

**On the Application of Betweenness Centrality in Chemical Network Analysis:
Computational Diagnostics and Model Reduction**

Peng Zhao¹, Samuel M. Nackman¹ and Chung K. Law^{1,2,†}

¹*Department of Mechanical and Aerospace Engineering,
Princeton University, Princeton, NJ 08544, USA*

²*Center for Combustion Energy, Tsinghua University,
Beijing 100084, China*

Abstract

In this work, the concept of shortest path and betweenness centrality (BC) is introduced in combustion systems, with its application demonstrated in chemical network analysis and skeletal mechanism generation. After establishing the graphical representation of a chemical network for given pressure, temperature, and species concentrations, a metric BC is defined to rank the shortest paths passed by other nodes between the pair of source and target node, and as such captures the dominant indirect kinetic pathways between any pair of nodes in the network for computational diagnostics. Consequently, the controlling pathways are retained by collecting the species with larger BC values. Such a notion then indicates that the index of importance of species could then be assigned based on their BC values to further guide skeletal mechanism generation. Unlike existing methods, the betweenness centrality approach takes into account of both the fluxes between species and their relative positioning within the chemical network. To demonstrate its potential utility to combustion studies, the approach was applied to the GRI-3.0, LLNL and USC-Mech II mechanisms to identify the important pathways in the chemical network at each local reaction state, and develop skeletal mechanisms from all reaction state samples in auto-ignition and perfectly stirred reactor (PSR) simulations. The performance of the BC ranking is compared to the methods of directed relation graph (DRG), DRG with error propagations (DRGEP) and sensitivity analysis (SA), and is shown to possess sufficient utility in producing skeletal mechanisms with good accuracy and flexibility for the cases studied.

Keyword: shortest path, betweenness centrality (BC), chemical reaction network, species ranking, mechanism reduction, computational diagnostics

† Corresponding author. E-mail: cklaw@princeton.edu

1. Introduction

Chemical kinetics of fuels oxidation invariably involve a large number of strongly-coupled species and reactions, which not only impose considerable difficulty in identifying the controlling entities, but also require high computational cost to attain the understanding, validation, and prediction of the underlying complex chemical and flame processes [1]. Consequently there exists considerable interest in developing computational tools for the diagnostics of the processes and phenomena of interest, and for the reduction of the size and complexity of the detailed reaction mechanisms describing them. For computational diagnostics, sensitivity analysis (SA) [2], computational singular perturbation (CSP) [3] and its extension to the chemical explosive mode analysis (CEMA) [4] have been developed to identify important species and reactions for global and local parameters. However, there are many situations for which these methods are either not feasible or capable of providing the information needed. For example, in the direct numerical simulation (DNS) of a turbulent flame, it is not practical to perturb a certain reaction rate constant and redo the simulation to obtain the sensitivity coefficients. CEMA is a useful tool to obtain the explosive mode of the local chemical Jacobian matrix and the corresponding contribution from each species and reaction; however, it is not able to be applied to mixtures without a chemical explosive mode or indicate the important pathway associated with the production of certain species.

Regarding mechanism reduction, methods based on graph theory, notably the directed relation graph (DRG) [5-7], and the subsequent extensions of path flux analysis (PFA) [8] and DRG with error propagation (DRGEP) [9], have been developed, resulting in substantial simplifications in reducing the model size and the subsequent application in combustion modeling [10-14]. These methods use different criteria to rank species based on the sampling of reaction states, locally characterized by temperature, pressure, and species concentrations. The species rankings are subsequently used to guide the elimination of species from a chemical mechanism. A systematic comparison of the DRG-based methods can be found in [15].

In the present work, the concept of betweenness centrality (BC), which measures the relative connectivity of a node in a network based on the shortest paths [16], is adopted to study the complex chemical networks in combustion systems [17]. Our interest in exploring the potential of this concept in chemical systems is motivated by recognizing that over the past decade this concept has been

widely and fruitfully employed in the analysis of equally complex networks that exist in, for example, social science [18], computer science [19], biology [20], transport [21], scientific cooperation [22], and interdisciplinary interactions [23]. In the following we shall first present the formulation of the concept of BC within the framework of the chemical network. We shall then demonstrate its utility in autoignition as well as perfectly stirred reactor (PSR) systems, and compare its performance with those of DRG, DRGEP and SA. It will be shown that the concept of BC, which considers both fluxes and the relative positioning of species within a chemical network, provides useful information of indirect kinetic influence between any pair of species, and further generates a species ranking that both effectively quantifies the relative importance of species and predicts target parameters in combustion simulations with considerable accuracy and flexibility.

2. Graphical Representation and Methodology

Different from graph representation in the DRG based methods, the forward and backward reactions of each elementary reaction are treated separately so as to retain as much kinetic information as possible. A graphical representation of a chemical network is chosen such that each node represents a chemical species and an edge of weight L_{AB} between nodes A and B is defined as the sum of the reaction rates, with species A being a reactant and B a product:

$$L_{AB} = \sum_{i=1}^N v_{A,i}(\omega_{f,i}\delta_{f,i}^{AB} + \omega_{b,i}\delta_{b,i}^{AB})$$

where $\omega_{f,i}$, $\omega_{b,i}$ are the forward and backward reaction rates, respectively, for the i^{th} reaction, and $v_{A,i}$ is the stoichiometric coefficient for species A in the i^{th} reaction. $\delta^{AB} = 1$ if and only if the forward or backward direction for the i^{th} reaction includes A as a reactant and B as a product, and $\delta^{AB} = 0$ otherwise. Such a definition denotes a lumped consumption rate of species A for the formation of species B and the adjacency matrix L is asymmetric.

Recognizing that a species can participate in many elementary reactions, for an arbitrary reversible reaction $R_{AB,l}$ with index l involving both A and B, which can be in the form of either $v_A A + v_M M \leftrightarrow v_B B + v_N N$ or $v_A A + v_B B \leftrightarrow v_M M + v_N N$, the consumption rate of A for the formation of B through either form of $R_{AB,l}$ can be expressed as: $v_{A,l}(\omega_{f,l}\delta_{f,l}^{AB} + \omega_{b,l}\delta_{b,l}^{AB})$.

Therefore, the ratio of the consumption rate of A due to the formation of B through reaction $R_{AB,l}$ to the total consumption rate of A is:

$$P(R_{AB,l}) = \frac{v_{A,l}(\omega_{f,l}\delta_{f,l}^{AB} + \omega_{b,l}\delta_{b,l}^{AB})}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_R} v_{A,j}(\omega_{f,j}\delta_{f,j}^{AM_i} + \omega_{b,j}\delta_{b,j}^{AM_i})}$$

where the subscripts i and j designate the i^{th} species and the j^{th} reaction, respectively. Assuming that each elementary reaction between A and B occurs independently, the normalized rate for the paths from nodes A to B can be measured as:

$$\begin{aligned} P_{AB} &= P(R_{AB,1}) + P(R_{AB,2}) + \dots + P(R_{AB,n}) \\ &= \frac{\sum_{j=1}^{N_R} v_{A,j}(\omega_{f,j}\delta_{f,j}^{AB} + \omega_{b,j}\delta_{b,j}^{AB})}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_R} v_{A,j}(\omega_{f,j}\delta_{f,j}^{AM_i} + \omega_{b,j}\delta_{b,j}^{AM_i})} = \frac{L_{AB}}{\sum_{i=1}^{N_s} L_{AM_i}} \end{aligned}$$

Although tracking the flux of element in each elementary reaction would be difficult, the above definition then shows the physical meaning of the fraction of the consumption of species A by the formation of species B. It is noted that there is a major difference in the manner of species B interacting with species A from DRG and its extensions. Specifically, species A must appear on the reactant side and B must appear on the product side in a reaction and hence it carries the information of the direction of flux from A to B. As a result, the normalized rate of such a transfer along any pathway between two nodes (from source M_l to target M_n) can be modeled as:

$$P(M_1 \dots M_n) = \prod_{i=1}^{n-1} P_{M_i M_{i+1}}$$

This quantity then measures the indirect influence of the source node on the target node in a chemical network. It is not unreasonable to expect that the larger the value of P along a certain path, the more possible is the transfer from the source to the target. Therefore, it would be of interest to investigate paths with large P values including the most probable path, connecting any two nodes of interest in the reaction network. As will be shown in the following sections, the pathways identified generate useful information of reaction paths especially for nodes without direct connections, which so far cannot be systematically obtained for computational diagnostic purposes.

We now introduce the BC concept in the adjacency matrix P to describe the effectiveness of a node in connecting any other two nodes as:

$$W_k = \sum_{s \neq t \neq k} \frac{p_{st}(k)}{p_{st}} = \sum_{s \neq t \neq k} \frac{p_{sk} \cdot p_{kt}}{p_{st}}$$

where p_{st} is the most probable path from nodes s to t when travelling through the adjacency matrix P by any possible path, *i.e.*, $p_{st} = \max P(s, \dots, t)$ for all possible paths from s to t . The BC is defined for each node k as the summation of the ratio of the most probable path from nodes s to t to the most probable path from nodes s to t passing through node k , which considers how frequently a node serves as a hub between other nodes.

In order to numerically evaluate BC efficiently, by further defining a new matrix A as $A_{ij} = -\ln P_{ij}$, it is readily shown that

$$e^{-\text{path length of } ABC} \equiv e^{-A_{AB}-A_{BC}} = e^{\ln P_{AB} + \ln P_{BC}} = P_{AB} P_{BC} = P_{ABC}$$

which means that the shortest path in A corresponds to the most probable path in P . And such a shortest path, σ , can be readily found by using Dijkstra's algorithm [24] with a time complexity of $O(|E| + |V| \log |V|)$, in which $|E|$ is the number of edges and $|V|$ the number of vertices. Consequently the metric of BC can be defined to rank the species for removal:

$$W_k = \sum_{k \neq s, t} \frac{e^{-\sigma(A, s, k)} \cdot e^{-\sigma(A, k, t)}}{e^{-\sigma(A, s, t)}} = \sum_{k \neq s, t} e^{\sigma(A, s, t) - \sigma(A, s, k) - \sigma(A, k, t)}$$

where $\sigma(A, i, j)$ is the length of the shortest path from node i to j in the adjacency matrix A , and W_k is the BC for the k^{th} node. Therefore, the rank of a node by its BC value implies the rank of the shortest path passing it, *i.e.*, with a larger BC value, a node lies on a shortest path with higher rank. Due to the large size of the adjacency matrix A and its asymmetry, if two nodes happen to have the same value of W , they probably lie on the same shortest path.

