



Max-Planck-Institut  
für Plasmaphysik

---

## Zerlegung von Massenspektren mit Hilfe der Bayes'schen Datenanalyse

**Thomas Schwarz-Selinger,**

H.D. Kang, R. Preuss, V. Dose

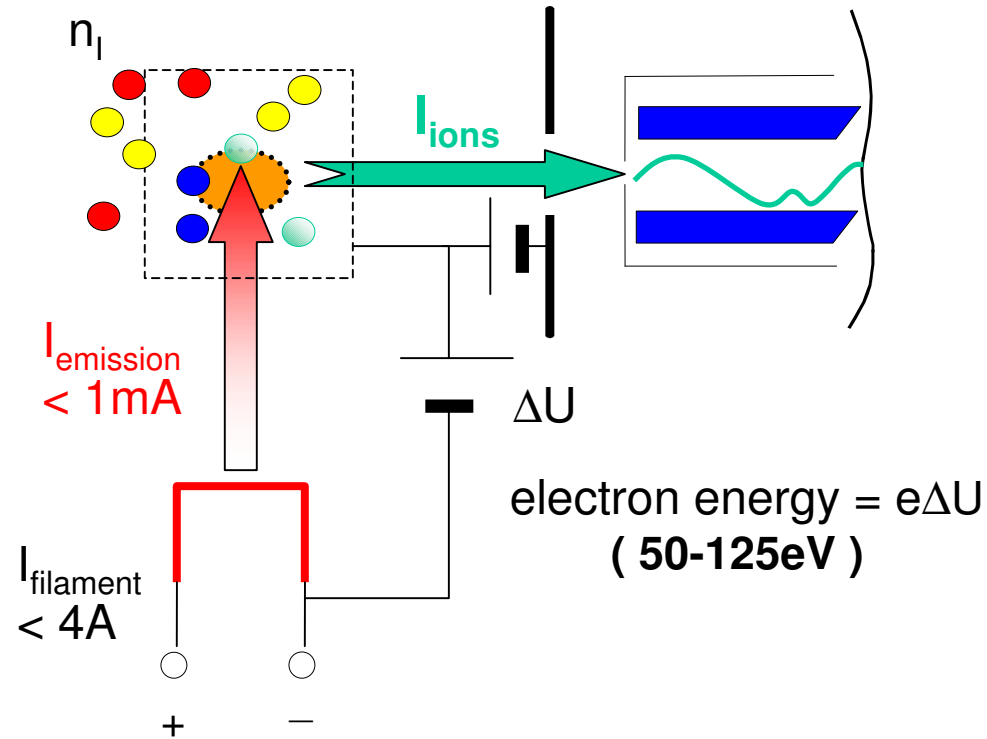
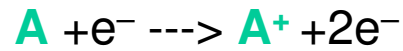
**Workshop Oberflächentechnologie mit Plasma- und Ionenstrahlprozessen  
Mühlleithen, March 2<sup>nd</sup>, 2004**



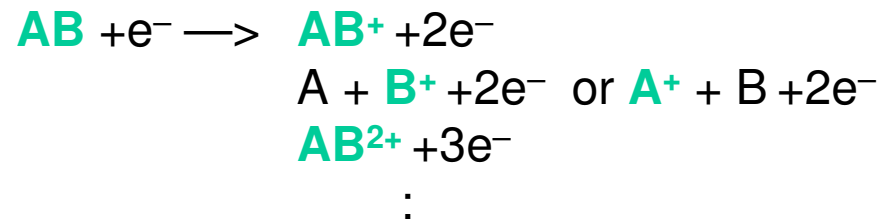
**CENTRE FOR INTERDISCIPLINARY PLASMA SCIENCE**



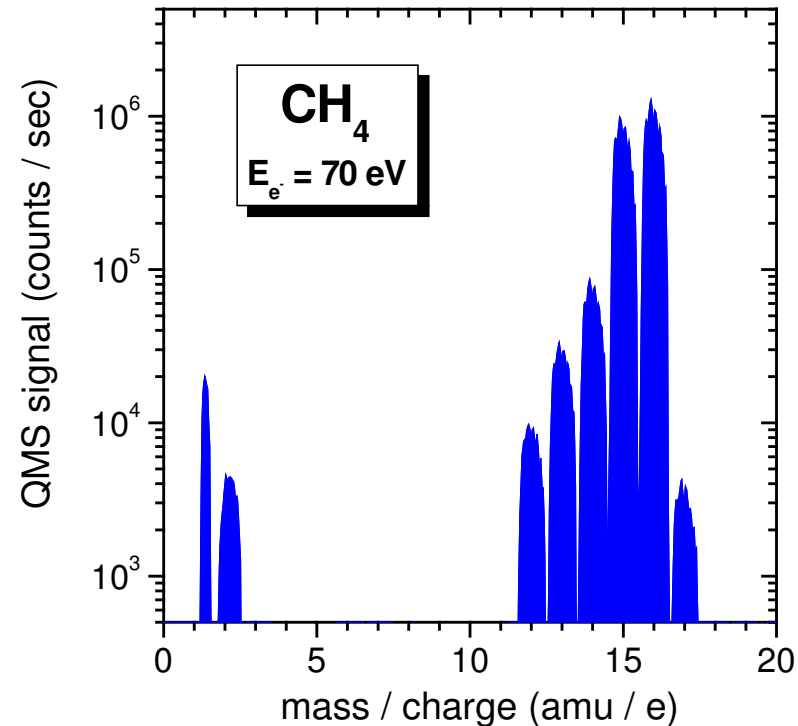
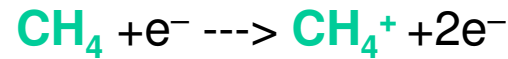
QMS ionizer:



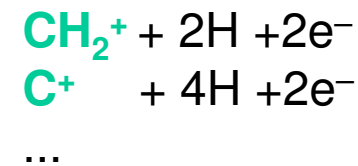
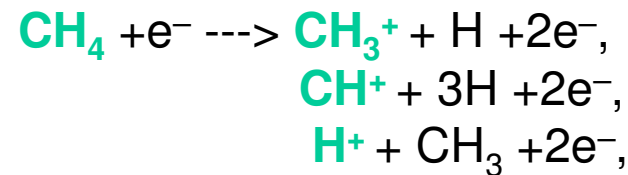
However:

