Analysis of kinetic reaction mechanisms

Version of 14 October, 2023

Chemical
Kinetics
Laboratory|ELTE

Tamás Turányi
Chemical Kinetics Laboratory
Institute of Chemistry, ELTE Eötvös Loránd University, Budapest, Hungary

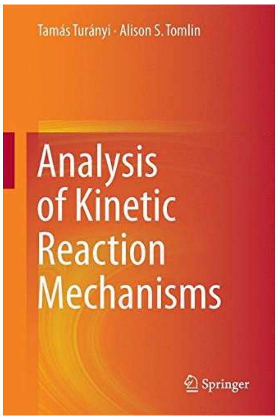
Analysis of kinetic reaction mechanisms – the book

Tamás Turányi and Alison S. Tomlin:

Analysis of Kinetic Reaction Mechanisms

Springer, 2014

(with 1025 references)



Tamás Turányi · Alison S. Tomlin

Analysis of Kinetic Reaction Mechanisms

Springer

Web page:
http://garfield.chem.elte.hu/Turanyi/KineticReactionMechanisms.html

- table of contents
- download the chapters
- references
- typos found

Topic 1: Reaction kinetics basics

stoichiometric equation, reaction mechanism,
parameterization of temperature dependence,

stoichiometric matrix, calculation of **J** and **F** matrices, general characteristics of the system of kinetic differential equations,

trajectories, conserved properties.

reaction kinetic simplifying principles:
- rate determining step,
- quasi steady-state approximation (QSSA),
- fast pre-equilibrium approximation,
- pool component approximation;

applications of reaction kinetics models

Reaction kinetics basics

Characterization of chemical changes with a stoichiometric (overall) equation:

- properly indicates the ratio of reactants and products
- usually there is no such a real chemical process

$2\ H_2 + O_2 = 2\ H_2O$

$0 = -2\ H_2 - 1\ O_2 + 2\ H_2O$

$$0 = \sum_j \nu_j A_j$$

$\nu_1 = -2$ $\quad A_1 =$ „H_2”

$\nu_2 = -1$ $\quad A_2 =$ „O_2”

$\nu_3 = +2$ $\quad A_3 =$ „H_2O”

ν_j stoichiometric coefficient (negative for reactants, positive for products)

Features:
- the order of the species is arbitrary
- the stoichiometric coefficients can be multiplied with the same real number

$H_2 + ½\ O_2 = H_2O$ is also a good overall equation

Reaction rate

production rate of a species: $$\frac{\mathrm{d}\,Y_j}{\mathrm{d}\,t}$$

reaction rate: $$r = \frac{1}{\nu_j}\frac{\mathrm{d}\,Y_j}{\mathrm{d}\,t}$$

Y_j is the molar concentration of species A_j *e.g.* [mole dm^{-3}]

in a small domain of concentrations always applicable: $$r = k\prod_j Y_j^{\alpha_j},$$

k	rate coefficient
α_j	reaction order with respect species *j*
$\alpha = \sum_j \alpha_j$	overall reaction order

Complex reaction mechanisms

Almost always there are many simultaneous reaction steps:

$$\sum_j \nu_{ij}^L \mathrm{A}_j = \sum_j \nu_{ij}^R \mathrm{A}_j$$

A reaction step
can be an elementary reaction (physically occurs this way) or
can be a non-elementary reaction lumped from elementary reactions

ν_{ij}^L	matrix of left hand side stoichiometric coefficients elementary: sum is not more than 2; zero or positive integer non-elementary: zero or positive integer
ν_{ij}^R	matrix of right hand side stoichiometric coefficients elementary: sum is not more than 2; zero or positive integer non-elementary: any real number (can be zero, negative, fraction)

$\Delta\nu_{ij} = \nu_{ij}^R - \nu_{ij}^L$ calculation of the (previous) stoichiometric matrix

Kinetic system of differential equations

law of mass action (Guldberg and Waage, 1865):

$$r_i = k_i \prod_j Y_j^{\nu_{ij}^L}$$

k_i	rate coefficient of reaction step *i*
r_i	rate of reaction step *i*

Definition of the kinetic system of differential equations:

$$\frac{\mathrm{d}Y_j}{\mathrm{d}t} = \sum_i \Delta\nu_{ij}\, r_i; \quad j = 1,\ 2,\ \ldots, n$$

The kinetic system of differential equations in matrix-vector form:

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = \boldsymbol{\nu}\mathbf{r}$$

Matrices to be mentioned frequently

Initial value problem in reaction kinetics:

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = \mathbf{f}(\mathbf{Y}, \mathbf{k}), \qquad \mathbf{Y}(t_0) = \mathbf{Y}_0$$

Jacobian: $\mathbf{J} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}$

The Jacobian usually changes with changing concentrations

matrix F: $\mathbf{F} = \left\{ \frac{\partial f_i}{\partial k_j} \right\}$

also depends on the concentrations

Kinetic system of differential equations: an example

The Oregonator model of the Belousov-Zhabotinskii oscillating reaction:

1.	$X + Y \rightarrow 2\,P$	k_1	$r_1 = k_1 xy$
2.	$Y + A \rightarrow X + P$	k_2	$r_2 = k_2 ya$
3.	$2\,X \rightarrow P + A$	k_3	$r_3 = k_3 x^2$
4.	$X + A \rightarrow 2\,X + 2\,Z$	k_4	$r_4 = k_4 xa$
5.	$X + Z \rightarrow 0.5\,X + A$	k_5	$r_5 = k_5 xz$
6.	$Z + M \rightarrow Y - Z$	k_6	$r_6 = k_6 zm$

$X = HBrO_2$
$Y = Br^-$
$Z = Ce^{4+}$
$A = BrO_3^-$
$P = HOBr$
M = malonic acid

The detailed 80-step reaction mechanism could be reduced to this 6 reaction step.

Note, that negative and fractional stoichiometric coefficients are present on the right hand side!

Kinetic system of differential equations: an example 2

$X = HBrO_2$	variable of a diff. equation
$Y = Br^-$	variable of a diff. equation
$Z = Ce^{4+}$	variable of a diff. equation
$A = BrO_3^-$	constant concentration
$P = HOBr$	product only
M = malonic acid	constant concentration

1.	$X + Y \rightarrow 2\,P$
2.	$Y + A \rightarrow X + P$
3.	$2\,X \rightarrow P + A$
4.	$X + A \rightarrow 2\,X + 2\,Z$
5.	$X + Z \rightarrow 0{,}5\,X + A$
6.	$Z + M \rightarrow Y - Z$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -1r_1 + 1r_2 - 2r_3 + 1r_4 - 0{,}5\,r_5 \quad \Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{d}t} = -k_1xy + k_2ya - 2k_3x^2 + k_4xa - 0{,}5\,k_5xy$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -1r_1 - 1r_2 + 1r_6 \quad \Rightarrow \quad \frac{\mathrm{d}y}{\mathrm{d}t} = -k_1xy - k_2ya + k_6zm$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = +2\,r_4 - 1r_5 - 2\,r_6 \quad \Rightarrow \quad \frac{\mathrm{d}z}{\mathrm{d}t} = 2k_4xa - k_5xz - 2k_6zm$$

$$\frac{dx}{dt} = -k_1 xy + k_2 ya - 2k_3 x^2 + k_4 xa - 0.5k_5 xz$$

$$\frac{dy}{dt} = -k_1 xy - k_2 ya + k_6 zm$$

$$\frac{dz}{dt} = 2k_4 xa - k_5 xz - 2k_6 zm$$

calculation of the Jacobian

$$\mathbf{J} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}$$

$\frac{\partial \frac{dx}{dt}}{\partial x} = -k_1 y - 4k_3 x + k_4 a - 0{,}5k_5 z$	$\frac{\partial \frac{dx}{dt}}{\partial y} = -k_1 x + k_2 a$	$\frac{\partial \frac{dx}{dt}}{\partial z} = -0.5k_5 x$
$\frac{\partial \frac{dy}{dt}}{\partial x} = -k_1 y$	$\frac{\partial \frac{dy}{dt}}{\partial y} = -k_1 x - k_2 a$	$\frac{\partial \frac{dy}{dt}}{\partial z} = +k_6 m$
$\frac{\partial \frac{dz}{dt}}{\partial x} = 2k_4 a - k_5 z$	$\frac{\partial \frac{dz}{dt}}{\partial y} = 0$	$\frac{\partial \frac{dz}{dt}}{\partial z} = -k_5 x - 2k_6 m$

$$\frac{dx}{dt} = -k_1 xy + k_2 ya - 2k_3 x^2 + k_4 xa - 0.5k_5 xz$$

$$\frac{dy}{dt} = -k_1 xy - k_2 ya + k_6 zm$$

$$\frac{dz}{dt} = 2k_4 xa - k_5 xz - 2k_6 zm$$

calculation of matrix F

$$\mathbf{F} = \left\{ \frac{\partial f_i}{\partial k_j} \right\}$$

$\frac{\partial \frac{dx}{dt}}{\partial k_1} = -xy$	$\frac{\partial \frac{dx}{dt}}{\partial k_2} = ya$	$\frac{\partial \frac{dx}{dt}}{\partial k_3} = -2x^2$	$\frac{\partial \frac{dx}{dt}}{\partial k_4} = xa$	$\frac{\partial \frac{dx}{dt}}{\partial k_5} = -0.5xz$	$\frac{\partial \frac{dx}{dt}}{\partial k_6} = 0$
$\frac{\partial \frac{dy}{dt}}{\partial k_1} = -xy$	$\frac{\partial \frac{dy}{dt}}{\partial k_2} = -ya$	$\frac{\partial \frac{dy}{dt}}{\partial k_3} = 0$	$\frac{\partial \frac{dy}{dt}}{\partial k_4} = 0$	$\frac{\partial \frac{dy}{dt}}{\partial k_5} = 0$	$\frac{\partial \frac{dy}{dt}}{\partial k_6} = zm$
$\frac{\partial \frac{dz}{dt}}{\partial k_1} = 0$	$\frac{\partial \frac{dz}{dt}}{\partial k_2} = 0$	$\frac{\partial \frac{dz}{dt}}{\partial k_3} = 0$	$\frac{\partial \frac{dz}{dt}}{\partial k_4} = 2xa$	$\frac{\partial \frac{dz}{dt}}{\partial k_5} = -xz$	$\frac{\partial \frac{dz}{dt}}{\partial k_6} = -2zm$

Properties of kinetic differential equations

- The system of differential equations contains only first order derivatives (d*c* / d*t*), which are usually nonlinear functions of the concentrations.
  ⇒ **first order nonlinear system of differential equations**
- In general, several other concentrations influence the production rate of each species.
  ⇒ **coupled differential equations**
- The reaction rates differ several orders of magnitude
  ⇒ **stiff differential equations**
- Simulation results of laboratory experiments do not depend on the wall clock time, **BUT** the results of atmospheric chemical models depend on the actual pressure, temperature and solar raditation ⇒ depend on the physical time.
  ⇒ **autonomous OR non-autonomous differential equations**
- Some laboratory reactions can be (approximately) spatially homogeneous, but outside the laboratories most chemical reactions are spatially inhomogeneous. In most cases the transport of species and heat have to be taken into account.
  ⇒ **partial system of differential equations, with chemical source term**

Conserved properties

Isolated system:
The total internal energy is constant

Constant volume closed system:
the sum of the concentrations is constant,
if each the change of the number of moles in each reaction step is zero.
e.g. for reaction $H_2+Cl_2 = 2\ HCl$

Closed system, elementary reactions only:
the number of moles of the elements is constant.

The moles of moieties (*e.g.* benzene ring) can remain constant

Example for conserved properties in a C_2H_4,CH_4 , C_6H_6 mixture:
C-atom → $2\ [C_2H_4] + 1\ [CH_4] + 6\ [C_6H_6]$ = constant
H-atom → $4\ [C_2H_4] + 4\ [CH_4] + 6\ [C_6H_6]$ = constant

Some linear combinations of the concentrations are constant.

N conserved property:
⇒ the rank of the stoichiometric matrix is lower by *N*
⇒ the system can be simulated **exactly** with (n-N) variables

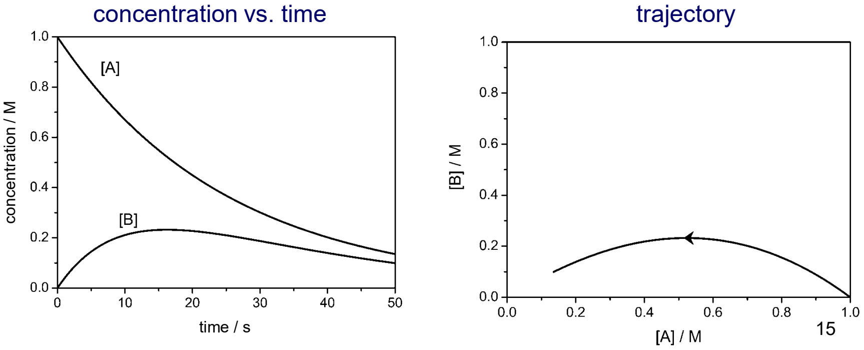
Trajectory

Actual state of the system: a point in the phase space

History of events: trace of this point in the phase space (a solid line)

In chemical kinetics this phase space can be the space of concentrations. This figure does not show time, but can be more intersting than the usual concentration vs. time curces.

Example: reaction $A \rightarrow B \rightarrow$

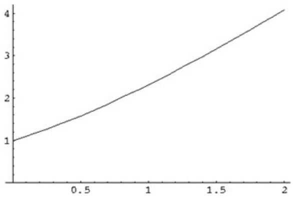


Trajectories 2

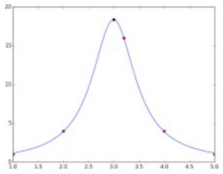
closed systems:
The concentrations are continously changing from the initial state to an equilibrium point

Possible variations:

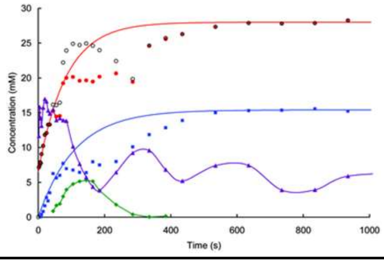
monotonic



maximum / minimum



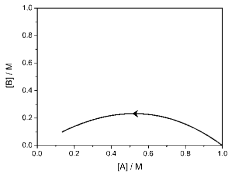
oligo oscillation



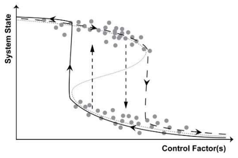
Trajectories 3

open systems:
The concentrations are continously changing from the initial state to a ...

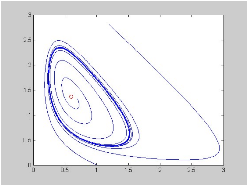
single stationary point



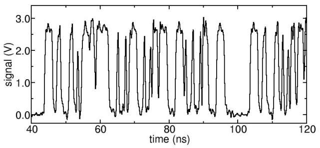
bistable bihaviour

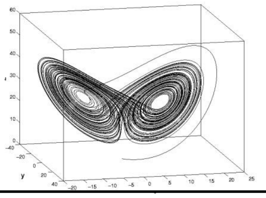


closed loop (limit cycle)



strange attractor





Temperature dependence of rate coefficient *k*

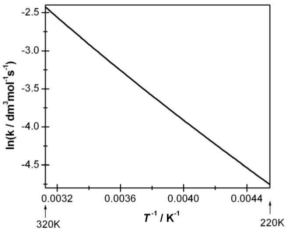
Described by the Arrhenius equation:

$$k = A\exp\left(-\frac{E_a}{RT}\right) \qquad \ln k = \ln A - \frac{E_a}{RT}$$

A preexponential factor
Ea activation energy

If the rate coefficient *k* is measured at several *T* temperatures and
ln *k* is plotted as a function of 1/*T*
the data fit to a line, if the (original) Arrhenius equation is valid
slope is $m = -E_a/R$ ⇒ determination of E_a

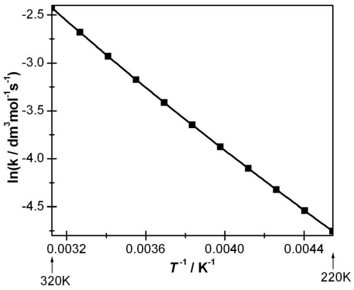
Arrhenius plot:



Example: reaction $CH_4 + OH \rightarrow CH_3 + H_2O$

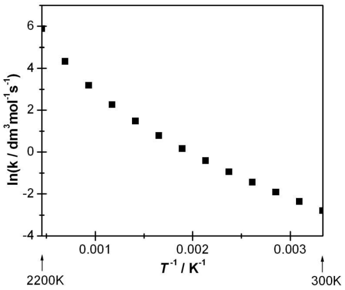
- the most important methane consuming reaction step in the troposphere
- one of the most important steps at methane combustion

Arrhenius plot between 220 K (– 53 °C) and 320 K (+ 47 °C)



the Arrhenius equation is usually very accurate in a small (few times 10 K) temperature range. (solution phase and atmospheric chemistry)

Arrhenius plot between 300 K (27 °C) and 2200 K (≈1930 °C)

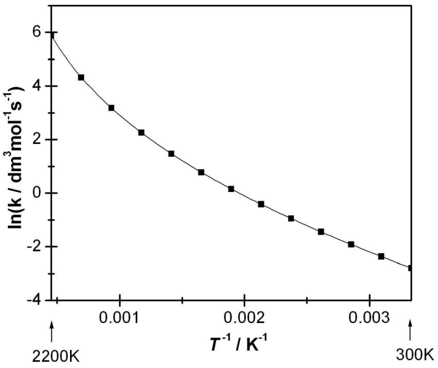


the original Arrhenius equation is usually not applicable in a wide temperature range (combustion and pyrolytic systems)

Temperature dependence of the rate coefficient *2*

$$k = BT^n \mathrm{e}^{-\frac{C}{RT}}$$

extended Arrhenius equation



Important!
If $n \neq 0$. then $A \neq B$ and $E_a \neq C$

General definition of activation energy:

$$E_a = -R \frac{\partial \ln k}{\partial (1/T)}$$

Reaction kinetics simplifying principles

Reaction kinetics simplifying principles: can be used for the simplification of a reaction mechanism (kinetic system of differential equations) in such a way that the obtained reaction mechanism (or system of differential equations) provides an almost identical (say within 1%) solution.

Reaction kinetic simplifying principles:

- rate determining step
- quasi steady state approximation (QSSA)
- fast pre-equilibrium approximation
- pool component approximation

Rate determining step

The case of consecutive first order reactions

Reaction step having the smallest rate coefficient is the rate determining.

The rate of production of the end product is equal to the rate coefficient of the rate determining step times the concentration of its reactant.

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D \xrightarrow{k_4} E \xrightarrow{k_5} F$$

$$k_2 << k_1 , k_3 , k_4 , k_5 \Rightarrow \qquad d\,[F]/d\,t = k_2\,[B]$$

In the case of any reaction mechanism: small increase of the rate coefficient of the rate determining step results in a large increase of the rate of production of the end product.

In general, the reaction step having the smallest rate coefficient is not the rate determining step!

Quasi steady-state approximation (QSSA)

Highly reactive and low-concentration species in a mechanism are selected.
These species (usually very reactive radicals) are called the QSSA species.
The left-hand-sides of the kinetic differential equations are zeroed, converting these differential equations to algebraic equations. The obtained set of algebraic equations describe the dependence of the concentrations of the QSSA species on the concentrations of the non-QSSA species. Solving together the sets of differential and algebraic equations provides a solution that is in good accordance with the solution of the original kinetic system of differential equations.

In most cases the system of algebraic equations can be solved analytically, the concentrations of the QSSA species can be given explicitly and this solution can be inserted to the remaining system of differential equations for the non-QSSA species.

⇒ smaller system of differential equations

The QSSA species are usually highly reactive and therefore low concentration species (*e.g.* radicals).

Application of the QSSA

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

if $k_1 << k_2$ B "QSSA species" A and C "non-QSSA species"

$$d[B]/d\,t = k_1\,[A] - k_2\,[B]$$

$$0 = k_1\,[A] - k_2\,[B]$$

$$[B] = k_1/k_2\,[A]$$

[B] can be calculated from [A]

Pre-equilibrium approximation

If the reactants of a fast equilibrium reaction is consumed by much slower reactions, then
the concentrations of the species participating in the fast equilibrium can be calculated from the equilibrium equations only
and the effects of other reactions should not be taken into account.

$$A \underset{k_2}{\overset{k_1}{\rightleftarrows}} B \xrightarrow{k_3} C \qquad\qquad K = k_1/k_2$$

if $k_3 << k_2$ and $d[B]/dt \approx 0$ (state of equilibrium)

$\Rightarrow k_1 [A] = k_2 [B]$

$\Rightarrow [B] = k_1/k_2 [A] = K [A]$

$\Rightarrow d[C]/dt = k_3 [B] = k_3 K [A]$

Pool component approximation

If the concentration of one of the reactants is much higher than those of the others, then this concentration will not change significantly during the reaction.

This way a second order reaction can be converted to an equivalent first order reaction by merging the rate coefficient and the concentration of the pool component.

$$\frac{d[C]}{dt} = k[A][B] = k'[A]$$

where $k' = k[B]$ is constant. ⇒ „pseudo first order reaction"

Example: „inversion of sucrose"

Hydrolysis of sucrose in an acidic solution. The products formed are optically active and their optical rotation can be determined by using a polarimeter. The decomposition of sucrose can be described by a first order decay reaction.

Applications of reaction kinetics models

- **modelling atmospheric chemical processes**
  - forecast of air pollution (weather forecast is needed!)
  - determination of emission limits

- **modelling of ignition and combustion**
  - modelling power stations, furnaces, engines
  - improving efficiency
  - elaboration of methods for the decrease of pollutant emiassion

- **process engineering; modelling of chemical engineering processes**
  - considering efficiency and the aspects of environment protection

- **systems biology: modelling biochemical processes within living organisms**
  - metabolic networks (*e.g.* medical drug decomposition in the body)
  - molecular signal transfer
  - modelling the cell cycle

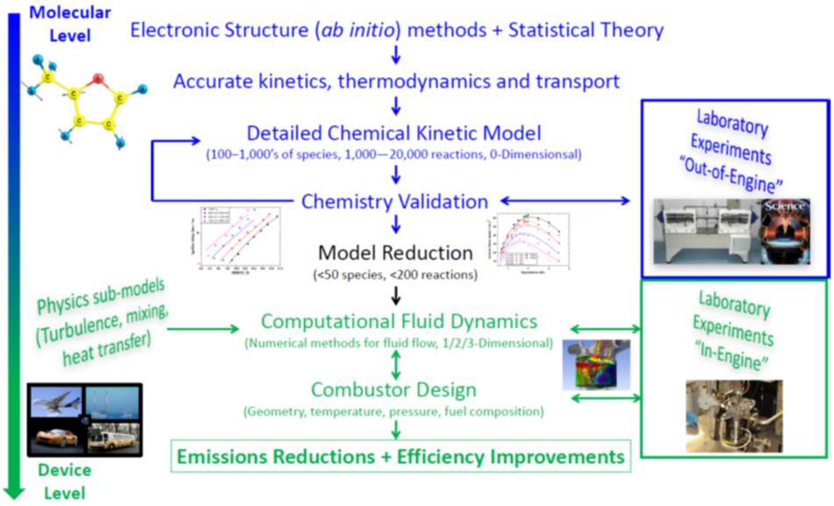
- **non-chemical models using reaction kinetic formalism**
  - predator-prey models
  - description of ecological systems

Topic 2: Construction of detailed reaction mechanisms and the reaction pathways

data sources,
traditional construction of reaction mechanisms,
automatic mechanism generation;

pathway analysis:
species conversion pathways, element following pathways,
pathways leading to the production of a given species

Hierarchy of combustion models



Molecular Level

Electronic Structure (*ab initio*) methods + Statistical Theory

↓

Accurate kinetics, thermodynamics and transport

↓

Detailed Chemical Kinetic Model

(100–1,000's of species, 1,000—20,000 reactions, 0-Dimensionsal)

↓

Chemistry Validation

↓

Model Reduction

(<50 species, <200 reactions)

↓

Physics sub-models (Turbulence, mixing, heat transfer)

Computational Fluid Dynamics

(Numerical methods for fluid flow, 1/2/3-Dimensional)

↕

Combustor Design

(Geometry, temperature, pressure, fuel composition)

↓

Emissions Reductions + Efficiency Improvements

Laboratory Experiments "Out-of-Engine"

Laboratory Experiments "In-Engine"

Device Level

Henry J. Curran: Developing detailed chemical kinetic mechanisms for fuel combustion
Proc.Combust.Inst., **37**, 57-81 (2019)

Source of chemical kinetic data

measured and calculated chemical kinetic data → journal publications

data compilation → books. data bases.
e.g. **NIST database**
www.nist.gov

data evaluation → review articles

reevaluation and comparison of several articles

evaluated/recommended data

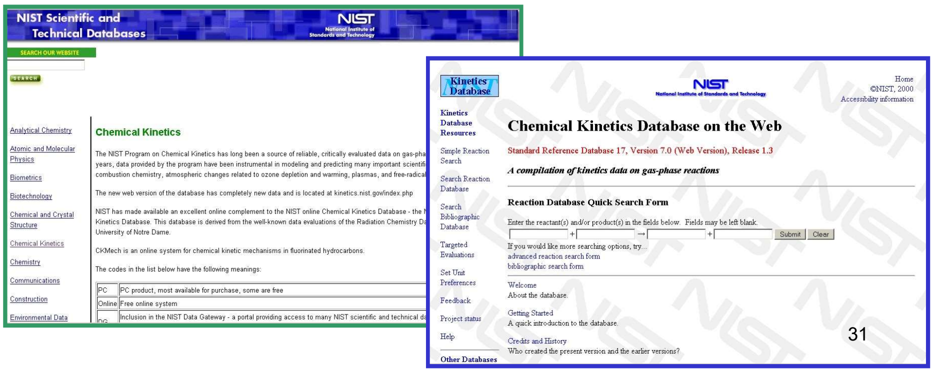
NIST Chemical Kinetics Database

www.nist.gov

⇒ Databases ⇒ Chemical Kinetics

⇒ NIST Chemical Kinetics Online (http://kinetics.nist.gov/kinetics/)

NIST Chemical Kinetics Database	11.700 gas phase reactions
	38.000 data entry
	12.000 referenced articles



Chemical Kinetics

The NIST Program on Chemical Kinetics has long been a source of reliable, critically evaluated data on gas-pha
years, data provided by the program have been instrumental in modeling and predicting many important scientifi
combustion chemistry, atmospheric changes related to ozone depletion and warming, plasmas, and free-radical

The new web version of the database has completely new data and is located at kinetics.nist.gov/index.php

NIST has made available an excellent online complement to the NIST online Chemical Kinetics Database - the N
Kinetics Database. This database is derived from the well-known data evaluations of the Radiation Chemistry Da
University of Notre Dame.

CKMech is an online system for chemical kinetic mechanisms in fluorinated hydrocarbons.

The codes in the list below have the following meanings:

PC	PC product, most available for purchase, some are free
Online	Free online system
DG	Inclusion in the NIST Data Gateway - a portal providing access to many NIST scientific and technical da

Chemical Kinetics Database on the Web

Standard Reference Database 17, Version 7.0 (Web Version), Release 1.3

A compilation of kinetics data on gas-phase reactions

Reaction Database Quick Search Form

Enter the reactant(s) and/or product(s) in the fields below. Fields may be left blank.

If you would like more searching options, try...
advanced reaction search form
bibliographic search form

Welcome
About the database.

Getting Started
A quick introduction to the database.

Credits and History
Who created the present version and the earlier versions?

NIST Chemical Kinetics Database 2

Author(s): Gierczak. T.; Talukdar. R.K.; Herndon. S.C.; Vaghjiani. G.L. Ravishankara. A.R.
Title: Rate coefficients for the reactions of hydroxyl radicals with methane and deuterated methanes
Journal: J. Phys. Chem. A:
Volume: 101
Page(s): 3125 - 3134
Year: 1997
Reference type: Journal article
Squib: 1997GIE/TAL3125-3134

Reaction: $CH_4 + \cdot OH \rightarrow \cdot CH_3 + H_2O$
Reaction order: 2
Temperature: 196 - 420 K
Pressure: 0.13 Bar
Rate expression: 1.76×10^{-13} (cm^3/molecule s) $(T/298\ K)^{2.82}\ e^{-1.96\ (\pm0.02\ \text{kcal/mole})/RT}$
Bath gas: He
Data type: Absolute value measured directly
Excitation technique: Flash photolysis (laser or conventional)
Analytical technique: Laser induced fluorescence

NIST Chemistry WebBook

Another important Web source:
Webbook (http://webbook.nist.gov/)

- thermochemical data for over 7000 compounds
- reaction thermochemistry data for over 8000 reactions.
- IR spectra for over 16.000 compounds.
- mass spectra for over 33.000 compounds.
- UV/Vis spectra for over 1600 compounds.
- gas chromatography data for over 27.000 compounds.
- electronic and vibrational spectra for over 5000 compounds.
- spectroscopic data for over 600 compounds.
- ion energetics data for over 16.000 compounds.
- thermophysical property data for 74 fluids.

