Range-Constrained Simultaneous Reaction and Species Elimination in Kinetic Mechanisms

Geoffrey M. Oxberry*, Alexander Mitsos†, Paul I. Barton*, William H. Green*
*Massachusetts Institute of Technology/Department of Chemical Engineering, Cambridge, MA, USA
†Massachusetts Institute of Technology/Department of Mechanical Engineering, Cambridge, MA, USA

Abstract— **Although many techniques exist to generate a reduced model from a large, detailed chemical model, few model reduction techniques generate reduced models that can reproduce the solution of the detailed model to within known error bounds. Furthermore, error bounds imposed on the reduced model at a finite set of reaction conditions cannot be propagated through a numerical method to yield error bounds on the numerical solution of the reduced model relative to the solution of the detailed model. In order to reproduce the numerical solution of the detailed model to within a known error tolerance using a reduced model, the reduced model must be generated with known error bounds that are satisfied over a range of reaction conditions.**

To generate a reduced model satisfying error bounds over a range of reaction conditions, we propose the method of range-constrained simultaneous reaction and species elimination. This method uses the solution of an integer linear program restriction of an integer linear semi-infinite program to determine a reduced model from a given detailed model, error tolerances on the time derivatives of state variables and range of reaction conditions. The reduced model obtained from the solution of the integer linear program can then be used within a numerical method to approximate faithfully the solution of the detailed model from which it was generated. Error bounds on the reduced model could then be propagated through the numerical method in order to obtain error bounds on the solution of the detailed model, yielding a more computationally efficient means of obtaining a numerical solution for currently intractable reacting flow problems to within known numerical error.

I. INTRODUCTION

The purpose of a kinetic model reduction technique is to generate an approximate, simplified chemical model from a more detailed chemical model in order to reduce the computational effort needed to simulate reacting flows. Many different methods exist to accomplish this task (such as [6], [14] and [7]), based on very different principles.

When carrying out model reduction, there are two competing objectives: reducing CPU effort and minimizing approximation error due to model reduction. Since the purpose of model reduction is to enable the simulation of computationally demanding reacting flow problems by generating less resource-intensive reduced models, we would like our model reduction technique to yield a reduced model that decreases as much as possible the CPU time required to simulate a reacting flow. Simultaneously, we would like to minimize the difference (or error) between the solution of our reduced model and the solution of the detailed model. Since CPU time limits the set of reacting flow problems that can be solved currently, our primary objective is to achieve maximum reduction of CPU effort subject to constraints that ensure that the error between the numerical solution of the detailed model and the solution of the reduced model is bounded to within acceptable limits over certain reaction conditions (temperatures and species mass fractions). In order to limit this error when using the reduced model in a numerical integration routine, it is necessary to bound the difference between the chemical source term of the detailed model and the chemical source term of the reduced model for each species. Although many methods have some form of error control (usually in the form of an adjustable parameter), few methods attempt to determine error bounds on the chemical source term of the reduced model with respect to the detailed model. Without these error bounds, it is extremely difficult to estimate the error in the solution obtained by applying a model reduction technique when solving a detailed model numerically.

It is also worth noting that bounding locally the error due to model reduction does not suffice [11]. Oluwole *et al.* have demonstrated that error bounds satisfied by a reduced model at points in state space do not necessarily hold within their convex hull. For this reason, a reduced model with error bounds satisfied at a finite collection of points in state space may not satisfy its stated error tolerances after one time step of numerical integration. Consequently, it is absolutely critical that reduced models be generated with error bounds on the time derivatives of state variables, and that these error bounds are satisfied over ranges in state space.

Optimization is a natural mathematical framework for model reduction because elements of model reduction problems can be adapted to an optimization formulation. Typically, the mode of model reduction can be cast in terms of decision variables of the optimization problem. The mode of model reduction can be thought of as the rules or allowable transformations that can take place in generating a reduced model from a detailed model. Reaction elimination [3] is one example of such rules. Error bounds can be formulated as constraints, and CPU effort can be formulated as an objective. Since simulation CPU time cannot be expressed directly as a continuous function of the decision variables, the number of reactions, species, or state variables is typically used as a proxy, because the CPU time needed to solve a reacting flow problem scales empirically as $O(N_R N_S^2)$, where N_R is the number of reactions and N_S is the number of species [13]. Adapting a model reduction technique to an optimization framework ensures that a given detailed model is maximally (or close to maximally) reduced, subject to the error constraints and model reduction rules supplied by the model reduction technique.