In combustion systems, since the global flux originates from reactants (Fuel, O₂) and terminates to major products (CO₂, H₂O, *etc.*) and important intermediates, it is reasonable and computationally convenient to just sum over a few pairs of (s, t) , such as (Fuel, CO₂), (O₂, H₂O), *etc.*, when evaluating W_k . To further reduce the overhead in calculating the shortest path from different starting

nodes, dummy source and target nodes are added in the adjacency matrix A . Specifically, a dummy source node (S) directly connects to the real nodes of origination (fuel and oxidizer) by equal edge weights. Likewise, a dummy target node (T) is directly connected to the real nodes of termination (major products CO_2 and H_2O or important intermediates) by equal edge weights. Therefore, the BC value can be simply evaluated from S to T , *i.e.*,

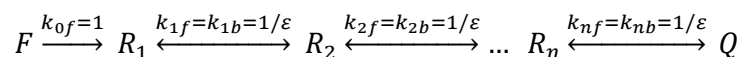
$$W_k = e^{\sigma(A,S,T) - \sigma(A,S,k) - \sigma(A,k,T)}$$

and both S and T can be disregarded after the calculation. The vector \mathbf{W} can then be used to identify different levels of shortest path and the importance of species. By collecting species from large to small BC values, the shortest paths connecting the source and target nodes are incorporated in the skeletal mechanism from high to low levels.

3. Demonstration of Shortest Path and BC

3.1 Generic Demonstration Problems

Reference [6] demonstrated the significance of combining the forward and reverse rates in terms of the consideration of partial equilibrium reactions, and raised the concern of the decay in the species relation for a long pathway in the DRGEP methods. To show the applicability of the current definitions, we consider an example similar to the one described in [6] and [9], namely a path from a fuel F to a product Q that goes through several intermediate species R_1 to R_n :



The first reaction is much slower, and all the subsequent reactions are fast ones to reach partial equilibrium after a short transient period. Then the first reaction becomes rate-controlling and the concentrations of R_i s and Q are approximately equal. The species correlations can be calculated as:

$$P_{FR_1} = \frac{\omega_{0f}}{\omega_{0f}} = 1, \quad P_{R_1R_2} = \frac{\omega_{1f}}{\omega_{1f}} = 1,$$

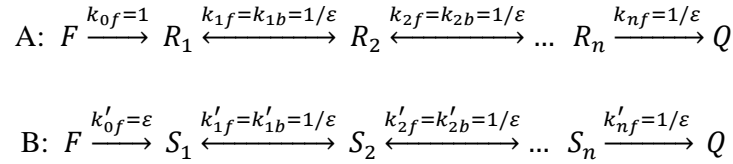
$$P_{R_2R_3} = \frac{\omega_{2f}}{\omega_{1b} + \omega_{2f}} = 0.5 = P_{R_iR_{i+1}} = P_{R_nQ} \quad (i = 2, \dots, n-1)$$

Similarly, $P_{R_2R_1} = P_{R_{i+1}R_i} = 0.5$ ($i = 1, \dots, n-1$), $P_{QR_n} = 1$. The most probable path for R_n from F is then $p(F, R_1, \dots, R_{n-1}, R_n) = 0.5^{n-2}$, which approaches 0 when n is large; however, the BC of species R_n between the pair (F, Q) is then calculated as:

$$W_{R_n} = \frac{p_{FR_n} \cdot p_{R_nQ}}{p_{FQ}} = \frac{0.5^{n-2} \cdot 0.5}{1 \cdot 1 \cdot 0.5^{n-1}} = 1,$$

which implies the importance of R_n to F is not underestimated. Actually, all the other species lie on the same shortest path, which is the only reaction path connecting F and Q . Therefore, all intermediates are equally important for the production of Q and no species can be removed.

We next examine a more complicated case with two paths from reactant F to product Q , specifically:



for both channels A and B; the first step is much slower than the subsequent reactions, which reach partial equilibrium after a short transient period. The difference is that channel B needs a much longer relaxation time than A. For this system, it is obvious that channel A dominates over channel B for the generation of product Q . By definition,

$$P_{FR_1} = \frac{\omega_{0f}}{\omega_{0f} + \omega'_{0f}} = \frac{1}{1 + \varepsilon}, \quad P_{FS_1} = \frac{\omega'_{0f}}{\omega_{0f} + \omega'_{0f}} = \frac{\varepsilon}{1 + \varepsilon}, \quad P_{R_1R_2} = P_{S_1S_2} = 1$$

Similarly, $P_{R_iR_{i+1}} = P_{S_iS_{i+1}} = P_{R_nQ} = P_{S_nQ} = 0.5$ ($i = 2, \dots, n-1$). The shortest path from F to Q is through channel A and therefore the BC value of species R_i is 1, the same as in the last case. Take S_3 for example, the BC value is then:

$$W_{S_3} = \frac{p_{FS_3} \cdot p_{S_3Q}}{p_{FQ}} = \frac{\frac{\varepsilon}{1 + \varepsilon}}{\frac{1}{1 + \varepsilon}} = \varepsilon$$

It is clear that channel B is the second shortest path connecting F and Q , and thus the species along it are not as important as those along A. This therefore further demonstrates that the BC concept

considers of both the flux and the relative positions within the network, and merits its application in more complex reaction systems.

3.2 Autoignition of methane

We first select GRI-3.0 [25] to demonstrate the shortest path and the BC concept. The mechanism consists of 53 species and 325 elementary reactions, with methane oxidation as well as the NO_x chemistry. The first condition tested is a stoichiometric mixture of methane and air at atmospheric pressure and initial temperature of 1600K, as shown in the mole fraction and temperature history plot in Fig.1. Two points A and B are marked to demonstrate the concepts, with point A chosen just before ignition, at $t=4.00\text{E-}4\text{s}$ and $T=1656.7\text{K}$, and point B chosen after ignition but before the system reaches chemical equilibrium, at $t=6.00\text{E-}4\text{s}$ and $T=2681.3\text{K}$. Local BC values of each species are evaluated for the pairs $(\text{CH}_4, \text{CO}_2)$ and $(\text{CH}_4, \text{C}_2\text{H}_2)$ at point A, and therefore different levels of shortest paths are identified between the source and target pairs. As shown in Fig. 2(a), CH_3 , HCO and CO have the largest BC value, forming the shortest path connecting the pair $(\text{CH}_4, \text{CO}_2)$; followed by CH_3O and CH_2O with the second largest BC value, along the second shortest path, etc. Hence, the results show the major pathway for CO_2 formation is to first form CH_3 through H abstraction, followed by further partial oxidation to form CH_3O , CH_2O and HCO and eventually CO and CO_2 . For the C_2H_2 formation shown in Fig. 2(b), CH_3 , C_2H_5 , and C_2H_4 have the largest BC value, serve as the shortest path: $\text{CH}_4\text{—CH}_3\text{—C}_2\text{H}_5\text{—C}_2\text{H}_4\text{—C}_2\text{H}_2$. With many indirect reaction channels, the shortest paths identified coincide with the existing understanding of methane oxidation process, such as in [26], and provide useful information of the dominant global reaction pathway at the local point. In many applications, a correct identification for the cause of pollutant (such as NO) is the prerequisite for its control, due to different working conditions and various formation mechanisms. For point B immediately after ignition, compared with the formation of the major species such as CO_2 and H_2O , it takes much longer for the formation of NO as seen in Fig. 1. Therefore, the BC value is evaluated for the pair (N_2, NO) to identify the major pathway for NO formation. It is found that the shortest path is $\text{N}_2\text{—NO}$ and the second shortest path is $\text{N}_2\text{—N—NO}$, which indicates that for the current case, the main NO formation is from the Zel'dovich mechanism $\text{N}_2+\text{O}=\text{NO}+\text{N}$, $\text{O}_2+\text{N}=\text{NO}+\text{O}$ and $\text{N}+\text{OH}=\text{NO}+\text{H}$.

With the shortest path and BC shown at an instant reaction state, we shall next illustrate its application in global sampling for the entire ignition process. Ignition delay τ is first defined as the time with the maximum temperature increase rate, and enough sampling points (35) are chosen in the sampling range from 0.001τ to 10τ . When evaluating the global BC, the dummy source node S is chosen to connect to CH_4 and O_2 and the dummy target node T is then connected to CO_2 and H_2O . The BC values of each species are then calculated at each sampling point and the maximum among all sampling points is taken as the overall BC for each species during ignition. Such overall BC of each species is normalized by the maximum value. As a comparison, an improved version of DRG [7] and DRGEP [9] are applied, in which the species rankings are given by the maximum DRG and DRGEP values among the same sampling points, respectively.

The results show that radicals such as OH, HO_2 , O, *etc.* generally have higher BC, which is reasonable since they involve many reactions and interact with many other species in the chemical network. On the other hand, species such as HOCN participates in only a few reactions; therefore the small BC indicates their weak roles in connecting the network. We next assess the ranking of BC and how it can be used in guiding the chemical network analysis in, for example, skeletal mechanism generation. Based on the sampling and calculation in Fig. 3, skeletal mechanisms of different sizes can be built up by sequentially keeping species with higher rankings of overall BC. The relative error of the ignition delay for each skeletal mechanism is then plotted in Fig. 4 for DRG, DRGEP and BC rankings. With the same sampling points and error evaluation for all sizes of skeletal mechanism, it is expected that the effectiveness of providing satisfactory index of importance by each method can be observed.

