Complex Reaction Networks in High Temperature Hydrocarbon Chemistry

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID:	CP-ART-10-2014-004736.R1
Article Type:	Paper
Date Submitted by the Author:	17-Oct-2014
Complete List of Authors:	Mutlay, Ibrahim; Grafen Chemical Industries Co., Restrepo, Albeiro; Universidad de Antioquia UdeA,

Complex Reaction Networks in High Temperature Hydrocarbon Chemistry

İbrahim Mutlay*[1], Albeiro Restrepo[2]

[1]Grafen Chemical Industries Co., Ankara, Turkey

[2]Instituto de Quimica, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellin, Colombia

January 30, 2015

Abstract

Complex chemical reaction mechanisms of high temperature hydrocarbon decomposition are represented as networks and their underlying graph topology are analyzed as a dynamic system. As model reactants, 1,3-butadiene, acetylene, benzene, ethane, ethylene, methane, methyl isobutyl ketone (MIBK) and toluene are chosen in view of their importance for global environment, energy technologies as well as quantum chemical properties. Accurate kinetic mechanisms are computationally simulated and converted to bipartite graphs for incremental conversion steps of the main reactant. Topological analysis of the resulting temporal networks reveals novel features unknown to classical chemical kinetics theory. Time-dependent percolation behavior of the chemical reaction networks shows infinite order phase transition and a unique correlation between the percolation thresholds and electron distribution of the reactants. These observations are expected to find important applications in development of a new theoretical perspective to chemical reactions and technological processes e.g. inhibition of greenhouse gases, efficient utilization of fossil fuels, and large scale carbon nanomaterial production.

Keywords: Chemical reaction networks, graph theory, combustion, pyrolysis.
PACS: 82.20.-w, 82.40.-g, 82.20.Wt, 64.60.aq.

I. Introduction

Chemical reactions are one of the primitive processes ruling the entire physical world. From biological life, environmental cycles to industrial technologies, energy generation, the majority of the evolution of matter can be reduced to chemical reactions. The extremely complex and hierarchic nature of chemical reactions involving quantum mechanical scattering of molecules, statistical physics of particle ensembles and non-equilibrium thermodynamics attracts still many researchers suggesting vast variety of models. Quantum mechanical techniques [1, 2, 3] are able to explain chemical reactions in a great detail, however their application to real-life chemical kinetic mechanisms containing thousands of species and reactions is a computationally difficult (or unfeasible) task. Although system wide calculations, e.g. for mass flux by classical mass action kinetics are possible, they still require solution of large number of ordinary differential equations (ODEs).

Theory of complex networks brings a new understanding on the large scale organization of interacting entities. Abstraction of the systems as topological graphs by representing their units as vertices and interactions as edges, one acquires extensive information of the system organization by minimal computational load. Complex network approach has been successfully implemented for surprisingly different systems including social organizations, the World Wide Web, earthquakes, electrical power grids, protein interactions, citations to scientific articles and metabolic chemistry [4, 5, 6]. Chemical reactions with their native pathways have been always viewed as networks. In this context, a genuine research direction on kinetic stability of chemical reaction networks (CRNs) has been developed by Horn-Jackson [7], Clarke [8] and Feinberg [9]. This research provides invaluable results however they are limited to small mechanisms and traditional view of mass action law. Application of modern complex network theory on CRNs was first studied in the field of biochemistry, such as in [10, 11]. However, in spite of the subsequent success in biochemical CRNs [12], out of life sciences, there are still only a few works concerning for example organic chemistry [13, 14, 15] and the Diels-Alder reaction [16].

In this work, we study complex network features of the CRNs emerging from high temperature reactions of hydrocarbons. This reaction class is im-

*Corresponding Author. E-mail: ibrahimmutlay@gmail.com, Phone: +903123948109, Fax: +903123945514

mensely important not only because of technological and industrial processes which rely on fossil fuels, but also because of atmospheric and environmental phenomena, especially global warming, which is governed by hydrocarbon chemistry. For example, combustion of hydrocarbon fuels is the universal energy source of the modern world from power plants to internal combustion engine and pyrolytic reactions of hydrocarbons are the backbone of petroleum industry by converting oil to large spectrum of chemicals [17, 18]. In order to efficiently control this reaction class and reduce undesirable by-products, vast variety of experimental and computational techniques have been using since early 20^{th} century. The standard experiments gas phase probe sampling of flames and shock tube techniques provide extensive kinetic data which are tabulated as mechanisms and libraries in literature [19, 20]. Computational approaches focus on electronic structure calculations of molecules to be used in the Transition State Theory (TST) [19]. Frenklach pioneers the usage of Kinetic Monte Carlo methods to simulate discrete reaction steps [21] and even extend to the growth of solid carbon structures from soot [22] to graphene [23]. In order to construct large reaction mechanisms artificially, automated reaction mechanism generators are developed which are expected to eliminate bias, errors and laborious preparation procedures of manual mechanism building [24]. Reaction mechanisms of this study are also constructed by this approach using the RMG (Reaction Mechanism Generator) software [25]. Reaction mechanisms under controlled conditions (e.g. temperature, pressure, mole fractions) are computationally generated and converted to bipartite graphs (for details see Method in Supplementary Information). In Section 2, graph topology of CRNs is introduced and chemical insights are elaborated for the time-dependent[1] evolution of graphs, centrality notion and graph motifs. In Section 3, we discuss percolation of CRNs by means of connected components, largest eigenvalue and average path length to understand efficiency and robustness of chemical reactions. An analogy between geometrical percolation threshold and chemical reactivity is suggested. In order to estimate chemical reactivity from electronic properties of molecules, we employ the Quantum Theory of Atoms in Molecules (QTAIM) [26]. The electron density and its Laplacian, $L(r) = -\nabla^2\rho(r)$ are the target subjects of study in QTAIM [27, 28]. $L(r)$ measures local concentration or depletion of the electron density [29]. $L(r)$ is seen to act as the local balance between kinetic and potential energy densities within a molecule; or simply as a local "quantum pressure" [27]. Local minima and maxima of the Laplacian is termed critical points (LCPs). An outstanding correspondence between LCPs with the idealized electron pairs of the Lewis model has also been shown [30]. Using this analogy, we find a correlation between LCPs of various hydrocarbon reactants and percolation threshold of their CRNs. Basic concepts and foundations of the graph theory are presented in Supplementary Information. Reader is highly recommended to view that text before proceeding to the next sections.

II. Topology of CRNs

Chemical reactions are transformations of molecules under the rigorously defined algebraic rules of stoichiometry. Our reaction system has m species M_i participating in r reversible reactions R_j with stoichiometric coefficients $s_{i,j}$:

$$\sum_{i=1,j=1}^{m,r} s_{i,j}^{-} M_i \overset{R_j}{\rightleftharpoons} \sum_{i=1,j=1}^{m,r} s_{i,j}^{+} M_i \tag{1}$$

If one assumes the mass action kinetics, the time evolution of the system is given by differential equations of the form:

$$\frac{d\mathbf{M}}{dt} = \mathbf{S} \cdot \nu(x) \tag{2}$$

where $\mathbf{M}_i(t)$ is species concentration and $\nu(x)$ is flux vector. $\mathbf{S}$ is stoichiometric matrix of entries $s_{i,j}$ corresponding to stoichiometric coefficient of the molecule M_i (ith row) at the reaction R_j (jth column). In the traditional notation of chemical reactions (i.e. Eq. (1)) $s_{i,j}^{-}$, $s_{i,j}^{+} > 0$, however stoichiometric matrix definition requires that $\mathbf{S}_{i,j} > 0$ ($\mathbf{S}_{i,j} < 0$) for produced (consumed) molecules while $\mathbf{S}_{i,j} = 0$ for absent species. A simple example can be found at Fig. 1a showing small portion of a giant reaction mechanism along with its stoichiometric matrix (Fig. 1b).

[1] We use the terms of conversion and time interchangeably. See Supplementary Information for the details.

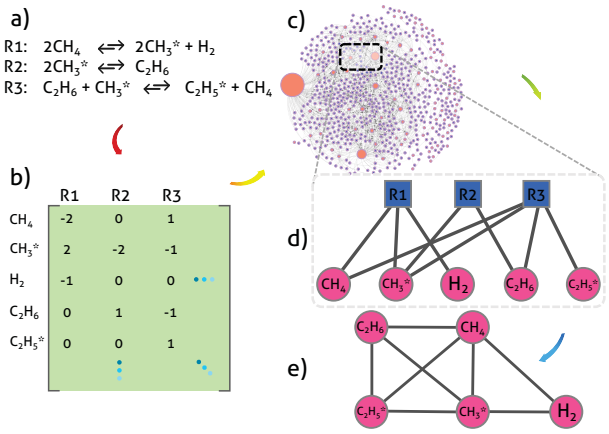


Figure 1: Fundamental definitions for bipartite representation of CRNs. a) initial reactions of CH_4 pyrolysis as a symbolic reaction mechanism. The species with asterisk are radicals. b) stoichiometric matrix of the reactions given in a). c) full bipartite CRN of CH_4 pyrolysis at 0.99 conversion. d) unweighted and undirected simple bipartite representation of CRNs which is in this case a subgraph sampled from c) and also is the exact bipartite graph of a) and b). e) unipartite species projection of the bipartite graph of d).

Stoichiometric matrix **S** which is an inherent incidence matrix of directed hypergraphs lies at the heart of topology of chemical reaction networks. Directed hypergraphs are the intuitive network representation of chemical reactions by their multiple edges between two vertices. Unfortunately, considering lack of current understanding on hypergraphs, one would not expect much functionality. In order to handle chemical reaction networks efficiently we choose bipartite graphs because they are among the most favored and widely employed representations for reaction networks [31, 32]. Bipartite graphs delineate interactions of two separate vertex sets, i.e. participation of "agent" vertices in "event" vertices which are allowed to connect only opposite vertex sets not themselves. Bipartite graphs are actually König representation of hypergraphs [31] and there is a direct conversion between them by using stoichiometric matrix as a bipartite adjacency matrix, **B** in the dimensions of agents $\times$ events [33]:

$$\mathbf{B}_{k,l} = \begin{cases} b_{k,l} \in \mathbb{R} \text{ if agent participates in event} \\ 0 \quad \text{if not} \end{cases}$$

After this transformation, agents are species, events are reactions, and $b_{k,l}$ is the stoichiometric coefficients obviously (see Fig. 1d). The numerical value and sign of $b_{k,l}$ reflect weight and direction of edges in bipartite CRNs, respectively. Frequently, two types of the nodes are called top and bottom vertices and can be projected into two unipartite graphs, i.e. "species graph" and "reaction graph" respectively which can be analyzed just like ordinary graphs. Fig. 1e illustrates this operation; if the vertices of agents (events) in the original bipartite graph have links to an opposite vertex, then these agent (event) vertices connect inter se.