In this work, we propose using an optimization framework to generate reduced models by simultaneous reaction and species elimination, such that the resulting models satisfy given error constraints over ranges in state space. This approach, called range-constrained simultaneous reaction and species elimination, is a natural extension of existing work on both reaction elimination [3], [10] and simultaneous reaction and species elimination [8]. In order to provide the necessary background for the development of range-constrained simultaneous reaction and species elimination, we first review the previous point-constrained reaction and species elimination by Mitsos *et al.* [8].

II. EXISTING POINT-CONSTRAINED SIMULTANEOUS REACTION AND SPECIES ELIMINATION FORMULATION

Bhattacharjee [1] proposed an integer nonlinear programming (INLP) formulation for simultaneous reaction and species elimination. Although this formulation was novel, it was also computationally demanding, since algorithms that solve INLPs are time-consuming. Mitsos *et al.* [8] discovered an equivalent integer linear program (ILP) reformulation of the original INLP that can be solved much more quickly. The resulting point-constrained reaction and species elimination ILP is presented below as in [8], restricting the formulation to the case where $\alpha_j = 1$ for all j and $\beta_i = 0$ for all i:

$$\min_{\mathbf{z},\mathbf{w}} \sum_{j=1}^{N_S} w_j, \tag{1a}$$

$$\text{s.t.} \left| \frac{\sum_{j=1}^{N_S} h_j(T_l) M_j \sum_{i=1}^{N_R} \nu_{ji} z_i r_i(\mathbf{x}_l, T_l)}{\rho_l (C_P)_l} - \Gamma_0^{ref}(\mathbf{x}_l, T_l) \right|$$
$$\leq (atol)_0 + (rtol)_0 |\Gamma_0^{ref}(\mathbf{x}_l, T_l)|, \quad l = 1, \ldots, N_t, \tag{1b}$$

$$\left| \frac{M_j \sum_{i=1}^{N_R} \nu_{ji} z_i r_i(\mathbf{x}_l, T_l)}{\rho_l} - \Gamma_j^{ref}(\mathbf{x}_l, T_l) \right|$$
$$\leq (atol)_j + (rtol)_j |\Gamma_j^{ref}(\mathbf{x}_l, T_l)|, \quad j = 1, \ldots, N_S; l = 1, \ldots, N_t, \tag{1c}$$

$$(N_R)_j w_j \geq \sum_{\{i \colon \nu_{ji} \neq 0\}} z_i, \quad j = 1, \ldots, N_S, \tag{1d}$$

$$z_i \in \{0, 1\}, \quad i = 1, \ldots, N_R, \tag{1e}$$

$$w_j \in \{0, 1\}, \quad j = 1, \ldots, N_S. \tag{1f}$$

The formulation in (1) requires some explanation. In (1), three subscripts are used. The subscript i indexes the N_R reactions present in the detailed mechanism, and is used in conjunction with reaction-based quantities, such as the rate of reaction. The subscript j indexes the N_S species present in the detailed mechanism, and is used in conjunction with species-based quantities, such as the species mass fractions. The subscript l indexes N_t user-supplied points in state space to be used as a reference data set for model reduction. This reference data set defines the conditions in state space over which the reduced model will be "valid," in that it satisfies error tolerances on the time derivatives of the state variables relative to the detailed model. In our case, state space is defined as the mass fractions of each species present in the detailed model and temperature; we assume that our detailed model describes the chemical source term in an adiabatic-isobaric batch reactor.

Reaction elimination is encoded by the binary decision variables z_i, and species elimination is encoded by the binary decision variables w_j. If $z_i = 1$, reaction i is included in the reduced model generated by this technique. If $z_i = 0$, reaction i is excluded from the reduced model generated by this technique. Similarly, if $w_j = 1$, species j is included in the reduced model generated by this technique. If $w_j = 0$, species j is excluded from the reduced model generated by this technique. Consequently, an optimal solution to (1) yields a reduced mechanism derived from the detailed mechanism supplied as input to point-constrained simultaneous reaction and species elimination. Given this interpretation of the binary decision variables, the objective function (1a) equals the number of species included in the reduced model.

