This is a repository copy of *Graph theory applied to plasma chemical reaction engineering*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/170931/

Version: Published Version

ORIGINAL PAPER

Graph Theory Applied to Plasma Chemical Reaction Engineering

Thomas D. Holmes[1] · Rachael H. Rothman[1] · William B. Zimmerman[1]

Abstract

This work explores the following applications of graph theory to plasma chemical reaction engineering: assembly of a weighted directional graph with the key addition of reaction nodes, from a published set of reaction data for air; graph visualisation for probing the reaction network for potentially useful or problematic reaction pathways; running Dijkstra's algorithm between all species nodes; further analysis of the graph for useful engineering information such as which conditions, reactions, or species could be enhanced or supressed to favour particular outcomes, e.g. targeted chemical formation. The use of reaction-nodes combined with derived parameters allowed large amounts of key information regarding the plasma chemical reaction network to be assessed simultaneously using a leading open source graph visualisation software (Gephi). A connectivity matrix of Dijkstra's algorithm between each two species gave a measure of the relative potential of species to be created and destroyed under specific conditions. Further investigation into using the graph for key reaction engineering information led to the development of a graph analysis algorithm to quantify demand for conditions for targeted chemical formation: Optimal Condition Approaching via Reaction-In-Network Analysis (OCARINA). Predictions given by running OCARINA display significant similarities to a well-known electric field strength regime for optimal ozone production in air. Time dependent 0D simulations also showed preferential formation for O· and O_3 using the respective conditions generated by the algorithm. These applications of graph theory to plasma chemical reaction engineering show potential in identifying promising simulations and experiments to devote resources.

Keywords Graph theory · Reaction engineering · Chemical engineering · Plasma chemistry · Optimization

✉ Thomas D. Holmes
  t.d.holmes@sheffield.ac.uk

  Rachael H. Rothman
  r.rothman@sheffield.ac.uk

  William B. Zimmerman
  w.zimmerman@sheffield.ac.uk

[1] Department of Chemical and Biological Engineering, The University of Sheffield, Sheffield, UK

Introduction

The prospect of chemical reaction engineering using atmospheric pressure plasmas is an increasingly attractive one. The wide range of species known to be present, speed with which the plasma energy can be altered, ability to attain high energy conditions without the need for expensive high pressure and temperature equipment, and the sheer abundance and availability of the "raw materials" (atmospheric gases, water, electrical energy and biomass) all justify the interest in this growing field [1, 2].

Whilst there are many applications of plasma in chemical engineering, including electrolysis cells, biomass treatment and packed bed reactors [3–6] to name but a few, one of the main obstacles limiting the more widespread use of plasmas for many chemistry and chemical engineering applications is their inherent lack of selectivity at atmospheric pressure [7, 8]. The energy supplied to the plasma manifests in the movement of charged particles, but where these particles move, and which other particles they collide with and when, is not trivial to control at the particle densities encountered at atmospheric pressure [9, 10].

Various possible solutions to this are being researched, such as plasma catalysis [11] and tailored waveforms [12]. One problem that each solution must encounter, however, is the ever-expanding complexity of plasma chemical reaction schemes. The number of possible reactions in air alone can be in excess of 1800 reactions, with over 70 different species being present [13]. This will only increase with the need to simulate plasma reactions involving more complex chemicals and mixtures.

For any chemical kinetic system, including plasma chemical kinetics, a number of questions need to be answered for the chemistry to be applied to make a particular chemical species. Some of the key questions are as follows:

1. Which chemical pathways are most important to our species of interest?
2. Which chemical species are potentially most disruptive to these pathways?
3. Could other chemical mechanisms potentially be better?
4. Is there other interesting plasma chemistry that could be possible?
5. Which conditions might affect those pathways?

To answer these questions, we need to be able to compare and contrast information about many different reactions simultaneously, as well as the relationships between them. Examples of more specific questions are:

(a) Which reactions are affected by temperature and which are not?
(b) Which are affected by electron energy and which are not? (both directly and indirectly)
(c) What are the relative rates of each of the reactions?
(d) How are the different reactions connected to each other?
(e) Are there any feedback loops with certain reactions?

In an ideal world, we could draw upon an exhaustive dataset to answer these questions; a complete experimental characterisation over all the combinations of variable values. However, even limiting each variable to only ten experimental values still means a tenfold increase in the number of experiments for each additional variable. To vary the concentrations of all the possible species or the rates of all the reactions in a plasma chemical kinetic scheme that could contain over 50 species and hundreds of reactions would require

multiple lifetimes of the universe even for rapid simplified simulations, let alone more time-consuming experiments.

Identifying patterns in the chemical kinetic data prior to running experiments or simulations could greatly assist in winnowing the number of simulations and experiments down to those which are potentially most relevant to a given application. One field of mathematics which has experienced many successful applications in finding patterns in large datasets of interrelated entities is Graph Theory.

Graph Theory is concerned with mathematical structures consisting of nodes (or vertices) connected by edges. The edges can also be directional, representing a connection from one node to another, but not the reverse. The nodes and edges can also be weighted, to quantify the relative significance of each connection between two nodes [14]. For example, in transport network analysis the relative speed of different routes from one point to another could depend on the rates of traffic, the size of the roads, the prevalence of one-way systems, etc. All this can be effectively modelled using graph theory in order to find the fastest route between two nodes, to highlight which routes should be avoided, and to identify possible changes that could be made to the system in order to increase the maximum convenience of the routes [15].

Visualisation of graphs can be especially useful as a means of comparing pathways. The visual representation of the data arranged in graph format has the advantage of being able to show the relationships between data entities (in this case, between species and reactions). Additional data associated with these entities and the relationships between them can also be included in the visualisation as attributes of the points or arrows (such as width, colour, size, etc.). The result is a highly information dense data visualisation, where the information may be more efficiently processed by the human mind [16].

Moreover, any of a growing number of numerical algorithms can be used to analyse a graph data structure (i.e. adjacency list, adjacency matrix) to assess the connectivity between entities, find the shortest pathways between entities, assess the number of feedback loops between entities, and many more.

Graph theory has already experienced some use in the field of chemical kinetics [17], but to date has experienced very little work in the field of plasma chemistry apart from the excellent work of Sakai et al. [18] which analysed the centrality index of a methane plasma chemical reaction system and also proposed the use of the reaction rate coefficients as edge weights, and Mizui et al. [19] who conducted more complex numerical analysis of the roles of the chemical species in the plasma chemical reaction network. Graph theory could prove to be a convenient tool in identifying potential ways to improve the selectivity and even the energy efficiency (which is strongly influenced by the direction of the reaction pathways and the plasma conditions [10]) of plasma chemical systems.

With modern computer processing power, the speed and effectiveness with which graph theory can be applied to solve problems in complex systems makes a strong case for further investigating how it could be applied to plasma chemical reaction systems to answer some of the engineering questions above. The following sections describe some first steps taken in investigating this.

Section "Methodology: Building the Graph of a Plasma Chemical Reaction System" describes a novel method of generating a graph from a plasma chemical kinetic dataset, using the chemical species as nodes and the chemical reactions as both nodes and edges. Section "Methodology: Visual Representation" then describes and demonstrates how this graph can be visualised for the purposes of plasma chemical reaction engineering.

Finally, "Mathematical Graph Theory Operations on the Plasma Chemistry Graph" section tests the application of mathematical graph algorithms on a plasma chemical graph

from an existing dataset [20] to gain useful information for the purposes of plasma chemical reaction engineering, with "Application of Dijkstra's Shortest Path Algorithm: Relative Potential Reaction-Chain Rate Connectivity Matrix" section demonstrating the application of Dijkstra's shortest path algorithm, and "Development of a Rudimentary Graph Algorithm for the Estimation of Optimal Condition Sequence for a Target Plasma Chemical Species" section detailing the development and operation of a novel algorithm for optimal condition proposal: Optimal Condition Approaching via Reaction-In-Network Analysis (OCARINA).

Methodology: Building the Graph of a Plasma Chemical Reaction System

This section details the necessary first step of forming a graph from a plasma chemical kinetic dataset. The data required for plasma chemical reaction engineering generally consists of a list of reactions with their corresponding reactants and products, along with the rate coefficients of such reactions. The rate coefficients are either constants, or are functions of one or more parameters, normally electron energy or temperature. Simulating the evolution of the concentrations of species over time consists of solving the rate equations for each reaction over discrete timesteps, summing the rates of creation and destruction reactions for each species, and then subtracting the latter from the former for each timestep. Whilst the concentration of species changes over time, the rate coefficients are constant for a given set of conditions and can be used as a measure of each reaction as a fast or slow step in the whole network.

