Radical detection in a methane plasma

H. D. Kang[a)] and V. Dose
Centre for Interdisciplinary Plasma Science, Max-Planck-Institut für Plasmaphysik, EURATOM Association, D-85748 Garching bei München, Germany

We present a method for detecting and quantifying the concentration of free radicals produced in a low temperature methane process plasma employing quadrupole mass spectrometry and Bayesian analysis. The latter includes a model comparison to determine species expected in the plasma and a parameter estimation of their concentrations and cracking coefficients. We find that the neutral mass spectra are best explained by incorporating the neutral molecules H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6 as well as the radicals H, CH_3, and C_2H_5 into the model. © *2003 American Vacuum Society.* [DOI: 10.1116/1.1621405]

Plasma-based surface processing has been widely used in the modern high tech industry, in particular in the microchip and display industry, where many manufacturing processes occur in plasma reactors. In most cases the form and the amount of reacting species are not sufficiently characterized and consequently the processing control relies on the heuristic approach. In this respect the identification and quantification of plasma products have become one of the urgent topics for plasma physicists. In particular, the detailed knowledge of concentrations of reactive particles like free radicals, especially in the immediate vicinity of a substrate, is needed to understand the underlying microprocesses at plasma/film boundary and thus to control manufacturing processes.[1] Many optical diagnostic tools to meet these demands, e.g., coherent anti-Stokes Raman spectroscopy,[2] laser-induced fluorescence spectroscopy,[3] infrared laser absorption spectroscopy,[4] and cavity ring down spectroscopy,[5] not only need expensive equipment but most of them are line-integrated and thus able to detect only the bulk gas-phase densities.

Mass spectrometry is a convenient technique to directly monitor the particle fluxes at the substrate sites. Its application is complicated by the fragmentation of analytes during the electron impact ionization. For overlapping cracking patterns subtraction methods have been devised to disentangle the measured spectra.[6] These methods suffer from excessive error buildup and are not applicable, when unstable constituents like radicals are assessed, due to the lack of cracking patterns. The derivation of radical concentration from ionization threshold mass spectrometry is also hampered by an absence of exact ionization cross sections.[7] An alternative rigorous analysis of composite mass spectra employs Bayesian probability theory[8,9] which also succeeds without exact cracking patterns. In this work we analyze quadrupole mass spectra from a capacitively coupled rf methane plasma using the Bayesian data analysis and concentrate on detection and quantification of radicals produced in the plasma. The derivation of our algorithm will be restricted to what is essential. For details of the Bayesian data analysis readers are referred to the recent reviews.[10,11]

Our stainless steel plasma reactor is similar in size to the Gaseous Electronic Conference (GEC) reference cell. The rf electrode consists of a pancake coil and is mounted into a dome-like dielectric window. The mass spectrometer unit (Balzers QMG422) is differentially pumped with a base pressure of $<10^{-7}$ hPa and connected with the plasma reactor by an orifice 50 μm in diameter. The mass spectra investigated in this work were taken differentially using a shutter equipped between the orifice and the ionization source of the mass spectrometer. In this way particles coming directly from the plasma could be separated from the rest gas. The measurement error was determined statistically and ranges from 0.2% to 10% depending on the absolute intensities of different mass channels. In order to collect a number of independent measurements, spectra were recorded under various discharge conditions through variation of the rf absorbed power and/or the residence time of the feed gas while the pressure in the reactor was fixed at 3 Pa. The spectra obtained result from a given number of species with varying relative concentrations. The number of truly independent measurements was inferred from the accumulated data set by a singular value decomposition and turned out to be 14 out of 20.[12] Though spectra of hydrocarbons extend out to 100 amu,[13] we limit our analysis to C_1H_x and C_2H_y molecules and radicals since already C_3 mass signals are very low and compare to the error of C_2 mass signals. Therefore our analysis deals with the stable molecules H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6 and the radicals H, CH, CH_2, CH_3, C_2H_3, and C_2H_5. We assume a linear response of the mass spectrometer and model the mass signal vector of measurement j, $\mathbf{d}_j$ as

$$\mathbf{d}_j = \mathbf{C}\mathbf{x}_j + \boldsymbol{\epsilon}_j, \tag{1}$$

where $\mathbf{C}$ is the cracking matrix containing the five columns for the stable molecules and zero to six columns for the radicals. $\mathbf{x}_j$ is the vector of species concentrations[14] in measurement j and $\boldsymbol{\epsilon}_j$ the vector of errors associated with the signal $\mathbf{d}_j$. The cracking column vectors are normalized to sum up to one. The same normalization has been applied to the data vector. It is easily shown that this implies the same sum norm for the concentrations. For the components of $\boldsymbol{\epsilon}_j$

