Interaction of CH$_3$ and H with amorphous hydrocarbon surfaces: estimation of reaction cross sections using Bayesian probability theory

Matthias Meier[1], Roland Preuss and Volker Dose

Centre for Interdisciplinary Plasma Science, Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstraße 2, 85748 Garching, Germany
E-mail: matthias.meier@ipp.mpg.de

Abstract. The interaction of CH$_3$ and H with amorphous hydrocarbon surfaces plays a central role during plasma deposition of such films. Recently, this interaction has been explored in particle beam experiments. A rate equation model has been proposed which explains the experimental observations on the basis of elementary surface reactions. This model includes several parameters which have the meaning of either a reaction cross section or a rate constant. The predictive power of the model and its applicability to more complex hydrocarbon deposition processes hinges on a reliable determination of the model parameters. In this paper, we develop a Bayesian analysis of the data. The result of this analysis are estimation distributions for each parameter rather than single numbers. We use this in-depth information to draw valuable conclusions about the ability of the model to describe the surface reactions. We find strong indications for a dependence of the reaction cross sections on the particles' angle of incidence.

---

[1] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Plasma deposited amorphous hydrocarbon films (a-C:H films) are commercially used for a large number of technological applications [1, 2]. These films combine many attractive properties such as infrared and optical transparency, chemical inertness and bio-compatibility. Depending on the deposition process they can have a very low friction coefficient and be extremely hard so that a-C:H films can be applied as wear protective coatings. Substantial empirical knowledge exists about the correlation between plasma parameters and resulting film properties. For example, it is known that typical hard films can be obtained by sufficient ion bombardment [3]–[5].

Although the empirical approach has proven quite successful in the past, the ultimate goal would be a microscopic understanding in terms of elementary reactions and reaction cross sections. This would allow us to tune a specific process in a *predictable* manner. Due to the complexity of the plasma and possible synergistic interactions between different species, particle beam experiments seem to be the only way to identify elementary reactions.

As a first step, the interaction of methyl radicals and atomic hydrogen with an a-C:H surface has been investigated recently using two radical beam sources [6]. A pronounced synergistic effect has been found: the effective sticking coefficient of methyl radicals increases by two orders of magnitude if a sufficiently high flux of atomic hydrogen is simultaneously directed towards the surface [7]–[10]. We will review the hydrogen flux dependence of the steady state growth rate as well as the dynamical behaviour of the growth rate upon switching on or off one of the radical beams. Based on these findings, a reaction scheme has been devised which breaks down the growth process into elementary surface reactions [6, 11] and leads to a simple rate equation model.

The main topic of this paper is the evaluation of the model and all experimental data using Bayesian probability theory. We show that this approach provides us with much more information than can be obtained by the conventional way, i.e. by simply fitting the model to the data: first, we can derive estimation distributions for each parameter, rather than only single numbers. Second, any kind of prior knowledge (e.g. values from literature) can be combined with the information that is contained in the experimental data. Third, we can consistently map
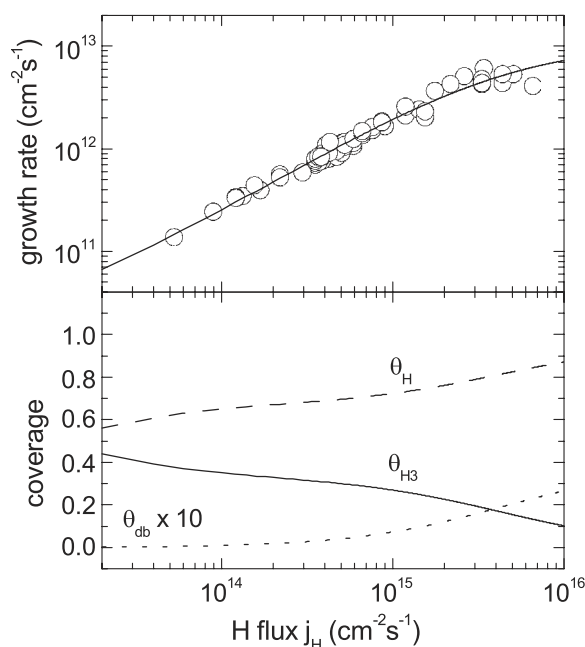
Figure 1. Steady state growth rate with simultaneous fluxes of CH_3 and H. For all data points, the CH_3 flux density is $j_C = 5.5 \times 10^{14}$ cm^{-2} s^{-1}. The solid curve results from a model calculation which involves three different types of surface site (lower panel). See the text for details.

distributions of experimental errors on the estimation of *distributions* of parameters. This avoids the truncation of information connected with simply calculating confidence intervals for the parameters from experimental error levels. The importance of this aspect is emphasized by the fact that the resulting parameter estimation distributions are in general asymmetric and cannot be characterized by a couple of numbers.

2. Growth synergism CH_3 + H

Hydrocarbon radicals hitting the surface of a hydrocarbon film are not necessarily very reactive: for methyl radicals an extremely low effective sticking coefficient of $10^{-5} \ldots 10^{-4}$ has been found [8, 12]. This quantity has been determined by ellipsometric measurement of the adsorption rate during exposure of an amorphous hydrocarbon film to a known flux density j_C of methyl radicals. The rate limiting step is the abstraction of surface-bonded hydrogen by impinging CH_3 (cross section σ_{abs}^C) which results in a dangling bond at the surface. Since these dangling bonds serve as chemisorption sites for CH_3 (chemisorption cross section σ_{add}^C), a very low value of σ_{abs}^C explains the low effective sticking coefficient. The rate for CH_3 chemisorption increases if an additional flux of atomic hydrogen j_H creates additional dangling bonds, again via hydrogen abstraction (cross section σ_{abs}^H) [8]–[10]. This enhancement of the effective CH_3 sticking coefficient is very pronounced, since the cross section for hydrogen abstraction is much larger for atomic hydrogen than for methyl radicals. For a constant flux density $j_C = 5.5 \times 10^{14}$ cm^{-2} s^{-1} of CH_3 radicals, the measured steady state growth rate is shown in figure 1 as a function of the hydrogen flux density j_H.

The calculated growth rate (solid curve in figure 1) results from a simple rate equation model which partitions the surface of the growing film into three types of surface site: Θ_{db} is the fraction of surface sites which carry a dangling bond, Θ_{H3} is the coverage of trihydride terminated sites (i.e. chemisorbed CH_3 radicals) and Θ_H refers to cross-linked surface sites. The model describes the incorporation of carbon atoms into the covalent network as a two-step process [11]: chemisorption of CH_3 (rate $j_C\sigma_{add}^C\Theta_{db}$) leads to a trihydride-terminated site (coverage Θ_{H3}). A continuation of the growth process then requires cross-linking of such hydrogen-rich surface groups in order to eliminate excess hydrogen from the film. This process is incorporated into the model as a direct process triggered by atomic hydrogen (rate $j_H\sigma_{elim}^H\Theta_{H3}$). This is justified by the fact that cross-linking occurs through recombination of dangling bonds which are in turn effectively created by atomic hydrogen. Also included in the model is the process of chemical erosion of the bulk film, which is described in the literature as thermally activated relaxation of a dangling bond via emission of a neighbouring hydrocarbon end group [13]–[15]. The rate for this process is $k_{eros}\Theta_{db}n_0$, where $n_0 \sim 1.35 \times 10^{15}$ cm^{-2} is the total number of surface sites per area (as deduced from the carbon density of such films [16]). The established model for chemical erosion by atomic hydrogen explicitly describes the formation of the dangling bond, but includes the rate for the preparation of end groups (this process involves breaking of CC bonds) implicitly in the rate constant k_{eros} and in the chemisorption and abstraction cross sections. Chemisorbed CH_3 groups are expected to be eroded faster, since they are already in the form of end groups: re-erosion only requires a dangling bond. Again, since it is hydrogen which creates dangling bonds effectively, the easiest way to incorporate this process into the model is a term $j_H\sigma_{eros}^H\Theta_{H3}$ which is proportional to the atomic hydrogen flux density. From these elementary reactions, the coverage state $\vec{\Theta} = (\Theta_{H3}, \Theta_H, \Theta_{db})$ of the surface can be calculated from the system of rate equations

$$\dot{\vec{\Theta}} = \mathbf{A}\vec{\Theta}, \tag{1}$$

where the matrix $\mathbf{A}$ has the following form:

$$\mathbf{A} = \begin{pmatrix} -a_1 & 0 & a_3 \\ a_1 & -a_2 & a_4 \\ 0 & a_2 & -a_3 - a_4 \end{pmatrix}$$

$$a_1 := j_H\sigma_{elim}^H + j_H\sigma_{eros}^H \qquad a_2 := j_C\sigma_{abs}^C + j_H\sigma_{abs}^H$$

$$a_3 := j_C\sigma_{add}^C \qquad\qquad a_4 := j_H\sigma_{add}^H + k_{eros}. \tag{2}$$

Since equation (1) is a linear differential equation, analytical time-dependent solutions $\vec{\Theta}(t)$ can be constructed from the eigenvectors and eigenvalues of the matrix $\mathbf{A}$. We remark that some of the terms in the model (cross-linking, erosion processes) do not represent elementary steps, but rather summarize a sequence of reaction steps. As an example, we discuss the cross-linking process: if we were to explicitly describe hydrogen abstraction from trihydride-terminated sites, we would have to keep track of the number of dihydride-terminated sites with a dangling bond and introduce a coverage Θ_{H2db}, i.e. our rate equation model would increase in dimension. Next, we would have to describe the rate for cross-linking as proportional to Θ_{H2db}^2, i.e. the model would become nonlinear. Therefore, we summarized this sequence of reactions by a single term in order to keep the model mathematically as simple as possible.