For all the graph theory work in this paper, the chemical kinetic reaction data so conveniently included in Sakiyama et al. [20] was used. The data was converted from tabular form into graph form (vertex and edge list form). As all the reactions in the dataset were forward reactions (i.e. there were no equilibrium reactions) a directional graph was selected as the most appropriate graph type. In this study, each individual reaction was connected to each of its reactants and products by setting the reactions as both edges and vertices, connecting the species which were set solely as vertices (Fig. 1a, b). This is different to previous applications of graph theory to plasma chemistry where only the chemical species were set as nodes,
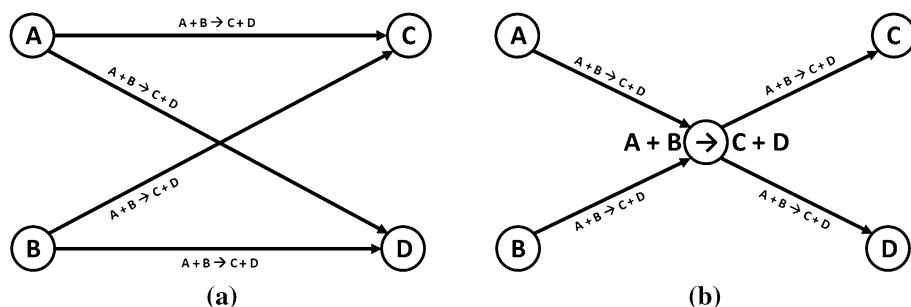

Fig. 1 **a** A simple graph of a single reaction A + B→C + D with reactants and products set as vertices and the reaction as directional edges from each reactant to each product (as used in Sakai et al. [18]). **b** A simple graph of a single reaction A + B→C + D with reaction as a central vertex between reactant and product vertices, and as directional edges from reactants to reaction and reaction to products (as was used in this study)

and the reactions only as edges [18] (Fig. 1a). Hence, the vertices and edges for the reaction A+B→C+D would be as shown in Table 1.

At a first glance this might seem counter intuitive, as the reactions are the paths linking the species to each other and should be thought of as edges. Whilst this is true, it is also true that a single species can be involved in many reactions, and to link each species by the reactions quickly makes visualisation of this graph impractical, and numerical operation on the graph more cumbersome, as an edge is defined by its start and end node thus any reaction with more than a single reactant and product cannot be considered as a single entity. Figure 1b illustrates how this clarifies visualising the relationships in the graph.

Variation of Rate Coefficients with Electron Temperature and Gas Temperature (Electronic and Temperature Susceptibility)

One of the key factors affecting the rates of reactions in plasma chemistry is the electron energy. As some reactions are affected by it, and some are not, the need for an expression of the sensitivity of each reaction to changes in the electron energy was essential in assessing chemical pathways.

The rate coefficients of many of the reactions are functions of the electron energy and of the electron temperature. Therefore, the values of the rate coefficients change if these variables change. However, a comparison of the variation of the rate coefficients over the ranges of these variables typically encountered in DBD plasmas (1–10 eV) shows that the vast majority of the rate coefficients simply increase or decrease over this range (or are not affected at all), albeit at varying rates. A linear approximation over this trend allows the effect of electron energy/electron temperature and/or gas temperature on the rate coefficients to be represented as singles values that we shall call Δk_{T_e} and $\Delta k_{T_{gas}}$ respectively.

The simplest way to quantify this was to use the difference between the rate coefficients at the maximum and minimum electron energies (within the energy range 1–10 eV). A matrix of rate constants was computed using a range of different electron energies (in this example from 1 to 10 eV) and the maximum and minimum values were found from this.

$$\Delta k_{T_e} = \frac{\log_{10}\left(k_{max}\right) - \log_{10}\left(k_{min}\right)}{T_{e_{k_{max}}} - T_{e_{k_{min}}}} \quad 1 \le T_e \le 10$$

$$\Delta k_{T_{gas}} = \frac{\log_{10}\left(k_{max}\right) - \log_{10}\left(k_{min}\right)}{T_{gas_{k_{max}}} - T_{gas_{k_{min}}}} \quad 1 \le T_{gas} \le 10$$

Table 1 Vertex and edge list of a simple graph in the reaction as node format

Vertex list	Edge list		
A	Source	Target	Label
B	A	A+B→C+D	A+B→C+D
A+B→C+D	B	A+B→C+D	A+B→C+D
C	A+B→C+D	C	A+B→C+D
D	A+B→C+D	D	A+B→C+D

Nearly all reactions that required electron energies to "go" (here called "plasma dependent reactions") tended to have their highest values of rate constants (and thus the most rapid reaction rates) at the highest electron energies, whilst the reactions that did not require electron energies to "go" (here called "ambient reactions") were not affected by changes in the electron temperature. A small number of reactions had their highest rate coefficients at the lowest electron energies, in which case the Δk_{T_e} values were slightly negative.

Methodology: Visual Representation

This section describes how a graph formed by the method outlined in the previous section can be visualised to display large amounts of key information simultaneously. It must be stressed that the data visualisation method illustrated here is not a method for deriving precise, quantitative information. Instead, it should be thought of as a method of gaining an overview of a plasma chemical reaction system under consideration, to identify aspects which may warrant further investigation, and which it may not be worthwhile to investigate.

A leading open source graph visualisation software, Gephi [21], was used. A directional weighted graph was generated from an existing plasma chemical reaction system for a humid air plasma [20] as described above. The thickness of each edge was used to show the rate of reaction, the fastest rates having the thickest edges. As the rate constants ranged over more than 30 orders of magnitude, log values of the rate constants were taken to create a visual logarithmic scale. These log values were then squared to produce proportional numbers that could span the maximum size range of the software. Other attributes of each reaction and species were weighted as other visual properties of the edges and nodes and the weighting factors used are included in Tables 2 and 3. The visual results of this are shown in Figs. 2 and 3. Though useful information can be gained from this (such as which species participate in reactions controlled by electron energy, and which reactions are influenced by electron energy as a proportion of the total number of reactions) more effective use of this "hairball" would require a very large high definition display to allow the individual nodes and edges to be more clearly distinguished from each other whilst still being able to visualise the whole system.

Using the filtering functions of Gephi, information relevant to single or small numbers of species of interest can be viewed in increasing levels of complexity. Of particular interest is the ability to view the relationships between a small number of species simultaneously using Gephi's ego network and union filters, which can then be conveniently arranged using a force projection algorithm, in this case a Yifan Hu algorithm [22] as shown in Fig. 3.

Example of Information Gained from this Visualisation: Ozone Formation

It can be seen from Figs. 2 and 3 that a large amount of key information on the plasma chemical system can be assessed simultaneously using graph visualisation.

It can be seen, for example, on inspection of Fig. 3 that ozone is not directly formed by "plasma sensitive" (electron impact) reactions, though its precursors are, and one ozone destruction reaction can be significantly accelerated by electron energy. It can also be seen that when the plasma is "active", the electrons react with many other species aside from those that lead directly to ozone formation.

Table 2 Criteria for specifying/scaling visual attributes of graph vertices

Graph attribute	Vertex type	Variable used as weight	Criteria
Node label size	Both	Identity	If the vertex is a species then it will be a constant large size, if a reaction then it will be a constant small size
Node size	Species	Identity	If vertex is a species then it will have a constant size (to avoid confusion it might be set as the mean value of the reaction node sizes, however it will still be possible to set it apart from the reactions by observing the node label size)
Node size	Reaction	Maximum rate coefficient	Size is proportional to rate coefficient as $\left[\log_{10}\left(k\right)+\log_{10}\left(k_{min}\right)\right]^2$
Node colour	Reaction	"Temperature susceptibility"	Blue = Negative temperature susceptibility (Rate constant is highest at low temperatures) Green = Neutral temperature susceptibility (the temperature has little or no effect on the reaction rate) Red = Positive temperature susceptibility (rate constant is highest at high temperatures)

Table 3 Criteria for specifying/scaling visual attributes of graph edges

Graph attribute	Edge type	Variable used as weight or other attribute	Criteria
Edge size	Reaction	Maximum rate coefficient over range of electron energies at constant temperature (300 K)	Edge size is proportional to rate constant
Edge colour	Reaction	"Electron energy susceptibility" (electron energy at which the rate coefficient experiences its maximum value) Maximum rate coefficient over range of electron energies at constant temperature (300 K)	Green = No electron energy susceptibility Yellow = Positive electron energy susceptibility Darker green = lower rate coefficient Lighter green = higher rate coefficient
Edge label	Reaction	Maximum rate coefficient over range of electron energies at constant temperature (300 K)	The value is set as a string variable and used to label the edge
Edge label text colour	Reaction	Source of the rate coefficient	Black = Rate coefficient from literature values or equations from literature values Blue = Rate coefficient derived from collision cross section data using a Boltzmann solver Red = Rate coefficient based on necessary assumptions
Edge label size	Reaction	Constant	Label sizes are constant, but small enough not to interfere with the visualisation of the overall scheme of things. On zooming in for closer inspection one can then read the actual values of the rate constants

This suggests, therefore, that from what need only be a very brief survey of Figs. 2 and 3, it can be deduced that maximising ozone formation requires "high plasma" conditions to form O, and then a subsequent "off period" (or "low plasma" period) to allow for the O to react and form O.

As this agrees with what is currently known about ozone generation, this visualisation technique may also be useful when applied to reaction systems about which relatively little is known. The efficiency with which these observations can be made suggests that more time spent examining such visualisations may yield yet more interesting observations, such as the identification of as-of-yet un-investigated reaction networks, including potential new catalytic pathways.

Mathematical Graph Theory Operations on the Plasma Chemistry Graph

Using the same graph described above, the following section describes applications of graph theory operations. As useful and applicable information can be gained from mathematical/algorithmic operations of graphs of similar complexity in other fields, it is reasonable to assume that information useful for plasma chemical reaction engineering can also be gleaned in this way.

Application of Dijkstra's Shortest Path Algorithm: Relative Potential Reaction-Chain Rate Connectivity Matrix

There may be a number of successive reactions between the plasma feed gas and a desired chemical product, and a number of different possible reaction pathways by which a product may be formed from a reactant. It is useful to know which one of these is the fastest. Whilst the concentration of reactants change the rate of a reaction, they do not determine which reaction pathway has the potential to be the fastest. It is the rate coefficients which are the best indicators of the potential for reactions to occur if the reactants are present.