In this study, we work on undirected and unweighted simple bipartite graphs to minimize computational workload. This is a reasonable approach since the reactions here are reversible which do not need any direction in the graph edges. Unfortunately, we are also limited to available theoretical tools of the literature covering mainly simple graphs (for definition of the "simple graphs" see Supplementary Information). In the preparation of CRNs, all the nonzero entries in the stoichiometric matrices are reduced to 1 and edge directions are omitted. We also divide bipartite CRNs into their projections and focus only on species graphs because reaction projection affords excessive number of edges and uncertain vertex identity. However this issue is due to our computational configurations and should not be generalized. Holme and Huss already suggested that species projection is the most convenient graph to interpret the parent CRN [34, 35].

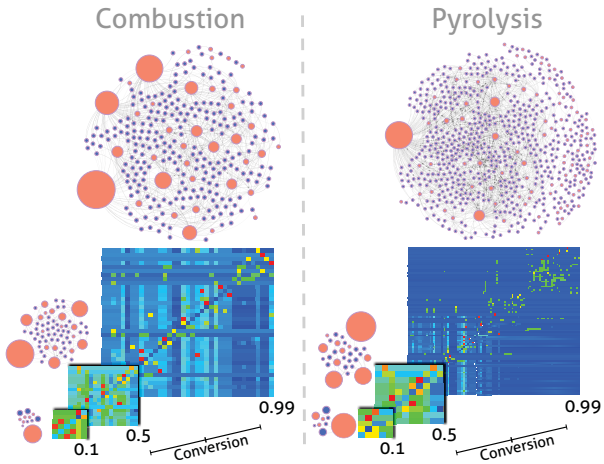


Figure 2: How CRNs grow in time? For CH_4 combustion and pyrolysis, time evolution (in reactant conversion dimension) of bipartite CRN graphs and adjacency matrices of their corresponding unipartite species graph projections. In bipartite CRN graphs, species and reaction vertices are salmon and light blue colored respectively. We emphasize species vertices with a nonlinear size proportionality to the degree (i.e. the total edge number connected to the vertex, for the definition see Supplementary Information). For the sake of visually enhanced representation of adjacency matrices we employ weighted unipartite species graphs only here: pixels of adjacency matrix plots are colored by continuously scaling weights of edges from blue $\equiv 0$ to red $\equiv$ max(weight).

In Fig. 2, we depict time evolution of CRNs with the help of visualizations of bipartite CRNs and their accompanying unipartite species graph adjacency matrices. After total consumption of the reactant, we have an enormous kinetic mechanism with several "social" species of high degree represented by the larger vertices. In that example of CH_4, combustion mechanism has many high degree species while in pyrolysis there is only one prominent species. Note that, dominant blue color of the adjacency matrices in 0.99 conversion implies that those species graphs are sparse in accordance with the suggestion of [36] and sparsity increases as the reaction progresses.

A graph topology is defined by its metrics or measures which quantify the network structure (these concepts are extensively explained in Supplementary Information). We monitor the major graph metrics over the course of conversion for all the reactants in both combustion and pyrolysis regimes (see Fig. 3). **Graph diameter**, i.e. the maximum shortest path between vertices, determines the navigation or transport efficiency of networks [4, 5]. An intriguing point is that diameters of biochemical reaction networks in 43 different organisms were found to be universally $\sim$3 [10, 38] smaller than our hydrocarbon mechanisms which could hint a difference in efficiency. In contrast to common temporal social networks [39], diameter of the species graphs does not fall, moreover increases slowly in time. The tendency of graphs to form cliques/clusters is measured by **the clustering coefficient**, especially transitivity i.e. the number ratio of triangles to triples [5, 38]:

$$T_{3,c} = \frac{3 \times n_{\triangle}}{n_{\diagup}} \tag{3}$$

Triangle and triple are vertex sets of three at which vertices of the former connect simply like a triangle and the latter connect by ensuring at least one of the vertices connects to the rest [4, 5]. Since transitivity is not valid for bipartite graphs that cannot accommodate triangles Lind et al. preferred to use the rectangular vertex sets of four [40]:

$$\langle C_4 \rangle = \frac{C_4(i)}{n_N} = \frac{n_{i,\boxdot}}{n_{\boxdot} n_N} \tag{4}$$

where $n_{i,\boxdot}$ is the number of squares of vertex i and $n_{\boxdot}$ is the total number of possible squares of i. Averaging over the nodes gives the global square clustering coefficient for bipartite graphs. We compare $T_{3,c}$ of projected species graphs with $\langle C_4 \rangle$ of bipartite CRNs. Tracking through the conversion reveals declining $\langle C_4 \rangle$ unlike constant $T_{3,c}$ as can be seen in Fig. 3. It is worth noting that there is little data on temporal profile of the clustering coefficients, for example [41] confirming our results. **Assortative mixing** is graph theory counterpart of the "like dissolves like" principle of chemistry and **assortativity** characterizes to what extent similar vertices connect. In assortative mixing, vertices select the other vertices to connect according to the similarity of their individual network measures [5]. It is possible to evaluate assortative mixing using any of the graph measures however in scientific literature the prevalent method is to calculate whether the

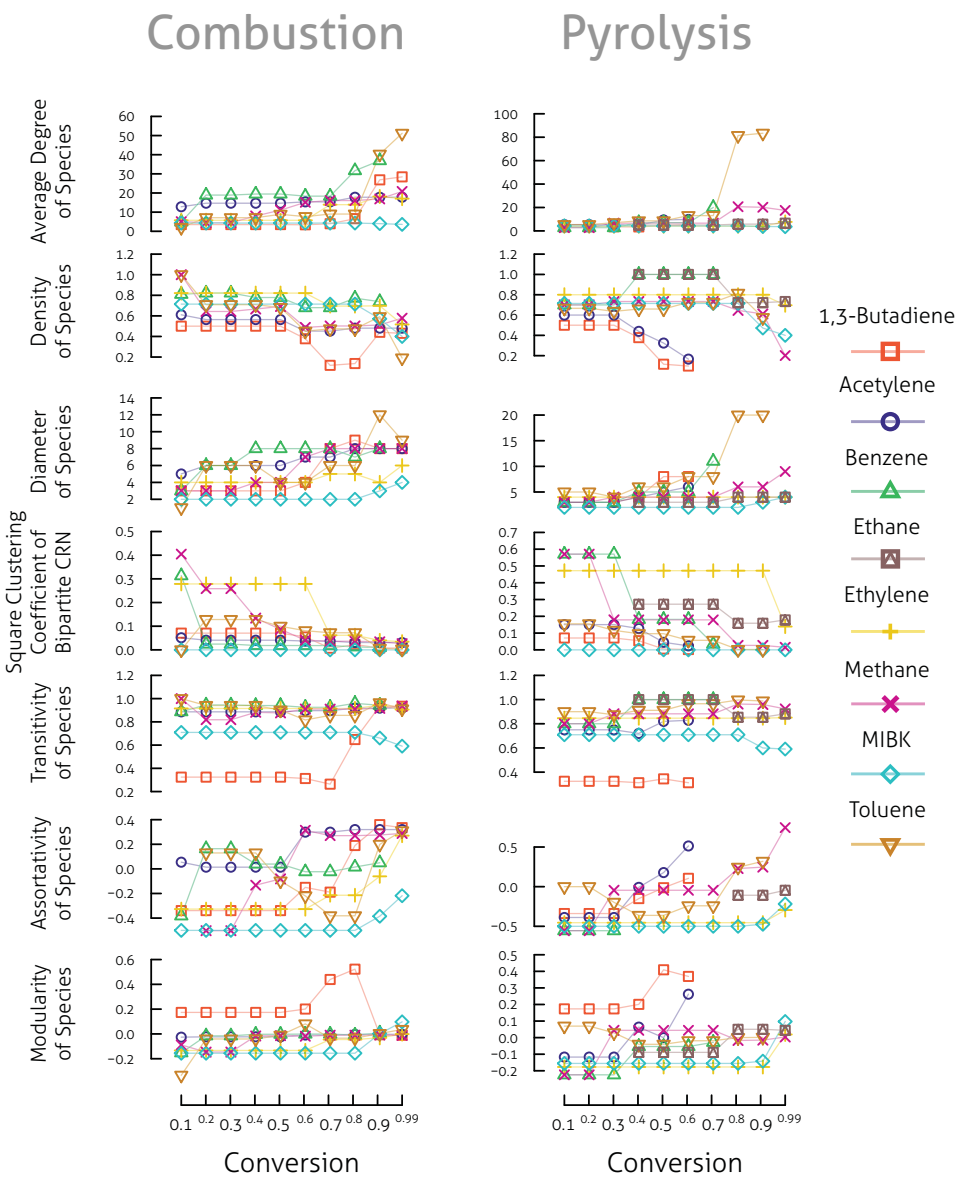


Figure 3: Time evolution of major graph metrics. All the curves are plotted for projected unipartite species graphs except square clustering coefficient which is for the original bipartite CRNs. Average vertex degree $\langle k \rangle$ defined on all vertex space is the simplest graph metric. Graph density is the ratio of total number to maximum number of edges. The density drops steadily by the conversion which is another reflection of sparsity increase [37]. Assortativity coefficient is calculated based on degree. See the main text and Supplementary Information for the detailed definitions and discussions.

vertices select other vertices to connect by considering their degree (i.e. degree based assortativity). In this study we also follow the same method to assess the time-dependent degree assortativity of reaction networks and the results are presented in Fig. 3 exhibiting surprising behavior: species graphs start as disassortative (negative values) but then turn out to be assortative (positive values) in late stages of the conversion, especially in the case of pyrolysis. This is highly distinctive in view of [42] reporting positive assortativity only for social networks and negative assortativity for biochemical networks. **Modularity** defines how communities, i.e. more densely connected vertex clusters are well separated [43]. Identification of communities requiring for further modularity assessment is accomplished by the WalkTrap algorithm [44] in this study. We find a generally increasing trend in Fig. 3 and the magnitude of modularity are significantly high, achieving nearly practical maximum [43] for 1,3-butadiene indicating strong communities in CRNs. We think that communities and modular topology correspond to reaction subclasses and have important applications in "lumping" techniques. Lumping is a model reduction method grouping compounds according to their physical/chemical properties in the form of pseudospecies which is able to save large amount of computational load [45]. In this regard, communities might help to find novel lumping techniques.