[a)]Electronic mail: hyungdong.kang@ipp.mpg.de

we assume that their expectation value $\langle \varepsilon_{jk} \rangle$ is zero and that their true variance $\langle \varepsilon_{jk}^2 \rangle$ is equal to the measurement error s_{jk}^2. Under these conditions the sampling distribution of the data given all parameters known (likelihood) becomes—on the basis of the principle of maximum entropy—a Gaussian[15]

$$p(\mathbf{D}|\mathbf{X},\mathbf{C},\mathbf{S},E,I) = \prod_j \frac{1}{\Pi_n s_{nj}\sqrt{2\pi}}$$
$$\times \exp\left\{ -\frac{1}{2}(\mathbf{d}_j - \mathbf{C}\mathbf{x}_j)^T \mathbf{S}_j^{-2}(\mathbf{d}_j - \mathbf{C}\mathbf{x}_j) \right\}. \tag{2}$$

We have collected the data vectors including the calibration measurements on the stable molecules in matrix $\mathbf{D}$ and correspondingly the concentration vectors $\mathbf{x}_j$ in matrix $\mathbf{X}$. $\mathbf{S}$ denotes the ensemble of diagonal matrices $\mathbf{S}_j$ with components $(\mathbf{S}_j^{-2})_{nn} = 1/s_{nj}^2$. E is the number of species (molecules + radicals) which we incorporate into the model. The proper choice of E is the central topic of the Bayesian analysis of our data. If the data would be noise-free, e.g., $\epsilon_j = 0$, then E would be the number of species which is required for a perfect fit. The perfect fit on the other hand would be indicated by a vanishing residue

$$\sum_j (\mathbf{d}_j - \mathbf{C}\mathbf{x}_j)^T \mathbf{S}_j^{-2}(\mathbf{d}_j - \mathbf{C}\mathbf{x}_j) = 0. \tag{3}$$

The number of species needed to formulate the model would then be equal to the number of components in $\mathbf{d}_j$. The noise on the measurements complicates the situation considerably. What we need is a set of species large enough to reproduce the data $\mathbf{D}$ satisfactorily and small enough to avoid fitting noise. This optimization is called Occam's principle. It is a built-in feature of Bayesian analysis. The probability for a particular set of E species in the model is given in terms of the data $\mathbf{D}$ and variances $\mathbf{S}$ by Bayes theorem[15]

$$p(E|\mathbf{D},\mathbf{S},I) = \frac{p(E|I)p(\mathbf{D}|E,\mathbf{S},I)}{p(\mathbf{D}|\mathbf{S},I)}. \tag{4}$$

For the prior probability $p(E|I)$ we choose a constant, $p(E|I) = 1/E_{max}$. E_{max} is the maximal number of components in the model. The right-hand side of Eq. (4) is then equal to a constant times the marginal likelihood $p(\mathbf{D}|E,\mathbf{S},I)$. This is obtained from

$$p(\mathbf{D}|E,\mathbf{S},I) = \int d\mathbf{C}d\mathbf{X}p(\mathbf{C}|E,I)p(\mathbf{X}|E,I)p(\mathbf{D}|\mathbf{C},\mathbf{X},E,\mathbf{S},I). \tag{5}$$

The evaluation of this integral is involved and will be described in detail elsewhere. The logarithm of Eq. (5) is shown in Fig. 1 for selected combinations out of six radicals along with the global residue $\Sigma_{j,k}\varepsilon_{jk}^2/s_{jk}^2$ which is a measure of the misfit. While the misfit decreases monotonically as more radicals are incorporated into the model, the evidence $p(E|\mathbf{D},\mathbf{S},I)$ rises steeply by incorporating C_2H_5 into the model, attains a maximum for inclusion of C_2H_5, CH_3, and H, and decays slowly for more complicated models. The set of the three radicals is therefore *necessary and sufficient* for
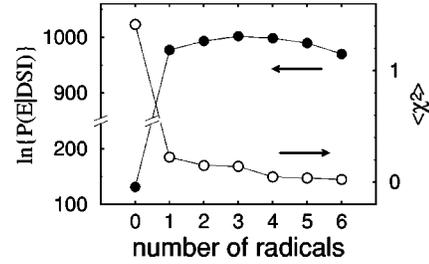

FIG. 1. Natural logarithm of $p(E|\mathbf{D},\mathbf{S},I)$ and the misfit $\langle \chi^2 \rangle$ between data and model for combinations of six free radicals (C_2H_5, CH_3, H, C_2H_3, CH, and CH_2). The x axis shows the number of radicals involved in the model, which were taken in the given order.

a satisfactory explanation of the data and anything beyond fits noise. This result is rather reasonable since these radicals are produced by breaking just one atomic bond from the stable and abundant molecules H_2, CH_4, and C_2H_6. In particular, the methyl radical has a very low sticking coefficient[16] ($\approx 10^{-5}$) and is therefore the most probable one to survive collisions in the plasma.