If we consider a rapid reaction between two species to be equivalent to a short pathway between two graph nodes, then we can use one of many shortest path algorithms to compare different sequences of reactions from one species (our potential "raw material") to another (our target "product") and compare it against the shortest path lengths to other products.

Figure 4 compares the outputs of Dijkstra's shortest path algorithm [23] run between each pair of species in the example plasma chemical reaction system. This gives an indication of the relative potential for one species to form any of the others, and to be formed by any of the others, in this specific chemical kinetic system. The path lengths are based on the rate coefficients of the reactions involved. Dijkstra's algorithm counts the weight value as numbers of steps, thus an edge with a weight of two counts as a path twice the length of an edge with a weight of one. As the largest rate coefficient represents the fastest reaction, it needs to be altered to compose the shortest path in that step to be used in Dijkstra's algorithm, and vice versa. This was achieved by taking the inverse of all rate coefficient values ($1/k$) and setting them as the reaction edge weights. Once computed, sorting the columns and rows in ascending order of their total values should allow convenient comparison of the relative potentials of one chemical to give rise to another in the reaction network. The

Fig. 2 **a** Visual representation of a graph of chemical reaction rate data from [20] using graph format from ▶
Table 1 and Fig. 1b, and visualisation criteria from Tables 2 and 3. **b** Close up of a section of figure to
clearly show visual attributes of the graph: (A) Species nodes and node labels (O_3 and O_3^-) above and to
the right, (B) Red reaction nodes with high $\Delta k_{T_{gas}}$ above and below right, (C) Yellow reaction edges with
high Δk_{T_e} above, (D) Edge labels with value of rate coefficient (maximum over all independent variables)
above, colour coded for origin of rate coefficient data from Boltzmann solver (blue), published literature
(black) or estimated value (red), (E) Reaction node, above, (F) Thick edges for faster reactions (on the left)
and thin edges for slower reactions (on the right) (Color figure online)

path nodes of each pathway were also retained for future reference, though to display this
information would clutter up the chart.

The resulting chart (Fig. 4) gives us an indication of which other species have the high-
est potential to be formed from a given species (via the most rapid pathways) at an electron
temperature of 1 eV, if every species starts with a concentration of 1. Clearly it will almost
certainly be the case that the species concentrations will not be equal, however, Fig. 4 does
tell us that, for a species with a high destruction potential, even if it is present in low con-
centrations, its fastest destruction pathway still has a very high potential of occurring, and
for species with low creation potentials, even if their reactants are present in very high
concentrations, their formation pathways are still going to be relatively slow, and will result
in relatively low species creation. It effectively gives us a measure of what the reaction
system "wants to do".

This is arguably useful information for plasma chemical reaction engineers consider-
ing altering the concentrations of reactants to try to increase yield, as they may be pushing
against a high destruction potential, or against a very low creation potential, or both. It may
also identify reactions which have the potential to be much more rapid, but for which the
necessary reactants are only present in tiny concentrations i.e. a main benefit of examining
the plasma chemical reaction system in this way is not to see which are the most prevalent
reactions, but rather, which can be the most rapid reactions if the concentrations of their
reactants could be altered accordingly.

Unfortunately Fig. 4 only compares the fastest reaction pathway (shortest path) between
each two species. Clearly, even if all things are equal, a species may be formed more rap-
idly through a combination of two or more slower pathways, for example. Moreover, each
reaction pathway can be affected differently by changes in conditions (temperature, elec-
tron energy, etc.), and this difference may not be apparent from a visual inspection of the
connectivity matrix, although a visualisation similar to Figs. 2 and 3 could shed light on
this. Nevertheless, the connectivity matrix may still be useful in some applications, espe-
cially given how rapidly it can be generated. And at the very least, this first example serves
to highlight how graph theory can be applicable to plasma chemical reaction engineering.

Development of a Rudimentary Graph Algorithm for the Estimation of Optimal Condition Sequence for a Target Plasma Chemical Species

From Target Chemical to Target Waveform

An application of plasmas may only require one or more specific chemicals. Due to their
low chemical selectivity, plasmas may end up being overlooked in applications for which
they would otherwise be ideal. Only one or more specific chemical species may be required
from the myriad of species known to be formed in the plasma. The large variable space of
plasma chemical systems suggests that, for every species, it is possible an optimal sequence
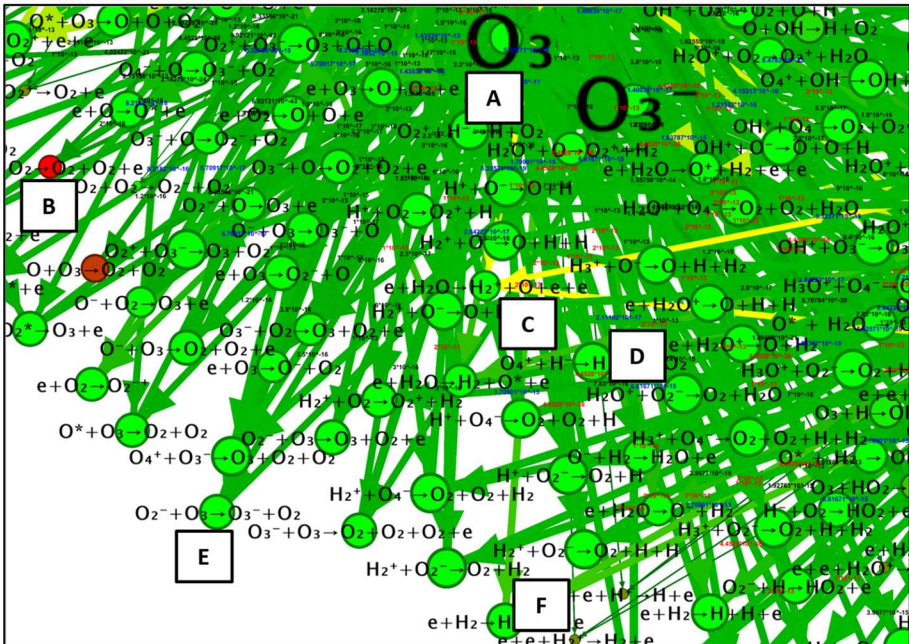
(a)
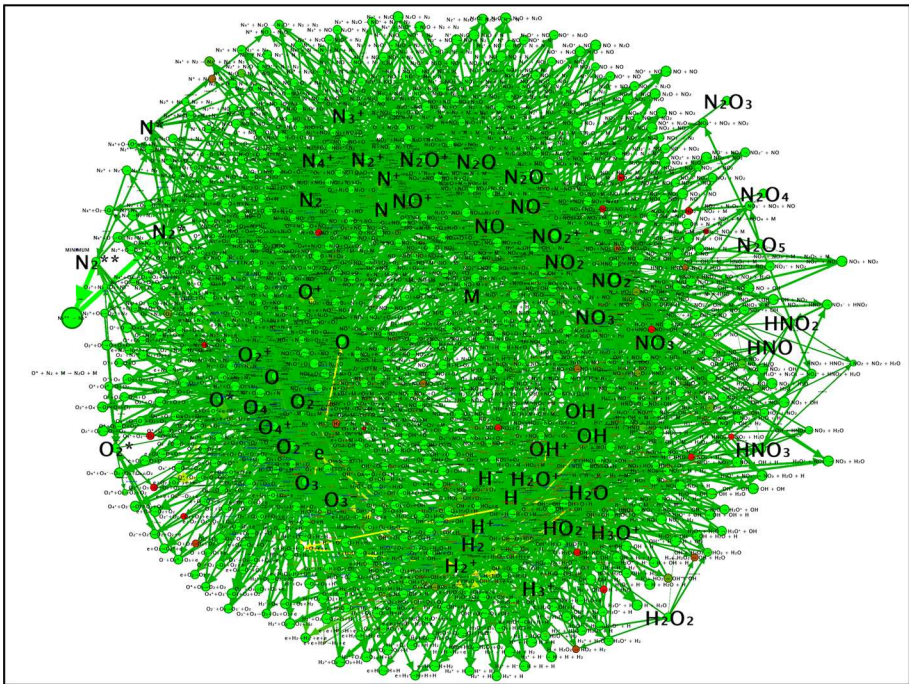
(b)

Fig. 3 a A Yifan Hu force directed projection of an ego network of four species vertices from the chemical ▶ kinetic graph and all reactions concerning those species. Edges and nodes are assigned a repellent "force" to position the nodes apart from each other in the visualisation based on their degree of relationship to one another. **b** Close up of the centre of figure showing the e, O_2, O, and O_3 species nodes, and the reaction nodes and edges involving two or more of these species. **c** Close up of outer (creation and destruction reactions of O that do not involve e, O2 or O3) neighbouring nodes and edges of atomic oxygen (to the top right), and inner neighbouring nodes (to the bottom left). **d, e** Close up of inner and outer neighbouring nodes and edges of O_3 and e species nodes and inner and outer neighbouring node of O2 species node, respectively

of conditions may exist that can allow it to be formed in the most appreciable quantities possible. What follows is the development of a rudimentary tool by which an "optimal sequence of conditions" may potentially be found (OCARINA).

Amongst the means at our disposal for affecting plasma selectivity, are varying the electric field, varying the temperature, the introduction of catalysts, and the introduction of electromagnetic radiation (e.g. photocatalysis).