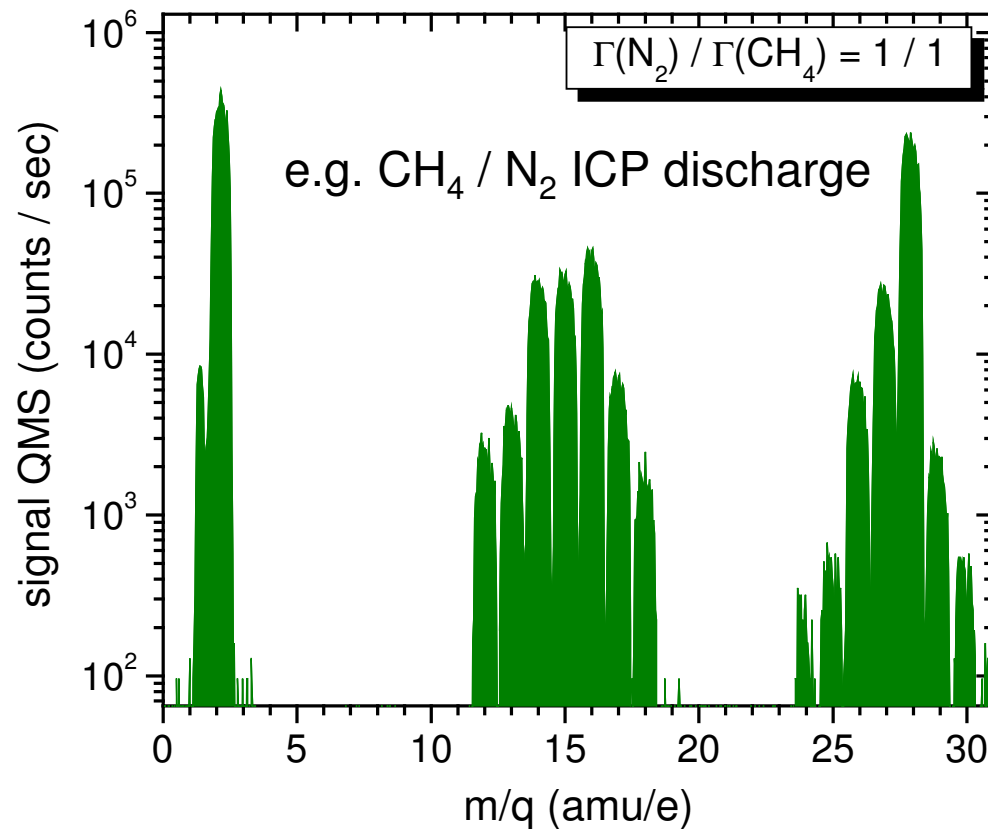


QMS ionizer:



However:





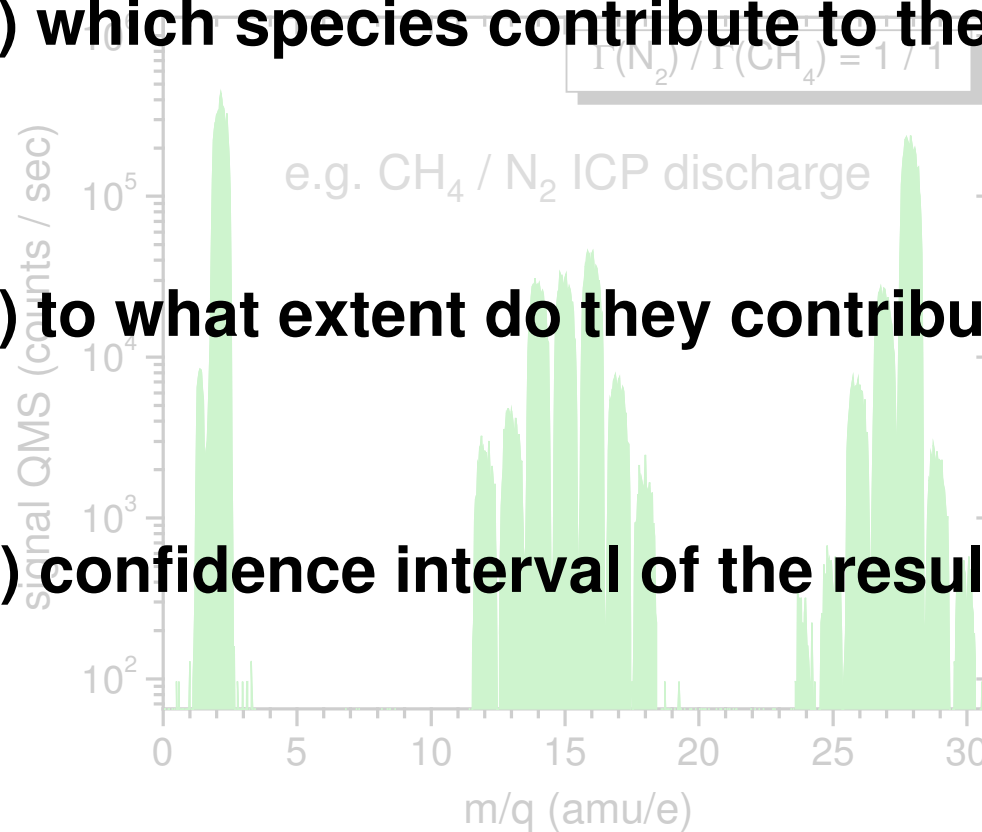
expected constituents:

$\text{H}_2$   
 $\text{CH}_4$   
 $\text{C}_2\text{H}_2$   
 $\text{C}_2\text{H}_4$   
 $\text{C}_2\text{H}_6$   
 $\text{NH}_3$   
 $\text{HCN}$   
 $\text{CH}_3$   
⋮  
⋮

1) which species contribute to the signal (identification)?

2) to what extent do they contribute (quantification)?

3) confidence interval of the result (accuracy)?