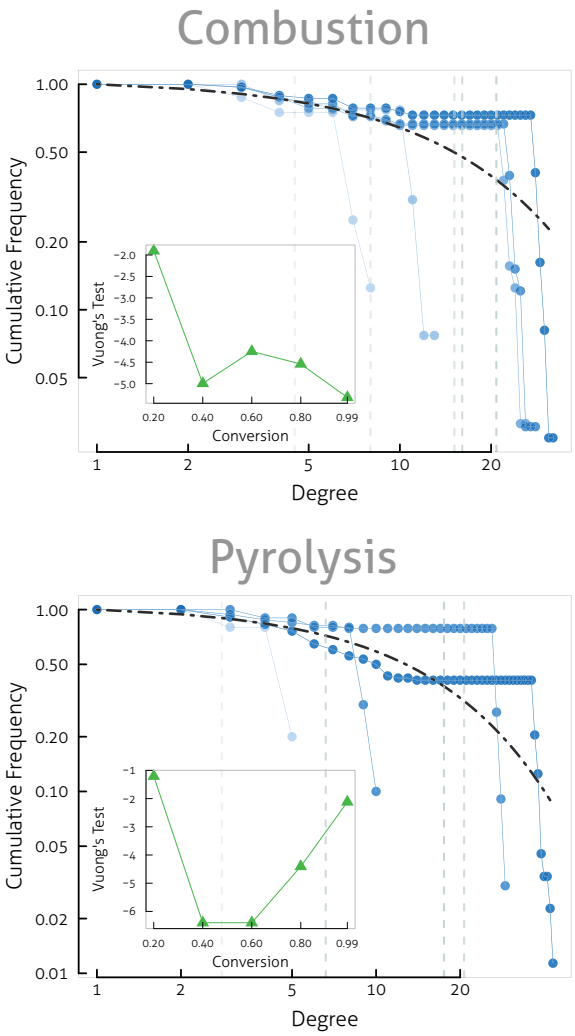


Figure 4: Log-log plot of the complementary cumulative distribution function (CCDF*) of the vertex degree distribution for species graphs of CH_4 combustion and pyrolysis along with their power-law analysis. Filled blue circles are the empirical degree distributions and vertical dashed lines are the average degrees. Both of them are shown with an increasing opacity order of the line color from the conversion of 0.2 to 0.99 in the interval of 0.2. Dash-point black curves are theoretical fit by the exponential degree distribution model to the final network (of the conversion 0.99). The insets are the time evolution of Vuong's Test results. *Note: CCDF curve at $x = k$ point represents the quantity of vertices with a degree of k or more.

Vertex degree distribution is one of the most informative topological features for complex networks (see Supplementary Information for underlying mathematics and the models). Commonly it is plotted as a cumulative distribution of vertex degrees since this offers visually more robust graphics [46]. What makes degree distribution so special is that its form is unique for each kind of network topology. In this context, it practically illuminates most of the network connectivity. Cumulative degree distributions of species graphs of our reaction mechanisms throughout the conversion window are presented in Fig. 4 and Fig. S1 (see Supplementary Information). After a qualitative examination of the shapes of the plots, one can easily deduct from their nonlinear decay profile in log-log axes that the best fitting theoretical probability distribution is the exponential distribution [47]. In order to prove this, we estimate the degree distribution fit of the species graphs for 0.99 conversion using the exponential distribution model [46]:

$$p(k) = \tau e^{\tau k_{min}} \left(e^{-\tau k}\right) \quad (5)$$

where τ is the rate parameter and k_{min} is the vertex degree cut-off value. Indeed, the fitting is acceptable for most of the reaction systems. However we need solid proof, hence we invoke Vuong's Test which is a model selection test based on likelihood ratio and the Kullback-Leibler criteria. The Vuong's Test of exponential distribution here is implemented against classical power-law model $p(k) \sim k^{-\gamma}$, γ is the scaling parameter. Sign and magnitude of the Vuong's Test results define which model is more suitable for the data. Negative and significantly low values of Vuong's Test are the final proof of exponential degree distribution in our species graphs. In fact, exponential degree distribution in temporal networks is considerably prevalent observation in the literature [48, 49] even including protein [50] and online social networks [51]. The rationale is that in non-equilibrium and/or time-varying networks, preferential attachment which is the major process leading to power-law degree distribution is hindered [49, 52, 53]. In the absence of scale-free topology, i.e. the graphs with power-law degree distribution, we seek for the other common topology; the small-worlds which are characterized by small diameter and high clustering coefficient. For the sake of objective evaluation, we follow the method of [54] suggesting a small-worldness index. Because of the unsuccessful calculations in some cases, we do not give the whole data however, generally speaking, the index indicates that all the species graphs here are small-worlds and their small-worldness increase by the conversion.

Centrality of a node is the quantitative measure of its influence and importance in the network. Intuition tells us that the more central a vertex is, the more prominent and functionally crucial a species is. This clearly serves as a solution to one of the most pursued questions in the theory of chemical kinetics, i.e. finding key species and pathways of reaction systems. Fig. 5 and Fig. S2 (see Supplementary Information) display bubble charts of fundamental centrality measures against gradually progressing conversion. **Authority (hub)** is an eigenvalue centrality which is determined by HITS algorithm calculating eigenvectors of the following matrices [55]:

$$\mathbf{M}_{Authority} = \mathbf{A}^T\mathbf{A} \quad \text{and} \quad \mathbf{M}_{Hub} = \mathbf{A}\mathbf{A}^T \quad (6)$$

where $\mathbf{A}$ is the adjacency matrix of, in our case, species graphs. Authority in conjunction with hub centrality, not surprisingly, delineates the vertices governing information transfer along the network. In order to see the performance of predictive capabilities of centrality measures, we compare them with chemical kinetics literature. The highest authority centrality in total burning of methane is detected for C_2HO* ketenyl radical which is also known as a critical species in this reaction [56]. The ruling role of C_2HO* in hydrocarbon combustion was already reported [57] particularly for acetylene proving our results. For combustion of all the reactants here except 1,3-butadiene, we observe that oxygen bounded primitive species are authorities. Monitoring authority in pyrolysis reactions unveils the importance of atomic hydrogen and methyl radical CH_3* mentioned in the literature [58, 59]. Richter-Howard molecular growth approach to polyaromatic hydrocarbon (PAH) and soot formation [60] is valid for our benzene and toluene pyrolysis. Betweenness and closeness both of which are geometric centralities based on shortest paths are analyzed next. **Betweenness** is an indicator of junction vertices conveying main information "traffic" in the network. This centrality therefore has

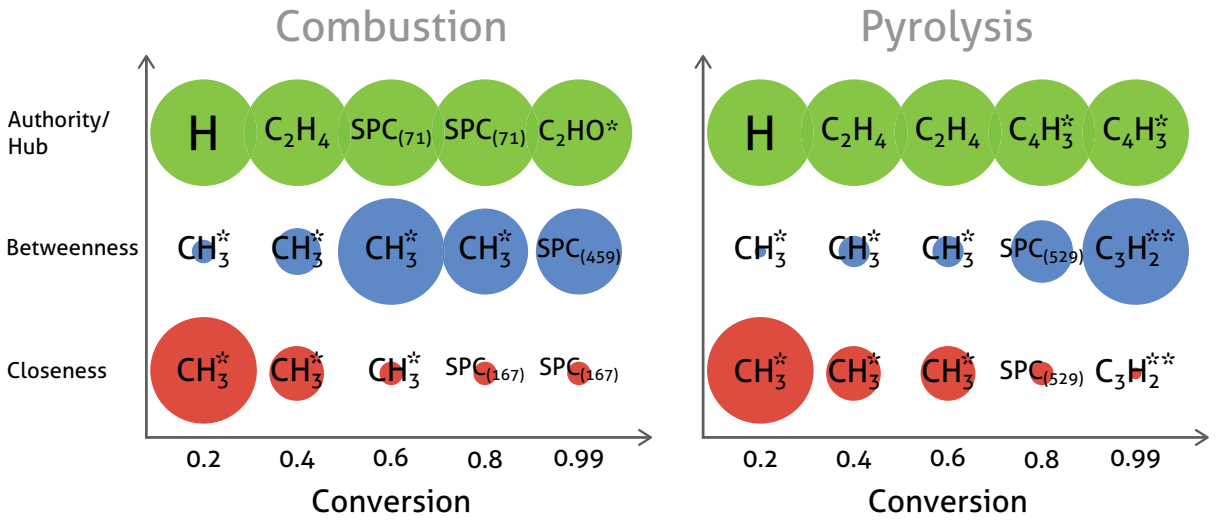


Figure 5: Time-dependent change of authority & hub, betweenness and closeness centralities for CH_4 combustion and pyrolysis. Labels in each circle correspond to the species with maximum centrality. Circle size is proportional to magnitude of the maximum centrality but not scaled. Since our networks are undirected, authority and hub centralities are exactly same and shown together. The species with asterisk symbol are radicals. $SPC_{(\#)}$ accounts for the uncertain isomers and long species names each of whose InChI codes are given below:
$SPC_{(71)}$: InChI=1/C2H2O/c1-2-3/h1-2H/mult3
$SPC_{(167)}$: InChI=1/C2H3O/c1-2-3/h1H3
$SPC_{(459)}$: InChI=1/C3H4O/c1-2-3-4/h1-2H2/mult3
$SPC_{(529)}$: InChI=1/C4H4/c1-3-4-2/h1,3H,2H2.

very important functions in graphs, particularly matter/information flow capacity and robustness. Betweenness centrality of a vertex i mathematically is as follows [6]:

$$\mathbb{C}_i^B = \sum_{j \in V} \sum_{\substack{k \in V \\ k \neq j}} \frac{\zeta_{jk}(i)}{\zeta_{jk}} \tag{7}$$

where $\zeta_{jk}(i)$ is the number of shortest paths starting from vertex j, crossing i and going to k and ζ_{jk} is the total number of the possible paths between $j-k$. CH_3* radical is a species of high betweenness in variety of mechanisms (see Fig. 5 and Fig. S2, Supplementary Information). We find acrolein ($SPC_{(459)}$) and C_3H_2** diradical as the key species of methane chemistry referred also by the literature [61, 62] but not studied in depth. Acrolein has a strong betweenness centrality in acetylene combustion. Maximum betweenness molecule hexadiyne ($SPC_{(298)}$) of 1,3-butadiene combustion at 0.99 conversion is an interesting molecule as a benzene precursor [63]. Ethylene chemistry is notably governed by radicals in view of betweenness, e.g. simple vinyl radical (C_2H_3*) form of ethylene (C_2H_4). Betweenness analysis of the aromatic reactants gives inherent species: C_6H_5* phenyl radical for benzene and C_7H_7* benzyl radical for toluene as well as O_2 for their combustion. Betweenness centrality, unlike authority/hub is variable, roughly increasing by conversion. **Closeness**, in a sense, is the center point of the network because it tells us how far away from other graph regions a vertex is. Calculation of closeness for a vertex i is quite straightforward [6]:

$$\mathbb{C}_i^C = \frac{N-1}{\sum_j d_{ij}} \tag{8}$$

i.e. inverse average of geodesic distances from i to j. The species with maximum closeness are almost same as that of the maximum betweenness.