It is very difficult to alter the temperature externally on the same timescale as the plasma reactions proceed. Also, the low density, even of atmospheric pressure plasmas, means electromagnetic irradiation of the plasma is likely to be a less efficient option for altering the chemistry. The use of catalysts in plasma conditions presents a selection problem as to which chemical reactions should be catalysed, which is something that may also be solved with the following method.

Conversely, the capabilities of high voltage ac power supplies, in both speed and variety of waveform produced, are continuously being improved, and in some cases the electric fields can be altered at speeds matching those of the plasma reaction rates. Energy can potentially be supplied to the electrons almost instantaneously via the electric field. However, knowing which of the myriad of possible waveforms may be most appropriate is still a problem, and any means of estimating how to manipulate the waveform (and therefore the electric field) to favour the selectivity of a particular species will be a worthwhile tool for plasma chemical engineering.

In mathematical kinetic modelling of plasma chemical reactions, some rate constants are a function of electron energy. From this it follows that which reactions occur, and which species are formed, should to a greater or lesser extent depend on how much energy is supplied to the electrons, and when. Supplying electron energy to the plasma in this sense is akin to opening and closing different pathways on a transport network.

As graph theory can be used to identify the best routes to take in a transport network (which has a fair degree of complexity) to get to a particular destination, it might be possible to identify the best chemical pathway to target for a particular chemical species and to recommend a set of conditions to promote that pathway.

Brief Description of the Algorithm

When the plasma chemistry begins in a plasma chemical reactor, we have only the initial reactants, then only the products of reactions of those initial reactants, and so on. It is therefore necessary to reduce the whole reaction network graph to distinct sub graphs for each "phase" (Fig. 5) that can be analysed individually.

Given that different reactions are affected differently by external parameters such as electron energy or temperature for example, how these parameters can be altered to favour the formation of different chemicals can be surmised as follows (see Fig. 5): say that the rate of reaction R_1 is either dependent or strongly enhanced by electron energy.

(a)

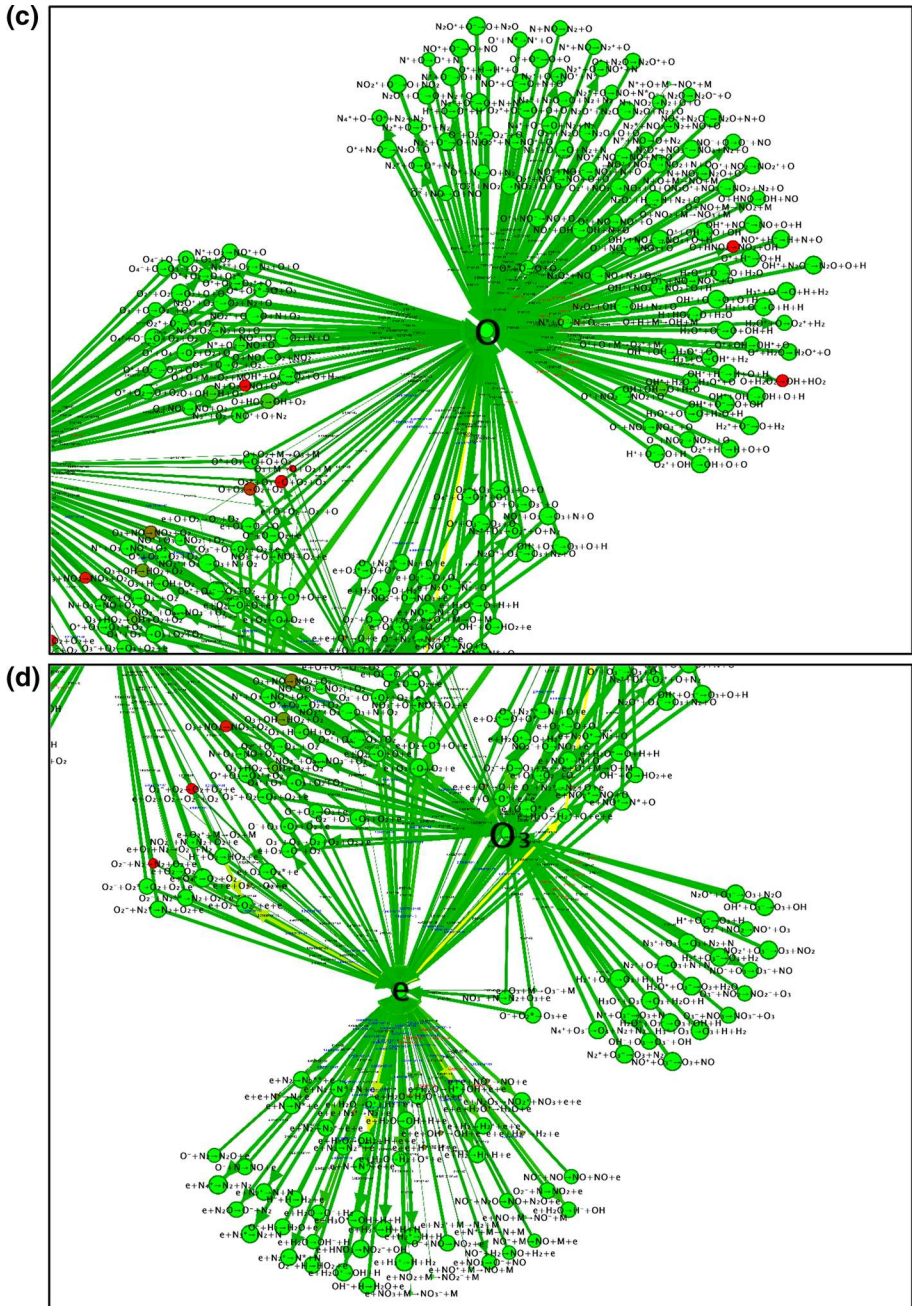**(b)**

(c)

(d)
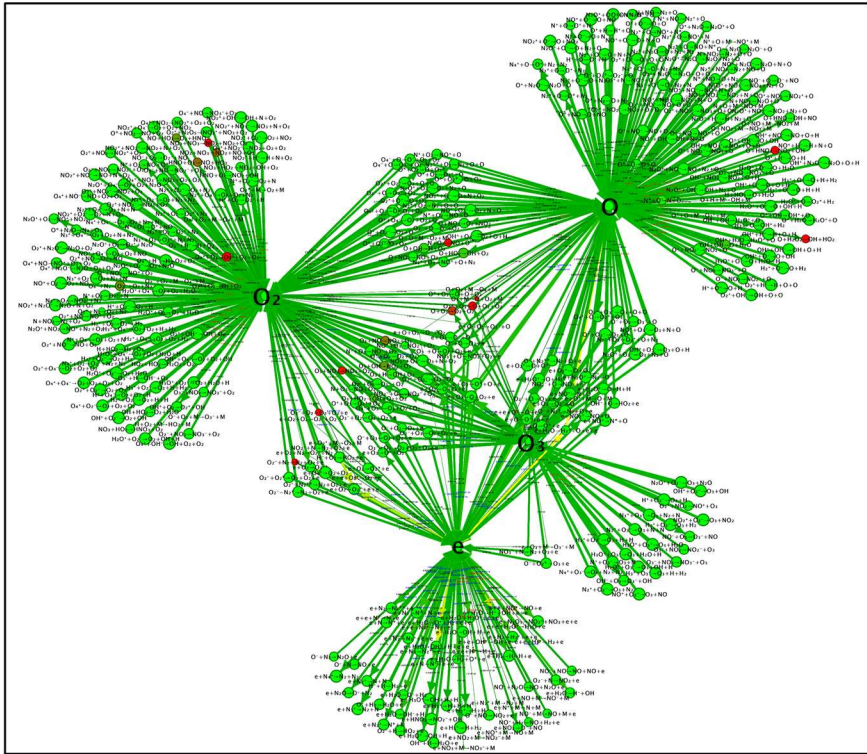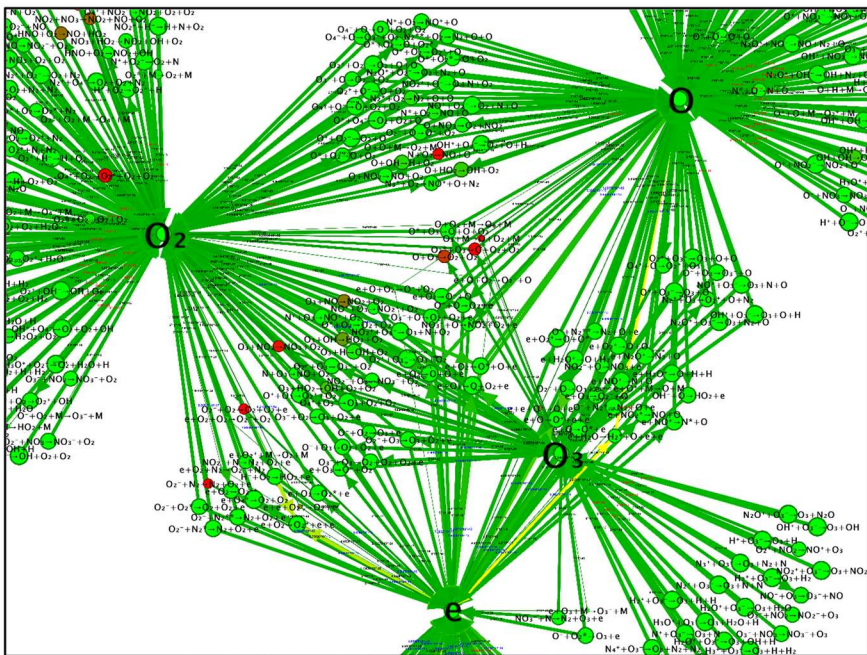
Fig. 3 (continued)