The processes which affect the balance of carbon atoms can be used to calculate the growth rate in terms of the number of incorporated or released carbon atoms per area and time:

$$\Gamma = n_0 (j_C \sigma_{add}^C \Theta_{db} - j_H \sigma_{eros}^H \Theta_{H3} - k_{eros} \Theta_{db}). \tag{3}$$

An important aspect of this model is the existence of a dynamic trihydride coverage of the surface (see the lower panel of figure 1): the model postulates that atomic hydrogen hitting a CH_3-covered area can only induce cross-linking. It cannot create chemisorption sites directly. This postulate is motivated by the fact that the growing film has optical properties that are typical for plasma-deposited polymer-like films. Such films have a stoichiometry of H:C $\sim$ 1 [16]. In view of the stoichiometry of the precursors CH_3 and H used in the radical beam experiment (H:C $>$ 3), we emphasized in our model the importance of the cross-linking process as a mechanism for eliminating the excess hydrogen from the film. In section 6 we will come back to the question of whether or not the cross-linking step is really *required* for further growth. In the framework of the proposed model, the trihydride coverage Θ_{H3} leads to a partial passivation of the surface against further growth. This effect is crucial for understanding the H-flux dependence of the steady state growth rate in figure 1. Furthermore, it explains the dynamical evolution of the growth rate that can be observed as the particle flux densities vary in time.

As an example, figure 2 shows a series of experiments where the methyl flux j_C is switched on at time $t = 0$ and switched off again a few thousand seconds later (we will call this type of experiment a 'CH$_3$-on/off' experiment). The flux of atomic hydrogen remains on during the whole experiment. The experimental time resolution is 5.9 s. In order to make the dynamical features more visible, we also display in figure 2 smoothed data (10-neighbour average) as a grey curve. The dynamics of the growth rate following the switching events is explained by the formation and removal of the trihydride coverage of the surface. Immediately after switching on the CH$_3$ flux, a large growth rate is observed. The number of dangling bonds at the surface is large as a result of the exclusive presence of H prior to the switching event. Due to the formation of passivating methyl surface groups, the growth rate then decreases towards a lower steady state value. After switching off the CH$_3$ flux, this trihydride coverage is removed by the processes of cross-linking and re-erosion. The latter reaction can be seen from the initial negative growth rate. The existence of a process of cross-linking into lower hydrides has been proven with infrared spectroscopy [6, 11]. The simulated growth rate shown in the middle panels of figure 2 demonstrates that the proposed model is capable of reproducing the experimentally observed dynamical behaviour.

For a detailed discussion of the experimental observations in the framework of the proposed model the reader is referred to our earlier publications [6, 11]. The above model is the simplest way to describe and explain the observations made by ellipsometry and infrared spectroscopy. On the other hand, it already requires seven parameters. Their proper estimation given the data is the non-trivial task which is the topic of this paper. Note that the model calculations shown in figures 1 and 2 already incorporate the results of our parameter estimation.

The measurements depicted in figures 1 and 2 are the only data material on which our analysis will be based. The experiments have been carried out under the following conditions: the steady state growth rates shown in figure 1 have been measured at a constant CH$_3$ flux density of 5.5×10^{14} cm^{-2} s^{-1}. The atomic hydrogen beam hits the film surface perpendicularly (90°) while the CH$_3$ radicals impinge at an angle of incidence of 45°. The 'CH$_3$-on/off' measurements (figure 2) have been performed with reversed geometry of the radical sources, i.e. with 90° and 45° angles of incidence for the CH$_3$ and the H beams, respectively. This experiment has been performed four times using different methyl fluxes (see figure 2). The flux density of atomic hydrogen is always $j_H = 9.0 \times 10^{14}$ cm^{-2} s^{-1}.
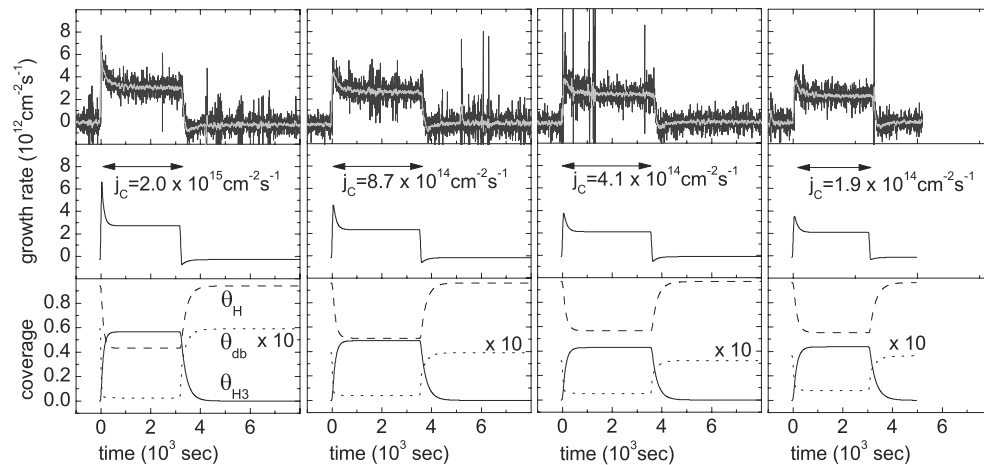
Figure 2. Time-resolved growth rate. An atomic hydrogen flux of $j_H = 9.0 \times 10^{14}$ cm^{-2} s^{-1} continuously interacts with the film surface. Starting at time $t = 0$, an additional CH$_3$ beam is switched on for the time period indicated by the arrow. The experiment is performed four times using different CH$_3$ flux densities as marked.

3. Bayesian data analysis

In this section we briefly describe the idea of Bayesian probability theory and its application to data analysis. An excellent introduction to this topic is given e.g. in the book of Sivia [17]. In the following we will use the notation $P(A|I)$ for the probability for the proposition A to be true. Although we will continue to speak of 'probabilities', it might be useful for the reader to think of P in terms of a 'measure of our belief' in A being true, especially if A represents a hypothesis rather than some random variable. Since a measure of belief is always conditional on some background knowledge I, we shall always display this information. In our notation, the background information I is separated from A by the vertical bar. We now ask for the probability for two propositions to be true simultaneously. In view of the following discussion we call these propositions H and D. This joint probability $P(H, D|I)$ can be factorized such that one of the propositions becomes part of the condition (i.e. moves to the right of the vertical bar). Due to the symmetry with respect to H and D, this can be done in two ways:

$$P(H, D|I) = P(D|I)P(H|D, I) = P(H|I)P(D|H, I). \tag{4}$$

This equation is called the *product rule*. We rearrange the second equality to get the so-called *Bayes theorem*:

$$P(H|D, I) = \frac{P(D|H, I)P(H|I)}{P(D|I)}. \tag{5}$$

The *sum rule* states that the probabilities of a proposition H and the proposition that H is false (i.e. for the proposition $\bar{H}$) add up to unity:

$$P(H|I) + P(\bar{H}|I) = 1. \tag{6}$$

By applying the sum rule to $P(H, D|I)$ and $P(\bar{H}, D|I)$, equation (4) can be used to obtain the so-called *marginalization rule*

$$P(D|I) = P(H, D|I) + P(\bar{H}, D|I). \tag{7}$$

As a generalization, the sum might not only involve a proposition H and its negative counterpart $\bar{H}$, but a continuous set of mutually exclusive and exhaustive possibilities for H. The marginalization rule then reads

$$P(D|I) = \int P(H, D|I)\, dH. \tag{8}$$

Bayes' theorem (equation (5)) and the *marginalization rule* (equation (7)) are the basis of Bayesian data analysis. For this purpose, we associate with D the meaning of a set of experimental data (growth rates as measured with ellipsometry). H shall be the hypothesis that the model parameters (reaction cross sections and rate constants) assume a certain value. Any additional background information (e.g. the information that the model itself is appropriate) shall be summarized by I. The second term in the numerator of equation (5) is called the *prior probability*, because it encodes the knowledge about the parameter values independent from the measured data, e.g. knowledge existing *prior* to the actual experiment. The first term in the numerator of equation (5) is the *likelihood function*. It represents the probability of measuring the data set D given H, e.g. given specific values for the model parameters. The result of this analysis is the *posterior probability* on the left-hand side. Equation (5) can be used for the purpose of parameter estimation by displaying it as a function of H: the left-hand side then can be viewed as a posterior *distribution*. As such it displays how much belief we can put on each value of the model parameters, given the data D as the outcome of our experiments. In that context, the denominator $P(D|I)$ (called the *global likelihood*, or *evidence*) on the right-hand side of equation (5) merely represents a normalization constant. Before we continue, we would like to add some important remarks concerning the difference between Bayesian parameter estimation and 'conventional' parameter determination.

The conventional way of assigning numerical values to the model parameters is to perform a 'least-squares fit'. By doing so, one ignores that in general, expert knowledge exists. The simplest example for expert knowledge is that reaction cross sections have to be positive. Within the Bayesian approach, this expert knowledge enters the analysis as prior distribution $P(H|I)$ and as such contributes to the shape of the posterior distribution. From simply fitting the model to the data (i.e. maximizing the likelihood), it might very well happen that some parameters appear to assume negative values. In that case, one would still apply expert knowledge, but in a rigorous and destructive way: one would reject the whole fitting result as being 'unphysical'. Consequently, one overlooks the existence of solutions that *are* 'physical'. Second, we consider it a trivial statement that displaying posterior distributions is much more informative and constructive than merely arriving at single numbers. Numbers are a handy way to communicate results. However, this compression of information is in many cases misleading: for example, a best fit value is meaningless if there are other local maxima of the posterior distribution that are comparable in height with the global maximum. An expectation value is meaningless e.g. in a case where the posterior distribution consist of two well separated peaks: the expectation value then lies somewhere in between these peaks and therefore represents a rather improbable parameter value. Finally, the conventional way of expressing the uncertainty with respect to some experimental variable is to assign an error bar to the best fit parameter values, e.g. by applying the Gaussian law of error propagation. We will show that within the Bayesian framework, an experimental variable can be formally treated as a model parameter and its uncertainty can be described by an appropriate prior distribution. From this formalism it follows that experimental sources of error

modify the shape of the posterior distribution and in general even its *position*. In that sense it seems not appropriate to separate the determination of best fit values from the 'error calculation'.

