

ABSTRACT

Purpose: To estimate rigorously and efficiently the range of behaviours that a dynamic model of a system may take given a set of uncertain parameters values.

Results: Using a novel bounding technique based on differential inequalities, a system of differential equations with optimization problems embedded is derived. This system is solved using a new algorithm developed at the Process Systems Engineering Laboratory at MIT to yield a tight enclosure of chemical kinetics models.

Conclusions: The algorithmic development has been key to implementing the existing bounding theory. Further development of the algorithm is under way so that it can be applied to more complicated models.

INTRODUCTION

Rigorous enclosure of the range of possible behaviours of a system is critical to understanding and safely operating the system. This range of behaviours, known as the design envelope, is a critical component of quality by design (QbD) initiatives. Of course, the inputs to a model are never known exactly; one example is the uncertainty in the rate constants in a kinetic model. Often these have huge variations when estimated from experimental data. Given a possible range of parameters, also called the design space, estimating the resulting design envelope is a nontrivial task, especially when the model involves differential equations.

The work presented here is potentially quite broad in its applicability to various unit operations and an entire chemical process in general; however, the current focus will be on chemical kinetics models. Results for a closed reacting system will be presented.

CONCEPTS

Model of interest:

The evolution of the species concentrations in a closed reacting system with uncertain parameters can be described by the initial value problem in n parametric ordinary differential equations (ODEs) of the form

$$\dot{\mathbf{x}}(t, \mathbf{p}) = \mathbf{S}\mathbf{r}(t, \mathbf{x}(t, \mathbf{p}), \mathbf{p}), \quad \mathbf{x}(t_0, \mathbf{p}) = \mathbf{x}_0(\mathbf{p}), \quad \forall (t, \mathbf{p}) \in [t_0, t_f] \times P,$$

where \mathbf{S} is the stoichiometry matrix, \mathbf{r} is the vector of rate functions, $[t_0, t_f]$ is the time interval of interest, and P is the bounded interval describing the uncertainty in the parameters. The stoichiometry of the reactions in the kinetic network implies constraints on the species concentrations. These constraints take the form of affine subspaces; the solution $\mathbf{x}(t, \mathbf{p})$ must lie in an affine subspace for all t and \mathbf{p} . More constraints can be deduced from the initial conditions in the form of upper and lower bounds for each species concentration called natural bounds. The intersection of all these constraints is a convex polyhedral set, G , called an invariant.

CONCEPTS (continued)

Bounding equations:

Using arguments based on differential inequalities, bounds are given by the solution of $2n$ ordinary differential equations with optimization problems embedded:

$$\dot{x}_i^L(t) = \min_{(\mathbf{p}, \mathbf{z})} \mathbf{s}_i \mathbf{r}(t, \mathbf{p}, \mathbf{z}), \quad x_i^L(t_0) = \min_{\mathbf{p} \in P} x_{0,i}(\mathbf{p}),$$

$$\text{s.t. } \mathbf{z} \in [\mathbf{x}^L(t), \mathbf{x}^U(t)] \cap G, \quad \mathbf{p} \in P,$$

$$z_i = \min_{\Psi} \psi_i$$

$$\text{s.t. } \Psi \in [\mathbf{x}^L(t), \mathbf{x}^U(t)] \cap G$$

$$\dot{x}_i^U(t) = \max_{(\mathbf{p}, \mathbf{z})} \mathbf{s}_i \mathbf{r}(t, \mathbf{p}, \mathbf{z}),$$

$$x_i^U(t_0) = \max_{\mathbf{p} \in P} x_{0,i}(\mathbf{p}),$$

$$\text{s.t. } \mathbf{z} \in [\mathbf{x}^L(t), \mathbf{x}^U(t)] \cap G, \quad \mathbf{p} \in P,$$

$$z_i = \max_{\Psi} \psi_i$$

$$\text{s.t. } \Psi \in [\mathbf{x}^L(t), \mathbf{x}^U(t)] \cap G$$

for each i , where \mathbf{s}_i is the i^{th} row of \mathbf{S} . The solution of this system gives rigorous interval bounds on the species concentrations for all parameter values in P . Precisely,

$$\mathbf{x}(t, \mathbf{p}) \in [\mathbf{x}^L(t), \mathbf{x}^U(t)], \quad \forall (t, \mathbf{p}) \in [t_0, t_f] \times P.$$

RESULTS & DISCUSSIONS

By taking affine relaxations of the rate expressions $\mathbf{s}_i \mathbf{r}$, one obtains a system of ODEs with embedded (hierarchical) linear programs (LPs). By hierarchical it is meant that the solution set of a “low” level LP influences the constraints of a “high” level LP, but not the other way around. A code for the efficient simulation of such systems has been developed.

This code relies on solving the embedded LPs once with a LP solver. Then the optimal “basis” is recorded. This basis is essentially an index set that allows the determination of the optimal solution set by solving a set of algebraic equations (rather than using the LP solver). Thus the above system of $2n$ equations can be converted into differential algebraic equations (DAEs). The result is a mathematical model that can be solved efficiently with numerical integrators. The recorded basis is in fact optimal for a finite period of time in the simulation; using event detection one can determine when the basis is no longer optimal and then re-solve the LPs to update the basis before continuing integration.

Example reaction network:

Consider the reaction of network of elementary reactions in a closed system



where the vector $\mathbf{p} = (p_1, p_2)$ is the vector of rate constants. Bounds for the species concentrations will be calculated when this p_1 ranges over $[150, 600] \text{ M}^{-1}\text{s}^{-1}$ and p_2 ranges over $[10, 300] \text{ M}^{-1}\text{s}^{-1}$.

