Knowledge-based probabilistic representations of branching ratios in chemical networks: The case of dissociative recombinations

Sylvain Plessis, Nathalie Carrasco, Pascal Pernot

Knowledge-based probabilistic representations of branching ratios in chemical networks: the case of dissociative recombinations

Sylvain Plessis

Univ Paris-Sud, Laboratoire de Chimie Physique, UMR 8000, Orsay, F-91405

Nathalie Carrasco

Laboratoire Atmosphères, Milieux, Observations Spatiales, Université de Versailles Saint-Quentin, UMR 8190, 91371 Verrières le Buisson cedex, France

Pascal Pernot[a)]

Univ Paris-Sud, Laboratoire de Chimie Physique, UMR 8000, Orsay, F-91405 and CNRS, Orsay, F-91405

Experimental data about branching ratios for the products of dissociative recombination of polyatomic ions is presently the unique information source available to modelers of natural or laboratory chemical plasmas. Yet, because of limitations in the measurement techniques, data for many ions are incomplete. In particular, the repartition of hydrogen atoms amongst the fragments of hydrocarbons ions is often not available. A consequence is that proper implementation of DR processes in chemical models is difficult, and many models ignore invaluable data. We propose a novel probabilistic approach based on Dirichlet-type distributions, enabling modelers to fully account for the available information. As an application, we consider the production rate of radicals through dissociative recombination in a ionospheric chemistry model of Titan, the largest moon of Saturn. We show how the complete scheme of dissociative recombination products derived with our method dramatically affects these rates in comparison with the simplistic H-loss mechanism implemented by default in all recent models.

[a)]Electronic mail: pascal.pernot@u-psud.fr

I. INTRODUCTION

Dissociative Recombination (DR) of positive ions with free electrons is an important process in highly rarefied and ionized media such as interstellar clouds,[1–4] upper planetary atmospheres,[5,6] and chemical plasmas.[7–9] DR is the major sink of electrons in ionized media,[10] it is a major pathway for the consumption of ionic species, and it is also an important source of energetic molecules and radicals which can influence significantly the chemical complexification of the system.[8,9] The formation of complex neutral species observed in those media is considered to be largely due to DR.

Models of ion-neutral chemistry generally include DR processes through the reaction rate constant

$$\alpha(T_e) = \alpha_0 \times (T_e/T_0)^{-\beta} \tag{1}$$

where α_0 is the rate constant at a reference electron temperature T_0 (typically 300 K), and T_e is the electron temperature of interest. When more than one product channel is accessible, the partial rate for channel i is $\alpha_i(T_e) = b_i \times \alpha(T_e)$, where $\{b_i, i = 1, N\}$ are the *branching ratios* of the reaction. It is important to note that the global DR rate $\alpha(T_e)$ and the branching ratios $\{b_i\}$ are typically derived from different experiments.

Despite many experimental results in the past years, there are still noticeable difficulties in the modeling of DR processes in complex chemical networks.[11] For instance, the rate of the process depends on the initial state of the ion (electronic and/or ro-vibrational). In non thermalized media, as planetary ionospheres, the temperature of electrons is not the only parameter controlling the DR rate and the temperature of ions should also be considered.[10,12] However, within large families of ions, DR rates typically vary within an order of magnitude only and might not be the major source of uncertainty in model predictions.

The main issue in DR modeling is certainly the lack of knowledge about branching ratios. The difficulty of the detection and quantification of neutral fragments, in various isomeric and/or electronic configurations, contributes to this state of affairs. In fact, measured branching ratios have often proved to be counterintuitive and at odd with the predictions of earlier theories.[13–15] Moreover, data about temperature effects on branching ratios (through collision and/or internal energy) are still very sparse.[11,16,17]

E. Herbst and coworkers have extensively explored the impact of DR branching ratios on the chemical composition of interstellar clouds.[2,18–22] Based on the theory of Bates,[13–15]

they implemented various branching ratios scenarii.[2,18,20,21] Herbst and Klemperer[18] first assumed in their model that all products resulting from exoergic channels were produced with equal efficiency, *i.e.* equal branching ratios. Following Bates,[14] Millar *et al.*[21] based new estimations of branching ratios on the distribution of charges in the parent ion with mitigated success. More recently, as discrepant experimental data about the importance of the H-loss channel became available, they used two alternative scenarii where they attributed a fraction of 0.30 or 0.05 to the H-loss channel, with equipartition amongst the other channels.[2] An important conclusion of these studies is that branching ratios have a strong impact on steady-state concentrations of neutral species in interstellar clouds. Without a doubt, this can be generalized to other systems where DR drives the fate of ions. Another conclusion is that prediction of DR branching ratios is not yet a routine task.

For the RD of H_3^+ Strasser *et al.*[23] proposed a statistical theory which reproduced nicely existing fragmentation data. However, the transferability of this model to much heavier ions is not straightforward,[11] and has not been attempted, to our knowledge.

In absence of simple predictive theory, high-level *ab initio* calculations would be the best choice. However, the implication in the process of highly excited electronic states, of many open electronic states and the necessity to explore complex potential energy surfaces make it a daunting task. In consequence, accurate quantum or semi-classical calculations are presently limited to very small ions.[17,24–27]

Considering the lack of tractable predictive theories for DR branching ratios of polyatomic ions, experimental data are at the moment the only source of information available to chemical plasma modelers. Several experimental groups made considerable advances in the past decades to produce reference data.[11,16,17] Despite these efforts, we still have very sparse data for molecules with more than four heavy atoms.[16] Even for smaller species, information is often incomplete: for hydrocarbon ions, the distribution of H atoms between fragments is not accessible, due to present experimental limitations.[16] This is illustrated in Fig. 1, summarizing the state of knowledge for the products of a middle-sized ion, $C_4H_9^+$. Also, the isomeric form of the parent ion and the isomeric form and electronic state of the neutral products are generally not known.[20] This lack of information about the chemical identity, and therefore the reactivity, of the products of DR can introduce substantial biases in the predictions of chemical models.[28]

It is therefore of utmost importance to be able to incorporate all the knowledge about

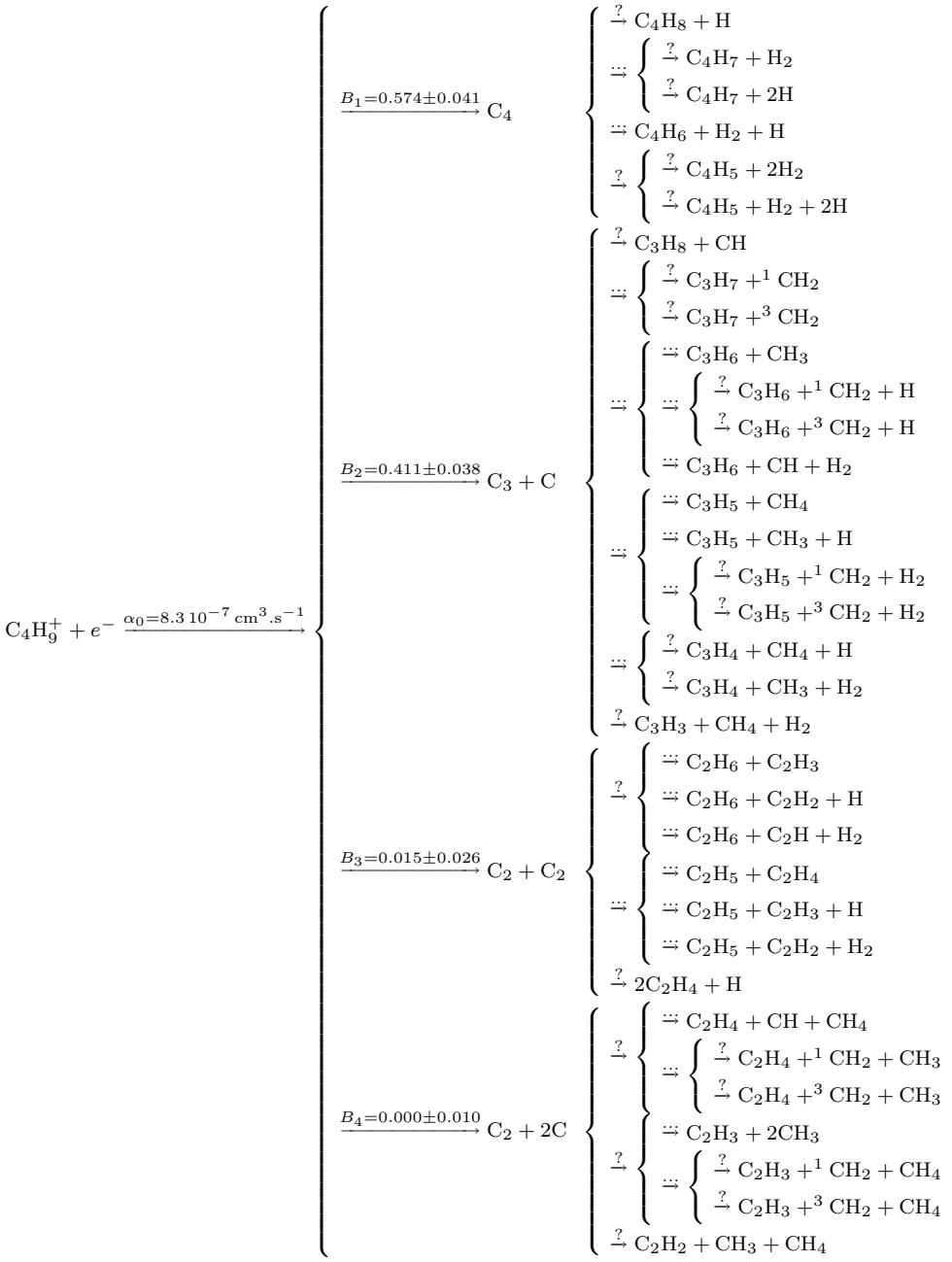

Figure 1. Tree structure of the putative products for the DR of $C_4H_9^+$. The first level implements available experimental data about the carbon fragmentation pattern;[16] deeper levels are built according to a list of exoergic channels and the list of species involved in the chemical system of interest. The deep tree structure presented here is obtained according to a heuristic explained in the text.

branching ratios into chemical models. This is not always straightforward, notably when the chemical formulas of the fragments are not fully elucidated.

The present paper addresses the problem of the uncertainty in experimental branching ratios and proposes solutions to enable a realistic implementation of DR in chemical plasma modeling. Probabilistic representations have been designed to account for the various uncertainty patterns identified through an extensive review of the literature (Section III). In order to deal with incomplete branching ratios data, an innovative representation is based on Nested Dirichlet distributions.[29,30] This work extends the previous study by Carrasco and Pernot,[31] about the representation of uncertain branching ratios of ion-neutral reactions, where the necessity of a good representation of correlations between these parameters was validated for uncertainty propagation and sensitivity analysis.[28] The theoretical developments are presented in Section II A.

The proposed methodology is illustrated on a series of representative ions and applied to the complex ionospheric chemistry of Titan (see Section III). In fact, the extreme complexity of DR products formation is in sharp contrast with the fact that none of the ion-neutral coupled models for Titan's ionosphere integrates the multi–pathway nature of this process.[32–38] With only few exceptions, the present paradigm is to consider the H-loss channel as the only pathway for all H-bearing ions

$$MH^+ + e^- \rightarrow M + H \tag{2}$$

which is at best a severe approximation. Recent experimental studies have clearly shown that DR could not only break bonds between heavy atoms efficiently, but also break more than one bond.[11,16,17] As our probabilistic approach enables to implement a DR scheme accounting for all plausible pathways, we tested it in a ionospheric chemistry model of Titan,[39,40] comparing the results to those of the H-loss approximation.

II. PROBABILISTIC REPRESENTATIONS OF BRANCHING RATIOS

A. Methods

The methodological framework is based on the representation of uncertain parameters by probability density functions (pdf). These pdfs are then used as inputs in chemical models and processed with Monte Carlo Uncertainty Propagation (MCUP),[41,42] as follows.

1. *Monte Carlo Uncertainty Propagation*

For a model F having as input a vector of uncertain variables $\mathbf{X} = (X_1, .., X_n)$,

$$Y = F(\mathbf{X}) = F(X_1, ..., X_n) \tag{3}$$

the probability density on the values of Y is obtained by the Markov integral

$$p\left(Y = y\right) = \int d\mathbf{X}\, \delta\left(y - F(\mathbf{X})\right)\, p\left(\mathbf{X}\right) \tag{4}$$

where $p\left(\mathbf{X}\right)$ is the joint pdf of the input variables. Equation (4) can rarely be solved analytically. Instead, we use the Monte Carlo approach which is defined as a standard by the Supplement 1 to the "Guide to the expression of Uncertainty in Measurement Data".[41] The principle is simple: one generates a set $\{x_{1,i}, \ldots, x_{n,i}\}$ of $i = 1,\ N_{run}$ random draws from the n-dimensional inputs pdf $p\left(\mathbf{X}\right)$, and for each element of this set, one evaluates the value of the model output, generating a representative sample of the output pdf $p(Y)$. This sample $\{y_i\}$ is then used to get statistical estimates of quantities of interest.

MCUP has been frequently used in chemical modeling to assess the effect of reaction rates uncertainty on the concentrations of species.[43–55]

2. *Probability density function of the uncertain variables*

The inputs for different reactions are typically independent,[56] which enables to factorize the inputs pdf reaction-wise

$$p\left(\mathbf{X}\right) = \prod_r p\left(\alpha_0^{(r)}, \beta^{(r)}, b_1^{(r)}, \ldots, b_{n_r}^{(r)}\right). \tag{5}$$

For a given DR reaction (r), branching ratios and reaction rates are measured independently, branching ratios are correlated by the sum-to-one constraint, and the rate parameters might also be correlated by the fitting procedure, leading to a further factorization

$$p\left(\mathbf{X}\right) = \prod_r p\left(\alpha_0^{(r)}, \beta^{(r)}\right) p\left(b_1^{(r)}, \ldots, b_{n_r}^{(r)}\right). \tag{6}$$

When treating individual DR processes, we will simplify the notations by omitting the (r) superscript. Our aim in this paper being to study branching ratios and their uncertainties, the DR rate coefficients were fixed at their nominal value, *i.e.*

$$p\left(\mathbf{X}\right) = \prod_r p\left(b_1^{(r)}, \ldots, b_{n_r}^{(r)}\right). \tag{7}$$

In the following, we consider the problem of finding adequate representations for the probability density function of individual sets of branching ratios $p(b_1, \ldots, b_n)$, which are consistent with available information and constraints.

B. Space of branching ratios and graphical representations

Branching ratios of a reaction form a *composition*, such as $0 \leq b_i \leq 1$ and $\sum_{i=1}^{n} b_i = 1$.[57] The sum-to-one is a major property to be preserved in the representation of branching ratios in chemical networks.[28,31] It introduces a strong negative correlation between the b_i, which cannot be ignored in statistical operations.