4. Likelihood

We associate with D the set of measured data (growth rates) $\{d_i\}$. We first comment on the forward calculation using the model, since this is necessary in order to evaluate the likelihood function. As already mentioned, we will consider two different kinds of data sets: in one case the index i labels individual measurements of the steady state growth rate which are carried out for different particle flux densities $j_i := (j_{C_i}, j_{H_i})$ (we use this notation despite the fact that the methyl flux was actually held constant). The expected steady state growth rate $\Gamma_i := \Gamma(j_i)$ can be calculated from equations (1) and (3) as an algebraic solution by setting $\dot{\vec{\Theta}} = 0$. If the data set $\{d_i\}$ represents a time-resolved measurement like the ones shown in figure 2, i enumerates the growth rate as measured at discrete time steps $t_i = \Delta t \times i$. As already mentioned, the time resolution was $\Delta t = 5.9$ s. The growth rate following the onset of the CH_3 beam is modelled by the time-dependent solution to equations (1) and (3) with the boundary conditions $\Theta(t = 0) = \Theta_0$, where Θ_0 is the steady state solution for $t < 0$. The analogous calculation has to be carried out for the growth dynamics following the switching-off of the CH_3 beam. In either of the two cases, the modelled growth rate Γ_i is a function of the particle flux densities j_i and the model parameters. In order to avoid confusion in notation with distribution variances, we do not use reaction cross sections σ_{reaction} but reaction probabilities $p_{\text{reaction}} = \sigma_{\text{reaction}}/A$ where $A = n_0^{-1}$ is the average area of a surface site. We combine all microscopic parameters in a vector p which shall be defined as

$$p := (p_{\text{add}}^C, p_{\text{abs}}^C, p_{\text{eros}}^H, p_{\text{elim}}^H, k_{\text{eros}}, p_{\text{add}}^H, p_{\text{abs}}^H). \tag{9}$$

We now construct the likelihood function for a single data point d_i. The error ϵ_i of the ellipsometric measurement causes a difference between the expected growth rate Γ_i and the observed growth rate d_i:

$$d_i = \Gamma_i + \epsilon_i. \tag{10}$$

We assume that on average the measurement error vanishes but the variance has a finite value σ_i^2:

$$\langle \epsilon_i \rangle = 0 \tag{11}$$
$$\langle \epsilon_i^2 \rangle = \sigma_i^2. \tag{12}$$

An estimation of the variance σ_i^2 is given by the experimentally observed scattering of the data. From all normalizable distributions respecting these moment conditions, the principle of maximum entropy singles out the Gaussian distribution as the one containing the least amount of additional information [17]:

$$P(\epsilon_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right). \tag{13}$$

The likelihood function for a single datum d_i is obtained by replacing ϵ_i with $d_i - \Gamma_i$,

$$P(d_i|p, j_i, I) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(d_i - \Gamma_i)^2}{2\sigma_i^2}\right). \tag{14}$$

A simplification is given by the fact that in the course of an experiment the variance remains constant ($\sigma_i \equiv \sigma$). The likelihood function of equation (14) reflects experimental errors with respect to the ellipsometric measurement. It still ignores the uncertainty with respect to the particle flux densities. This uncertainty can be included in the following way: in the above notation j_i are the particle flux densities which actually hit the sample surface. However, only the expectation values j_i^0 are known from the operating parameters of the particle beam sources and from calibration experiments. Therefore, in the above likelihood function j_i has to be replaced by j_i^0. The actual fluxes j_i are parameters which are necessary for performing the model calculation but their estimation is of no intrinsic interest. Integration eliminates such *nuisance parameters* as variables but the effect of their existence on the considered distribution is properly accounted for:

$$P(d_i|\boldsymbol{p}, j_i^0, I) = \int P(d_i, j_i|\boldsymbol{p}, j_i^0, I) \, dj_i \tag{15}$$

$$= \int P(d_i|\boldsymbol{p}, j_i, j_i^0, I) P(j_i|\boldsymbol{p}, j_i^0, I) \, dj_i. \tag{16}$$

In the second line, we used the product rule with respect to the variables d_i and j_i. As already mentioned, the first factor of the integrand is equal to the right-hand side of equation (14). The second factor $P(j_i|\boldsymbol{p}, j_i^0, I) \equiv P(j_i|j_i^0, I)$ is the (two-dimensional) probability distribution for the actual flux densities given their expectation values. The same argumentation holds as for the case of the measurement error: for the distribution of the error $\epsilon_j = \boldsymbol{j} - \boldsymbol{j^0}$ with respect to the flux densities of methyl radicals and hydrogen atoms we assume a zero expectation value and finite second moments σ_C^2, σ_H^2. The least informative distribution fulfilling these requirements is the (two-dimensional) Gaussian distribution

$$P(j_i|j_i^0, I) = \frac{1}{2\pi} \frac{1}{\sigma_C \sigma_H} \exp\left(-\frac{(j_i - j_i^0)^{\mathrm{T}} \cdot \mathbf{S}_j^{-2} \cdot (j_i - j_i^0)}{2} \right), \tag{17}$$

where the reciprocal variances $\sigma_C^{-2}, \sigma_H^{-2}$ are the non-zero elements of the diagonal matrix $\mathbf{S}_j^{-2}$. The integral in equation (16) can be calculated analytically if the model function $\Gamma_i \equiv \Gamma(j_i)$ in equation (14) is expanded around j_i^0 up to first order:

$$\Gamma(j_i) \approx \Gamma(j_i^0) + \boldsymbol{\nabla}\Gamma(j_i^0)^{\mathrm{T}} \cdot \delta j_i \tag{18}$$

$$=: \Gamma_i^0 + (\boldsymbol{\nabla}\Gamma_i^0)^{\mathrm{T}} \cdot \delta j_i. \tag{19}$$

For simplicity we introduced the notations

$$\Gamma_i^0 := \Gamma(j_i^0) \tag{20}$$

$$(\boldsymbol{\nabla}\Gamma_i^0)^{\mathrm{T}} := \left(\frac{\partial \Gamma}{\partial j_C}(j_i^0), \frac{\partial \Gamma}{\partial j_H}(j_i^0) \right). \tag{21}$$

In this approximation, the exponent of the integrand in equation (16) is a quadratic form in $\delta j_i := j_i - j_i^0$ which can be rearranged as

$$-\frac{1}{2\sigma_i^2}\{(d_i - \Gamma_i^0 - (\boldsymbol{\nabla}\Gamma_i^0)^{\mathrm{T}} \cdot \delta j_i)^2 + \delta j_i^{\mathrm{T}} \cdot \sigma^2 \mathbf{S}_j^{-2} \cdot \delta j_i\}$$

$$= : -\frac{1}{2\sigma_i^2}\{(\delta j_i - \delta j_i^0)^{\mathrm{T}} \cdot \mathbf{Q} \cdot (\delta j_i - \delta j_i^0) + R\}. \tag{22}$$

We remark that the left-hand side is not a complete second order approximation for $(d_i - \Gamma_i)^2$. Equation (22) does not contain the term $(2\sigma_i^2)^{-1}(d_i - \Gamma_i^0)\delta j_i^T \mathbf{H} \delta j_i$ which arises if the expansion of Γ_i in equation (19) were continued up to second order ($\mathbf{H}$ being the Hessian matrix of Γ with respect to j). However, in practice this term can be neglected for the following reason: the factor $d_i - \Gamma_i^0$ is the random measurement error, which can have either sign. Evaluation of a whole data set $\{d_i\}$ (see below) will involve the summation over the index i so that the terms containing the second derivatives of Γ should tend to cancel out. From the quadratic form in equation (22), the integral can be evaluated as

$$P(d_i|\mathbf{p}, j_i^0, I) = \frac{1}{\sqrt{2\pi}\sigma_i} \frac{1}{2\pi\sigma_C\sigma_H} \frac{2\pi\sigma_i^2}{\sqrt{\det \mathbf{Q}}} \exp\left(-\frac{R}{2\sigma_i^2}\right). \tag{23}$$

The matrix $\mathbf{Q}$, the displacement δj_i^0 and the residue R are obtained by comparing coefficients on both sides of equation (22):

$$\mathbf{Q} = \nabla\Gamma_i^0 \cdot (\nabla\Gamma_i^0)^T + \sigma^2 \mathbf{S}_j^{-2} \tag{24}$$

$$\delta j_i^0 = (d_i - \Gamma_i^0)(\mathbf{Q}^{-1})^T \cdot \nabla\Gamma_i^0 \tag{25}$$

$$R = (d_i - \Gamma_i^0)^2 \{1 - (\nabla\Gamma_i^0)^T \cdot (\mathbf{Q}^{-1})^T \cdot \nabla\Gamma_i^0\}. \tag{26}$$

A straightforward calculation of R and $\det \mathbf{Q}$ leads to

$$P(d_i|\mathbf{p}, j_i^0, I) = \frac{1}{\sqrt{2\pi}\sigma_{\text{eff}\,i}} \exp\left(-\frac{(d_i - \Gamma^0)^2}{2\sigma_{\text{eff}\,i}^2}\right) \tag{27}$$

$$\sigma_{i,\text{eff}}^2 := \sigma_i^2 + \left(\frac{\partial\Gamma_i^0}{\partial j_C}\right)^2 \sigma_C^2 + \left(\frac{\partial\Gamma_i^0}{\partial j_H}\right)^2 \sigma_H^2. \tag{28}$$

From this result for a single data point d_i we can now construct a likelihood function for a whole data set $\{d_i\}$:

$$P(\{d_i\}|\mathbf{p}, \{j_i^0\}, I) = \prod_i \int P(d_i|\mathbf{p}, j_i, j_i^0, I) P(j_i|\mathbf{p}, j_i^0, I)\, dj_i. \tag{29}$$

Two cases have to be distinguished.

Case (i). We apply equation (29) to the set of steady state growth rates in figure (1). The individual measurements i can then be considered independent from each other. Each integral can be evaluated as shown above and the likelihood for this data set can be written as

$$P(\{d_i\}|\mathbf{p}, \{j_i^0\}, I) = \prod_i P(d_i|\mathbf{p}, j_i^0, I), \tag{30}$$

where each factor has the form given in equations (27) and (28).