However the time-dependent profile of closeness appears to diminish in contrast to betweenness. Investigation of the brief results above and detailed kinetic mechanisms which are provided in Supplementary Information reveals a practical, universal principle: a species is said to be central to its species graph if it participates in any of the first three reactions of its kinetic mechanism.

Network motifs is a notion of great interest believed to be abstract building blocks of networks each of which has specific real-life function. Motifs which are isomorphic subgraphs with high cardinality [6] have been a matter of debate to understand their correspondence relations with physical and especially biochemical phenomena, such as biochemical reactions and neural patterns in brain [38]. Although motifs do not have to be specific reactions or chemical entities in CRNs [64], it is tempting to explore possible relationships to the hierarchy of kinetic mechanisms. Toward that goal, we extract size-3 and size-4 motifs from species graphs of all reactants throughout the conversion range (see Fig. 6). Remarkably, we reveal that the most abundant motifs, M2, M5 and M8, denote the unipartite species graph projections of the following reaction system:

$$\begin{aligned} X_i &\rightleftharpoons X_{i+1} && \text{(I)} \\ X_{i+1} + X_{i+2} &\rightleftharpoons X_{i+3} && \text{(II)} \\ X_{i+4} + X_{i+5} &\rightleftharpoons X_{i+6} + X_{i+7} && \text{(III)} \end{aligned}$$

Bearing the rules of bipartite CRN building and unipartite projection in mind, reaction (II) produces M2 and combination of (I)-(II) M5. The last reaction (III) gives a closed triplet, i.e. M8. Our number distribution of motifs is unique in comparison to [65] acquiring completely different findings for several social, spatial and biological networks. The listed mechanism of (I), (II) and (III) is obviously among the fundamental reactions of chemical kinetics, hence this fact would have significant outcomes in the context of chemistry which will be discussed in Section 3.

III. Percolation in CRNs

Percolation, the evolution in properties of a thermodynamic system by the emergence of a large spanning cluster of the system units, is among the universal phase transitions. The recent decades have witnessed comprehensive research on this model finding a great variety of applications in phase transitions of superconductivity, magnetization, electrical/thermal conductivity of granular materials and composites, colloidal formations, epidemics, fracture, and even in socioeconomic organizations [66, 67, 68, 69]. Theoretical background of percolation phenomenon is inherently compatible with graphs. Connectivity of networks, under gradual node/edge addition or removal, experiences an abrupt transition at some point corresponding to the verge of appearance or disappearance of the giant connected component which is the largest cluster spreading on majority of the network. This critical onset is called percolation threshold at which global transport capacity and robustness, in other words, function and survival state of network is switched. Since a network growing (or shrinking by disconnections and attacks) beyond the percolation threshold would excel in its function (or disrupt and be annihilated), percolation plays a dramatic role for real-life CRNs by characterizing their chemical yield and efficiency.

In order to interpret percolation phenomenon for CRNs successfully, one would need to understand the analogy between abstract CRNs and real chemical reaction systems. Chemical reactions are the alteration of electronic and nuclear wavefunctions via different processes, particularly atomic redistribution of molecules. In another view adopted by biochemistry society referring holistic principles, CRNs are supposed to be mass flux networks. Clearly, both of the approaches point out that practical productivity of a chemical reaction mechanism relies on (information or mass) transport capability of its CRN and of course its percolation transition.

Fundamental laws governing the percolation differentiate for static and dynamic networks. Static networks are in the class of equilibrium percolation obeying finite size scaling and power law $X \sim (p - p_c)^{\omega}$ at $N \to \infty$ where X is any global network feature, p is the fraction of nodes/edges added or removed and p_c is the percolation threshold. Unfortunately, powerful techniques of finite size scaling fail for dynamic networks, e.g. CRNs and there is little work on non-equilibrium percolation. Callaway et al. [48] studied dynamic networks by adding a single vertex at each time step and

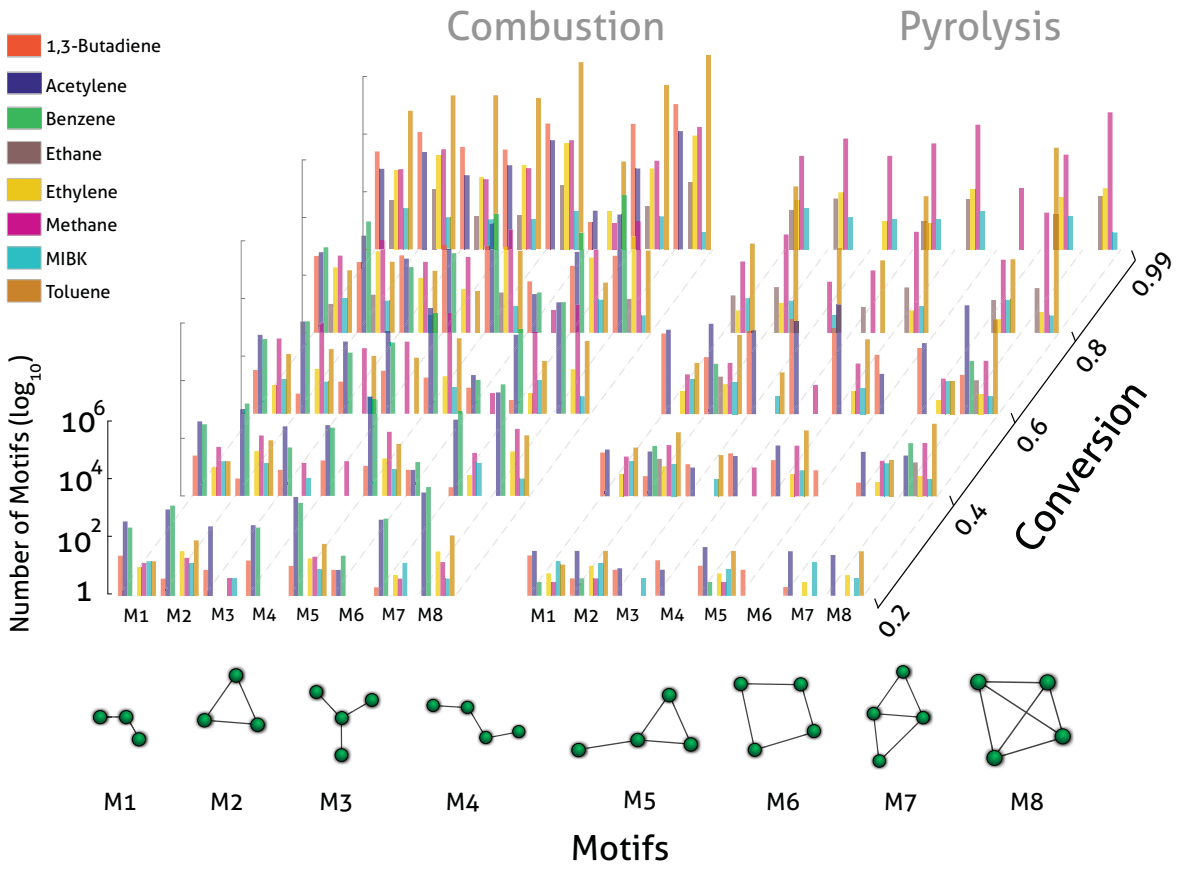


Figure 6: Time-dependent number distribution of network motifs. M1-2 are size-3 and M3-8 are size-4 motifs. y-axis is scaled to $\log_{10}$ but the actual numbers of motifs are given. Note that, the distribution for pyrolysis shows much less data because of the failed simulations of reaction mechanism generations and not related to any known physical process.

found that non-equilibrium percolation is in the class of infinite order phase transition. They also successfully approximated the percolative growth of dynamic graph by the relation below:

$$S(\delta) \sim e^{\alpha(\delta-\delta_c)^{-\beta}} \tag{9}$$

as $\delta \to \delta_c$ where δ is edge formation probability, δ_c percolation threshold, α and β are critical exponents [48]. Percolation in growing protein networks evolved by mutations [70] and synthetically generated time-varying graphs [71] were also identified as infinite order.

We study percolation analysis of our species graphs projected from bipartite CRNs in Fig. 7 displaying peculiar percolation behavior. The giant connected component, GCC emerges suddenly just at the beginning of conversion. This is surprising because in classical percolation theory there is no GCC before percolation threshold. Moreover, size of the GCC exactly equals to network size throughout the whole range of conversion in contrast to classical GCC that can cover a graph only partially. Nevertheless, percolation trends of our species graphs are quite similar to the results of the conventional systems and experiences a critical point as a growth burst in network size. Although we are not able to simulate conversion (ξ) at $\xi \to 0$ limit, this unusual percolation is prevalent for all the reactants in the range of $\xi > 0.01$. In fact, CRNs are not the only systems showing such an always percolated phase