H<sub>2</sub>  
CH<sub>4</sub>  
C<sub>2</sub>H<sub>2</sub>  
C<sub>2</sub>H<sub>4</sub>  
C<sub>2</sub>H<sub>6</sub>  
NH<sub>3</sub>  
CH<sub>3</sub>  
⋮  
⋮

# Decomposition of Multicomponent Mass Spectra applying Bayesian Data Analysis

## Outline:

- Treating the Inverse Problem of Mass Spectrometry
- Basic Rules of Bayesian Probability Theory
- Nitrogen Containing Methane Plasma
- Summary

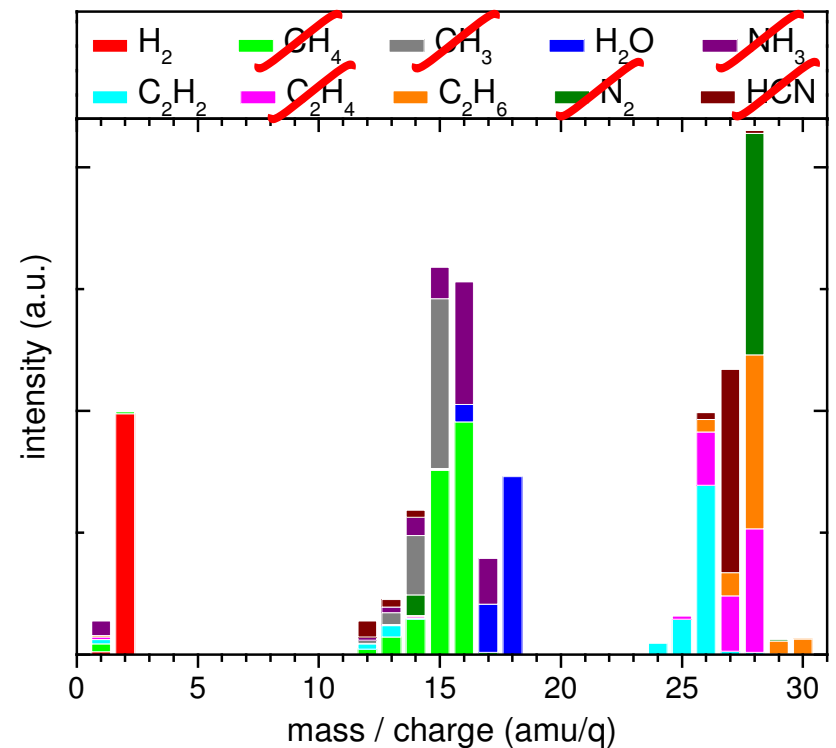
## 1) successive subtraction of constituents signal by signal

⇒ only applicable if there is one channel with no overlap

⇒ exact cracking patterns of constituents are needed

⇒ error propagation

⇒ not applicable for radicals (no cracking pattern, low concentration)



2) matrix inversion

$$\underline{d} = \underline{C} \cdot \underline{x} + \underline{\varepsilon}$$

$$\underline{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}, \quad \underline{C} = \begin{bmatrix} C_{1,1} & \cdot & \cdot & \cdot \\ \cdot & C_{2,2} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ C_{m,1} & \cdot & \cdot & C_{m,n} \end{bmatrix}, \quad \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Definitions:

- data  $\underline{d}$  with  $m$  elements, representing  $m$  mass channels
- concentrations  $\underline{x}$  with  $n$  elements, representing  $n$  species
- cracking matrix  $\underline{C}$  with  $n$  columns consisting of  $m$  elements
- measurement error  $\underline{\varepsilon}$  with  $m$  elements (counting statistics, drifts...)

$\underline{x}$  : relative concentrations depending on normalization of cracking matrix



2) matrix inversion:

$$\underline{d} = \underline{C} \cdot \underline{x} + \underline{\varepsilon}$$

e.g. H<sub>2</sub>, CH<sub>4</sub>:

$$\begin{array}{l}
 1 \text{ amu} \rightarrow \\
 2 \text{ amu} \rightarrow \\
 \vdots \\
 15 \text{ amu} \rightarrow \\
 16 \text{ amu} \rightarrow
 \end{array}
 \begin{pmatrix}
 0.015 \\
 0.691 \\
 \vdots \\
 0.12 \\
 0.15
 \end{pmatrix}
 =
 \begin{array}{cc}
 \underline{C}_{\text{H}_2} & \underline{C}_{\text{CH}_4} \\
 \downarrow & \downarrow \\
 \begin{bmatrix}
 0.015 & 0.015 \\
 0.985 & 0.004 \\
 \vdots & \vdots \\
 0 & 0.4 \\
 0 & 0.5
 \end{bmatrix}
 \cdot
 \begin{pmatrix}
 0.7 \\
 0.3
 \end{pmatrix}
 \begin{array}{l}
 \leftarrow x_{\text{H}_2} \\
 \leftarrow x_{\text{CH}_4}
 \end{array}
 \end{array}$$

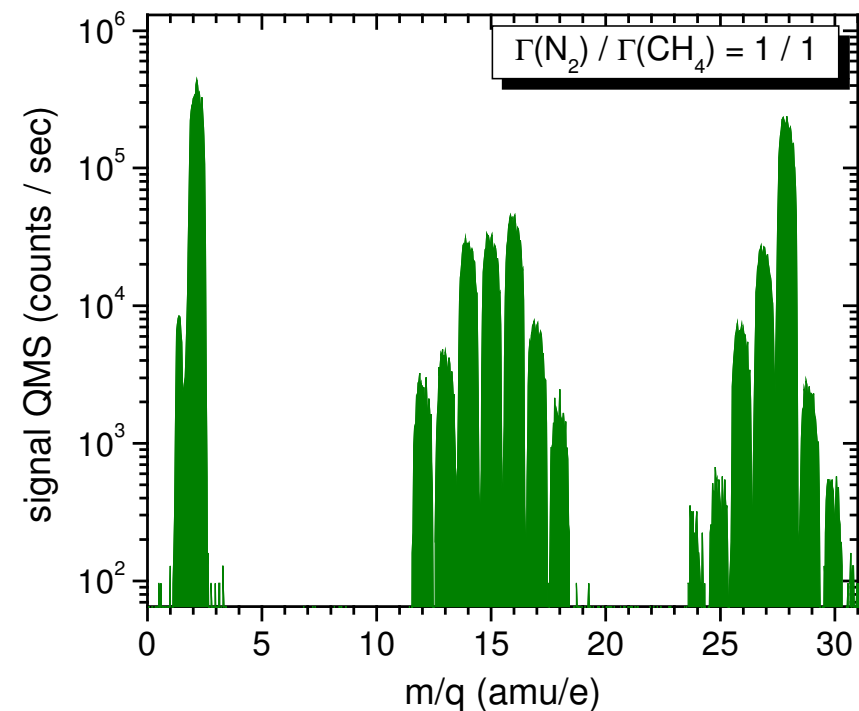
Definitions:

data  $\underline{d}$  with  $m$  elements, representing  $m$  mass channels  
 concentrations  $\underline{x}$  with  $n$  elements, representing  $n$  species  
 cracking matrix  $\underline{C}$  with  $n$  columns consisting of  $m$  elements  
 measurement error  $\underline{\varepsilon}$  with  $m$  elements (counting statistics, drifts...)