It is seen that with the same sampling, all methods can predict the global parameter of the ignition delay time quite well, and the obvious trend that a larger mechanism size results in smaller error. More importantly, however, the results show that the ranking given by BC is more accurate than DRG and DRGEP for almost all sizes of the skeletal mechanisms generated. To explore the cause of this improved performance, Fig. 5 shows the sensitivity of ignition delay to each species, which is defined as the relative error in the ignition delay by removing only the corresponding species from the detailed mechanism. It is then demonstrates that, as a general trend, those species with higher and lower sensitivities are respectively ranked relatively high and low by the BC value;

while in the DRG ranking, some species with high sensitivities are ranked relatively low. If this is the reason for the improved performance of BC, then it would be of interest to evaluate the performance of ranking by the sensitivity of each species, although it is computationally expensive. The relative error in the ignition delay by the sensitivity analysis (SA) ranking is also shown as circles in Fig. 4. Not surprisingly, among all four rankings, the ranking by SA has the lowest error for almost all mechanism sizes.

Further studies with different temperatures, pressures and equivalence ratios show that the ranking by BC predicts ignition delay time more accurately, which largely coincides with the species sensitivity ranking, as detailed in the *Supplementary Materials*. From the definition of BC, such a correspondence is reasonable in that a species with a higher value of BC generally sits along the shortest paths connecting a major pair of species, and once such a species is removed, many important paths passing it are consequently removed, leading to a “detour” in major paths and therefore flux into secondary channels and affecting the radical pool built-up process during ignition. To further show the chemical fidelities of the skeletal mechanisms generated, Fig. 5 compares selected species concentrations calculated by the detailed and skeletal mechanisms with 25 species by all rankings. As a general trend, all mechanisms perform well to capture the change of the major and minor species in the phase plot. However, compared with the discrepancy from the CH₄, CH₃, HO₂, OH profiles predicted by the DRG and DRGEP skeletal mechanisms of the same size, such difference for the BC mechanism is hardly discernible.

To give a better sense of skeletal mechanism generation over a wider range of system parameters, using the BC method, a skeletal mechanism for CH₄/air mixture with 32 species is produced based on GRI 3.0, yielding accurate ignition delays over the parametric space (pressure ranging from 0.1 to 30 atm, equivalence ratio from 0.5 to 1.5, initial temperature from 1000 to 1600 K). In addition, we have also validated it for the NO history, PSR extinction and laminar flame speed, demonstrating satisfactory performance as shown in Fig. S3-S12 in the *Supplementary Materials*.

3.3 Two-Stage Ignition of *n*-Heptane

A more complicated case for autoignition is the two-stage ignition phenomena in the negative-temperature coefficient (NTC) regime [27], with the first-stage caused by the low

temperature chemistry and producing intermediate species such as aldehydes and ketones, followed by the second stage driven by H_2O_2 chain branching and small radical reactions. From previous work on model reduction, this is usually the regime where the largest reduction error in the ignition delay occurs. Consequently, we have selected the autoignition of stoichiometric *n*-heptane/air mixture at the initial temperature 650K and atmospheric pressure within the NTC regime to demonstrate the application of shortest paths and BC, using the LLNL *n*-heptane mechanism [27] of 561 species and 2539 reactions. The mole fractions of selected species during the two-stage autoignition are shown in Fig. 7. It is shown that both the reactant depletion and the major species accumulation processes exhibit the two-stage behavior. Since aldehydes are generally considered as the major products for low-temperature chain branching reactions [28] and CO is a good indicator for the low-temperature reactivity [29], it would be of interest to apply the current method to identify the major reaction pathway for their formation in light of the existing understanding. Consequently the local BC of each species connecting the pairs (O_2 , $\text{C}_2\text{H}_5\text{CHO}$) and ($n\text{C}_7\text{H}_{16}$, CO) are evaluated respectively at Point C just before the first stage ignition with $t=5.74\text{E-}2\text{s}$ and $T=800.0\text{K}$, as shown in Fig. 7. Figure 8(a) show that the formation of $\text{C}_2\text{H}_5\text{CHO}$ is mainly from the low-temperature chemistry, yet from two different major channels identified from the shortest path: the first is through the formation of ketohydroperoxides and its decomposition; and the second is from the second and third shortest path, representing the decomposition of cyclic ether formed from the QOOH radicals [27]. For the CO formation from *n*-heptane at low temperatures, the major pathway as shown in Fig. 8(b) is similar to that identified in the high-temperature methane autoignition case, through the formation of HCO radicals. However, because of the long chain structure of *n*-heptane and the various decomposition products, there are additional major pathways for CO formation, such as those from the decomposition of propionyl ($\text{C}_2\text{H}_5\text{CO}$) radicals.

For the global sampling of the two-stage ignition, the major intermediates and products vary with time due to the transition from low to high temperature chemistry. Therefore in order to evaluate the BC more properly, a dynamic evaluation process is introduced. Specifically, at each sampling point during ignition, according to the ranking of the local concentrations, the top 5% species (excluding original fuel, oxidizer, and major inert) will be chosen as the major intermediates and to connect to the dummy target node T , while the dummy source node S is still connected to the original fuel and oxidizer. Such treatment could improve the capturing of the controlling chemistry

locally without increasing much the computational cost. By the same sampling and error evaluation process, the relative errors in the ignition delay as a function of the skeletal mechanism size based on the BC/DRG/DRGEP rankings are compared in Fig. 9. It is shown that with dynamic evaluation, BC ranking can capture the ignition delay time more accurately, with the lowest error for almost all mechanism sizes. Figures 10(a) and (b) compare the selected species concentrations calculated by the detailed and the skeletal mechanisms with 140 species for all rankings. It is seen that, compared with the DRG and DRGEP methods, better agreement is achieved with the detailed model for both major species and selected intermediates predicted by the BC ranking, hence further showing its capability of retaining major formation channels to major intermediates and products.

3.4 PSR Studies

While auto-ignition describes the low through moderately high-temperature ignition chemistry, PSR captures the high-temperature chemistry relevant for intense burning and extinction. To show the application of the BC in PSR, a case of stoichiometric 1-butene/air mixture with inlet temperature 1000K is calculated with USC-Mech II [30], which consists of 111 species and 784 reactions. The S-curve response of the reactor temperature to residence time is shown in Fig. 11.

To show the application for computational diagnostics, a local point with temperature 2262.8K on the upper branch of the S-curve is selected, and the shortest paths to be evaluated are set between the fuel molecule to benzene, and such pathway could provide useful information for the PAH and soot precursor formation. Figure 12 shows that the major pathways for the benzene formation are largely through the formation of C_3H_3 radicals, hence further supporting the well-accepted notion for their role as benzene precursors [31].

For global sampling, to represent all specific reaction states and relations among the species, about 20 points are sampled along the S-curve including the upper intense burning branch and the intermediate branch. When evaluating the BC, the dummy source node connects to the original reactants 1-butene and O_2 , while CO_2 and H_2O are chosen as the major products to connect to the dummy target node. Similar to the autoignition case, the overall importance of the species is then ranked by its maximum BC value among all the sampling points. Therefore, skeletal mechanism with different sizes can be constructed by keeping species from high to low in the ranking. In Fig. 11, it is

seen that the two typical skeletal mechanisms with various sizes from 40 to 80 can all capture the entire S-curve well. As a comparison, DRG and DRGEP are applied with the same sampling points, and the relative error in residence time is calculated for every sampling point by keeping the temperature fixed. For each skeletal mechanism, the worst error is given by the maximum of the relative error in the residence time among all the sampling points, as in Fig. 13. It is seen that the ranking by BC can generate mechanisms with the same size but improved accuracy for almost all sizes, showing its effectiveness in capturing the major species along the dominant chemical pathways.

4. Conclusions

The concept of betweenness centrality (BC) is introduced in chemical networks for the purposes of computational diagnostics and skeletal mechanism generation, recognizing that this concept not only considers the flux between nodes, but it also accounts for the relative positions of nodes in the network. By applying it at the local points of autoignition and PSR conditions, it is shown that the metric BC can satisfactorily detect the local controlling chemistry, such as intermediate species formation, low-temperature chemistry, NO and soot precursor formation, etc. Therefore, it is a useful tool to identify the indirect chemical effect between any nodes, complementing the existing tools such as sensitivity analysis and chemical explosive mode analysis. For global sampling, it is found that the BC concept could properly rank the importance of the species, as shown in both high/low temperature autoignition and PSR combustion systems. It is demonstrated that the skeletal mechanisms can be generated quite flexibly by BC ranking and can capture the target parameters as well as the detailed kinetics of important intermediates even more accurately as compared to the DRG and DRGEP with the same sampling and error evaluation. It is nevertheless also emphasized that the present study is exploratory in nature, focusing on the applicability of the BC concept in combustion systems. The results are encouraging, meriting further investigations to possibly enhance its utility and robustness as a computational tool for reaction network analysis and model reduction.

Acknowledgments

This work was supported by the Combustion Energy Frontier Research Center, an Energy Frontier Research Center funded by the US Department of Energy, Office of Basic Energy Sciences under

Award Number DESC0001198. The authors sincerely thank Professor Tianfeng Lu of the University of Connecticut for helpful discussions.