Traditional way for the development of detailed reaction mechanisms

1. List of elementary reactions is generated

2. Determination of the rate parameters one-by-one:

   Based on direct measurements
   Using chemical kinetic databases
   Calculation/estimation of rate parameters

3. Comparison of the simulation results
   with the results of indirect measurements.
   Indirect measurements: time-to-ignition, flame velocity, concentration–time or concentration–distance profiles.

   No good agreement in most cases

4. Identification of the most important reactions by sensitivity analysis at the experimental conditions.

5. Tuning the rate parameters of the most important reactions, till the model reproduces the experimental data.

Different authors tune different parameters
⇒ different mechanisms

Computer generation of mechanisms

General ideas:

1 starting from some reactants (for example: fuel molecule + O_2)

2 elementary reactions are generated according to reaction types
(*e.g.* H-abstraction, additions, fission of radicals)

3 irrealistic elementary reactions are not considered
(based on various filtering principles)

4 getting rate and thermodynamic parameters from databases

5 approximate automatic calculation of the missing parameters

6 new list of intermediates **GOTO 2**

Stopping of mechanism generation at some complexity/number of reactions.

Investigation of the computer generated mechanism by

- **comparison with human generated mechanisms**
- **testing against experimental data.**

Codes for mechanism generation – some examples

EXGAS (Nancy)
includes mechanism generator, kinetic data base and estimation of thermochemical parameters
wider classes of fuels: heavy alkanes, oxygenated species, biomass fuels

F. Battin-Leclerc, P. A. Glaude, V. Warth, R. Fournet, G. Scacchi, G. M. Côme: Computer tools for modelling the chemical phenomena related to combustion. *Chem. Eng. Sci.* **55**, 2883-2893 (2000)

MAMOX (Milano)
- automatic generation of mechanisms
- considering isomers with similar kinetic behaviour as a single lumped
- species lumping parallel reaction pathways for similar isomers
- fitting lumped reaction rates to predictions from the full scheme

Ranzi, E., Faravelli, T., Gaffuri, P., Sogaro, A.: Low-temperature combustion: Automatic generation of primary oxidation reactions and lumping procedures. *Combust. Flame* **102**, 179-192 (1995)

KUCRS (University of Tokyo)
automatic generation of mechanisms

A. Miyoshi, KUCRS - Detailed kinetic mechanism generator for versatile fuel components and mixtures, Proceedings of the 8th International Conference on Modeling and Diagnostics for Advanced Engine Systems, COMODIA **2012**, 116–121 (2012)

Codes for mechanism generation – some examples 2

Genesys (Ghent University)
includes mechanism generator, kinetic data base and estimation of thermochemical parameters

N.M. Vandewiele, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, Genesys: Kinetic model construction using chemo-informatics, *Chemical Engineering Journal*. **207–208** , 526–538 (2012)

RGM - Reaction Mechanism Generator (MIT) http://rmg.sourceforge.net/
„RMG is an automatic chemical reaction mechanism generator that constructs kinetic models composed of elementary chemical reaction steps using a general understanding of how molecules react."

W. H. Green, P. I. Barton, B. Bhattacharjee, D. M. Matheu, D. A. Schwer, J. Song, R. Sumathi, H. H. Carstensen, A. M. Dean, J. M. Grenda: Computer construction of detailed chemical kinetic models for gas-phase reactors. *Ind. Eng. Chem. Res.* **40**, 5362–5370 (2001)

Pathway analysis

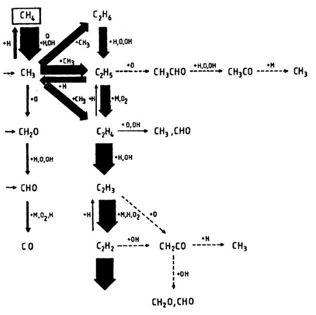
species conversion pathways
pathways leading to the production of a given species

Reaction pathways

Conversion of one species to another

Reaction fluxes:
The width of the arrows is proportional
to the interconversion rate

Several textbooks contain pathway figures,
But usually the exact calculation of
The width of the arrows is not revealed.



pathways in a rich methane-air flame

Warnatz J., Maas U., Dibble R. W.
Combustion. Physical and chemical fundamentals, modeling and simulation, experiments, pollutant formation
Springer, New York, 1996

Reaction fluxes

S.R. Turns:
An introduction to combustion.
Concepts and applications.
second edition,
Boston, McGraw-Hill, 2000.

„each arrow indicates a reaction step”

„the width of the arrows is proportional to the consumption rate of the reactant”

Not a good idea, since consecutive arrows having different width may belong to identical fluxes.

⇒ **Flux of a conserved property Has to be plotted!!!**

⇒ Fluxes of elements (Revel *et al.*, 1994)

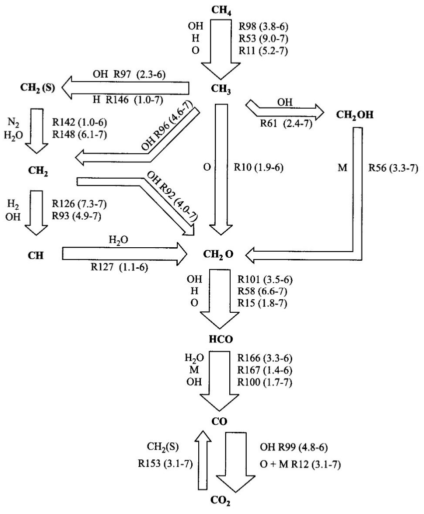


Figure 5.4 High-temperature reaction pathway diagram for combustion of methane in a well-stirred reactor at $T = 2200$ K and $P = 1$ atm for a 0.1-s residence time. Reaction numbers refer to Table 5.3, while reaction rates are shown in parentheses. For example, 2.6-7 implies $2.6 \cdot 10^{-7}$ (gmol/cm^3-s).

Fluxes of elements

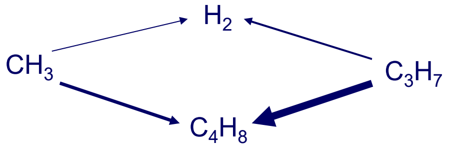
$CH_3 + C_3H_7 \Rightarrow C_4H_8 + H_2$ reaction rate= r_1

number of H-atoms: 3 7 8 2

number of H-atoms on the left hand side: 10

flux of H-atoms from one species to another:

CH_3	➔	C_3H_7	0	
CH_3	➔	C_4H_8	$3/10*8*r_1$ =	$2.4*r_1$
CH_3	➔	H_2	$3/10*2*r_1$ =	$0.6*r_1$
C_3H_7	➔	CH_3	0	
C_3H_7	➔	C_4H_8	$7/10*8*r_1$ =	$5.6*r_1$
C_3H_7	➔	H_2	$7/10*2*r_1$ =	$1.4*r_1$



$$A_{ijk} = \frac{n_{\mathrm{A},j}\, n_{\mathrm{A},k}\, r_i}{N_{\mathrm{A},i}}$$

J. Revel, J. C. Boettner, M. Cathonnet, J. S. Bachman: Derivation of a global chemical kinetic mechanism for methane ignition and combustion. *J. Chim. Phys.* **91**, 365-382 (1994)

Calculation of element fluxes using KINALC

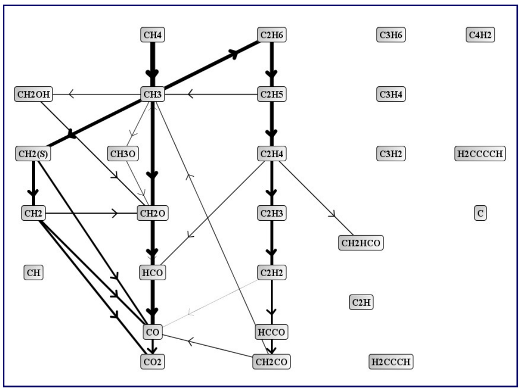
```
c         ATOMFLOW Fluxes of elements from species to species are investigated
c         The name(s) of elements are listed after the keyword.
c         Usage: ATOMFLOW <element1> <element2> ...

ATOMFLOW C H
```

```
=== ATOMFLOW ===============================

         Fluxes of elements from species to species

Net fluxes of element H                        absolute                         rel.

     1  H2                => H2O                6.843E-02  mole/(cm3 sec)       1.0000
     2  H2                => H                  4.584E-02  mole/(cm3 sec)        .6699
     3  OH                => H2O                4.034E-02  mole/(cm3 sec)        .5895
     4  H                 => OH                 3.360E-02  mole/(cm3 sec)        .4910
     5  H                 => HO2                2.370E-02  mole/(cm3 sec)        .3463
     6  OH                => H                  2.302E-02  mole/(cm3 sec)        .3364
     7  HO2               => OH                 1.797E-02  mole/(cm3 sec)        .2626
     8  H2                => OH                 1.162E-02  mole/(cm3 sec)        .1699
     9  H                 => H2                 6.346E-03  mole/(cm3 sec)        .0927
    10  H                 => H2O                4.084E-03  mole/(cm3 sec)        .0597
    11  HO2               => H2                 3.334E-03  mole/(cm3 sec)        .0487
    12  HO2               => H2O                2.689E-03  mole/(cm3 sec)        .0393
    13  OH                => H2                 1.049E-03  mole/(cm3 sec)        .0153
    14  H2O               => OH                 7.891E-04  mole/(cm3 sec)        .0115
    15  OH                => H2O2               7.617E-04  mole/(cm3 sec)        .0111
    16  H2O               => H2                 7.010E-04  mole/(cm3 sec)        .0102
    17  H2O2              => H2O                5.367E-04  mole/(cm3 sec)        .0078
    18  H2O2              => OH                 2.131E-04  mole/(cm3 sec)        .0031
```

KINALC → FluxViewer

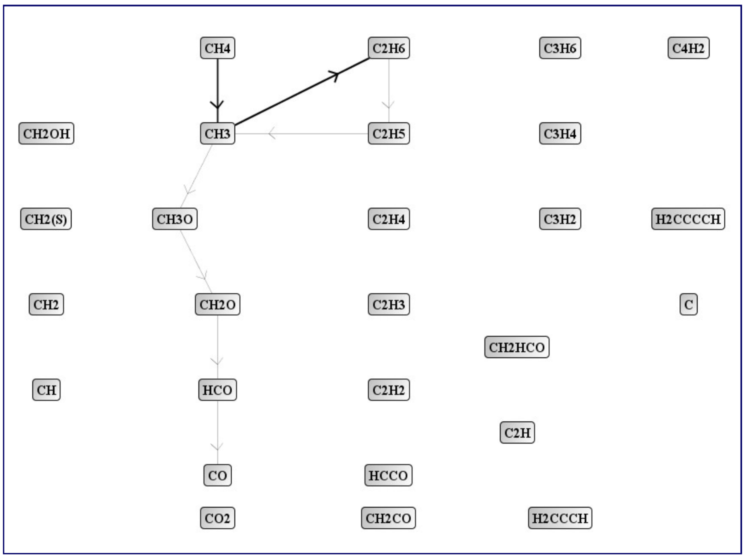
FluxViewer: JAVA code for the visualization of the element fluxes

- the labels of the species can be moved („drag-and-drop”)
- the width of the arrows is proportional to the log of the element fluxes
- the width of the arrows can be hanged
- creation of drawings or movie films



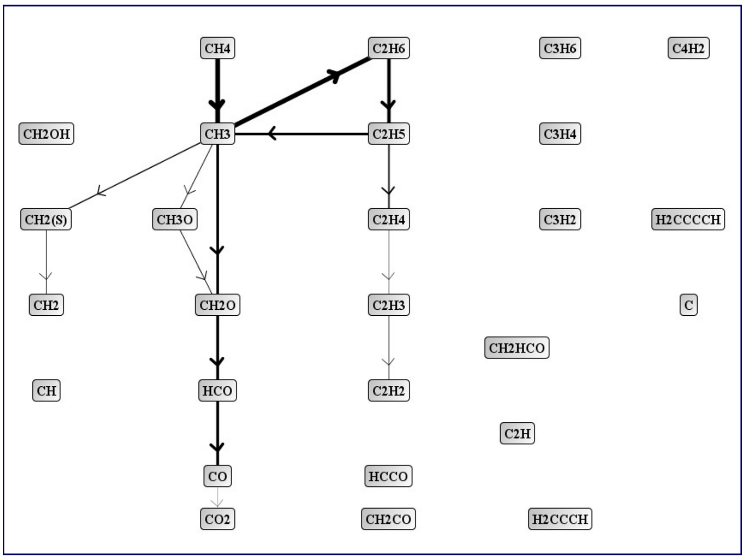
I. Gy. Zsély, I. Virág, T. Turányi:
Investigation of a methane oxidation mechanism via the visualization of element fluxes
Paper IX.4 in: Proceedings of the 4th Mediterranean Combustion Symposium,
Lisbon, Portugal, 5-10 October, 2005, Eds: F. Beretta, N. Selçuk, M.S. Mansour

C-atom fluxes, φ =1.0, methane–air flame



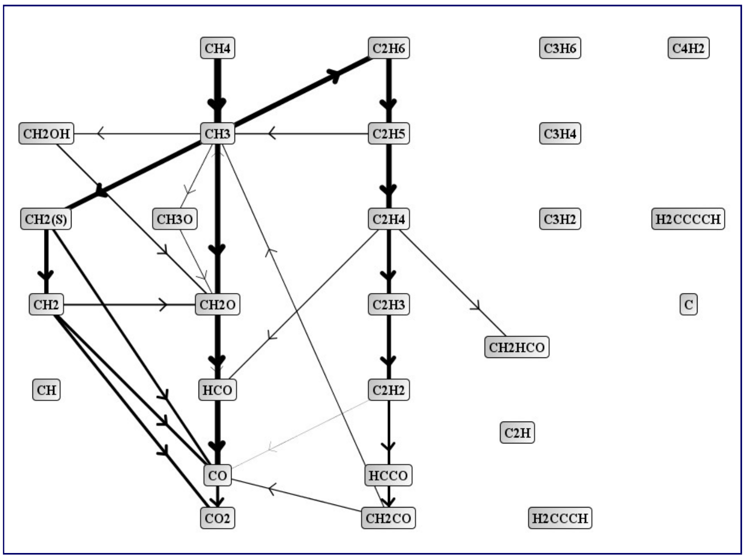
T= 815 K

C-atom fluxes, φ =1.0, methane–air flame



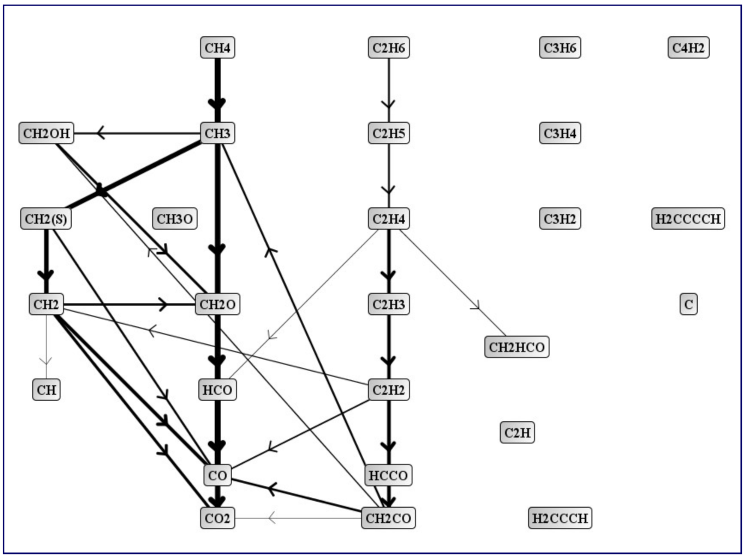
T=1155 K

C-atom fluxes, φ =1.0, methane–air flame



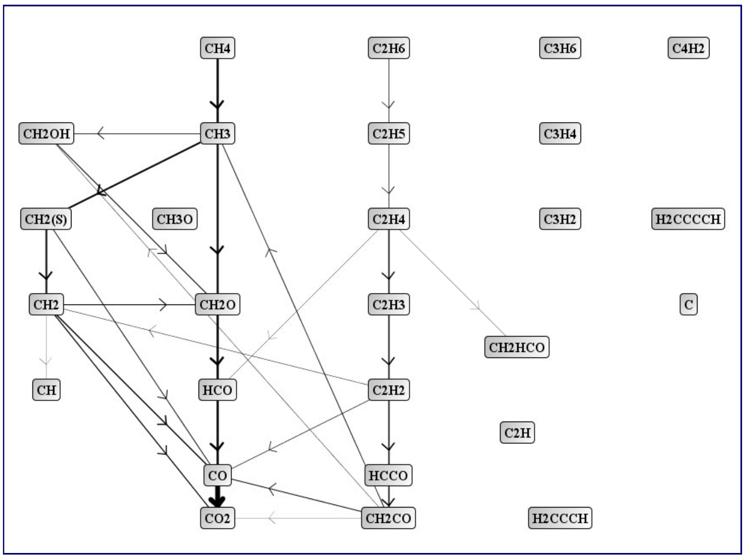
T=1500 K

C-atom fluxes, φ =1.0, methane–air flame



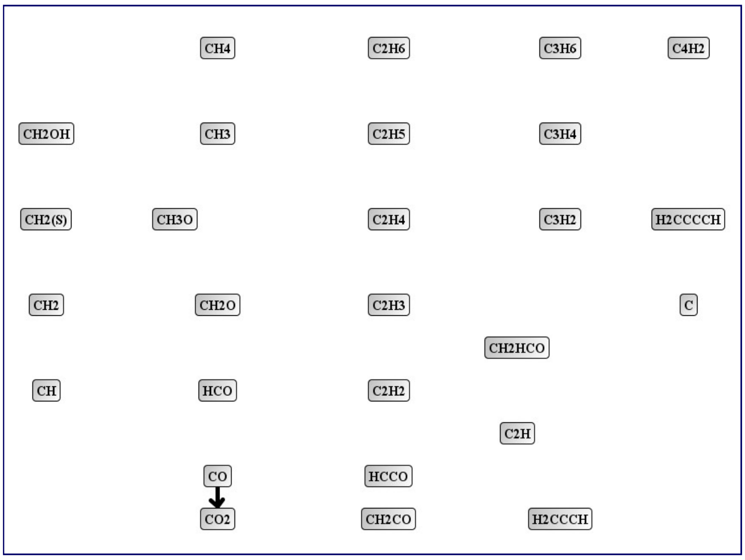
T=1805 K

C-atom fluxes, φ =1.0, methane–air flame



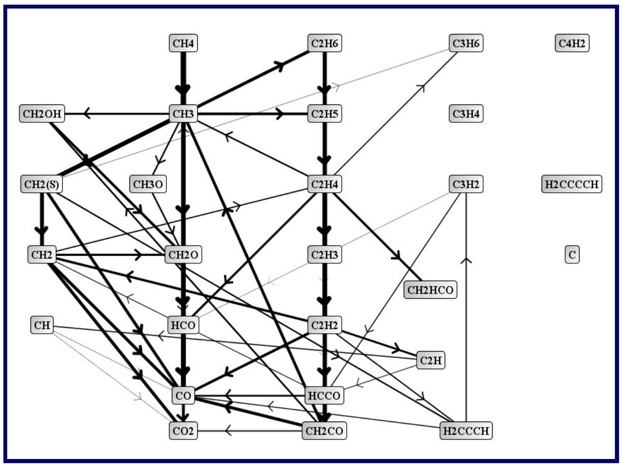
T=1865 K

C-atom fluxes, φ =1.0, methane–air flame



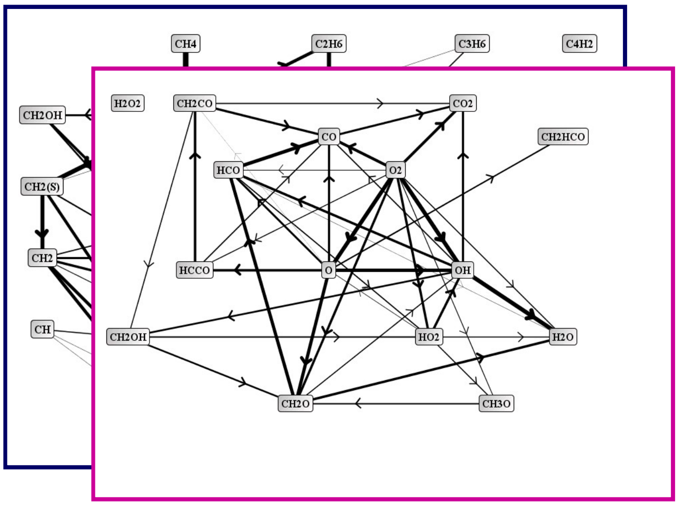
T=1915 K

Comparison of element fluxes in a methane–air explosion (φ =1.0, *T*=1800 K)



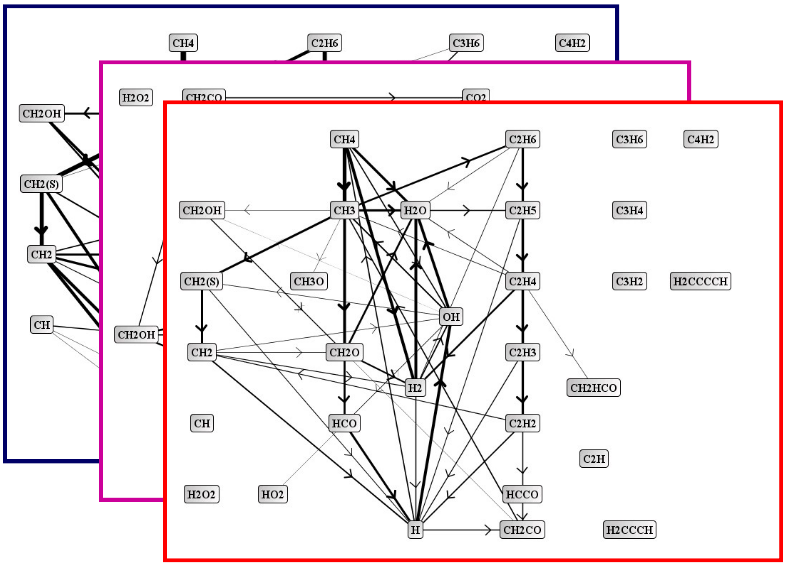
C-fluxes

Comparison of element fluxes in a methane–air explosion (φ =1.0, *T*=1800 K)



O-fluxes

Comparison of element fluxes in a methane–air explosion (φ =1.0, *T*=1800 K)

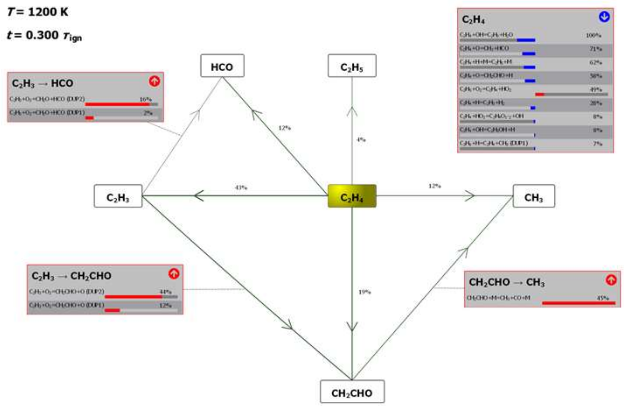


H-fluxes

Optima++ → FluxViewer++

FluxViewer++: C++ code for the visualization of the element fluxes and interactive analysis of chemical kinetic systems

Papp, Máté: Code FluxViewer++, 2023
https://respecth. hu/



Pathways for the consuption/production of a given species

Problem: What is the sequence of reactions that leads to the consuption (or production) of a given species? The detailed reaction mechanism is known.

Example: What is the sequence of reactions that leads to the consuption of methane in the stratosphere?

Answer:

$$\begin{aligned}
&CH_4 + OH + O_2 \rightarrow CH_3O_2 + H_2O\\
&CH_3O_2 + NO + O_2 \rightarrow CH_2O + HO_2 + NO_2\\
&CH_2O + h\upsilon \rightarrow CO + H_2\\
&NO + HO_2 \rightarrow NO_2 + OH\\
&NO_2 + h\upsilon \rightarrow NO + O\\
&O + O_2 \rightarrow O_3
\end{aligned}$$

The global reaction for CH_4 consumption:

$$CH_4 + 4\ O_2 \rightarrow 2\ O_3 + H_2O + H_2 + CO$$

R. Lehmann: An algorithm for the determination of all significant pathways in chemical reaction systems. *J. Atm. Chem.* **47**, 45-78 (2004)

Topic 3: Local sensitivity analysis 1

Local sensitivity coefficient and its interpretation,

system of differential equations
for local sensitivity coefficients,

initial concentration sensitivity coefficients,

calculation of sensitivity coefficients with
finite difference approximation,

calculation of sensitivity coefficients with
the Direct Method and the Decoupled Direct Method

automatic differentiation

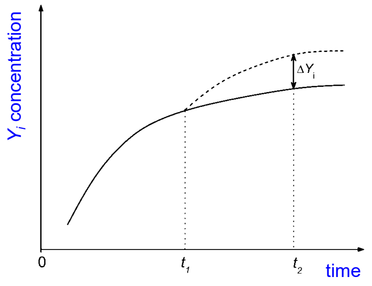
Local sensitivity analysis

Sensitivity analysis is a family of mathematical methods.
It investigates the dependence of the model results
on the values of the parameters

Local sensitivity analysis: investigates the effect of the small change of parameters

Local sensitivity coefficients can be investigated by a finite difference approximation:

$$\frac{\partial Y_i}{\partial p_j}(t_1,t_2) \approx \frac{\Delta Y_i(t_2)}{\Delta p_j} = \frac{Y_i'(t_2) - Y_i(t_2)}{\Delta p_j}$$



parameter is changed at time t_1
the result is observed at time t_2

T. Turányi: Sensitivity analysis of complex kinetic systems:
Tools and applications, *J. Math. Chem.*, **5**, 203-248(1990)

Local sensitivity analysis 2

Another approach: Taylor series expansion

$$Y_i(t,\mathbf{p}+\Delta\mathbf{p}) = Y_i(t,\mathbf{p}) + \sum_{j=1}^{m}\frac{\partial Y_i}{\partial p_j}\Delta p_j + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 Y_i}{\partial p_k \partial p_j}\Delta p_k \Delta p_j + \ldots$$

Local sensitivity coefficient: $s_{ik} = \dfrac{\partial Y_i}{\partial p_k}$

Local sensitivity matrix: $\mathbf{S} = \left\{\dfrac{\partial Y_i}{\partial p_k}\right\}$

The effect of parameter changes can be estimated using local sensitivities:

Changing a single parameter: $Y_i'(t_2) = Y_i(t_2) + \dfrac{\partial Y_i}{\partial p_j}\Delta p_j$

Changing several parameters: $\mathbf{Y}'(t_2) = \mathbf{Y}(t_2) + \mathbf{S}(t_1,t_2)\Delta\mathbf{p}(t_1)$

Local sensitivity analysis 3

$$\frac{\mathrm{d}\,\mathbf{Y}}{\mathrm{d}\,t} = \mathbf{f}(\mathbf{Y},\mathbf{p}) \qquad \mathbf{Y}(t_0) = \mathbf{Y}_0$$

Differentiation with respect p_j

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\frac{\partial \mathbf{Y}}{\partial p_j} = \mathbf{J}\frac{\partial \mathbf{Y}}{\partial p_j} + \frac{\partial \mathbf{f}}{\partial p_j} \qquad \frac{\partial\,\mathbf{Y}}{\partial\,p_j}(t_0) = 0 \qquad j = 1, 2, \ldots, m$$

The same equation with matrix-vector notation:

$$\dot{\mathbf{S}} = \underbrace{\mathbf{J}\,\mathbf{S}}_{\text{indirect effect}} + \underbrace{\mathbf{F}}_{\text{direct effect}}, \quad \mathbf{S}(0) = \mathbf{0}$$

where $\mathbf{J} = \left\{\dfrac{\partial f_i}{\partial Y_j}\right\}$ $\quad \mathbf{F} = \left\{\dfrac{\partial f_j}{\partial p_k}\right\}$

Initial concentration sensitivities

initial concentration sensitivities: the consequence of changing the initial conc. can be calculated with finite differences:

$$\frac{\partial Y_i}{\partial Y_j(t_1)}(t_2) \approx \frac{\Delta Y_i(t_2)}{\Delta Y_j(t_1)} = \frac{Y_i'(t_2) - Y_i(t_2)}{\Delta Y_j(t_1)}$$

kinetic system of ODEs: $\frac{\mathrm{d}\,\mathbf{Y}}{\mathrm{d}\,t} = \mathbf{f}(\mathbf{Y},\mathbf{p}) \qquad \mathbf{Y}(t_0) = \mathbf{Y}_0$

Differentiating it with respect to $Y_j(t_1)$:

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\frac{\partial\,\mathbf{g}}{Y_j^0(t_1)} = \mathbf{J}\frac{\partial\,\mathbf{g}}{\partial\,Y_j^0(t_1)} \qquad \frac{\partial\,\mathbf{Y}}{\partial\,Y_j^0(t_1)}(t_1) = \boldsymbol{\delta}_j \qquad j = 1, 2, \ldots, n$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{G}(t,t_1) = \frac{\partial\mathbf{f}}{\partial\mathbf{Y}}(t)\,\mathbf{G}(t,t_1) \qquad \mathbf{G}(t_1,t_1) = \mathbf{I}$$

the initial value of variable *j* is changed at time t_1 and the effect is read at time t_2

$$g_{ij}(t,t_1) = \frac{\partial Y_i(t)}{\partial Y_j^0(t_1)} \qquad \mathbf{g}_j(t,t_1) = \frac{\partial\mathbf{Y}(t)}{\partial Y_j^0(t_1)}$$

Green function matrix **G**

Calculation of local sensitivity coefficients

1 Brute force method (finite difference approximation)

$$\frac{\partial Y_i}{\partial p_j(t_1)}(t_2) \approx \frac{\Delta Y_i(t_2)}{\Delta p_j(t_1)} = \frac{Y_i'(t_2) - Y_i(t_2)}{\Delta p_j(t_1)}$$

Δp_j small: large error due to the representation of numbers

Δp_j large: large error due to nonlinearity

2 Direct method

2a. Coupled Direct Method:

coupled solution of the kinetic and sensitivity differential equations:

$$\frac{\mathrm{d}\,\mathbf{Y}}{\mathrm{d}\,t} = \mathbf{f}(\mathbf{Y},\mathbf{p}) \qquad \mathbf{Y}(t_0) = \mathbf{Y}_0$$

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\frac{\partial\mathbf{Y}}{\partial p_j} = \mathbf{J}\frac{\partial\mathbf{Y}}{\partial p_j} + \frac{\partial\mathbf{f}}{\partial p_j} \qquad \frac{\partial\,\mathbf{Y}}{\partial\,p_j}(t_0) = 0$$

The coupled solution is repeated for each parameter: $j = 1, 2, \ldots, m$

Lots of unnecessary calculations.