Case (ii). By contrast, if $\{d_i\}$ ($i = 1, \ldots, N$) represents a time resolved measurement (the index i corresponding to time) as in figure 2, one has to account for the fact that the set $\{j_i\}$ of particle flux densities is highly correlated: after the switching event the flux densities remain constant for all subsequent i. This knowledge can be expressed by the distributions $P(j_i|j_i^0, I)$:

$$P(j_1|j_1^0, I) = \frac{1}{2\pi\sigma_C\sigma_H} \exp\left(-\frac{(j_1 - j_0)^T \cdot S_j^{-2} \cdot (j_1 - j_0)}{2}\right) \tag{31}$$

$$P(j_i|j_i^0, I) = \delta(j_i - j) \qquad (i \neq 1). \tag{32}$$

Here $\delta(.)$ is the delta function. Hence, in equation (29) integration over j_i $(i \neq 1)$ is trivial and we need to evaluate

$$
\begin{aligned}
P(\{d_i\}|p, \{j_i^0\}, I) &= \int \left\{\prod_{i=1}^{N} P(d_i|p, j_1, I)\right\} P(j_1|j_1^0, I) \, dj_1 \\
&= \left\{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i}\right\} \int \exp\left(-\sum_i \frac{(d_i - \Gamma_i)^2}{2\sigma_i^2}\right) P(j_1|j_1^0, I) \, dj_1 \\
&= \left\{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i}\right\} \frac{1}{2\pi\sigma_C\sigma_H} \int \exp(A(j_1)) \, dj_1.
\end{aligned}
\tag{33}
$$

We simplify notation by introducing dimensionless growth rates

$$d_i' := d_i/\sigma_i, \qquad \Gamma_i' := \Gamma_i/\sigma_i. \tag{34}$$

The exponent of the integrand in equation (33) can then be rewritten as

$$A(j_1) = -\frac{\|\{d_i' - \Gamma_i'\}\|^2}{2} - \frac{(j_1 - j_0)^T \cdot S_j^{-2} \cdot (j_1 - j_0)}{2}. \tag{35}$$

In analogy to equation (22), each component of the vector $\{\Gamma_i'\} = \{\Gamma_i'(j_1)\}$ is expanded around $j_1 = j_0$

$$\{\Gamma_i'\} \approx \{\Gamma_i'^0\} + \{\nabla\Gamma_i'\} \cdot \delta j_1 \tag{36}$$

so that $A(j_1)$ is a quadratic form in $\delta j_1 = j_1 - j_0$,

$$A(j_1) =: -\frac{1}{2\sigma^2}\{(\delta j_1 - \delta j_1^0)^T \cdot Q \cdot (\delta j_1 - \delta j_1^0) + R\}. \tag{37}$$

As explained earlier, this quadratic form neglects a second order term containing the Hessian matrix of Γ with respect to j. The likelihood function can be written as

$$P(\{d_i\}|p, \{j_i^0\}, I) = \left\{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i}\right\} \frac{1}{2\pi\sigma_C\sigma_H} \frac{2\pi}{\sqrt{\det Q}} \exp\left(-\frac{R}{2}\right). \tag{38}$$

Comparing coefficients between equations (35), (36) and (37) yields

$$Q = \begin{pmatrix} \|\{(\frac{\partial\Gamma'}{\partial j_C})_i\}\|^2 + \sigma_C^{-2} & \{(\frac{\partial\Gamma'}{\partial j_C})_i\}^T\{(\frac{\partial\Gamma'}{\partial j_H})_i\} \\ \{(\frac{\partial\Gamma'}{\partial j_H})_i\}^T\{(\frac{\partial\Gamma'}{\partial j_C})_i\} & \|\{(\frac{\partial\Gamma'}{\partial j_H})_i\}\|^2 + \sigma_H^{-2} \end{pmatrix}$$

$$\delta j^0 = Q^{-1} \cdot \{\nabla\Gamma_i'\}^T \cdot \{d_i' - \Gamma_i'^0\} \tag{39}$$

$$R = \|\{d_i' - \Gamma_i'^0\}\|^2 - \{d_i' - \Gamma_i'^0\}^T \cdot \{\nabla\Gamma_i'\} \cdot Q^{-1} \cdot \{\nabla\Gamma_i'\}^T \cdot \{d_i' - \Gamma_i'^0\}. \tag{40}$$

In order to reduce computer time, we implemented this result after approximating the partial derivatives $\nabla\Gamma_i'$ by their steady state values for $t \to \infty$. The difference from case (i) immediately becomes clear by considering the example of an N-fold repetition of measuring a steady state growth rate $\Gamma_i^0 \equiv \Gamma^0$. In that case, the likelihood reduces to

$$P(\{d_i\}|\boldsymbol{p}, \{\boldsymbol{j}_i^0\}, I) \propto \exp\left(-\frac{(\bar{d} - \Gamma^0)^2}{2\sigma_{\text{eff}}{}^2}\right) \tag{41}$$

$$\sigma_{\text{eff}}{}^2 := \frac{\sigma^2}{N} + \left(\frac{\partial\Gamma^0}{\partial j_C}\right)^2 \sigma_C^2 + \left(\frac{\partial\Gamma^0}{\partial j_H}\right)^2 \sigma_H^2. \tag{42}$$

The best estimate for the expectation value Γ^0 is the measured mean value $\bar{d} = N^{-1}\sum_i d_i$. It can be seen that the contribution of the particle flux uncertainty to the effective quadratic width of this distribution does not scale with N^{-1} as does the contribution of the measurement error. In that sense, case (ii) takes into account that by increasing the number of measurements the particle flux error cannot be averaged out since it is a systematic one.

5. Prior distributions

In order to apply Bayes' theorem (equation (5)) the prior distribution $P(H|I)$ has to be specified. We associate with H the vector $\boldsymbol{p} = \{p_k\}$ ($k = 1\ldots 7$) containing the reaction probabilities and rate constants (see equation (9)). Prior knowledge about each of these model parameters has to be encoded by a probability distribution $P(p_k|I)$. Due to logical independence the prior distribution $P(H|I) \equiv P(\boldsymbol{p}|I)$ factorizes,

$$P(\boldsymbol{p}|I) = \prod_k P(p_k|I). \tag{43}$$

We apply the principle of maximum entropy and distinguish between two cases:

(A) we require a parameter p_k to have a certain expectation value $p_k^{(1)}$. In that case, an exponential function

$$P(p_k|\lambda, I) = \frac{1}{\lambda}\exp\left(-\frac{p_k}{\lambda}\right) \tag{44}$$

is the appropriate distribution with $\lambda := p_k^{(1)}$ if $P(p_k|I)$ is defined on the support $0 \le p_k < \infty$.

(B) In addition to an expectation value $p_k^{(1)}$, we want to specify the uncertainty of the parameter p_k by a certain value of its second central moment $p_k^{(2)}$ (also known as variance). The principle of maximum entropy then leads to a Gaussian distribution

$$P(p_k|a, b, I) = \frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(p_k - a)^2}{2b^2}\right). \tag{45}$$

On the support $-\infty < p_k < \infty$, a and b^2 are the first and second central moments. However, we know that the variable p_k represents either a reaction cross section or a reaction rate constant and as such a positive quantity. Therefore, on the support $0 \le p_k < \infty$, a and b have to be chosen such that

$$\frac{1}{p_k^{(0)}} \int_0^\infty p_k \exp\left(-\frac{(p_k - a)^2}{2b^2}\right) \mathrm{d}p_k = p_k^{(1)} \tag{46}$$

$$\frac{1}{p_k^{(0)}} \int_0^\infty (p_k - p_k^{(1)})^2 \exp\left(-\frac{(p_k - a)^2}{2b^2}\right) \mathrm{d}p_k = p_k^{(2)} \tag{47}$$

where the constant $p_k^{(0)}$ normalizes the distribution $P(p_k|I)$. This system of equations can be solved e.g. by numerical iteration taking $(p_k^{(1)}, p_k^{(2)})$ as a starting point for (a, b). We now apply either case (A) or case (B) to each of the model parameters p_k:

$p_{\mathrm{add}}^{\mathrm{H}}$, $p_{\mathrm{abs}}^{\mathrm{H}}$: quantitative knowledge about these two parameters can be found in the literature. The group of Küppers studied chemical erosion of amorphous hydrocarbon films with atomic hydrogen [14]. From a microscopic modelling of their results, they derived the reaction cross sections $\sigma_{\mathrm{add}}^{\mathrm{H}} = 1.3$ Å^2 and $\sigma_{\mathrm{abs}}^{\mathrm{H}} = 0.05$ Å^2. Given the area density of surface sites of $n_0 = 1.35 \times 10^{15}$ cm^{-2} s^{-1}, this corresponds to reaction probabilities of $p_{\mathrm{add}}^{\mathrm{H}} = 0.175$ and $p_{\mathrm{abs}}^{\mathrm{H}} = 6.75 \times 10^{-3}$. Although only a point estimate (i.e. an estimation in terms of a single numerical value), we do not apply case (A) for the following reasons: the exponential distribution of equation (44) is maximized by the value $p_k = 0$. On the other hand, published parameter values are commonly obtained from a *best fit* i.e. they maximize some sort of likelihood. In addition, the relative uncertainty of experimentally determined quantities is not completely unknown (as case (A) would imply) but in most cases is in the range of 0 to 100%. We therefore use the mentioned values for $p_{\mathrm{add}}^{\mathrm{H}}$ ($p_{\mathrm{abs}}^{\mathrm{H}}$) as the first moments in case (B) together with second moments of 50% (60%). As mentioned, the parameters of the Gaussian distribution differ from its first and second moments due to the non-negativity of the variable: we arrive at $(a, b) = (1.66 \times 10^{-1}, 0.96 \times 10^{-1})$ for the parameter $p_{\mathrm{add}}^{\mathrm{H}}$ and $(a, b) = (5.5 \times 10^{-3}, 5.0 \times 10^{-3})$ for the parameter $p_{\mathrm{abs}}^{\mathrm{H}}$.

k_{eros}: as a prior distribution for this parameter, we take the posterior distribution obtained from Bayesian analysis of an experiment where the CH$_3$ beam is switched off and only atomic hydrogen interacts with the film surface. Such an experiment offers the most direct access to k_{eros}, since the only parameters remaining for the model calculation are $p_{\mathrm{add}}^{\mathrm{H}}$, $p_{\mathrm{abs}}^{\mathrm{H}}$ and k_{eros}. We analysed one of the corresponding sections in the series of CH$_3$-switching experiments, taking a flat prior distribution for k_{eros}. A nice Gaussian posterior distribution for k_{eros} resulted with the maximum $a = 3.5 \times 10^{-3}$ s^{-1} and a width $b = 1 \times 10^{-3}$ s^{-1}. Referring to case (B), we take this Gaussian as the common prior distribution for all data sets to be analysed. Note that this parameter does not depend on the angle of incidence of the particle beams (which was different for the two data sets to be analysed): this parameter appears only in a reaction rate that does not scale with any particle flux density.