Having identified the number of species contributing to our set of measurements we proceed to the estimation of their relative concentrations under the chosen discharge conditions and their cracking patterns. Bayes theorem yields the required probability distribution

$$p(\mathbf{X},\mathbf{C}|\mathbf{D},\mathbf{S},E,I) = p(\mathbf{X}|E,I)p(\mathbf{C}|E,I)p(\mathbf{D}|\mathbf{C},\mathbf{X},\mathbf{S},E,I)/Z. \tag{6}$$

Z is the normalization factor. Therefore after specification of the prior probabilities $p(\mathbf{X}|E,I)$ and $p(\mathbf{C}|E,I)$ expectation values and variances for all elements of $\mathbf{X}$ and $\mathbf{C}$ can be calculated by Markov Chain Monte Carlo techniques.[17] Note that this approach does not even require the knowledge of the normalization constant Z. The prior knowledge about the cracking coefficients is drawn from the tables of Cornu and Massot for the stable molecules.[18] Such point estimates, though not directly applicable to the mass spectrometer in our experiment, can be used to formulate a prior distribution for the true, unknown cracking coefficients c in terms of the approximate tabulated value c_0 as an exponential on the support [0, 1] with expectation value $\langle c \rangle = c_0$.[9] Problematic is the choice of priors for the radicals. For CH_3 there is information available from an earlier experiment.[9] For C_2H_5 we use the tabulated pattern of C_2H_6 as a very coarse estimate. This is indeed vague prior knowledge, but if the data are sufficiently informative on the cracking pattern of C_2H_5 they will overrule the prior information completely. Prior information on the species concentrations $\mathbf{X}$ is drawn from the experience that in a CH_4 plasma the main neutral constituents are CH_4 ($\sim 90\%$) and H_2 ($\sim 10\%$). Reaction products like C_2H_6 and radicals are in the fractional percent range (0.1%). Again we code this information in exponential distributions.

Table I shows the cracking coefficients of CH_3 and C_2H_5 derived in this work in the column marked "posterior." The numbers in brackets are the standard deviations. The third column in the CH_3 table shows cracking coefficients ob-

TABLE I. Cracking coefficients of methyl(CH_3) and ethyle(C_2H_5) radicals determined from the CH_4 rf plasma.

Mass	CH_3			C_2H_5	
	posterior	prior	Ref. 9	posterior	prior
1	0.014(10)	0.007			
12	0.028(17)	0.008	0.008(1)		
13	0.087(36)	0.030	0.030(2)		
14	0.246(55)	0.261	0.263(11)	0.011(11)	0.021
15	0.625(68)	0.693	0.698(11)	0.010(10)	0.025
26				0.003(2)	0.131
27				0.376(8)	0.18
28				0.486(12)	0.54
29				0.115(9)	0.10

tained in an earlier experiment.[9] Note the fairly good agreement with the present results. Perfect agreement cannot be expected since the two experiments were run with different quadrupole spectrometers. For the same reason, we considered only the point estimates and not the standard deviation of the earlier experiment for setting up a prior for the present evaluation. There is a considerable difference in the precision of the CH_3 cracking coefficients obtained in this work compared to Ref. 9. The reason is simple. In this work we have to calculate the cracking pattern of CH_3 in a situation where the concentration of CH_4 which contributes to the same mass channels is 500 times higher than that of CH_3 ($\sim 0.2\%$). In the experiment of Ref. 9 these concentrations were nearly equal and slightly greater than 10%. The situation for C_2H_5 is more favorable since both C_2H_5 and C_2H_6 are formed by neutral-radical reactions in the plasma and turn out to have by and large the same concentrations. Moreover, there is no other species with overwhelming concentrations contributing to the mass channels of C_2H_5 and C_2H_6. This is reflected in
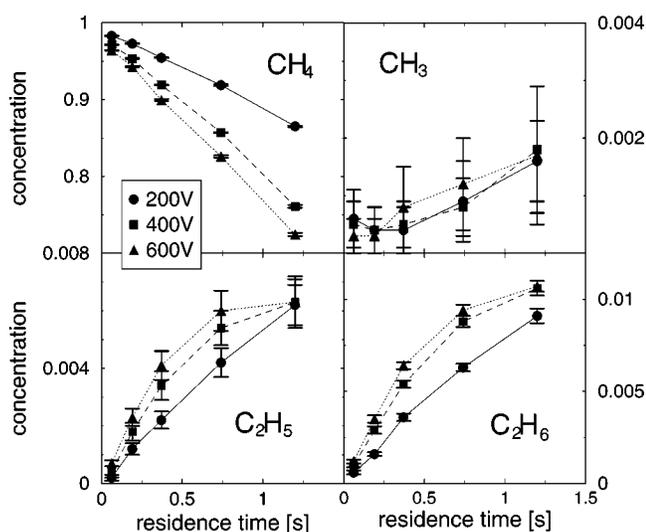
the rather precise determination of the C_2H_5 cracking coefficients even from trace concentration smaller than 1%.