Calculation of local sensitivity coefficients 2

2b. Decoupled Direct Method (DDM):

joint solution of the kinetic and sensitivity diff. equations in each step:

$$\frac{\mathrm{d}\,\mathbf{Y}}{\mathrm{d}\,t} = \mathbf{f}(\mathbf{Y},\mathbf{p}) \qquad \mathbf{Y}(t_0) = \mathbf{Y}_0$$

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\frac{\partial \mathbf{Y}}{\partial p_j} = \mathbf{J}\frac{\partial \mathbf{Y}}{\partial p_j} + \frac{\partial \mathbf{f}}{\partial p_j} \qquad \frac{\partial\,\mathbf{Y}}{\partial\,p_j}(t_0) = 0 \qquad j = 1, 2, \ldots, m$$

The Jacobian of these equations are identical, therefore in each step

⇒ transformation of the Jacobian to a triangle matrix

⇒ selection of stepsize Δt based on the Jacobian

⇒ solution of the stiff ODE: calculation of new **Y**

⇒ calculation of the new sensitivity vector for parameter j =1
using the same triangle matrix

⇒ ⇒ ⇒ ⇒ ⇒ repeating for all parameters $j = 1, 2, \ldots, m$

⇒ ⇒ repeating for new time steps from the transformation of **J**

features:

- very fast method; the computer time only slightly increases with the number of parameters *m* (because the transformation of **J** is the most time-consuming)
- the accuracy of the solution can be controlled

Automatic differentiation

The simulation result calculated on a computer is obtained by a sequence of simple operations such as additions, multiplications, and elementary functions such as sines and cosines.
By applying the chain rule over and over again to these simple operations it is possible to calculate the derivatives to machine precision
in an automatic way.

ADIFOR: a Fortran simulation code is converted by a program
to a modified code for the calculation of the local sensitivity coefficients

Bischof, C., Carle, A., Khademi, P., Mauer: The ADIFOR 2.0 system for the automatic differentiation of FORTRAN 77 programes. *IEEE J. Comput. Sci. Eng.* **3**, 18-32. (1996)

ADIC: a C simulation code is converted by a program
to a modified code for the calculation of the local sensitivity coefficients

Bischof, C.H., Roh, L., Mauer-Oats, A.J.: ADIC: an extensible automatic differentiation tool for ANSI-C. *Soft. Pract. Exper.* **27**, 1427-1456 (1997)

Topic 4: Local sensitivity analysis 2

original and normalized sensitivities,

principal component analysis of the sensitivity matrix,

local uncertainty analysis,

applications of local sensitivities

Interpretation of local sensitivity coefficients

$$s_{ik} = \frac{\partial Y_i}{\partial p_k}$$

(Original) local sensitivity coefficients:
the parameter is changed by one unit
inspected: the result is changed by how many units
[unit of result / unit of parameter]

Normalized local sensitivity coefficients:

$$\tilde{s}_{ik} = \frac{p_k}{Y_i}\frac{\partial Y_i}{\partial p_k} = \frac{\partial \ln Y_i}{\partial \ln p_k}$$

investigates relative changes
How much % change of the result
due to 1 % change of the parameter?
dimension free

So far: single parameter is changed
effect on a single model result is investigated

Further information can also be extracted from sensitivity matrix **S** using principal component analyis, like the case when
several parameters are changed simultaneously, and
the effect on multiple model results is investigated.

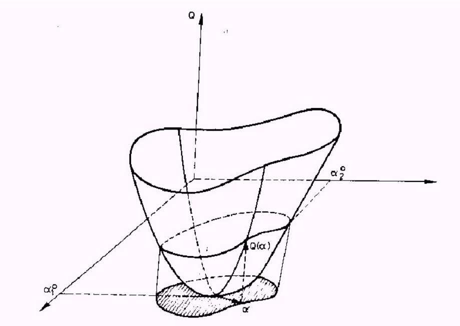
PCAS: principal component analysis of the sensitivity matrix S

Several parameters are changed simultaneously and the effect on several model outputs is investigated.

The effect of changing parameters is measured by a

Célfüggvény:

$$e(\mathbf{p}) = \int_{t_1}^{t_2} \sum_{i=1}^{m} \left(\frac{Y_i^*(t) - Y_i(t)}{Y_i(t)} \right)^2 \mathrm{d}t$$

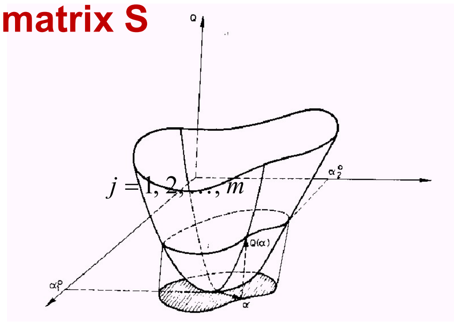


S. Vajda, P. Valkó, T. Turányi: Principal component analysis of kinetic models
Int. J. Chem. Kinet., **17**, 55-81(1985)

PCAS: principal component analysis of the sensitivity matrix S

The objective function
can be approximated by:

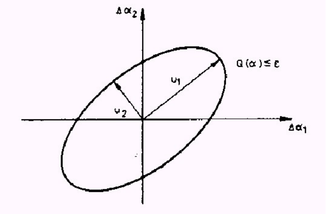
$$e(\boldsymbol{\alpha}) = (\Delta\boldsymbol{\alpha})^{\mathrm{T}} \tilde{\mathbf{S}}^{\mathrm{T}} \tilde{\mathbf{S}} (\Delta\boldsymbol{\alpha})$$



$j = 1, 2, \ldots, m$

where

$$\Delta\boldsymbol{\alpha} = \Delta \ln \mathbf{p} \qquad \tilde{\mathbf{S}} = \begin{bmatrix} \tilde{\mathbf{S}}_1 \\ \tilde{\mathbf{S}}_2 \\ \vdots \\ \tilde{\mathbf{S}}_n \end{bmatrix}$$

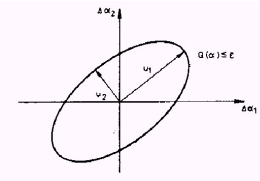


And the normed sensitivity matrix
belonging to time t_r

$$\tilde{\mathbf{S}}_r = \{(p_k / Y_i)(\partial Y_i(t_r) / \partial p_k)\}$$

PCAS: principal component analysis of the sensitivity matrix S

$$e(\boldsymbol{\alpha}) = (\Delta\boldsymbol{\alpha})^{\mathrm{T}} \tilde{\mathbf{S}}^{\mathrm{T}} \tilde{\mathbf{S}} (\Delta\boldsymbol{\alpha})$$



This quadratic form determines a (hyper) ellipsoid:

2D	ellipse
3D	ellipsoid (rugby ball shape)
4D	hyper ellipsoid

Another characterization of the hyper ellipsoid:
- lengths of the axes
- directions of the axes

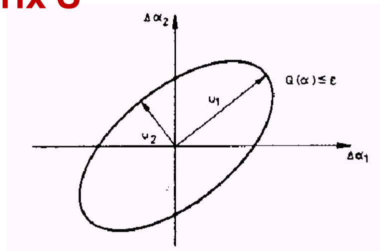
Eigenvalue-eigenvector decomposition of matrix $\tilde{\mathbf{S}}^{\mathrm{T}} \tilde{\mathbf{S}}$

λ_i eigenvalue *i* = length of axis *i*
if λ_i is small: the objective function increases rapidly to this direction
= parameter group *i* is highly influental

$\mathbf{u}_i$ eigenvector *i* = direction of axis *i*

PCAS: principal component analysis of the sensitivity matrix S

$$e(\boldsymbol{\alpha}) = (\Delta\boldsymbol{\alpha})^{\mathrm{T}} \tilde{\mathbf{S}}^{\mathrm{T}} \tilde{\mathbf{S}} (\Delta\boldsymbol{\alpha})$$



An alternative form of the objective function:

$$e(\boldsymbol{\alpha}) = \sum_{i=1}^{r} \lambda_i (\Delta\Psi_i)^2$$

where $\Delta\Psi_i = \mathbf{u}_i \boldsymbol{\alpha}$ transformed parameters called principal components

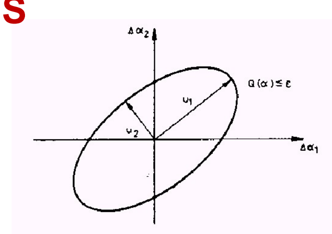
In the figure above:

λ_1 small ⇒ axis 1 is long; $\mathbf{u}_1 = (0.707, 0.707)$

λ_2 large ⇒ axis 2 is short; $\mathbf{u}_2 = (-0.707, 0.707)$

Note: the eigenvectors are unit vectors, therefore $0{,}707^2 + 0{,}707^2 = 1$

PCAS: principal component analysis of the sensitivity matrix S



Example 1:

λ_1 small ⇒ axis 1 is long; $\mathbf{u}_1 = (0.707, 0.707)$

λ_2 large ⇒ axis 2 is short; $\mathbf{u}_2 = (-0.707, 0.707)$

axis 1 is long → changing the parameters to direction $\mathbf{u}_1$
the objective function changes little

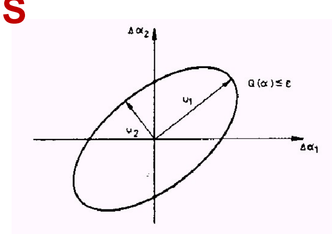
→ if α_2-α_1 = ln p_2 - ln p_1 = ln (p_2/p_1) constant, ⇒ little change of the objective function

→ if p_2/p_1 constant, ⇒ little change of the objective function

Thus, eigenvector $\mathbf{u} = (0.707, 0.707)$ means that keeping the ratio of the corresponding two parameters constant the inspected result(s) of simulation do not change.

Chemistry: the model results do not change if we keep the equilibrium constant $K=k_1/k_2$ fixed.

PCAS: principal component analysis of the sensitivity matrix S



Example 2:

$\mathbf{u}_1 = (0.707, 0.707, 0)$	large eigenvalue
$\mathbf{u}_2 = (-0.707, 0.707, 0)$	small eigenvalue
$\mathbf{u}_3 = (0, 0, 1)$	large eigenvalue

Interpretation of the eigenvectors:

p_1/p_2 and p_3 can be determined from the experimental data

p_3 can be determined independently

Only the ratio of p_1 and p_2 can be determined.

T. Perger, T. Kovács, T. Turányi, C. Treviño:
Determination of adsorption and desorption parameters from ignition temperature measurements in catalytic combustion systems, *J. Phys. Chem. B*, **107**, 2262-2274 (2003)

Local uncertainty analysis

If the parameters are correlated, then using the rule of spread of errors
the uncertainty of model results
can be calculated from the correlation matrix of parameters:

$$\boldsymbol{\sigma}_{\mathbf{Y}} = \mathbf{S}^{\mathrm{T}}\boldsymbol{\Sigma}_p\mathbf{S}$$

Here $\Sigma_{\mathbf{p}}$ is the covariance matrix of parameters, **S** is the sensitivity matrix and σ_{Y} is the variance of simulation results.

If the parameters are uncorrelated, then variance $\sigma^2(y)$ of model result *y*

can be calculated from the variance of parameters: $\sigma^2(p_k)$

$\sigma_k^2(y)$ is the contribution of parameter *k* to the variance of model result *y*

$$\sigma_k^2(y) = \sigma^2(p_k)\left(\frac{\partial y}{\partial p_k}\right)^2 \qquad \sigma^2(y) = \sum_k \sigma_k^2(y)$$

T. Turányi, L. Zalotai, S. Dóbé, T. Bérces: Effect of the uncertainty
of kinetic and thermodynamic data on methane flame simulation results
Phys. Chem. Chem. Phys., **4**, 2568-2578 (2002)

Local uncertainty analysis 2

- Linear approximation of the variance of the model result
- Does not take into account the nonlinear effects
- The result belongs to the nominal set of model parameters
- Realistic results, if the model behaves qualitatively similarly in the whole domain of parameters
- Non-realistic results, if the model is qualitatively different in the various parts of the parameter domain
- Provides separately the contribution of parameters
- Can be calculated fast

Applications of local sensitivities

1. Analysis of models
   - Estimation of the effect of parameter perturbation
   - Identification of cooperating parameters
2. Reduction of models
   - Identification of ineffective parameters;
     production of a simpler model with less parameters, but almost identical results
3. Local uncertainty analysis
   - May replace global uncertainty analysis:
     less accurate, much faster
4. Parameter estimation
   - All gradient methods are based on the (hidden) application of local sensitivity coefficients
   - Identification of effective parameters
   - Experimental design

Topic 5: Global uncertainty analysis and global sensitivity analysis

local and global uncertainty analysis

Morris' screening method,
Monte Carlo method,
Latin hypercube sampling,
Fourier Amplitude Sensitivity Test (FAST) method,
sensitivity indices,

surface response methods:
polynomial chaos expansion (PCE) method,
high-dimensional model representation (HDMR) method.

What is uncertainty analysis generally good for?

Global uncertainty analysis

Local uncertainty analysis
Provides information at the nominal parameter set
- well applicable, if the model behaves qualitatively similarly in the various regions of parameter space
- exact for linear models

Global uncertainty analysis
the whole physically possible region of parameters is investigated

Global *vs.* local uncertainty analysis
global methods require much more computer time
acquired information ~ computer time

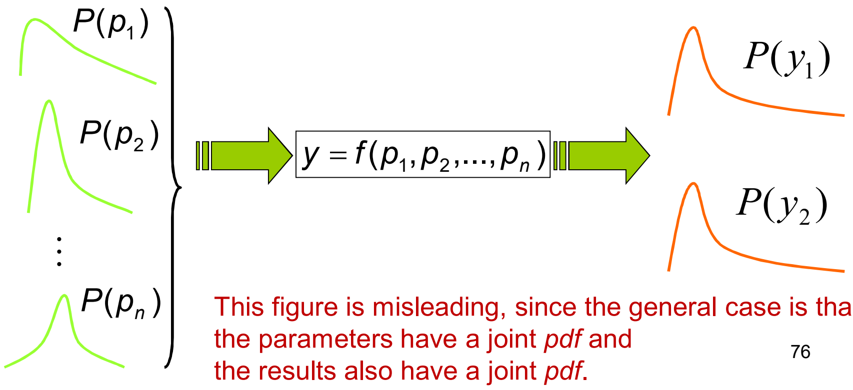
global uncertainty analysis	calculation of the uncertainty of model results from the uncertainty of model parameters
global sensitivity analysis	**as above** + identification of the individual contribution of the uncertainty of model parameters

---

Global uncertainty analysis 2

The uncertainty of parameters can be characterized by their probability density function (*pdf*)

The aims of global uncertainty analysis:
1. Calculation of the *pdf* of the results on the basis of the *pdf* of parameters
2. Determination of the contribution of the individual parameters to the standard deviation of model results

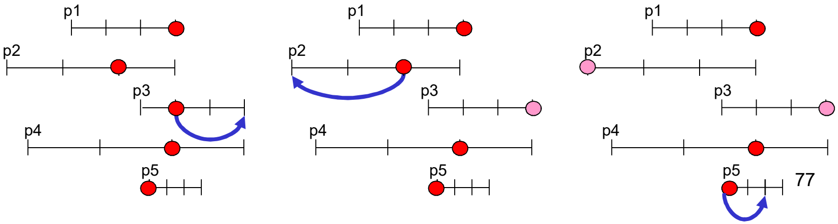


This figure is misleading, since the general case is that the parameters have a joint *pdf* and the results also have a joint *pdf*.

Morris method

Screening methods provide approximate information quickly
The Morris method allows the investigation of the effect of large parameter changes

- lower and upper uncertainty limits are assigned to each parameter.
- the uncertainty interval is divided to *n* parts for each parameter
- random parameter set is selected
- one parameter is changed at each run
- statistical interpretation of the results
- assumes uniform distribution of the parameters
- does not provide the *pdf* of the results
- intermediate computer time

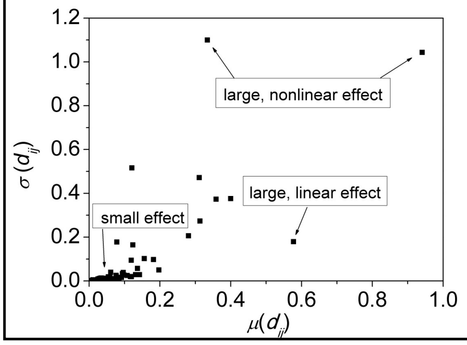


Morris method 2

Value d_{ij} shows the influence of parameter p_j at the random values of all other parameters within their uncertainty interval:

$$d_{ij} = \frac{Y_i\left(p_1^z, p_2^z, \ldots, p_j^z + \Delta, \ldots, p_N^z\right) - Y_i\left(\mathbf{p}^{z-1}\right)}{\left|\Delta\right|}$$

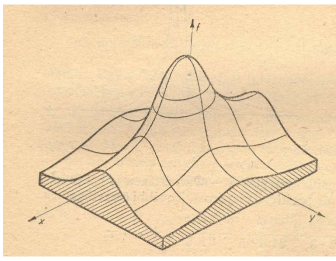
The d_{ij} values are calculated many times in a random calculation and the expected value and standard deviation of d_{ij} is determined.



M. D. Morris:
Factorial sampling plans for preliminary computational experiments.
Technometrics **33**, 161-174 (1991)

F. Campolongo, J. Cariboni, A. Saltelli:
An effective screening design for sensitivity analysis of large models.
Env.Model. Softw. **22**, 1509-1518 (2007)

Monte Carlo method



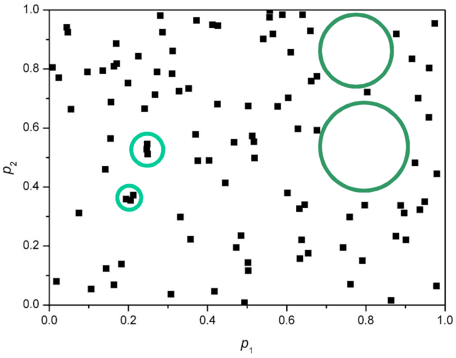
Several thousands of random parameter sets are generated in accordance with the joint *pdf* of the parameters.

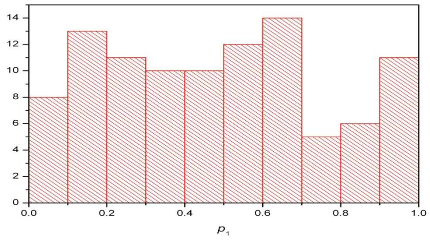
The simulations are carried out at these parameter sets.

Statistical analysis of the simulation results:
- determination of the histogram of a result
- calculation of the expected value and standard deviation

Problems:
- requires much computer time
- it is not easy to trace the effect of individual parameters

Monte Carlo method 2



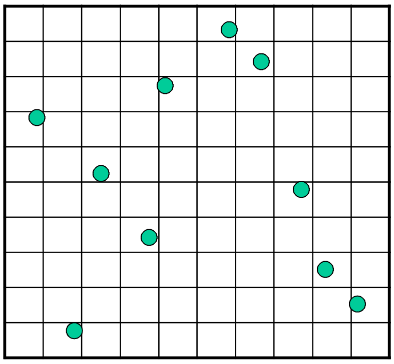


histogram of p_1

Generation of 100 (ξ.γ) points
ξ and γ are random numbers with [0.1] even distribution
$\Rightarrow$ used for the generation of parameter values p_1 and p_2 , respectively.

Clogging and empty spots

Monte Carlo method with Latin hypercube sampling



even distribution

M. D. McKay, R. J. Beckman, W. J. Conover: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **42**, 55-61 (2000)

J. C. Helton, F. J. Davis: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Engng Syst. Safety* **81**, 23–69 (2003)

- stripes („*strata*”) with equal probability are designated
- within each stripe a point is placed randomly
- if a stripe already contains a point, another point is not placed there

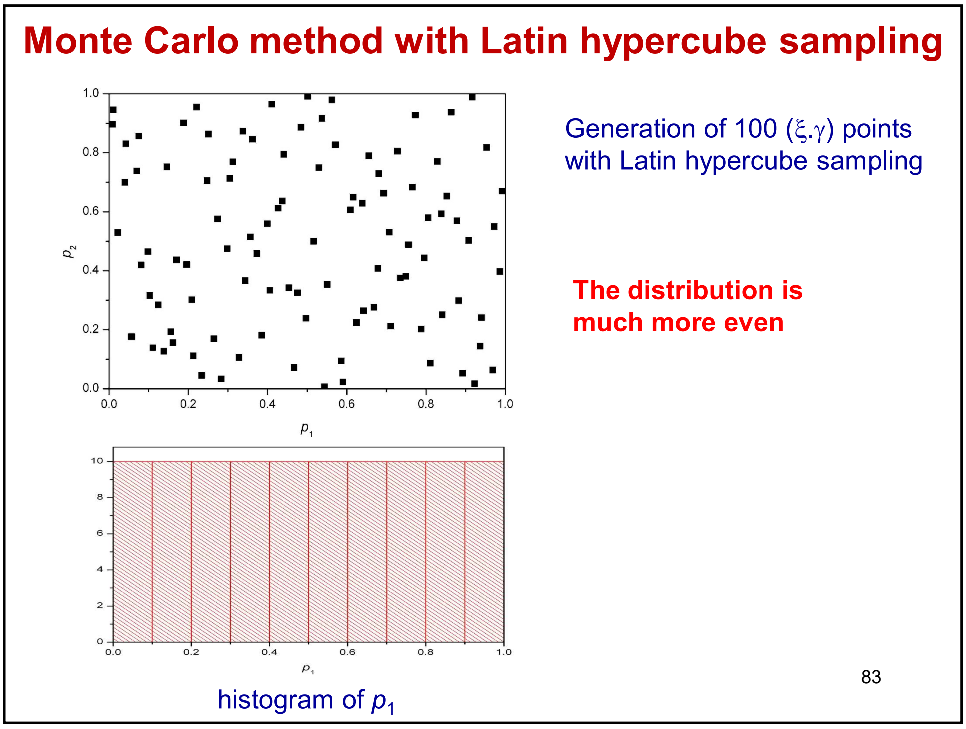
Monte Carlo method with Latin hypercube sampling

Sir Ronald Aylmer Fisher
(17 February 1890 – 29 July 1962)
British statistician and geneticist.

He has been described as
"a genius who almost single-handedly created the foundations for modern statistical science„.

His contributions to statistics include the maximum likelihood, the derivation of various sampling distributions, founding principles of the design of experiments, and much more. He developed the analysis of variance (ANOVA) method.

Stained glass window in the dining hall of Caius College, in Cambridge, commemorating Ronald Fisher and representing a Latin square, discussed by him in *The Design of Experiments*

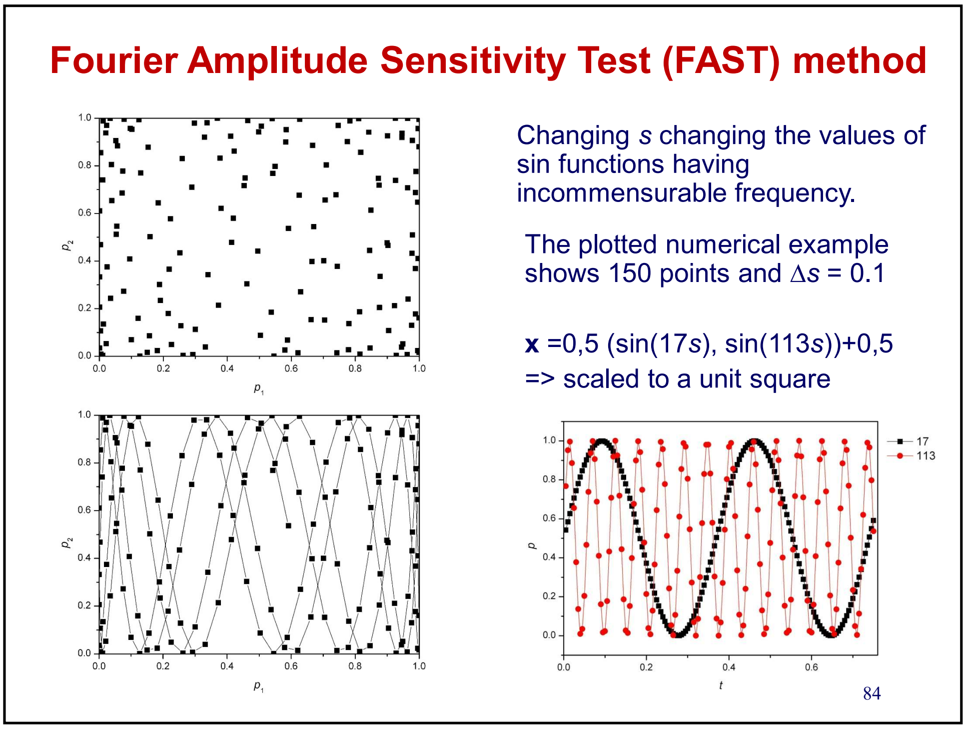


Monte Carlo method with Latin hypercube sampling

Generation of 100 (ξ.γ) points with Latin hypercube sampling

The distribution is much more even

histogram of p_1



Fourier Amplitude Sensitivity Test (FAST) method

Changing s changing the values of sin functions having incommensurable frequency.