$p_{\mathrm{elim}}^{\mathrm{H}}$, $p_{\mathrm{eros}}^{\mathrm{H}}$: the only information about these parameters exists in terms of consistency with the model picture: in the framework of our model, both parameters belong to processes which are triggered by hydrogen abstraction due to impinging atomic hydrogen. Accordingly, we take the expectation value $\langle p_{\mathrm{abs}}^{\mathrm{H}} \rangle = 6.75 \times 10^{-3}$ as a point estimate for the parameters $p_{\mathrm{elim}}^{\mathrm{H}}$ and $p_{\mathrm{eros}}^{\mathrm{H}}$. We apply case (A) in order to express our ignorance about the confidence of this estimate.

$p_{\mathrm{abs}}^{\mathrm{C}}$: the most direct access to this quantity is the observation of the adsorption rate of CH$_3$ radicals without additional atomic hydrogen flux. In that case hydrogen abstraction by incoming CH$_3$ is the rate-limiting step and the only channel for creating dangling bonds. Von Keudell *et al* [8] performed CH$_3$-adsorption experiments with an angle of incidence of 45° [18]

(in contrast to the description in [8] which quoted an angle of incidence of 90°). He derived a value of $\sigma_{abs}^C = 1.6 \times 10^{-3}$ (which corresponds to $p_{abs}^C = 2 \times 10^{-4}$). By contrast, we performed CH_3-adsorption experiments at normal incidence and found a value of $p_{abs}^C = 2 \times 10^{-5}$. We apply case (A) and match the geometries of the CH_3-adsorption experiments with those of the experiments under investigation here: we take $\langle p_{abs}^C \rangle = 2 \times 10^{-4}$ as the expectation value for the data set consisting of the steady state growth rate as a function of the atomic hydrogen flux. For the time-resolved measurements (CH_3 radicals hit the film surface perpendicularly) we choose $\langle p_{abs}^C \rangle = 2 \times 10^{-5}$ as the expectation value.

p_{add}^C: in the case of the steady state measurements in figure 1, we express full ignorance about the value of this parameter by taking the prior distribution to be flat, i.e. $P(p_{add}^C|I) = \text{constant}$, but restricted to positive values ($P(p_{add}^C|I) = 0$ for $p_{add}^C < 0$). As we will show, a posterior expectation value of $\langle p_{add}^C \rangle = 1.22$ results. We shall take this expectation value as prior knowledge (case (A)) for evaluating the set of CH_3-on/off experiments (figure 2). This is a reasonable choice because of the following factor: the flux ratio $j_H : j_C$ in those experiments is so low that the limiting factor for the chemisorption rate $j_C p_{add}^C \Theta_{db}$ is the coverage with dangling bonds Θ_{db}, the main source of which is hydrogen abstraction by atomic hydrogen. In other words, the influence of the factor $j_C p_{add}^C$ on the net growth rate is only weak, as can be already seen directly from the data in figure 2: the steady state growth rate in the four measurements changes only slightly although j_C is varied within one order of magnitude. Likewise, a rather flat posterior distribution for the parameter p_{add}^C would result if no additional information existed. However, since the collection of steady state growth rates in figure 1 supplies us with such information, it seems to suggest itself to take an exponential prior distribution with expectation value $\langle p_{add}^C \rangle = 1.22$ as a reference.

6. Results

We briefly summarize our considerations concerning the application of Bayes' theorem to our problem: the full information about the value of the model parameters in the light of the data and prior knowledge is contained in the posterior distribution

$$P(\boldsymbol{p}|\{d_i\}, \{\boldsymbol{j}_i^0\}, I) \propto P(\{d_i\}|\boldsymbol{p}, \{\boldsymbol{j}_i^0\}, I) \times P(\boldsymbol{p}|I). \tag{48}$$

We have discussed the adaptation of likelihood function $P(\{d_i\}|\boldsymbol{p}, \{\boldsymbol{j}_i^0\}, I)$ and prior distributions $P(\boldsymbol{p}|I)$ for the data set of steady state growth rates (figure 1), respectively for the time-resolved 'CH_3-on/off' measurements shown in figure 2. In addition to the measurement error, we took into account the uncertainty with respect to the particle beam flux densities by analytic marginalization. From this combined posterior distribution for the vector $\boldsymbol{p}$, we now construct posterior distributions $P(p_k|\{d_i\}, \{\boldsymbol{j}_i^0\}, I)$ for each microscopic parameter p_k by marginalization:

$$P(p_k|\{d_i\}, \{\boldsymbol{j}_i^0\}, I) = \int P(\boldsymbol{p}|\{d_i\}, \{\boldsymbol{j}_i^0\}, I)\, \mathrm{d}p_{l \neq k}. \tag{49}$$

As indicated by the notation $p_{l \neq k}$, the integration extends over all parameters except p_k. This integration is done numerically with a Monte Carlo Markov chain (MCMC) computer code [19]: a random walk $\{\boldsymbol{p}_n\}$ in parameter space is constructed such that the sampling density is given by $P(\boldsymbol{p}|\{d_i\}, \{\boldsymbol{j}_i^0\}, I)$. The marginal posterior distribution $P(p_k|\{d_i\}, \{\boldsymbol{j}_i^0\}, I)$ can then be obtained from plotting the kth component of the random walk $\{\boldsymbol{p}_n\}_k$ as a histogram. We calculated 5×10^4 steps for each data set so that stationary conditions were achieved.

6.1. Steady state growth rates versus time-resolved measurements

We first estimate the model parameters from the set of steady state growth rates given in figure 1. In figure 3, the marginal posterior distributions for the seven microscopic parameters are displayed. Comparing the histograms with the prior distributions (dashed curves) gives an estimate of how much information beyond prior knowledge is contained in the data. For this comparison, we have also quoted the expectation value of each marginal posterior distribution (the first moments of the prior distributions were discussed in the previous section).

Apparently, not much new information about the parameters k_{eros}, p_{eros}^H, p_{abs}^C and p_{add}^H can be inferred from the data. This can be easily explained: the posterior distribution for k_{eros} is almost identical to the prior estimate, because the rate for erosion of bulk material $n_0 k_{eros} \Theta_{db}$ only makes a minor contribution to the net growth rate. This can be seen already from the 'CH$_3$-off' phase in the time-resolved measurements, where this erosion rate turns out to be below 5×10^{11} cm^{-2} s^{-1}. If an additional flux of methyl radicals is present (as is the case for the data of figure 1), the contribution of the term $n_0 k_{eros} \Theta_{db}$ is actually much less, because CH$_3$ chemisorption decreases the number of dangling bonds Θ_{db}. Likewise, the data are not informative with respect to p_{abs}^C. According to our prior information, the cross section for hydrogen abstraction by CH$_3$ is much smaller than the one for hydrogen abstraction by H. The latter process therefore dominates the rate for dangling bond creation, i.e. the reactivity of the surface. We conclude that the process of hydrogen abstraction by CH$_3$ would be somewhat unnecessary if we were to explain solely the data given in figure 1. The data make also only a tiny change in our knowledge about the parameter p_{eros}^H. This is not surprising since from observing the net growth rate it is difficult to tell whether the rate for CH$_3$ chemisorption is partly compensated by erosion of chemisorbed CH$_3$ groups. In principle this distinction should be possible due to the fact that the former process depends on $j_C \Theta_{db}$ whereas the latter scales with $j_H \Theta_{H3}$. However, the observed H-flux dependence seems to be not significant enough. Finally, a similar argument holds for the parameter p_{add}^H: the data do not allow us to distinguish the rate for dangling bond creation from the rate for dangling bond annihilation (via chemisorption by H). Although the first process scales with Θ_H and the second one with Θ_{db}, the flux dependent data contain no information that would substantially deviate from what is known *a priori* about the cross section p_{add}^H.

By contrast, the experiment is obviously quite informative with respect to the cross sections p_{abs}^H, p_{elim}^H and p_{add}^C. These parameters correspond to elementary reactions which the model proposes to be necessary even if only steady state growth rates are to be explained: hydrogen abstraction by H is the main source that provides dangling bonds, i.e. chemisorption sites. From the observation of nonzero steady state growth rates follows the parameter p_{abs}^H being strictly positive. Indeed, the posterior distribution indicates a finite minimum value for p_{abs}^H despite the fact that our prior distribution erroneously gives a small chance for the possibility of p_{abs}^H being zero. In the light of the data, our prior knowledge is also overruled in the sense that the Gaussian shape is not preserved. Also the process of cross-linking of chemisorbed CH$_3$ groups into lower hydrides is mandatory within the framework of the model. Again, the prior distribution for the corresponding parameter (p_{elim}^H) does not exclude the possibility $p_{elim}^H = 0$; in fact it even ascribes the highest probability to this case. Data analysis tells us, however, that values less than $p_{elim}^H = 2.5 \times 10^{-3}$ can be safely ruled out and that the expectation value has to be corrected from $\langle p_{elim}^H \rangle = 6.75 \times 10^{-3}$ to a higher value of $\langle p_{elim}^H \rangle = 9.5 \times 10^{-3}$. The most explicit information obtained from the experimental observations results for the cross section for CH$_3$ chemisorption. Although nothing was known *a priori* about the parameter p_{add}^C except non-negativity, Bayesian
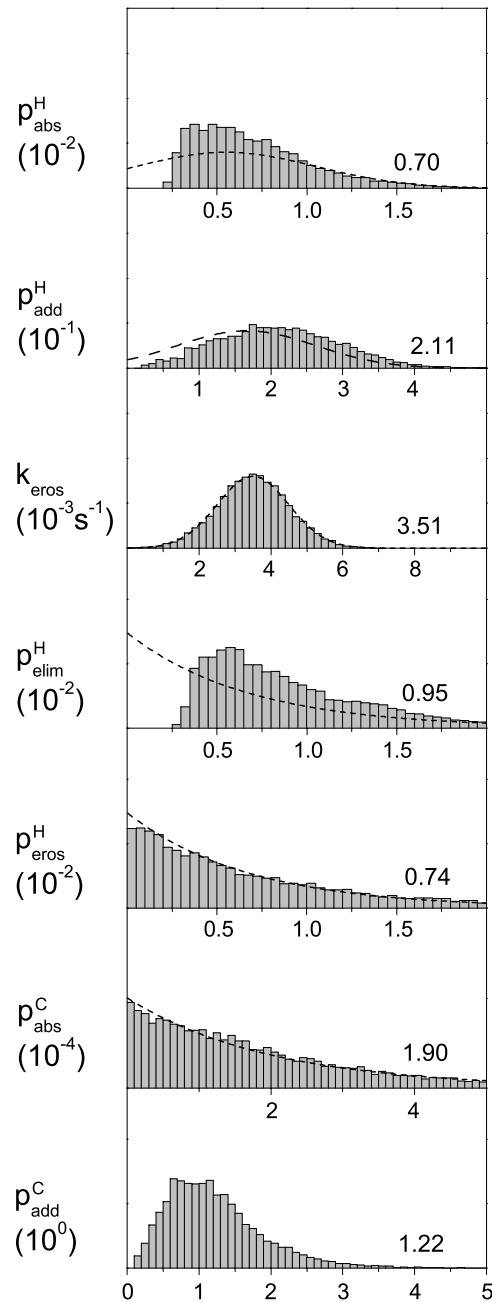
Figure 3. Bayesian analysis of the steady state growth rate in figure 1: the marginal posterior distributions of the seven model parameters are given by the histograms. Corresponding expectation values are shown. Prior distributions are drawn as dashed curves.

parameter estimation provides detailed information: p_{add}^C is of the order of unity with expectation value $\langle p_{add}^C \rangle = 1.22$. The posterior distribution has a non-Gaussian, asymmetric shape with a pronounced right wing.