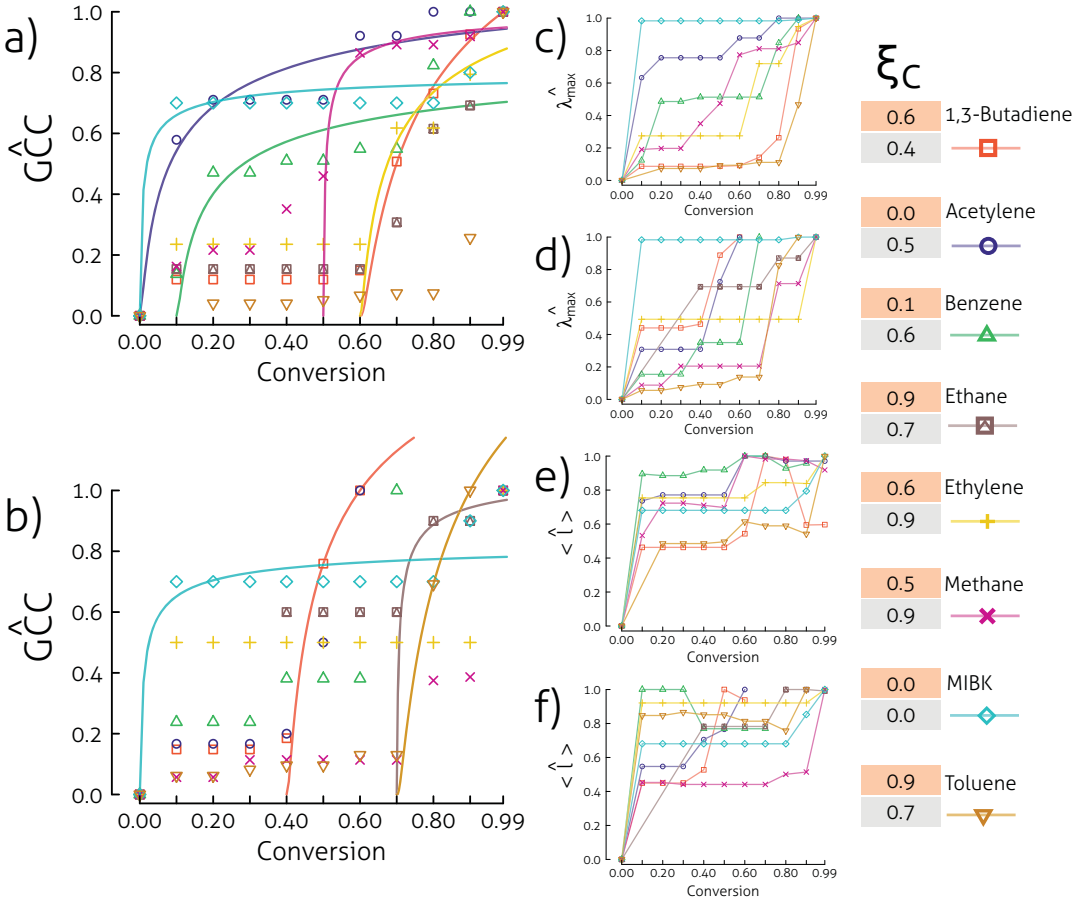


Figure 7: Percolation analysis of species graphs of combustion and pyrolysis CRNs. The plots are GCC size (a)-(b) , largest eigenvalue (c)-(d) and average path length (e)-(f) against conversion. Data points are empirical data whilst solid curves are theoretical fit based on Eq. (10). First plots (a, c, e) are for combustion, second plots (b, d, f) are for pyrolysis. Intensity of the percolative graph measures (y-axes) for all graphics are normalized over their maximum values. In order to estimate α and ξ_c of Eq. (10) which is needed to plot $S(\xi)$ time-dependent GCC size curves of (a)-(b), we fit Eq. (10) by a linearized log-log model for each reactant separately. Right column ξ_c is the percolation thresholds for combustion (orange box, upper) and pyrolysis (gray box, lower) of the corresponding reactants.

[72]. Dynamically growing protein networks of [70], self-similar networks [73] and graphs from Achlioptas process [74] all exhibit a percolation threshold at the stage zero of their evolution. Similarly, GCCs in the size of their parent networks require special conditions to grow. For instance, in particle contact networks of granular materials, gravity ensures close packing of the particles leading to a GCC whose size equals to full network [75]. Facebook, thanks to its extremely connected social web, is another interesting example at which GCC to network size ratio achieves 0.999 [76].

Let us recall our main goal in this paper, we try to comprehend how and why percolation of CRNs belongs to that strange phase transition class. The graph motif distribution we have discussed in the previous Section comes into play at this point. As one can see from the tabulated reaction mechanisms in Supplementary Information, almost all of the reactions are in the forms of $X + Y \rightleftharpoons Z$ or $X + Y \rightleftharpoons W + Z$ which also correspond to the primary motifs M2 and M8, respectively. Furthermore, in those mechanisms, every new reaction added in each step contains species from previous steps. Intuitively, this process is supposed to be as follows: by increasing ξ, in each step, species graph grows by M2 and M8 motifs without any isolated vertex or cluster considering absence of quasi-reactions e.g. $0 \rightleftharpoons X$, and the fact that each new reaction is tied to previous ones by their reactants or products. Thus, the reaction system is thoroughly evolved as a single unique cluster, i.e. the GCC. The growth mechanism of CRNs depicted above clarifies the strange percolation trend observed here. Among all other complex networks, it is fairly natural to assume that such kind of percolation behavior pertains to CRNs which are able to ensure a strongly connected growth based on non-isolated triplets.

Dynamic percolation of CRNs is investigated via infinite order phase transition model of [48]. We utilize asymptotic form of Eq. (9) by a slight abuse of notation:

$$S(\xi) \sim e^{\alpha/\sqrt{(\xi - \xi_c)}} \tag{10}$$

to fit the time-dependent empirical GCC size data (see Fig. 7a-b). In Eq. (10), $S(\xi)$ is the size of giant connected component, ξ is the conversion, ξ_c is the conversion at the percolation threshold and α is the critical exponent. Agreement between the fit and data points indicates that infinite order phase transition is also valid for CRNs. Note that, although original GCC size data differentiates in a broad range, after normalization we observe a consistent and similar trend for all the reactants which is discriminated only by percolation thresholds.

Strong change of spreading capacity and navigation efficiency related to global connectivity occur during percolation. In the case of positive direction for system evolution, fine synchronization of spreading and navigation at percolation provides the best way of boosting network functionality. However, for the situations where critical point coherence is lost, the network is adversely affected by percolation transition. Spreading capacity of network, like mass flux of reactions or information transport in our CRNs, is controlled by the largest eigenvalue λ_{max} [77]. λ_{max} undergoes a transition, p_c^λ leading to rise in spreading capacity that is similar to usual percolation. In this regard, relative position of p_c^λ is the critical parameter affecting spreading efficiency. $p_c^\lambda < p_c$ is highly unlikely, however if $p_c^\lambda \gg p_c$ then one can have a GCC but in fact the network is not capable of spreading. Hence in order to compare spreading efficiency of CRNs we propose the measure in terms of conversion below:

$$\epsilon_\lambda = \frac{\xi_c}{\xi_c^\lambda} \tag{11}$$

In the case of navigation efficiency ϵ_l, we simply calculate inverse of the average shortest path length:

$$\epsilon_l = \frac{1}{\langle l \rangle} \tag{12}$$

Spontaneous increase of $\langle l \rangle$ in the vicinity of percolation threshold hindering network transport and navigability has been already reported by Lopez et al. [78]. Recall that CRNs are abstract grids for wave function routing on molecules, navigation efficiency is a vital factor determining the yield and productivity of chemical reactions [79]. We apply spreading and navigation efficiency notions to our CRNs. Almost all the reaction systems studied here are highly efficient spreaders with $\epsilon_\lambda \approx 1$ (see Fig. 7). On the other hand, average path length does not have a clearly tractable trend. Navigation efficiency of ethane as well as of acetylene, ethylene, and methane, rapidly drops after percolation threshold. Even worse, navigation efficiency for the majority

of reactants does not recover to its original high level in contrast to [78] in which navigability falls only at percolation threshold locally.

Percolation threshold of physical systems e.g. electrically conductive composites [80] or granular packings [81] is known to have a strong correlation to the geometry of system units. This universal correlation is

$$p_c \propto \frac{1}{\phi} \tag{13}$$

where ϕ is the aspect ratio, the size ratio of two major topological axes of the geometrical objects. Eq. (13) is practically very important by enabling estimation of percolation threshold via only simple geometrical data of the system building-blocks. Although it is proposed geometrically, in order to benefit from its utility, we try to extend it to our CRNs by establishing an analogy. The major feature governing aspect ratio influence on percolation is the coordination number or vertex degree in graph terms. More anisotropic granular particles are able to build system-wide contact paths using less particles. Thus, at percolation, there are smaller number of particles (smaller coordination number) giving reduced percolation threshold. In chemical reactions, system unit is molecules and connectivity is expected as the function of reactivity. However one can clearly discern from basic chemistry that geometrical aspect ratio of molecules or any other pure geometrical characteristic is not sufficient to estimate their reactivity. Quantum mechanical electronic and nuclear structure of molecules is the actual parameter defining chemical reactivity which is analyzed here in the framework of QTAIM. The critical points of the Laplacian of the electron densities (the LCPs) have been successfully used in a wide variety of chemical reactivity situations [28], including the correct prediction of potential sites for nucleophilic attack [82], the structures of hydrogen bonded complexes [83], the preferred attacking sites for electrophilic substitution in substituted aromatic compounds [84], and the relative susceptibility of activated double bonds to Michael addition [85]. The collected evidence of the ability of the Laplacian of the electron density to correctly predict chemical interactions lead to the formulation of the Laplacian complementary principle [86], which states that "A maximum in the valence shell of an atom in one molecule combines with a minimum in $L(r)$ in the valence shell of an atom in another molecule". In this work, we use the properties of the LCPs of the electron density (those for which $\vec{\nabla} L(r) = 0$) to gain information about chemical reactivity. In this context, we suggest that LCPs for the reacting molecules is the equivalent of geometrical aspect ratio for the particles in granular percolation. To be precise, the aspect ratio analogy is the ratio of local Laplacian maxima to minima:

$$\frac{|L(r)^{max}|}{|L(r)^{min}|} \equiv \phi \tag{14}$$

We illustrate this analogy in Fig. 8 indicating a correlation highly similar to the literature [80, 87]. Simple inverse proportionality like Eq. (13) can be clearly seen. Following [80], we attempt to fit the empirical data using Pade approximation but the results are not convincing. Nevertheless, it supports the evidence of Eq. 13 and 14 which, to our knowledge, is the first demonstration of a relationship between the percolation of abstract CRNs and real-life ab initio principles. Percolation threshold for CRNs, like global reactivity measures, could be used to evaluate activity of overall reaction mechanisms. Besides, mathematical approach of this study is far more unsophisticated than common reactivity estimation techniques which will help potential technological applications.

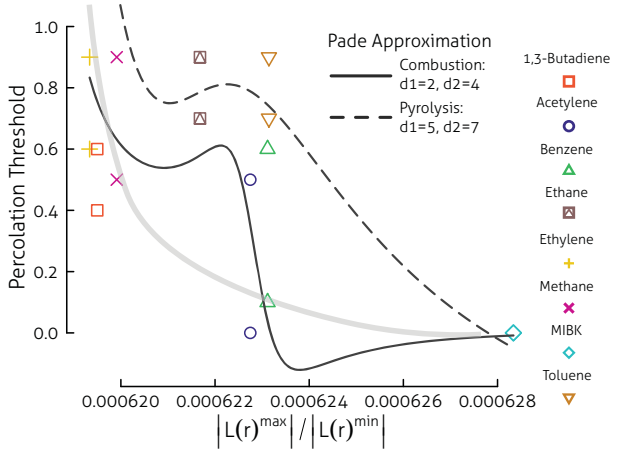


Figure 8: The correlation between percolation threshold and the Laplacian of electron density of the reactant molecules. We test traditional Pade approximation fit of aspect ratio to percolation threshold observed in granular matter [80]. Light gray line is guide to the eye to represent inverse proportionality. In order to find which data point is for combustion or pyrolysis reader can check the percolation thresholds. For calculation details and fitting parameters of Pade approximants see Supplementary Information.