The points of a n-dimensional set of branching ratios lie in a subspace of dimension $n-1$ called a *simplex* (see Fig. 2(a)). In a 3-dimensional(3D) space, the simplex is an equilateral triangle. The corner labelled b_i corresponds to the composition $(b_i = 1, b_{j\neq i} = 0)$. In this 3D case, probability density functions or representative random samples can be conveniently plotted in a ternary graph, where each point corresponds to an element of the sample (Fig. 2(b)). For spaces of higher dimensions, we rather use alternative representations such as parallel graphs where the coordinates of each point are drawn on separate parallel vertical axes and linked by a line (Fig. 2(c)). This kind of graph can convey a lot of information, but makes it difficult to appreciate, for instance, the sampling uniformity in the simplex.

A third representation used in the following is built on the parallel graph basis, but displays boxplots of the individual one-dimensional (marginal) densities instead of line sets (Fig. 2(d)). It is quite convenient to use for comparison of samples, but all information about correlations between the branching ratios is erased.

Another point to have in mind when appreciating these graphs is that the one-dimensional marginal densities from a uniform pdf in a simplex are *not* uniform. It is enlightning to compare the information conveyed by the ternary graph of Fig.2(b) and the boxplot graph of Fig. 2(d), obtained from identical samples.

C. Elicitation of probability density functions for branching ratios

The design of probability density functions from available information is called *elicitation*.[58] It is a key stage in the MCUP process, for which the principle of maximum entropy (PME)

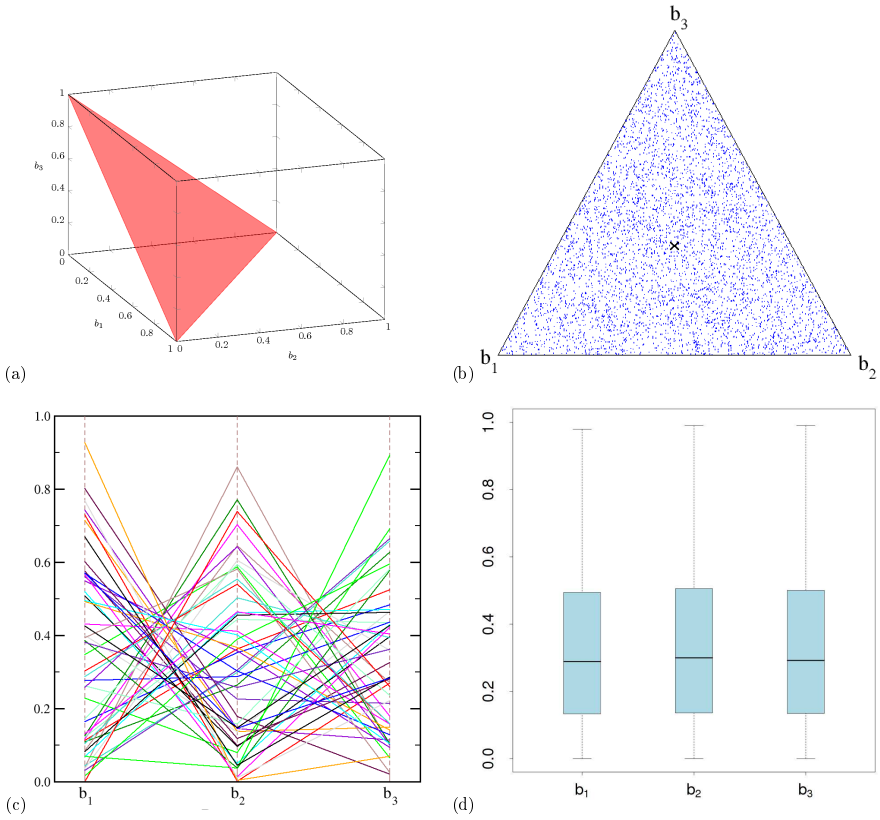

Figure 2. Visualisation of the simplex in a 3-dimensional space (b_1, b_2, b_3) (a) and alternative representations of a sample of the uniform distribution within this simplex: (b) ternary graph; (c) parallel graph and (d) boxplots for the one-dimensional marginal densities $p(b_i),\ i = 1, 3$.

provides a reference tool.[59] As there are typically an infinity of pdfs obeying to a limited set of constraints, the PME enables to disambiguate the problem in providing a unique pdf. Standard elicitations based on the PME have been defined, mostly for single variable problems.[41] For instance, if an average value and a standard deviation are available, the reference pdf is a Gaussian/Normal distribution. For correlated parameters, the multivariate Gaussian is obtained from a vector of best estimates and a strictly positive variance/covariance matrix. With minimal constraints of positivity and sum-to-one, the PME defines the Dirichlet distribution.[59–61]

A review of DR reactions for ions included in state-of-the-art models of Titan's ionosphere was performed in order to identify representative uncertainty/information patterns for branching ratios. These cases range from sets of preferred values and standard deviations, to intervals, to... no information, *i.e.* unidentified products or unknown branching ratios for known products. All these cases can be successfully handled through distributions in the Dirichlet family and combinations thereof. We consider successively the cases where we have (i) preferred values and the associated uncertainties; (ii) intervals; (iii) no information; (iv) no numerical information but an ordering scheme; and (v) sets of incomplete or heterogeneous data. A decision tree (Fig. 3) has been designed to help the user to choose the appropriate distribution. Sampling methods for these distributions are described in the Appendix.

1. Preferred values and uncertainties: Dirg representation

To sample uncertain branching ratios for ion-molecule reactions, Carrasco and Pernot[31] used the standard Dirichlet distribution (Diri)

$$(b_1, \ldots, b_n) \sim \mathrm{Diri}(\mu_1, .., \mu_n \,;\, \gamma), \tag{8}$$

where μ_i is the preferred value for branching ratio b_i, and γ is a global accuracy factor optimized to reproduce, at best in the least-squares sense, the global relative uncertainty r on the branching ratios proposed in reference databases[62,63]

$$\gamma = \frac{4}{r^2}\left(\frac{\sum_i \mu_i(1-\mu_i)}{\sum_i \mu_i\sqrt{\mu_i(1-\mu_i)}}\right)^2 - 1. \tag{9}$$

The Diri distribution represents the fluctuations of quantities independent of each other, under the condition that their sum remains fixed.[60]

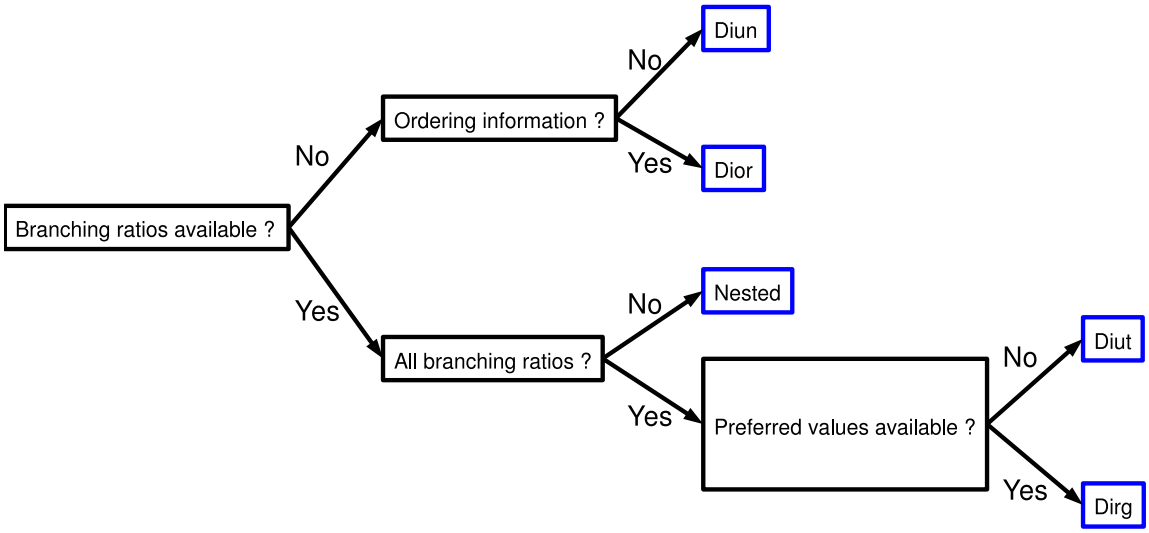

Figure 3. Decision tree for the choice of a Dirichlet-type distribution.

This elicitation method reproduces exactly the mean values, but it cannot perfectly reproduce a uniform relative uncertainty for all channels. For dissociative recombination channels, uncertainty statements are often more detailed than a global relative uncertainty; in the best cases, one even has a standard uncertainty per channel, u_i. Being isotropic in the $(n-1)$-simplex (see Fig. 4(a)), the standard Dirichlet distribution cannot properly account for this information, and we found necessary to use a more flexible distribution.

Generalization of the Dirichlet distribution is an active research area in statistics.[64–66] We use here the generalization proposed in Lingwal *et al.*.[66] Compared to the Dirichlet distribution, the generalized Dirichlet distribution

$$(b_1, \ldots, b_n) \sim \mathrm{Dirg}(\mu_1, .., \mu_n \,;\, u_1, .., u_n) \tag{10}$$

enables more anisotropy in the $(n-1)$-simplex, through a differentiated treatment of the generating Gamma distributions (see Appendix). Here again, the sum-to-one constraint prevents to fully preserve the input uncertainties u_i, but we found this sampling technique to be satisfactory for our purpose.

a. Example. Considering a set of three branching ratios $b_1 = 0.30 \pm 0.30$, $b_2 = 0.50 \pm 0.04$ and $b_3 = 0.20 \pm 0.01$, a comparison of samples in the 2-dimensional simplex for the corresponding Diri$(0.3, 0.5, 0.2\,;\, 13.7)$ (the value $\gamma = 13.7$ results from Eq. 9) and Dirg$(0.3, 0.5, 0.2\,;\, 0.3, 0.04, 0.01)$ distributions is shown in Fig. 4(a,b), as ternary graphs and parallel graphs (the latter representation is introduced here for reference; it is more useful for displaying samples in dimensions higher than 3). One can see how, starting from the same information, the Dirg distribution enables an anisotropic sampling more representative of the data constraints than the Diri distribution.

2. Intervals: Diut representation

In some instances branching ratios are defined by limits, without explicit reference to a preferred value: $b_i \in [b_i^{min}, b_i^{max}]$. The corresponding distribution is called Diut (from Dirichlet Uniform Truncated) and noted

$$(b_1, \ldots, b_n) \sim \text{Diut}\left(\left[b_1^{min}, b_1^{max}\right], \ldots, \left[b_n^{min}, b_n^{max}\right]\right) \tag{11}$$

a. Example. For comparison of Diut sampling with Dirg, we build intervals from the mean values μ_i and standard uncertainties u_i provided in the previous example, using 3σ intervals, *i.e.* $b_i = \mu_i \pm u_i \longrightarrow b_i \in [\mu_i - 3u_i, \mu_i + 3u_i]$. The corresponding distribution is therefore Diut$([0.00, 1.00], [0.38, 0.62], [0.17, 0.23])$; a sample is shown in Fig. 4(c). It can be seen that the initial interval constraints cannot be globally verified within the simplex, leading to reduced intervals for some variables. In this example, the interval for b_1 is much smaller than prescribed. This is a consequence of the prevailing constraint that the variables should have a null probability to be *outside* the prescribed intervals, *i.e.* the smaller intervals impose their rule. This type of elicitation should thus be avoided when the prescribed intervals have very different ranges $(b_i^{max} - b_i^{min})$.

3. No information: Diun representation

In the absence of information about a set of n branching ratios, one uses a *uniform* pdf over the $(n-1)$-dimensional simplex, *i.e.* a distribution which does not favor any value or set of values (see Fig. 2(b)). The uniform Dirichlet (Diun) can be obtained as a special case

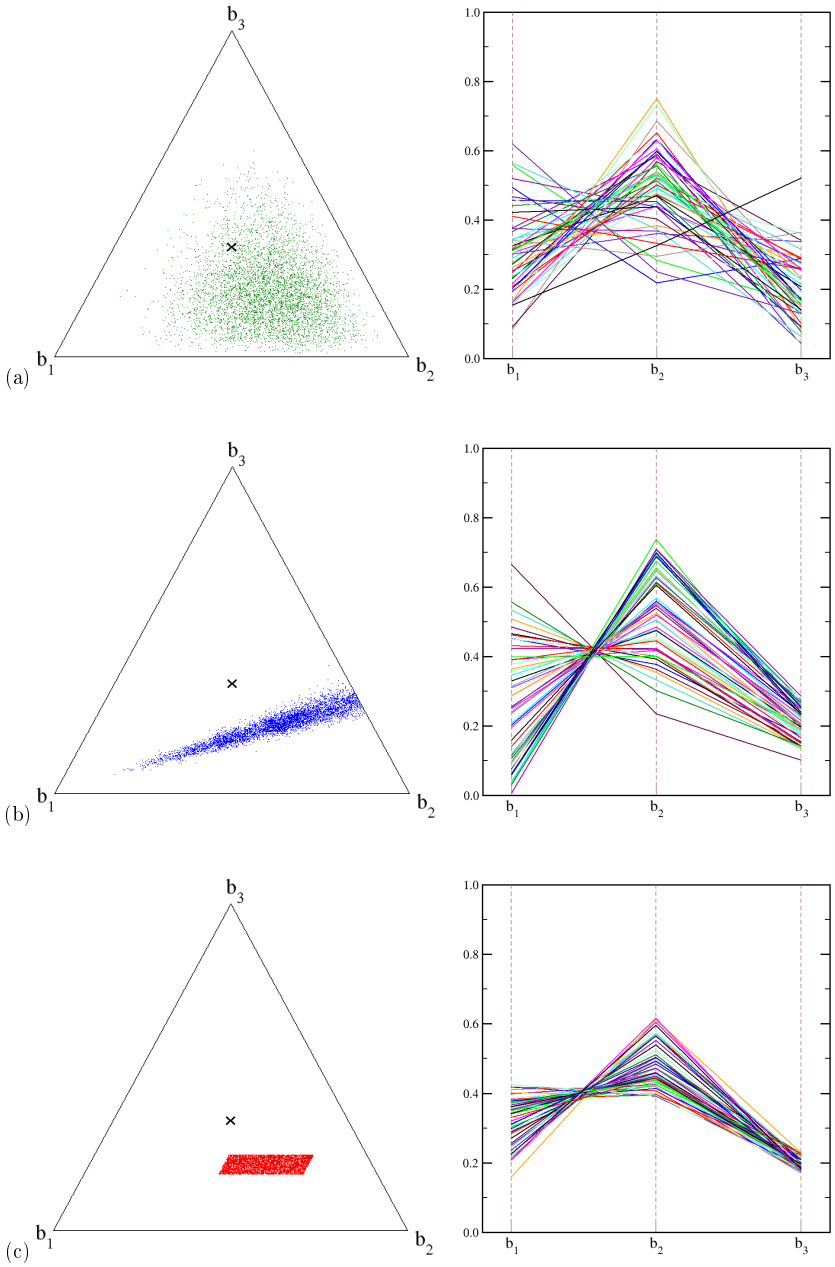

Figure 4. Ternary (left) and parallel (right) graphs of samples of Dirichlet-type distributions for a set of three branching ratios (b_1, b_2, b_3) to represent the following information $b_1 = 0.30 \pm 0.30$, $b_2 = 0.50 \pm 0.04$ and $b_3 = 0.20 \pm 0.01$: (a) Diri$(0.3, 0.5, 0.2\,;\, 13.7)$; (b) Dirg$(0.3, 0.5, 0.2\,;\, 0.3, 0.04, 0.01)$; and (c) Diut$([0.00, 1.00]\,, [0.38, 0.62]\,, [0.17, 0.23])$.

of the Diri distribution

$$(b_1, \ldots, b_n) \sim \mathrm{Diun}(n) \equiv \mathrm{Diri}(1/n, ..., 1/n\,;\, n). \tag{12}$$

Note that the Diri-type notation for the Diun distribution is preserved for use in the case of nested distributions, as shown below.