We now focus on estimating the model parameters from the set of time-resolved measurements shown in figure 2. We analysed the four experiments separately for reasons which

will become clear later. Nevertheless, for the moment we would like to draw the reader's attention to the posterior distribution features displayed in figure 4 (e.g. the shape) that are common to all four measurements. Similar to the previous case, the prior distributions of the parameters k_{eros}, p_{abs}^C and p_{add}^H are more or less preserved by the posterior distributions. According to the prior information, the cross section for hydrogen abstraction by CH_3 is even smaller compared to the steady state data (again we remark that the angle of incidence for the methyl radicals is different in the two cases). Consequently, there is even less information that the measurements would add to the existing knowledge. The Gaussian shape of the prior distribution for k_{eros} survives at least qualitatively after analysing the data. A more pronounced deviation between prior and posterior distribution can be seen for the cross sections p_{add}^H and p_{abs}^H. As already mentioned, a lower flux ratio $j_H : j_C$ puts more weight on the process of dangling bond creation via hydrogen abstraction by H as being rate-limiting. The data are therefore informative about the cross section p_{abs}^H. Furthermore, it is the ratio of p_{abs}^H and p_{add}^H which determines the number of dangling bonds which are available at time $t = 0$ when the CH_3 flux is switched on. In that sense the data also provide some information about p_{add}^H by means of the initial surface reactivity at $t = 0$. The posterior distributions for p_{abs}^H and p_{add}^H broaden out as the flux ratio $j_H : j_C$ is increased from the left to the right. Consistently with what is said above, the trend has the opposite direction for p_{add}^C: from the experiment with $j_C = 2.0 \times 10^{15}$ cm^{-2} s^{-1} one can basically only learn that the expectation value $\langle p_{add}^C \rangle$ is larger than the prior point estimate (namely $\langle p_{add}^C \rangle = 1.4$ instead of $\langle p_{add}^C \rangle = 1.22$) and that—contrary to the prior information—$p_{add}^C = 0$ is certainly not the most likely value. However, as the flux ratio $j_H : j_C$ is increased from the left to the right, the measurements get more and more explicit about the cross section for CH_3 chemisorption being different from the prior distribution.

Obviously, the cross sections p_{elim}^H and p_{eros}^H can be determined very precisely from the Bayesian analysis of the 'CH_3-on/off' experiments: the resulting posterior distributions are very narrow (note that the abscissa has been stretched compared to figure 3!) and there is hardly anything left from the asymmetry of the exponential prior distributions. This resembles the fact that the cross section p_{elim}^H has a strong influence on the slowing down of the initially fast growth process following the onset of the CH_3 beam as well as the resulting steady state level. The cross section p_{eros}^H is responsible for the increased etching rate immediately after switching the CH_3 beam off. We compare the posterior distribution for the parameter p_{eros}^H with that for the parameter p_{abs}^H: from all four data sets we derive the expectation value $\langle p_{eros}^H \rangle$ to be substantially smaller (by a factor of about 5–10) than the expectation value $\langle p_{abs}^H \rangle$, although *a priori* we had assumed equal expectation values. Since there is no overlap between the posterior distributions for p_{eros}^H and those for p_{abs}^H, one can even conclude with confidence that p_{eros}^H has a smaller value than p_{abs}^H. We can think of two possible reasons for this finding. In our model, the reaction probability p_{eros}^H is meant to summarize a sequence of reactions: first, atomic hydrogen abstracts surface-bonded hydrogen. The resulting dangling bond recombines via emission of a neighbouring surface-bonded CH_3 group. We assumed the first step to be rate limiting. Consequently, we chose the prior distribution for p_{eros}^H to have the same expectation value as that for p_{abs}^H. The posterior result '$p_{eros}^H < p_{abs}^H$' can be interpreted as overruling our prior assumption: if the rate-limiting step is not the abstraction reaction but the activated release of the neighbouring surface-bonded CH_3 group, then the total reaction probability is indeed overestimated by the reaction probability for the first reaction step. Alternatively, one might speculate about the existence of a different reaction pathway, e.g. the direct abstraction of surface-bonded CH_3 by atomic hydrogen: in the case of CH_3 groups chemically bonded to a Si(100) surface, a beam of atomic hydrogen leads
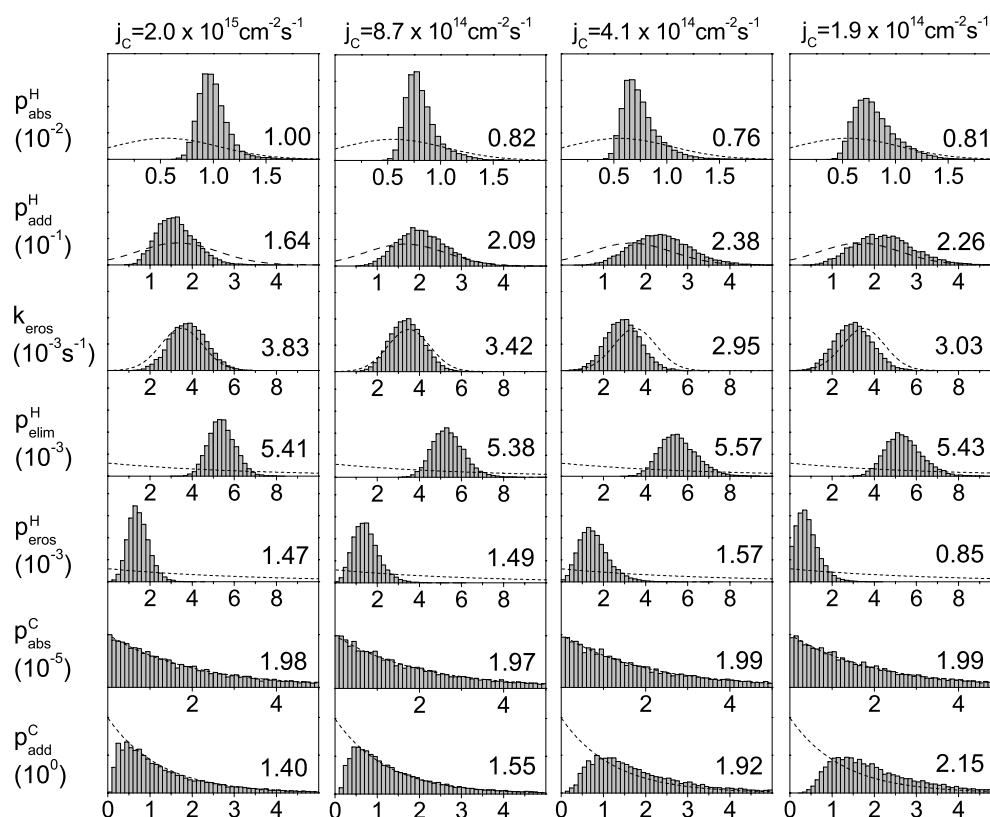
Figure 4. Bayesian analysis of the series of time-resolved measurements in figure 2: each of the four panels belongs to the corresponding panel in figure 2. The marginal posterior distributions of the seven model parameters are given by the histograms. Corresponding expectation values are given as a number. Prior distributions are drawn as dashed curves.

to the formation of CH_3SiH_3 as the only carbon-containing erosion product [20]. However, Loh *et al* [21] and Foord *et al* [22] showed by mass spectrometry that CD_3 groups bonded to a polycrystalline diamond surface can indeed be abstracted by atomic hydrogen as CD_xH_{4-x} molecules. A rate equation model for the isotopic exchange at the surface and the abstraction of surface-bonded CD_xH_{3-x} species by H was set up. From comparing the model with the mass spectrometric measurements of CD_xH_{4-x}, the probability for hydrogen abstraction was found to be five times larger than the probability for abstraction of a methyl group. This finding is consistent with our posterior estimation of the parameters p_{eros}^H and p_{abs}^H.

6.2. Beyond the present model

In this section, we compare the posterior distributions resulting from the four 'CH$_3$-on/off' experiments (figure 4) with each other. Our model is devised to predict the time-dependent growth rate as a function of the time-dependent flux densities $\Gamma(t) = \Gamma(j_C(t), j_H(t))$. We proposed the microscopic cross sections p to be a set of constants. The model therefore simplifies reality by claiming that the complex interaction of CH_3 and H at the surface of an amorphous hydrocarbon film can be completely described on the basis of a few (seven) numbers. *Conversely, we can*

judge the quality of this simplification from the degree to which the parameter estimations are (or are not) flux dependent.

Before we discuss systematic variations with respect to varying fluxes, we note that some of these trends may be inconsistent with the fourth data set (corresponding to $j_C = 1.9 \times 10^{14}$ cm^{-2} s^{-1}). However, this can be traced back to the fact that this fourth measurement has been terminated too early: the 'CH$_3$-off' phase has not yet safely reached a steady state. The distinction between the initial phase of erosion of chemisorbed CH$_3$ groups and the phase of steady state bulk erosion is somewhat ambiguous; consequently, so is the determination of the corresponding parameters p_{eros}^H and k_{eros}. This also has an influence on the parameters p_{abs}^H and p_{add}^H.