References

- [1] T. Lu, and C.K. Law, *Prog. Energy. Combust. Sci.*, 35(2009)192-215.
- [2] H. Rabitz, M. Kramer, and D. Dacol, *Ann. Rev. Phys. Chem.*, 34(1983) 419-461.
- [3] S. H. Lam, *Combustion Science and Technology*, 89 (1993) 375-404.
- [4] T.F. Lu, C.S. Yoo, J.H. Chen, C.K. Law, *J. Fluid Mech.*, 652 (2010) 45–64
- [5] T. Lu, and C.K. Law, *Proc. Combust. Inst.*, 30(2005) 1333-1341.
- [6] T. Lu, and C.K. Law, *Combust. Flame*, 146(2006) 472-483.
- [7] Z. Luo, T.F. Lu, M.J. Maciaszek, S. Som, D.E. Longman, *Energy & Fuels*, 24(2010) 6283 – 6293.
- [8] W. Sun, Z. Chen, X. Gou, Y. Ju, *Combust. Flame*, 157 (2010) 1298–1307.
- [9] P. Pepiot, H. Pitsch, *Combust. Flame* 154 (2008) 67–81.
- [10] C.K. Law, P. Zhao, *Combust. Flame*, 159 (3) (2012), pp. 1044–1054.
- [11] P. Zhao, C.K. Law, *Combust. Flame*, 160 (11) (2013), pp. 2352–2358.
- [12] C.S. Yoo, T.F. Lu, J.H. Chen, C.K. Law, *Combust. Flame*, 158 (2011), pp. 1727–1741.
- [13] A.P. Kelley, W. Liu, Y.X. Xin, A.J. Smallbone, C.K. Law, *Proc Combust Inst*, 33 (2011), pp. 501–508
- [14] S. Deng, P. Zhao, D. Zhu, C.K. Law, *Combust. Flame*, 161 (8), 1993-1997
- [15] L. Tosatto, B. Bennett, M.D. Smooke, *Combust. Flame*, 160 (2013), 1572–1582.
- [16] M.E.J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.
- [17] Samuel M. Nackman, *The Development of a Probability Betweenness Method for Chemical Kinetic Mechanism Reduction*, M.S. Thesis, Princeton University, NJ, USA, 2013.
- [18] T. Opsahl, F. Agneessens, J. Skvoretz, *Social Networks* 32(2010) 245–251.
- [19] X. Bei, W. Chen, S. Teng, J. Zhang, J. Zhu, *Theoretical Computer Science* 412(2011) 7147–7168
- [20] H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, *Nature* 411(2001): 41.
- [21] R. Puzis, Y. Altshuler, Y. Elovici, S. Bekhor, Y. Shifan, A. Pentland, *Journal of Intelligent Transportation Systems*, 17(2013) 91-105.
- [22] A. Abassi, L. Hossain, L. Leydesdorff, *Journal of Informetrics* 6(2012): 403–412.
- [23] L. Leydesdorff. *Journal of the American Society for Information Science and Technology* 58(2007) 1303-1319.
- [24] E. W. Dijkstra, *Numerische Mathematik* 1(1959) 269-271.
- [25] G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriarty, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, W.C. Gardiner, V.V. Lissianski, Z. Qin, <http://www.me.berkeley.edu/gri_mech/>.
- [26] C.K. Law, *Combustion Physics*, Cambridge University Press, New York (2006)
- [27] H.J. Curran, P. Gaffuri, W.J. Pitz, C.K. Westbrook, *Combust. Flame*, 114 (1998), pp. 149–177.
- [28] A. Kazakov, M. Chaos, Z.W. Zhao, F.L. Dryer, *J. Phys. Chem. A*, 110 (2006), pp. 7003–7009
- [29] R.D. Wilk, D.N. Koert, N.P. Cernansky, *Energy Fuels* 3 (1989) 292-298.

- [30] H. Wang, X. You, A.V. Joshi, S.G. Davis, A. Laskin, F. Egolfopoulos & C.K. Law, USC Mech Version II. High-Temperature Combustion Reaction Model of H₂/CO/C₁-C₄ Compounds. <http://ignis.usc.edu/USC_Mech_II.htm>, May 2007.
- [31] J.A. Miller, M.J. Pilling, J. Troe, Proc. Combust. Inst., 30(2005) 43-88.

Figure captions

Fig. 1. Temperature and mole fraction history of selected species for stoichiometric CH₄/air autoignition at initial temperature 1600K, pressure 1 atm.

Fig. 2(a). Normalized BC and shortest paths from CH₄ to CO₂ at Point A in Fig. 1.

Fig. 2(b). Normalized BC and shortest paths from CH₄ to C₂H₂ at Point A in Fig. 1.

Fig. 2(c). Normalized BC and shortest paths from N₂ to NO at Point B in Fig. 1.

Fig. 3. Rankings of species by maximum DRG value, maximum DRGEP value, and normalized maximum BC value over all sampling points for the autoignition case in Fig. 1.

Fig. 4. Relative error in ignition delay time for skeletal mechanism of all sizes by BC, DRG, DRGEP and sensitivity analysis (SA) methods for the same condition as in Fig.1.

Fig 5. Sensitivity of each species for rankings by DRG and BC method as in Fig.1, showing better correspondence between BC and sensitivity.

Fig. 6(a). Comparison of major species concentrations calculated using the detailed and skeletal mechanisms of 25 species by different methods for auto-ignition of stoichiometric CH₄-air under the condition of Fig. 1.

Fig. 6(b). Comparison of selected intermediate species concentrations calculated using the detailed and skeletal mechanisms of 25 species by different methods for auto-ignition of stoichiometric CH₄-air under the condition of Fig. 1.

Fig. 7. Species history in a two-stage ignition process of stoichiometric n-heptane/air with initial temperature 650 K under 1 atm.

Fig. 8(a). Normalized BC and shortest paths connecting O₂ and C₂H₅CHO at t=5.74E-2s and T=800.0K in Fig. 7.

Fig. 8(b). Normalized BC and shortest paths connecting fuel and CO at t=5.74E-2s and T=800.0K in Fig. 7.

Fig. 9. Relative error in ignition delay time for skeletal mechanism of all sizes by BC, DRG, DRGEP (starting from H and nC₇H₁₆) methods for the same condition as in Fig.7.

Fig. 10(a). Comparison of major species concentrations calculated using the detailed and skeletal mechanisms of 140 species by different methods for auto-ignition of stoichiometric n-heptane/air under the condition of Fig. 7.

Fig. 10(b). Comparison of intermediate species concentrations calculated using the detailed and skeletal mechanisms of 140 species by different methods for auto-ignition of stoichiometric n-heptane/air under the condition of Fig. 7.

Fig. 11. PSR calculation by USC II and skeletal mechanism of different sizes by BC method for stoichiometric 1-butene/air mixture with initial temperature 1000K and pressure 1atm.

Fig. 12. Normalized BC and shortest paths connecting 1-butene and benzene at local point $T=2262.8\text{K}$ on the upper branch of Fig. 11.

Fig. 13. Worst error in residence time predicted by skeletal mechanisms of all sizes by BC, DRG, DRGEP methods, for all sampling points along the S-curve of Fig. 11.

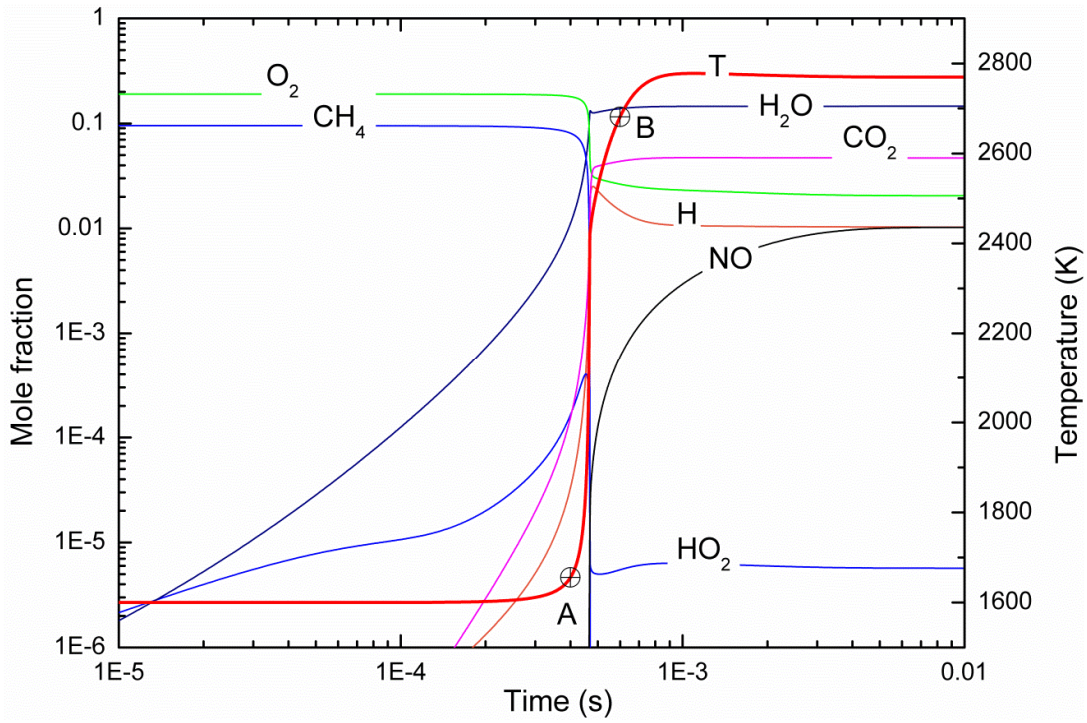


Fig. 1. Temperature and mole fraction history of selected species for stoichiometric CH₄/air autoignition at initial temperature 1600K, pressure 1 atm.

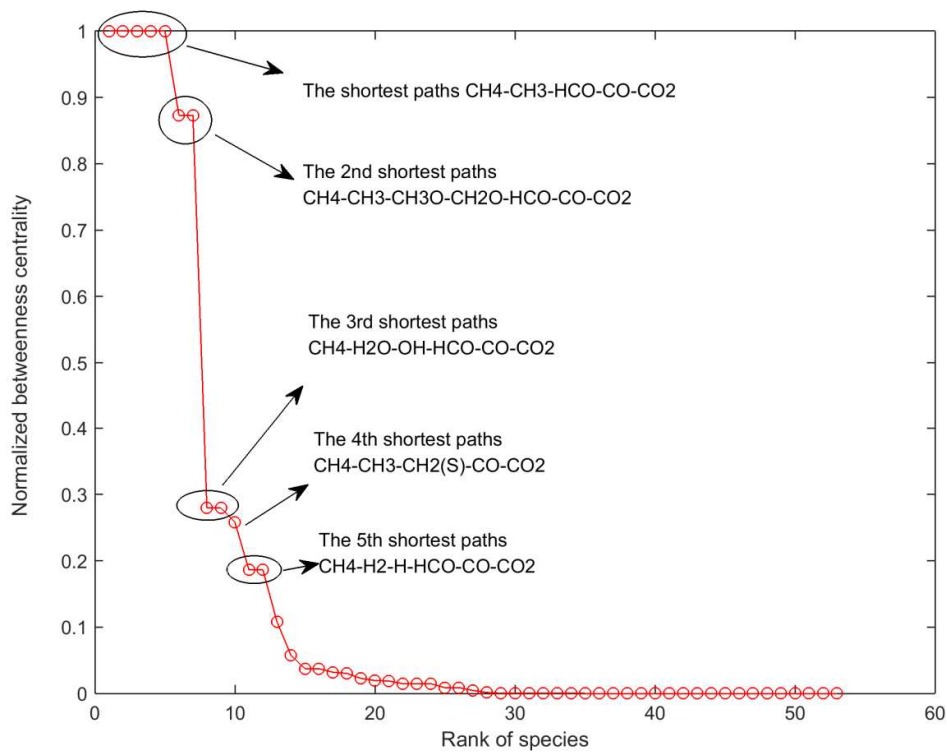


Fig. 2(a). Normalized BC and shortest paths from CH₄ to CO₂ at Point A in Fig. 1.