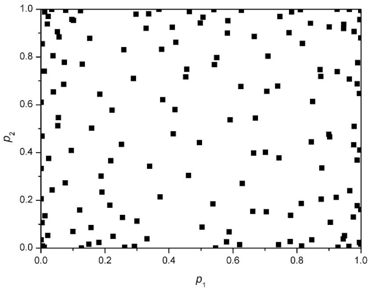
The plotted numerical example shows 150 points and Δs = 0.1

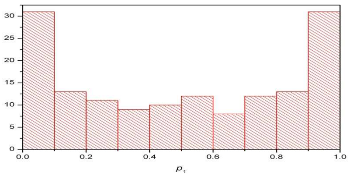
$\mathbf{x}$ =0,5 (sin(17s), sin(113s))+0,5
=> scaled to a unit square

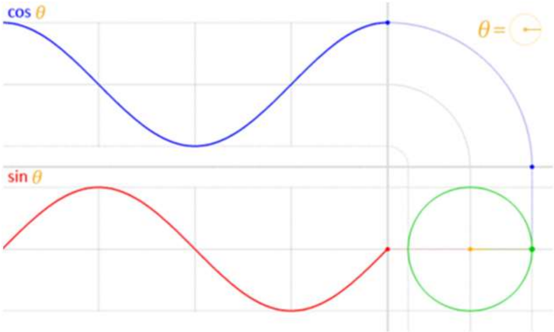
Fourier Amplitude Sensitivity Test (FAST) method 2

If the G_j functions are linear, then
(due to the nature of the sine function)
there are more points at the edges

Using appropriate G_j functions,
an *pdf* can be reconstructed
(for independent parameters)







Fourier Amplitude Sensitivity Test (FAST) method 3

E(Y_i) is the expected value of model result Y_i :

$$E(Y_i) = \iint \ldots \int h_i(p_1, p_2, \ldots, p_N)\, P(p_1, p_2, \ldots, p_N)\, \mathrm{d}\, p_1\, \mathrm{d}\, p_2 \ldots \mathrm{d}\, p_N$$

Y_i is the value of function h_i; *P* is the joint *pdf* of parameters **p**

Parameter p_j is changed by changing scalar *s*

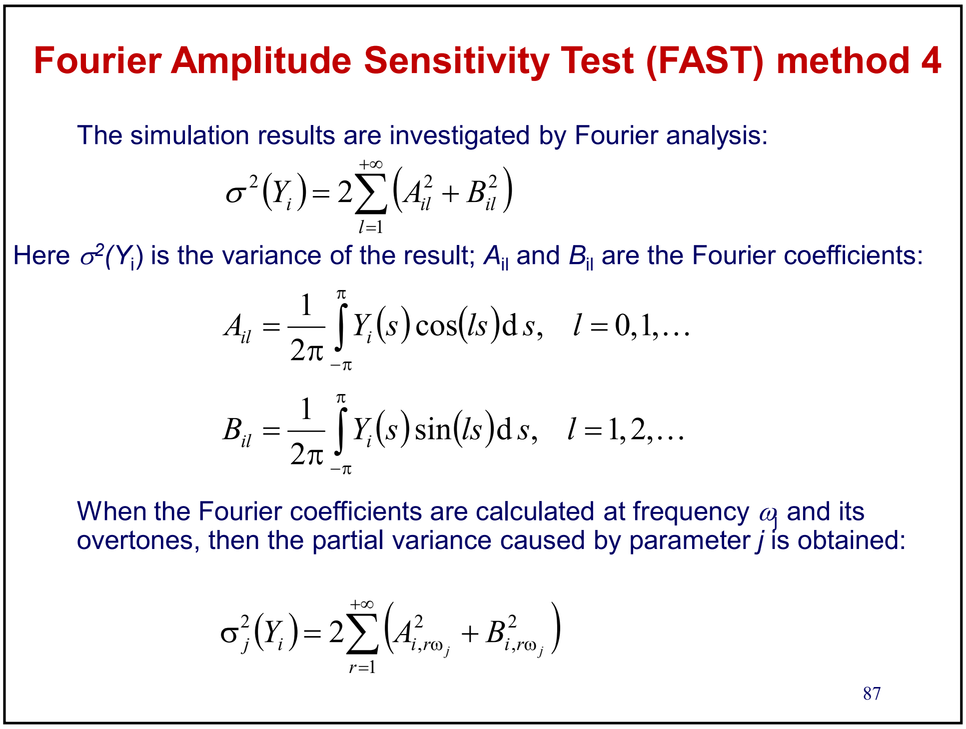
$$p_j(s) = G_j(\sin \omega_j s)$$

Function G_j has to be selected to to reproduce the joint probability density of parameters **P**

ω_j is the frequency assigned to parameter p_j
The frequencies have to be relative primes.

If $-\pi < s < \pi$ and $\Delta s = 2\pi/N$ ⇒ *N* points are located in the space of parameters; the local density corresponds to the *pdf*

R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petschek, J. H. Schaibly:
Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients I. Theory.
J. Chem. Phys. **59**, 3873-3878 (1973)



Fourier Amplitude Sensitivity Test (FAST) method 4

The simulation results are investigated by Fourier analysis:

$$\sigma^2(Y_i) = 2\sum_{l=1}^{+\infty}\left(A_{il}^2 + B_{il}^2\right)$$

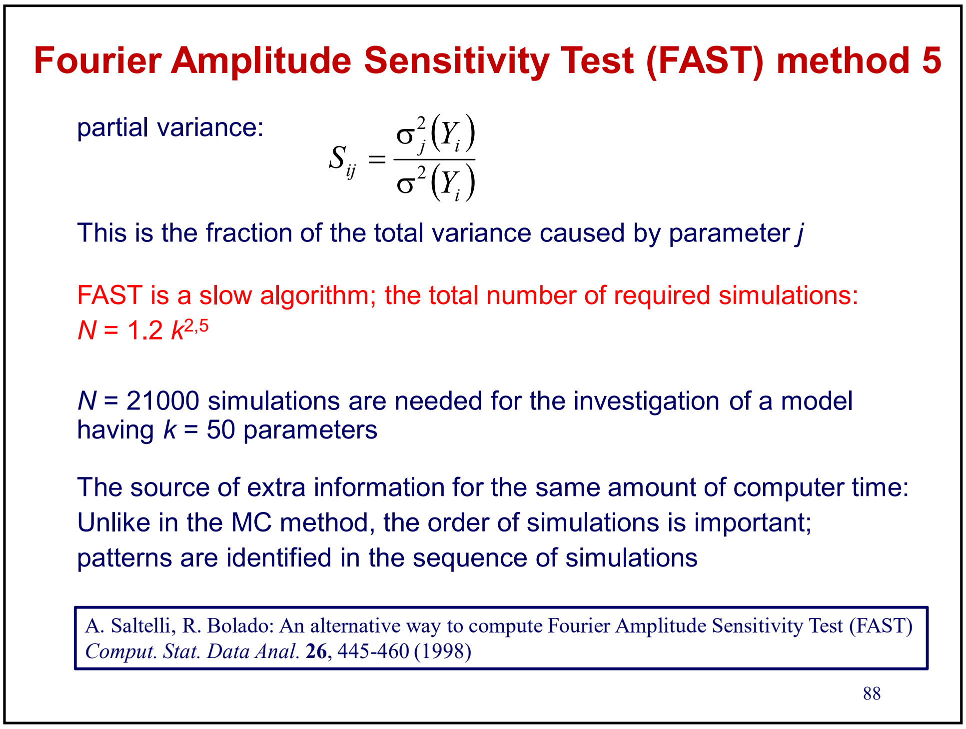
Here $\sigma^2(Y_\mathrm{i})$ is the variance of the result; A_il and B_il are the Fourier coefficients:

$$A_{il} = \frac{1}{2\pi}\int_{-\pi}^{\pi} Y_i(s)\cos(ls)\,\mathrm{d}\,s, \quad l = 0,1,\ldots$$

$$B_{il} = \frac{1}{2\pi}\int_{-\pi}^{\pi} Y_i(s)\sin(ls)\,\mathrm{d}\,s, \quad l = 1,2,\ldots$$

When the Fourier coefficients are calculated at frequency ω_j and its overtones, then the partial variance caused by parameter j is obtained:

$$\sigma_j^2(Y_i) = 2\sum_{r=1}^{+\infty}\left(A_{i,r\omega_j}^2 + B_{i,r\omega_j}^2\right)$$



Fourier Amplitude Sensitivity Test (FAST) method 5

partial variance:

$$S_{ij} = \frac{\sigma_j^2(Y_i)}{\sigma^2(Y_i)}$$

This is the fraction of the total variance caused by parameter j

FAST is a slow algorithm; the total number of required simulations:
$N = 1.2\ k^{2,5}$

N = 21000 simulations are needed for the investigation of a model having k = 50 parameters

The source of extra information for the same amount of computer time:
Unlike in the MC method, the order of simulations is important;
patterns are identified in the sequence of simulations

A. Saltelli, R. Bolado: An alternative way to compute Fourier Amplitude Sensitivity Test (FAST)
Comput. Stat. Data Anal. **26**, 445-460 (1998)

Sensitivity indices

May be considered as a further developed version of FAST:
The expected value of Y_i :

$$E(Y_i)=\iint\ldots\int f_i(p_1,p_2,\ldots,p_N)P(p_1,p_2,\ldots,p_N)\quad \mathrm{d}\,p_1\,\mathrm{d}\,p_2\ldots\mathrm{d}\,p_N$$

The variance of Y_i :

$$V(Y_i)=\iint\ldots\int\left(f_i(p_1,p_2,\ldots,p_N)-E(Y_i)\right)^2 P(p_1,p_2,\ldots,p_N)\mathrm{d}\,p_1\,\mathrm{d}\,p_2\ldots\mathrm{d}\,p_N=$$
$$=\iint\ldots\int f_i^2(p_1,p_2,\ldots,p_N)P(p_1,p_2,\ldots,p_N)\mathrm{d}\,p_1\,\mathrm{d}\,p_2\ldots\mathrm{d}\,p_N-E^2(Y_i)$$

Variance of Y_i , if parameter p_j is fixed: $V(Y_i|p_j)$
Its expected value: $E(V(Y_i|p_j))$
Variance of Y_i caused by p_j : $V(E(Y_i|p_j))=V(Y_i)-E(V(Y_i|p_j))$

First order uncertainty index (similar to FAST partial variance):

$$S_{j(i)}=\frac{V\left(E\left(Y_i|p_j\right)\right)}{V(Y_i)}$$

A. Saltelli: Making best use of model evaluations to compute sensitivity indices, *Comput. Phys Commun.*, **145**, 280-297 (2002)

Sensitivity indices 2

Variance of Y_i , caused by parameters p_j and p_k together: $V(E(Y_i|p_j,p_k))$

It can be used for the calculation of the *second order uncertainty index*:

$$S_{kj(i)}=\frac{V\left(E\left(Y_i|p_k,p_j\right)\right)-V\left(E\left(Y_i|p_k\right)\right)-V\left(E\left(Y_i|p_j\right)\right)}{V(Y_i)}$$

This index shows the interaction of parameters p_j and p_k

The *n*-th order uncertainty index can be obtained in a similar way.

Example: a model has three parameters: *a, b, c*
The total index: $S^{\mathrm{tot}}_{a(i)}=S_{a(i)}+S_{ab(i)}+S_{ac(i)}+S_{abc(i)}$

parameter *j* has no correlations: $S_{j(i)}=S^{\mathrm{tot}}_{j(i)}$

interactions of parameter *j*: $S^{\mathrm{tot}}_{j(i)}-S_{j(i)}$

Sensitivity indices 3

- Global method
- Application of pseudo random numbers allow the fast calculation of the integrals
- Calculates the the first and higher order effects
- Calculates the total effect
- Takes into account the *pdf* of the parameters
- Requires much computer time (about 25000 runs for 50 parameters)

Surface response methods (SRMs)

A whole family of the global sensitivity analysis methods is based on the idea that the original model (*e.g.* an ODE or PDE based complex model) is approximated by a simpler function and the global sensitivity analysis is carried out with the help of the simpler model.

Gaussian process emulator methods

Uses metamodels based on the assumption that for target outputs **Y** = **f**(**x**), the value of Y at an unknown value of **x** follows a multivariate Gaussian distribution. Suitable for systems with a small number of main effects and only weak parameter interactions.

Polynomial chaos expansion (PCE) methods

Polynomial chaos (PC; also called Wiener chaos expansion) is a non-sampling-based method to determine evolution of uncertainty in dynamical system, when there is probabilistic uncertainty in the system parameters. (Wiener, 1938)

High-dimensional model representation (HDMR) methods

To be discussed in details later.

Polynomial chaos expansion (PCE)

It is not related to „chaos” in the dynamical systems theory way.
It has been used several times in combustion modelling.

M. T. Reagan, H. N. Najm, B. J. Debusschere, O. P. Le Maitre, O. M. Knio, R. G., Ghanem: Spectral stochastic uncertainty quantification in chemical systems. *Combust. Theor. Model.* **8**, 607-632 (2004)

D.A. Sheen, X. You, H. Wang, T. Løvås: Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. Proc. Combust. Inst. 32, 535-542 (2009)

x_i, x_j uncertain Arrhenius parameters *A*, scaled to interval [–1,+1]
description of model response $\eta_r(\mathbf{x})$ with a second order polynomial:

$$\eta_r(\mathbf{x}) = \eta_{r,0} + \sum_{i=1}^{m} a_{r,i} x_i + \sum_{i=1}^{m} \sum_{j \ge i}^{m} b_{r,i,j} x_i x_j$$

The uncertainty in **x** may be expressed as a polynomial expansion of basis random variables **ξ**,

$$\mathbf{x} = \mathbf{x}^{(0)} + \sum_{i=1}^{m} \boldsymbol{\alpha}_i \xi_i + \sum_{i=1}^{m} \sum_{j \ge i}^{m} \boldsymbol{\beta}_{ij} \xi_i \xi_j + \ldots,$$

93

Polynomial chaos 2

Combining the previous two equations:

$$\eta_r(\boldsymbol{\xi}) = \eta_r(\mathbf{x}^{(0)}) + \sum_{i=1}^{m} \hat{\boldsymbol{\alpha}}_{r,i} \xi_i + \sum_{i=1}^{m} \sum_{j \ge i}^{m} \hat{\boldsymbol{\beta}}_{r,ij} \xi_i \xi_j + \ldots,$$

where $\hat{\boldsymbol{\alpha}}_r = {}^1\!/_2 \mathbf{I}_m \mathbf{a}_r$ $\qquad \hat{\boldsymbol{\beta}}_r = {}^1\!/_4 \mathbf{I}_m^T \mathbf{b}_r \mathbf{I}_m$

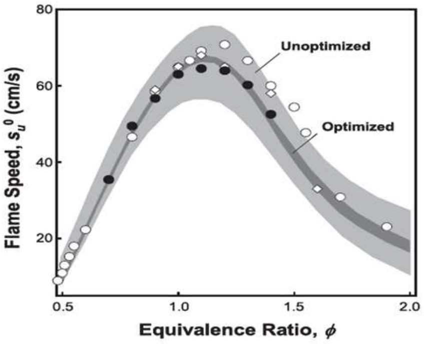
⇒ the overall model prediction is given by its nominal value plus uncertainty contributions from each rate coefficient. The overall output variance may then be represented as the sum over terms involving the coefficients of the equivalent expansion:

$$\sigma_r(\boldsymbol{\xi})^2 = \sum_{i=1}^{m} \hat{\boldsymbol{\alpha}}^2{}_{r,i} + 2\sum_{i=1}^{m} \hat{\boldsymbol{\beta}}^2{}_{r,ij} + \sum_{i=1}^{m} \sum_{j > i}^{m} \hat{\boldsymbol{\beta}}^2{}_{r,ij}$$

D.A. Sheen, X. You, H. Wang, T. Løvås: Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. *Proc. Combust. Inst.* **32**, 535-542 (2009)

Polynomial chaos 3

A typical result from Sheen *et al.* (2009):



Light grey: prior uncertainty of ethylene–air flame velocity from the uncertainty factors *f* of Baulch *et al.* (2004).
Dark grey: posterior sensitivities from the optimized model.
Symbols: experimental data.

D.A. Sheen, X. You, H. Wang, T. Løvås: Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. *Proc. Combust. Inst.* **32**, 535-542 (2009)

HDMR method

High Dimensional Model Representation

The simulation results are approximated by a polynomial of the parameters:

$$Y(\mathbf{x}) = Y_0 + \sum_{i=1}^{n} Y_i(x_i) + \sum_{1 \le i < j \le n} Y_{ij}(x_i, x_j) + \ldots$$

Y(x_i) the only variable is parameter x_i
But the function can be even an 8th order polynomial!

Y(x_i, x_j) the variables are parameters x_i and x_j
Two variables only, but it can also be a high-order polynomial!

T. Ziehn, A. S. Tomlin: GUI-HDMR - A software tool for global sensitivity analysis of complex models. *Environ. Model. Soft.* **24**, 775-785 (2009)

HDMR-method 2

Types:
cut HDMR
The polynomial is generated from a reference point

random sampling HDMR (RS-HDMR):
Generation of random points in a parameter domain,
fitting polynomials to these points

Examples:

Approximation with base functions: $Y_i(x_i) = \sum_{r=1}^{k} \alpha_r^i \, \varphi_r(x_i)$

Partial variances: $D_i = \sum_{r=1}^{k_i} \left(\alpha_r^i\right)^2$

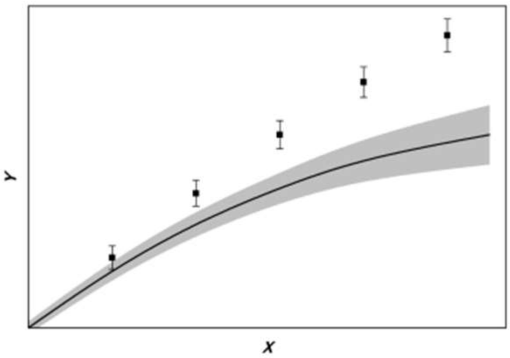
Sensitivity indices: $S_{i_1,\dots,i_s} = \frac{D_{i_1,\dots,i_s}}{D}, \qquad 1 \le i_1 < \dots < i_s \le m$

Comparison of the methods

	local	Morris	MC LHS	sens. index
input variance	✓	✓	✓	✓
input *pdf*	×	×	✓	✓
output *pdf*	×	×	✓	×
output variance	✓(linear)	×	✓	✓(biased)
CPU requirement?	✓ **1**	**2110**	**3000**	**16280**
Individual contributions	✓ (linear)	✓(only qualitative)	×	✓
global?	×	✓	✓	✓
info about the non-linearities	×	✓(only qualitative)	×	✓

Uncertainty analysis: What is it good for?

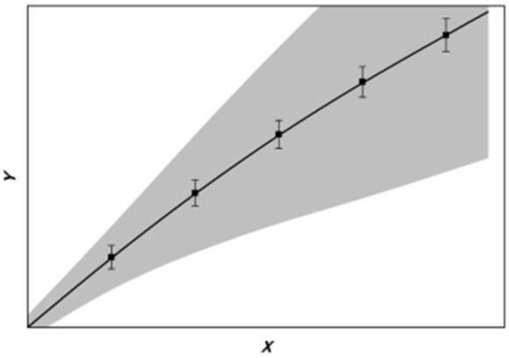
Is the structure of the model OK?



if the experimental data are correct and
if the uncertainty regions of the parameters are well known
BUT the uncertainty regions of measured and simulated results
do not overlap ⇒ the structure of the model is bad (wrong equations!)

What is it good for? 2

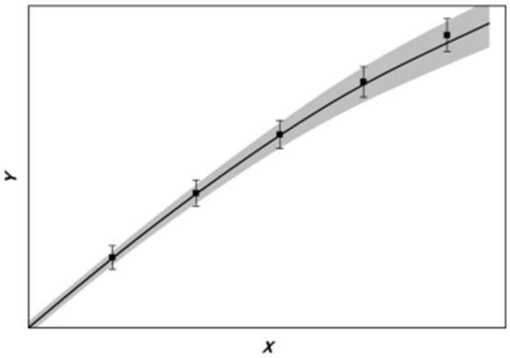
Is the model well established?



if the uncertainty of the simulated results is much wider than
the uncertainty of the data
⇒ any simulation result can be obtained with the parameters
⇒ the model is not useful

What is it good for? 3

Are the model parameters well known?



the uncertainty ranges of the data overlap with the uncertainty of the simulation results; the two uncertainty ranges are similar

⇒ the model is OK,
but the uncertainty of the simulation results can be decreased,
if the critical parameters (to be identified by uncertainty analysis)
are determined with smaller uncertainty

Summary of uncertainty analysis methods

Local uncertainty analysis

one parameter is changed at a time;
based on partial derivatives
can be calculated quickly

Screening methods

several parameters are changed in wide parameter ranges
intermediate computer time requirement
Morris method

Global uncertainty analysis

all parameters are changed simulataneously
according to their joint *pdf*
requires much computer time
e.g. Monte Carlo method (with Latin hypercube sampling)

Topic 6: Uncertainty of the thermodynamic and kinetic parameters

uncertainty of thermodynamic data,
Active Thermochemical Tables (ATcT),

direct and indirect measurements,

estimation of uncertainty for gas kinetic rate coefficients,

EXAMPLE:
applications of several uncertainty analysis methods to a methane flame model

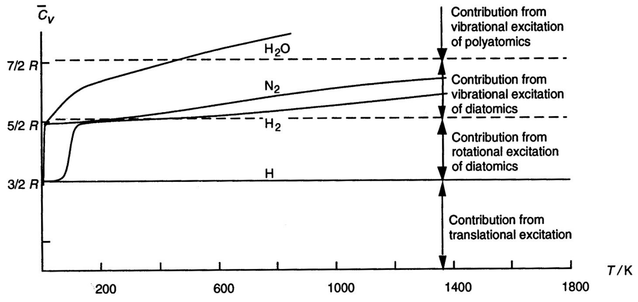
Temperature dependence of thermodynamic data

NASA polynomials

$$\frac{H^{\theta}}{RT} = a_1 + \frac{a_2}{2}T + \frac{a_3}{3}T^2 + \frac{a_4}{4}T^3 + \frac{a_5}{5}T^4 + \frac{a_6}{T}$$

$$\frac{c_p}{R} = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4$$

$$\frac{S^{\theta}}{R} = a_1 \ln T + a_2 T + \frac{a_3}{2}T^2 + \frac{a_4}{3}T^3 + \frac{a_5}{4}T^4 + a_7$$



Using thermodynamic data in combustion simulations

ΔH_f ⇒ calculation of heat production in a reacting mixture
→ calculation of temperature changes
⇒ calculation of $\Delta_r G^0$

c_p ⇒ calculation of temperature changes

ΔS ⇒ calculation of $\Delta G = \Delta H - T\Delta S$
→ calculation of the equilibrium constant
→ calculation of the rate coefficient of reverse reactions

Uncertainty of thermodynamic data

thermodynamic data influence the reaction kinetic calculations in two ways:
- Calculated temperature
- Calculation of the rate coefficients of backward reaction steps

Thermodynamic data used:
- heat capacity (can be calculated using statistical thermodynamics)
- entropy (can be calculated using statistical thermodynamics)
- standard enthalpy of formation (measurement or high level calculation)

↓

- The databases contain the recommended values and variances
- Are the enthalpies of formation correlated?

Uncertainty of thermodynamic data

c_p and ΔS can be calculated from the IR spectrum
using methods of statistical thermodynamics

ΔH_f - can be computed
(for small molecules only; not easy)
- can be determined experimentally by

- measuring the equilibrium constant of a reaction
  → reaction enthalpy → enthalpy of formation
- measuring ionization energy by mass spectrometry

Uncertainty of thermodynamic data 2

Typical uncertainty of ΔH_f (1σ):

molecules and small radicals: 0.1-0.5 kJ/mole
e.g. CO= 0.17 kJ/mole, CH_4= 0.4 kJ/mole, CH_3=0.4 kJ/mole

large radicals: 1.0 – 5.0 kJ/mole
e.g. HO_2= 3.35 kJ/mole, CH_2OH= 4.2 kJ/mole

less known radicals: 8-10 kJ/mole
e.g. HCCO= 8.8 kJ/mole, CH_2HCO= 9.2 kJ/mole

Determination of the enthalpies of formation

Methods for the determination of enthalpies of formation ΔH_f :

1) direct experimental determination:
   calorimetry; synthesis from reference state elements
   $H_2 + ½ O_2 = H_2O$ applicable for few compounds only

2) direct experimental determination from MS ionization energies
   applicable for few compounds only; not very accurate

3) direct theoretical calculation
   high level *ab initio* method required: accurate for small molecules only

4) generally applicable method:
   indirectly from experimentally measured reaction enthalpies $\Delta_r H^\Theta$
   determination of ΔH_f after a chain of calculations

Determination of the enthalpies of formation 2

„4) determination of ΔH_f after a chain of calculations"
- starting from directly determined ΔH_f values

$$\Delta_r H^\Theta = \sum_j \nu_j H_f^\Theta(j)$$

a) combining it with a $\Delta_r H^\theta$ value provides a new ΔH_f^θ value
   $\Rightarrow$ indirectly determined ΔH_f^θ
b) GO TO a) until we get the required ΔH_f^θ

$\Rightarrow$ the chain of calculation provides the required ΔH_f^θ

PROBLEMS:
- Going on in the chain of calculations, the errors are accumulated
  ΔH_f^θ values at the end of a long chain are not very accurate.
- ΔH_f^θ values for the same species can be obtained at the ends of two different calculation chains $\Rightarrow$ different ΔH_f^θ values are obtained **?????**

Active Thermochemical Tables (ATcT)

Idea of Branko Ruscic http://atct.anl.gov/

the determination of many enthalpies of formation ΔH_f^{θ} in one step:

using *n* direct experimental determination: $H_f^{\Theta}(j) = A_j$ $j = 1, \ldots, n$

using *m* measured $\Delta_r H^{\theta}$ values: $\Delta_r H_i^{\theta} = \sum_j \nu_{ij} H_f^{\Theta}(j)$ $j = n+1, \ldots, n+m$

The aim is the determination of *k* values of ΔH_f^{θ} :

- if $k > n+m$ ⇒ not enough info
- if $k < n+m$ ⇒ overdetermined linear algebraic system of equations
  ⇒ determination of the ΔH_f^{θ} values by the least squares method

If the errors of the measurements are also taken into account

⇒ **weighted least squares method**

B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszevski, S. J. Bittner, S. G. Nijsure, K. A.Amin, M. Minkoff, A. F. Wagner: Introduction to Active Thermochemical Tables: Several „key" enthalpies of formation revisited. *J. Phys. Chem. A* **108**, 9979-9997 (2004)

Active Thermochemical Tables (ATcT) 2

NOTES:

The reason of the name: the original idea was that the tables would be „active": on a Web site adding new measurement data all enthalpies of formation would be recalculated.

It never worked this way: Dr. Ruscic is continously adding new measurements and sometimes publishes $\Delta_r H^{\theta}$ values.