4. No information, except for an order constraint: Dior representation

Statistical and/or thermodynamical considerations are sometimes used to order the putative branching ratios of a set of channels. In such cases where the only information is provided by an ordering of the branching ratios $b_1 \geq b_2 \geq \ldots \geq b_n$, we use a uniform Diun sampling and reorder the outputs to conform with the constraint:

$$(b_1, \ldots, b_n) \sim \mathrm{Dior}(n) = \mathrm{sort}(\mathrm{Diun}(n)). \tag{13}$$

Fig. 5(b) presents the Dior distribution for three variables, with $b_1 \geq b_2 \geq b_3$.

5. Heterogeneous data sources or incomplete information about branching ratios subsets: Nested Dirichlet

In some instances, the full set of branching ratios for a reaction comes from separate experiments, and in many instances, the set of branching ratios is incomplete, with total indetermination between subsets of products. Such cases cannot be handled by any of the elementary distributions presented above and require a specific treatment.

Let us consider an example where one experiment measured the relative efficiencies of the productions of $\mathrm{M}_1+\mathrm{M}_2$ (B_1) and M_3 ($B_3 = 1-B_1$). Another experiment, independent of the first one, was able to measure the branching ratios between M_1 (B_{11}) and M_2 ($B_{12} = 1-B_{11}$). Each experiment comes with its own set of uncertainty ΔB_i, which should be preserved as well as possible when generating branching ratios for the whole set of products. A schematic representation of this information would be

$$I^+ + e^- \begin{cases} \xrightarrow{B_1 \pm \Delta B_1} \begin{cases} \xrightarrow{B_{11} \pm \Delta B_{11}} M_1 \\ \xrightarrow{B_{12} \pm \Delta B_{12}} M_2 \end{cases} \\ \xrightarrow{B_3 \pm \Delta B_3} M_3 \end{cases} \tag{14}$$

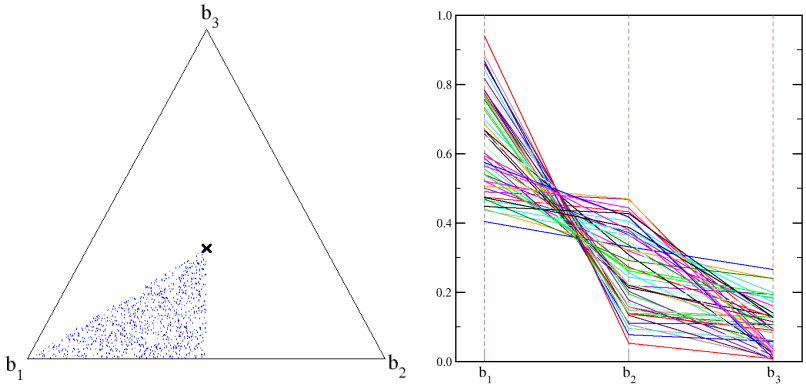

Figure 5. Dirichlet distributions for a set of three branching ratios (b_1, b_2, b_3) when only ordering information $(b_1 \geq b_2 \geq b_3)$ is available: Dior(3).

The final branching ratios for M_1, M_2 and M_3 can be computed from both data sets

$$\begin{aligned} \mu_1 &= B_1 \times B_{11}, \\ \mu_2 &= B_1 \times B_{12}, \\ \mu_3 &= B_3. \end{aligned} \tag{15}$$

An option to calculate the uncertainty on the final branching ratios is to use the standard law of combination of variances,[67] *i.e.*

$$\frac{u_i}{\mu_i} = \sqrt{\left(\frac{\Delta B_1}{B_1}\right)^2 + \left(\frac{\Delta B_{1i}}{B_{1i}}\right)^2};\ i = 1,\, 2 \tag{16}$$

and to transform the nested scheme (Eq. 14) into a "one-level" scheme

$$I^+ + e^- \begin{cases} \xrightarrow{\mu_1 \pm u_1} M_1 \\ \xrightarrow{\mu_2 \pm u_2} M_2 \\ \xrightarrow{\mu_3 \pm u_3} M_3 \end{cases} \tag{17}$$

with the distribution

$$\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\left(\mu_1, \mu_2, \mu_3\,;\, u_1, u_2, u_3\right). \tag{18}$$

We stress out that this method is not reliable when large uncertainties are involved, as is often the case for DR branching ratios. As the original correlations between the pathways are wiped out, the structure of the distribution is affected, with unpredictable effects on uncertainty propagation and sensitivity analysis.[28,31]

In order to ensure the preservation and proper treatment of the initial information tree, we propose to use a Nested Dirichlet distribution:[29,30]

$$\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\left(B_1 \otimes \mathrm{Dirg}(B_{11}, B_{12}\,;\, \Delta B_{11}, \Delta B_{12}), B_3\,;\, \Delta B_1, \Delta B_3\right) \tag{19}$$

reproducing the tree structure of the data (Eq. 14). We use the $\otimes$ symbol to emphasize that the previous notation does imply a specific function composition, *i.e.* the values sampled from the external Dirg distribution *do not* depend on the values sampled from the internal distribution. In practice, both sets of random numbers are sampled independently, and then, the necessary products are performed according to Eqs (15).

This construction becomes mandatory when the set of branching ratios contains subsets with total uncertainty. In our example, let us assume that only the set $\{B_1, B_3\}$ is

characterized by experimental data, and that the subset of products $\{M_1, M_2\}$ is inferred from energetics considerations without information on the values of the branching ratios $\{B_{11}, B_{12}\}$. This structure can be represented by the distribution

$$\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\left(B_1 \otimes \mathrm{Diun}(2), B_3\,;\, \Delta B_1, \Delta B_3\right). \tag{20}$$

Fig. (6) shows the difference between the Nested Dirichlet sampling and the "one-level" sampling (Eq.18) with the values $B_1 = 0.6 \pm 0.1$, $B_3 = 0.4 \pm 0.05$ and $B_{11} \in [0,1]$, $B_{12} \in [0,1]$. It compares samples of the distributions

$$\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\left(0.6 \otimes \mathrm{Diun}(2), 0.4\,;\, 0.1, 0.05\right) \tag{21}$$

and

$$\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\left(0.30, 0.30, 0.40\,;\, 0.18, 0.18, 0.05\right) \tag{22}$$

where the parameters are derived using Eqns (15,16) and $B_{11} = B_{12} = 0.5$ as mean value and $\Delta B_{11} = \Delta B_{12} = \sqrt{1/12}$ as standard deviation for the unit square distribution.[41] Due to the large and dominant uncertainty on b_1 and b_2, the "one-level" model (Eq. 22) is unable to preserve the accurate information about b_3, and enables a too high proportion of M_3, incompatible with the initial data. In contrast, this accurate information about b_3 is preserved by the Nested Dirichlet representation, along with the indetermination between b_1 and b_2.

The Nested Dirichlet approach is very flexible and offers a very powerful technique: (i) to preserve the statistical independence of complementary experimental information about branching ratios; and (ii) to implement partial knowledge into kinetic modeling. All the Dirichlet-type distributions presented above can be nested, to any level, as required by experimental data.

III. APPLICATIONS

This work was motivated by the study of the impact of uncertain chemical parameters on the predicted densities of ions and neutral species in Titan's ionosphere.[28,31,39,54,68,69]

In a first part, we present the guidelines we followed to build our database of dissociative recombinations for Titan's ionosphere. This is illustrated by representative examples of individual reactions extracted from the database. In a second part, we turn to the model

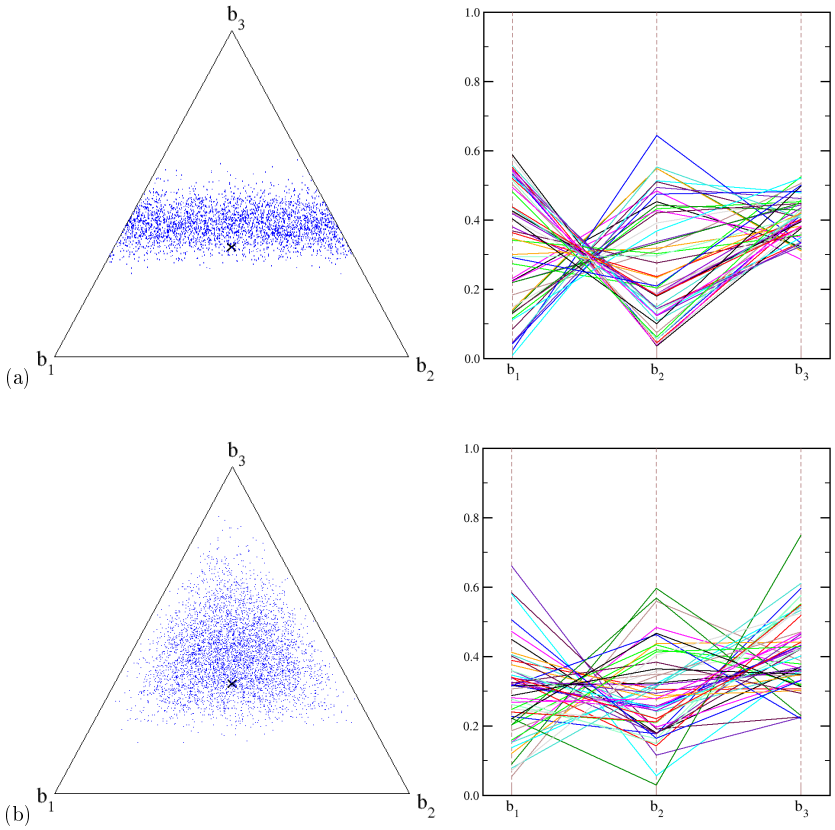

Figure 6. Comparison of two representations for the consolidation of complementary branching ratios information (see text for details): (a) $\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\,(0.6 \otimes \mathrm{Diun}(2), 0.4; 0.1, 0.05)$ and (b) $\{b_1, b_2, b_3\} \sim \mathrm{Dirg}\,(0.30, 0.30, 0.40; 0.18, 0.18, 0.05)$.

of Titan's ionospheric chemistry and compare the production rates of neutral species calculated with the "full model", as enabled by the Dirichlet approach, with the simplistic H-loss scenario, generated as a subset of the previous one.

A. A database of DR for Titan's ionosphere

Because uncertainty about branching ratios is not readily available in reference databases or review articles, a thorough review of the literature has been done to build the dataset used in this work. The focus has been put on the ions identified in Titan's ionosphere, as listed in Table II.

The starting points for our literature search were the UMIST database,[70] the OSU database,[71] the Dissociative Recombination database of the Molecular Physics group of Stockholm University,[72] and the review articles by Florescu and Mitchell,[16] and Adams *et al.*[73]. A search of various bibliographic databases was performed to find the most recent references. The full dataset, including DR rate constants and the relevant references, is provided as Supplementary Material.[74]

1. Rate constants.

The present article being focussed on branching ratios and their uncertainties, the DR rate coefficients were taken at their nominal value in the following simulations. The various issues in the treatment of uncertainty of DR rate constants will be treated in a future article. The guidelines we followed to establish the nominal values for both rate parameters α_0 and β in Eq. 1 are described here.

a. DR Rate at reference temperature: α_0

- when a single reference value was available, it was taken at face value;
- when several reference values were found:
    1. with discrepant values: the most recent was considered;
    2. with compatible values: the mean value of a loguniform distribution covering all values was used. This was retained to compensate the absence of uncertainty statement in some references.

- when no reference values were available, we used the mean values of loguniform distributions covering intervals with limits defined by existing rates for similar ions:

    1. for light species (less than 3 heavy atoms) the interval is $[5 \times 10^{-8},\ 1 \times 10^{-6}]\,\mathrm{cm^3.s^{-1}}$, and the nominal value is $2.24 \times 10^{-7}\,\mathrm{cm^3.s^{-1}}$; *e.g.* NH_2^+;
    2. for heavy species, presenting generally enhanced reaction rates, the interval is $[5 \times 10^{-7},\ 3 \times 10^{-6}]\,\mathrm{cm^3.s^{-1}}$, and the nominal value is $1.22 \times 10^{-6}\,\mathrm{cm^3.s^{-1}}$; *e.g.* $C_3H_5^+$, $C_4H_7^+$...

b. Temperature dependence, β

- when a single reference value was available, it was taken at face value;
- when several reference values were found, we used the mean of a uniform interval covering these values, *i.e.* $\beta = (\beta_{min} + \beta_{max})/2$;
- when no reference values were available, the mean value of the largest interval as defined by the theoretical values of β for direct and indirect processes was used, *i.e.* $\beta = (0.5 + 1.5)/2 = 1$.

We want to emphasize that these choices can certainly be improved. Evaluation by commitees of experts, as in other fields of chemical kinetics, would be most welcomed. Moreover, our deliberate use of intervals as the basis of reference value evaluation is motivated by our needs to consider uncertainties in these rate parameters for future work.