IV. Discussion

Abstraction of CRNs and analysis of their topology via methods of complex network theory provides rich data and could change the view on chemical kinetics which has been already proven in biochemistry. However a more complementary discussion to assert the graph theoretical results here needs to evoke classical chemical kinetics. Our reaction mechanisms generated artificially by the RMG calculating also their mass fluxes by Eq. (2) and Arrhenius parameters are given in Supplementary Information. Determination of the key elementary steps of a reaction mechanism is a significant part of kinetics research. Considering practical yield, we assume the chemical reactions of maximum flux are the major pathways which are listed in Table 1 for our results. We compare those reactions with the species found in centrality analysis (see Fig. 5 and Fig. S2). It is observed that central species predicted by graph theory are in excellent agreement with the major pathways. These topologically identified species are same as the ones participating the key elementary reactions which is proof of how the centrality is a powerful indicator. Some of the difference in results can be easily attributed to the further reaction steps for example C_2H_4 and C_4H_3* of methane pyrolysis can be emanated from CH_3* by consecutive reactions. Similarly conversion between C_2H_3O and C_2HO* of acetylene is possible with a H_2 reaction. In an attempt to support our findings, we also test the Transition State Theory calculations (see Fig. S3, Supplementary Information) however extremely high temperature, unimolecular and free radical dominant reaction mechanisms in our study prevent us to use that technique which could fail for these situations.

In addition to the theoretical importance, high temperature hydrocarbon chemistry has substantial implications for the Earth and humankind. 70% of global electricity generation still relies on thermal decomposition of fossil fuels[2]. Furthermore, transportation, commodity chemicals and even many advanced materials are dependent on petroleum. Critical carbon allotropes fullerene [88], carbon nanotube [89] and graphene [90] as well as pyrolytic graphite [91] are produced by the high temperature reactions of hydrocarbons. Unfortunately, those chemical processes pose a great danger to the environment and human health. Atmospheric carbon black (or soot) and greenhouse gases which are the two major causes of the global warming, and PAHs which are highly mutagenic and carcinogenic to biological life originate from anthropogenic hydrocarbon decomposition [92, 93]. In this regard, design and control of hydrocarbon CRNs is a two sided problem: elimination of the hazardous effects require inhibition and disruption of the reaction networks whilst robust and efficient networks are vital for the processes of power generation or material production. Evolution of hydrocarbon reactants to PAHs and soot consists of nearly countless and complex pathways which is still a controversial research topic. The leading models in that field are proposed by Howard [60, 94, 95] and Frenklach [96, 97, 98] independently. According to the general picture developed in those models [60, 96] high temperature hydrocarbon reactions start with the addition of C_2 species, particularly acetylene C_2H_2,

[2]The World Bank and U.S. Energy Information Administration data.

Table 1: Chemical reactions with maximal flux generated in this study. If fluxes of the 1st and 2nd maximum flux reactions for a reactant are nearly equal, we present both of them. Pure H-OH reactions without carbon species are omitted. A, β and E_a are respectively the pre-exponential factor, the temperature exponent and the activation energy (kcal/mol) of the modified Arrhenius equation: $k = AT_{gas}^{\beta} exp\left(\frac{-E_a}{RT_{gas}}\right)$ with the gas constant R and the reaction temperature T_{gas}.

	Reactants	Reactions	A	β	E_a
Combustion	C_4H_6	$C_3H_3^* + CH_3^* \rightleftharpoons C_4H_6$	7.230×10^{13}	0.00	0.00
	C_2H_2	$C_2H_2 + OH \rightleftharpoons C_2H_3O$	1.520×10^{8}	1.70	1.00
	C_6H_6	$C_6H_5^* + O_2 \rightleftharpoons C_6H_5O_2^*$	6.030×10^{12}	0.00	0.32
		$C_6H_5^* + H \rightleftharpoons C_6H_6$	2.200×10^{14}	0.00	0.00
	C_2H_4	$C_2H_2 + H \rightleftharpoons C_2H_3^*$	5.500×10^{12}	0.00	2.42
	CH_4	$CH_4 + H \rightleftharpoons CH_3^* + H_2$	1.224×10^{8}	1.87	10.59
		$C_2H_2 + H \rightleftharpoons C_2H_3^*$	5.500×10^{12}	0.00	2.42
	$C_6H_{12}O$	$CO^{**} + CH_3^* \rightleftharpoons C_2H_3O^*$	3.060×10^{6}	1.89	4.82
	C_7H_8	$C_6H_5^* + O_2 \rightleftharpoons C_7H_5O_2^*$	6.030×10^{12}	0.00	0.32
Pyrolysis	C_4H_6	$C_3H_3^* + CH_3^* \rightleftharpoons C_4H_6$	7.230×10^{13}	0.00	0.00
		$C_4H_6^{**} \rightleftharpoons C_4H_6$	1.132×10^{11}	1.00	9.50
	C_2H_2	$C_2H^* + H_2 \rightleftharpoons C_2H_2 + H$	1.080×10^{13}	0.00	2.17
		$C_2H_2 + C_2H^* \rightleftharpoons C_4H_3^*$	1.000×10^{13}	0.00	0.00
	C_6H_6	$C_6H_5^* + H \rightleftharpoons C_6H_6$	2.200×10^{14}	0.00	0.00
		$C_6H_5^* + C_6H_5^* \rightleftharpoons C_{12}H_{10}$	5.700×10^{12}	0.00	0.00
	C_2H_4	$C_2H_2 + H \rightleftharpoons C_2H_3^*$	5.500×10^{12}	0.00	2.42
	CH_4	$CH_4 + H \rightleftharpoons CH_3^* + H_2$	1.224×10^{8}	1.87	10.59
	$C_6H_{12}O$	$CO^{**} + CH_3^* \rightleftharpoons C_2H_3O^*$	3.060×10^{6}	1.89	4.82
	C_7H_8	$C_6H_5^* + CH_3^* \rightleftharpoons C_7H_8$	1.380×10^{13}	0.00	0.05

to the small molecules (C_1-C_4 species). First of the critical points in that reaction scheme is the formation of initial aromatic rings which is followed by the growth to PAHs via aliphatic growth by addition of C_2 species and/or direct aromatic growth. The final stage is the physical agglomeration of the giant PAHs and evolution of these PAH clusters to larger soot particles. Although there are various views on the formation of first aromatic rings, the following reactions are accepted as the key pathways [60, 96]:

$$C_2H_2 + C_2H_3^* \rightleftharpoons 1,3-C_4H_5^* \quad (1)$$
$$1,3-C_4H_5^* + C_2H_2 \rightleftharpoons C_6H_6 + H \quad (2)$$
$$n-C_4H_3^* + C_2H_2 \rightleftharpoons C_6H_5^* \quad (3)$$
$$C_3H_3^* + C_2H_2 \rightleftharpoons c-C_5H_5 \quad (4)$$

Since $C_4H_3^*$ and C_2H_2 are abundant species in our mechanisms it is thought that reaction (3) could be a crucial pathway in our conditions. Along with C_2H_2 we detect many C_2 species supporting C_2-based evolution theory. C_2s also play an essential role in PAH growth formulated by the HACA (H-abstraction-C_2H_2-addition) reactions [96]:

$$A_i + H \rightleftharpoons A_i^* + H_2 \quad (5)$$
$$A_i^* + C_2H_2 \rightleftharpoons PAHs \quad (6)$$

where A_i is an aromatic molecule and A_i^* is the aromatic radical of A_i. Prominence of this type of reactions prompts us to investigate the topological routes involving C_2 species in our CRNs. We measure the average shortest path between the main reactant and C_2 molecules in species graphs (see Fig. 9). Profiles and average levels of $d_{s_1 \rightarrow C_{2,i}}$ are highly consistent and similar for combustion and pyrolysis. $d_{s_1 \rightarrow C_{2,i}}$ levels of each reactant surprisingly follow a rule of hydrocarbon chemistry: for the order of reactants of alkanes, mono-olefins, di-olefins, benzenes and naphthalenes soot yielding capacity

increases [60]. This refers to a correlation, the CRNs with higher $d_{s_1 \to C_{2,i}}$ tend to produce more soot. The smallest $d_{s_1 \to C_{2,i}}$ for MIBK might be related to its molecular structure which is the only ketone in this work.

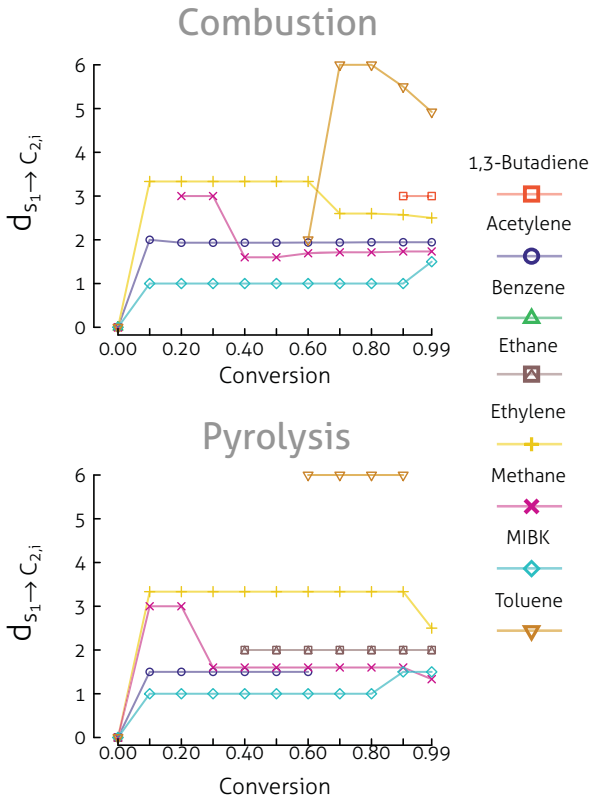


Figure 9: Average of the shortest paths between the reactant (main input molecule) of CRNs and C_2 species in their mechanisms, i.e. the molecules with two carbon atoms.

Theory of CRNs in combination with chemical kinetics knowledge could help the design of reaction systems satisfying both of the targets mentioned above. Quenching techniques like carbon sequestration is believed to be used in reduction of the global warming. However enormous amount of the atmospheric matter to be stored is a serious technological barrier. At that point, the centrality analysis of graph theory might determine the low concentration key species and reactions which can be more easily handled. In the case of positive targets like chemical synthesis, CRNs can be optimized to improve the production efficiency and enhance the robustness against disturbances in the reaction conditions which is also the future research goal of this study.