Please observe the similarity and difference between ATcT and the optimization of kinetic reaction mechanisms:

- using both direct and indirect measurements
- the error of measurements is used for the calculation of the uncertainty of parameters
- **ATcT**: the simulated data are a linear functions of the parameters

kinetics: the simulated data are obtained by solving ODEs or PDEs (strongly nonlinear functions of parameters)

B. Ruscic, R. E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, A. F. Wagner, Active Thermochemical Tables: Thermochemistry for the 21st Century. *J. Phys. Conf. Ser.* **16**, 561-570 (2005)

Rate coefficient uncertainties

Uncertainty factor f_j as defined in data evaluations
(Tsang, Warnatz, Baulch, Konnov):

uncertainty factor u_j uncertainty parameter f_j

$$u_j = \frac{k_j^0}{k_j^{\min}} = \frac{k_j^{\max}}{k_j^0} \qquad f_j = \log_{10}\left(u_j\right)$$

k_j^0	recommended value of the rate coefficient of reaction j
$k_j^{\min}$	possible minimal value of k_j
$k_j^{\max}$	possible maximal value of k_j

$\Rightarrow$ $[k_j^{\min}, k_j^{\max}]$ is the physically realistic range for the rate coefficients

assume that ln $k^{\min}$ and ln $k^{\max}$ deviate 3σ from ln k^0

$$\Rightarrow \sigma^2(\ln k_j) = ((f_j \ln 10)/3)^2$$

1σ uncertainty limit (assuming that u corresponds to 3σ): $l = 10^{f/3}$

Uncertainty of *k* at a given temperature

Uncertainty of (direct) rate coefficient measurements:

very high quality data	uncertainty factor u= 1.26 ⇔ f=0.1 ⇔ ± 8 % (1σ)
typical good data	uncertainty factor u= 2.00 ⇔ f=0.3 ⇔ ±26 % (1σ)
typical data	uncertainty factor u= 3.16 ⇔ f=0.5 ⇔ ±47 % (1σ)

(high level) theoretical determinations:
TST/master equation calculations

best systems	uncertainty factor u= 2.00 ⇔ f=0.3 ⇔ ±26 % (1σ)
multi well, main channels	uncertainty factor u= 3.16 ⇔ f=0.5 ⇔ ±47 % (1σ)
multi well, minor channels	uncertainty factor u=10 ⇔ f=1.0

C. F. Goldsmith, A. S. Tomlin, S. J. Klippenstein: Uncertainty propagation in the derivation of phenomenological rate coefficients from theory: A case study of *n*-propyl radical oxidation
Proc. Combust. Inst., **34**, 177-185 (2013)

J. Prager, H. N. Najm, J. Zádor: Uncertainty quantification in the *ab initio* rate-coefficient calculation for the $CH_3CH(OH)CH_3$+OH → $CH_3C.(OH)CH_3$+H_2O reaction, *Proc. Combust. Inst.*, **34**, 583-590 (2013)

Local uncertainty analysis of chemical kinetic models

$f_j \rightarrow \sigma^2(\ln k_j)$ uncertainty parameter f_i is transformed to the variance of ln k_j

$\partial Y_i / \partial \ln k_j$ seminormalized local sensitivity coefficients

$$\sigma^2_{\mathrm{K}j}(Y_i) = \left(\partial Y_i / \partial \ln k_j\right)^2 \sigma^2\left(\ln k_j\right)$$

contribution of the uncertainty of parameter k_j to the variance of result Y_i

$\sigma^2_{\mathrm{K}}(Y_i) = \sum_j \sigma^2_{\mathrm{K}j}(Y_i)$ variance of result Y_i due to kinetic uncertainties

$$\sigma^2_{Tj}(Y_i) = \left(\partial Y_i / \partial \Delta_f H^{\circ}_{298}(j)\right)^2 \sigma^2\left(\Delta_f H^{\circ}_{298}(j)\right)$$

contribution of the uncertainty of the enthalpy of formation of species *j* to the variance of result Y_i

$$\sigma^2(Y_i) = \sigma^2_K(Y_i) + \sigma^2_T(Y_i) = \sum_j \sigma^2_{Kj}(Y_i) + \sum_j \sigma^2_{Tj}(Y_i)$$

estimated total variance of result Y_i from both kinetic and thermodynamic uncertainties

T. Turányi, L. Zalotai, S. Dóbé, T. Bérces:
Effect of the uncertainty of kinetic and thermodynamic data on
methane flame simulation results, *Phys.Chem.Chem.Phys.*, **4**, 2568-2578 (2002)

Example: the uncertainty of methane flame simulation results

The investigated methane flames:

- one dimensional, adiabatic, freely propagating, laminar, premixed stationary flame investigated at equivalence ratios φ = 0.70 (lean), 1.00 (stoichiometric), and 1.20 (rich)
- cold boundary conditions p = 1.0 atm and T = 298.15 K

Monitored outputs:

- laminar flame velocity
- maximum temperature
- maximum species concentration of H, O, OH, CH, CH_2

Uncertainty analysis of a laminar methane flame

Leeds Methane Oxidation Mechanism:
37 species and 175 reversible reactions
stationary, laminar 1D simulations

37 species: the recommended values of the enthalpies of formation and their variance was calculated from thermodynamic databases

175 reactions: uncertainty parameteres *f* were collected from Baulch *et al.*

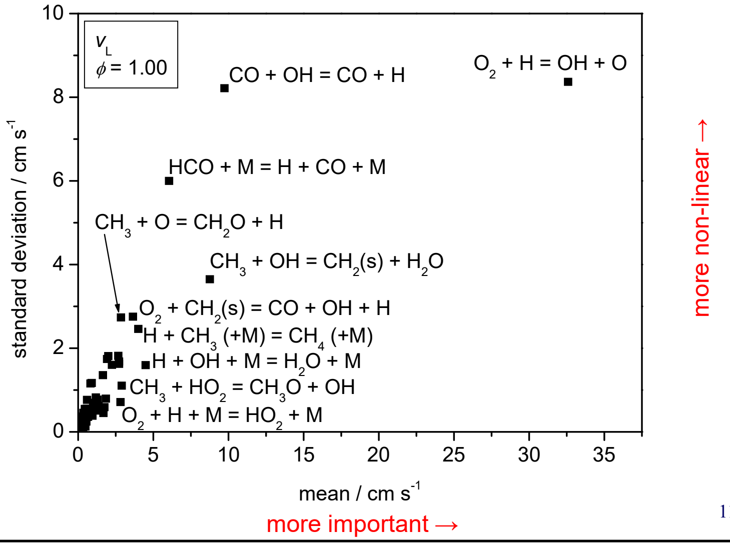
The investigated simulation results:
maximal flame temperature, laminar flame velocity,
maximal concentrations of radicals H, O, OH, CH, CH_2

Uncertainty analysis methods:
local uncertainty analysis, Morris' method,
Monte Carlo with Latin Hypercube sampling, sensititivty indices

J. Zádor, I. Gy. Zsély, T. Turányi, M. Ratto, S. Tarantola, A. Saltelli: Local and global uncertainty analyses of a methane flame model, *J. Phys. Chem. A*, **109**, 9795-9807 (2005)

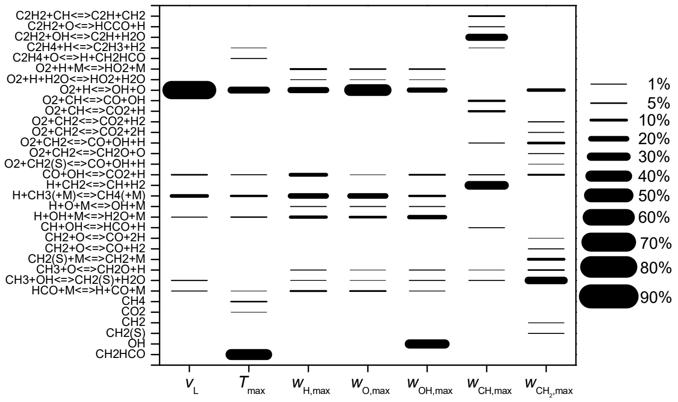
Morris' method results (φ=1)

contribution of parameter uncertainties to the uncertainty of the flame velocity



Local uncertainty analysis results

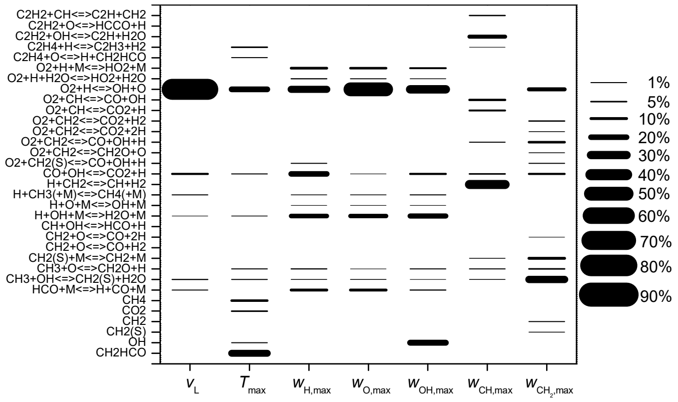
% contribution of parameter uncertainties to the uncertainty of the simulated results



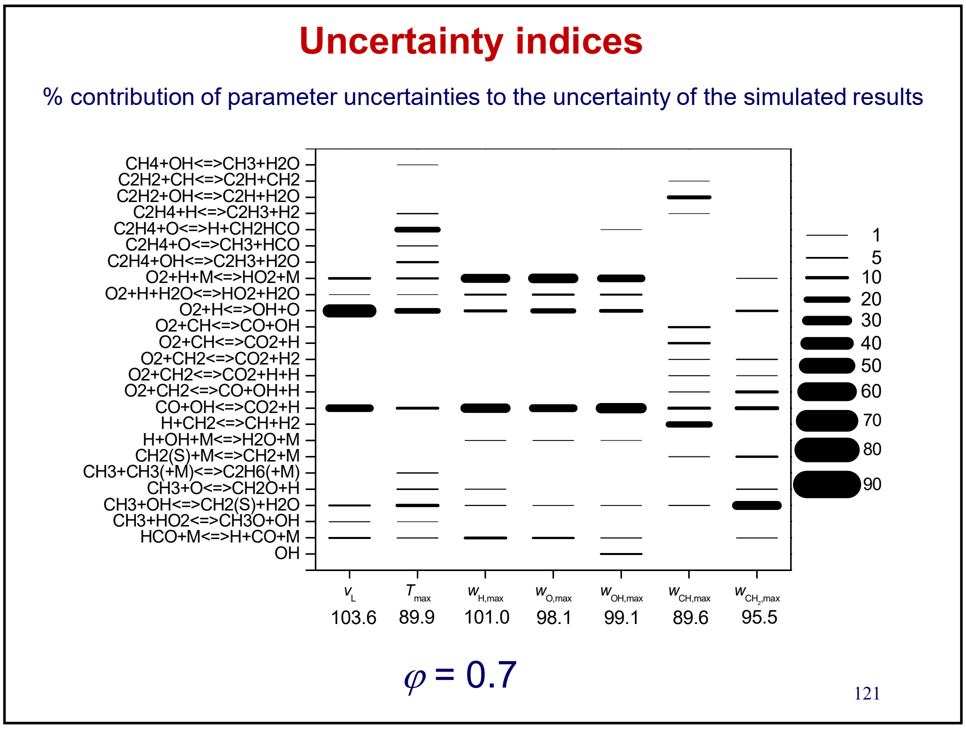
$\varphi = 1.0$

Uncertainty indices

% contribution of parameter uncertainties to the uncertainty of the simulated results



$\varphi = 1.0$

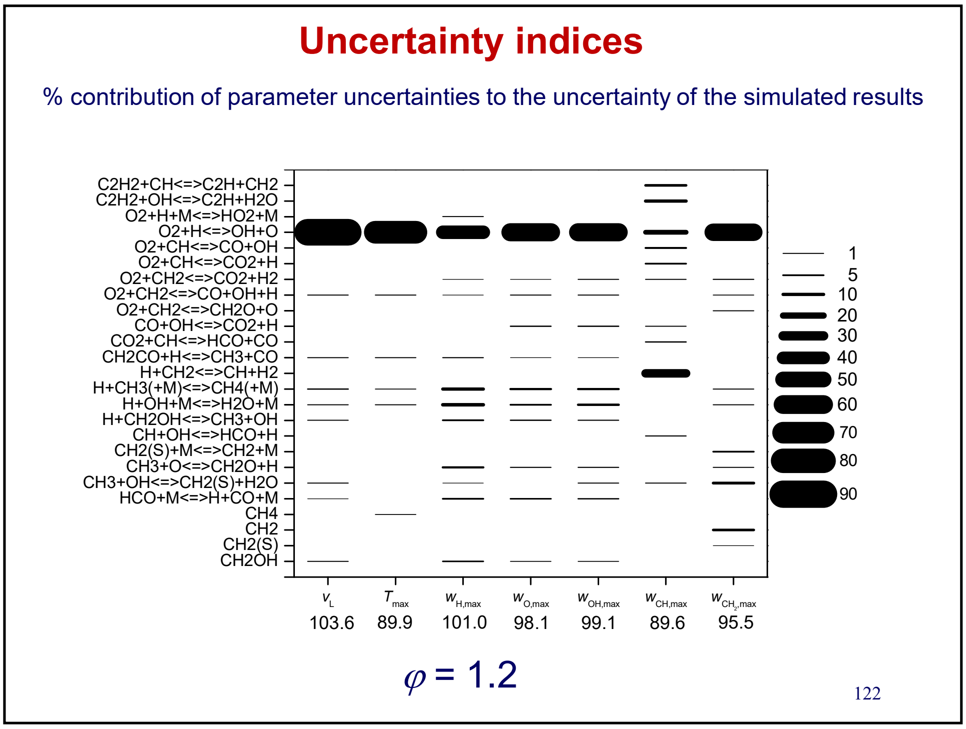


Uncertainty indices

% contribution of parameter uncertainties to the uncertainty of the simulated results

$\varphi = 0.7$

121



Uncertainty indices

% contribution of parameter uncertainties to the uncertainty of the simulated results

$\varphi = 1.2$

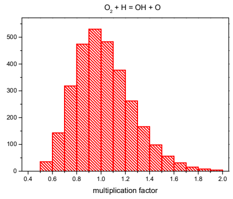
122

Assumed probability density functions of kinetic and thermodynamic parameters

The Monte Carlo and the sensitivity index methods require an assumption on the probability density functions (*pdf*s) of parameters

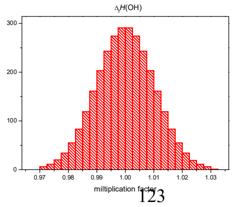
Rate coefficients:

- log-normal distribution
- σ_j was calculated from the f_j uncertainty factor
- the log-normal distribution is clipped at ±3σ (ln k_j)



Enthalpies of formation:

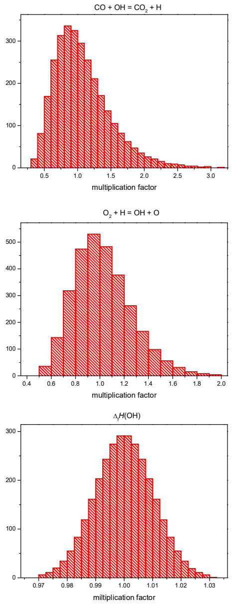
- normal distribution
- σ is assessed on the basis of thermodynamic tables
- the normal distribution is clipped at ±3σ



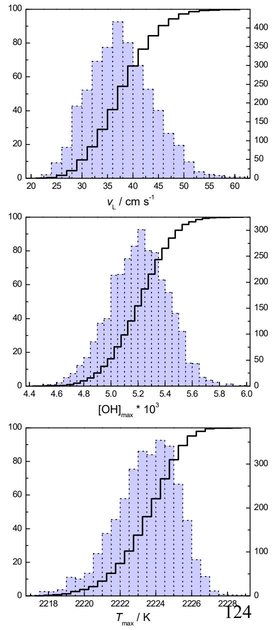
---

pdf of parameters

pdf of simulation results



3,000 Monte Carlo simulations



Comparison of the results of local and global (Monte Carlo) uncertainty analyses for a stoichiometric, stationary, flat methane-air flame

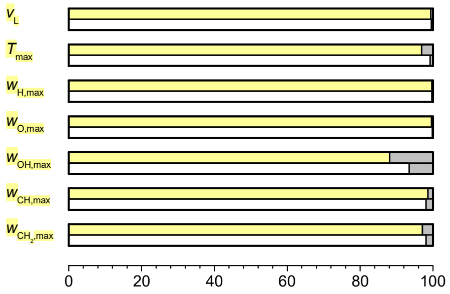
	result	calculated variances from	
		local	Monte Carlo
		uncertainty analyses	
flame velocity	38.1 cm/s	4.6 cm/s	6.2 cm/s
max. T	2224.2 K	2.8 K	1.7 K
max. w_H	2.14×10^{-4}	14.7%	12.6%
max. w_O	1.74×10^{-3}	13.3%	10.4%
max. w_{OH}	5.27×10^{-3}	3.6%	4.0%
max. w_{CH}	8.07×10^{-7}	46.3%	49.2%
max. w_{CH2}	2.54×10^{-5}	23.8%	24.0%

Largest and smallest results that can be achieved with any parameter combination, selected from the domain of uncertainty of the parameters

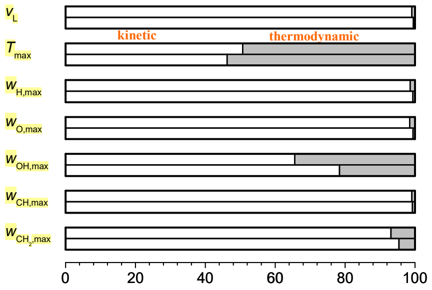
	nominal simulation result	minimal	maximal
		achievable result	
flame velocity	38.1 cm/s	21.3 cm/s	61.6 cm/s
max. T	2224.2 K	2217.4 K	2228.6 K
max. w_H	2.14×10^{-4}	63.1%	144.4%
max. w_O	1.74×10^{-3}	66.9%	136.1%
max. w_{OH}	5.27×10^{-3}	86.4%	114.8%
max. w_{CH}	8.07×10^{-7}	15.5%	474.6%
max. w_{CH2}	2.54×10^{-5}	37.9%	219.5%

Conclusion:
Physically irrealistic results can be obtained, even if the parameters were randomly selected from the uncertainty ranges recommended by the gas kinetics databases.
Reasons: (1) these uncertainties are based on direct measurements;
(2) correlations of uncertainties are not taken into account.

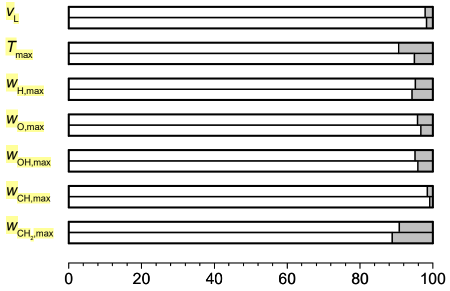
Kinetic *vs.* thermodynamic contributions



ϕ = 0.70



ϕ = 1.00



ϕ = 1.20

upper bar: local results
lower bar: sensitivity index results

Share of thermodynamic uncertainty: 1%-12%. except for the calculated temperature and w_{OH} at stoichiometric conditions

---

Global uncertainty analysis of local sensitivity coefficients

The calculated local sensitivity coefficients belong to a nominal parameter set
different parameter sets ⇒ different local sensitivity coefficients
(this is why these are called local ...)

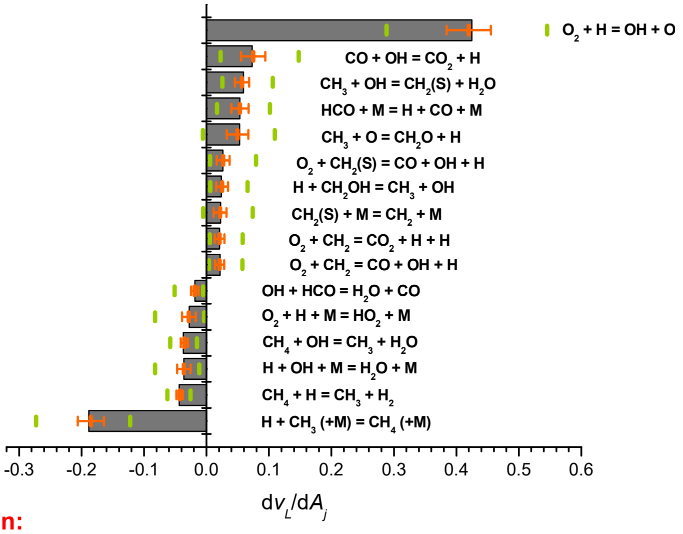
local sensitivity coefficients of flame velocity
were calculated at 3000 randomly selected parameter sets

J. Zádor, I. Gy. Zsély, T. Turányi, M. Ratto, S. Tarantola, A. Saltelli: Local and global uncertainty analyses of a methane flame model, *J. Phys. Chem. A*, **109**, 9795-9807 (2005)

The plotted results:

- local sensitivity coefficients at the nominal set ▭
- standard deviation of sensitivities ▬
- the largest and the smallest calculated sensitivity coefficient ❙ ❙

Global uncertainty analysis of local sensitivity coefficients 2



Conclusion:

There is a surprisingly little variation of the calculated local sensitivity coefficients with the changing parameters. The order of sensitivity coefficients is almost identical in the whole parameter space.

Methane flame uncertainty analysis: general conclusions

Good agreement between the calculated total variances by the local uncertainty analysis and the Monte Carlo method. (surprise)

Good agreement between the importance of parameters assessed by the local uncertainty analysis and the sensitivity indices. (surprise)

Better simulation results can be achieved, if the rate coefficients of a few reactions and the enthalpies of formation of a few species are known better (= with smaller variance)
These represent a small fraction of the total number of species/reactions.

Significant rate coefficients:

$O_2 + H = OH + O$
$O_2 + H + M = HO_2 + M$
$CO + OH = CO_2 + H$
$H + CH_3 + M = CH_4 + M$
$CH_3 + OH = CH_2(S) + H_2O$
$C_2H_2 + OH = C_2H + H_2O$
$C_2H_2 + CH = C_2H + CH_2$
$H + CH_2 = CH + H_2$

Significant enthalpies of formation:

OH
$CH_2(S)$
CH_2
CH_2OH
CH_2CHO

Uncertainty analysis study of the laminar methane flame the points to be corrected

Which were the weak points of
the previously discussed uncertainty analysis study?

„uncertainty of the rate coefficient" = uncertainty of Arrhenius parameter *A* was considered only

What is the uncertainty of each Arrhenius parameter?

The used uncertainty parameters *f* were based on the direct measurements
Considering also the indirect measurements decreases the uncertainty

At the development of detailed reaction mechanisms the direct and indirect experimental results are both considered; the nominal parameter set contains correlations that have to be taken into account.
Considering the parameter correlations is needed

Topic 7: Uncertainty of the Arrhenius parameters

temperature dependence of uncertainty factor *f*,

domain of uncertainty of the Arrhenius parameters,

joint uncertainty of the Arrhenius parameters,

calculation of the covariance matrix of the Arrhenius parameters

determination of the covariance matrix of the Arrhenius parameters from literature measurements

Temperature dependence of uncertainty factor *f*

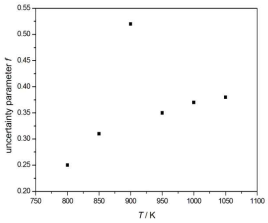
Uncertainty parameter *f* is either constant (Tsang, Warnatz, Konnov)

or defined in temperature regions (Baulch *et al.* evaluations):

$$f(T)=\begin{cases} f_1 & \text{if } T\in(T_1,T_2) \\ f_2 & \text{if } T=T_3 \\ \vdots & \qquad\vdots \end{cases}$$

f_1, f_2, f_3, … corresponds to the actual scatter of measurements in this temperature region.

We will call them f_{original} values.

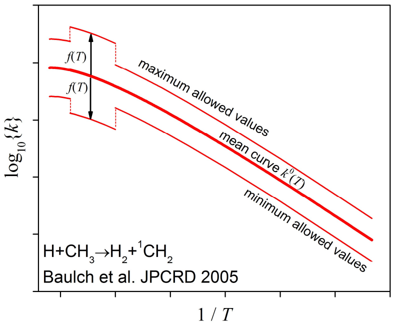


The temperature dependence of the rate coefficents imposes a relation among the uncertainty parameter *f* values at different temperatures.

The f_{original} values are not in accordance with the temperature dependence of the rate coefficient *k*

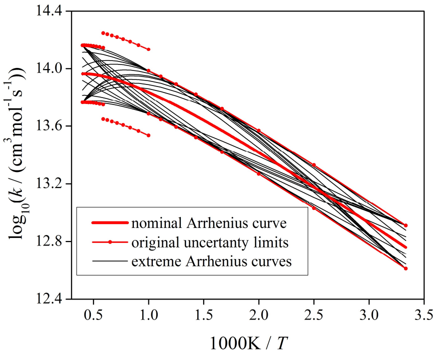
Getting consistent uncertainty factors *f*

Example from the Baulch *et al.* (2005) evaluation:



Solution:

Drawing the possible extreme Arrhenius curves span the realistic uncertainty limits of the rate coefficient *k*:



The corrected *f* values are called f_{extreme}

Domain of uncertainty of Arrhenius parameters

**The f_{extreme} (T) values define the uncertainty domain
of the rate coefficient k in interval [T_1, T_2]
with the temperature dependence of the rate coefficient k**

The evaluations provide the uncertainty of k,
but the real parameters of the model are Arrhenius parameters A, n, E

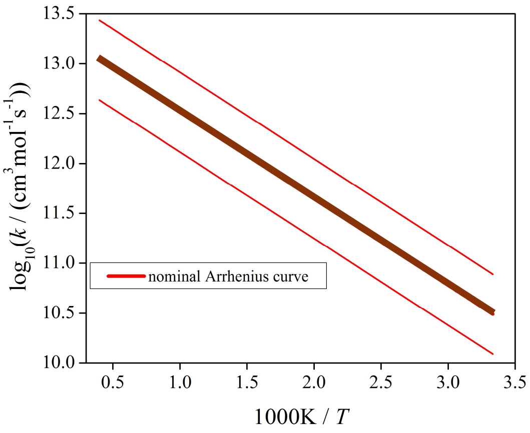
Better to deal with the transformed Arrhenius parameters ln A, n, E/R

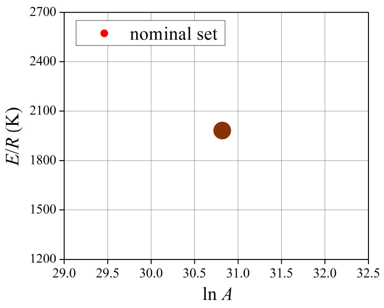
**Statement: the extreme Arrhenius curves span
the domain of uncertainty of the Arrhenius parameters.**

Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of k

two Arrhenius parameters ln A, E/R

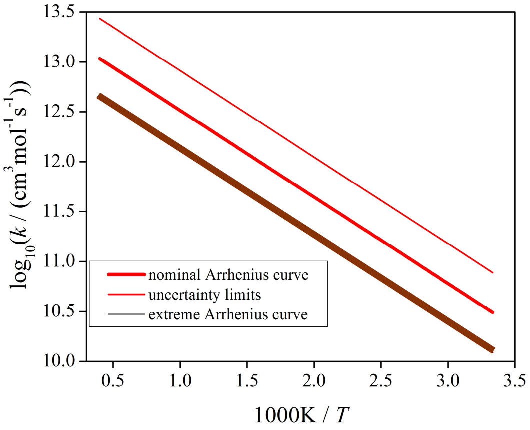


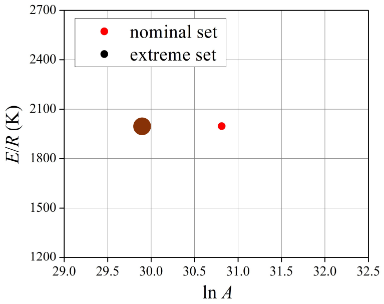


Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of *k*

two Arrhenius parameters ln *A*, *E*/*R*

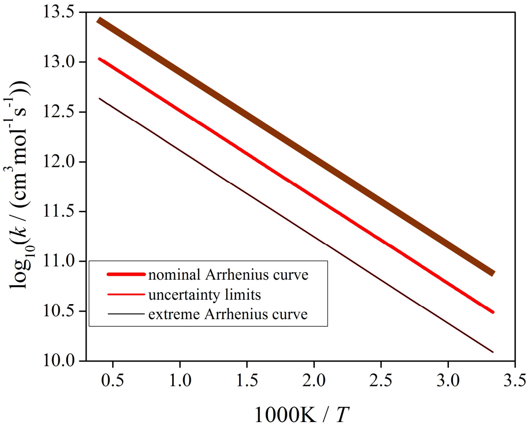


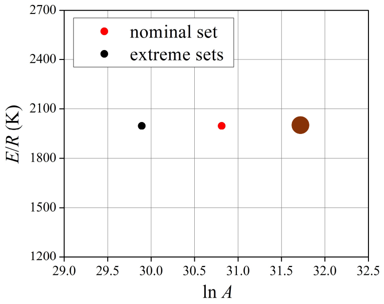


Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of *k*

two Arrhenius parameters ln *A*, *E*/*R*

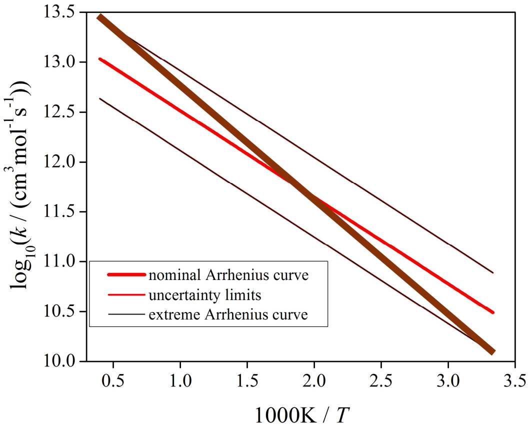


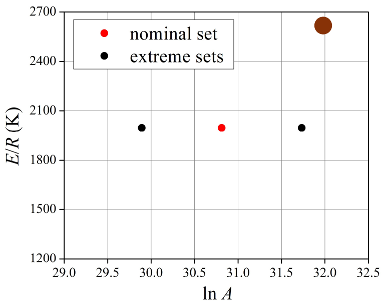


Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of *k*

two Arrhenius parameters ln *A*, *E*/*R*

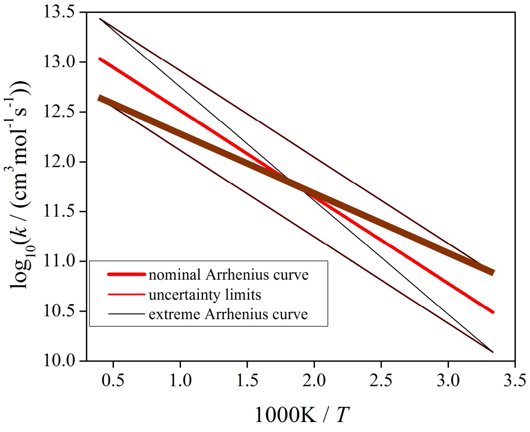


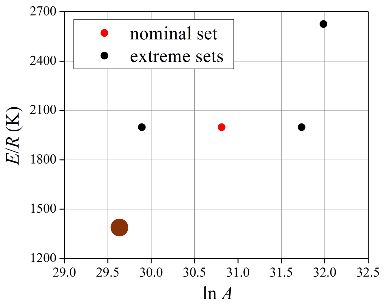


Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of *k*

two Arrhenius parameters ln *A*, *E*/*R*

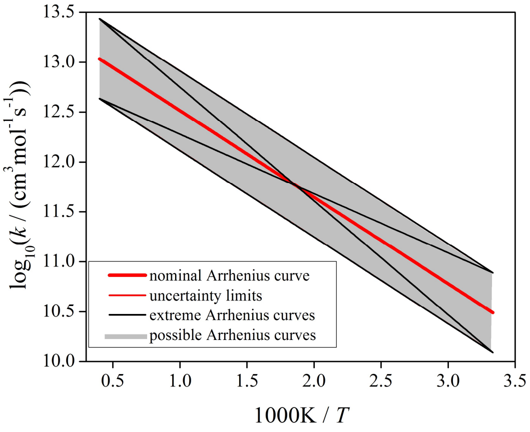


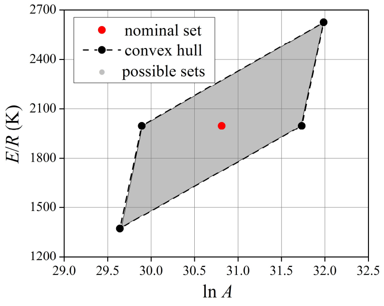


Domain of uncertainty of Arrhenius parameters

The simplest case: temperature independent uncertainty of *k*

two Arrhenius parameters ln *A*, *E*/*R*



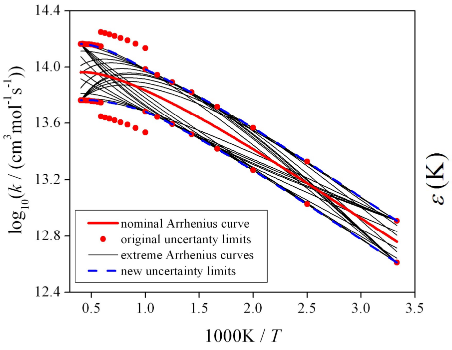


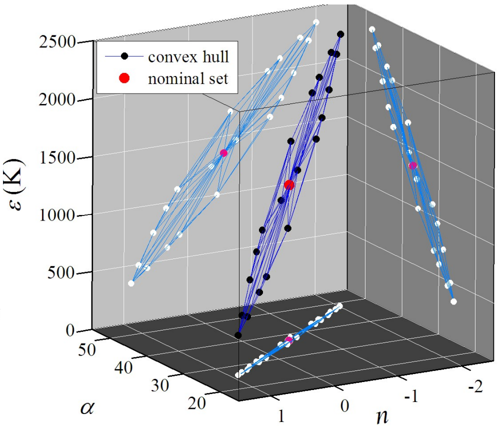
3D domain of Arrhenius parameters

The most complex case: temperature dependent uncertainty of *k*

three Arrhenius parameters α= ln *A*, *n*, ε= *E*/*R*

the extreme Arrhenius curves = vertices of the convex hull





Uncertainty parameter *f*

Definition of uncertainty factor *f*:

$$f(T) = \log_{10}\left(k^0(T)/k^{\min}(T)\right) = \log_{10}\left(k^{\max}(T)/k^0(T)\right)$$