Assuming that our detailed model describes the chemical source term in an adiabatic-isobaric batch reactor, the error constraints in (1b) and (1c) limit the difference between the time derivatives of the state variables in the detailed model and the time derivatives of the corresponding state variables in the reduced model. In these equations, h_j is the specific enthalpy of species j, M_j is the molar mass of species j, ν_{ji} is the stoichiometric coefficient of species j in reaction i (using the standard sign convention that $\nu_{ji} > 0$ if species j is produced in reaction i, and $\nu_{ji} < 0$ if species j is consumed in reaction i), r_i is the molar rate of reaction i, $\mathbf{x}_l$ is the vector of species mass fractions in the batch reactor at reference point l, T_l is the temperature in the batch reactor at reference point l, ρ_l is the density of the gas in the batch reactor at reference point l, $(C_P)_l$ is the specific heat capacity of the mixture in the batch reactor at reference point l, Γ_0^{ref} is the chemical source term for the temperature evaluated at a reference point specified as an argument, and Γ_j^{ref} is the chemical source term for species j evaluated at a reference point specified as an argument. The absolute error tolerance for the time derivative of species j is defined as $(atol)_j$, and the absolute error tolerance for the time derivative of temperature is defined as $(atol)_0$; these tolerances are set by the user. Corresponding relative tolerances are defined as $(rtol)_j$ for species j and $(rtol)_0$ for temperature.

Mitsos *et al.* [8] recognized that in order to avoid the production or destruction of mass via the reactions of the reduced mechanism, a species may only be eliminated from the detailed mechanism if all of the reactions in which it participates (in the detailed mechanism) are also eliminated. The mass conservation constraint (1d) encodes this condition, where $(N_R)_j$ is the number of reactions of the detailed mechanism in which species j participates. In the case of unimolecular, bimolecular, and explicit termolecular reactions, participation of a chemical species is unambiguous. However, some reaction mechanisms contain third body reactions, in which a molecule called a third body acts upon the reactants to give them enough kinetic energy for reaction to occur. If a third body reaction is treated as a bimolecular reaction, in that the third body species are not considered to participate in that reaction, then the estimates of the time derivatives of the state variables in (1d) may be inaccurate. For most practical purposes, these inaccuracies are insignificant and can be ignored. If the inaccuracies due to neglecting the participation of third body species are significant, Mitsos *et al.* [8] propose alternate treatments for third body species.

Point-constrained species elimination has been applied to mechanisms as large as the LLNL n-heptane mechanism [4] successfully. Due to the $O(N_R N_S^2)$ scaling of reacting flow solvers, simultaneous elimination of reactions and species reduces the computational effort of reacting flow solvers to a greater extent than elimination of reactions only. However, since the chemical source term is, in general, a non-convex function of the state variables over the convex hull of the reference points supplied to (1), the error in the reduced model is also a non-convex function over the convex hull of the reference points. Consequently, if a reduced model is generated satisfying error bounds at the reference points, it is not necessarily true that the reduced model also satisfies the same error bounds at any point in state space within the convex hull of the reference points. For this reason, if error-controlled reduced models are desired over regions in state space (for example, in a reacting flow solver using an adaptive chemistry algorithm [12]), either a valid range must be determined from a point-constrained reduced model by using the range-finding algorithm of Oluwole *et al.* [11], or the constraints in the point-constrained model reduction formulation (1) must be revised to limit errors over ranges in state space.

