR case, owing to the use of an exactly lumpable differential hull. Finally, Table V shows the error statistics for varying W. The behavior with varying n is similar to the QN case, hence we do not report it here; notice, however, that unlike the differential hull for QN, the one for the SIR model does not depend on n explicitly.

Instead, for a visual appreciation of the bounds, Figure 1 compares the solution of the differential hulls against the original model in selected cases characterized by increasing tolerance errors with $W = 0.15$.

Summary: Although the accuracy is model- and parameterization-dependent, taken together these result suggest good applicability of aggregated differential hulls for models with a degree of heterogeneity within a few percent.

V. Conclusion

I n this technical note we have considered a class of nonlinear ordinary differential equation models for heterogeneous systems, where heterogeneity is due to the fact that different classes have the same structure, but they are characterized by different parameters. We have proposed a method to perform model reduction by collapsing variables related to distinct classes into the same one. The reduced model requires no information about the trajectories of the original system, only the availability of its vector field. Its solution provides upper and lower bounds for all original variables, which have shown satisfactory tightness in our numerical validation with case studies of queuing and epidemiological networks. A relevant subject of future work is to investigate algorithms that detect these *near-symmetries* between state variables and produce lumpable differential hulls automatically.

Acknowledgement

This work is partially supported by the EU project QUANTICOL, 600708, and by the DFG SPP-1593 project DAPS.

References

[1] G. Fibich, A. Gavious, and E. Solan, "Averaging principle for second-order approximation of heterogeneous models with homogeneous models," *PNAS*, vol. 109, no. 48, pp. 19 545–19 550, 2012.

[2] N. Ramdani, N. Meslem, and Y. Candau, "A hybrid bounding method for computing an over-approximation for the reachable set of uncertain nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 54, no. 10, pp. 2352–2364, 2009.

[3] M. S. Okino and M. L. Mavrovouniotis, "Simplification of mathematical models of chemical reaction systems," *Chemical Reviews*, vol. 2, no. 98, pp. 391–408, 1998.

[4] J. Toth, G. Li, H. Rabitz, and A. S. Tomlin, "The effect of lumping and expanding on kinetic differential equations," *SIAM Journal on Applied Mathematics*, vol. 57, no. 6, pp. 1531–1556, 1997.

[5] M. Tschaikowski and M. Tribastone, "Exact fluid lumpability for Markovian process algebra," in *23rd International Conference on Concurrency Theory*, ser. LNCS, 2012, pp. 380–394.

[6] J. Hillston, *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.

[7] R. K. Watson, "On an epidemic in a stratified population," *Journal of Applied Probability*, vol. 9, no. 3, pp. 659–666, 1972.

[8] G. Bolch, S. Greiner, H. de Meer, and K. Trivedi, *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. Wiley, 2005.

[9] M. Müller, "Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen," *Mathematische Zeitschrift*, vol. 26, pp. 619–645, 1927. [Online]. Available: http://eudml.org/doc/167946

[10] J. K. Scott and P. I. Barton, "Bounds on the reachable sets of nonlinear control systems," *Automatica*, vol. 49, no. 1, pp. 93 – 100, 2013.

[11] A. B. Singer and P. I. Barton, "Bounding the solutions of parameter dependent nonlinear ordinary differential equations." *SIAM J. Scientific Computing*, vol. 27, no. 6, pp. 2167–2182, 2006.

[12] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*, ser. Advances in Design and Control. SIAM, 2005.

[13] A. Girard and G. Pappas, "Approximate bisimulations for nonlinear dynamical systems," in *CDC-ECC*, 2005, pp. 684–689.

[14] T. Ishizaki, K. Kashima, A. Girard, J.-i. Imura, L. Chen, K. Aihara *et al.*, "Clustering-based $\mathcal{H}_2$-state aggregation of positive networks and its application to reduction of chemical master equations," in *51st IEEE Conference on Decision and Control*, 2012, pp. 4175–4180.

[15] T. Ishizaki, K. Kashima, J.-I. Imura, and K. Aihara, "Model reduction and clusterization of large-scale bidirectional networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 1, pp. 48–63, Jan 2014.

[16] G. Iacobelli and M. Tribastone, "Lumpability of fluid models with heterogeneous agent types," in *The 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 2013, pp. 1–11.

[17] J. D. Murray, *Mathematical Biology I: An Introduction*, 3rd ed. Springer, 2002.

[18] D. Qiu and R. Srikant, "Modeling and performance analysis of BitTorrent-like peer-to-peer networks," in *Proceedings of ACM SIGCOMM 2004*, 2004, pp. 367–378.

[19] L. Bortolussi, J. Hillston, D. Latella, and M. Massink, "Continuous approximation of collective system behaviour: A tutorial," *Performance Evaluation*, vol. 70, no. 5, pp. 317–349, 2013.

[20] D. L. Eager and J. N. Lipscomb, "The AMVA priority approximation," *Perf. Eval.*, no. 8, pp. 173–193, 1988.

[21] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A.*, vol. 115, no. 772, pp. 700–721, 1927.