Calculation of the variance of ln *k* from uncertainty factor *f*:

(assuming 3σ deviation between $\log_{10} k^0$ and $\log_{10} k^{\max}$)

$$\sigma(\ln k) = \frac{\ln 10}{3} f$$

Instead of temperature dependent $\sigma(\ln k)$

the covariance matrix of the Arrhenius parameters is needed!

extended Arrhenius expression: $\boxed{k(T) = AT^n \exp(-E/RT)}$

linearized form: $\boxed{\underbrace{\ln\{k(T)\}}_{\kappa(\theta)} = \underbrace{\ln\{A\}}_{\alpha} + \underbrace{n}_{n} \cdot \underbrace{\ln\{T\}}_{\theta} - \underbrace{\{E/R\}}_{\varepsilon} \cdot \underbrace{\{T\}^{-1}}_{\theta}}$

Relation between the σ of the rate coefficient and the covariance matrix of the Arrhenius parameters

Matrix-vector form of the linearized Arrhenius equation: $\boxed{\kappa(\theta) = \mathbf{p}^{\mathrm{T}}\boldsymbol{\theta}}$

$\boxed{\mathbf{p}^{\mathrm{T}} := [\alpha \ \ n \ \ \varepsilon]}$ $\boxed{\boldsymbol{\theta}^{\mathrm{T}} := [1 \ \ \ln\theta \ \ -\theta^{-1}]}$

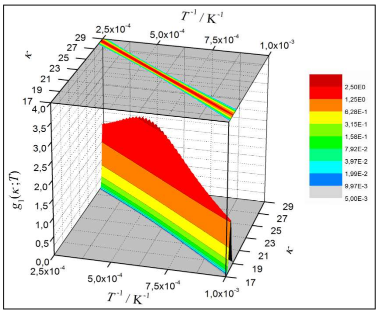
The covariance matrix of the Arrhenius parameters and its relation to the uncertainty of the rate coefficient:

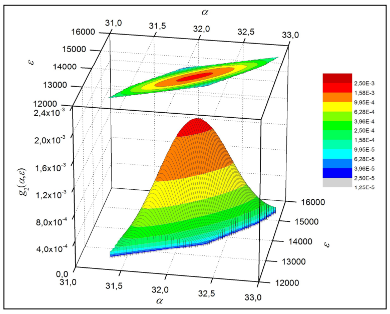
$$\boxed{\boldsymbol{\Sigma}_{\mathbf{p}} = \overline{(\mathbf{p}-\bar{\mathbf{p}})(\mathbf{p}-\bar{\mathbf{p}})^{\mathrm{T}}} = \begin{bmatrix} \sigma_\alpha^2 & r_{\alpha n}\sigma_\alpha\sigma_n & r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon \\ r_{\alpha n}\sigma_\alpha\sigma_n & \sigma_n^2 & r_{n\varepsilon}\sigma_n\sigma_\varepsilon \\ r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon & r_{n\varepsilon}\sigma_n\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix}} \qquad \boxed{\sigma_\kappa(\theta) = \sqrt{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{p}}\boldsymbol{\theta}}}$$

⇒ the temperature dependent standard deviation of *k* can be calculated from a quadratic form.

Nagy, T.; Turányi, T. Uncertainty of Arrhenius parameters
Int. J. Chem. Kinet., **43**, 359-378 (2011)

Example: reaction O + N_2O → NO + NO

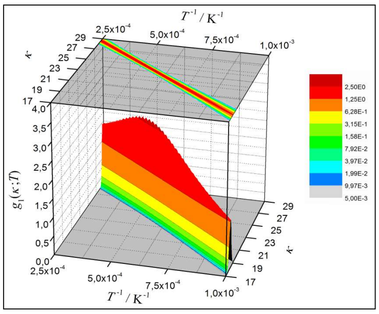


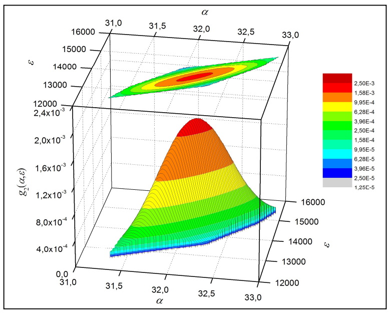


Temperature dependent
1D normal
probability density function (*pdf*)
of the rate coefficient
(1000 K – 4000 K)

Temperature independent
multivariate joint normal
probability density function (*pdf*)
of the Arrhenius parameters

Example: reaction O + N_2O → NO + NO





Temperature dependent
1D normal
probability density function (*pdf*)
of the rate coefficient
(1000 K – 4000 K)

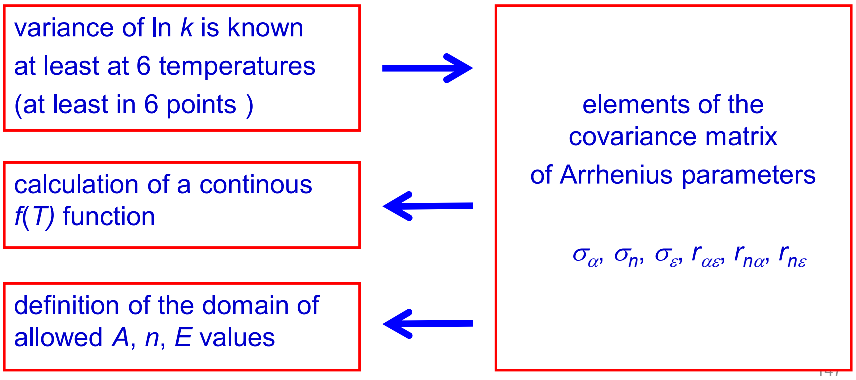
Temperature independent
multivariate joint normal
probability density function (*pdf*)
of the Arrhenius parameters

Calculation of the covariance matrix of the Arrhenius parameters

$$\boxed{\sigma_\kappa(\theta)=\sqrt{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{p}}\boldsymbol{\theta}}}$$

For the 3-parameter Arrhenius equation:

$$\boxed{\sigma_\kappa(\theta)=\sqrt{\sigma_\alpha^2+\sigma_n^2\ln^2\theta+\sigma_\varepsilon^2\theta^{-2}+2r_{\alpha n}\sigma_\alpha\sigma_n\ln\theta-2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon\theta^{-1}-2r_{n\varepsilon}\sigma_n\sigma_\varepsilon\ln\theta\cdot\theta^{-1}}}$$



variance of ln *k* is known
at least at 6 temperatures
(at least in 6 points)

calculation of a continous
f(*T)* function

definition of the domain of
allowed *A*, *n*, *E* values

elements of the
covariance matrix
of Arrhenius parameters

σ_α, σ_n, σ_ε, $r_{\alpha\varepsilon}$, $r_{n\alpha}$, $r_{n\varepsilon}$

Features of the uncertainty parameter *f*

Baulch et al. (2005):

temperature independent *f* (constant *f*(*T*) function) about 50%

OR

a verbally defined f(T) function about 50%

"*f* = 0.1 at 800 K raising to 0.2 at 2000 K "

Other sources:

NIST Chemical Kinetics Database, Tsang, Warnatz, Konnov
temperature independent *f* values.

Good features:

- *f* factors are available for several hundred reactions
- *f* factors are very realistic (to our experience)

Bad features:

- derivation of the *f* parameter is not documented
- temperature dependence is missing or not well defined
  = cannot be used for the calculation of the uncertainty of the Arrhenius parameters

⇒ Reassessment of the uncertainty parameters is needed!

Reassessment of the $f(T)$ functions is needed!

We have written a MATLAB + Fortan code
for the semiautomatic calculation of the $f(T)$ functions.

Major steps for a given elementary reaction:

1 collection of all direct measurements and theoretical calculations
source: NIST Chemical Kientics Database + recent reviews

2 foreward direction: selected (direction with more data)
backward direction: converted to forward direction Arrhenius parameters

3 preparation of a datafile: each line one measurement/calculation
squib + temperature range + Arrhenius parameters

4 selection of a mean line ($\ln k - 1/T$)
in the middle of uncertainty band: almost always Baulch *et al.*, 2005

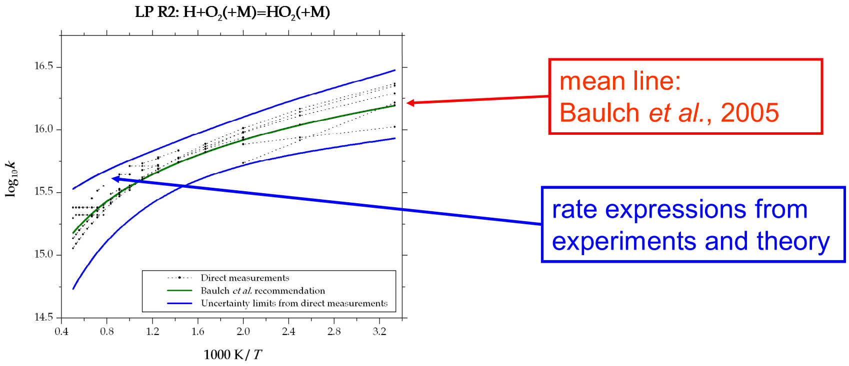
5 MATLAB code: interactive elimination of outliers

6 automatic calculation of „empirical” *f* points at several temperatures
fitting the elements of the covariance matrix to these points (Fortran code)
plotting the experimental/theoretical results + the recalculated $f(T)$

Example: reaction $H + O_2 + M = HO_2 + M$

We found about 60 experimental/theoretical rate expressions.
After the selection remained:
Ar bath gas: 9 experimentally determined and
1 theoretically calculated rate coefficient expressions
N_2 bath gas: 10 experimentally determined and
2 theoretically calculated rate coefficient expressions

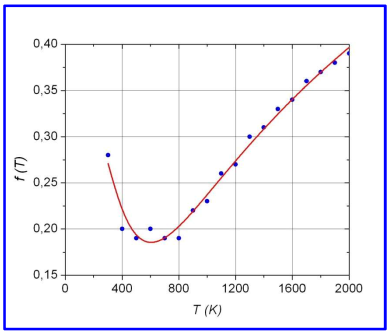
used together assuming m=0.5 (relative collision efficiency Ar to N_2)

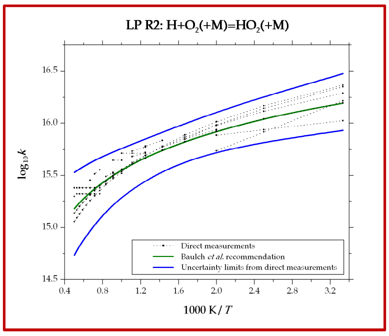


mean line:
Baulch *et al.*, 2005

rate expressions from
experiments and theory

Example: reaction $H + O_2 + M = HO_2 + M$





blue dots:
distance from the extreme
experimental/theoretical values
(„empirical *f* points”)

red line:
calculated from the fitted
elements of the covariance matrix
⇒ *a priori* $f(T)$ function

blue line:
a priori
temperature dependent
k^{min} and k^{min} bounds
calculated from the
covariance matrix
of Arrhenius parameters

Topic 8: Mechanism optimisation and determination of the posterior parameter uncertainties

steps of mechanism optimisation,

relations between the following uncertainty domains:

prior uncertainty of the input parameters,
uncertainty of the model results calculated from
the prior uncertainty of the input parameters,
uncertainty of model results measured by indirect experiments,
prior uncertainty of the input parameters,
uncertainty of the model results calculated from
the posterior uncertainty of the input parameters;

results of mechanism optimization

Direct and indirect measurements

Direct measurements:
- determination of the rate coefficient of a single elementary reaction at a given temperature, pressure, and bath gas
- the rate coefficient values are published

Theoretical (direct) determinations:
- TST/master equation calculations
- the rate coefficients are published at given T, p
- parameterised T, p dependence of rate coefficient k

Indirect measurements:
- a property of the whole combustion system is measured
- interpretation is based on a detailed mechanism
- e.g. laminar flame velocities, ignition delays, concentration profiles

T. Turányi, T. Nagy, I. Gy. Zsély, M. Cserháti, T. Varga, B.T. Szabó,
I. Sedyó, P. T. Kiss, A. Zempléni, H. J. Curran:
Determination of rate parameters based on both direct and indirect measurements
Int.J.Chem.Kinet., **44**, 284–302 (2012)

Uncertainty of reaction rate parameters

Reaction rate parameters:
Arrhenius parameters ***A*, *n*, *E*** ,
3rd body collision efficiencies,
(parameters of presure dependence: Lindemann and Troe parameters)
(enthalpies of formation)

a priori **uncertainty of reaction rate parameters:**
uncertainty of reaction rate parameters, deduced from available direct measurement data and theoretical calculations

a posteriori **uncertainty of reaction rate parameters:**
uncertainty of reaction rate parameters, deduced from fitting to direct measurement data + theoretical calculations results + indirect measurement data

The Michael Frenklach approach („GRI-Mech method”)

1. Starting from a carefully evaluated mechanism.
   The Arrhenius parameters are based on the latest direct measurements.
2. A set of „targets” are collected. The targets are based on the results of indirect measurements (*e.g.* flame velocity, time-to-ignition)
3. The most important reactions at the experimental conditions of the targets are identified by sensitivity analysis. The *A*-factors of these reactions are called „active parameters”.
4. Limiting A_{min} and A_{max} values are identified for each reaction.
   These limits are usually based on collections of evaluated data, using uncertainty parameters *f*.
5. The active *A* factors are fitted within the limits to give a good least squares approximation to the targets.

⇒ **Good agreement for most targets**

The further developments of the GRI method

1. The objective function was extended:
   penalized if the fitted *A* factor deviates from the evaluated value
   - You, X.; Russi, T.; Packard, A.; Frenklach, M.
     Proc.Combust.Inst., **33**, 509-516 (2011)
   - Sheen, D.; Wang, H.
     Combust. Flame, **158**, 645-656(2011)

2. Model optimisation as a **basis of Data Collaboration**

many Frenklach *et al.* papers, including:

Russi, T.; Packard, A.; Feeley, R.; Frenklach, M.
J. Phys. Chem. A, **112**, 2579-2588(2008).

Applications and extensions by Hai Wang

1. Application of the method for a series of the subsystems of the USC Mech/JetSurF mechanism

propane	Z. Qin, V. Lissianski, H. Yang, W.C. Gardiner, S. Davis, H. Wang *Proc.Combust. Inst.,* 2000, 28, 1663-1669.
H_2/CO	S. Davis, A. Joshi, H. Wang, F. Egolfopoulos *Proc.Combust.Inst.,* 2005, 30, 1283-1292.
ethylene	D.A. Sheen, X. You, H. Wang, T. Løvås, *Proc. Combust. Inst.* 2009, 32, 535-542.
n-heptane	D.A. Sheen, H. Wang, *Combust. Flame*, 2011, 158, 645-656.

2. Method of Uncertainty Analysis using Polynomial Chaos Expansions (MUM-PCE)

Sheen, D. A.; You, X.; Wang, H.; Løvås, T.
Proc. Combust. Inst., 32, 535-542(2009).

3. Fitting to many original experimental datapoints
(not only fitting to selected targets)

Sheen, D.; Wang, H.
Combust. Flame, **158**, 645-656(2011)

Method of Uncertainty quantification and Minimization using Polynomial Chaos Expansions:
MUM-PCE method of David Sheen and Hai Wang

The approximation of the kinetic model (called „response surface” or „surrogate model”) and the calculation of uncertainty propagation is done in a single step.

Scaling the uncertainty region of a parameter to interval [-1, +1]

$$x_i = \frac{\ln k_i / k^0}{\ln f_i}$$

k_i	actual value of the rate coefficient
k^0	nominal value of the rate coefficient
f	uncertainty parameter

$x_i = 0$	*k* is at the nominal value
$x_i = -1$	*k* is at the lower bound
$x_i = +1$	*k* is at the upper bound

MUM-PCE 2

The approximation of the solution of the kinetic model (*e.g.* a calculated flame velocity) as a function of parameters
within their range of uncertainty, using a 2nd order polynomial:

$$\eta_r(\mathbf{x}) = \eta_{r,0} + \sum_{i=1}^{m} a_{r,i} x_i + \sum_{i=1}^{m}\sum_{j\ge i}^{m} b_{r,i,j} x_i x_j$$

Consider x_i as a random variable (denoted by ξ_i)
with unit normal distribution.

The model result will also be a random variable, calculated by

$$\eta_r(\boldsymbol{\xi}) = \eta_r(\mathbf{x}^{(0)}) + \sum_{i=1}^{m} \hat{\boldsymbol{\alpha}}_{r,i}\xi_i + \sum_{i=1}^{m}\sum_{j\ge i}^{m} \hat{\boldsymbol{\beta}}_{r,ij}\xi_i\xi_j + \ldots,$$

The parameters of this expansion can be calculated from the coefficients of the response surface polynomial:

$$\hat{\boldsymbol{\alpha}}_r = {}^1\!/_2\,\mathbf{I}_M \mathbf{a}_r \qquad \hat{\boldsymbol{\beta}}_r = {}^1\!/_4\,\mathbf{I}_M^T \mathbf{b}_r \mathbf{I}_M$$

MUM-PCE 3

The new coefficients provide the variance of the fitted parameters:

$$\sigma_r(\boldsymbol{\xi})^2 = \sum_{i=1}^{m} \hat{\boldsymbol{\alpha}}^2{}_{r,i} + 2\sum_{i=1}^{m} \hat{\boldsymbol{\beta}}^2{}_{r,ij} + \sum_{i=1}^{m}\sum_{j>i}^{m} \hat{\boldsymbol{\beta}}^2{}_{r,ij}$$

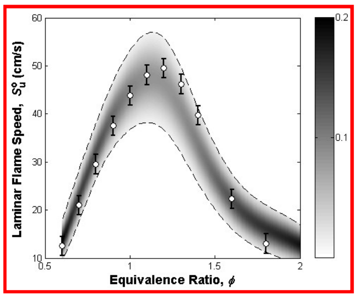
a priori uncertainty of the rate coefficients:
uncertainty as we know it **before the optimisation**.

a posteriori uncertainty of the rate coefficients:
much smaller uncertainty, obtained
as a result of optimisation.

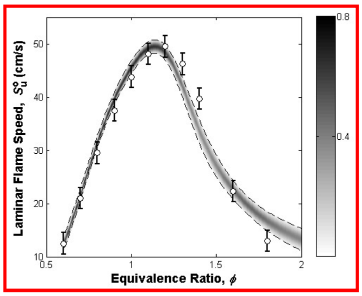
Uncertainty of simulation results before and after the optimisation

experimental uncertainty
of flame velocity measurements

before the optimisation,
the solution of the kinetic model
has larger uncertainty
(grey shading shows the *pdf*)
than the experimental data



after mechanism optimisation:
- new parameter set
- better known parameters
  with lower uncertainty
- the solution of the model is
  different now
- the solution has lower uncertainty



Mechanism optimization at ELTE Chemical Kinetics Laboratory

1 all indirect measurement data should be collected
that are applicable for testing a mechanism.

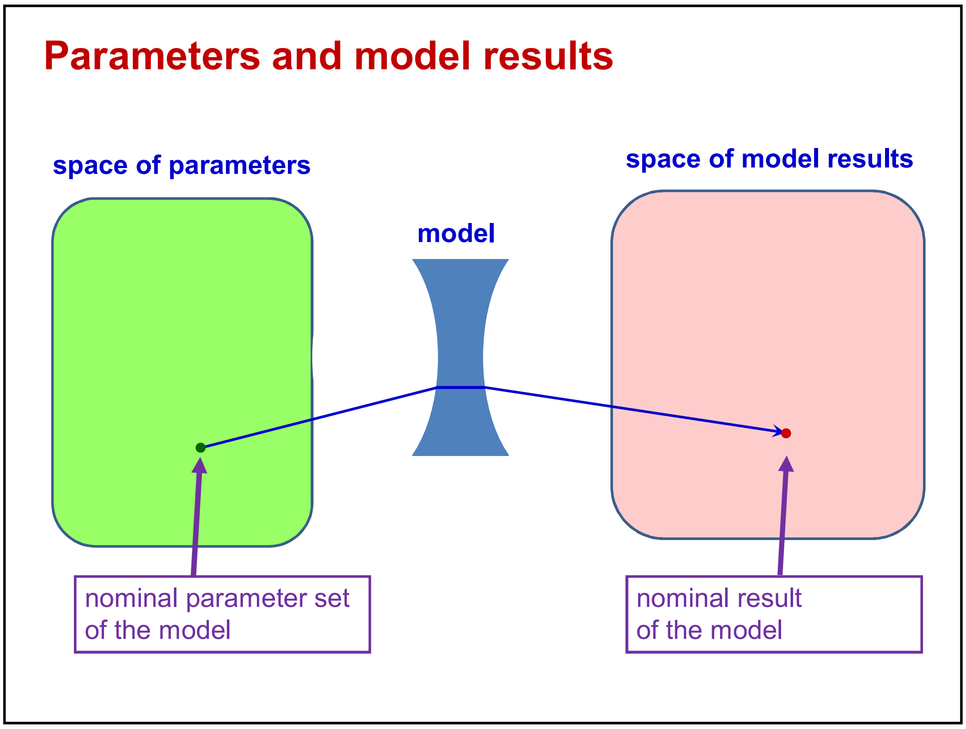
2 sensitivity analysis for finding the important reaction steps
(simulated data points wiith respect to the rate parameters)
the rate parameters of these reactions will be optimised

3 determination of the *a priori* uncertainty of the rate parameters
(= determination of the domain of allowed parameter values)

4 all reliable direct measurement data related to the important reactions
are collected

5 global parameter optimisation
considering both the indirect and direct measurement data
⇒ new rate parameters with physical meaning
⇒ ***a posteriori* uncertainty domain of rate parameters**

Turányi T, Nagy T, Zsély IGy, Cserháti M, Varga T, Szabó B,
Sedyó I, Kiss P, Zempléni A, Curran H J
Determination of rate parameters based on both direct and indirect measurements.
Int. J. Chem. Kinet. **44**, 284–302 (2012)



Parameters and model results

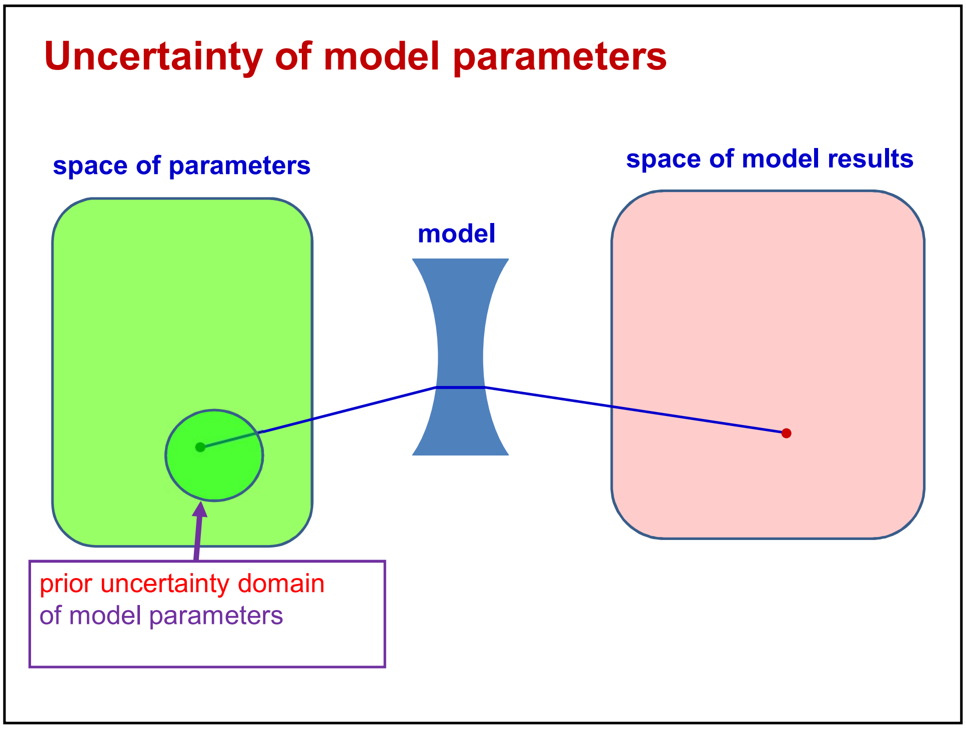
space of parameters

model

space of model results

nominal parameter set
of the model

nominal result
of the model



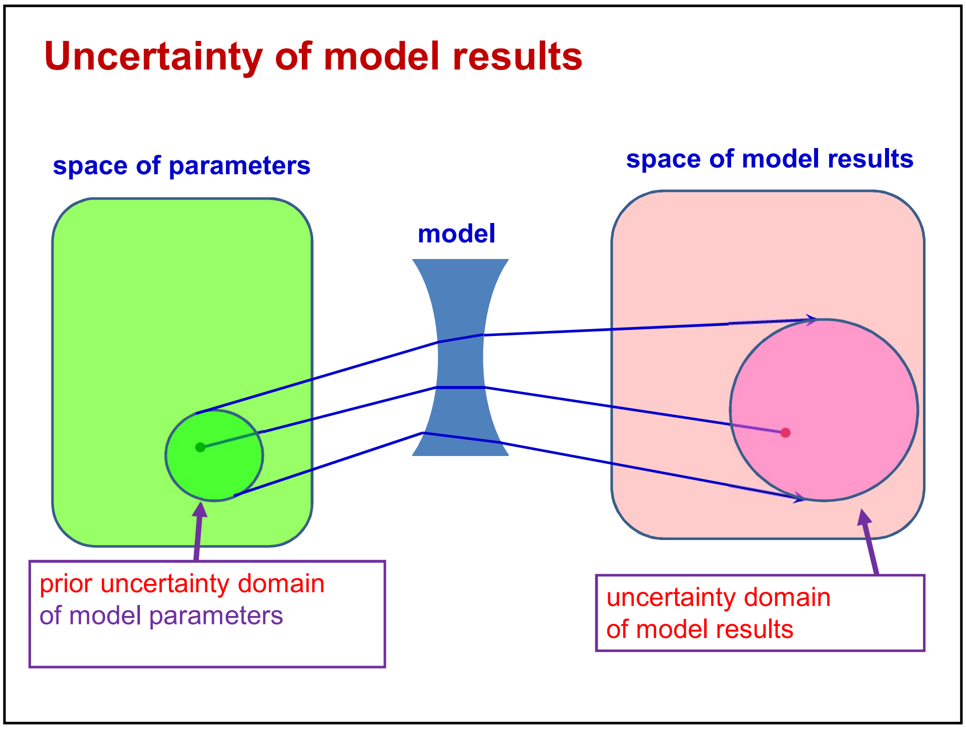
Uncertainty of model parameters

space of parameters

model

space of model results

prior uncertainty domain
of model parameters



Uncertainty of model results

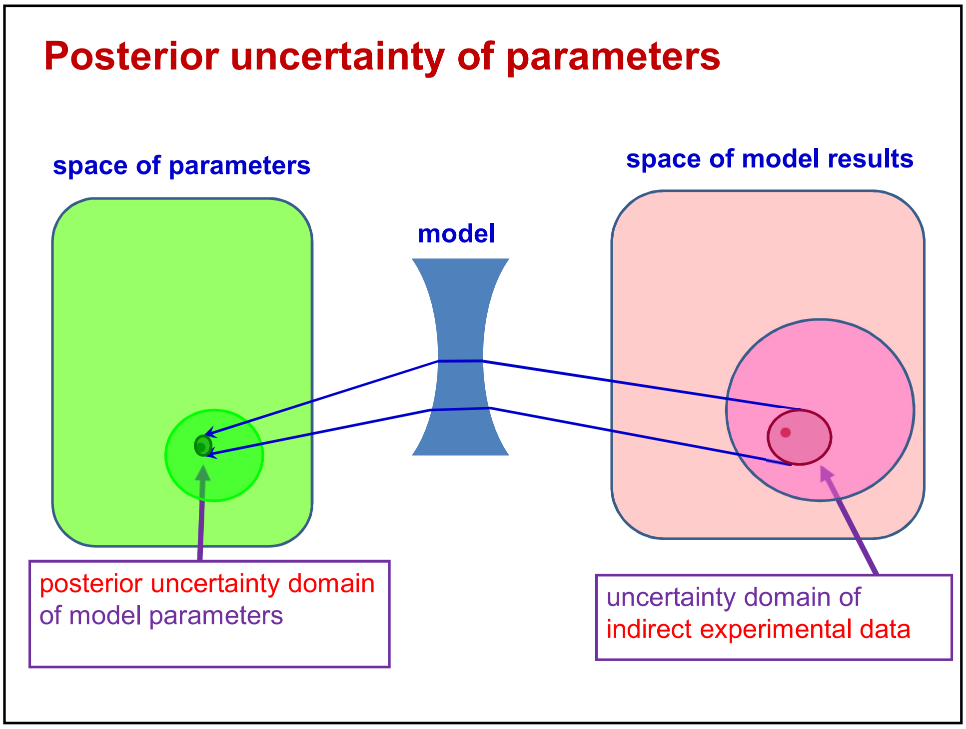
space of parameters

model

space of model results

prior uncertainty domain
of model parameters

uncertainty domain
of model results



Posterior uncertainty of parameters

space of parameters

model

space of model results

posterior uncertainty domain
of model parameters

uncertainty domain of
indirect experimental data

Optimisation and uncertainty calculation

Optimisation - minimisation of this error function by fitting the parameters **within their prior uncertainty domain**

$$\boxed{E(\mathbf{p}) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(\frac{Y_{ij}^{mod}(\mathbf{p}) - Y_{ij}^{exp}}{\sigma(Y_{ij}^{exp})}\right)^2}$$

$$Y_{\mathrm{ij}} = \begin{cases} y_{\mathrm{ij}} & \text{if } \sigma(y_{\mathrm{ij}}^{\mathrm{exp}}) \approx \text{constant} \\ \ln y_{\mathrm{ij}} & \text{if } \sigma(\ln y_{\mathrm{ij}}^{\mathrm{exp}}) \approx \text{constant} \end{cases}$$

y_{ij}	measured/calculated rate coefficient OR measured/calculated ignition time/flame velocity in data point *j* of data series *i*
σ	standard deviation of the measured data
N_i	number of data points in data series *i*
N	number of data series (different experiments)