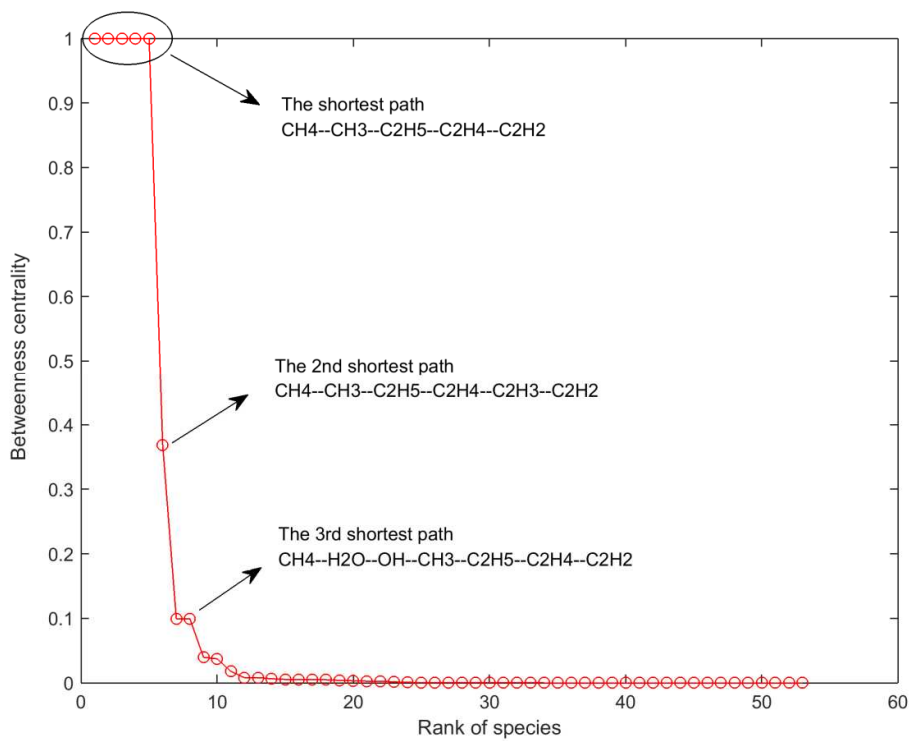


Fig. 2(b). Normalized BC and shortest paths from CH₄ to C₂H₂ at Point A in Fig. 1.

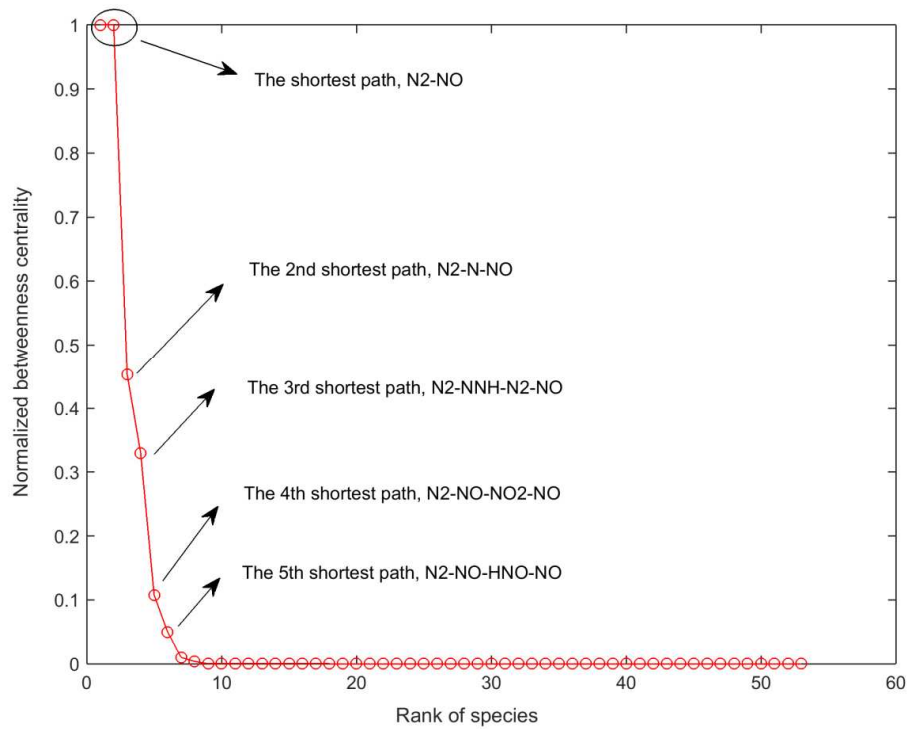


Fig. 2(c). Normalized BC and shortest paths from N₂ to NO at Point B in Fig. 1.

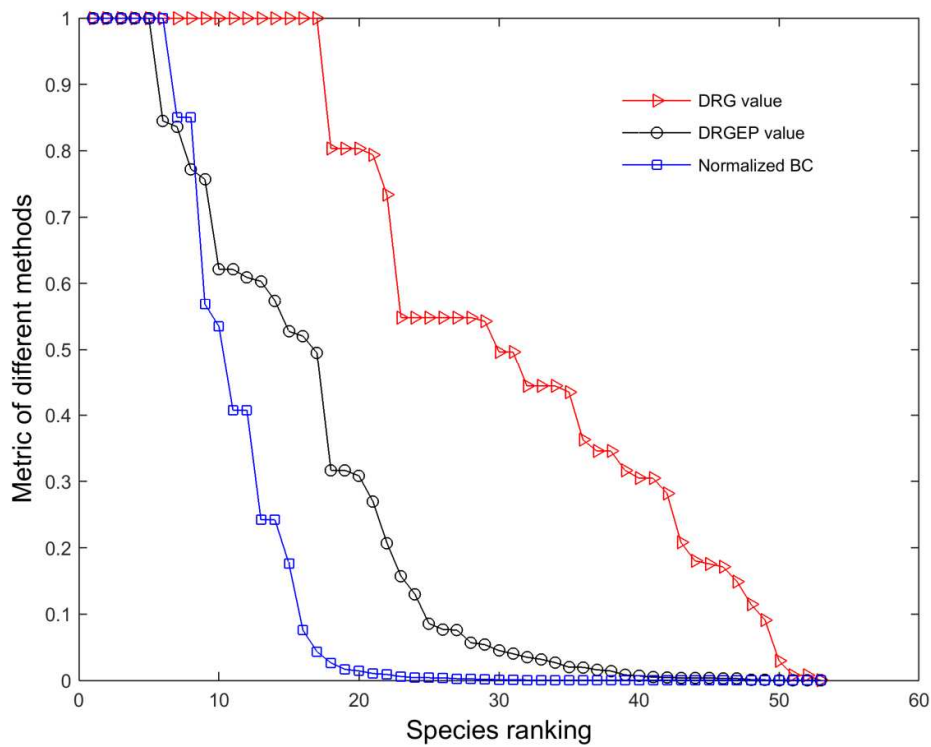


Fig. 3. Rankings of species by maximum DRG value, maximum DRGEP value, and normalized maximum BC value over all sampling points for the autoignition case in Fig. 1.

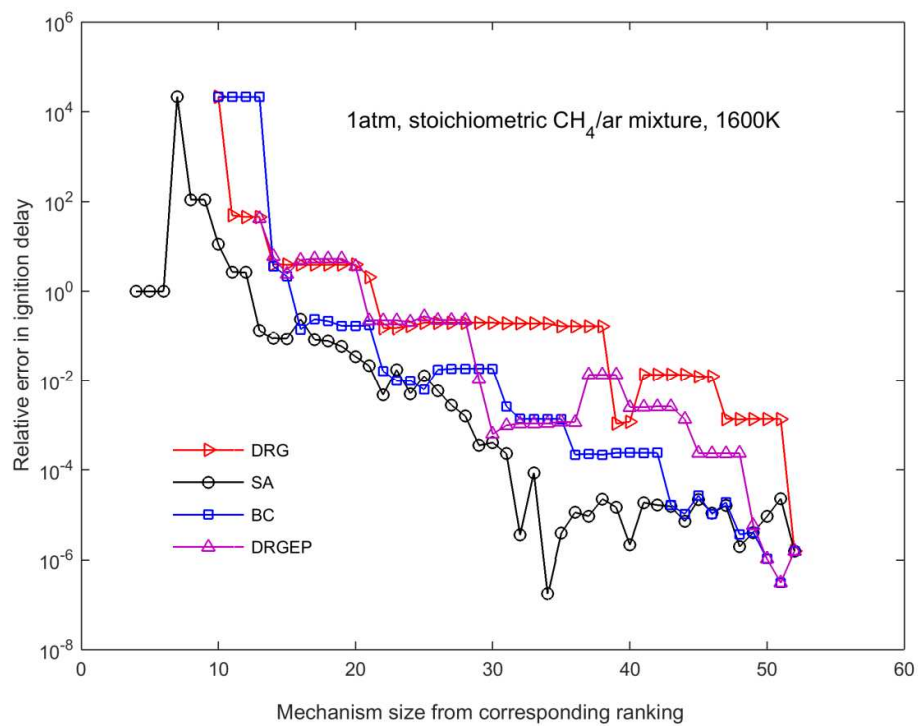


Fig. 4. Relative error in ignition delay time for skeletal mechanism of all sizes by BC, DRG, DRGEP and sensitivity analysis (SA) methods for the same condition as in Fig.1.