2. Branching ratios.

Information on branching ratios is typically sparser than on reaction rates, and it is exceptional to have to consider conflicting data. The assignments are therefore mostly based on the rules defined in the decision tree (Fig. 3). However, the following set of considerations is used in order to define the structures of the (Nested) Dirichlet distributions.

a. Maximum number of fragments. In absence of experimentally characterized products, one should consider all the exoergic channels and state a total lack of knowledge on the corresponding branching ratios (Diun distribution). It is not possible to make further hypothesises wrt. the relative stability of the products. From Ref.[16], we know for instance

that the measured branching ratios have no definite correlation with the exoergicity of the pathways. An issue in building a list of exoergic pathways is the number of fragments that can be accepted in each pathway. For instance, on the basis of their previous results, for $C_3H_7^+$ Ehlerding *et al.*[75] stop at a three body breakup pattern, even though some four body channels are opened ($C_3H_3 + H_2 + 2\,H$, $C_2H_2 + CH_3 + 2\,H$). Other authors consider also four body breakup patterns in their analysis: CD_3CDO^+,[76] CD_3CND^+,[77] and CH_2CHCNH^+.[78] In the present treatment, we favored exhaustivity and enabled four body breakups when possible.

b. H_2 vs. 2H. An empirical rule appears throughout the database of branching ratios: the loss of 2 H atoms is generally more probable than the loss of an H_2 molecule. The only measured exception are NH_4^+ and to some extent $C_2D_5^+$. As shown by Strasser *et al.*[79] through statistical modeling of the branching ratios of H_3^+, a fraction of the observed H atoms might come from the fragmentation of the H_2 product, which is created with enough internal energy to breakup through predissociative states. This rule can be used to reduce uncertainty by nesting both pathways within a Dior distribution.

c. Heavy fragments as a basis for nesting. A corrolary of the previous rule is that we are often induced to nest together pathways involving the same heavy fragment (e.g. X + 2H and X + H_2). For hydrocarbons, and in absence of experimental values, we adopted this as a general guideline to ensure an equilibrated treatment of the heavy fragments. This rule becomes ambiguous for N-bearing molecules and was not generalized to them.

The impact of these rules on the production rates of neutral species is dicussed in the next sections.

It is to be noted that, by default, all branching ratios are given for zero collision energy. If these data were to be used for systems where this is not representative, the list of open dissociation channels might have to be reconsidered and an additional uncertainty/bias factor should be attached to the measured data. At the moment, we have no information on how this contribution could be designed: it is probably very ion-dependent.

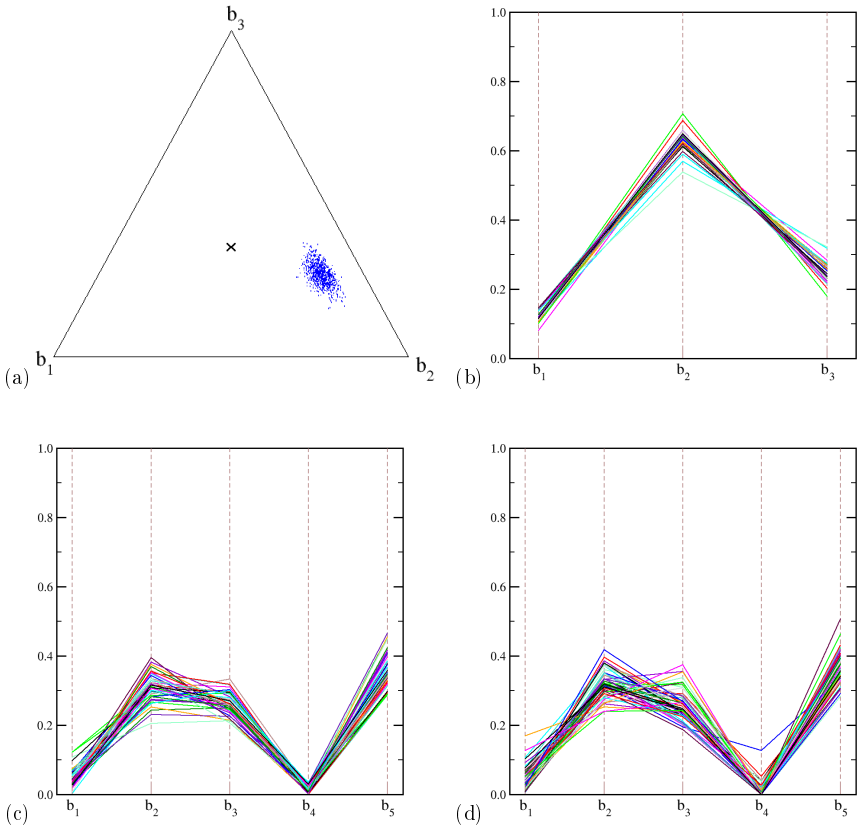

Figure 7. Samples of uncertain branching ratios for some species of interest: (a,b) CH_2^+ with Nested Dirg ; (c,d) NH_2^+ with Nested Dirg (c) and one-level Dirg (d).

B. Case studies

1. CH_2^+

a. Information. Larsson and collaborators measured the branching ratios for the DR of CH_2^+ in the CRYRING Storage Ring, and reported preferred values with uncertainty for all the pathways: $b_1(C + H_2) = 0.12 \pm 0.02$, $b_2(C + 2H) = 0.63 \pm 0.06$ and $b_3(CH + H) = 0.25 \pm 0.04$.[80]

b. Representation. As we have the complete information on branching ratios and their uncertainties, this is a straightforward case where a Dirg distribution is well adapted

$$\{b_1, b_2, b_3\} \sim \text{Dirg}\,(0.12, 0.63, 0.25\,;\, 0.02, 0.06, 0.04) \tag{23}$$

A sample of this distribution can be seen in Fig. 7(a,b).

2. NH_2^+

a. Information. This is an interesting case where we have to combine two sets of branching ratios coming from different experiments. The main channels have been measured first $B_1(N+H_2) = 0.04 \pm 0.03$, $B_2(N+2H) = 0.58 \pm 0.09$ and $B_3(NH+H) = 0.38 \pm 0.06$.[81,82] In a second stage, the spin state of N through the second channel was elucidated: $B_{21}(({}^4S)N + 2H) = 0.53 \pm 0.04$, $B_{22}(({}^2D)N + 2H) = 0.45 \pm 0.05$ and $B_{23}(({}^2P)N + 2H) = 0.02 \pm 0.02$.[82]

b. Representation. Separately, both sets of data can be handled with Dirg distributions. The global data/uncertainty pattern can be preserved via a Nested Dirichlet structure:

$$\begin{aligned}\{b_1, \ldots, b_5\} \sim \text{Dirg}\,(0.04, 0.58 \otimes \text{Dirg}(0.53, 0.45, 0.02\,;\, 0.04, 0.05, 0.02), 0.38\,;\\ 0.03, 0.09, 0.06)\,.\end{aligned} \tag{24}$$

For comparison, the one-level representation with uncertainties calculated through the standard law of combination of variances (Eq. 16) would be

$$\{b_1, \ldots, b_5\} \sim \text{Dirg}\,(0.04, 0.31, 0.26, 0.01, 0.38\,;\, 0.03, 0.05, 0.05, 0.02, 0.06)\,. \tag{25}$$

Samples from both distributions are shown in Fig. 7(c,d). As expected for small standard deviations, both representations look similar. A closer inspection reveals differences for

b_3 and b_4, for which the one-level distribution overestimates the dispersion. To be fully consistent, in absence of information of the spin state distribution of N in the first channel, one could also use a nested Diun(3) distribution over three spin state channels.

3. N_2H^+

a. Information. This is a case where the values of branching ratios are stated as intervals[83]

$$N_2H^+ + e^- \longrightarrow \begin{cases} \xrightarrow{b_1 \in [0.95,1.00]} N_2 + H \\ \xrightarrow{b_2 \in [0.00,0.05]} NH + N \end{cases} \tag{26}$$

b. Representation. As the ranges of the intervals are similar, one can use a Diut representation

$$\{b_1, b_2\} \sim \text{Diut}([0.95, 1.00], [0.00, 0.05]) \tag{27}$$

As in the previous example, an additional nesting level could be introduced in the second channel to account for the unresolved spin states of N.

4. $C_3H_2^+$

This small ion has partially unspecified products and requires the use of a Nested Dirichlet representation.

a. Information. One has measurements for $B_1(C_3)$ and $B_2(C_2 + C)$.[84,85] For the C_3 pathway, three exoergic channels are considered with unknown branching ratios ($C_3 + H_2$), ($C_3 + 2H$) and ($C_3H + H$)), and for the $C_2 + C$ pathway, two channels are invoked ($C_2H_2 + C$) and ($C_2 + CH_2$), also with unknown proportions.

b. Representation. A *flat* representation of the information provides the following scheme

$$C_3H_2^+ + e^- \longrightarrow \begin{cases} \xrightarrow{B_1 = 0.875 \pm 0.017} \begin{cases} \xrightarrow{B_{11} \in [0,1]} C_3 + 2H\,(b_1) \\ \xrightarrow{B_{12} \in [0,1]} C_3 + H_2\,(b_2) \\ \xrightarrow{B_{13} \in [0,1]} C_3H + H\,(b_3) \end{cases} \\ \xrightarrow{B_2 = 0.125 \pm 0.021} \begin{cases} \xrightarrow{B_{21} \in [0,1]} C_2H + CH\,(b_4) \\ \xrightarrow{B_{22} \in [0,1]} C_2H + C + H\,(b_5) \end{cases} \end{cases} \tag{28}$$

and the corresponding Nested Dirichlet distribution is

$$\{b_1, .., b_5\} \sim \text{Dirg}(0.875 \otimes \text{Diun}(3), 0.125 \otimes \text{Diun}(2)\,;\, 0.017, 0.021). \tag{29}$$

Implementation of the "$2\text{H} \geq \text{H}_2$" rule leads to a further nesting pattern

$$\text{C}_3\text{H}_2^+ + e^- \longrightarrow \begin{cases} \xrightarrow{B_1=0.875\pm0.017} \begin{cases} \xrightarrow{B_{11}\in[0,1]} \begin{cases} \xrightarrow{B_{111}\in[0,1]} \text{C}_3 + 2\text{H}\,(b_1) \\ \xrightarrow{B_{112}\leq B_{111}} \text{C}_3 + \text{H}_2\,(b_2) \end{cases} \\ \xrightarrow{B_{12}\in[0,1]} \text{C}_3\text{H} + \text{H}\,(b_3) \end{cases} \\ \xrightarrow{B_2=0.125\pm0.021} \begin{cases} \xrightarrow{B_{21}\in[0,1]} \text{C}_2\text{H} + \text{CH}\,(b_4) \\ \xrightarrow{B_{22}\in[0,1]} \text{C}_2\text{H} + \text{C} + \text{H}\,(b_5) \end{cases} \end{cases} \tag{30}$$

and the updated distribution

$$\{b_1, .., b_5\} \sim \text{Dirg}(0.875 \otimes \text{Diri}(0.5 \otimes \text{Dior}(2), 0.5\,;\, 2), 0.125 \otimes \text{Diun}(2)\,;\, 0.017, 0.021). \tag{31}$$

Note that for the needs of nesting within a uniform distribution, we used the notation $\text{Diri}(0.5, 0.5\,;\, 2)$ to represent $\{B_{11}, B_{12}\} \sim \text{Diun}(2)$.

We want to emphasize that in the absence of experimental data, the choice of a nesting scheme is not without consequences. The average probability of a product depends on the number of channels where it occurs. In the first case (Eq.28), one will produce C_3 and C_3H with mean probability B_1*2/3 and B_1*1/3, respectively, whereas in the second case (Eq. 30), the mean probabilities are B_1*1/2 each. However, this effect is strongly compensated by the dispersion of the samples due to uniform sampling, as shown in Fig. (8).

5. C_2H_4^+

This small ion had all its pathways experimentally characterized.[86] Starting from an aggregated version of the data, we use it to show what can be expected from the Nested Dirichlet representation, as used above for C_3H_2^+.

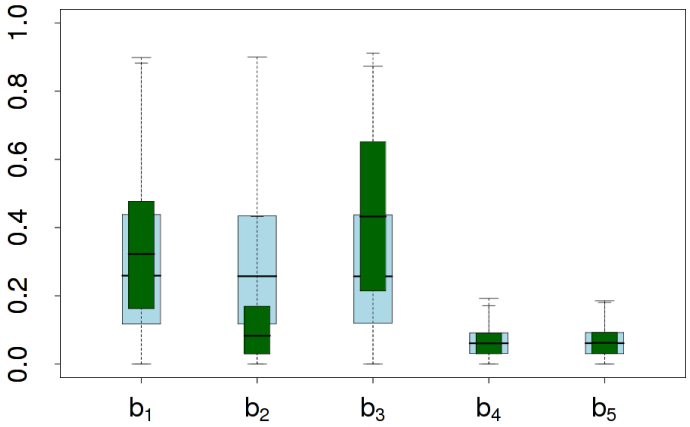

Figure 8. Influence of the nesting pattern on the distribution of branching ratios $\{b_1, \ldots, b_5\}$ for the DR of $C_3H_2^+$; (blue/light) flat representation of the sub-channels; (green/dark) nesting of auxillary information.

a. Information. The full information available is[86]

$$C_2H_4^+ + e^- \longrightarrow \begin{cases} \xrightarrow{b_1=0.11\pm0.07} C_2H_3 + H \\ \xrightarrow{b_2=0.66\pm0.06} C_2H_2 + 2H \\ \xrightarrow{b_3=0.06\pm0.03} C_2H_2 + H_2 \\ \xrightarrow{b_4=0.10\pm0.04} C_2H + H_2 + H \\ \xrightarrow{b_5=0.01\pm0.01} CH_4 + C \\ \xrightarrow{b_6=0.02\pm0.02} CH_3 + CH \\ \xrightarrow{b_7=0.04\pm0.02} 2CH_2 \end{cases} \tag{32}$$

Mock-up data are generated by aggregating the $C_2(B_1)$ and $C + C\,(B_2)$ pathways and using representative measurement uncertainties: $B_1 = 0.93 \pm 0.02$ and $B_2 = 0.07 \pm 0.03$.

b. Representation. The complete experimental data shown above can be represented by a flat Dirg distributions, as in the case of CH_2^+. A sample is represented in Fig. 9. Starting from the aggregated data, we build a nested representation based on our set of empirical rules (as they have not been measured, we did not include spin isomers of CH_2).

$$\mathrm{C_2H_4^+ + e^- \longrightarrow} \begin{cases} \xrightarrow{B_1=0.93\pm0.02} \begin{cases} \xrightarrow{B_{11}\in[0,1]} \mathrm{C_2H_3 + H}\,(b_1) \\ \xrightarrow{B_{12}\in[0,1]} \begin{cases} \xrightarrow{B_{121}\in[0,1]} \mathrm{C_2H_2 + 2H}\,(b_2) \\ \xrightarrow{B_{122}\leq B_{121}} \mathrm{C_2H_2 + H_2}\,(b_3) \end{cases} \\ \xrightarrow{B_{13}\in[0,1]} \mathrm{C_2H + H_2 + H}\,(b_4) \end{cases} \\ \xrightarrow{B_2=0.07\pm0.03} \begin{cases} \xrightarrow{B_{21}\in[0,1]} \mathrm{CH_4 + C}\,(b_5) \\ \xrightarrow{B_{22}\in[0,1]} \mathrm{CH_3 + CH}\,(b_6) \\ \xrightarrow{B_{22}\in[0,1]} \mathrm{2CH_2}\,(b_7) \end{cases} \end{cases} \tag{33}$$

The corresponding Nested Dirichlet distribution is

$$\{b_1, .., b_7\} \sim \mathrm{Dirg}(0.93 \otimes \mathrm{Diri}(1/3, 1/3 \otimes \mathrm{Dior}(2), 1/3\,;\,3), 0.07 \otimes \mathrm{Diun}(3)\,;\,0.02, 0.03). \tag{34}$$

A sample of this distribution is represented in Fig. 9 and compared to the original experimental data. The Nested Dirichlet representation partitions correctly the probability between the C_2 ($\{b_1, \ldots, b_4\}$) and C + C ($\{b_5, \ldots, b_7\}$) pathways. The $\{b_1, \ldots, b_4\}$ space is sampled uniformly, except for the "$b_2 \geq b_3$" constraint. Although this might be difficult to perceive from the one-dimensional marginal densities (boxplots), this sample *does contain* the sample representing the experimental data. This comparison illustrates an essential feature of our method: it is not expected to be predictive, but rather comprehensive/exhaustive, in the sense that it is designed to contain the "correct" branching ratios in agreement with the input data.