$\underline{x}$  : relative concentrations depending on normalization of cracking matrix

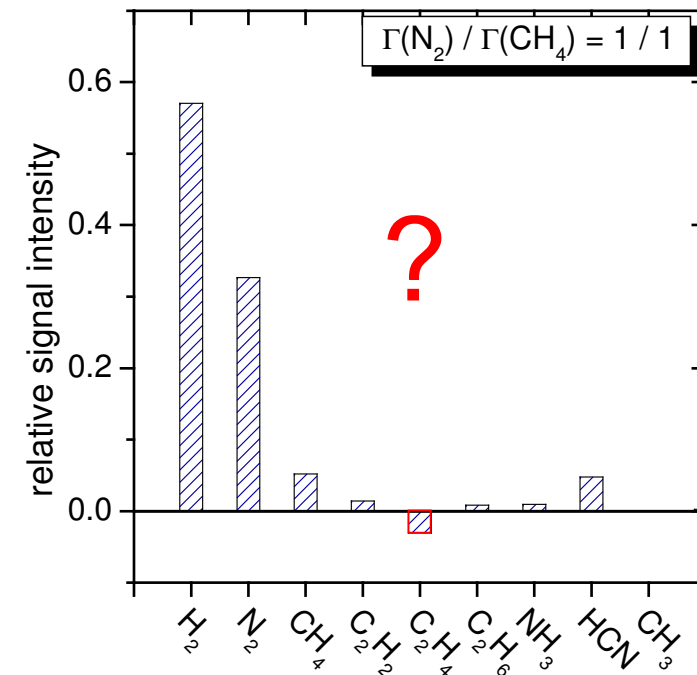
2) matrix inversion if:  $\underline{d} = \underline{C} \cdot \underline{x} + \underline{\varepsilon}$  :  $\underline{d} = \underline{C} \cdot \underline{x}$  :  $\underline{x} = \underline{C}^{-1} \cdot \underline{d}$   
 (singular value decomposition)

- ⇒ exact cracking patterns of constituents are needed
- ⇒ negative concentrations are possible
- ⇒ not applicable for radicals  
 (no cracking pattern, low concentration)



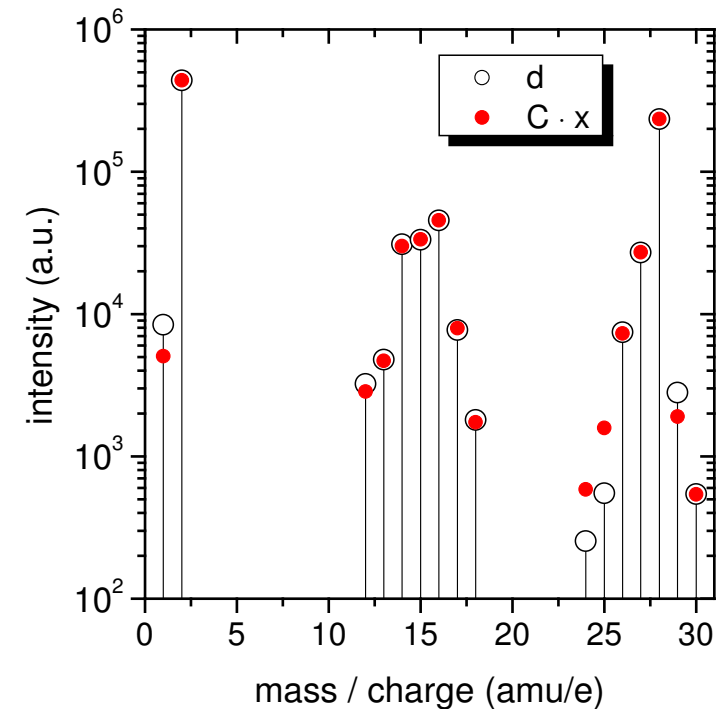
2) matrix inversion if:  $\underline{d} = \underline{C} \cdot \underline{x} + \underline{\varepsilon}$  :  $\underline{d} = \underline{C} \cdot \underline{x}$  :  $\underline{x} = \underline{C}^{-1} \cdot \underline{d}$   
 (singular value decomposition)

- ⇒ exact cracking patterns of constituents are needed
- ⇒ negative concentrations are possible
- ⇒ not applicable for radicals  
 (no cracking pattern, low concentration)



2) matrix inversion if:  $\underline{d} = \underline{C} \cdot \underline{x} + \varepsilon$  :  $\underline{d} = \underline{C} \cdot \underline{x}$  :  $\underline{x} = \underline{C}^{-1} \cdot \underline{d}$   
(singular value decomposition)

- ⇒ exact cracking patterns of constituents are needed
- ⇒ negative concentrations are possible
- ⇒ not applicable for radicals  
(no cracking pattern, low concentration)



3) least square evaluation ( $\chi^2$ -fits):  $\underline{d} = \underline{C} \cdot \underline{x} + \varepsilon$  :