If, for example, the target chemical is D, and reaction R_3 is not affected by electron energy, then the most effective use of electrical energy is in phase 1, whereas in phase 2 an application of electrical energy would be counterproductive (Table 4). Alternatively,
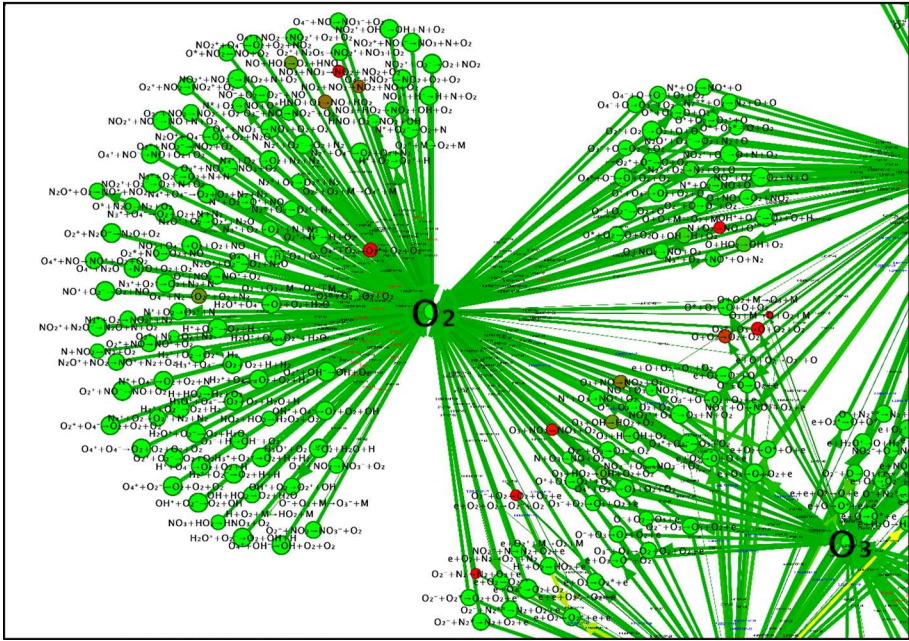
(e)

Fig. 3 (continued)

if E is the target chemical and reaction R_2 is also either dependent or strongly enhanced by electron energy then it is necessary to supply electron energy in this phase too (Table 4). The resulting targeted "Boolean waveforms" for either scenario is illustrated in Fig. 6.

In most scenarios more detail will be required than a simple on or off. It may be that, for example, reaction R_2 does not require as much electron energy as R_1 in order to proceed just as quickly. An engineer examining a simple plasma chemical kinetic scheme, would make a series of decisions when comparing the rate constants of the reactions, which could be expressed as a series of logical functions. Likewise, a series of computational logical statements are required to appropriately estimate a relative electron energy "demand" of all the reactions for more complex chemical kinetic schemes. It is unlikely that a set of logical statements suitable for all chemical kinetic schemes can be built up without much more long-term scrutiny than is available in the scope of this work. A preliminary set of some of the most important logical statements, however, are listed as follows:

1. If a reaction creates the target node then it scores positively ($+1$).
2. If a reaction creates a creator of the target node then it scores positively ($+1$).
3. If a reaction destroys a destroyer of the target node then it scores positively ($+1$).
4. If a reaction destroys the target node then it scores negatively (-1).
5. If a reaction creates a destroyer of the target node then it scores negatively (-1).
6. If a reaction destroys a creator of the target node then it scores negatively (-1).
7. If none of the above is true then the reaction does not score ($+0$).
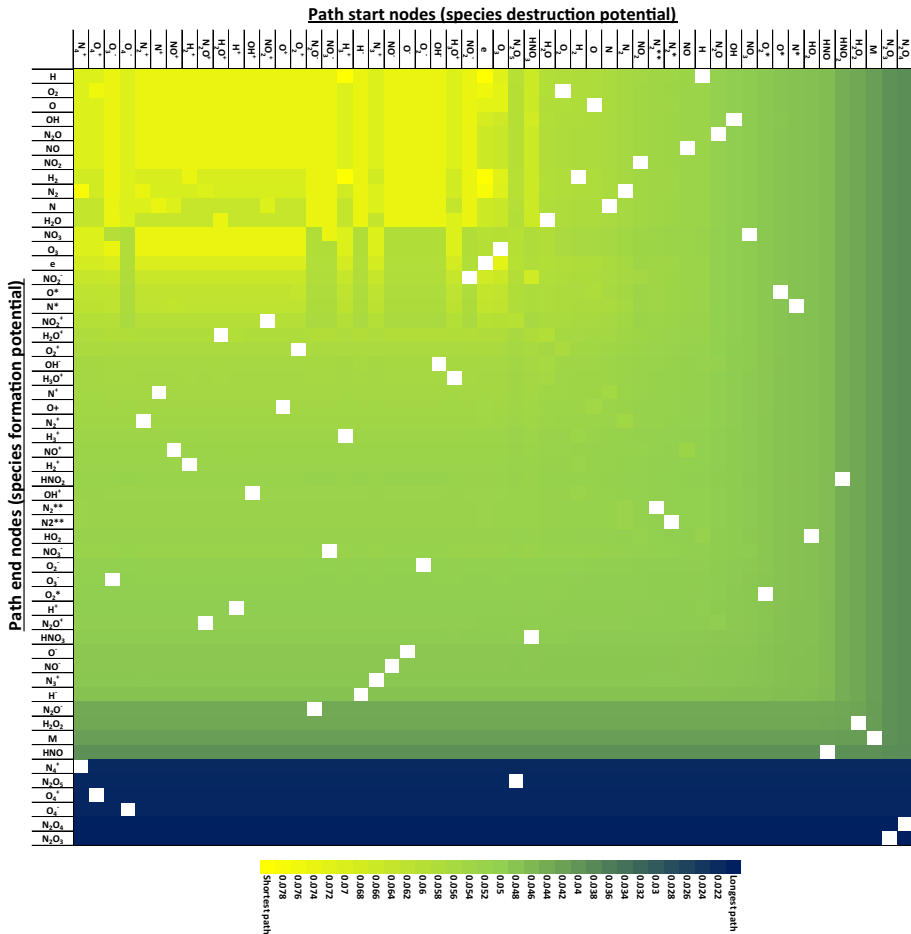
Fig. 4 A Connectivity Matrix of relative fastest pathways (using Dijkstra's algorithm) between each chemical species in the plasma chemical kinetic scheme. The lightest areas are the "shortest paths". The values are arrived at by taking the inverse of the logarithm to base 10 of the computed shortest path lengths, the range of values is shown on the scale at the foot of the figure. The darker blocks represent the longest path lengths, the lighter blocks represent the shortest. In the case of the path of any species to itself, the block is coloured white. The rows are arranged by the sum of their path lengths, shortest sum of path lengths at the top, longest at the bottom. The columns are arranged in the same way, shortest at the left, longest at the right. The shortest collection of routes from all species is to H, the longest is to N_2O_3, suggesting that H has the highest potential to be formed and N_2O_3 the lowest potential to be formed based on the shortest paths alone (starting from equal concentrations of all species)

These steps should be modified, and additional steps added, as and when flaws in the logic are revealed. Figure 7 shows how these logical statements can be computed to determine the value of a conditional modifier (cμ) for each reaction.

Using these conditional modifiers for each reaction in the network, for each graph phase i the *demand* of our target species formation for electron energy can be estimated as follows:

Fig. 5 A simple example of different phases of reaction groups. Reactions are only included in a phase if their reactants are products of the reactions in the preceding phase
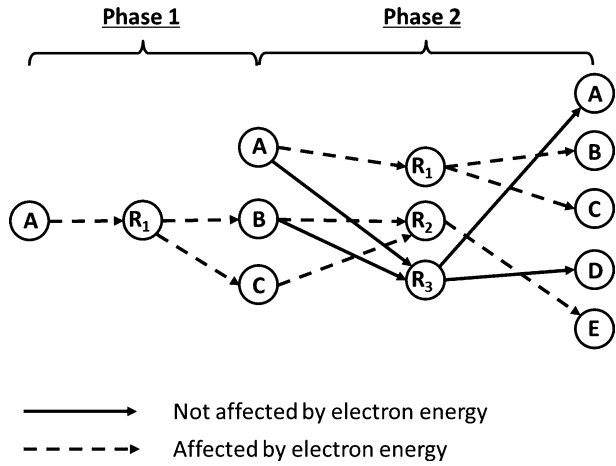
Table 4 An example justification of a simple "on or off" waveform in Fig. 6

Target chemical	Phase 1	Phase 2	Reason
D	Electron energy supplied	No electron energy supplied	R3 not affected by electron energy
E	Electron energy supplied	Electron energy supplied	R2 strongly enhanced by electron energy