6. $CH_2NH_2^+$

a. Information. There are no data about the DR rate and fragmentation pattern for this ion, which has been recently postulated as a main contributor to the formation of ammonia in Titan's ionosphere through its NH_2 fragment.[87]

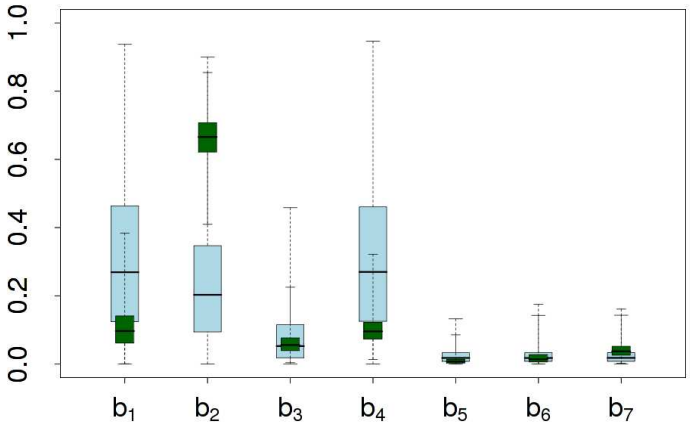

Figure 9. Comparison of branching ratios $\{b_1, \ldots, b_7\}$ for the DR of $C_2H_4^+$, as measured (dark green boxplots) and recovered from an aggregated version of these data (light blue boxplots).

b. Representation. We consider a uniform distribution over the heavy fragments with exoergic pathways, with an additional nesting for both isomers of HCN

$$\mathrm{CH_2NH_2^+} + e^- \longrightarrow \begin{cases} \xrightarrow{B_1 \in [0,1]} \mathrm{CH_3N} + \mathrm{H} \\ \xrightarrow{B_2 \in [0,1]} \mathrm{CH_2} + \mathrm{NH_2} \\ \xrightarrow{B_3 \in [0,1]} \begin{cases} \xrightarrow{B_{31} \in [0,1]} \mathrm{HCN} + \mathrm{H_2} + \mathrm{H} \\ \xrightarrow{B_{32} \in [0,1]} \mathrm{HNC} + \mathrm{H_2} + \mathrm{H} \end{cases} \\ \xrightarrow{B_4 \in [0,1]} \mathrm{CN} + 2\mathrm{H_2} \end{cases} \tag{35}$$

with distribution

$$\{b_1, .., b_5\} \sim \mathrm{Diri}(1/4, 1/4, 1/4 \otimes \mathrm{Diun}(2), 1/4\,;\, 4). \tag{36}$$

7. $HCCCNH^+$

a. Information. Geppert *et al.*[88] measured partial branching ratios for the deuterated molecule $DCCCND^+$, providing probabilities $B_1 = 0.52 \pm 0.05$ for channels $\{DC_3N + D,\ C_3N + D_2\}$ and $B_2 = 0.48 \pm 0.05$ for $\{DCN + CD_2,\ CN + C_2D_2\}$. An additional and important information is that isotope effects are expected to be small,[88] *i.e.* we can transpose this information to the hydrogenated ion $HCCCNH^+$. The experimental setup does not allow to distinguish between isomers, and Osamura *et al.*[89] have shown that there exist interconversion barriers

low enough to enable the formation of different isomers of HC_3N. They also indicate that the formation of HC_3N is more likely than HC_2NC.

b. Representation. The information about this system can be translated into the following scheme

$$\mathrm{HCCCNH^+ + e^- \longrightarrow} \begin{cases} \xrightarrow{B_1=0.52\pm0.05} \begin{cases} \xrightarrow{B_{11}\in[0,1]} \begin{cases} \xrightarrow{B_{111}\in[0,1]} \begin{cases} \xrightarrow{B_{1111}\in[0,1]} \mathrm{HC_3N + H} \\ \xrightarrow{B_{1112}\in[0,1]} \mathrm{C_3NH + H} \end{cases} \\ \xrightarrow{B_{112}\leq B_{111}} \begin{cases} \xrightarrow{B_{1121}\in[0,1]} \mathrm{C_2NCH + H} \\ \xrightarrow{B_{1122}\in[0,1]} \mathrm{HC_2NC + H} \end{cases} \end{cases} \\ \xrightarrow{B_{12}\in[0,1]} \mathrm{C_3N + H_2} \end{cases} \\ \xrightarrow{B_2=0.48\pm0.05} \begin{cases} \xrightarrow{B_{21}\in[0,1]} \begin{cases} \xrightarrow{B_{211}\in[0,1]} \mathrm{HCN + C_2H} \\ \xrightarrow{B_{212}\in[0,1]} \mathrm{HNC + C_2H} \end{cases} \\ \xrightarrow{B_{22}\in[0,1]} \mathrm{CN + C_2H_2} \end{cases} \end{cases} \tag{37}$$

In the branch B_{111}, we explicited isomers based on the position of the H atom, as proposed by Osamura *et al.*[89] Note that without explicit information about isomers of C_3N (branch B_{12}), we did not introduce additional species.

The expression of the distribution becomes very intricated and illegible and is not presented here. It involves a Dirg distribution a the first level, and then sequentially nested two-elements Diun distribution, at the exception of branches B_{111} and B_{112} which are nested into a Dior node.

C. Titan ionospheric chemistry

Prior to arrival of Cassini in Titan's atmosphere it was thought that only relatively small compounds, with masses lower than 100 u (maximum mass chosen for the Ion Neutral Mass Spectrometer instrument), would be detected in the upper atmosphere by the Cassini orbiter, with small contributions of heavy species. However both the INMS and the Cassini Plasma Spectrometer (CAPS) sensors revealed a large amount of heavy positive ions in the ionosphere, with significant ion densities above 100 u.[90–92] Waite *et al.*[91] suggested that chemical growth in the upper atmosphere is initiated by the ion-neutral chemistry, in agreement with the work of Carrasco *et al.*[93] showing the important role of the condensation reactions (bond rearrangement reactions) for ion growth. These latter processes are very important to

understand the organic complexification of the compounds in the upper atmosphere, which cannot be explained only by protonation of neutrals as suggested earlier by Vuitton *et al.*[94]

A few heavy neutrals, among them benzene, and numerous negatively charged species have been detected by INMS and CAPS in significant amounts in the thermosphere. The abundance of benzene in Titan's upper atmosphere raises up numerous issues. This stable aromatic compound constitutes another key species towards organic growth and has been observed in Titan atmosphere by different instruments: the Infrared Space Observatory,[95] the Composite InfraRed Spectrometer on board Cassini,[96] and the neutral INMS mass spectrometer.[97] The upper atmospheric chemistry seems to produce efficiently high amounts of benzene, with mole fractions of about $\sim 10^{-6}$cm^{-3}, larger than in the stratosphere $\sim 10^{-10}$ to 10^{-9}.[98] This production, specific to the thermosphere, is suspected to be due to the electron recombination of the positive ion $C_6H_7^+$. More generally, the production of all the heavy species (neutral or negatively charged) is presently understood as coming from the ultimate electron recombination of heavy positive ions. Unfortunately, as described previously, the products from these dissociative processes are poorly known, and only the H-loss co-products are implemented in the recent models.[32–37]

Exceptions to this exclusive implementation of the H-loss channels exist, but are often introduced *ad hoc*. For instance, Krasnopolsky[38] uses in his coupled model 32 DR, with only one having several branching ratios: $C_4H_5^+ + e^-$ has two channels, with branching ratios different from the measured values (Tab. I). Another exception to the H-loss scenario is $C_4H_3^+$, condemned by this author to dissociate through a unique CH–loss channel, which is known to account for less than 7 % of the measured products.[16] Another oddity concerns the major ion of Titan's ionosphere, $HCNH^+$, for which the selected H-loss channel (HCN + H) accounts at best for 40% of the products of this ion. Let us also mention that a few ions have been implemented with a "H_2-loss" process, neglecting the *a priori* more favorable "2H-loss" process, without theoretical or experimental evidence ($CHCCNH^+$, CH_3CNH^+ and $C_3H_5^+$).

In astrochemistry and ionospheric chemistry, it is well acknowledged that the "benefit" of the H-loss mechanism is to produce large neutral species.[19,100] This choice introduces therefore a bias towards the production of heavy neutrals by deprotonation of heavy ions, enforcing ion-molecule reactivity as the main engine for molecular complexification. Considering that major ions in Titan's ionosphere are protonated species (*e.g.* $HCNH^+$, $C_2H_5^+$), the H-loss DR mechanism is singularly producing stable neutral molecules (HCN, C_2H_4).

Reaction	Nb. of open channels	Selected channels (b_i) by Krasnopolsky[38]	Measured contribution of selected channels
$HCNH^+ + e^-$	4	HCN + H (1.0)	0.26 – 0.41[99]
$C_4H_3^+ + e^-$	14	C_3H_2 + CH (1.0)	0.00 – 0.07[16]
$C_4H_5^+ + e^-$	15	C_3H_4 + CH (0.7)	0.00 – 0.11[16]
		C_4H_4 + H (0.3)	0.00 – 0.51[16]
$CH_4^+ + e^-$	8	CH_3 + H (1.0)	unknown

Table I. Comparison of the DR treatment of Krasnopolsy[38] and the available literature for some representative ions.

This deprives the chemical system of an important source of highly energetic radicals.

The previous examples show that the treatment of DR in coupled ionospheric models deserves further attention. It is therefore important to check how a more complete implementation of the present knowledge about DR products can improve on the H-loss paradigm. We focus here on the production rates of neutral species from stationary ion densities. Implementing a fully coupled ionospheric model is beyond the scope of this study.

1. The kinetic model

The production rate v_M for a neutral species M through DR is

$$v_M = \frac{d\,[\mathrm{M}]}{dt} = \sum_r \left[\alpha_0^{(r)} \times \left(\frac{T_e}{300} \right)^{-\beta^{(r)}} \times \left[I^+ \right]_r \times d_e \times \sum_{j=1}^{n_r} b_j^{(r)} \times \nu_{j,\mathrm{M}} \right], \tag{38}$$

where [M] is the concentration of the neutral species of interest, $\alpha_0^{(r)}$ the thermal rate constant at 300 K of DR r, $\beta^{(r)}$ the power parameter of DR r, $[I^+]_r$ the density of the ion destroyed by DR r, $[e^-]$ the electron density, $b_j^{(r)}$ the j^{th} branching ratio of DR r and $\nu_{j,\mathrm{M}}$ the stoechiometric coefficient of the neutral species in channel j.

a. Initial conditions. The production rates were calculated with ionic densities taken from the T19 flyby INMS ionic day spectrum at 1100 km of altitude from Ref.[91] and Ref.[100] peaks assignment (see Tab. II). Electron density d_e is equal to the sum of the ions density.[101] A representative value of the electrons temperature (570 K) at 1100 km has been used.

b. Chemical schemes. Our "full scheme" implements DR for the ions listed in Table II. This scheme includes a detailed treatment for 58 ions, corresponding to 448 partial reactions,

m/z (u)	Density (cm^{-3})	Ions and predicted fractional contribution					
12.	4.31	C^+	1.00000				
13.	6.44	CH^+	1.00000				
14.	20.15	N^+	0.80000	CH_2^+	0.20000		
15.	60.62	NH^+	0.00099	CH_3^+	0.99901		
16.	17.87	CH_4^+	0.99310	NH_2^+	0.00690		
17.	90.47	CH_5^+	0.99854	NH_3^+	0.00146		
18.	155.35	NH_4^+	1.00000				
25.	7.11	C_2H^+	1.00000				
26.	9.80	$C_2H_2^+$	0.99954	CN^+	0.00046		
27.	59.42	$C_2H_3^+$	0.55195	HCN^+	0.44805		
28.	3258.69	$HCNH^+$	0.95518	N_2^+	0.02772	$C_2H_4^+$	0.01710
29.	582.39	$C_2H_5^+$	0.97422	N_2H^+	0.02526	H_3CN^+	0.00052
30.	1271.59	$CH_2NH_2^+$	1.00000				
31.	30.08	$C_2H_7^+$	0.94778	$CH_3NH_2^+$	0.05222		
32.	5.71	$CH_3NH_3^+$	1.00000				
36.	2.37	C_3^+	1.00000				
37.	2.83	C_3H^+	1.00000				
38.	11.97	C_2N^+	0.99865	$C_3H_2^+$	0.00135		
39.	2318.55	$C_3H_3^+$	0.82189	$CHCN^+$	0.17811		
40.	110.60	CH_2CN^+	0.82123	$C_3H_4^+$	0.17877		
41.	630.96	$C_3H_5^+$	0.98318	CH_3CN^+	0.01682		
42.	786.41	CH_3CNH^+	0.99955	$C_3H_6^+$	0.00045		
43.	52.69	$C_3H_7^+$	0.96577	$C_2H_3NH_2^+$	0.03423		
44.	5.06	$C_3H_8^+$	1.00000				
45.	0.79	$C_3H_9^+$	1.00000				
49.	0.40	C_4H^+	1.00000				
50.	36.02	$C_4H_2^+$	0.99878	C_3N^+	0.00122		
51.	209.77	$C_4H_3^+$	0.99887	$CHCCN^+$	0.00113		
52.	1083.38	$CHCCNH^+$	0.99620	$C_4H_4^+$	0.00364	$C_2N_2^+$	0.00016
53.	146.29	$C_4H_5^+$	0.85226	$HC_2N_2^+$	0.12801	$C_2H_3CN^+$	0.01973
54.	755.55	CH_2CHCNH^+	0.99997	$C_4H_6^+$	0.00003		
55.	63.10	$C_4H_7^+$	1.00000				
56.	67.00	$C_2H_5CNH^+$	0.99998	$C_4H_8^+$	0.00002		
57.	4.67	$C_4H_9^+$	1.00000				

Table II: Peaks assignment of the T19 flyby. Spectrum values taken from Waite *et al.*,[91] assignment from Vuitton *et al.*[100]. In absence of data about their DR, species in italics have not been considered in the model.

forming 62 different neutral species. The H-loss scheme was obtained by pruning the full scheme from irrelevant pathways and ensuring adequate normalization of the remaining branching ratios. This scheme comprises 63 partial reactions producing a subset of 48 neutral species. Note that for some ions, more than one H-loss pathways are possible (due to products isomers or spin states), and that the detailed branching ratios for all species not bearing H have been maintained. Differences observed between both models can therefore be attributed to H-bearing species only.