RESULTS & DISCUSSIONS (continued)

The mathematical equations which describe the evolution of the concentrations x_A, x_B, x_C are

$$\begin{aligned} \dot{x}_A(t, \mathbf{p}) &= -p_1 x_A(t, \mathbf{p}), & x_A(t_0, \mathbf{p}) &= 1.5\text{M} \\ \dot{x}_B(t, \mathbf{p}) &= p_1 x_A(t, \mathbf{p}) - p_2 x_B(t, \mathbf{p}), & x_B(t_0, \mathbf{p}) &= 0.5\text{M} \\ \dot{x}_C(t, \mathbf{p}) &= p_2 x_B(t, \mathbf{p}), & x_C(t_0, \mathbf{p}) &= 0.0\text{M}. \end{aligned}$$

By mass balance arguments, it is known that the sum of the concentrations must always equal the total concentration at the initial time. This provides the affine subspace on which the solution must lie. Similar arguments gives upper and lower bounds on the concentration of each species that are valid for all times. Together this knowledge furnishes the invariant G .

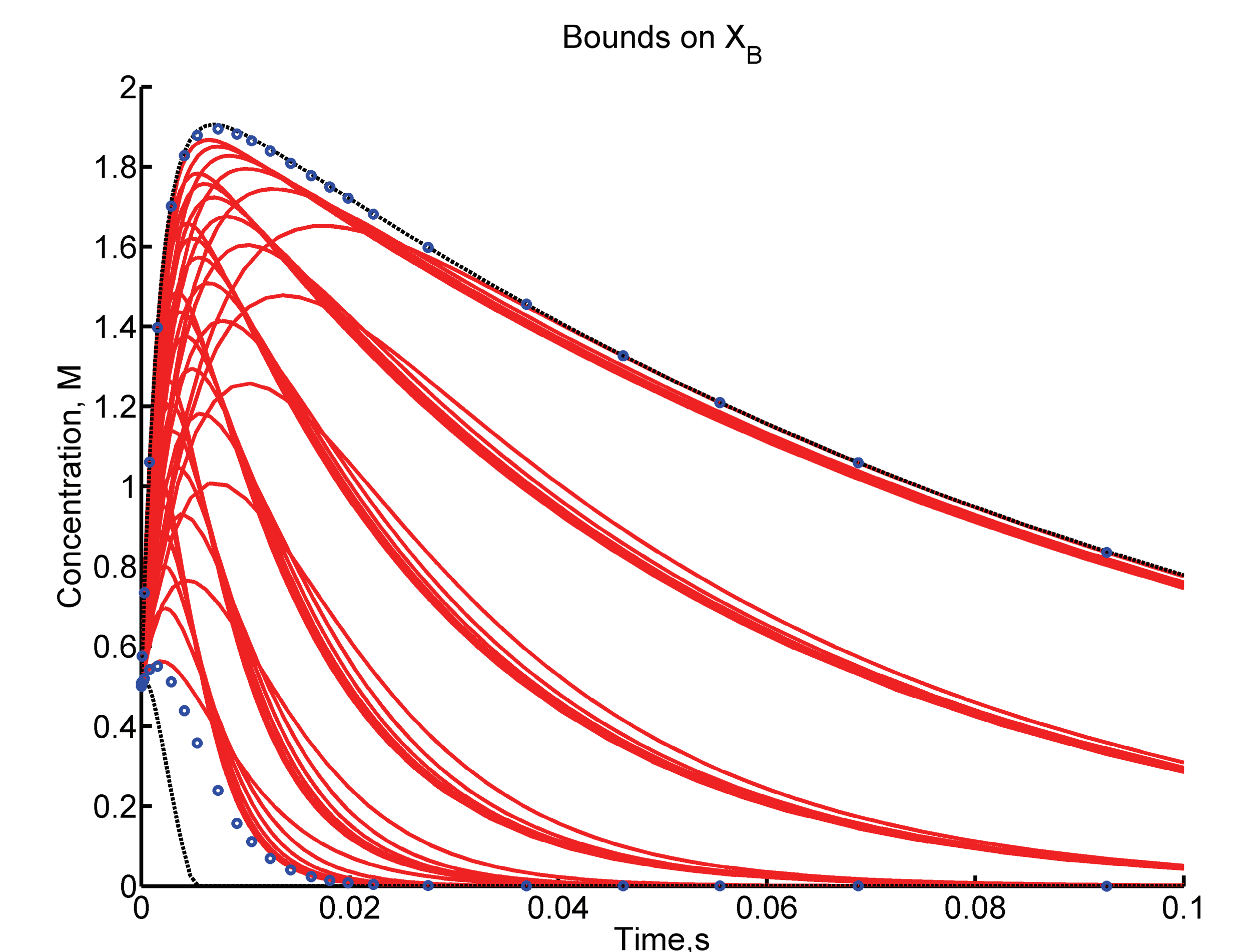


Figure 1

Figure 1 shows the bounds for the concentration of species B from the example reaction network. The solid lines are sampled trajectories, while the dashed lines represent bounds from a previous bounding method. The circles are the bounds from the current theory and algorithm.

One can see that the bounds generated by the theory and algorithm presented in this work are quite tight, even when the parameters may vary by over an order of magnitude.

CONCLUSIONS

A new bounding theory and efficient implementation have been presented and applied to an example reaction network. The bounds produced are extremely tight even when the rate parameters of the system vary by over an order of magnitude. The bounds produced are better than previous bounding procedures. This theory and algorithm are being applied to more complicated reaction systems to further investigate and improve its performance.