The concentrations of CH_4, CH_3, C_2H_5, and C_2H_6 as a function of the residence time in the reactor are shown in Fig. 2 for three different values of the deposited rf power. The depletion of CH_4 upon increase of the two parameters is clearly visible. As expected the concentration of the reaction products increases with the residence time. The methyl radical exists to a very small amount ($<0.2\%$) and shows a delayed rise. From the presence of C_2H_5 and C_2H_6 it is clear that we must consider neutral-radical reactions as, for example, $CH_3 + CH_4 \rightarrow C_2H_6 + H$. Such reactions consume CH_3 for the build-up of higher molecules and radicals and therefore we observe an increase of the CH_3 concentration where the signals of C_2H_5 and C_2H_6 start to indicate saturation from a balance of production and loss. A final comment concerns errors. The errors on the CH_3 concentrations are considerably larger than those for C_2H_5 and C_2H_6. This reflects again the difficulty of detecting traces of radicals in a mother molecule environment 500 times more dense. In this case the conventional subtraction method will completely fail to handle the data since the measurement error is higher than 1/500.

In summary we have shown that detection and quantification of radicals effusing from a low temperature process plasma can be achieved by high precision quadrupole mass spectrometry combined with a fairly involved Bayesian analysis.

FIG. 2. Estimated concentrations of CH_4, CH_3, C_2H_5, and C_2H_6 as a function of the plasma power and the residence time. The number in the legend is the self-bias on the electrode, which is a monotonic function of the absorbed rf power.

[1]A. von Keudell, Thin Solid Films **402**, 1 (2002).
[2]S. O. Hay, W. C. Roman, and M. B. Colket, J. Mater. Res. **5**, 2387 (1990).
[3]G. Hancock, J. Sucksmith, and M. J. Toogood, J. Phys. Chem. **94**, 3269 (1990).
[4]M. Haverlag, E. Stoffels, W. W. Stoffels, G. M. W. Kroesen, and F. J. de Hooog, J. Vac. Sci. Technol. A **12**, 3102 (1994).
[5]A. O'keefe and D. Deacon, Rev. Sci. Instrum. **59**, 2544 (1988).
[6]R. Dobrozemsky and G. Schwarzinger, J. Vac. Sci. Technol. A **10**, 2661 (1992).
[7]T. Schwarz-Selinger, V. Dose, W. Jacob, and A. von Keudell, J. Vac. Sci. Technol. A **19**, 101 (2001).
[8]H. D. Kang, R. Preuss, T. Schwarz-Selinger, and V. Dose, J. Mass Spectrom. **37**, 748 (2002).
[9]T. Schwarz-Selinger, R. Preuss, V. Dose, and W. von der Linden, J. Mass Spectrom. **36**, 866 (2001).
[10]V. Dose, Rep. Prog. Phys. (to be published).
[11]G. D'Agostini, Rep. Prog. Phys. (to be published).
[12]W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, England, 1992).
[13]W. Jacob, Thin Solid Films **326**, 1 (1998).
[14]For the absolute composition of a mixture $\mathbf{x}_j$ has to be scaled by a species-dependent sensitivity factor. Its evaluation is beyond the scope of this work. We just call $\mathbf{x}_j$ the concentration of a mixture.
[15]D. Sivia, *Data Analysis: A Bayesian Tutorial* (Oxford University, New York, 1996).
[16]A. von Keudell, T. Schwarz-Selinger, and W. Jacob, Appl. Phys. Lett. **76**, 676 (2000).
[17]*Markov Chain Monte Carlo in Practice*, edited by W. Gilks, S. Richardson, and D. Spiegelhalter (Chapman and Hall, London, 1996).
[18]A. Cornu and R. Massot, *Compilation of Mass Spectral Data* (Heyden, London, 1979).