As figure 4 shows, the posterior distribution for p_{abs}^H shifts to larger values as the CH$_3$ flux is increased. This can be seen from the corresponding expectation values or the p_{abs}^H-values which maximize the posterior distributions (we have already explained why the distribution also gets broader). We attribute this effect to the simplification that the model makes about possible chemisorption sites: the model requires the process of cross-linking as mandatory. It ignores the possibility that the sequence of dangling bond creation and CH$_3$ chemisorption could also take place directly at chemisorbed CH$_3$ groups which are not yet cross-linked with each other. However, this process might indeed happen as the CH$_3$ flux gets more and more intense compared to the atomic hydrogen flux. In this case, dangling bonds created at methyl groups have no time to recombine with each other. Therefore, if the measured growth rate is to be explained exclusively by abstraction/chemisorption events at cross-linked sites (Θ_H), then a larger cross section for dangling bond creation (p_{abs}^H) is needed as Θ_H decreases, i.e. as the CH$_3$ flux is increased. Consistent with this explanation, it is no surprise that the distribution for k_{eros} moves to higher values with increasing CH$_3$-flux j_C (i.e. from right to left in figure 4): a less cross-linked hydrocarbon film is expected to be eroded more easily by atomic hydrogen, since fewer CC bonds need to be broken in order to produce the necessary end groups. Indeed, the measured steady-state erosion rates during the 'CH$_3$-off' phases increase from right to left in figure 4, although the scaling of the graph makes it impossible to discern this effect. As already explained in section 2, the bond-breaking process is phenomenologically accounted for by the value of the parameter k_{eros}. Summarizing, if during the 'CH$_3$-on' phase the CH$_3$ flux is higher, a less cross-linked film grows that can be eroded at a higher rate in the subsequent 'CH$_3$-off' erosion phase, as is reflected by a higher value of k_{eros}. In that sense, the marginal posterior distributions reveal that the model makes an approximation by describing the structural properties of the growing bulk film as being independent from the fluxes j_C and j_H.

A pronounced flux dependence can be discerned for the parameters describing chemisorption cross sections: the posterior distributions both for CH$_3$ chemisorption (p_{add}^C) and for H chemisorption (p_{add}^H) shift to lower values upon increasing the CH$_3$ flux, already indicated by the expectation values. Analogous to the preceding discussion, we interpret this trend as arising not from the fluxes themselves, but from the resulting coverages on the surface, i.e. the chemisorption cross sections decrease for larger values of Θ_{H3}. This finding supports theoretical studies of Träskelin *et al* [23]: molecular dynamics simulations based on a tight-binding model showed that the cross section for chemisorption of methyl radicals at an isolated dangling bond on a monohydride-terminated diamond(111) surface is $\sigma_{add}^C = (11 \pm 2)$ Å^2. The authors find a guiding effect by surrounding surface-bonded hydrogen as responsible for this cross section being larger than the surface site area (5.7 Å^2). But if the dangling bond is surrounded by three neighbouring methyl groups, then the chemisorption cross section is

only $\sigma_{\text{add}}^{\text{C}} = (0.2 \pm 0.1)$ Å^2. This means that the attractive potential of a dangling bond can be substantially shielded by nearby CH_3 groups. As a result of our Bayesian analysis we found an important effect: each term appearing in the mathematical formulation of our model depends linearly on the coverages Θ_{H3}, Θ_{H} and Θ_{db}. The linearity of the model with respect to the coverages is both conceptually attractive (it contains the minimum number of parameters to describe the experimental observations qualitatively) and very advantageous in terms of numerical effort (there exists an analytic solution to the set of rate equations). However, this approach cannot account for any proximity effects like shielding of a dangling bond by neighbouring groups.

6.3. Angular dependence

In this section we try to answer the question of whether the reaction cross sections of the model depend on the angle of incidence of the impinging reactants. Indeed, one pronounced angular dependence was known *a priori*: the expected cross section for hydrogen abstraction by CH_3 is $\langle p_{\text{abs}}^{\text{C}} \rangle \sim 2 \times 10^{-4}$ for an angle of incidence of 45° (steady state data, figures 1 and 3), whereas for normal incidence this cross section has a substantially smaller value of $\langle p_{\text{abs}}^{\text{C}} \rangle \sim 2 \times 10^{-5}$ (time-resolved data, figures 2 and 4). This has been derived from measuring the adsorption rate during exposure of the film surface to a quantified beam of CH_3 (without additional hydrogen flux) and a simple rate equation model [8, 12]. In these CH_3-adsorption experiments, the reaction probability $p_{\text{abs}}^{\text{C}}$ is the rate limiting parameter. The experiments discussed here bear no new information about $p_{\text{abs}}^{\text{C}}$, as we have shown above. It can be expected that the cross section for CH_3-chemisorption also depends on the angle of incidence. Alfonso and Ulloa [24] performed molecular dynamics simulations of CH_3 chemisorption on a (2×2)-reconstructed (100) diamond surface with 75% of the surface sites being hydrogen terminated. At perpendicular incidence, CH_3 chemisorbed in 50% of all cases. At an angle of incidence of 45°, only 10% of the attempts lead to a chemisorption event. The authors point out that the chemisorption cross section at 45° is smaller due to the lower amount of kinetic energy of the motion component perpendicular to the surface. One cannot expect that the results of a calculation on a crystalline model surface with a very small number of surface sites can be transferred quantitatively on the real surface of an amorphous hydrocarbon film. But qualitatively this effect appears in the posterior estimates of the chemisorption cross section $p_{\text{add}}^{\text{C}}$: we showed that from the steady state data an expectation value of $\langle p_{\text{add}}^{\text{C}} \rangle = 1.22$ can be derived for CH_3 radicals hitting the film surface at 45°. For perpendicular angle of incidence, we conclude from Bayesian analysis of the 'CH$_3$-on/off' experiments substantially larger expectation values of $\langle p_{\text{add}}^{\text{C}} \rangle \geq 1.40$. Recall that the chemisorption cross section $p_{\text{add}}^{\text{C}}$ decreases for a decreasing flux ratio $j_{\text{H}}:j_{\text{C}}$. But the higher expectation values $\langle p_{\text{add}}^{\text{C}} \rangle$ for perpendicular incidence have *not* been determined from experiments employing higher flux ratios $j_{\text{H}}:j_{\text{C}}$. The steady state data of figure 1 cover flux ratios of roughly $0.1 < j_{\text{H}}:j_{\text{C}} < 15$; the flux ratios used in the time-resolved measurements of figure 2 are $j_{\text{H}}:j_{\text{C}} = (0.45, 1.03, 2.20, 4.74)$, i.e. they lie within that interval. We can therefore identify the angular dependence and the flux dependence (which is better described as coverage dependence) of $\langle p_{\text{add}}^{\text{C}} \rangle$ as two distinct effects.

It is important to note that in the preceding discussion we talked about parameter *expectation values* and found a dependence on the angle of particle incidence. This conclusion can be drawn not only for the parameters $p_{\text{abs}}^{\text{C}}$ and $p_{\text{add}}^{\text{C}}$, but also for other model parameters. For example, for the parameter $p_{\text{eros}}^{\text{H}}$ we find a smaller expectation value for hydrogen atoms hitting the surface

at an angle of 45° compared to normal incidence. However, given the spread of the posterior distributions, one cannot put much confidence in the existence of an angular dependence. In the following, we will present additional experimental data and investigate whether these can increase our confidence in the predictive power of the expectation values we have derived so far.

Let $\langle p \rangle_1$ be the set of the seven parameter expectation values (see figure 3) which we obtained from analysing the steady state data of figure 1. We will refer to the corresponding experimental geometry (hydrogen atoms hit the film surface perpendicularly, CH_3 under 45°) as geometry 1. Likewise, $\langle p \rangle_2$ shall denote the set of expectation values we derived from the leftmost time-resolved experiment of figure 2 (see figure 4). The corresponding geometry is interchanged with respect to the H and CH_3 beam and will be referred to as geometry 2. Using these parameter sets, we will try to simulate a time-resolved measurement which has been performed for both geometries. The time-resolved measurement in the left-hand panel of figure 5 shows the growth rate in an experiment where a hydrogen flux of $j_H = 1.4 \times 10^{15}$ cm^{-2} s^{-1} is switched on at time $t = 0$. During the whole experiment, a CH_3 flux of $j_C = 2 \times 10^{15}$ cm^{-2} s^{-1} interacts with the film surface. This experiment was performed in geometry 2. A similar experiment was done by Schwarz-Selinger [25] in geometry 1. It is shown in the right-hand panel of figure 5. Here, the methyl flux is $j_C = 2.1 \times 10^{14}$ cm^{-2} s^{-1} and at time $t = 0$ an atomic hydrogen flux of $j_H = 4 \times 10^{15}$ cm^{-2} s^{-1} is switched on. An initial negative spike appears before synergistic growth sets in. This phenomenon can be easily explained in terms of our proposed model [11]: as a result of the interaction with the CH_3 beam, the film surface is fully covered with chemisorbed methyl groups ($\Theta_{H3} = 1$). As soon as the hydrogen flux is switched on, the processes of cross linking (probability p_{elim}^H) and etching of trihydride-terminated groups (probability p_{eros}^H) sets in. Both processes lower the coverage Θ_{H3} which is necessary for steady state net growth. We now ask why the latter process appears directly in the experiment shown on the right-hand panel, whereas no indication for an initial negative growth rate can be seen in the left-hand panel. This question can be answered by simulating these experiments by a model calculation for the time-dependent particle fluxes used in both experiments. We remark that for experimental reasons there is always a small flux of atomic hydrogen remaining in the 'H-off' phase: the CH_3-radical source also emits hydrogen molecules which can be dissociated at the hot filaments and capillary of the atomic hydrogen source. Consequently, switching off the H_2 flow through the latter device is not sufficient in order to exclude any atomic hydrogen from the process. The following simulations therefore assume a flux of $j_H = 1 \times 10^{13}$ cm^{-2} s^{-1} for times $t < 0$.