Acknowledgement

We are grateful to Prof. W.H. Green and his research team from MIT, USA for their essential support on the RMG software. The authors also wish to thank the referees for their invaluable comments and recommendations enabling pretty refined final form of this study.

Supplementary Content

An introductory text on the foundations of graph theory, detailed description of computational techniques and methods as well as complete list of reaction mechanisms are available in Supplementary Information.

References

[1] Henry Eyring. Quantum mechanics and chemical reactions. *Chemical Reviews*, 10(1):103–123, 1932.

[2] William H. Miller. Spiers memorial lecture quantum and semiclassical theory of chemical reaction rates. *Faraday Discuss.*, 110:1–21, 1998.

[3] Stuart C. Althorpe and David C. Clary. Quantum scattering calculations on chemical reactions. *Annual Review of Physical Chemistry*, 54(1):493–529, 2003. PMID: 12651964.

[4] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002.

[5] M. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[6] Petter Holme and Jari Saramäki. *Temporal networks*. Springer, 2013.

[7] Fritz Horn and Roy Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47(2):81–116, 1972.

[8] Bruce L Clarke. *Stability of complex reaction networks*. Wiley Online Library, 1980.

[9] Martin Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors - 1. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 42(10):2229 – 2268, 1987.

[10] Hawoong Jeong, Bálint Tombor, Réka Albert, Zoltan N Oltvai, and A-L Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.

[11] Andreas Wagner and David A. Fell. The small world inside large metabolic networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1478):1803–1810, 2001.

[12] Frank J. Bruggeman and Hans V. Westerhoff. The nature of systems biology. *Trends in Microbiology*, 15(1):45 – 50, 2007.

[13] Chris M. Gothard, Siowling Soh, Nosheen A. Gothard, Bartlomiej Kowalczyk, Yanhu Wei, Bilge Baytekin, and Bartosz A. Grzybowski. Rewiring chemistry: Algorithmic discovery and experimental validation of one-pot reactions in the network of organic chemistry. *Angewandte Chemie*, 124(32):8046–8051, 2012.

[14] Mikolaj Kowalik, Chris M. Gothard, Aaron M. Drews, Nosheen A. Gothard, Alex Weckiewicz, Patrick E. Fuller, Bartosz A. Grzybowski, and Kyle J. M. Bishop. Parallel optimization of synthetic pathways within the network of organic chemistry. *Angewandte Chemie International Edition*, 51(32):7928–7932, 2012.

[15] Patrick E. Fuller, Chris M. Gothard, Nosheen A. Gothard, Alex Weckiewicz, and Bartosz A. Grzybowski. Chemical network algorithms for the risk assessment and management of chemical threats. *Angewandte Chemie*, 124(32):8057–8061, 2012.

[16] Gil Benkö. A toy model of chemical reaction networks. *Master's thesis, Universität Wien*, 2002.

[17] Phillip E. Savage. Mechanisms and kinetics models for hydrocarbon pyrolysis. *Journal of Analytical and Applied Pyrolysis*, 54(1–2):109 – 126, 2000.

[18] Eliseo Ranzi, Alessio Frassoldati, Silvia Granata, and Tiziano Faravelli. Wide-range kinetic modeling study of the pyrolysis, partial oxidation, and combustion of heavy n-alkanes. *Industrial - Engineering Chemistry Research*, 44(14):5170–5183, 2005.

[19] Michael J. Pilling. From elementary reactions to evaluated chemical mechanisms for combustion models. *Proceedings of the Combustion Institute*, 32(1):27 – 44, 2009.

[20] John M. Simmie. Detailed chemical kinetic models for the combustion of hydrocarbon fuels. *Progress in Energy and Combustion Science*, 29(6):599 – 634, 2003.

[21] Michael Frenklach. Monte carlo simulation of diamond growth by methyl and acetylene reactions. *The Journal of Chemical Physics*, 97(8):5794–5802, 1992.

[22] Michael Balthasar and Michael Frenklach. Monte-carlo simulation of soot particle coagulation and aggregation: the effect of a realistic size distribution. *Proceedings of the Combustion Institute*, 30(1):1467 – 1475, 2005.

[23] Russell Whitesides and Michael Frenklach. Detailed kinetic monte carlo simulations of graphene-edge growth. *The Journal of Physical Chemistry A*, 114(2):689–703, 2010. PMID: 20000728.

[24] Hans-Heinrich Carstensen and Anthony M. Dean. Rate constant rules for the automated generation of gas-phase reaction mechanisms. *The Journal of Physical Chemistry A*, 113(2):367–380, 2009. PMID: 19090679.

[25] William H. Green, Joshua W. Allen, Beat A. Buesser, Robert W. Ashcraft, Gregory J. Beran, Caleb A. Class, Connie Gao, C. Franklin Goldsmith, Michael R. Harper, Amrit Jalan, Murat Keceli, Gregory R. Magoon, David M. Matheu, Shamel S. Merchant, Jeffrey D. Mo, Sarah Petway, Sumathy Raman, Sandeep Sharma, Jing Song, Yury Suleymanov, Kevin M. Van Geem, John Wen, Richard H. West, Andrew Wong,

Hsi-Wu Wong, Paul E. Yelvington, Nathan Yee, and Joanna Yu. RMG - Reaction Mechanism Generator v4.0.1. `http://rmg.sourceforge.net/`, 2013.

[26] Richard FW Bader. Atoms in molecules: a quantum theory. international series of monographs on chemistry 22, 1990.

[27] PLA Popelier. On the full topology of the laplacian of the electron density. *Coordination Chemistry Reviews*, 197(1):169–189, 2000.

[28] Nathaniel OJ Malcolm and Paul LA Popelier. On the full topology of the laplacian of the electron density 2: umbrella inversion of the ammonia molecule. *The Journal of Physical Chemistry A*, 105(32):7638–7645, 2001.

[29] P.L.A. Popelier. *Atoms in Molecules: An Introduction*. Pearson Education. Prentice Hall, 2000.

[30] RFW Bader, PJ MacDougall, and CDH Lau. Bonded and nonbonded charge concentrations and their relation to molecular geometry and reactivity. *Journal of the american Chemical Society*, 106(6):1594–1605, 1984.

[31] O.N. Temkin, A.V. Zeigarnik, and D.G. Bonchev. *Chemical Reaction Networks: A Graph-Theoretical Approach*. Taylor & Francis, 1996.

[32] AV Zeigarnik and ON Temkin. A graph-theoretical model of complex reaction mechanisms: Bipartite graphs and the stoichiometry of complex reactions. *Kinetics and catalysis*, 35(5):647–655, 1994.

[33] M. Newman. *Networks: An Introduction*. OUP Oxford, 2010.

[34] Petter Holme and Mikael Huss. Substance graphs are optimal simple-graph representations of metabolism. *Chinese Science Bulletin*, 55(27-28):3161–3168, 2010.

[35] Wynand Winterbach, Huijuan Wang, Marcel Reinders, Piet Van Mieghem, and Dick de Ridder. Metabolic network destruction: Relating topology to robustness. *Nano Communication Networks*, 2(2):88–98, 2011.

[36] Baruch Barzel, Ofer Biham, Raz Kupferman, Azi Lipshtat, and Amir Zait. Dimensional reduction of the master equation for stochastic chemical networks: The reduced-multiplane method. *Physical Review E*, 82(2):021117, 2010.

[37] Satu Elisa Schaeffer. Graph clustering. *Computer Science Review*, 1(1):27–64, 2007.

[38] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.

[39] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.

[40] Pedro G Lind, Marta C Gonzalez, and Hans J Herrmann. Cycles and clustering in bipartite networks. *Physical review E*, 72(5):056127, 2005.

[41] Henry Dorrian, Kieran Smallbone, et al. Size dependent growth in metabolic networks. *arXiv preprint arXiv:1210.2550*, 2012.

[42] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, Oct 2002.

[43] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, Feb 2004.

[44] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006.

[45] Teh C. Ho. Kinetic modeling of large`[U+2010]`scale reaction systems. *Catalysis Reviews*, 50(3):287–378, 2008.

[46] A. Clauset, C. Shalizi, and M. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[47] Gourab Ghoshal and Albert-László Barabási. Ranking stability and super-stable nodes in complex networks. *Nature communications*, 2:394, 2011.

[48] Duncan S. Callaway, John E. Hopcroft, Jon M. Kleinberg, M. E. J. Newman, and Steven H. Strogatz. Are randomly grown graphs really random? *Phys. Rev. E*, 64:041902, Sep 2001.

[49] Weibing Deng, Wei Li, Xu Cai, and Qiuping A. Wang. The exponential degree distribution in complex networks: Non-equilibrium network theory, numerical simulation and empirical data. *Physica A: Statistical Mechanics and its Applications*, 390(8):1481 – 1485, 2011.

[50] JC Nacher, T Ochiai, M Hayashida, and T Akutsu. A mathematical model for generating bipartite graphs and its application to protein networks. *Journal of Physics A: Mathematical and Theoretical*, 42(48):485005, 2009.

[51] Chu-Xu Zhang, Zi-Ke Zhang, and Chuang Liu. An evolving model of online bipartite networks. *Physica A: Statistical Mechanics and its Applications*, 392(23):6100 – 6106, 2013.

[52] Luıs A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences*, 97(21):11149–11152, 2000.

[53] Matú š Medo, Giulio Cimini, and Stanislao Gualdi. Temporal effects in the growth of networks. *Phys. Rev. Lett.*, 107:238701, Dec 2011.

[54] Mark D Humphries and Kevin Gurney. Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PloS one*, 3(4):e0002051, 2008.

[55] E.D. Kolaczyk. *Statistical Analysis of Network Data: Methods and Models*. Springer Series in Statistics. Springer, 2009.

[56] Boxiong Shen, Qiang Yao, and Xuchang Xu. Kinetic model for natural gas reburning. *Fuel Processing Technology*, 85(11):1301 – 1315, 2004.

[57] Hong-bin Xie, Yi-hong Ding, and Chia-chung Sun. Reaction of ketenyl radical with acetylene: a promising route for cyclopropenyl radical. *The Journal of Physical Chemistry A*, 110(22):7262–7267, 2006. PMID: 16737278.

[58] Wing Tsang. Chemical kinetic data base for hydrocarbon pyrolysis. *Industrial & Engineering Chemistry Research*, 31(1):3–8, 1992.

[59] Vladimir S Arutyunov and Vladimir I Vedeneev. Pyrolysis of methane in the temperature range 1000–1700 k. *Russian Chemical Reviews*, 60(12):1384, 1991.