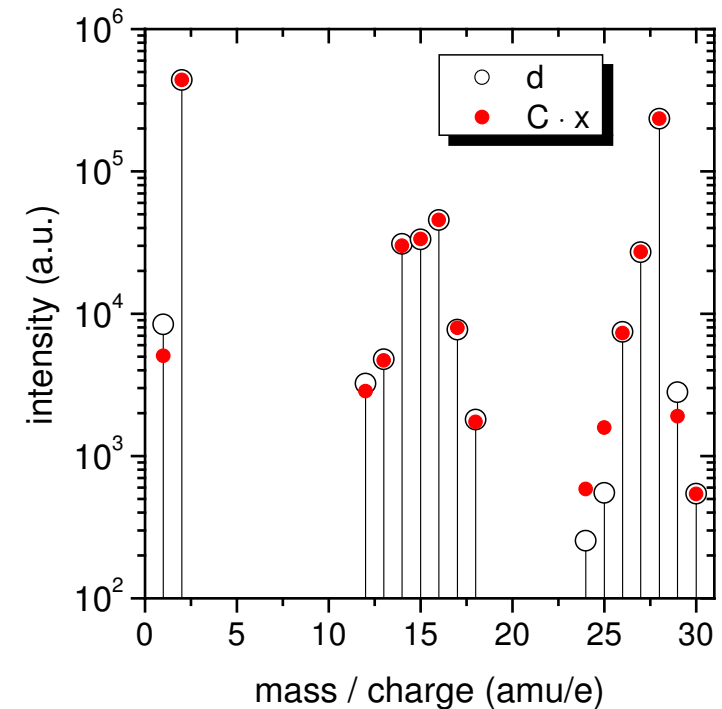
forward calculation: minimize:  $\underline{d} - \underline{C} \cdot \underline{x}$

⇒ assumes exact measurements

⇒ needs exact cracking patterns of constituents

⇒ best fit  $\neq$  best result **fitting noise ?**

⇒ not applicable for radicals  
(no cracking pattern, low concentration)



- powerful tool to solve **inverse problems**, incorporating consistently further information
- provides the **most probable result under the current state of knowledge**:

$$\langle x \rangle = \frac{\int x \cdot p(x) dx}{\int p(x) dx} \quad (\text{expectation values})$$

- powerful tool to solve **inverse problems**, incorporating consistently further information
- provides the **most probable result under the current state of knowledge**:

$$\langle x_k \rangle = \frac{\int x_k p(x_k | \vec{d}, \sigma_d, I) dx_k}{\int p(x_k | \vec{d}, \sigma_d, I) dx_k} \quad (\text{expectation values})$$

- provides the **confidence interval of the result**:

$$\sigma^2(x_i) = \langle (x_i - \langle x_i \rangle)^2 \rangle = \langle x_i^2 \rangle - \langle x_i \rangle^2 \quad (\text{standard deviation})$$

**Strategy:** evaluate  $p(x_k | \vec{d}, \sigma_d, I)$  with the help of simple rules

## Product Rule



$$P(\text{rot}, \text{Ferrari} | I) = P(\text{rot} | I, \text{Ferrari}) \cdot P(\text{Ferrari} | I) = P(\text{Ferrari} | I, \text{rot}) \cdot P(\text{rot} | I)$$



## Product Rule

$$P(A, B | I) = P(A | B, I) \cdot P(B | I) = P(B | A, I) \cdot P(A | I)$$

⇒ **Bayes' theorem**

$$P(A | B, I) = \frac{P(A | I) \cdot P(B | A, I)}{P(B, I)}$$

A → *H*ypothesis, B → *D*ata

$$P(H | D, I) = \frac{P(H | I) \cdot P(D | H, I)}{P(D | I)}$$

tells us how to update our **prior knowledge**  $P(H|I)$

about the physical **hypothesis**  $H$  in the light of **data**  $D$

which we collected from our experiment

## Sum Rule

$$P(A_1 + A_2 | I) = P(A_1 | I) + P(A_2 | I) - P(A_1, A_2 | I)$$

⇒ **marginalization**

$$P(A | I) = \int P(A, B | I) dB$$

getting rid of nuisance parameters

likelihood: (foreward calculation)

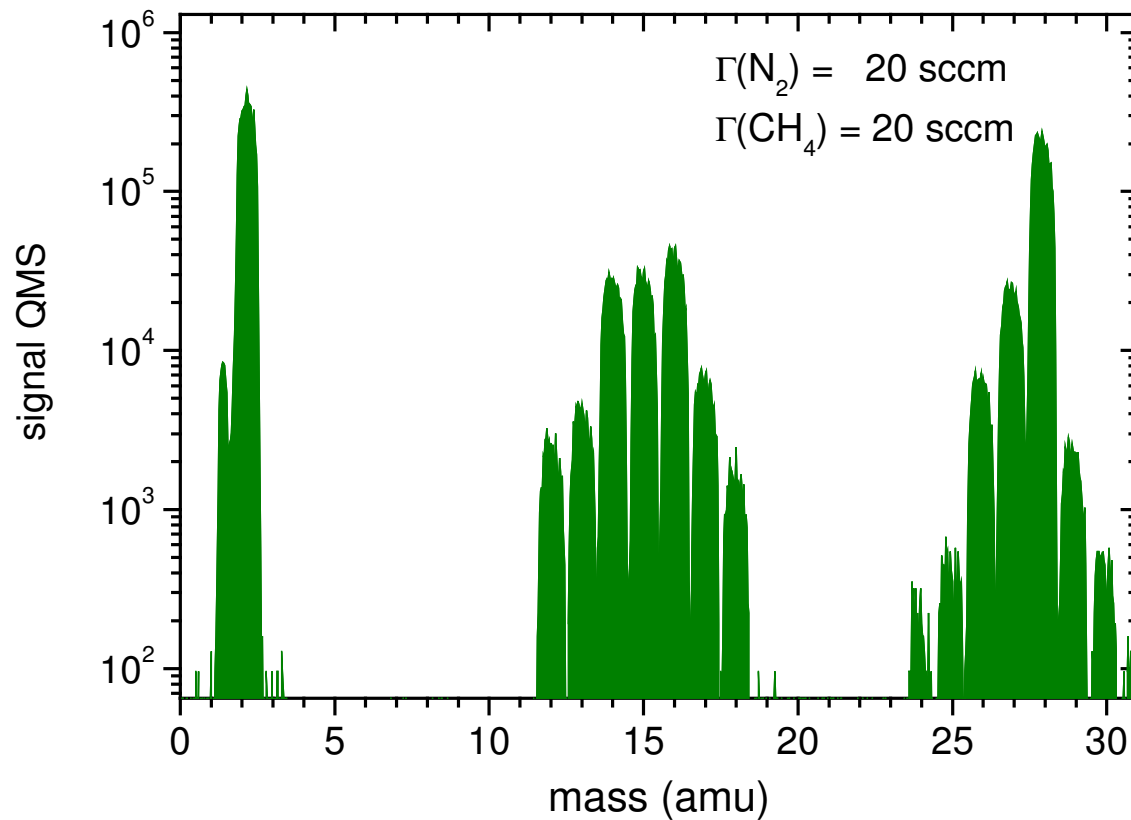
$$P(D | H, I) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(D - M)^2}{2\sigma^2}\right)$$