III. PROPOSED RANGE-CONSTRAINED SPECIES ELIMINATION FORMULATION

By analogy to the previous work on reaction elimination by Bhattacharjee *et al.* [1], [3] and Oluwole *et al.* [10], we propose the following range-constrained simultaneous reaction and species elimination formulation as an extension of a combination of the work of [8] and [10]:

$$\min_{\mathbf{z},\mathbf{w}} \sum_{j=1}^{N_S} w_j, \tag{2a}$$

$$\text{s.t.} \left| \frac{\sum_{j=1}^{N_S} h_j(T) M_j \sum_{i=1}^{N_R} \nu_{ji} z_i r_i(\mathbf{x}, T)}{\rho(\mathbf{x}, T) C_P(\mathbf{x}, T)} - \Gamma_0^{ref}(\mathbf{x}, T) \right|$$
$$\leq (atol)_0 + (rtol)_0 |\Gamma_0^{ref}(\mathbf{x}, T)|, \quad \forall (\mathbf{x}, T) \in \Phi, \tag{2b}$$

$$\left| \frac{M_j \sum_{i=1}^{N_R} \nu_{ji} z_i r_i(\mathbf{x}, T)}{\rho(\mathbf{x}, T)} - \Gamma_j^{ref}(\mathbf{x}, T) \right|$$
$$\leq (atol)_j + (rtol)_j |\Gamma_j^{ref}(\mathbf{x}, T)|, \quad j = 1, \ldots, N_S; \forall (\mathbf{x}, T) \in \Phi, \tag{2c}$$

$$(N_R)_j w_j \geq \sum_{\{i :\, \nu_{ji} \neq 0\}} z_i, \quad j = 1, \ldots, N_S, \tag{2d}$$

$$z_i \in \{0, 1\}, \quad i = 1, \ldots, N_R, \tag{2e}$$

$$w_j \in \{0, 1\}, \quad j = 1, \ldots, N_S, \tag{2f}$$

where Φ is the Cartesian product of intervals $\Phi = [\mathbf{x}^L, \mathbf{x}^U] \times [T^L, T^U]$ in mass fraction-temperature space. Note that the set Φ in (2) replaces the finite set of reference points in (1). Even though (2) is linear in the decision variables, it contains infinitely many constraints, indexed by the set Φ, and is classified as a semi-infinite program (SIP) [10].

Solving SIPs to global optimality is computationally demanding, and current algorithms cannot determine an optimal solution for large problem instances, corresponding to large kinetic mechanisms. Bhattacharjee *et al.* [2] developed an algorithm using interval extensions [9] that determines guaranteed feasible points for an SIP by formulating a restriction of the SIP. The optimal objective function value of the resulting program is an upper bound on the solution of the exact formulation of the original SIP. In the case of the SIP in (2), this restriction will be an ILP.

Oluwole *et al.* [10] used ideas from the SIP restriction algorithm of Bhattacharjee *et al.* within a range-constrained reaction elimination formulation to overestimate the difference between the detailed model source terms and the reduced model source terms with constraints similar to (2b) and (2c) and determine reduced models guaranteed to satisfy the error constraints for reaction elimination. Our proposed formulation extends the work of Oluwole *et al.* by adding the binary variables w_j and the mass conservation constraint in (2d); it extends the work of Mitsos *et al.* by replacing the finite set of reference points with an interval in state space and modifying the appropriate constraints accordingly. In applying the approach of Oluwole *et al.* to (2), DAEPACK [15] is used to generate Taylor model interval extensions that overestimate the error between the detailed model source terms and reduced model source terms. These overestimates are used to formulate the ILP restriction of (2), which is then solved to global optimality using CPLEX [5]. Since an optimal solution of the ILP restriction of (2) corresponds to a feasible point of the range-constrained SIP formulation (2), the objective function value at this point is an upper bound on the optimal objective function value for the SIP. Cast in terms of the problem at hand, our proposed algorithm will determine a reduced model with fewer species that is guaranteed to satisfy error bounds on the source term in the reduced model over a specified range of conditions of interest, but it will not necessarily determine the reduced model with the fewest species satisfying those error bounds over the range of interest.

IV. CONCLUSIONS AND FUTURE WORK

A method for automatically generating error-controlled range-validated reduced kinetic mechanisms by simultaneous reaction and species elimination is presented. The method follows either by extending the point-constrained simultaneous species and reaction elimination formulation to a range-constrained formulation, or by extending the range-constrained reaction elimination formulation to include simultaneous reaction and species elimination. The resulting range-constrained simultaneous reaction and species elimination formulation is a semi-infinite program solved approximately by constructing an ILP restriction using interval extensions and solving that ILP to global optimality. The solution obtained by this algorithm is a feasible point for the original SIP, but it is not necessarily the best possible.