[60] H Richter and J.B Howard. Formation of polycyclic aromatic hydrocarbons and their growth to soot—a review of chemical reaction pathways. *Progress in Energy and Combustion Science*, 26(4–6):565 – 608, 2000.

[61] Rubik Asatryan, Gabriel da Silva, and Joseph W. Bozzelli. Quantum chemical study of the acrolein (CH2CHCHO) + OH + O2 reactions. *The Journal of Physical Chemistry A*, 114(32):8302–8311, 2010.

[62] James R. Fincke, Raymond P. Anderson, Timothy A. Hyde, and Brent A. Detering. Plasma pyrolysis of methane to hydrogen and carbon black. *Industrial & Engineering Chemistry Research*, 41(6):1425–1435, 2002.

[63] D.J. Goebbert. *Bonding Interactions Through Hydrogen*. PhD Thesis Purdue University, 2006.

[64] Conner I. Sandefur, Maya Mincheva, and Santiago Schnell. Network representations and methods for the analysis of chemical and biochemical pathways. *Mol. BioSyst.*, 9:2189–2200, 2013.

[65] Almerima Jamakovic, Priya Mahadevan, Amin Vahdat, Marián Boguná, and Dmitri Krioukov. How small are building blocks of complex networks. *arXiv preprint arXiv:0908.1143*, 2009.

[66] M. B. Isichenko. Percolation, statistical topography, and transport in random media. *Rev. Mod. Phys.*, 64:961–1043, Oct 1992.

[67] Tsuneyoshi Nakayama, Kousuke Yakubo, and Raymond L. Orbach. Dynamical properties of fractal networks: Scaling, numerical simulations, and physical realizations. *Rev. Mod. Phys.*, 66:381–443, Apr 1994.

[68] Scott Kirkpatrick. Percolation and conduction. *Rev. Mod. Phys.*, 45:574–588, Oct 1973.

[69] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 60:7332–7342, Dec 1999.

[70] J. Kim, P. L. Krapivsky, B. Kahng, and S. Redner. Infinite-order percolation and giant fluctuations in a protein interaction network. *Phys. Rev. E*, 66:055101, Nov 2002.

[71] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Anomalous percolation properties of growing networks. *Phys. Rev. E*, 64:066110, Nov 2001.

[72] S. A. Trugman and Abel Weinrib. Percolation with a threshold at zero: A new universality class. *Phys. Rev. B*, 31:2974–2980, Mar 1985.

[73] M. Ángeles Serrano, Dmitri Krioukov, and Marián Boguñá. Percolation in self-similar networks. *Phys. Rev. Lett.*, 106:048701, Jan 2011.

[74] Y. S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim. Percolation transitions in scale-free networks under the achlioptas process. *Phys. Rev. Lett.*, 103:135702, Sep 2009.

[75] Roberto Arévalo, Luis A. Pugnaloni, Diego Maza, and Iker Zuriguel. Tapped granular packings described as complex networks. *Philosophical Magazine*, 93(31-33):4078–4089, 2013.

[76] Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow. The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*, 2011.

[77] Ning Ning Chung, Lock Yue Chew, and Choy Heng Lai. Spectral analysis on explosive percolation. *EPL (Europhysics Letters)*, 101(6):66003, 2013.

[78] Eduardo López, Roni Parshani, Reuven Cohen, Shai Carmi, and Shlomo Havlin. Limited path percolation in complex networks. *Phys. Rev. Lett.*, 99:188701, Oct 2007.

[79] Masoud Mohseni, Patrick Rebentrost, Seth Lloyd, and Alán Aspuru-Guzik. Environment-assisted quantum walks in photosynthetic energy transfer. *The Journal of Chemical Physics*, 129(17):–, 2008.

[80] E. J. Garboczi, K. A. Snyder, J. F. Douglas, and M. F. Thorpe. Geometrical percolation threshold of overlapping ellipsoids. *Phys. Rev. E*, 52:819–828, Jul 1995.

[81] David M. Walker and Antoinette Tordesillas. Topological evolution in dense granular materials: A complex networks perspective. *International Journal of Solids and Structures*, 47(5):624 – 639, 2010.

[82] Richard FW Bader and Preston J MacDougall. Toward a theory of chemical reactivity based on the charge density. *Journal of the American Chemical Society*, 107(24):6788–6795, 1985.

[83] Marshall T Carroll, Cheng Chang, and Richard FW Bader. Prediction of the structures of hydrogen-bonded complexes using the laplacian of the charge density. *Molecular Physics*, 63(3):387–405, 1988.

[84] Richard FW Bader and Cheng Chang. Properties of atoms in molecules: electrophilic aromatic substitution. *The Journal of Physical Chemistry*, 93(8):2946–2956, 1989.

[85] Marshall T Carroll, James R Cheeseman, Roman Osman, and Harel Weinstein. Nucleophilic addition to activated double bonds: predictions of reactivity from the laplacian of the charge density. *The Journal of Physical Chemistry*, 93(13):5120–5123, 1989.

[86] RFW Bader, PLA Popelier, and C Chang. Similarity and complementarity in chemistry. *Journal of Molecular Structure: THEOCHEM*, 255:145–171, 1992.

[87] Y.-B. Yi and A. M. Sastry. Analytical approximation of the percolation threshold for overlapping ellipsoids of revolution. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2048):2353–2380, 2004.

[88] Christopher J. Pope, Joseph A. Marr, and Jack B. Howard. Chemistry of fullerenes c60 and c70 formation in flames. *The Journal of Physical Chemistry*, 97(42):11001–11013, 1993.

[89] Vincent Jourdain and Christophe Bichara. Current understanding of the growth of carbon

nanotubes in catalytic chemical vapour deposition. *Carbon*, 58(0):2 – 39, 2013.

[90] Michael Frenklach and Jonathan Ping. On the role of surface migration in the growth and structure of graphene layers. *Carbon*, 42(7):1209 – 1212, 2004. Carbon '03 Conference.

[91] Koyo Norinaga and Olaf Deutschmann. Detailed kinetic modeling of gas-phase reactions in the chemical vapor deposition of carbon from light hydrocarbons. *Industrial - Engineering Chemistry Research*, 46(11):3547–3557, 2007.

[92] T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552, 2013.

[93] S.O. Baek, R.A. Field, M.E. Goldstone, P.W. Kirk, J.N. Lester, and R. Perry. A review of atmospheric polycyclic aromatic hydrocarbons: Sources, fate and behavior. *Water, Air, and Soil Pollution*, 60(3-4):279–300, 1991.

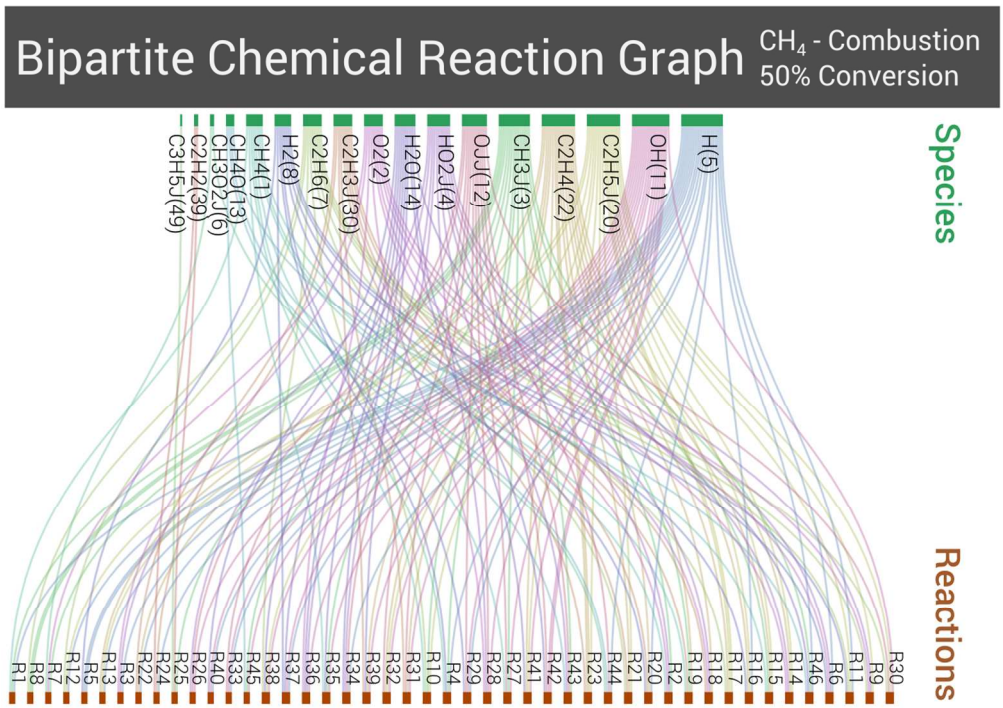
[94] Henning Richter and Jack B. Howard. Formation and consumption of single-ring aromatic hydrocarbons and their precursors in premixed acetylene, ethylene and benzene flames. *Phys. Chem. Chem. Phys.*, 4:2038–2055, 2002.

[95] Jack B Howard. Carbon addition and oxidation reactions in heterogeneous combustion and soot formation. In *Symposium (International) on Combustion*, volume 23, pages 1107–1127. Elsevier, 1991.

[96] Michael Frenklach. Reaction mechanism of soot formation in flames. *Phys. Chem. Chem. Phys.*, 4:2028–2037, 2002.

[97] Michael Frenklach and Hai Wang. Detailed modeling of soot particle nucleation and growth. *Symposium (International) on Combustion*, 23(1):1559 – 1566, 1991. Twenty-Third Symposium (International) on Combustion.

[98] Michael Frenklach and Hai Wang. Aromatics growth beyond the first ring and the nucleation of soot particles. In *Preprints of the 202nd ACS National Meeting (American Chemical Society, Washington, DC 1991) Vol*, volume 36, page 1509, 1991.


Bipartite Chemical Reaction Graph
CH_4 - Combustion
50% Conversion
Species
Reactions
H(5)
OH(11)
C2H5J(20)
C2H4(22)
CH3J(3)
OJJ(12)
HO2J(4)
H2O(14)
O2(2)
C2H3J(30)
C2H6(7)
H2(8)
CH4(1)
CH4O(13)
CH3O2J(6)
C2H2(39)
C3H5J(49)
R30
R9
R11
R6
R46
R14
R15
R16
R17
R18
R19
R2
R20
R21
R44
R23
R43
R42
R41
R27
R28
R29
R4
R10
R31
R32
R39
R34
R35
R36
R37
R38
R45
R33
R40
R26
R25
R24
R22
R3
R13
R5
R12
R7
R8
R1