2. Impact of nesting schemes

The full model was run with two extreme nesting patterns: a flat pattern where all nesting options were ignored, except for the spin states of CH_2, and our working pattern, with implementation of the nesting rules defined in Section III A 2. The ratios of the production rates of neutral species by the flat (v_M^{Flat}) and the full, nested, (v_M^{Full}) mechanisms are shown in Fig. 10. As a control, we also calculated the ratio v_M^{Full}/v_M^{Full} using two different Monte Carlo samples. The 50 percent symmetric probability interval corresponding to the inner part of our boxplots has been plotted in Fig. 10 to enable a visual appreciation of the statistical significance of the deviations of production rates ratios from 1. For v_M^{Full}/v_M^{Flat} the mean densities of all neutral species vary by less than a factor two, which is much smaller than the global uncertainty on the production rates of most species. Statistically, the choice of a nesting pattern can thus be considered as a secondary effect.

3. Comparison of DR models

The ratios of the production rates of neutral species by the H-loss (v_M^{H-loss}) and the full (v_M^{Full}) mechanisms are shown in Fig. 10. It is to be noted first that the effects observed here are much larger than those due to the choice of a nesting scheme. The differences between production rates for the H-loss and full models can thus be considered as independent from the implementation details of the full model.

The H-loss mechanism favors 15 species among the 45 neutrals common to both schemes, all of which are stable neutral species. Among these 15 significantly overestimated species, 7 are nitrogen containing:

- **stable nitriles** such as HC_3N, CH_3CN, C_3H_3N. Chemistry of nitriles is a present major concern for the better understanding of Titan's aerosol formation. Vuitton *et al.*[100] deduced unsuspected high densities of nitriles in the upper atmosphere from an ion mass spectrum analysis measured by the Cassini INMS instrument, and propose them as very plausible aerosol precursors. From the Huygens ACP experiment, Israel *et al.*[102] identified -CN groups as one major constitutive function in the aerosol composition. Moreover, recent works on Titan's lab analogs analysis also highlighted patterns involving an unsaturated C,N combination for tholin molecular growth.[103,104] However, photochemistry and neutral-neutral chemistry of nitriles have already been identified as incomplete and to be improved (see the review of Hébrard *et al.*[52]). Here, we point out that the H-loss model tends to exagerate the nitrile production rates in Titan's upper atmosphere. As an example, Vigren *et al.*[77] showed that the dissociative recombination of CD_3CND^+ led to 35% of bond rupture between heavy atoms.

- **one imine**, the methanimine CH_3N. Balucani *et al.*[105] identified this species as a good candidate for polymerization or co-polymerization in Titan's atmosphere, suggesting new paths for nitrogen rich aerosols production. According to their experimental and theoritical study, methanimine would also be produced by the reaction between $N(^2D)$ and ethane. It would thus be important to compare these two methanimine production pathways in a complete ion-neutral Titan's atmospheric model.

- and three **excited isomers** of HC_3N: C_3NH, C_2NCH and HC_2NC, possibly produced by the DR of HC_3NH^+.[89]. These very reactive nitrogen-bearing species are not known to be produced by any other process in Titan's atmosphere. Moreover their own reactivity remains completely unknown. As a matter of fact, these three reactive species are very possibly produced in Titan's atmosphere, are very reactive, nitrogen-containing, but their becoming in Titan's atmosphere is completely unknown. This highlights a challenging case for Titan's chemistry understanding.

The production rates of 6 species are indifferent to the choice of DR mechanism, among them N_2, H (logically well represented in a H-loss mechanism) and ammonia. Considering the poor knowledge on the dissociative recombination of larger protonated amines, an enhancement of the production rate of ammonia could be expected when more data get available.

The most salient feature of this comparison is that the full mechanism enhances significantly the production rate for 24 of the 41 species common to both schemes. The enhancement factor goes from 2-3 to 10^4. We consider below a few cases which are of high interest for Titan's atmospheric chemistry.

- One major result of this simulation is that the H-loss mechanism does not account for the rich reactivity enhancement due to the formation of radicals by DR. The H-loss model strongly underestimates the production rates of radicals in comparison with the stable neutral species (except C_3H_7 which is not significantly modified): 16 of the 24 underestimated species are radicals. This may affect the neutral growth pathways. Their underestimation by DR processes may lead to improper neutral production rates in photochemical models. A first example is the methyl radical CH_3, an important node in the Titan neutral chemical network for hydrocarbon neutral growth. Its DR production rate is underestimated by a factor of ten in the H-loss mechanism, which might be significant for night chemistry. Another striking example is the C_4H_y family or radicals. Those are the main compounds presently taken into account as "soot" in the photochemical models to initiate aerosol nucleation in the stratosphere, whereas their production rates by DR in the H-loss mechanism are largely underestimated, with factors above 100 for C_4H_5 and C_4H_7.

- Nitrogen atoms production rate is also slightly enhanced in the full-model. Actually the increase itself is not important, but the spin states of these atoms are not systematically well-quantified in the DR studies and cannot therefore be presently taken into account in the DR models, except maybe through a Diun construction as we did for the spin states of CH_2. And yet, the long-lived metastable $N(^2D)$ atoms undergo efficient reactions with all the abundant stable neutrals in Titan's atmosphere, whereas the ground state $N(^4S)$ atoms are almost not reactive.[106] Their respective production from DR pathways have then to be carefully measured in order to be compared to molecular nitrogen dissociation production.

- The production rate of C_2H_2 is underestimated in the H-loss model by a factor of a few tens. C_2H_2 is a saturated hydrocarbon invoked to explain polymerization mechanisms in Titan's stratosphere.[38,107–109] An additional production rate by the DR full

mechanism could thus be important in comparison with the neutral production rates in coupled ion-neutral chemistry models.

- It is also to be noted that H_2 is not formed through the H-loss scenario. The underestimation of H_2 could also have an important impact in the models, this species being suspected to participate to heterogeneous reactions with the aerosols.[110–112]

For all these compounds, we highlight here the importance of the calculation of their production rates with a full-DR model instead of a simplified H-loss model to be properly compared with the neutral production rate contribution in coupled ion-neutral chemistry models.

IV. CONCLUSION

Important results about the distribution of products of dissociative recombination have been gathered by several teams in the last years. Although partial, this information deserves to be considered in detailed chemical models of ionized media. However, the nominal approach of chemistry modeling is not able to cope consistently with partial data. We have shown that a probabilistic approach based on Nested Dirichlet distributions, encompassing plausible cases in conformance with experimental data, enables to deal with this situation. An inconvenience of this approach is the additional computational cost of Monte Carlo simulations, but in such conditions of uncertainty, deterministic modeling is unarguably a poor option.

Distributions of the Dirichlet family are generic tools for sum-to-one variables, easy to implement in a Monte Carlo uncertainty propagation framework, providing a flexible treatment of the available data about branching ratios. We have shown that they enable to unlock the modeling of complex chemical networks involving partially known dissociative recombination products. The representation of uncertain branching ratios proposed in this work is compact, self-contained and suited for implementation in kinetics databases such as the Kinetic Database for Astrochemistry (KIDA).[113]

When compared to the H-loss mechanism still in use in many ionospheric chemistry models for Titan, the Dirichlet modeling provides a spectacular enrichment in the chemodiversity and in the production rates of highly reactive neutral radicals. Where the H-loss scenario, forming mainly stable neutral species, can be considered as damping the reactivity

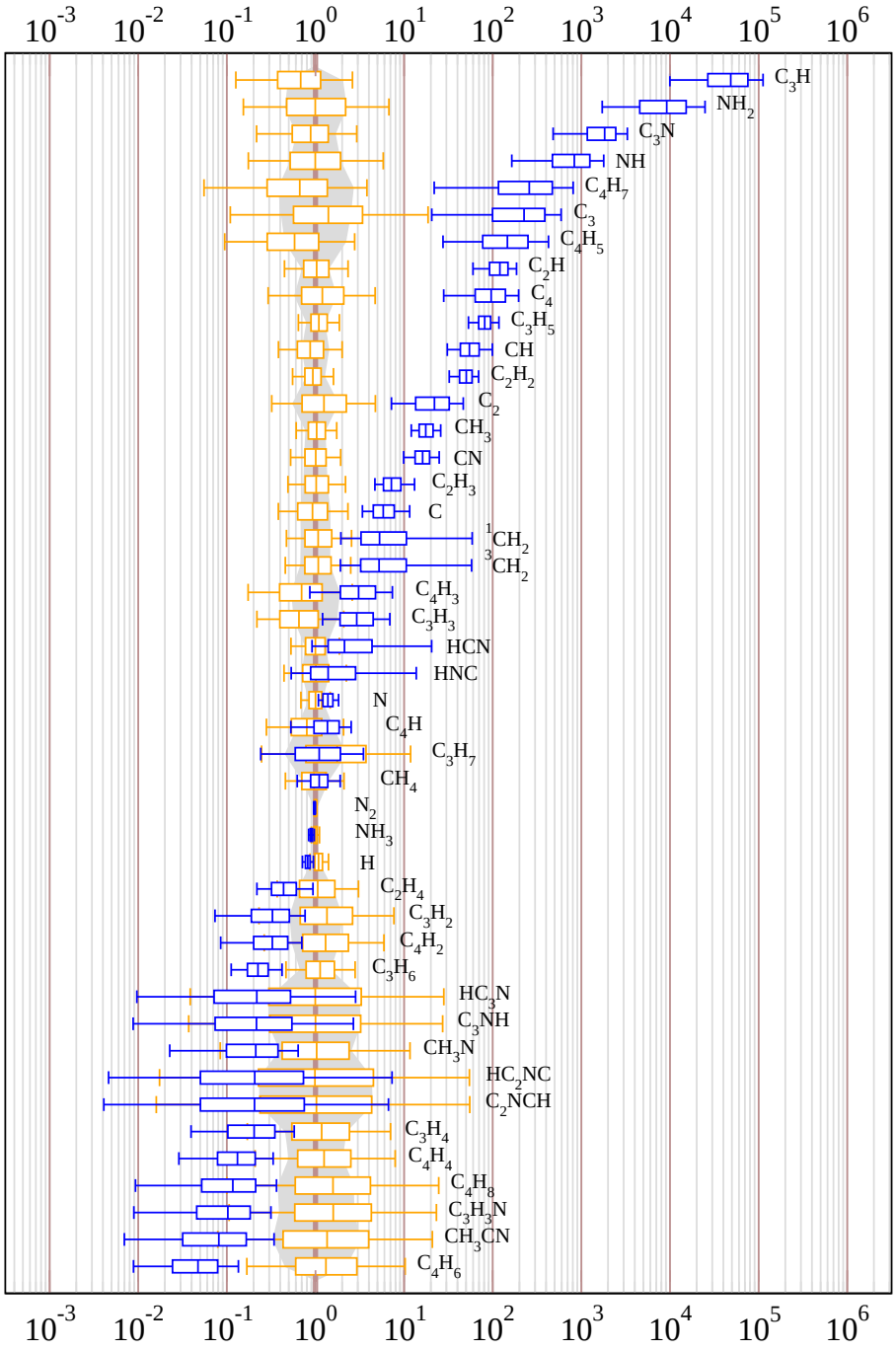

Figure 10. Comparison of the production rates of neutral species between different DR model implementations: (brown) ratio v_M^{Full}/v_M^{Flat} between the nested and flat versions of the "Full" model (see text); (blue) ratio v_M^{Full}/v_M^{H-loss} between the "Full" and "H-loss" models. They grey zone depicts the intrinsic dispersion due to the Monte Carlo estimation of production rates ratios: symmetric 50 percent probability interval for the ratio v_M^{Full}/v_M^{Full} estimated from two independent samples.

of neutral species, the full model can be seen as boosting this reactivity, and contributing to molecular growth through radical chemistry. The effect of these radicals on the chemistry of neutral species has now to be quantified by implementation of the enhanced DR scheme in a ion-neutral coupled model (work in progress).

SUPPLEMENTARY MATERIAL

The database of DR rate constants and branching ratios elaborated for this article and used in the applications is provided in a separate document, with additional notes and complete references.

ACKNOWLEDGMENTS

The authors thank J.B. Mitchell, K. Béroff and M. Chabot for fruitful discussions.

Appendix A: Sampling of Dirichlet-type distributions

We present here the basic techniques to generate random samples of Dirichlet-type distributions used in this article. A Fortran code and R interface to generate random samples from these distributions are available by simple request to the contact author.

1. The Dirichlet distribition: Diri

The Dirichlet distribution is efficiently sampled using sum-normalized Gamma random numbers:[31,114]

1. independent random values $x_1,\ldots, x_n$ are drawn from standard gamma distributions, defined by a single parameter α_i

$$\mathrm{Sgamma}(\alpha_i) = \frac{1}{\Gamma(\alpha_i)}\, x_i^{\alpha_i - 1} e^{-x_i};\ \text{where}\, \alpha_i = \mu_i \times \gamma, \tag{A1}$$

   where $\Gamma(.)$ is the gamma function,[115] and

2. the x_i are normalized

$$b_i = x_i / \sum_{j=1}^{n} x_j. \tag{A2}$$

2. The generalized Dirichlet distribution: Dirg

For each branching ratio, one defines two parameters, α_i (shape parameter) and β_i (scale parameter) through

$$\alpha_i = \left(\frac{\mu_i}{u_i}\right)^2 \text{ and } \beta_i = \frac{\mu_i}{u_i^2}, \tag{A3}$$

one generates n independent random samples x_1,..., x_n from gamma distributions

$$\text{Gamma}(\alpha_i, \beta_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} x_i^{\alpha_i - 1} e^{-\beta_i x_i}, \tag{A4}$$

and one normalizes them (Eq. A2).

3. Uniform Dirichlet: Diun

The recipe is to generate n independent random samples x_1,..., x_n from unit-scaled exponential distributions $\text{Expon}(1) = e^{-x_i}$ and normalize them (Eq. A2).