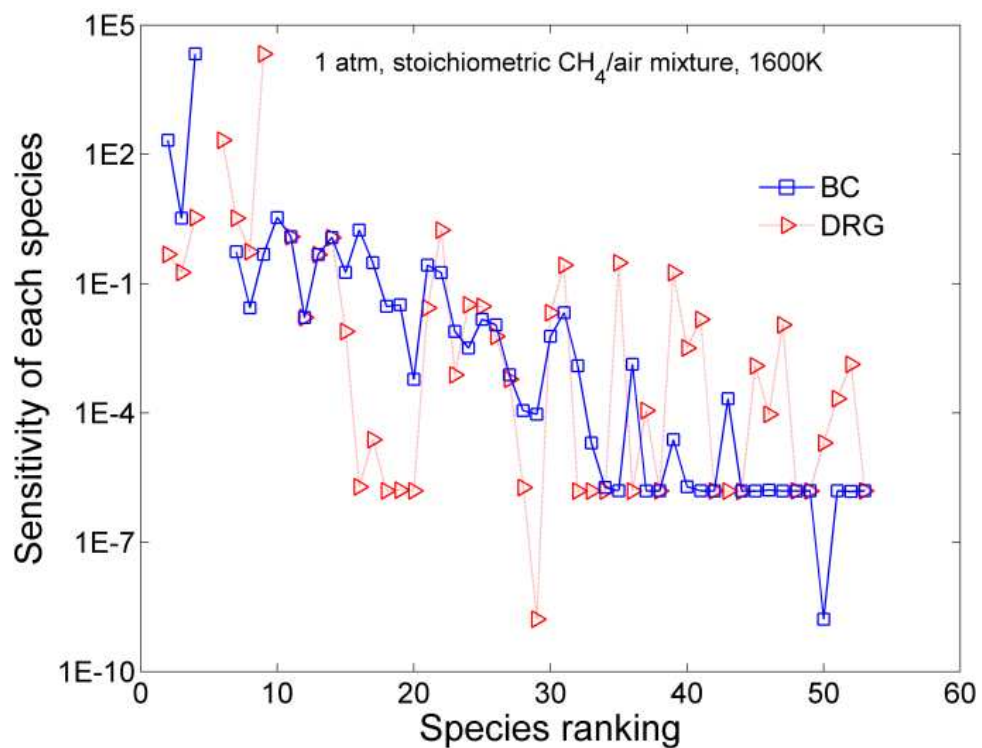


Fig 5. Sensitivity of each species for rankings by DRG and BC method as in Fig.1, showing better correspondence between BC and sensitivity.

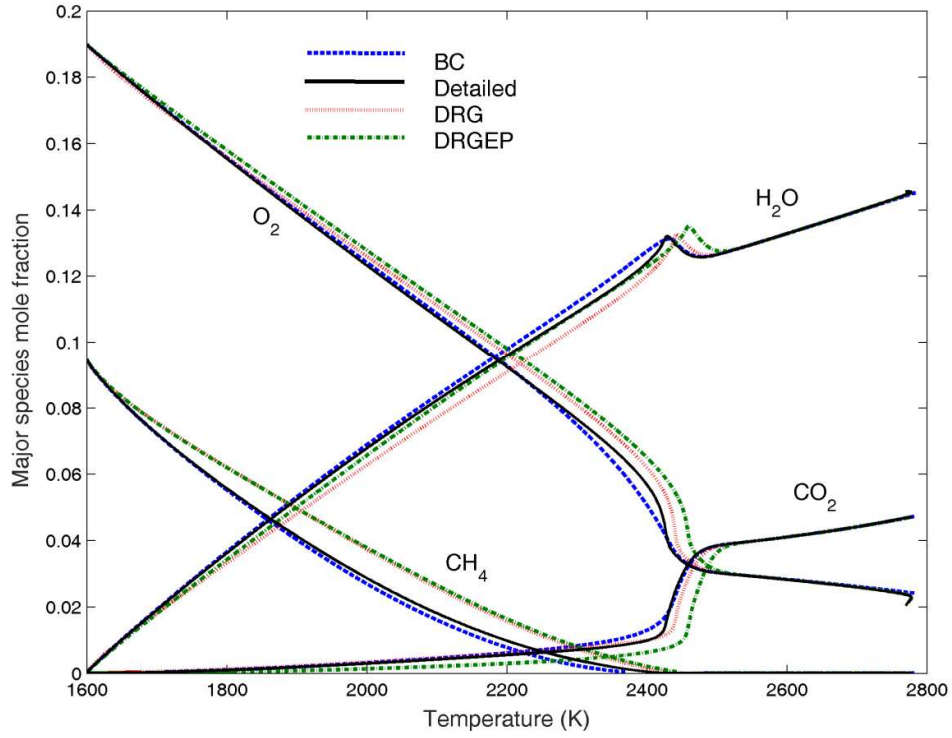


Fig. 6(a). Comparison of major species concentrations calculated using the detailed and skeletal mechanisms of 25 species by different methods for auto-ignition of stoichiometric CH₄-air under the condition of Fig. 1.

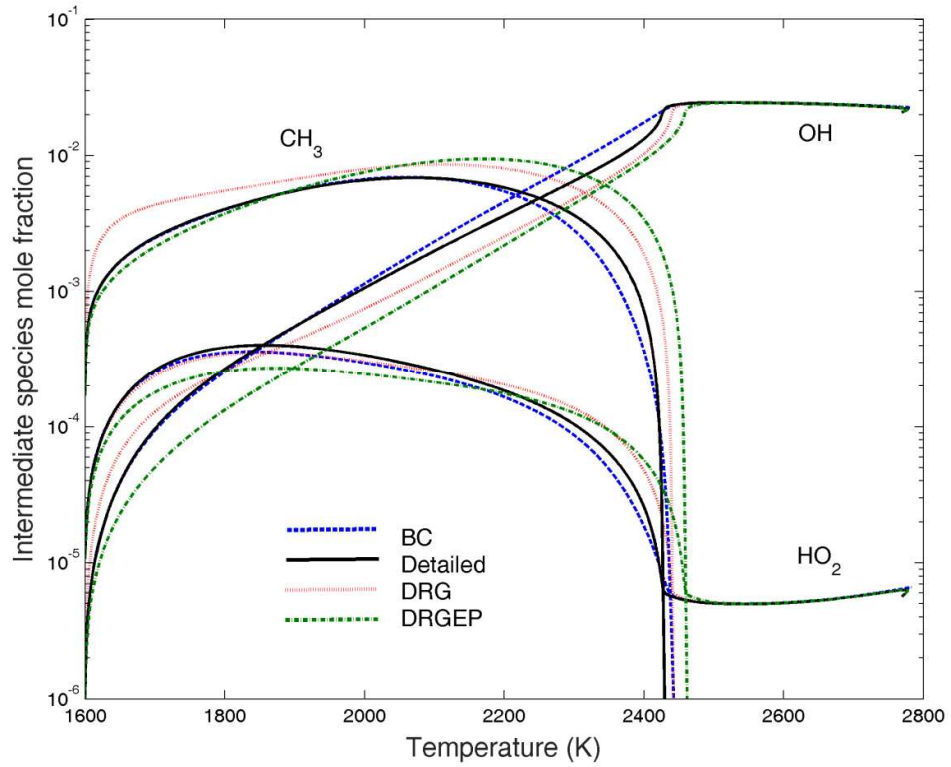


Fig. 6(b). Comparison of intermediate species concentrations calculated using the detailed and skeletal mechanisms of 25 species by different methods for auto-ignition of stoichiometric CH_4 -air under the condition of Fig. 1.

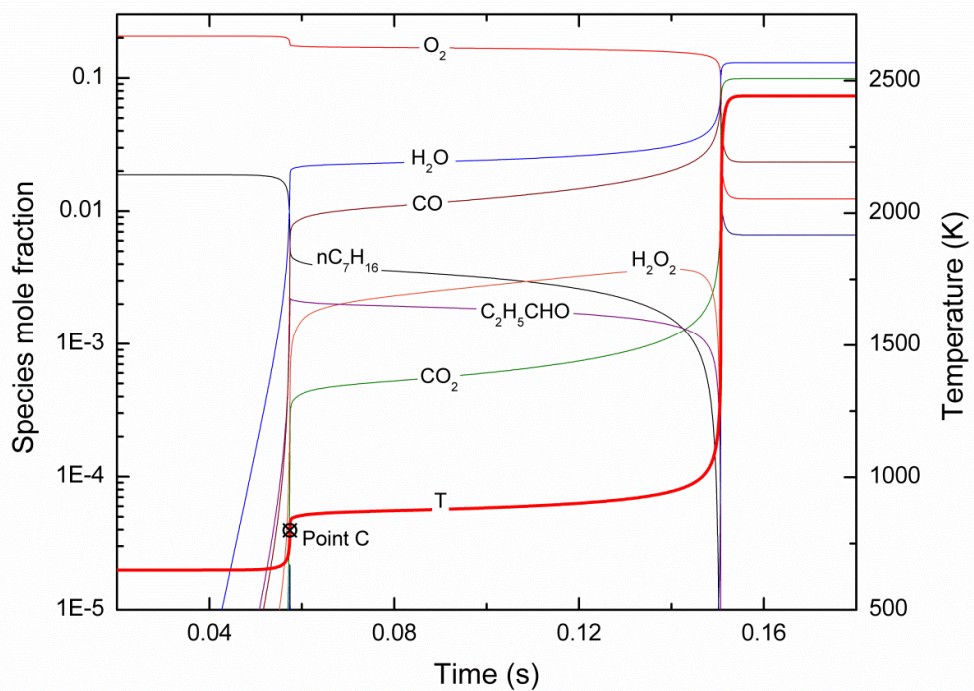


Fig. 7. Species history in a two-stage ignition process of stoichiometric n-heptane/air with initial temperature 650 K under 1 atm.