Calculation of the covariance matrix of the estimated parameters:

Covariance matrix of experiments

Discrepancy between experiments and model

$$\boxed{\mathbf{\Sigma_p} = \left[\left(\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{J}_{\mathrm{o}}\right)^{-1}\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\right]\left(\mathbf{\Sigma}_{\mathbf{Y}} + \mathbf{\Sigma}_{\Delta}\right)\left[\left(\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{J}_{\mathrm{o}}\right)^{-1}\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\right]^{\mathrm{T}}}$$

Results of optimisation

- **Optimised combustion model**
    - Better overall performance than any previously published model
- **Set of optimised rate parameters**
    - Optimized Arrhenius parameters
    - Optimized third body collision efficiency parameters
    - Optimized enthalpies of formation
- **Posterior covariance matrix of the optimised parameters**
    - Temperature independent
    - Uncertainty (i.e. estimated scatter) of each optimised parameter
    - Correlation coefficients between the parameter pairs

Example: optimization of a methanol and formaldehyde combustion mechanism

- **Methanol** is an alternative automotive fuel, fuel additive and feedstock in various industrial processes
- **Model system** for studies of **C_1 combustion**: important radicals include CH_2OH and CH_3O
- **Relevance** for the oxidation of higher hydrocarbons/ oxygenates
- **Not all** experimentally observed **combustion characteristics** (*e.g.* ignition, flame propagation, speciation profiles) are **well-described** by available kinetic mechanisms

Data collection of methanol combustion

- **Indirect measurements**
- **Direct measure-ments** of rate coefficients (926 data points/ 66 data sets)
- **Theoretical rate determinations** (33 data sets)

Type of measurement experimental facility	**Data sets**	**Data points**	***p* / atm**	***T* / K**
Ignition delay times	**81**	**574**		
Shock tube	67	421	0.3–51.7	963–2180
Shock tube (CH_2O)	7	99	1.6	1363–2304
Rapid compression machine	7	54	9.3–40.6	817–980
Burning velocity measurements	**87**	**632**		
Outwardly/ spherically propagating flame	35	170	0.5–9.9	298–500
Counterflow twin-flame	5	90	1	298–368
Heat flux method / laminar flat flame	41	280	0.2–1	298–358
Flame-cone method	6	92	1	298–413
Concentration measurements	**97**	**23,694**		
Flow reactor concentration–time profiles	18	1,452	1–20	752–1043
Flow reactor conc.–time profiles (CH_2O)	13	462	1–6	852–1095
Flow reactor outlet concentrations	13	444	1–98.7	600–1443
Flow reactor outlet concentrations (CH_2O)	3	156	1.05	712–1279
Jet-stirred reactor outlet concentrations	9	711	1–20	697–1200
Shock tube concentration–time profiles	14	12,756	0.3–2.5	1266–2100
Shock tube conc.–time profiles (CH_2O)	27	7,713	1.5–2.0	1244–1907

Initial mechanism for optimization

Starting point: $CH_3OH/CH_2O/CO$ mechanism of Li *et al.* (2007)

J. Li; Z. W. Zhao; A. Kazakov; M. Chaos; F. L. Dryer; J. J. Scire Jr.,
Int. J. Chem. Kinet. 39 (2007) 109–136

Update of rate coefficients in the H_2/CO sub-mechanism

using values from our recently optimized joint hydrogen and syngas mechanism

T. Varga; C. Olm; T. Nagy; I. Gy. Zsély; É. Valkó; R. Pálvölgyi; H.J. Curran; T. Turányi, *Int. J. Chem. Kinet.* **48** (2016) 407–422

Further modifications:

- Thermochemistry updated

  S. M. Burke; J. M. Simmie; H. J. Curran, *J. Phys. Chem. Ref. Data* 44 (2015) 013101

  Burcat database (March 13, 2015)

- $CH_3OH+H\dot{O}_2 = CH_3\dot{O}+H_2O_2$ abstraction channel added

  S. J. Klippenstein; L. B. Harding; M. J. Davis; A. S. Tomlin; R. T. Skodje, *Proc. Combust. Inst.* 33 (2011) 351–357

Local sensitivity analysis

Important reactions in formaldehyde and methanol combustion

Brute force method (parameters varied by +5%) using the initial mechanism, shown: ($|S|_{avg} \geq 0.1$)

CH_3OH shock tube IDTs

🟩	$CH_3OH+H\dot{O}_2 = \dot{C}H_2OH+H_2O_2$	0.80
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.54
▨	$\dot{H}+O_2 = \ddot{O}+\dot{O}H$	0.41
🟩	$\dot{O}H+\dot{C}H_3\ (+M) = CH_3OH(+M)$	0.26
🟩	$CH_3OH+\dot{O}H = CH_3\dot{O}+H_2O$	0.24
🟩	$CH_3OH+\dot{H} = \dot{C}H_2OH+H_2$	0.20
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.18
▨	$H\dot{O}_2+\dot{O}H = H_2O+O_2$	0.17
▨	$\dot{H}+H\dot{O}_2 = H_2+O_2$	0.14
▨	$H\dot{C}O\ (+M) = \dot{H}+CO\ (+M)$	0.12
🟩	$\dot{O}H+\dot{C}H_3(+M) = CH_3OH(+M)$	0.11
🟩	$CH_3OH+O_2 = \dot{C}H_2OH+H\dot{O}_2$	0.10
🟩	$CH_3OH+\dot{O}H = \dot{C}H_2OH+H_2O$	0.10

CH_2O shock tube IDTs

▨	$H\dot{C}O(+M) = \dot{H}+CO(+M)$	0.86
	$CH_2O+H\dot{O}_2 = H\dot{C}O+H_2O_2$	0.53
🟩	$CH_2O+O_2 = H\dot{C}O+H\dot{O}_2$	0.43
🟩	$CH_2O+\dot{H} = H\dot{C}O+H_2$	0.42
▨	$\dot{H}+O_2 = \ddot{O}+\dot{O}H$	0.41
	$H\dot{C}O+O_2 = CO+H\dot{O}_2$	0.40
	$CH_2O+M = H\dot{C}O+\dot{H}+M$	0.32
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.23
▨	$H\dot{C}O+\dot{H} = CO+H_2$	0.18
🟩	$CH_2O+\dot{O}H = H\dot{C}O+H_2O$	0.16
	$\dot{C}H_2OH+H\dot{C}O = CH_2O+CH_2O$	0.15
🟩	$CH_2O+M = CO+H_2+M$	0.13
	$CH_2O+\ddot{O} = H\dot{C}O+\dot{O}H$	0.12
	$CH_2O+\dot{C}H_3 = H\dot{C}O+CH_4$	0.10
▨	$\dot{H}+H\dot{O}_2 = H_2+O_2$	0.10
	$\dot{C}H_2OH+M = CH_2O+\dot{H}+M$	0.10

CH_3OH RCM IDTs

🟩	$CH_3OH+H\dot{O}_2 = \dot{C}H_2OH+H_2O_2$	1.00
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.55
	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.31
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.28
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.19
	$CH_2O+H\dot{O}_2 = H\dot{C}O+H_2O_2$	0.13
🟩	$CH_3OH+O_2 = \dot{C}H_2OH+H\dot{O}_2$	0.10

CH_3OH burning velocities

▨	$\dot{H}+O_2 = \ddot{O}+\dot{O}H$	0.99
▨	$H\dot{C}O(+M) = \dot{H}+CO(+M)$	0.60
▨	$CO+\dot{O}H = CO_2+\dot{H}$	0.52
▨	$H\dot{O}_2+\dot{H} =\dot{O}H+\dot{O}H$	0.38
🟩	$CH_3OH+\dot{O}H = CH_3\dot{O}+H_2O$	0.35
▨	$H\dot{O}_2+\dot{O}H = H_2O+O_2$	0.32
▨	$\dot{H}+H\dot{O}_2 = H_2+O_2$	0.30
	$H\dot{C}O+O_2 = CO+H\dot{O}_2$	0.29
▨	$\dot{H}+\dot{O}H+M = H_2O+M$	0.23
▨	$\dot{H}+O_2(+M) = H\dot{O}_2(+M)$	0.23
	$H\dot{C}O+\dot{O}H = CO+H_2O$	0.18
🟩	$CH_3\dot{O}+M = CH_2O+\dot{H}+M$	0.18
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.17
🟩	$CH_3OH+\dot{H} = \dot{C}H_2OH+H_2$	0.16
▨	$\dot{O}H+H_2 = \dot{H}+H_2O$	0.15
▨	$H\dot{C}O+\dot{H} = CO+H_2$	0.14

CH_3OH speciation data

🟩	$CH_3OH+H\dot{O}_2 = \dot{C}H_2OH+H_2O_2$	0.36
🟩	$CH_3OH+\dot{O}H = CH_3\dot{O}+H_2O$	0.29
🟩	$CH_3OH+O_2 = \dot{C}H_2OH+H\dot{O}_2$	0.25
🟩	$CH_3OH+\dot{O}H = \dot{C}H_2OH+H_2O$	0.25
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.23
▨	$H\dot{O}_2+\dot{O}H = H_2O+O_2$	0.23
🟩	$CH_2O+\dot{O}H = H\dot{C}O+H_2O$	0.21
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.20
▨	$\dot{H}+O_2 = \ddot{O}+\dot{O}H$	0.20
▨	$CO+\dot{O}H = CO_2+\dot{H}$	0.20
🟩	$CH_3\dot{O}+M = CH_2O+\dot{H}+M$	0.20
🟩	$CH_3OH+\dot{H} = \dot{C}H_2OH+H_2$	0.19
	$CH_2O+H\dot{O}_2 = H\dot{C}O+H_2O_2$	0.19
🟩	$\dot{C}H_3+H\dot{O}_2 = CH_4+O_2$	0.18
▨	$\dot{H}+O_2(+M) = H\dot{O}_2(+M)$	0.18
	$\dot{C}H_2OH+M = CH_2O+\dot{H}+M$	0.17
🟩	$CH_2O+\dot{H} = H\dot{C}O+H_2$	0.16
▨	$H\dot{C}O(+M) = \dot{H}+CO(+M)$	0.16
🟩	$\dot{C}H_3+H\dot{O}_2 = CH_3\dot{O}+\dot{O}H$	0.15
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.15
🟩	$\dot{O}H+\dot{C}H_3(+M) = CH_3OH(+M)$	0.15

CH_2O speciation data

▨	$H\dot{C}O(+M) = \dot{H}+CO(+M)$	0.46
🟩	$CH_2O+\dot{H} = H\dot{C}O+H_2$	0.38
🟩	$CH_2O+O_2 = H\dot{C}O+H\dot{O}_2$	0.34
	$H\dot{C}O+O_2 = CO+H\dot{O}_2$	0.30
	$CH_2O+H\dot{O}_2 = H\dot{C}O+H_2O_2$	0.30
🟩	$CH_2O+M = CO+H_2+M$	0.27
	$\dot{C}H_2OH+H\dot{C}O = CH_2O+CH_2O$	0.27
	$\dot{C}H_2OH+M = CH_2O+\dot{H}+M$	0.19
	$CH_3OH+H\dot{C}O = \dot{C}H_2OH+CH_2O$	0.19
▨	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.17
▨	$\dot{H}+O_2 = \ddot{O}+\dot{O}H$	0.17
	$CH_2O+M = H\dot{C}O+\dot{H}+M$	0.17
🟩	$CH_2O+\dot{O}H = H\dot{C}O+H_2O$	0.13
▨	$CO+\dot{O}H = CO_2+\dot{H}$	0.13
	$CH_2O+\ddot{O} = H\dot{C}O+\dot{O}H$	0.12
	$\dot{O}H+\dot{O}H(+M) = H_2O_2(+M)$	0.12
	$H\dot{C}O+H\dot{O}_2 => CO_2+\dot{H}+\dot{O}H$	0.12
▨	$H\dot{O}_2+\dot{H} = \dot{O}H+\dot{O}H$	0.12
▨	$H\dot{O}_2+\dot{O}H = H_2O+O_2$	0.12
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.11
	$H\dot{C}O(+M) = \dot{H}+CO(+M)$	0.11
	$\dot{C}H_2OH+O_2 = CH_2O+H\dot{O}_2$	0.11
▨	$H\dot{C}O+\dot{H} = CO+H_2$	0.11
🟩	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.10

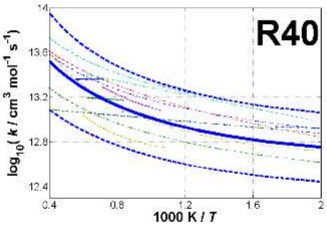
🟩 **Optimized reactions (this work)**
▨ **Already optimized H_2/CO reactions**
☐ Not optimized reactions

Mechanism optimization

- 57 Arrhenius parameters of **17 reactions** optimized

No.	Reaction	f_{prior}
R14/R15	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.30–0.70
R37 LPL	$CH_2O+M = CO+H_2+M$	0.50
R38	$CH_2O+\dot{H} = H\dot{C}O+H_2$	0.60
R40	$CH_2O+\dot{O}H = H\dot{C}O+H_2O$	0.34–0.43
R41	$CH_2O+O_2 = H\dot{C}O+H\dot{O}_2$	1.20
R47	$\dot{C}H_3+H\dot{O}_2 = CH_3\dot{O}+\dot{O}H$	0.46–0.76
R53	$\dot{C}H_3+H\dot{O}_2 = CH_4+O_2$	1.00
R60	$\dot{C}H_2OH+O_2 = CH_2O+H\dot{O}_2$	0.50
R67 LPL	$CH_3\dot{O}+M = CH_2O+\dot{H}+M$	0.84–1.24
R77 HPL	$\dot{O}H+\dot{C}H_3 = CH_3OH$	0.34–0.84
R77 LPL	$\dot{O}H+\dot{C}H_3+M = CH_3OH+M$	1.20
R80	$CH_3OH+\dot{H} = \dot{C}H_2OH+H_2$	0.44–1.07
R81	$CH_3OH+\dot{H} = CH_3\dot{O}+H_2$	1.70
R83	$CH_3OH+\dot{O}H = CH_3\dot{O}+H_2O$	0.70
R84	$CH_3OH+\dot{O}H = \dot{C}H_2OH+H_2O$	0.46–0.87
R85	$CH_3OH+O_2 = \dot{C}H_2OH+H\dot{O}_2$	0.80
R87	$CH_3OH+H\dot{O}_2 = \dot{C}H_2OH+H_2O_2$	1.10
R88	$CH_3OH+H\dot{O}_2 = CH_3\dot{O}+H_2O_2$	0.70

T. Nagy; É. Valkó; I. Sedyó; I. Gy. Zsély; M. J. Pilling; T. Turányi, *Combust. Flame* 162 (2015) 2059–2076



Mechanism optimization 2

- **Optimization targets:**
  - 517 Shock tube, 59 RCM ignition delay points
  - 153 Laminar burning velocity points
  - 2,508 Flow reactor species concentration points
  - 706 Jet-stirred reactor species concentration points
  - 20,460 Shock tube species concentration points
  - 926 Direct measurements of reaction rate coefficients
  - 33 Theoretical determinations of reaction rate coefficients

Polynomial surrogate model (“response surfaces”)
used for computationally expensive **flame simulations**

Hierarchical optimization strategy:
Step-by-step inclusion of reactions and optimization targets

Response surfaces

- **Polynomial approximation** of the simulations (up to 4th order)
- Response surfaces were generated for each flame velocity experiment — **up to 10,000x times faster simulations**
- **20,000 samples** of the active parameters
  - Different parameters at each condition, based on sensitivity analysis
- Accuracy of the polynomials was tested against a new sample, which had not been used for polynomial creation
  - **Maximum allowed deviations:** $\Delta\ \ln(\tau/\mathrm{s})_{\max} = 0.2;\ \Delta\ v_{l\ \max} = 2\ \mathrm{cm/s}$
- About **80% of all data points** could be fitted accurately
  - Only the flame response surfaces were used
    The other simulated data were calculated with direct integration

Mechanism optimization results

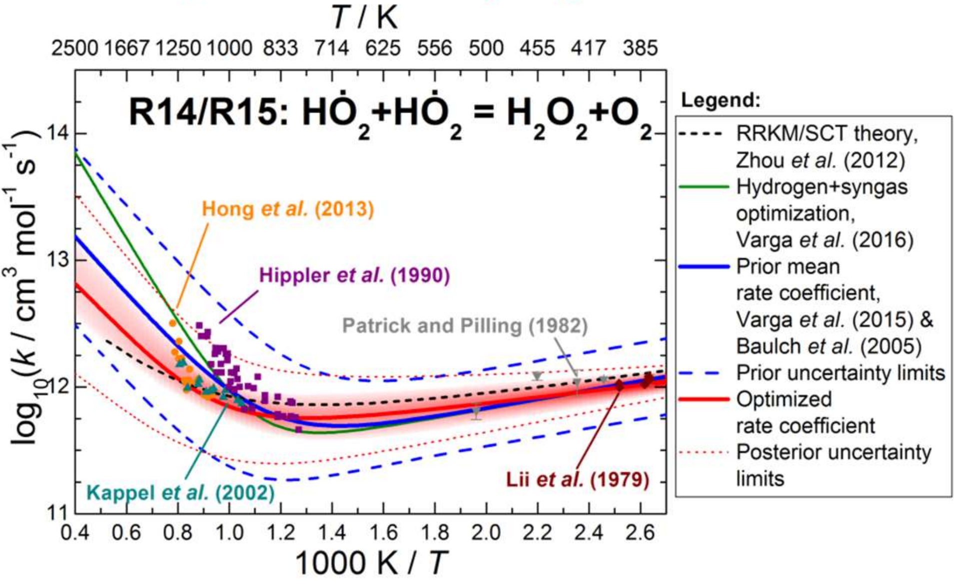
No.	Reaction	f_{prior}	$f_{posterior}$	$f_{posterior+H2/CO}$	
R14/R15	$H\dot{O}_2+H\dot{O}_2 = H_2O_2+O_2$	0.30–0.70	0.08–0.71	0.09–0.57	↘
R37 LPL	$CH_2O+M = CO+H_2+M$	0.50	0.09–0.12	0.09–0.12	→
R38	$CH_2O+\dot{H} = H\dot{C}O+H_2$	0.60	0.08–0.10	0.09–0.11	→
R40	$CH_2O+\dot{O}H = H\dot{C}O+H_2O$	0.34–0.43	0.22–0.30	0.19–0.27	↘
R41	$CH_2O+O_2 = H\dot{C}O+H\dot{O}_2$	1.20	0.19–0.20	0.18–0.20	→
R47	$\dot{C}H_3+H\dot{O}_2 = CH_3\dot{O}+\dot{O}H$	0.46–0.76	0.26–0.49	0.26–0.38	↘
R53	$\dot{C}H_3+H\dot{O}_2 = CH_4+O_2$	1.00	0.25–0.45	0.25–0.32	↘
R60	$\dot{C}H_2OH+O_2 = CH_2O+H\dot{O}_2$	0.50	0.27–0.38	0.27–0.33	↘
R67 LPL	$CH_3\dot{O}+M = CH_2O+\dot{H}+M$	0.84–1.24	0.16–0.26	0.16–0.21	→
R77 HPL	$\dot{O}H+\dot{C}H_3 = CH_3OH$	0.34–0.84	0.20–0.31	0.19–0.28	↘
R77 LPL	$\dot{O}H+\dot{C}H_3+M = CH_3OH+M$	1.20	0.07–0.39	0.07–0.27	↘
R80	$CH_3OH+\dot{H} = \dot{C}H_2OH+H_2$	0.44–1.07	0.18–0.27	0.17–0.24	→
R81	$CH_3OH+\dot{H} = CH_3\dot{O}+H_2$	1.70	0.24–0.38	0.24–0.31	→
R83	$CH_3OH+\dot{O}H = CH_3\dot{O}+H_2O$	0.70	0.13–0.44	0.12–0.25	↘
R84	$CH_3OH+\dot{O}H = \dot{C}H_2OH+H_2O$	0.46–0.87	0.19–0.41	0.18–0.40	→
R85	$CH_3OH+O_2 = \dot{C}H_2OH+H\dot{O}_2$	0.80	0.78–1.01	0.72–0.91	↘
R87	$CH_3OH+H\dot{O}_2 = \dot{C}H_2OH+H_2O_2$	1.10	0.20–0.25	0.16–0.21	↘
R88	$CH_3OH+H\dot{O}_2 = CH_3\dot{O}+H_2O_2$	0.70	0.15–0.42	0.15–0.26	↘

When also considering
all sensitive H_2/CO reactions:

and the corresponding
H_2/CO data

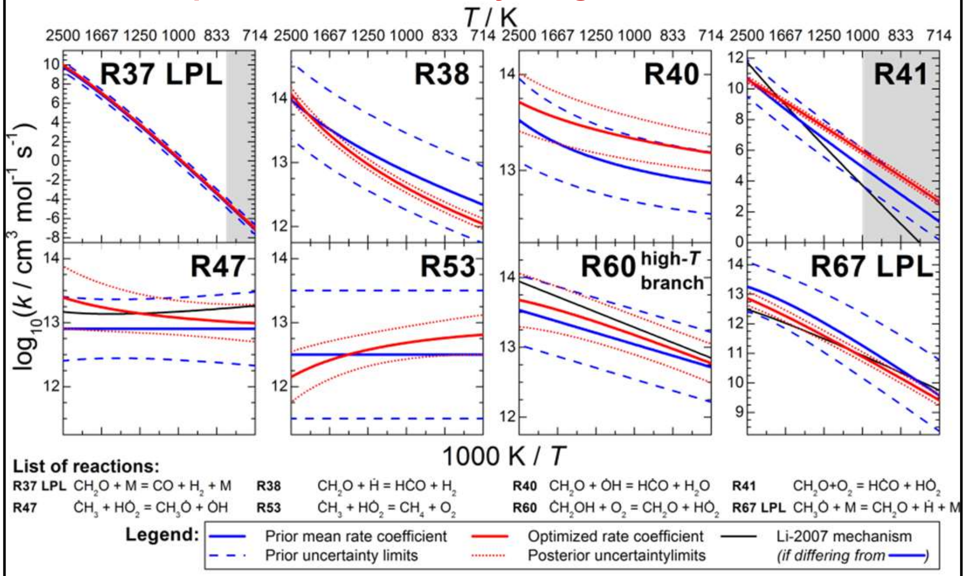
Arrhenius plots (I)

Prior and posterior uncertainty ranges



Arrhenius plots (II)

Prior and posterior uncertainty ranges



List of reactions:

R37 LPL $CH_2O + M = CO + H_2 + M$
R38 $CH_2O + \dot{H} = H\dot{C}O + H_2$
R40 $CH_2O + \dot{O}H = H\dot{C}O + H_2O$
R41 $CH_2O+O_2 = H\dot{C}O + H\dot{O}_2$
R47 $\dot{C}H_3 + H\dot{O}_2 = CH_3\dot{O} + \dot{O}H$
R53 $\dot{C}H_3 + H\dot{O}_2 = CH_4 + O_2$
R60 $\dot{C}H_2OH + O_2 = CH_2O + H\dot{O}_2$
R67 LPL $CH_3\dot{O} + M = CH_2O + \dot{H} + M$

Mechanism comparison results

Error function values for each type of data and overall

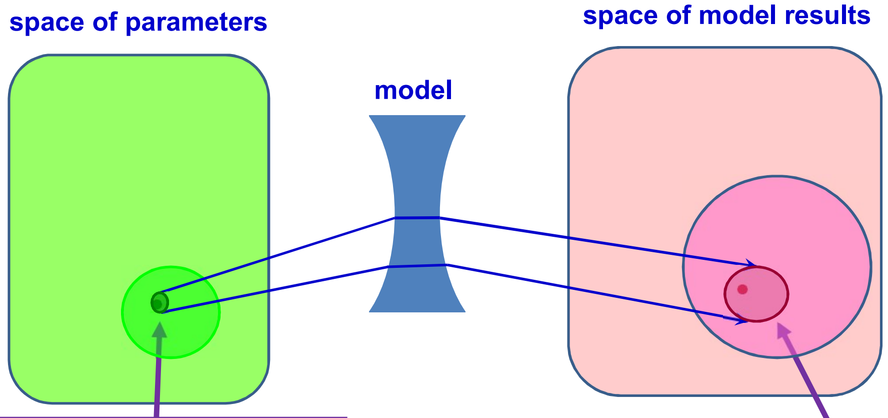
Mechanism	Average error function value Ignition delay times CH_3OH	 Ignition delay times CH_2O	 Burning velocities CH_3OH	 Concentration profiles CH_3OH	 Concentration profiles CH_2O	Overall
AAU-2008	28.4	2.6	no transport	21.3	18.9	–
Alzueta-2001	103.9[a]	11.3	no transport	32.1	15.9	–
Christensen-2016	35.4	3.3	5.2	15.7	8.4	**16.3**
Hamdane-2012	41.2	6.4	(72.2)	109.6	12.9	**(59.8)**
Klippenstein-2011	131.8	3.4	1.7	14.9	6.4	**41.5**
Li-2007	7.6	3.4	1.9	14.8	6.4	**6.9**
Rasmussen-2008	32.2	5.2	17.9	40.9	12.3	**25.3**
CaltechMech2.3-2015	51.4	2.5	3.0	15.9	6.4	**19.7**
Johnson-2009	11.8	10.9	no transport	19.9	13.1	–
Kathrotia-2011	10.7	8.2	[3.0]	60.9	43.0	**[23.6]**
Konnov-2009	54.6	7.5	[76.0]	25.6[b]	37.2	**[51.7][b]**
Leplat-2011	410.5[a]	6.1	22.2	52.6	38.1	**131.6[a]**
Marinov-1999	260.7	20.7	(14.1)	36.7	30.4	**(90.4)**
SanDiego-2014	24.7	1.6	3.8	31.4	10.0	**16.2**
SaxenaWilliams-2007	86.2	1.7	2.3	17.3	9.6	**30.0**
USC-II-2007	602.8[a]	2.3	(8.5)	27.8	19.1	**(178.4)[a]**
Aramco1.3-2013	(41.3)	(11.1)	(4.3)	(16.0)	(12.7)	**(18.6)**
NUIG-16.09-2016	(51.6)	(11.0)	(4.2)	(20.0)	(12.4)	**(22.2)**
Initial mechanism	8.1	2.3	2.1	15.1	12.4	**8.1**
Optimized mechanism	6.8	2.0	1.6	12.0	6.3	**5.9**
No. of data sets	**74**	**7**	**87**	**54**	**43**	**265**
No. of data points	**475**	**99**	**632**	**15,363**	**8,331**	**24,900**

$$E = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(\frac{Y_{ij}^{\mathrm{mod}} - Y_{ij}^{\mathrm{exp}}}{\sigma(Y_{ij}^{\mathrm{exp}})}\right)^2$$

$E = 9$: data can be described within a 3σ uncertainty

Improvement for all types of data!