4. Truncated uniform Dirichlet: Diut

In order to obtain a uniform sampling in a subspace of the simplex, it is possible to start from the uniform Diun distribution, and use a rejection algorithm. This approach becomes very inefficient for small intervals and/or high dimensions. Efficient direct generation of uniform samples on restricted spaces is an active research area in experimental design.[116,117] A direct algorithm based on conditional distributions, proposed by Fang and Yang,[118] has been used in this work.

5. Nested Dirichlet

Practical implementation is based on the building of a tree structure in which one generates at each node the random numbers for its children nodes and then calculate the probabilities of the terminal leaves by recursive products along the tree branches.

REFERENCES

[1]T.J. Millar, D.J. DeFrees, A.D. McLean, and E. Herbst. The sensitivity of gas-phase models of dense interstellar clouds to changes in dissociative recombination branching ratios. *Astron. Astrophys.*, 194:250–256, 1988.

[2]E. Herbst and H.-H. Lee. New dissociative recombination product branching fractions and their effect on calculated interstellar molecular abundances. *The Astrophysical Journal*, 485:689–696, 1997.

[3]W.D. Geppert, W.D. Thomas, A. Ehlerding, J. Semaniak, F. Österdahl, M. af Ugglas, N. Djurić, A. Paál, and M. Larsson. Extraordinary branching ratios in astrophysically important dissociative recombination reactions. *Faraday Discuss.*, 127:425–437, 2004.

[4]W.D. Geppert, R. Thomas, A. Ehlerding, F. Hellberg, F Österdahl, M. Hamberg, J. Semaniak, V. Zhaunerchyk, M. Kaminska, A Källberg, A. Paál, and M. Larsson. Dissociative recombination branching ratios and their influence on interstellar clouds. *J. Phys.: Conf. Ser.*, 4:26–31, 2005.

[5]J.L. Fox. Effects of dissociative recombination on the composition of planetary atmospheres. *Journal of Physics: Conference Series*, page 32, 2005.

[6]D.L. Huestis, S.W. Bougher, J.L. Fox, M. Galand, R.E. Johnson, J.I. Moses, and J.C. Pickering. Cross section and reaction rates for comparative planetary aeronomy. *Space Science Reviews*, 139:63–105, 2008.

[7]M. Masuzaki and T. Tsuzuki. Dissociative recombination as a possible cause of the axial density decay in a long diffused hydrogen plasma column. *Physics Letters A*, 38:441 – 442, 1972.

[8]D.C. Schram. Plasma processing and chemistry. *Pure Appl. Chem.*, 74:369–380, 2002.

[9]D.C. Schram, R.A.B. Zijlmans, O. Gabriel, and R. Engeln. Dissociative recombination as primary dissociation channel in plasma chemistry. *Journal of Physics: Conference Series*, 192:012012, 2009.

[10]C.H. Sheehan and J.P. St.-Maurice. Dissociative recombination of the methane family ions: rate coefficients and implications. *Advances in Space Research*, 33:216–220, 2004.

[11]R.D. Thomas. When electrons meet molecular ions and what happens next: Dissociative recombination from interstellar molecular clouds to internal combustion engines. *Mass Spectrometry Reviews*, 27(5):485–530, 2008.

[12]J.L. McLain and N.G. Adams. Flowing afterglow studies of temperature dependencies for electron dissociative recombination of $HCNH^+$, CH_3CNH^+ and $CH_3CH_2CNH^+$ and their symmetrical proton-bound dimers. *Planetary and Space Science*, 57:1642–1647, 2009.

[13]D.R. Bates. Dissociative recombination. *Phys. Rev.*, 78:492–493, 1950.

[14]D.R. Bates. Products of dissociative recombination of polyatomic ions. *Astrophysical Journal*, 306:L45–L47, 1986.

[15]D.R. Bates. Dissociative recombination of polyatomic ions - Curve crossing. *Astrophysical Journal*, 344:531–534, 1989.

[16]A.I. Florescu-Mitchell and J.B.A. Mitchell. Dissociative recombination. *Physics Reports*, 430:277–374, 2006.

[17]M. Larsson and A.E. Orel. *Dissociative Recombination of Molecular Ions.* Cambridge University Press, 2008.

[18]E. Herbst and W. Klemperer. The formation and depletion of molecules in dense interstellar clouds. *Astrophysical Journal*, 185:505–534, 1973.

[19]E. Herbst. What are the products of polyatomic ion-electron dissociative recombination reactions ? *Astrophysical Journal*, 222:508–516, 1978.

[20]S. Green and E. Herbst. Metastable isomers - A new class of interstellar molecules. *Astrophysical Journal*, 229:121–131, 1979.

[21]T.J. Millar, D.J. Defrees, A.D. McLean, and E. Herbst. The sensitivity of gas-phase models of dense interstellar clouds to changes in dissociative recombination branching ratios. *Astronomy and Astrophysics*, 194:250–256, 1988.

[22]E.T. Galloway and E. Herbst. Can phase space theory reproduce experimental neutral product branching ratios for dissociative recombination reactions? *Astrophysical Journal*, 376:531–539, 1991.

[23]D. Strasser, J. Levin, H. B. Pedersen, O. Heber, A. Wolf, D. Schwalm, and D. Zajfman. Branching ratios in the dissociative recombination of polyatomic ions: The H_3^+ case. *Phys. Rev. A*, 65:010702, 2001.

[24]D. Talbi, A. Le Padellec, and J.B.A. Mitchell. Quantum chemical calculations for the dissociative recombination of HCN^+ and HNC^+. *Journal of Physics B: Atomic, Molecular and Optical Physics*, page 3631, 2000.

[25]A.P. Hickman, R.D. Miles, C. Hayden, and D. Talbi. Dissociative recombination of e + $HCNH^+$: Diabatic potential curves and dynamics calculations. *Astronomy & Astro-*

physics, 438:31–37, 2005.

[26]D. Talbi. A quantum chemical study of the $N_2H^+ + e^- \longrightarrow N_2 + H$ reaction I: The linear dissociation path. *Chemical Physics*, 332:298–303, 2007.

[27]H.-G. Yu. A spherical electron cloud hopping model for studying product branching ratios of dissociative recombination. *The Journal of Chemical Physics*, 128:194106, 2008.

[28]N. Carrasco, C. Alcaraz, O. Dutuit, S. Plessis, R. Thissen, V. Vuitton, R. Yelle, and P. Pernot. Sensitivity of a Titan ionospheric model to the ion-molecule reaction parameters. *Planetary and Space Science*, 56:1644–1657, 2008.

[29]B. Null. The Nested Dirichlet distribution: properties and applications. Working paper, Department of Management Science and Engineering, Stanford University, october 2008.

[30]G.-L. Tian, M.-L. Tang, K.C. Yuen, and K.W. Ng. Further properties and new applications of the Nested Dirichlet distribution. *Computational Statistics & Data Analysis*, 54:394–405, 2010.

[31]N. Carrasco and P. Pernot. Modeling of branching ratio uncertainty in chemical networks by Dirichlet distributions. *Journal of Physical Chemistry A*, 11(18):3507–3512, 2007.

[32]C.N. Keller, T.E. Cravens, and L. Gan. A model of the ionosphere of Titan. *Journal of Geophysical Research*, 97:12117, 1992.

[33]N.G. Keller, V.G. Anicich, and T.E. Cravens. Model of Titan's ionosphere with detailed hydrocarbon ion chemistry. *Planetary and Space Science*, 46:1157, 1998.

[34]M. Banaszkiewicz, L.M. Lara, R. Rodrigo, J.J. Lopez-Moreno, and G.J. Molina-Cuberos. A coupled model of Titan's atmosphere and ionosphere. *Icarus*, 147:386–404, 2000.

[35]E.H. Wilson and S.K. Atreya. Current state of modeling the photochemistry of Titan's mutually dependent atmosphere and ionosphere. *Journal of Geophysical Research*, 109:E06002, 2004.

[36]V. De La Haye. *Corona Formation and Heating Efficiencies in Titan's Upper Atmosphere: Construction of a coupled Ion, Neutral and Thermal Structure Model To Interpret the First INMS Cassini Data.* PhD thesis, The University of Michigan, 2005.

[37]V. De La Haye, J.H. Waite, T.E. Cravens, I.P. Robertson, and S. Lebonnois. Coupled ion and neutral rotating model of Titan's upper atmosphere. *Icarus*, 197:110–136, 2008.

[38]V.A. Krasnopolsky. A photochemical model of Titan's atmosphere and ionosphere. *Icarus*, 201:226–256, 2009.

[39]N. Carrasco, O. Dutuit, R. Thissen, M. Banaszkiewicz, and P. Pernot. Uncertainty analysis of bimolecular reactions in Titan ionosphere chemistry model. *Planetary and Space Science*, 55:141–157, 2007.

[40]N. Carrasco, C. Alcaraz, O. Dutuit, S. Plessis, R. Thissen, V. Vuitton, R. Yelle, and P. Pernot. Sensitivity of a Titan ionospheric model to the ion-molecule reaction parameters. *Planetary and Space Science*, 56:1644 – 1657, 2008.

[41]BIPM, IEC, IFCC, ISO, IUPAP, and OIML. Evaluation of measurement data - Supplement 1 to the GUM: propagation of distributions using a Monte Carlo method. Technical report, International Organization for Standardization (ISO), Geneva, 2008.

[42]J.C. Helton, J.D. Johnson, C.J. Sallaberry, and C.B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety*, 91:1175–1209, 2006.

[43]A.M. Thompson and R.W. Stewart. Effect of chemical kinetics uncertainties on calculated constituents in a tropospheric photochemical model. *J. Geophys. Res.*, 96:13089–13108, 1991.

[44]M. Dobrijevic and J.P. Parisot. Effect of chemical kinetic uncertainties on hydrocarbon production in the stratosphere of Neptune. *Planetary and Space Science*, 46:491–505, 1998.

[45]B.D. Phenix, J.L. Dinaro, M.A. Tatang, J.W. Tester, J.B. Howard, and G.J. McRae. Incorporation of parametric uncertainty into complex kinetic mechanisms: Application to hydrogen oxidation in supercritical water. *Combustion and flame*, 112:132–146, 1998.

[46]G.E. Moore and R.J. Londergan. Sampled Monte Carlo uncertainty analysis for photochemical grid models. *Atmospheric Environment*, 35:4863–4876, 2001.

[47]D. Bose, M.V.V.S. Rao, T.R. Govindan, and M. Meyyappan. Uncertainty and sensitivity analysis of gas-phase chemistry in a CHF_3 plasma. *Plasma Sources Science and Technology*, 12:225–234, 2003.

[48]M. Dobrijevic, J.L. Ollivier, F. Billebaud, J. Brillet, and J.P. Parisot. Effect of chemical kinetic uncertainties on photochemical modeling results: Application to Saturn's atmosphere. *A & A*, 398:335–344, 2003.

[49]M.T. Reagan, H.N. Najm, P.P. Pebay, O.M. Knio, and R.G. Ghanem. Quantifying uncertainty in chemical systems modeling. *Int. J. Chem. Kinet.*, 37:368–382, 2005.

[50]V. Wakelam, F. Selsis, E. Herbst, and P. Caselli. Estimation and reduction of the uncertainties in chemical models: application to hot core chemistry. *A. & A.*, 444:883–891, 2005.

[51]V. Wakelam, E. Herbst, and F. Selsis. The effect of uncertainties on chemical models of dark clouds. *A. & A.*, 451:551–562, 2006.

[52]E. Hébrard, M. Dobrijevic, Y. Bénilan, and F. Raulin. Photochemical kinetics uncertainties in modeling Titan's atmosphere: a review. *Journal of Photochemistry and Photobiology A: Chemistry*, 7:211–230, 2006.

[53]E. Hébrard, M. Dobrijevic, Y. Bénilan, and F. Raulin. Photochemical kinetics uncertainties in modeling Titan's atmosphere: First consequences. *Planetary and Space Science*, 55:1470–1489, 2007.

[54]M. Dobrijevic, N. Carrasco, E. Hébrard, and P. Pernot. Epistemic bimodality and kinetic hypersensitivity in photochemical models of Titan's atmosphere. *Planetary and Space Science*, 56:1630–1643, 2008.

[55]E. Hébrard N. Carrasco Z. Peng, M. Dobrijevic and P. Pernot. Photochemical modeling of titan atmosphere at the "10 percent uncertainty horizon". *Faraday Discuss.*, In press, 2010.

[56]An exception would be the case where a group of rate constants has been adjusted simultaneously to fit some data, leading to correlated uncertainties. However, there is no traceability for such cases in reference chemical reactions databases.

[57]J. Aitchison. *The Statistical Analysis of Compositional Data.* Monographs on Statistics and Applied Probability. Chapman and Hall, London, 1986.

[58]A. O'Hagan. Research in elicitation. Research Report No.557/05. Technical report, Department of Probability and Statistics, University of Sheffield, 2005. Invited article for a volume entitled Bayesian Statistics and its Applications.

[59]J.N. Kapur. *Maximum entropy models in science and engineering.* John Wiley and Sons, 1989.

[60]M. Evans, N. Hastings, and B. Peacock. *Statistical Distributions.* Wiley-Interscience, 3rd edition, 2000.

[61]M.O. Vlad, M. Tsuchiya, P. Oefner, and J. Ross. Bayesian analysis with random chemical composition: Renormalization-group approach to Dirichlet distributions and the statistical theory of dilution. *Phys. Rev. E*, 65:011112, 2002.

[62]V.G. Anicich. Evaluated bimolecular ion-molecule gas phase kinetics of positive ions for use in modelling planetary atmospheres, cometary comae and interstellar clouds. *J. Phys. Chem. Ref. Data*, 22(6):1469–1569, 1993.

[63]V.G. Anicich. An index of the literature for bimolecular gas phase cation-molecule reaction kinetics. *JPL Publication*, 03-19:1–1194, 2003.

[64]T.-T. Wong. Generalized Dirichlet distribution in Bayesian analysis. *Applied Mathematics and Computation*, 97:165–181, 1998.

[65]A. Ongaro, S. Migliorati, and G.S. Monti. A new distribution on the simplex containing the Dirichlet family. CODAWORK'08, May 2008. Girona: La Universitat.

[66]J.W. Lingwall, W.F. Christensen, and C.S. Reese. Dirichlet based bayesian multivariate receptor modeling. *Environmetrics*, 19:618–629, 2008.

[67]BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM). Technical report, International Organization for Standardization (ISO), Geneva, 2008.

[68]N. Carrasco, E. Hébrard, M. Banaszkiewicz, M. Dobrijevic, and P. Pernot. Influence of neutral transport on ion chemistry uncertainties in Titan ionosphere. *Icarus*, 192:519–526, 2007.

[69]E. Hébrard, M. Dobrijevic, P. Pernot, N. Carrasco, A. Bergeat, K. M. Hickson, A. Canosa, S. D. Le Picard, and I. R. Sims. How measurements of rate coefficients at low temperature increase the predictivity of photochemical models of Titan's atmosphere. *The Journal of Physical Chemistry A*, 113:11227–11237, 2009.