Fig. 6 An example of two hypothetical "Boolean waveforms" to target optimum formation of chemicals D or E based on the phases in Fig. 5
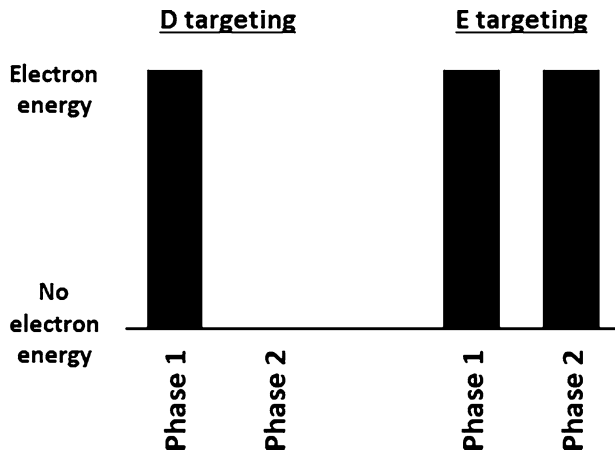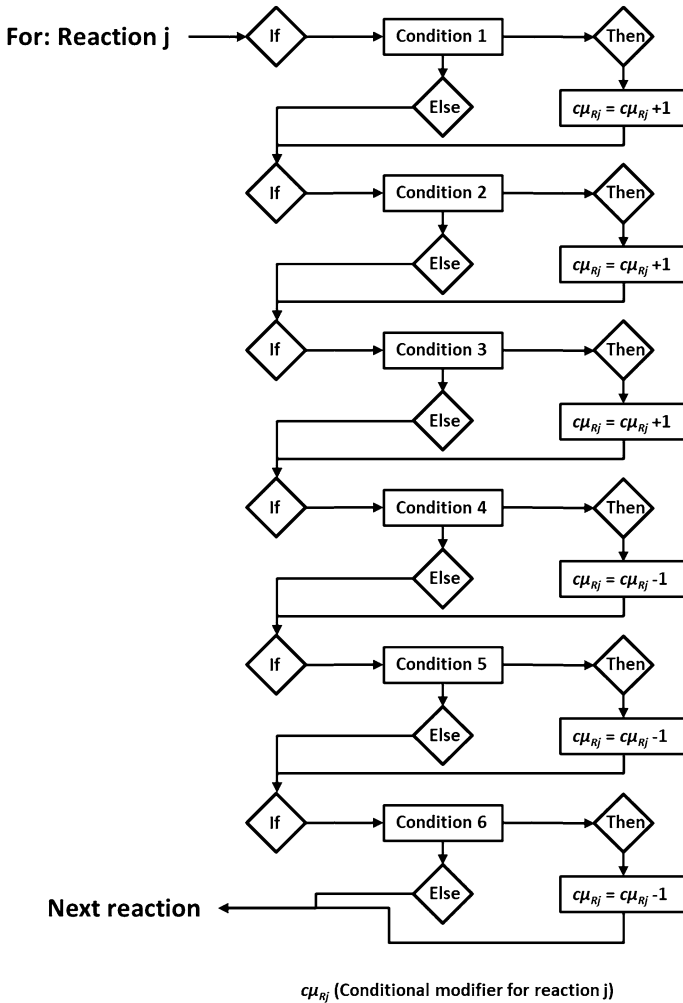
Fig. 7 A delineation of the logical process by which the Conditional Modifier (cμ) value for each reaction can be arrived at by assessing the significance of the reaction in the network relative to the target chemical species

$$D_e i = \sum_{j=1}^{n} \left(\langle k_j \rangle \cdot c\mu_j \cdot \Delta k_{T_e j} \right)$$

where $D_e i$ is the electron energy demand for the ith phase, n is the number of reactions in the ith phase, k_j is the mean rate coefficient for the jth reaction over the range electron energy values (1–10 eV), $c\mu_j$ is the conditional modifier (see Fig. 7) for the jth reaction, and $\Delta k_{T_e j}$ is the electron energy susceptibility of the jth reaction.

Each phase graph is specified as follows:

$$G = \text{full graph of reactions}$$

G_1 = initial species, and direct successors, and all edges between them = $(V_1, E_1) = G$—all that is not G_1

$$G_2 = (V_2, E_2) = (V_1 + V_1 \ successors, E_1 + E_1 \ successors)$$
$$G_3 = (V_3, E_3) = (V_2 + V_2 \ successors, E_2 + E_2 \ successors)$$

Generally :

$$G_{i+1} = (V_{i+1}, E_{i+1}) = (V_i + V_i \ successors, E_i + E_i \ successors)$$

The "demand" for other variables can also be estimated by slightly modifying the above equation, for example to estimate the "temperature demand" D_e could be replaced by D_T (the gas temperature demand). The equation can also be modified to estimate the "demand" for other chemical species by our species of interest, which could be used as a means of identifying key reactions that would be useful to catalyse in plasma reactors.

Testing the Algorithm on the Plasma Chemical Graph

The above logical statements were incorporated into an algorithm programmed in the python programming language using the networkx library. The script was then run to generate a series of "predictions" using the algorithm incorporating the above methodology. The resulting "relative electron energy demand" values for the optimum selectivity of O· and O_3 generated using this script are shown in Fig. 8. The relative electron energy demand values on the charts in Fig. 8 could be met by electric field strengths that could be generated from a high voltage–power supply, for example. The time lengths of the phase are identical to each other, as an acceptable method for estimating a optimum phase times from the graphs of each of these phases could not be found within the scope of this work.

We have called the above methodology: Optimal Condition Approaching via Reaction-In-Network Analysis (OCARINA). A 0D time dependent simulation of successive 100 ns electron energy pulses in a 4:1 N_2:O_2 mixture was run to test the OCARINA predictions for O· and O_3 by setting the magnitude of each peak in proportion to the electron energy demand for its corresponding phase. The electron energy for any phase with negative
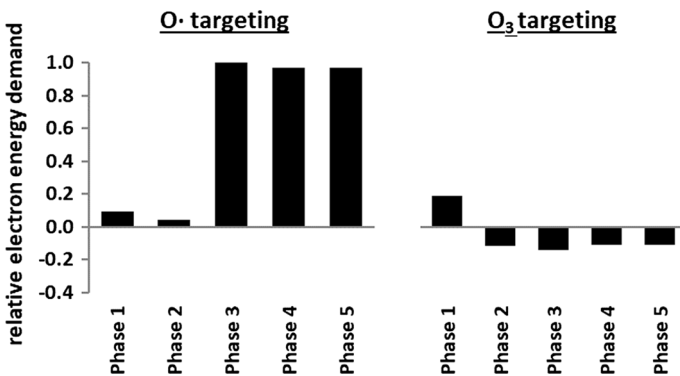

Fig. 8 Two different "relative electron energy demands" for two different target species, O· and O_3, generated by the OCARINA algorithm. The negative values represent the extent to which electron energy is predicted to adversely affect this phase of the process. The time duration of each phase is not estimated by the algorithm

electron energy demand values was set to the lowest electron energy value of 0.038 eV, the mean molecular kinetic energy at room temperature and pressure. The reactions and rate coefficients from Sakiyama et al. [20] were used, with the initial concentrations of N_2, O_2 and electrons set to 1.93×10^{25} m^{-3}, 4.82×10^{24} m^{-3} and 1×10^{21} m^{-3} respectively. The initial concentrations of all other species were set to 10 m^{-3}, and all hydrogen containing species were excluded from the simulation. The reaction rates and species concentrations were computed in 1×10^{-13} s timesteps for a total of 0.5 μs. The resulting concentrations of O· and O_3 for each targeted electron energy regime shown in Figs. 9 and 10 (respectively) clearly suggest preferential formation of O· or O_3 under their corresponding targeted electron energy regimes.

Discussion

Agreement with Literature Values

In addition to the simulated results shown above, the energy level regime predicted by the O_3 targeted run of OCARINA also agrees with a recent study by Seri et al. [24] of the effect of different waveforms on ozone generation. The study observed that the best efficacy of energy input for ozone generation was achieved with relatively large spaces between on times to prevent ozone destruction reactions from occurring. This also agrees with the general understanding of ozone generation in the field [10]. However, it is not clear from a literature summary [10] whether the energy inputs predicted by OCARINA for the synthesis
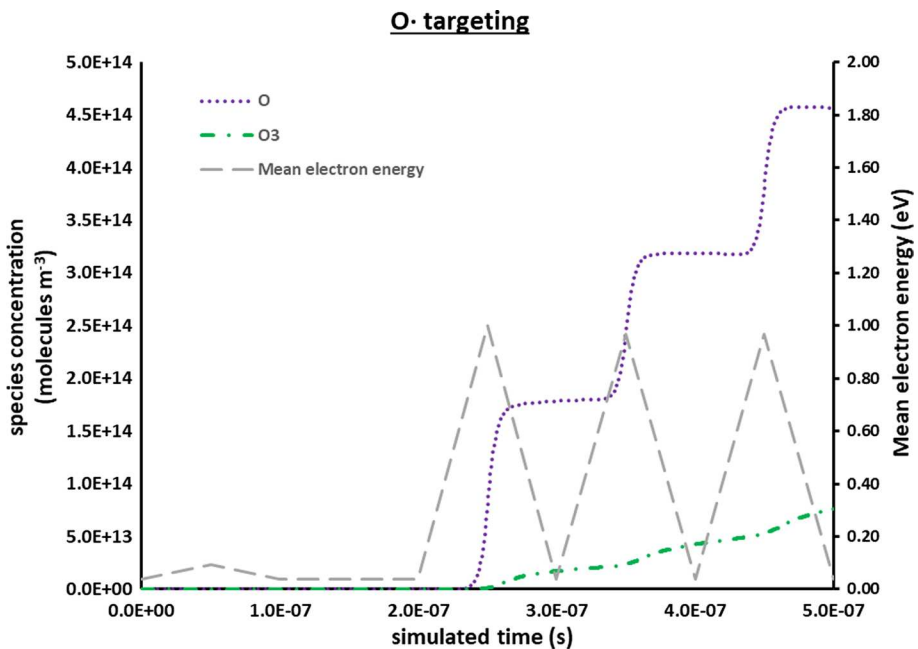