prior: e.g. if you know only a point estimate  $\mu$  ( from maximum Entropy)

$$p(x | \mu, I) = \frac{1}{\mu} \exp\left(-\frac{x}{\mu}\right)$$

final goal: decomposing mass spectra of reactive plasmas

e.g. CH<sub>4</sub> / N<sub>2</sub> ICP discharge



expected constituents:

- H<sub>2</sub>
  - CH<sub>4</sub>
  - C<sub>2</sub>H<sub>2</sub>
  - C<sub>2</sub>H<sub>4</sub>
  - C<sub>2</sub>H<sub>6</sub>
  - NH<sub>3</sub>
  - HCN
  - CH<sub>3</sub>
  - ⋮
- no calibration available**

final goal: decomposing mass spectra of reactive plasmas

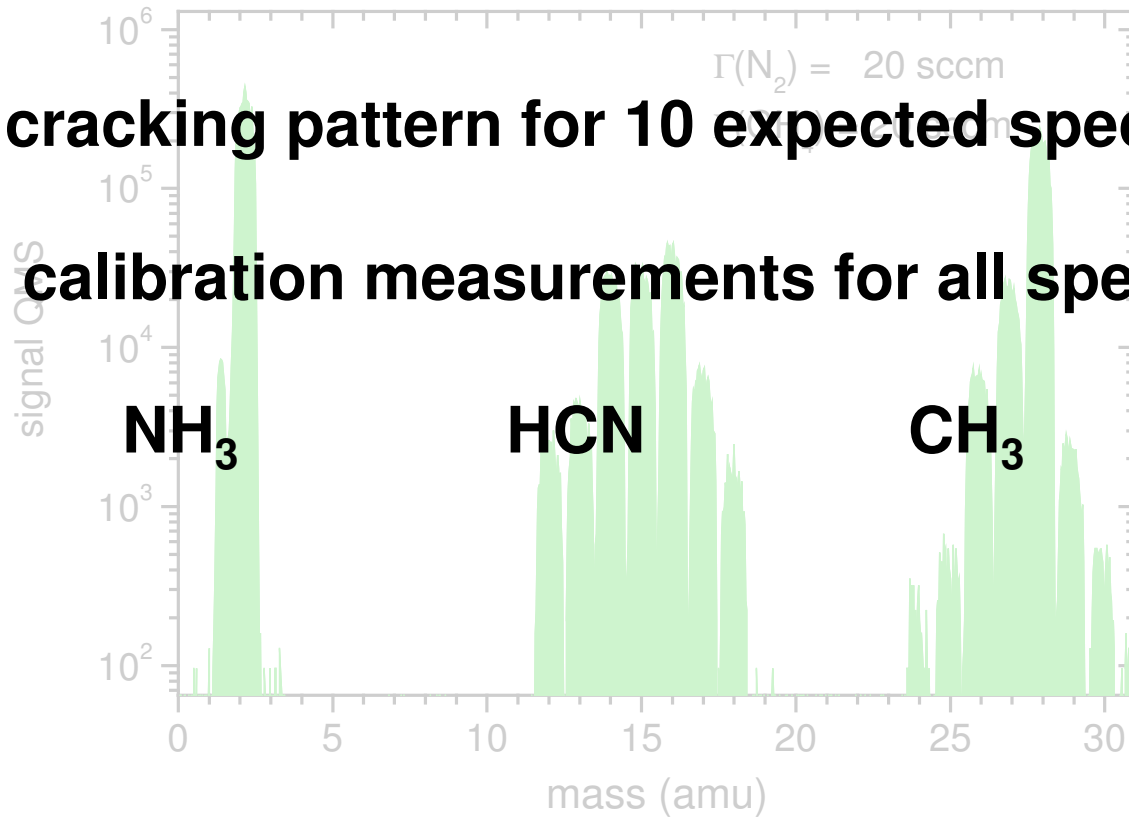
**input:**

e.g. CH<sub>4</sub> / N<sub>2</sub> ICP discharge

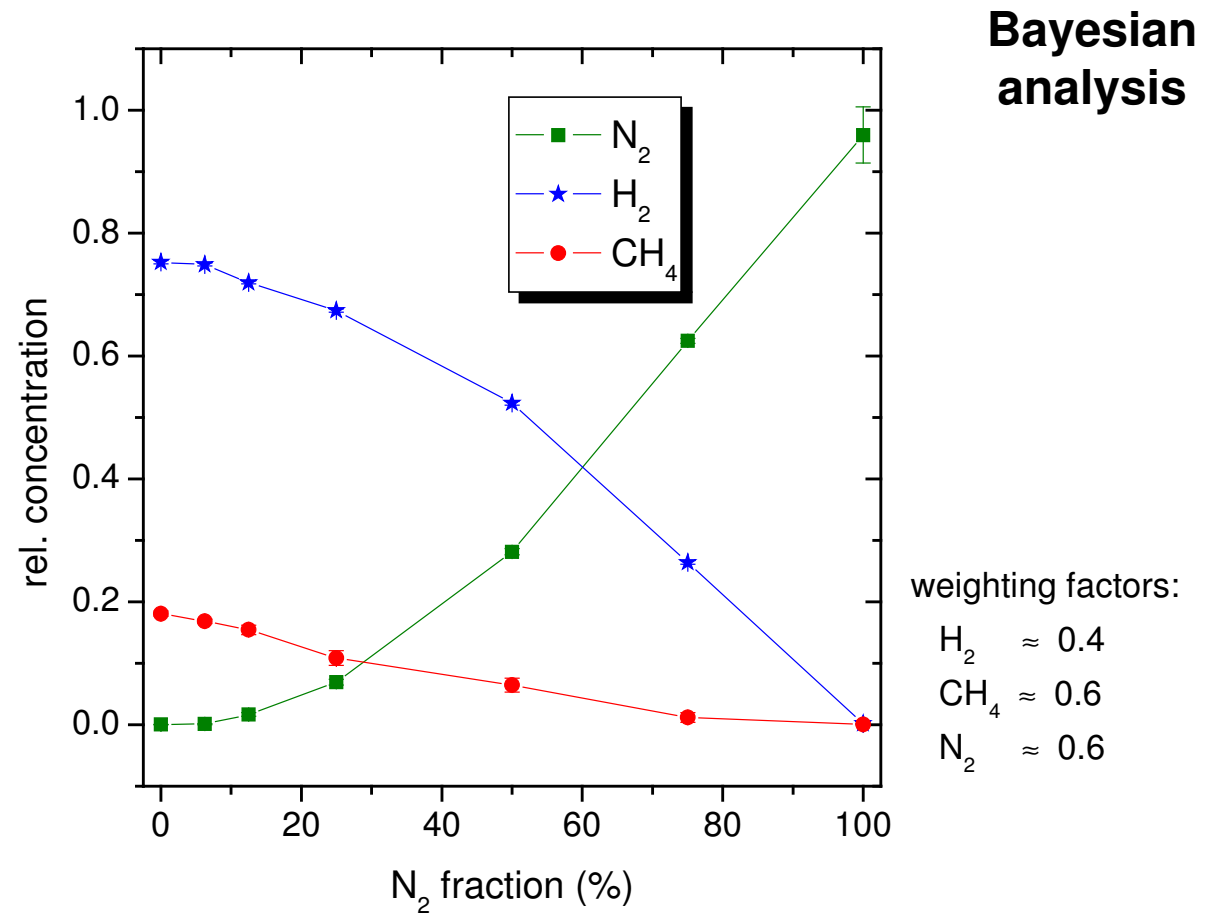
**1) intensities of 16 mass channels for 7 mixtures**

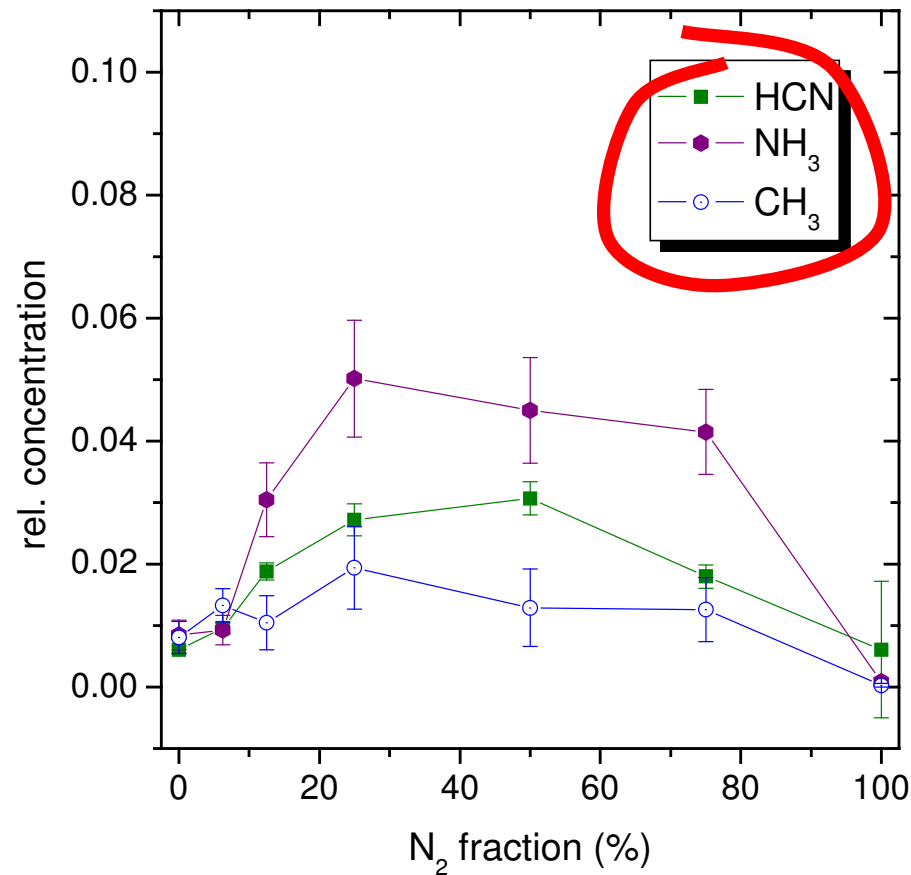
**2) cracking pattern for 10 expected species (tables)**

**3) calibration measurements for all species except:**



- H<sub>2</sub>
  - CH<sub>4</sub>
  - C<sub>2</sub>H<sub>2</sub>
  - C<sub>2</sub>H<sub>4</sub>
  - C<sub>2</sub>H<sub>6</sub>
  - NH<sub>3</sub>
  - HCN
  - CH<sub>3</sub>
  - ⋮
- no calibration available**