[70]J. Woodall, M. Agúndez, A.J. Markwick-Kemper, and T.J. Millar. The UMIST database for astrochemistry 2006. *Astron. Astrophys.*, 466:1197–1204, 2007.

[71]E. Herbst. Gas-phase model. http://www.physics.ohiostate.edu/ eric/research.html.

[72]Molecular Physics Group. Dissociative recombination database. http://mol.physto.se/DRdatabase/.

[73]N.G. Adams, V. Poterya, and L.M. Babcock. Electron molecular ion recombination: Product excitation and fragmentation. *Mass Spectrometry Reviews*, 25:798–828, 2006.

[74]See Supplementary Material Document No._____ for the full dataset of branching ratios, along with DR rate constants and the relevant references.

[75]A. Ehlerding, S.T. Arnold, A.A. Viggiano, S. Kalhori, J. Semaniak, A.M. Derkatch, S. Rosen, M. af Ugglas, and M. Larsson. Rates and products of the dissociative re-

combination of $C_3H_7^+$ in low-energy electron collisions. *Journal of Physical Chemistry A*, 107:2179–2184, 2003.

[76]E. Vigren, M. Kaminska, M. Hamberg, V. Zhaunerchyk, R.D. Thomas, J. Semaniak, M. Danielsson, M. Larsson, and W.D. Geppert. Dissociative recombination of the deuterated acetaldehyde ion, CD_3CDO^+: product branching fractions, absolute cross sections and thermal rate coefficient. *Physical Chemistry Chemical Physics*, 9:2856–2861, 2007.

[77]E. Vigren, M. Kaminska, M. Hamberg, V. Zhaunerchyk, R.D. Thomas, M. Danielsson, J. Semaniak, P.U. Andersson, M. Larsson, and W.D. Geppert. Dissociative recombination of fully deuterated protonated acetonitrile, CD_3CND^+: product branching fractions, absolute cross section and thermal rate coefficient. *Physical Chemistry Chemical Physics*, 10:4014–4019, 2008.

[78]E. Vigren, M. Hamberg, V. Zhaunerchyk, M. Kaminska, R. D. Thomas, M. Larsson, T. J. Millar, C. Walsh, and W. D. Geppert. The dissociative recombination of protonated acrylonitrile, CH_2CHCNH^+, with implications for the nitrile chemistry in dark molecular clouds and the upper atmosphere of Titan. *The Astrophysical Journal*, 695:317–324, 2009.

[79]D. Strasser, L. Lammich, S. Krohn, M. Lange, H. Kreckel, J. Levin, D. Schwalm, Z. Vager, R. Wester, A. Wolf, and D. Zajfman. Two- and three-body kinematical correlation in the dissociative recombination of H_3^+. *Phys. Rev. Lett.*, 86:779–782, 2001.

[80]A. Larson, A. Le Padellec, J. Semaniak, C. Stromholm, M. Larsson, S. Rosen, R. Peverall, H. Danared, N. Djuric, G.H. Dunn, and S. Datz. Branching fractions in dissociative recombination of CH_2^+. *The Astrophysical Journal*, 505:459–465, 1998.

[81]L. Vikor, A. Al-Khalili, H. Danared, N. Djurić, G.-H. Dunn, M. Larsson, A. Le Padellec, S. Rosén, and M. Af Ugglas. Branching fractions in the dissociative recombination of NH_4^+ and NH_2^+ molecular ions. *Astron. Astrophys.*, 344:1027–1033, 1999.

[82]R.D. Thomas, F. Hellberg, A. Neau, S. Rosén, M. Larsson, C.R. Vane, M.E. Bannister, S. Datz, A. Petrignani, and W.J. van der Zande. Three-body fragmentation dynamics of amidogen and methylene radicals via dissociative recombination. *Phys. Rev. A*, 71:032711–16, 2005.

[83]C.D. Molek, V. Poterya, N.G. Adams, and J.L. McLain. Development of a novel technique for quantitatively determining the products of electron-ion dissociative recombination. *International Journal of Mass Spectrometry*, 285:1–11, 2009.

[84]G. Angelova, O. Novotny, J.B.A. Mitchell, C. Rebrion-Rowe, J.L. Le Garrec, H. Bluhme, A. Svendsen, and L.H. Andersen. Branching ratios for the dissociative recombination of hydrocarbon ions: III: the cases of $C_3H_n^+$ (n=1-8). *International Journal of Mass Spectrometry*, 235:7–13, 2004.

[85]J.B.A. Mitchell, G. Angelova, C. Rebrion-Rowe, O. Novotny, J.L. Le Garrec, H. Bluhme, K. Seiersen, A. Svendsen, and L.H. Andersen. Branching ratios for the dissociative recombination of hydrocarbon ions. *Journal of Physics: Conference Series*, 4:198–204, 2005.

[86]A. Ehlerding, F. Hellberg, R. Thomas, S. Kalhori, A.A. Viggiano, S.T. Arnold, M. Larsson, and M. af Ugglas. Dissociative recombination of C_2H^+ and $C_2H_4^+$: absolute cross sections and product branching ratios. *Phys. Chem. Chem. Phys.*, 6:949–954, 2004.

[87]R. Yelle, V. Vuitton, P. Lavvas, M. Smith, S. Horst, and J. Cui. Synthesis of NH_3 in Titan's upper atmosphere. In *AAS/Division for Planetary Sciences Meeting Abstracts*, volume 41 of *AAS/Division for Planetary Sciences Meeting Abstracts*, page 17.07, September 2009.

[88]W.D. Geppert, A. Ehlerding, F. Hellberg, J. Semaniak, F. Osterdahl, M. Kaminska, A. Al-Khalili, V. Zhaunerchyk, R. Thomas, M. af Ugglas, A. Kallberg, A. Simonsson, and M. Larsson. Dissociative recombination of nitrile ions: $DCCCN^+$ and $DCCCND^+$. *The Astrophysical Journal*, 613:1302–1309, 2004.

[89]Y. Osamura, K. Fukuzawa, R. Terzieva, and E. Herbst. A molecular orbital study of the $HC_3NH^+ + e^-$ dissociative recombination and its role in the production of cyanoacetylene isomers in interstellar clouds. *The Astrophysical Journal*, 519:697–704, 1999.

[90]D. Young, J. Berthelier, M. Blanc, J. Burch, A. Coates, R. Goldstein, M. Grande, T. Hill, R. Johnson, V. Kelha, D. Mccomas, E. Sittler, K. Svenes, K. Szego, P. Tanskanen, K. Ahola, D. Anderson, S. Bakshi, R. Baragiola, B. Barraclough, R. Black, S. Bolton, T. Booker, R. Bowman, P. Casey, F. Crary, D. Delapp, G. Dirks, N. Eaker, H. Funsten, J. Furman, J. Gosling, H. Hannula, C. Holmlund, H. Huomo, J. Illiano, P. Jensen, M. Johnson, D. Linder, T. Luntama, S. Maurice, K. Mccabe, K. Mursula, B. Narheim, J. Nordholt, A. Preece, J. Rudzki, A. Ruitberg, K. Smith, S. Szalai, M. Thomsen, K. Viherkanto, J. Vilppola, T. Vollmer, T. Wahl, M. Wuest, T. Ylikorpi, and C. Zinsmeyer. Cassini plasma spectrometer investigation. *Space Science Reviews*, 114:1–112, 2004.

[91]J.H. Waite, D.T. Young, T.E. Cravens, A.J. Coates, F.J. Crary, B. Magee, and J. Westlake. The process of tholin formation in Titan's upper atmosphere. *Science*, 316:870–875, 2007.

[92]R. Brown, J-P. Lebreton, and H. Waite. *Titan from Cassini-Huygens*. Springer, 2009.

[93]N. Carrasco, S. Plessis, and P. Pernot. Towards a reduction of the bimolecular reaction model for Titan ionosphere. *International Journal of Chemical Kinetics*, 40:699–709, 2008.

[94]V. Vuitton, R. V. Yelle, and M.J. McEwan. Ion chemistry and N-containing molecules in Titan's upper atmosphere. *Icarus*, 191:722–742, 2007.

[95]A. Coustenis, A. Salama, B. Schulz, S. Ott, E. Lellouch, T. Encrenaz, D. Gautier, and H. Feuchtgruber. Titan's atmosphere from ISO mid-infrared spectroscopy. *Icarus*, 161:383–403, 2003.

[96]A. Coustenis, R.K. Achterberg, B.J. Conrath, D.E. Jennings, A. Marten, D. Gautier, C.A. Nixon, F.M. Flasar, N.A. Teanby, B. Bézard, R.E. Samuelson, R.C. Carlson, E. Lellouch, G.L. Bjoraker, P.N. Romani, F.W. Taylor, P.G.J. Irwin, T. Fouchet, A. Hubert, G.S. Orton, V.G. Kunde, S. Vinatier, J. Mondellini, M.M. Abbas, and R. Courtin. The composition of Titan's stratosphere from Cassini/CIRS mid-infrared spectra. *Icarus*, 189:35–62, 2007.

[97]J.H. Waite, H. Niemann, R.V. Yelle, W.T. Kasprzak, T.E. Cravens, J.G. Luhmann., R.L. McNutt, W.-H. Ip, D. Gell, V. De La Haye, I. Muller-Wordag, B. Magee, N. Borggren, S. Ledvina, G. Fletcher, E. Walter, R. Miller, S. Scherer, R. Thorpe, J. Xu, B. Block, and K. Arnett. Ion Neutral Mass Spectrometer Results from the First Flyby of Titan. *Science*, 308:982–986, 2005.

[98]V. Vuitton, R.V. Yelle, and J. Cui. Formation and distribution of benzene on Titan. *J. Geophys. Res.*, 113:E05007, 2008.

[99]J. Semaniak, B.F. Minaev, A.M. Derkatch, F. Hellberg, A. Neau, S. Rosen, R. Thomas, M. Larsson, H. Danared, A. Paal, and M. af Ugglas. Dissociative recombination of $HCNH^+$: Absolute cross-sections and branching ratios. *The Astrophysical Journal Supplement Series*, 135:275–283, 2001.

[100]V. Vuitton, R.V. Yelle, and V.G. Anicich. The nitrogen chemistry of Titan's upper atmosphere revealed. *The Astrophysical Journal*, 647:L175–L178, 2006.

[101]Note that there is a running controversy about a systematic discrepancy factor of about 3 between ion densities measured by INMS and those measured by the CAPS instrument.[92] Whereas in our model the absolute production rates of neutrals species by DR scale as the squared value of this factor, all the results presented in this work are based on ratios

of production rates, and therefore totally independent of it.

[102]G. Israel, C. Szopa, F. Raulin, M. Cabane, H.B. Niemann, S.K. Atreya, S.J. Bauer, J.-F. Brun, E. Chassefière, P. Coll, E. Condé, D. Coscia, A. Hauchecorne, P. Millian, M.-J. Nguyen, T. Owen, W. Riedler, R.E. Samuelson, J.-M. Siguier, M. Steller, R. Sternberg, and C. Vidal-Madjar. Complex organic matter in Titan's atmospheric aerosols from in situ pyrolysis and analysis. *Nature*, 438:796–799, 2005.

[103]N. Carrasco, I. Schmitz-Afonso, J-Y. Bonnet, E. Quirico, R. Thissen, O. Dutuit, A. Bagag, O. Laprévote, A. Buch, A. Giulani, G. Adandé, F. Ouni, E. Hadamcik, C. Szopa, and G. Cernogora. Chemical characterization of Titan's tholins: solubility, morphology and molecular structure revisited. *Journal of Physical Chemistry A*, 113:11195–11203, 2009.

[104]P. Pernot, N. Carrasco, R. Thissen, and I. Schmitz-Afonso. Tholinomics - chemical analysis of nitrogen-rich polymers. *Analytical Chemistry*, 82:1371–1380, 2010.

[105]N. Balucani, F. Leonori, R. Petrucci, M. Stazi, D. Skouteris, M. Rosic, and P. Casavecchia. Formation of nitriles and imines in the atmosphere of Titan: combined crossed-beam and theoretical studies on the reaction dynamics of excited nitrogen atoms N(^{2}d) with ethane. *Faraday Discussions*, 147, 2010. In press.

[106]N. Balucani, A. Bergeat, L. Cartechini, G.G. Volpi, P. Casavecchia, D. Skouteris, and M. Rosi. Combined Crossed Molecular Beam and Theoretical Studies of the N(2D) + CH4 Reaction and Implications for Atmospheric Models of Titan. *The Journal of Physical Chemistry A*, 113:11138–11152, 2009.

[107]S. Lebonnois, E.L.O. Bakes, and C.P. McKay. Transition from gaseous compounds to aerosols in Titan's atmosphere. *Icarus*, 159(2):505–517, 2002.

[108]S. Lebonnois. Benzene and aerosol production in Titan and Jupiter's atmospheres: a sensitivity study. *Planetary and Space Science*, 53:486–497, 2005.

[109]P.P. Lavvas, A. Coustenis, and I.M. Vardavas. Coupling photochemistry with haze formation in Titan's atmosphere, Part II: results and validation with Cassini/Huygens data. *Planetary and Space Science*, 56:67–99, 2008.

[110]S. Lebonnois, E.L.O. Bakes, and C.P. McKay. Atomic and molecular hydrogen budget in Titan's atmosphere. *Icarus*, 161:474–485, 2003.

[111]H.L. DeWitt, M.G. Trainer, A.A. Pavlov, C.A. Hasenkopf, A.C. Aiken, J.L. Jimenez, C.P. McKay, O.B. Toon, and M.A. Tolbert. Reduction in haze formation rate on prebiotic Earth in the presence of hydrogen. *Astrobiology*, 9:447–453, 2009.

[112]E. Sciamma-O'Brien, N. Carrasco, C. Szopa, A. Buch, and G. Cernogora. Titan's atmosphere: an optimal gaz mixture for aerosol production ? *Icarus*, 2010. In press.

[113]V. Wakelam. KIDA: A Kinetic Database for Astrochemistry. In *American Astronomical Society Meeting Abstracts*, volume 214, page 402.15, May 2009.

[114]A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. Chapman & Hall, London, 1995.

[115]M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth Dover printing, tenth GPO printing edition, 1964.

[116]T.J. Santner, B.J. Williams, and W.I. Notz. *The Design and Analysis of Computer Experiments*. Springer-Verlag, New York, 2003.

[117]G.-L. Tian, H.-B. Fang, M. Tan, H. Qin, and M.-L. Tang. Uniform distributions in a class of convex polyhedrons with applications to drug combination studies. *Journal of Multivariate Analysis*, 100:1854–1865, 2009.

[118]K.-T. Fang and Z.-H. Yang. On uniform design of experiments with restricted mixtures and generation of uniform distribution on some domains. *Statistics & Probability Letters*, 46:113–120, 2000.