Fig. 9 Concentration of atomic oxygen and ozone over time with simulated electron energy pulse magnitudes set in accordance with the results of the OCARINA graph algorithm set for O· targeting
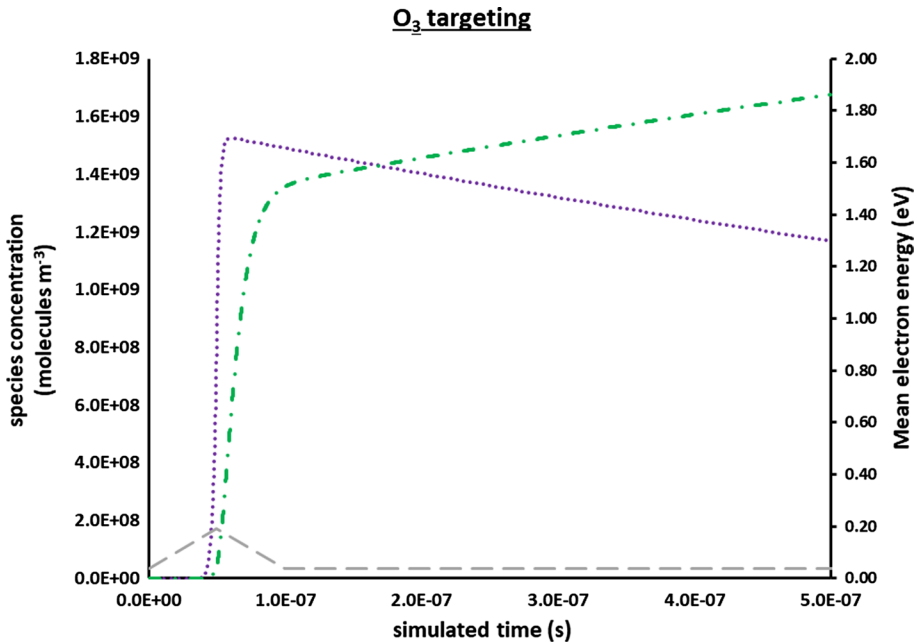
Fig. 10 Concentration of atomic oxygen and ozone over time with simulated electron energy pulse magnitudes set in accordance with the results of the OCARINA graph algorithm set for O_3 targeting

of many other species would be optimal or not. It would certainly be beneficial to investigate this both computationally and experimentally to assess if these energy regimes are more optimal or not, and if not, what alternative conditional statements or methodologies might improve this.

Utility of Graphs for Plasma Chemical Reaction Engineering

Whether for visualising large amounts of information, or for mathematical operations for making predictions, it is clear that graph theory has useful applications to plasma chemical reaction engineering. It is also clear that there must be limitations of these applications. The accuracy of using graph theory as a tool for plasma chemical reaction engineering depends on the following: how accurately the graph itself represents the system, how effectively the graph is operated on for a given application, and how the results of such operations are interpreted.

How Accurately the Graph Represents the System

A clear limitation on how accurately the graph method used in this paper can represent reality is the use of the Δk_{T_e} and $\Delta k_{T_{gas}}$ values, which reduce the relationship between the rate coefficient and one of its independent variables to a single numerical value. Node and edge weights are single numerical values in many graph algorithms, and the use of single numerical values was therefore a logical starting point in this study. However, there is no real limit on how many values can theoretically be attributed to a node

or an edge, and it may therefore be more accurate to increase the "resolution" of the graph edges by, for example, having Δk_{T_e} and values for every 1 eV interval, instead of a single values to approximate the whole range. This would also address the issue of what to do when the rate coefficients do not conveniently simply increase over the range of 1–10 eV, but instead increase and decrease at varying intervals. Increasing the number of attributes of the edges and nodes will inevitably make the graphs and graph operations more complex, and it is therefore essential that care must be taken to avoid the same impractical situation described in the introduction where every additional variable increases the computing time exponentially. That said, one of the strengths of using single numerical values for edge or node attributes are that these can be visualised as a single visual effect. It will certainly be challenging to visualise networks where each edge has ten different Δk_{T_e} values represented visually, by colour for example.

Furthermore, there are other key variables which may need to be included in order for the graph to more accurately represent reality, such as the effect of concentration on reaction rate. It should be noted that the incorporation of concentration into the graph does present problems, in that concentration and reaction rate change over time in any process, and are also dependent on each other, meaning it may not be possible to incorporate concentration, even starting concentrations, and still represent the whole system with a single graph. It may also not be necessary, as it may be possible to incorporate the effects of concentration into the operations that are performed on the graph instead.

How Effectively the Graph is Operated on for a Given Application

In the case of operating on the graph for the purposes of visualisation, some compromises must be made. For example, logarithmic values of edge weights must be used to visualise the rate coefficients for all reactions, many of which are already mean values of a large range, and whilst this allows a clear contrast between the extremely slow, and extremely fast, reactions, it makes it hard to distinguish between reactions which may differ by only two orders of magnitude, as a reaction would appear almost the same as one a hundred times faster or slower at first glance. How much of a problem this is depends on the purpose of the graph operation, and so it follows that clarity of purpose of any graph visualisation is key in deciding which compromises can and cannot be made. For example, if the purpose of a graph is to identify slow reactions which could play a key role if their rates could be increased via catalysis, then it is clear that visualisation of reaction rates takes priority, and the slowest reactions must be easily distinguishable, with the need to distinguish between progressively less slow reactions once the very slowest reactions have been considered.

For the more mathematical/algorithmic operations on the graphs, compromises for the sake of human visualisation are not important, as the very reason for using computable algorithms is to avoid having to make such compromises. However, the effectiveness of these operations depend on the quality of the graph, as described in the previous section. The selection of appropriate logical statements, in the case of OCARINA, is also critical, and as mentioned previously, should be modified when flaws in the logical statements are found. New logical statements can also be added as required.

Interpretation of Results

Care should be taken when interpreting the results of a graph application to plasma chemical engineering, that a relatively accurate graph or graph operation is not rejected due to a relatively inaccurate interpretation of the results.

In the case of the OCARINA algorithm, for example, it might be tempting to think of the electron energy demand values as a measure of how much electron energy should be provided at each point, but the numbers are more of a response to the question "how important is electron energy for the formation of the target species at this point?". This does not mean electron energy is the most important factor, although it is clear from the results in Figs. 9 and 10 that applying the output of OCARINA directly to the electron energy favours the target species in these two specific cases. It is likely that in both cases electron energy has a significant effect relative to other factors. In other cases, however, the electron energy demand may also be high, but providing a high electron energy alone may not produce significant results, as other factors such as temperature, or the presence of a particular species, may outweigh any effect of electron energy.

In the case of the graph visualisation, there is clearly a more direct connection between the accuracy of the graph itself and how it is interpreted. Some operations will need to be performed for the visualisation to be effective, and correct interpretation of the visualisation must rely on being able to know which operations have been performed. If a logarithmic scale has been used for size, for example, including the original values as labels will aid the user in correct interpretation of the visualisation. Similar verification should be possible for all aspects of the visualisation, including colour, text size, node size, etc. A useful feature of the more mathematical operations on the graphs would be a similar ability to check what the operations have been doing when analysing their results, as opposed to a simple "black box" output. Some crossover from the visualisation application might be necessary for this.

Alternatives to the Methods Described in this Work

The most reliable and well-known method for optimisation of conditions is still the running of repeated simulations over ranges of variables [25]. As mentioned above, this is meticulously effective in small variable spaces but computationally impractical as the number and ranges of variables increase.

Neural networks have been used by some researchers in order to identify experimental conditions [26] and even to optimise an existing plasma chemical reaction system [27], though there are no published articles at the time of writing that have applied this to plasma chemical kinetic reaction systems in order to infer which conditions might be most optimal.

Two notable papers have used graph theory applied to plasma chemical kinetics [18, 19], but it should be made clear that the methods described in this paper differ from previous work principally in the emphasis on the utility of graph visualisation as a tool in itself, in the addition of reaction-nodes, and in the use of reaction pathways for algorithmic operations. The OCARINA algorithm is also not previously published (at the time of writing).

Areas for Further Development

As mentioned above, the effectiveness of any of these applications all ultimately depends on how accurately the graph represents the system. Whilst 0D simulations can be useful approximations in cases where the plasma can be treated as homogeneous, 1D time dependent simulations which include movement of particles in response to electric fields and diffusion gradients are much more useful for DBD modelling. A key next step would be the development of a graph of a 1D system, which would necessarily include some measure of species transport and electric field effects. Whilst this would undoubtedly be challenging, it would also offer significant advantages, especially given the much higher computational requirements of 1D simulations. Development of a graph of a 1D model could also open up the possibility of 2D and even 3D model graphs.

The utility of the graph visualisation may be limited by the size and resolution of the monitor, as mentioned previously, especially with larger numbers of plasma chemical reactions or graphs of higher dimensional models. Trialling plasma chemical graph visualisations in a studio specifically designed for visualisation of large graphs would be a useful next step in further investigating ways in which it can be used. Another interesting application may be time evolution of the graph visualisation as the concentrations are changed over the course of a 0D simulation, which could also be applied to time dependent chemical kinetic simulations. This may be especially useful in identifying the potential role of minor reactions or species as the concentrations are changed.