Given the recent proliferation of model reduction methods in the literature, it would be interesting to attempt to cast some of the existing model reduction methods into the form of optimization problems for the sake of comparison. Model reduction methods could then be compared using comparable error constraints in order to determine the extent to which they reduce the computational requirements of reacting flow solvers. This endeavor would also be useful in that it could facilitate the combination of different model reduction methods in order to further reduce the computational requirements of reacting flow solvers.

Finally, it would be interesting to examine how error-controlled model reduction methods interact with different numerical methods, as well as different problem formulations. Currently, little theory exists [10] to bound the errors in the numerical solution of a reduced model, as compared to the numerical solution of its corresponding detailed model. A theory explaining how error-controlled model reduction methods interact with different numerical methods for solving reacting flow problems could then be applied to range-constrained, error-controlled model reduction techniques to solve reduced models to within known error bounds, greatly enhancing the utility of model reduction techniques by clearly quantifying the error incurred by model reduction.

ACKNOWLEDGMENTS

We would like to thank Joseph K. Scott and Dr. Ajay Selot for helpful discussions. GMO gratefully acknowledges support under a DOE CSGF fellowship provided under grant DE-FG02-97ER25308. Financial support from the U.S. Department of Energys Office of Basic Sciences, Division of Chemical Sciences, Geosciences and Biosciences through Grant DE-FG02-98ER14914 is gratefully acknowledged.

REFERENCES

[1] B. Bhattacharjee. *Kinetic Model Reduction using Integer and Semi-infinite Programming*. PhD thesis, MIT, December 2003.

[2] B. Bhattacharjee, W. H. Green, and P. I. Barton. Interval methods for semi-infinite programs. *Computational Optimization and Applications*, 30:63–93, 2005.

[3] B. Bhattacharjee, D. A. Schwer, P. I. Barton, and W. H. Green. Optimally-reduced kinetic models: reaction elimination in large-scale kinetic mechanisms. *Combustion and Flame*, 135:191–208, 2003.

[4] H. J. Curran, P. Gaffuri, W. J. Pitz, and C. K. Westbrook. A comprehensive modeling study of n-heptane oxidation. *Combustion and Flame*, 114:149–177, 1998.

[5] ILOG. *ILOG CPLEX 11.0 User's Manual*.

[6] S. H. Lam. Using CSP to understand complex chemical kinetics. *Combustion Science and Technology*, 89(5):375–404, 1993.

[7] T. Lu and C. K. Law. Systematic approach to obtain analytic solutions of quasi steady state species in reduced mechanisms. *Journal of Physical Chemistry A*, 110:13202–13208, 2006.

[8] A. Mitsos, G. M. Oxberry, P. I. Barton, and W. H. Green. Optimal automatic reaction and species elimination in kinetic mechanisms. *Combustion and Flame*, 155(1-2):118–132, 2008.

[9] R. E. Moore. *Methods and Applications of Interval Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 1987.

[10] O. O. Oluwole, P. I. Barton, and W. H. Green. Obtaining accurate solutions using reduced chemical kinetic models: a new model reduction method for models rigorously validated over ranges. *Combustion Theory and Modelling*, 11(1):127–146, 2007.

[11] O. O. Oluwole, B. Bhattacharjee, J. E. Tolsma, P. I. Barton, and W. H. Green. Rigorous valid ranges for optimally reduced kinetic models. *Combustion and Flame*, 146:348–365, 2006.

[12] D. A. Schwer, P. Lu, and W. H. Green. An adaptive chemistry approach to modeling complex kinetics in reacting flows. *Combustion and Flame*, 133:451–465, 2003.

[13] D. A. Schwer, J. E. Tolsma, W. H. Green, and P. I. Barton. On upgrading the numerics in combustion chemistry codes. *Combustion and Flame*, 128:270–291, 2002.

[14] S. Singh, J. M. Powers, and S. Paolucci. On slow manifolds of chemical reactive systems. *Journal of Chemical Physics*, 117(4):1482–1496, 2002.

[15] J. Tolsma and P. I. Barton. DAEPACK: An open modeling environment for legacy models. *Industrial and Engineering Chemistry Research*, 36(6):1826–1839, 2000.