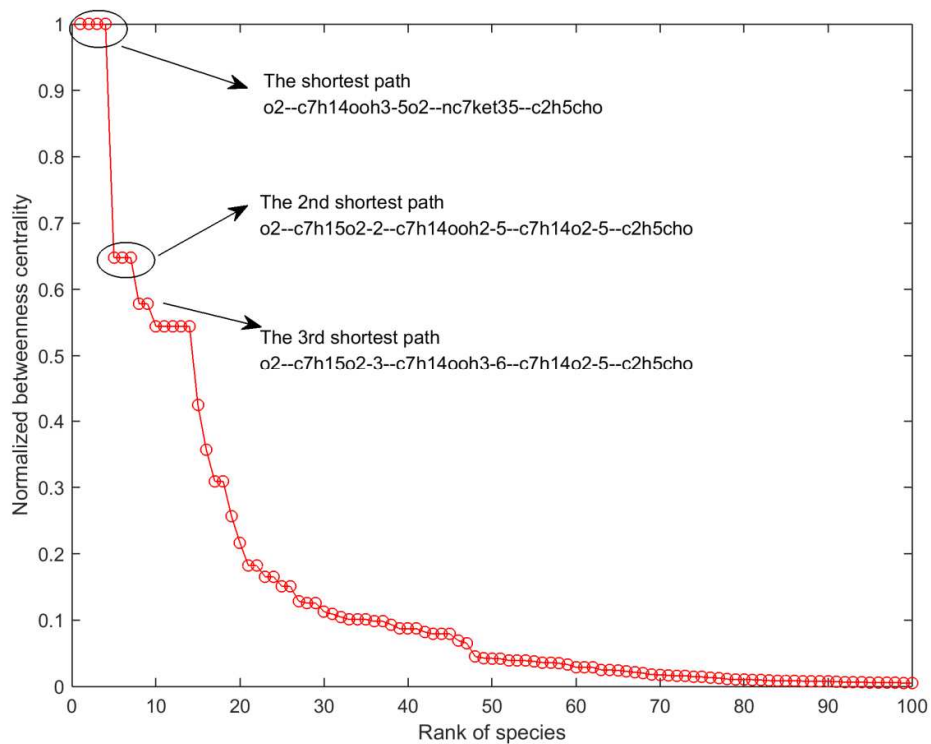


Fig. 8(a). Normalized BC and shortest paths connecting O_2 and C_2H_5CHO at Point C $t=5.74E-2s$ and $T=800.0K$ in Fig. 7.

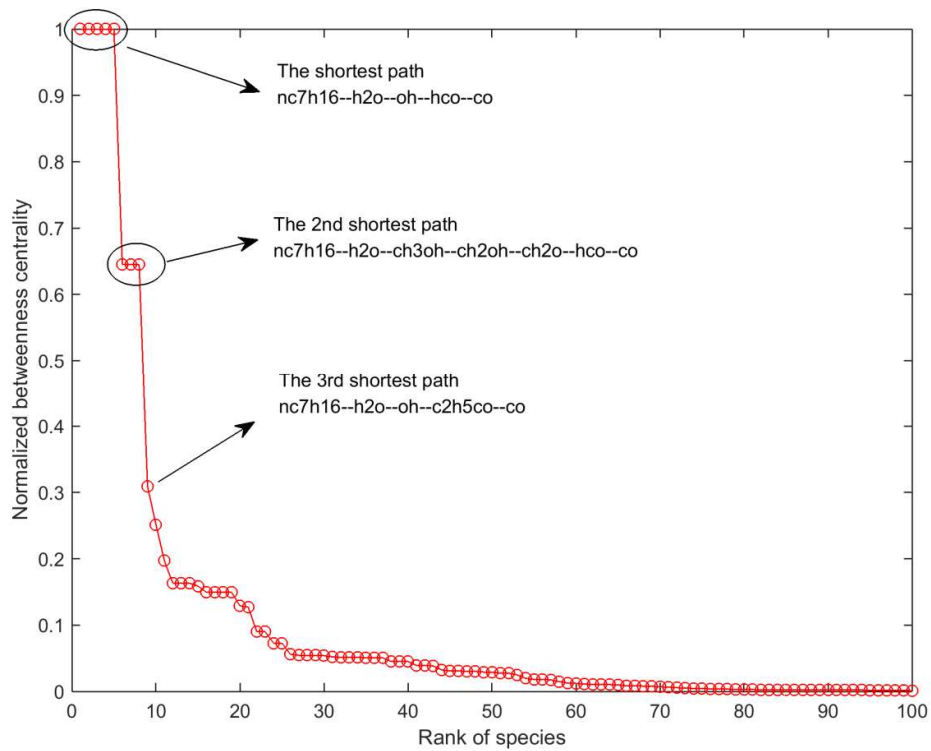


Fig. 8(b). Normalized BC and shortest paths connecting fuel and CO at Point C $t=5.74E-2s$ and $T=800.0K$ in Fig. 7.

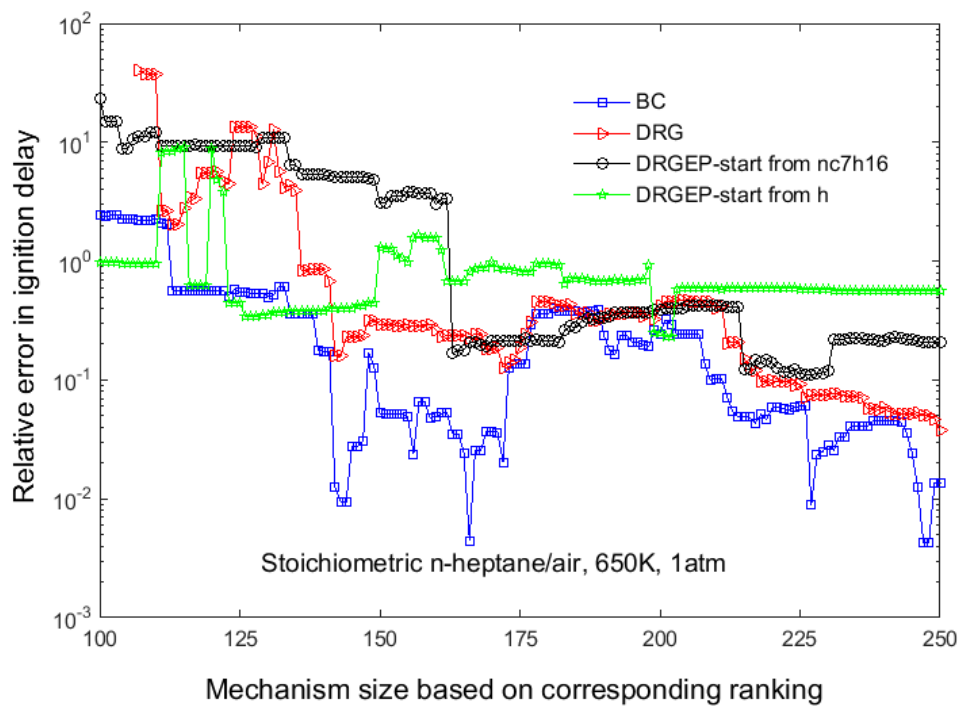


Fig. 9. Relative error in ignition delay time for skeletal mechanism of all sizes by BC, DRG, DRGEP (starting from H and nC_7H_{16}) methods for the same condition as in Fig.7.

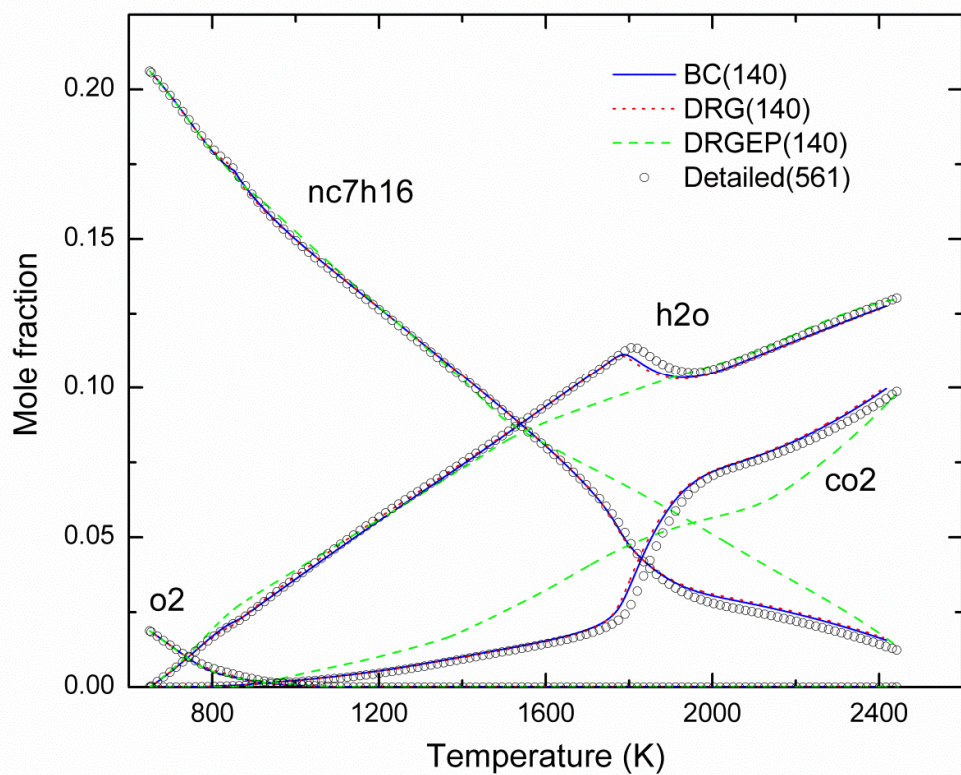


Fig. 10(a). Comparison of major species concentrations calculated using the detailed and skeletal mechanisms of 140 species by different methods for auto-ignition of stoichiometric n-heptane/air under the condition of Fig. 7.