A clear limitation of the OCARINA method is that it cannot yet estimate the actual duration (in units of time) of the phases. Whilst it is unlikely that a precise prediction of the exact times for which the electron energy should be raised or lowered, some idea of how the durations of the phases compare to one another would be advantageous. Such a comparison will most likely come from some statistical treatment of reciprocals of the rate coefficients, the reciprocal of the mean rate coefficient would at least make a good starting point for future studies. Additionally, as mentioned previously, the OCARINA algorithm can only work if the change in the rate coefficients with temperature or electron energy can be approximated as linear. This is most certainly not the case with some reaction cross sections, which increase and decrease over greater electron energy ranges. There may be some patterns even in these cross sections which can also be approximated, or it may be that a multi-dimensional graph may need to be generated. Whilst both of these potential solutions are possible, they will each need in-depth study before they can be applied to methods outlined above.

Improvements in optimisation of efficiency or yield can be made by identifying where process input energy and materials are going, or where new process outputs might come from. The graphs shown above could certainly be adapted to include Gibbs free energy values as attributes of the graph edges.

Furthermore, a machine-learning feedback approach could be integrated with the applications such as the OCARINA algorithm. The rates of important reactions, for example, could be increased or decreased to generate new graphs for re-analysis, generating other important reactions, generating a new graph, etc. This would, once again, require in-depth study to be developed and tested and was certainly well outside the scope of this study.

All the above points to increasing the complexity of the graphs, and of the applications. Nevertheless, there is enough potential in these methods to make it worth the necessary time and effort. Whilst each of the applications need to be further investigated

and developed, the results of this study suggest it may be possible to rapidly hone in on favourable conditions for plasma processes (based on analysis of graphs of the chemical kinetic reaction data) which can then be tested via simulation or experiment. This could reduce the number of experiments/simulations that it may be necessary to try in order to increase the efficiency of a plasma chemical process, thereby increasing the rate of development of plasma chemical processes as a whole. These potential benefits are a good pretext for further investigation.

Finally, the methods described in this work could most likely be adapted for other chemical systems aside from plasmas to identify ways to also increase their effectiveness. It may not be an advantage to use the methods outlined in this paper if the chemical system is not too complex to prevent it being conveniently assessed without computational assistance, but it is potentially applicable to any chemical system where stimuli can be applied to the system to affect some or all of the chemistry (provided this can be modelled mathematically). In the case of plasma, the external stimulus is the electron energy provided by the plasma power supply. Other stimuli could be rapid heating/cooling in a thermochemical system, electromotive force in an electrolysis cell, or electromagnetic radiation in the form of photons in a photochemical system, to give a few examples. A stimulus could also be the addition or removal of different chemicals from the system. In addition to complex chemical systems with multiple parallel reactions, it could also be modified to identify potential optimal conditions for relatively long reaction sequences such as polymerization processes, as each step of a polymerization process has specific conditions that must be optimised in favour of the precursors of the following stage, and to ensure that net destruction of the polymer does not occur at any of the stages. In theory there is no limitation on the number of phases that could be included in the algorithm, so any long/multi-step process could include one or more phases for each process step. Other applications of these methods may also present themselves, and further investigation of these in addition to the above should most certainly be an objective of future studies.

Conclusions and Outlook

Some graph theory techniques applied to plasma chemistry can be used to conveniently and expediently examine a plasma chemical reaction system. Visualisation of a graph of the plasma chemical kinetic data can be used to simultaneously display which reactions/reaction networks are the slowest and fastest and which are and are not affected by electron energy or temperature, and to what extent, with the crucial factor of being able to see how all these reactions and species are related to each other. Whilst the exact quantitative accuracy of this kind of visualisation has its limitations, a large amount of qualitative information useful to plasma chemical reaction engineering can still be gained from it.

Graph theory operations, such as Dijkstra's shortest path algorithm, operated on the plasma graph can be used in combination with the visualisation, as well as other techniques, to analyse the plasma chemical kinetic scheme without needing to run simulations.

The OCARINA algorithm was developed and shown to predict, at least for ozone formation, points in the reaction chronology where it might be most useful to increase or decrease the electron energy (even if this might not be possible in practice).

Use of all the methods above, applied to an air plasma reaction set from [20], suggests that an optimal electron energy regime for ozone formation is a high amount of electron energy in the initial phase, and then no electron energy for the successive phases. This

corresponds with evidence in the literature [10, 24]. Use of the above methods also suggests that ozone destruction reactions occur much more rapidly at high electron energies, and at high temperatures (Fig. 3), which is also well known. A 0D simulation using the output of the OCARINA algorithm for targeted formation of O· and O_3 indicated preferential formation of these species using the electron energy values suggested by the algorithm.

This all suggests that the application of graph theory to plasma chemical reaction engineering can be a useful tool in the field of plasma chemistry, and potentially any field where similar systems exist.

References

1. Adamovich I et al (2017) The 2017 Plasma Roadmap: low temperature plasma science and technology. J Phys D Appl Phys 50(32):323001
2. Brandenburg R et al (2019) White paper on the future of plasma science in environment, for gas conversion and agriculture. Plasma Process Polym 16(1):1–18
3. Call AV, Holmes TD, Yanallah K, Desai PD, Zimmerman WB, Rothman RH (2019) Improving the efficiency of high-temperature electrolysis of carbon dioxide in a solid oxide cell. ECS Trans 91:2623–2630
4. Wright A et al (2018) Dielectric barrier discharge plasma microbubble reactor for pretreatment of lignocellulosic biomass. AIChE J 64(11):3803–3816
5. Butterworth T, Elder R, Allen R (2016) Effects of particle size on CO_2 reduction and discharge characteristics in a packed bed plasma reactor. Chem Eng J 293:55–67
6. Butterworth T, Allen RWK (2017) Plasma-catalyst interaction studied in a single pellet DBD reactor: dielectric constant effect on plasma dynamics. Plasma Sources Sci Technol 26(6):065008
7. Weltmann KD et al (2019) The future for plasma science and technology. Plasma Process Polym 16(1):1–29
8. Wende K, Lackmann J-W, Jablonowski H, Stapelmann K, von Woedtke T, Bekeschus S (2018) Can we achieve selectivity in plasma medicine? Clin Plasma Med 9:43
9. Lieberman MA, Lichtenberg AJ (2005) Principles of plasma discharges and materials processing, 2nd edn. Wiley, New York
10. Fridman A (2008) Plasma chemistry. Cambridge University Press, Cambridge
11. Whitehead JC (2016) Plasma-catalysis: the known knowns, the known unknowns and the unknown unknowns. J Phys D Appl Phys 49(24):243001
12. Doyle SJ, Lafleur T, Gibson AR, Tian P, Kushner MJ, Dedrick J (2017) Enhanced control of the ionization rate in radio-frequency plasmas with structured electrodes via tailored voltage waveforms. Plasma Sources Sci Technol 26(12):125005
13. Van Gaens W, Bogaerts A (2013) Kinetic modelling for an atmospheric pressure argon plasma jet in humid air. J Phys D Appl Phys 46:275201
14. Wilson RJ (1996) Introduction to graph theory, 4th edn. Prentice Hall, Harlow
15. Derrible S, Kennedy C (2011) Applications of graph theory and network science to transit network design. Transp Rev 31(4):495–519

16. Ware C (2013) Information visualisation: perception for design, 3rd edn. Morgan Kaufman-Elsevier, Amsterdam
17. Temkin ON, Bonchev DG (1992) Application of graph theory to chemical kinetics: part 1. Kinetics of complex reactions. J Chem Educ 69:544
18. Sakai O, Nobuto K, Miyagi S, Tachibana K (2015) Analysis of weblike network structures of directed graphs for chemical reactions in methane plasmas. AIP Adv 5(10):107140
19. Mizui Y, Kojima T, Miyagi S, Sakai O (2017) Graphical classification in multi-centrality-index diagrams for complex chemical networks. Symmetry 9(12):309
20. Sakiyama Y, Graves DB, Chang H-W, Shimizu T, Morfill GE (2012) Plasma chemistry model of surface microdischarge in humid air and dynamics of reactive neutral species. J Phys D Appl Phys 45(42):425201
21. Jacomy M, Bastian M, Heymann S et al (2009) Gephi: an open source software for exploring and manipulating networks. In: Proceedings of international AAAI Conference on Web and Social Media
22. Hu Y (2006) Efficient, high-quality force-directed graph drawing. Math J 10(1):37–71
23. Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1:269–271
24. Seri P et al (2019) Influence of the voltage waveform's shape and on-time duration on the dissolved ozone produced by a DBD bubble reactor. Plasma Sources Sci Technol 28:035001
25. Snoeckx R, Zeng YX, Tu X, Bogaerts A (2015) Plasma-based dry reforming: improving the conversion and energy efficiency in a dielectric barrier discharge. R Soc Chem 5:29799–29808
26. Gonoskov A, Wallin E, Polovinkin A, Meyerov I (2019) Employing machine learning for theory validation and identification of experimental conditions in laser-plasma physics. Sci Rep 9:7043
27. Istadi, Amin NAS (2006) Hybrid artificial neural network genetic algorithm technique for modeling and optimization of plasma reactor. Ind Eng Chem Res 45:6655–6664