In the upper panels of figure 5, we show as black solid curves the model calculations which result for the parameter set $\langle p \rangle_2$. This parameter set assumes the geometry used in the experiment shown on the left-hand side. This experiment is perfectly reproduced: the model calculation confirms that there is no initial etching spike prior to the onset of synergistic growth. But according to the model, there also should not be a spike for the time-dependent fluxes used in the experiment on the right-hand side! In addition, the calculation overestimates the resulting steady state growth rate significantly. In the lower panels of figure 5, we repeat these simulations, but with $\langle p \rangle_1$ as parameter values. Now the model erroneously predicts an initial etching phase for the experiment on the left-hand side, but perfectly matches the experiment on the right-hand side, both with respect to the etching effect and the synergistic growth rate in steady state.

It is important to note that the experiments in figure 5 have not been used in the course of our parameter estimation procedure. We obtained our knowledge about the microscopic parameters from steady state measurements and time-resolved 'CH_3-on/off' measurements. We apply this knowledge here to a different type of experiment, a 'H-on' experiment with its own dynamical
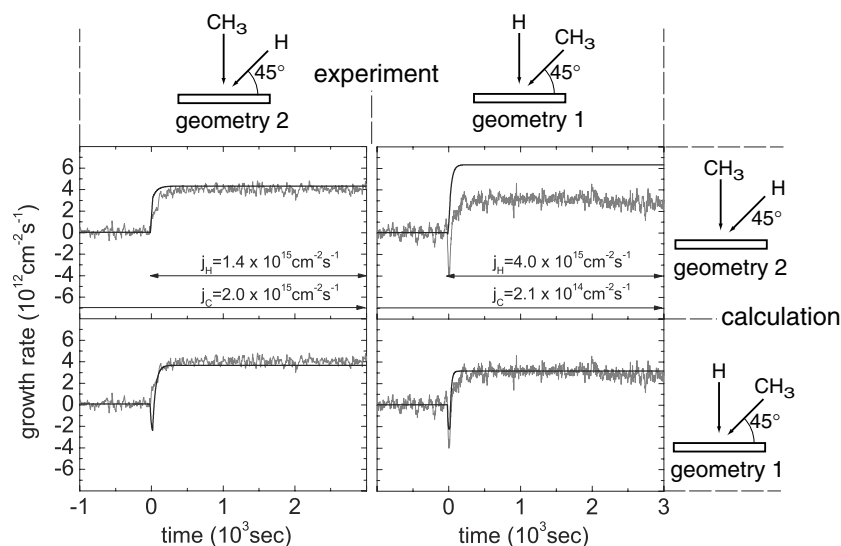
Figure 5. Time-resolved growth rates. In the left (right) panels, a flux of atomic hydrogen of $j_H = 1.4 \times 10^{15}$ cm^{-2} s^{-1} ($j_H = 4.0 \times 10^{15}$ cm^{-2} s^{-1}) is switched on at time $t = 0$, while a methyl flux of $j_C = 2.0 \times 10^{15}$ cm^{-2} s^{-1} ($j_C = 2.1 \times 10^{14}$ cm^{-2} s^{-1}) continuously interacts with the surface. The model calculation (black solid curve) in the upper panels (lower panels) is performed using the set of parameter expectation values determined for geometry 2 (geometry 1).

characteristics. The calculations in figure 5 show that the growth rate behaves differently in the two experiments not because different time-dependent fluxes have been used, but because of the geometry: each experiment can be described only by the set of parameter expectation values which corresponds to the same angle of incidence of the particle beams. This adds substantial confidence in our earlier hypothesis that elementary reaction cross sections depend on the angle of incidence of the reactant.

7. Summary and outlook

In this paper we have analysed the synergistic interaction of CH$_3$ radicals and atomic hydrogen with the surface of an amorphous hydrocarbon film. The basis for this analysis is a rate equation model for film growth which contains a number of elementary reaction steps. Applying the concepts of Bayesian probability theory, we derived marginal posterior distributions for the corresponding elementary reaction cross sections. Distributions encoding prior knowledge are an important contribution to these posterior distributions. This formalism also allows us to properly account for experimental sources of error: we mapped onto the posterior distributions uncertainties with respect to all experimental variables, namely the flux densities of the particle beams as well as the ellipsometric measurement of film thickness. We showed that the Bayesian posterior distributions for the model parameters are much more informative than simple maximum-likelihood estimates as obtained from standard parameter fitting. The comparison of posterior distributions with prior distributions reveals the amount of information that is contained in the analysed data set. As an example, we showed that the parameters p_{elim}^H and p_{eros}^H can be

determined very precisely by measuring the transient growth rate after switching on or off the particle beams. By analysing data sets corresponding to different flux densities, we were able to detect effects which go beyond the linear approximation of our model: dangling bonds can be sterically blocked by neighbouring CH_3 groups at the surface. This provides support for a key assumption in our proposed model: we postulated that chemisorbed methyl groups need to be cross-linked by atomic hydrogen, before such sites can be activated for further growth by dangling bond creation. Our analysis also indicated that the reaction cross sections depend on the angle of particle incidence. This was corroborated by an independent measurement which was performed for two different experimental geometries. Different sets of parameter expectation values are able to explain different (even *qualitatively* different) experimental observations.

An important implication of this work is that even narrower posterior distributions and more stringent tests of the growth model could be obtained by measuring a larger set of transient growth conditions. With this paper, we pursued two purposes: first, we wanted to report on our current state of knowledge concerning synergistic film growth with CH_3 and H and the growth model. Second, we hope to stimulate the use of the Bayesian data analysis formalism as a versatile tool to handle all kinds of parameter estimation problem. It combines prior knowledge and knowledge about experimental error sources with the experimental data. As a result, the *complete* information about the model parameters is displayed. We are convinced that this approach is very useful and instructive, not only in our specific case.

References

[1] Grill A 1999 Diamond-like carbon: state of the art *Diamond Relat. Mater.* **8** 428
[2] Lettington A H 1998 Applications of diamond-like carbon thin films *Carbon* **36** 555
[3] Jacob W 1998 Surface reactions during growth and erosion of hydrocarbon films *Thin Solid Films* **326** 1
[4] Koidl P, Wild Ch, Dischler B, Wagner J and Ramsteiner M 1989 Plasma deposition, properties and structure of amorphous hydrogenated carbon films *Mater. Sci. Forum* **52/53** 41
[5] von Keudell A, Jacob W and Fukarek W 1995 Role of hydrogen ions in plasma-enhanced chemical vapor deposition of hydrocarbon films, investigated by *in situ* ellipsometry *Appl. Phys. Lett.* **66** 1322
[6] von Keudell A 2002 Formation of polymer-like hydrocarbon films from radical beams of methyl and atomic hydrogen *Thin Solid Films* **402** 1
[7] von Keudell A, Meier M and Schwarz-Selinger T 2001 Simultaneous interaction of methyl radicals and atomic hydrogen with amorphous hydrogenated carbon films, as investigated with optical *in situ* diagnostics *Appl. Phys.* A **72** 551
[8] von Keudell A, Schwarz-Selinger T and Jacob W 2001 Simultaneous interaction of methyl radicals and atomic hydrogen with amorphous hydrogenated carbon films *J. Appl. Phys.* **89** 2979
[9] von Keudell A, Schwarz-Selinger T, Meier M and Jacob W 2000 Direct identification of the synergism between methyl radicals and atomic hydrogen during growth of amorphous hydrogenated carbon films *Appl. Phys. Lett.* **76** 676
[10] von Keudell A, Schwarz-Selinger T, Meier M and Jacob W 2000 Direct identification of the synergism between methyl radicals and atomic hydrogen during growth of amorphous hydrogenated carbon films *Appl. Phys. Lett.* **77** 459 (erratum)

Institute *of* **Physics** Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

[11] Meier M and von Keudell A 2001 Hydrogen elimination as a key step for the formation of polymerlike hydrocarbon films *J. Appl. Phys.* **90** 3585

[12] Meier M 2002 Elementary mechanisms during interaction of methyl radicals and hydrogen atoms with the surface of amorphous hydrocarbon films *PhD Thesis* University Bayreuth

[13] Horn A, Schenk A, Biener J, Winter B, Lutterloh C, Wittmann M and Küppers J 1994 H atom impact induced chemical erosion reaction at C:H film surfaces *Chem. Phys. Lett.* **231** 193

[14] Küppers J 1995 The hydrogen surface chemistry of carbon as a plasma facing material *Surf. Sci. Rep.* **22** 249

[15] Vietzke E, Flaskamp K, Phillips V, Esser G, Wienhold P and Winter J 1987 Chemical erosion of amorphous hydrogenated carbon films by atomic and energetic hydrogen *J. Nucl. Mater.* **145–147** 443

[16] Schwarz-Selinger T, von Keudell A and Jacob W 1999 Plasma chemical vapor deposition of hydrocarbon films: the influence of hydrocarbon source gas on the film properties *J. Appl. Phys.* **86** 3988

[17] Sivia D S 1996 *Data Analysis—a Bayesian Tutorial* 1st edn (Oxford: Clarendon)

[18] Schwarz-Selinger 2002 private communication

[19] Chib S and Greenberg E 1995 Understanding the Metropolis–Hastings algorithm *Am. Stat.* **49** 327

[20] Cheng C C, Lucas S R, Gutleben H, Choyke W J and Yates J T Jr 1992 Interaction of atomic hydrogen with the surface methyl group on Si(100)—removal of surface carbon *Surf. Sci. Lett.* **273** L441

[21] Loh K P, Foord J S and Jackman R B 1997 Diamond growth chemistry: its observation using real time *in situ* molecular beam scattering techniques *Diamond. Relat. Mater.* **6** 219

[22] Foord J S, Loh K P and Jackman R B 1998 Surface studies of the reactivity of methyl, acetylene and atomic hydrogen at CVD diamond surfaces *Surf. Sci.* **399** 1

[23] Träskelin P, Salonen E, Nordlund K, Krasheninnikov A V, Keinonen J and Wu C H 2003 *J. Appl. Phys.* **93** 1826–31

[24] Alfonso D R and Ulloa S E 1993 Molecular-dynamics simulations of methyl-radical deposition on diamond (100) surfaces *Phys. Rev.* B **48** 12235

[25] Schwarz-Selinger T 2000 Bestimmung des Haftkoeffizienten von Methyl-Radikalen auf amorphen, wasserstoffhaltigen Kohlenstoffschichten unter Verwendung von quantifizierten Radikalteilchenstrahlquellen *PhD Thesis* University Bayreuth