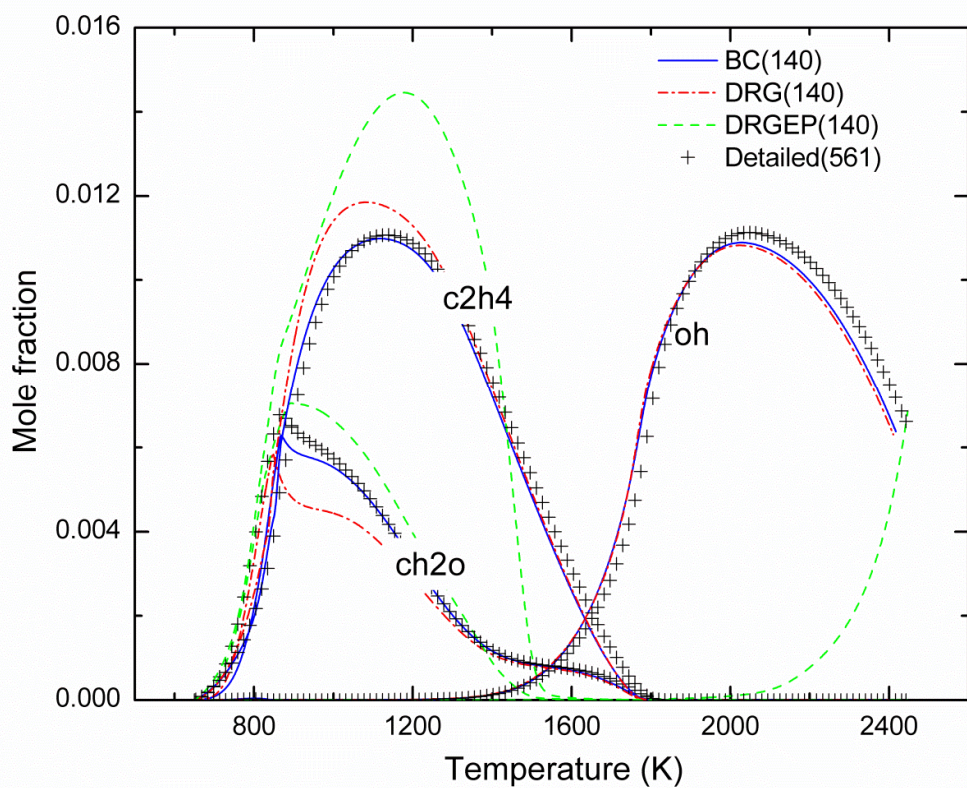


Fig. 10(b). Comparison of intermediate species concentrations calculated using the detailed and skeletal mechanisms of 140 species by different methods for auto-ignition of stoichiometric n-heptane/air under the condition of Fig. 7.

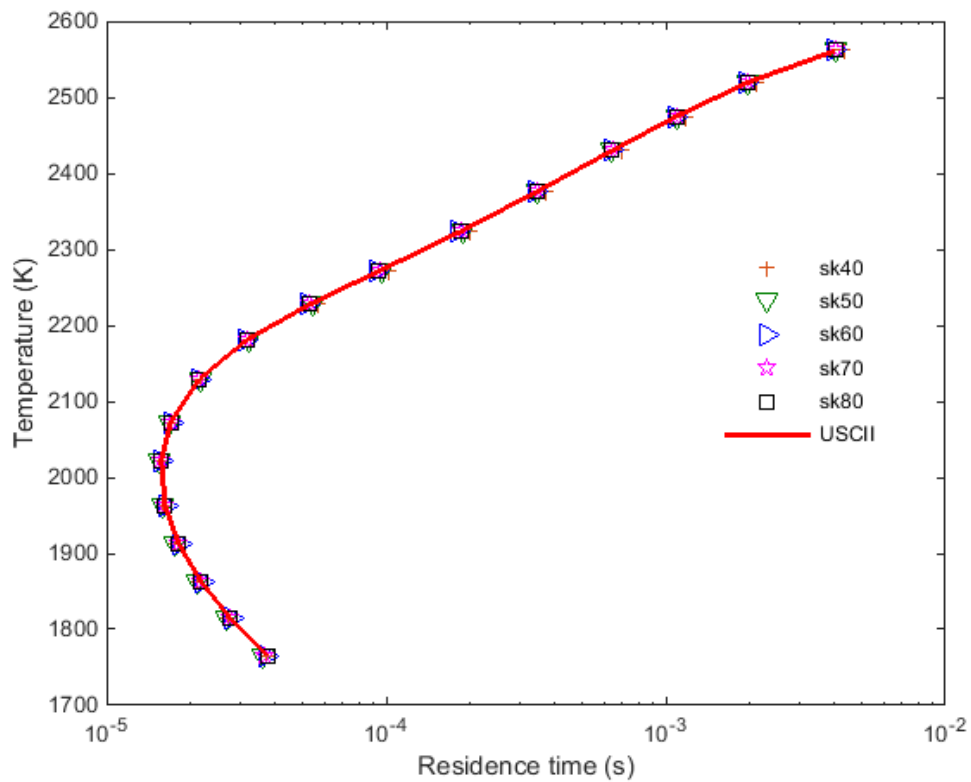


Fig. 11. PSR calculation by USC II and skeletal mechanism of different sizes by BC method for stoichiometric 1-butene/air mixture with initial temperature 1000K and pressure 1atm.

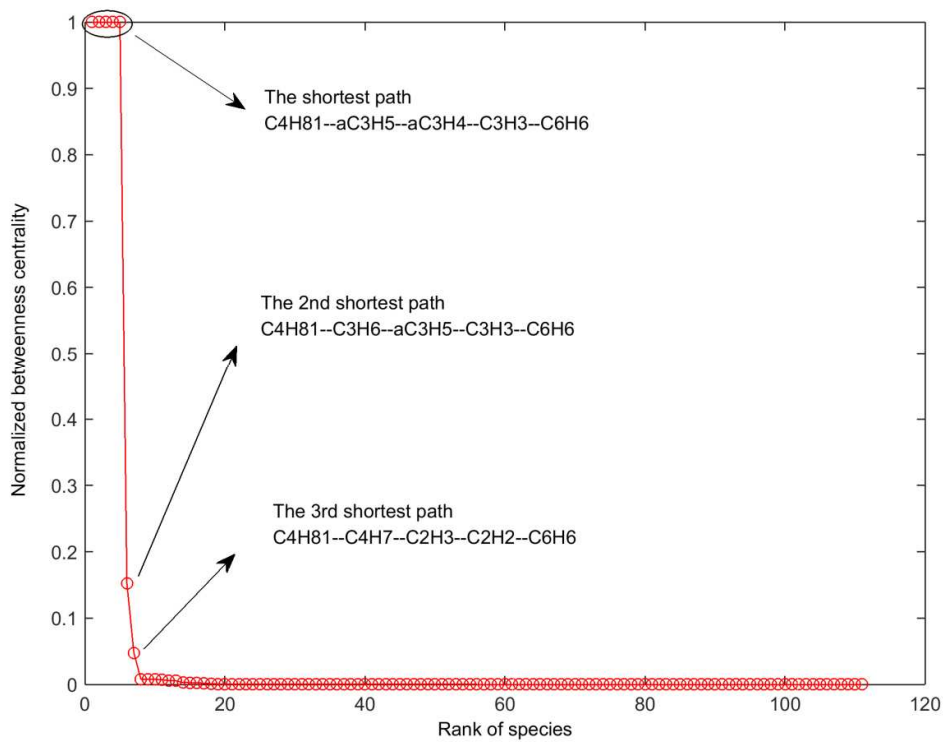


Fig. 12. Normalized BC and shortest paths connecting 1-butene and benzene at local point $T=2262.8\text{K}$ on the upper branch of Fig. 11.

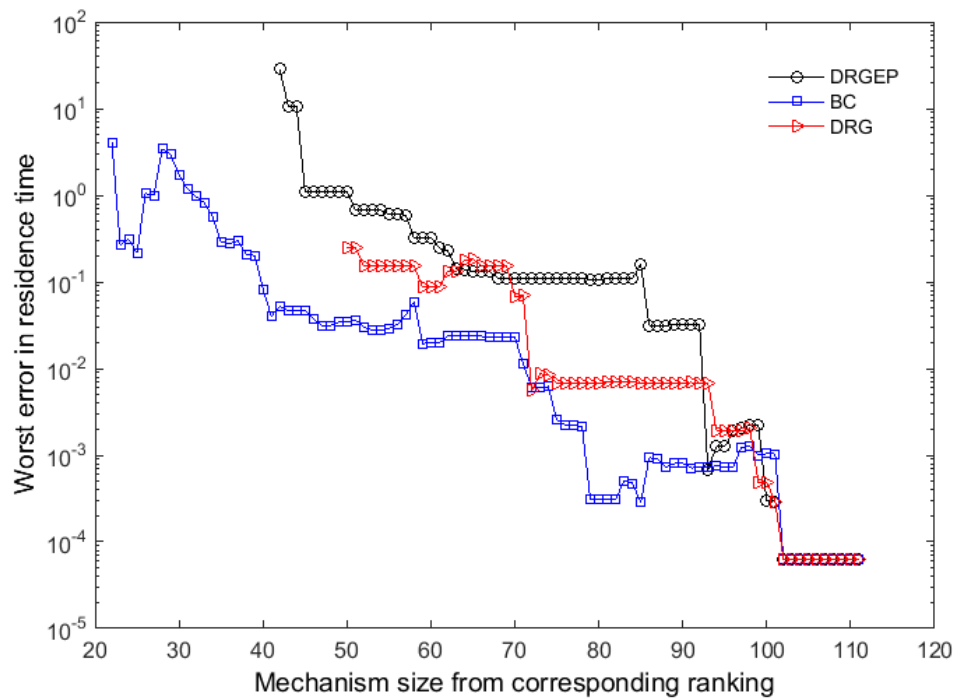


Fig. 13. Worst error in residence time predicted by skeletal mechanisms of all sizes by BC, DRG, DRGEP methods, for all sampling points along the S-curve of Fig. 11.