**Bayesian analysis**

**no calibration available!**

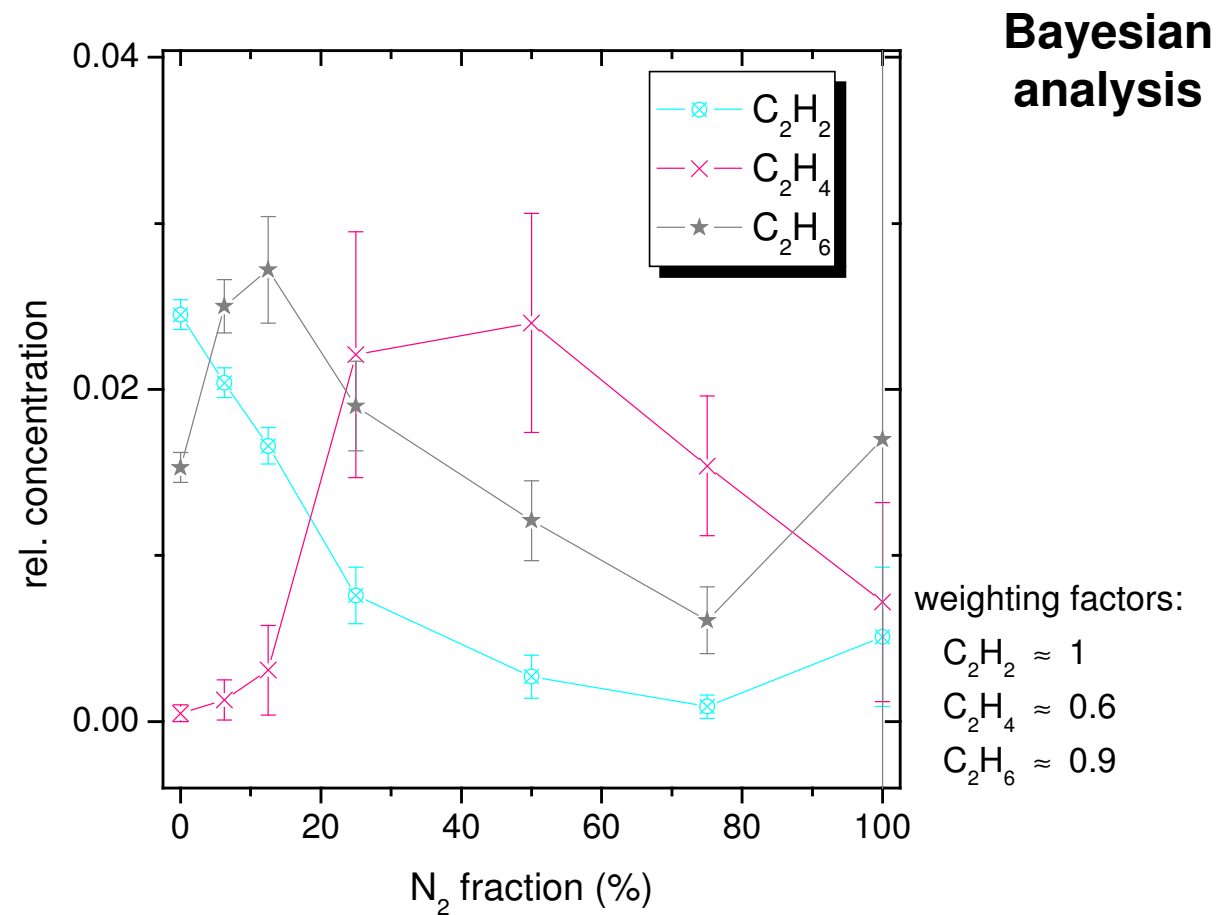
weighting factors:

HCN ≈ 0.3

NH<sub>3</sub> ≈ 0.8

CH<sub>3</sub> ≈ 0.6

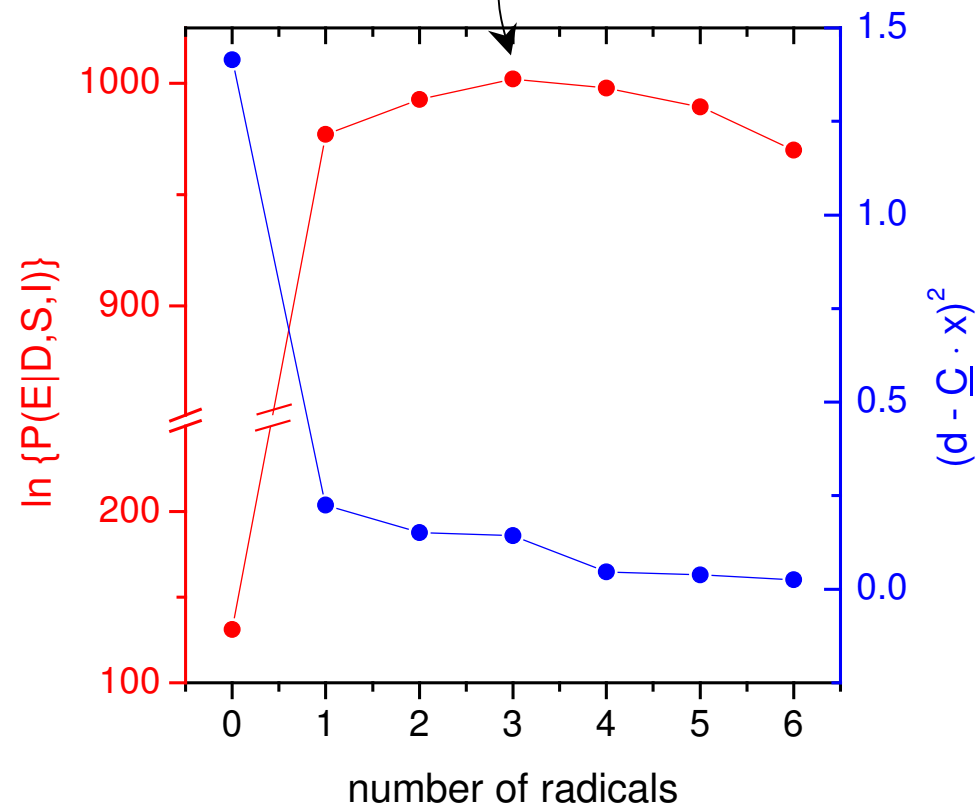




output:

model comparison with „Occams Razor“

CH<sub>3</sub>, C<sub>2</sub>H<sub>5</sub>, H



- Bayesian data analysis allows to decompose multicomponent mass spectra incorporating further information (e.g. calibration measurements)
- parameter estimation delivers the expectation value as well as the confidence interval for the relative signal intensities as well as the cracking patterns
- species with unknown cracking pattern can be included
- model comparison allows to determine the species present in the mixture
- the analysis can help to optimise the experiment

- **the method has to fail, when**
  - **we under- / overestimate our errors**
  
  - **we make wrong assumptions like**  
**inappropriate calibration measurements**  
**excluding species**
  
- **the method cannot do wonders**
  
- **its simply “common sense reduced to calculation”**



***"a mass spectrometrists is someone,  
who figures out what something is,  
by smashing it with a hammer  
and looking at the pieces"***