Posterior uncertainty of parameters

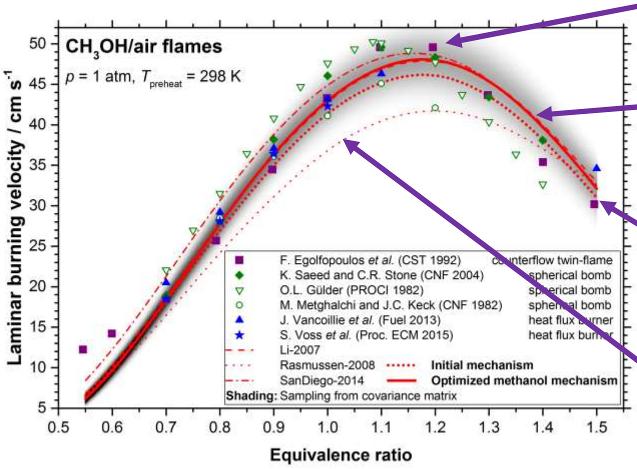


posterior uncertainty domain of model parameters

uncertainty of simulations obtained from posterior uncertainties

Comparison of simulation results

Propagation of rate coefficient posterior uncertainties



dots: measured flame velicity

simulated flame velocity optimized mech

shading: *pdf* of the simulated flame velocity optimized mech

simulated flame velocity using other mech's

Summary

- **New optimized mechanism** for **methanol** and **formaldehyde combustion** simulations
- **Best reproduction** of indirect experimental data, while optimized rate coefficients are **consistent with direct measurements** and **theoretical calculations** within their uncertainty limits
- Determination of the **posterior uncertainty domain** of the rate parameters

C. Olm, T. Varga, É. Valkó, H. J. Curran, T. Turányi:
Uncertainty quantification of a newly optimized methanol and formaldehyde combustion mechanism
Combust. Flame, **186**, 45-64 (2017)

Topic 9: Time-scale analysis

Lifetime and its interpretation for various systems,

stiff systems, slow and fast variables,

slow manifolds in dynamical systems,

modes, calculation of the dynamical dimension,

stability analysis of stationary and dynamic systems

Lifetimes and time scales

Half life:

Time period needed for the concentration of a species to decrease to ½,
if during this time it is not produced and
the concentrations of all other species remains identical.

Lifetime:

Time period needed for the concentration of a species to decrease to 1/*e*,
if during this time it is not produced and
the concentrations of all other species remains identical.

Single first order reaction : $Y = Y_0\,\mathrm{e}^{-kt}$

A → P

lifetime: $\tau = \dfrac{1}{k}$ half life: $\tau_{1/2} = \dfrac{\ln 2}{k}$

Several first order reactions
(*e.g.* in photochemistry the reactions of an excited species):

$$\begin{aligned} \mathrm{A} &\rightarrow \mathrm{P}_1 \\ &\rightarrow \mathrm{P}_2 \\ &\rightarrow \mathrm{P}_3 \end{aligned}$$

$Y = Y_0\,\mathrm{e}^{-(k_1+k_2+k_3)t}$ lifetime: $\tau = \dfrac{1}{(k_1+k_2+k_3)}$

Lifetime

Atmospheric chemistry:
small radical concentrations ⇒ radical-radical reactions are missing (*e.g.* 2 $CH_3 \rightarrow C_2H_6$) ⇒ no Y_i^2 type terms in the kinetic system of ODEs

P_i	effect of producing steps
L_i	effect of consuming steps

Production rate of Y_i : $$\mathrm{d}Y_i/\mathrm{d}t = P_i - L_i Y_i$$

Y_i lifetime: $$\tau = \frac{1}{L_i}$$

General reaction mechanism:

Y_i lifetime: $\tau = -\frac{1}{j_{ii}}$ where $j_{ii} = \frac{\partial f_i}{\partial Y_i}$

j_{ii} is the i^{th} diagonal element of the Jacobian

Example

A	→ B	k_1
A + C	→ D	k_2
B	→ A	k_3

Production rate of species A :

$$\mathrm{d}a/\mathrm{d}t = -k_1 a - k_2\, ac + k_3\, b$$
$$\mathrm{d}a/\mathrm{d}t = k_3\, b - (k_1 + k_2\, c)a$$

„Atmospheric chemical" lifetime:

$$\mathrm{d}Y_i/\mathrm{d}t = P_i - L_i Y_i \qquad \tau_{\mathrm{A}} = 1/L_{\mathrm{A}} = 1/(k_1 + k_2\, c)$$

General lifetime:

$$j_{\mathrm{AA}} = \frac{\partial(\mathrm{d}a/\mathrm{d}t)}{\partial a} = -(k_1 + k_2\, c) \qquad \tau_{\mathrm{A}} = -1/j_{\mathrm{AA}} = 1/(k_1 + k_2\, c)$$

Slow variables and fast variables

The concentration of a single species is changed by $\Delta y'_i$
and the concentration change of the other species is negligible:

$$\Delta y_i'(t) = \Delta y_i'^{0} \, \mathrm{e}^{j_{ii} t}$$

short lifetime species = the effect of perturbation decreases rapidly =
the perturbed trajectory converges very fast to the the original trajectory
⇒ **fast variable**

long lifetime species = the effect of perturbation does not decrease rapidly
= the original and the perturbed trajectory are „parallel”
⇒ **slow variable**

Consequences:

- fast variables „forget” their initial value
- the values of fast variables are determined by the values of the other variables
- the „slow/fast variable” classification is independent of the actual change of the variables in time ($\mathrm{d}Y_i/\mathrm{d}\,t$)

Eigenvector-eigenvalue decomposition of the Jacobian

$$\mathbf{\Lambda} = \mathbf{W}\,\mathbf{J}\,\mathbf{V} \qquad\qquad \mathbf{J} = \left\{\frac{\partial f_i}{\partial Y_j}\right\}$$

$\mathbf{\Lambda}$ diagonal matrix that contains the eigenvalues (complex eigenvalues!)

$\mathbf{W}$ matrix of left eigenvectors (row vectors)

$\mathbf{V}$ matrix of right eigenvectors (column vectors)

denote $\mathbf{W}_\mathrm{f}$ the matrix of left eigenvectors, related to small negative Re(λ)
(„eigenvectors related to fast directions”)

Features: the left and right eigenvectors are orthonormed:

$\mathbf{I} = \mathbf{W}\,\mathbf{V}$ $\quad$ $\mathbf{I} = \mathbf{V}\,\mathbf{W}$ $\quad$ therefore: $\mathbf{J} = \mathbf{V}\,\mathbf{\Lambda}\,\mathbf{W}$

Stiff systems 1

The eigenvalues define the time scales of a model: $t_i = 1/|\mathrm{Re}(\lambda_i)|$

Very different time scales ⇒ stiff mathematical models

Mathematicians' definition of stiffness:
the ratio of the longest and shortest time scale

stiffness ratio S_1

$$S_1 = \underbrace{\frac{1}{\min_i |\mathrm{Re}(\lambda_i)|}}_{\text{longest time scale}} \Big/ \underbrace{\frac{1}{\max_i |\mathrm{Re}(\lambda_i)|}}_{\text{shortest time scale}}$$

$$S_1 = \max_i |\mathrm{Re}(\lambda_i)| \,/\, \min_i |\mathrm{Re}(\lambda_i)|$$

Gear, C. W., Numerical Initial-Value Problems in Ordinary Differential Equations, Englewood Cliffs: Prentice Hall, 1971

Stiff systems 2

Physicists' and chemists' definition of stiffness:
the ratio of the characteristic („typical") time scale and the shortest time scale

stiffness ratio S_2

$$S_2 = \underbrace{\tau}_{\text{characteristic time scale}} \Big/ \underbrace{\frac{1}{\max_i |\mathrm{Re}(\lambda_i)|}}_{\text{shortest time scale}}$$

stiffness ratio $S_2 = \tau \max_i |\mathrm{Re}(\lambda_i)|$
τ is the characteristic time scale of the system

A modell is considered stiff, if the stifness ratio is large (*e.g.* 10^8-10^{12})

The stiff systems of differential equations:

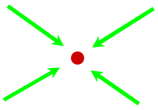
⇒ can be solved with special algorithms only
(„backward differentiation formulas", „implicit solvers")
⇒ stiffness changes with changing concentrations

Stability analysis of a stationary system

differential equation of the system $\mathrm{d}\mathbf{Y}/\mathrm{d}t = \mathbf{f}(\mathbf{Y},\mathbf{p})$

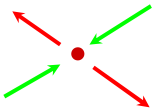
in the stationary point $\mathrm{d}\mathbf{Y}/\mathrm{d}t = \mathbf{0}$

Stable stationary point: we move it out, goes back



Jacobian $\mathbf{J} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}$ the real parts of all eigenvalues are negative

Unstable stationary point: we move it out, goes away

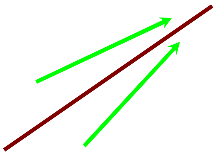


There exist at least one eigenvalue of the Jacobian having positive real part

Stability analysis of a moving system

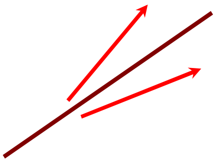
differential equation of the system $\mathrm{d}\mathbf{Y}/\mathrm{d}t = \mathbf{f}(\mathbf{Y},\mathbf{p})$

Stable trajectory: we move it out, goes back to the original trajectory



Jacobian $\mathbf{J} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}$ the real parts of all eigenvalues are negative

Unstable trajectory: we move it out, goes away



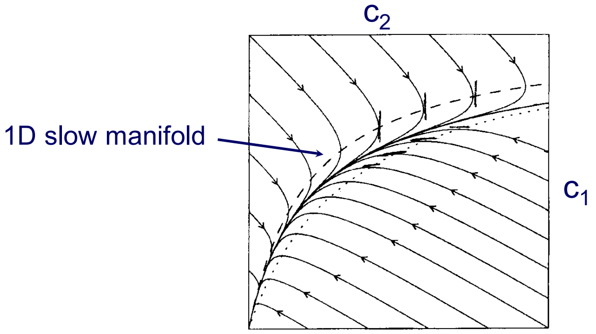
The exist at least one eigenvalue of the Jacobian that the real part of it is positive

Chemical example: autocatalytic runaway, explosions

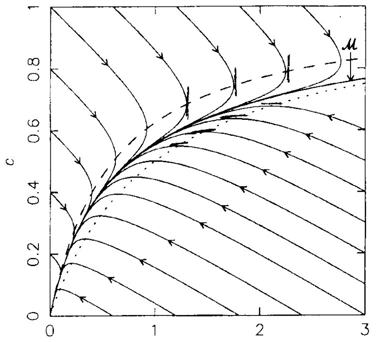
Slow manifolds in dynamical systems

Assume that the manifolds are attracting
(no strong autocatalysis, no explosion)
In the space of variables (in chemical kinetics: concentrations) the fast moving points approach a given object (called slow manifolds) and move slowly along it.

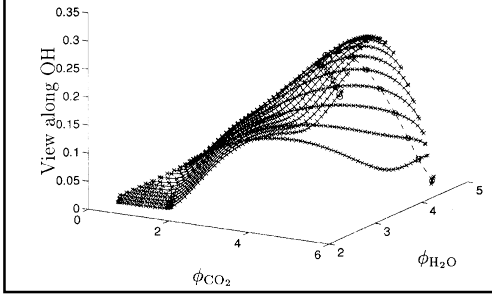
2D manifold : the trajectories approach a plane
1D manifold: the trajectories approach a (curved) line
0D manifold: the trajectories approach the equilibrium point

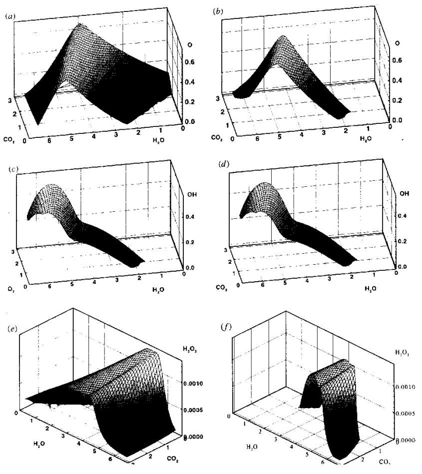


Calculated slow manifolds in combustion systems



O Gicquel et al





Modes

Previous assumption: „The concentration of a single species is changed by $\Delta y'_i$ and the concentrations of the other species are not changed"
This is approximately true if the concentration of the changed species is low and its lifetime is short.

Usually changing the concentration of a single species induces the change of the concentrations of several other species.
⇒ A more general approach:
the concentrations of several species are changed simultaneously by $\Delta\mathbf{Y}$

linear approach
to the change of the perturbation:

$$\frac{\mathrm{d}(\Delta\mathbf{Y})}{\mathrm{d}t} \approx \frac{\partial\mathbf{f}}{\partial\mathbf{Y}}\Delta\mathbf{Y} = \mathbf{J}\Delta\mathbf{Y}$$

The solution of this ODE: (assuming that the Jacobian does not change during the short time of relaxation)

$$\Delta\mathbf{Y}(t) = e^{\mathbf{J}\tau}\Delta\mathbf{Y}(t_1) \qquad \Rightarrow \qquad \Delta Y_i = C_1 e^{\lambda_1\tau} + C_2 e^{\lambda_2\tau} + C_3 e^{\lambda_3\tau} + \ldots + C_n e^{\lambda_n\tau}$$

where $\tau = t - t_1$ is the elapsed time from the start of perturbation

Modes 2

introducing new variables: **z** vector of modes
c.f. normal coordinates in spectroscopy

calculation of mode z_i : $z_i = \mathbf{w}_i\ \mathbf{Y}$ $\qquad \mathbf{z} = \mathbf{W}\ \mathbf{Y}$

where $\mathbf{w}_i$ is the *i*-th row of matrix **W**

calculation of concentration y_i : $Y_i = \mathbf{v}_i\ \mathbf{z}$ $\qquad \mathbf{Y} = \mathbf{V}\ \mathbf{z}$

where $\mathbf{v}_i$ is the *i*-th column of matrix **V**

the transformated kinetic system of ODEs: $\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \mathbf{W}\ \mathbf{f}(\mathbf{V}\mathbf{z}),\qquad \mathbf{z}_0 = \mathbf{W}\ \mathbf{Y}_0$

If perturbation $\Delta\mathbf{Y}$ is towards direction $\mathbf{w}_1$, then $z_1 \neq 0$,
but for all the other *j* directions $z_j = 0$.

The change of z_i after perturbation $\qquad \Delta z_i(t) = \Delta z_i^0\ \mathrm{e}^{\lambda_i t}$

IMPORTANT: the Jacobian changes in the space of concentrations, therefore transformation **Y** → **z** also changes with **Y**!

Modes 3

Denote $\mathbf{z}_f$ the fast modes:

$$\tilde{\mathbf{Y}}(t) = \mathbf{Y}(t) + \Delta\mathbf{Y}(t)$$

$$\frac{\mathrm{d}\tilde{\mathbf{Y}}}{\mathrm{d}t} = \frac{\mathrm{d}(\mathbf{Y}+\Delta\mathbf{Y})}{\mathrm{d}t} = \mathbf{f}(\mathbf{Y},\mathbf{p}) + \frac{\mathrm{d}\Delta\mathbf{Y}}{\mathrm{d}t}$$

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \mathbf{W}\,\mathbf{f}(\mathbf{V}\mathbf{z}), \qquad \mathbf{z}_0 = \mathbf{W}\,\mathbf{Y}_0$$

$$\frac{\mathrm{d}z_i}{\mathrm{d}t} = \mathbf{w}_i\mathbf{f} = \mathbf{w}_i\mathbf{f}(\mathbf{Y}^m) + \mathbf{w}_i\frac{\mathrm{d}\Delta\mathbf{Y}}{\mathrm{d}t} = \frac{\mathrm{d}\Delta z_i}{\mathrm{d}t}$$

Let *i* belong to a fast mode and let $\mathbf{Y}^m$ be a point on the manifold.

$$\mathbf{W}_i\mathbf{f}(\mathbf{Y}^m) = \mathbf{0}$$

Therefore, for mode *i*:

$$\frac{\mathrm{d}\Delta z_i}{\mathrm{d}t} = \mathbf{w}_i\mathbf{f}$$

Modes 4

From the previous page: $\frac{\mathrm{d}\Delta z_i}{\mathrm{d}t} = \mathbf{w}_i\mathbf{f}$

The modes can be perturbed independently from each other:

$$\frac{\mathrm{d}\Delta z_i}{\mathrm{d}t} = \lambda_i\,\Delta z_i$$

Comparing the two equations above:

$$\mathbf{w}_i\mathbf{f} = \lambda_i\Delta z_i$$

$$\Delta z_i = \mathbf{w}_i\mathbf{f}/\lambda_i$$

where Δz_i is the distance from manifold *i*

Calculation of the dynamical dimension

$$\Delta z_i = \mathbf{w}_i \mathbf{f} / \lambda_i$$

Δz_i distance from the slow manifold towards direction $\boldsymbol{w}_j$

If $|\Delta z_i| < z_{\text{threshold}}$ $\Rightarrow$ this mode is (almost) on the manifold

Assume that there are n_r modes
relaxed onto the corresponding slow manifold.

n	number of variables in the model
n_c	number of conserved properties = number of zero eigenvalues of the Jacobian
n_r	number of relaxed modes

actual dynamical dimension: $n_D = n - n_c - n_r$.

Change of the dynamical dimension during the course of a chemical reaction

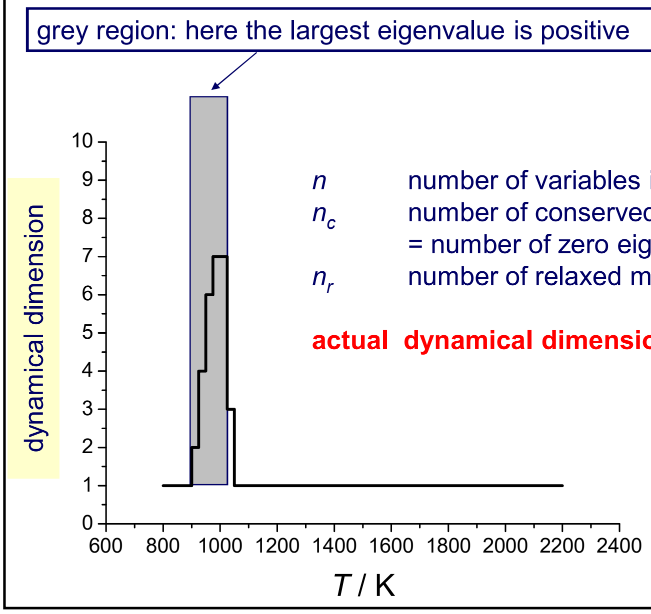
n	number of species in the mechanism
n_c	number of conserved properties (*e.g.* number of elements in a closed system)

If the initial composition is randomly selected,
initially the trajectory is moving in a (n - nc) dimensional space
(change of elementary composition by chemical reactions is forbidden)

If all eigenvalues of the Jacobian are negative:
- all manifolds are attractive
- the actual dynamical dimension is continously decreasing
- finally it nD= 1 (moving along a line) and the nD=0 (equilibrium point)

If some eigenvalue(s) of the Jacobian are positive:
- some manifolds are repellent
- the actual dynamical dimension is increasing
- the case of autocatalytic runaways (including explosions)

Change of dimension during adiabatic hydrogen/air explosion

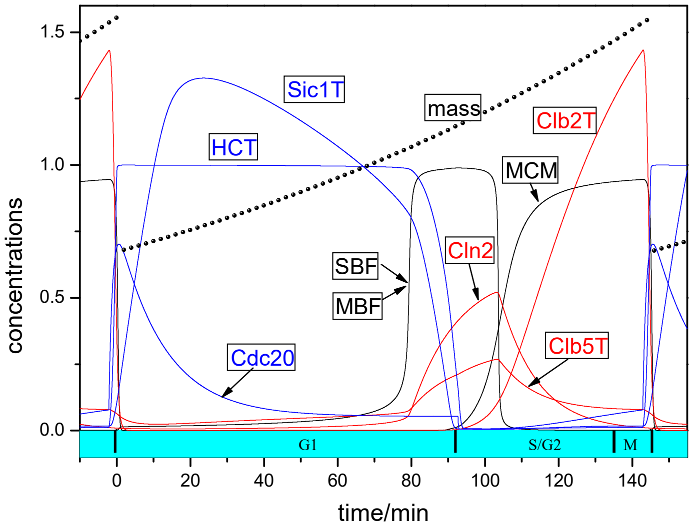


n	number of variables in the model
n_c	number of conserved properties = number of zero eigenvalues of the Jacobian
n_r	number of relaxed modes

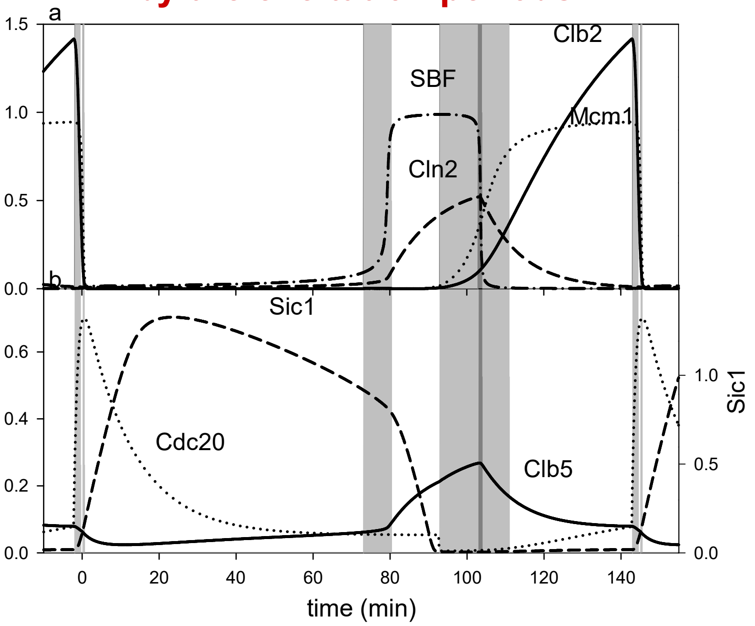
actual dynamical dimension: $n_D = n - n_c - n_r$.

201

Change of protein concentrations during a cell cycle



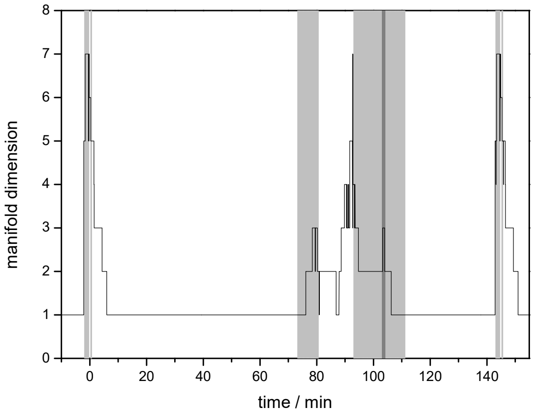
Concentration changes are affected by the excitation periods



Calculation the the dynamic dimension 2

n	number of ODE variables
n_c	number of conserved properties = number of the zero eigenvalues of the Jacobian
n_r	number of relaxed modes

Actual dynamical dimension: $n_D=n-n_c-n_r$



Topic 10: Reduction of reaction mechanisms 1: Elimination of redundant reactions

General principles, characteristics of detailed mechanisms.

Why is it permitted and useful to reduce reaction mechanisms?

Rate-of-production analysis,

principal component analysis of matrices **S** and **F**,

integer programming methods,

genetic algorithm-based methods.

Need for the reduction of large reaction mechanisms

Increasing knowledge in chemical kinetics
⇒ increasing size of reaction mechanisms

Typical sizes of detailed reaction mechanisms:

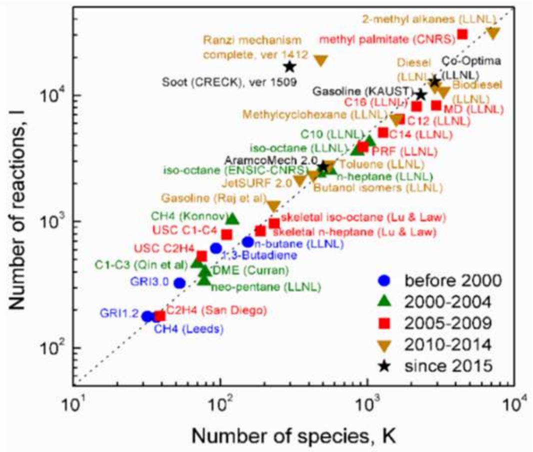
high temperature combustion:	30-100 species, 200-500 reaction steps
low temperature combustion:	500 species, 10000 reaction steps
tropospheric chemistry:	500 species, 10000 reaction steps

Large mechanisms are not for humans:

Small reduced mechanisms are needed for the
- understanding of the chemical processes
- fast calculation of chemistry

Reaction mechanisms are getting larger

The reaction mechanisms are getting larger, because
– increasing chemical knowledge
– faster computers with larger memory are available



F.N. Egolfopoulos, N. Hansen, Y. Ju, K. Kohse-Höinghaus, C.K. Law, F. Qi, *Prog. Energy Combust. Sci.* **43** (2014) 36–67.

Need for the reduction of large reaction mechanisms 2.

Practical requirement:
simulation of spatially distributed systems with complex geometry.
The simulations must be fast and accurate.

The usual strategy: mechanism reduction at a series of simplified conditions, like plug-flow reactors, PSRs, 1D laminar flames.

These simple models together cover all the conditions in the model having complex geometry.

Is this a good approximation or an oversimplification?

I. Gy. Zsély, T. Turányi: Thermal and diffusion couplings do not affect mechanism reduction *Phys.Chem.Chem.Phys.*, **5**, 3622-3631 (2003)

Advantages of mechanism reduction

1 Spatially homogeneous system

- described by an ordinary system of differential equations
  fast numerical simulation
  a few minutes simulation time of a model having
  several thousand species and several ten thousands of reactions
- nobody understands such a big mechanism.
  a skeleton model is needed to understand the chemistry

2 Spatially inhomogeneous system

- described by a partial system of differential equations
  very slow numerical simulation
  method of operator splitting is frequently used
  the chemistry and the advection/diffusion is simulated separately:

$$\boxed{\frac{\mathrm{d}\mathbf{c}(\mathbf{r},t)}{\mathrm{d}t} = \mathbf{f}\left(\mathbf{r},t,\mathbf{c}\right) + \Theta(\mathbf{r},t,\mathbf{c},\nabla\mathbf{c},\nabla^2\mathbf{c})}$$

Most of the computer time (*e.g.* 99%) is used for solving the chemical kinetic equations

Mechanism reduction offers a good possibility for speeding up the calculations!

Advantages of mechanism reduction 2

The original mechanism is not the „real one", because …

- the reaction mechanisms contain thermal average rate coefficients and the concentrations represent the thermal average quantum state of species
- selection of the list of species depend on the sensitivity of the applied chemical analytical methods
- some species are known to be present, but most models do not contain them. Example: species CHO^+ in hydrocarbon flames
- almost all mechanisms contain redundant species and reactions, because presence of such species/reactions do not cause error
- none of the original mechanisms are general, but these have been created for a given domain of (**y**, *p*, *T*)

it is not „unfair" to reduce reaction mechanisms, since…

- the domain of applicability of a reaction mechanism may be smaller than the original domain of validity
- it is OK if the simulation results change less due to reduction than the accuracy of the validation experiments (say 5% or 10%)
- we interest in the concentrations of some species only, not all calculated species concentrations.

Overview of mechanizmus reduction methods

I. without considering time scale separation

1. **selection of a submechanism**
   elimination of redundant species and redundant reactions
   ⇒ skeleton mechanism; smaller ODE

2. lumping of reactions and species
   ⇒ smaller mechanism, smaller ODE

II. based on time scale separation

1. Simplification of the ODE: QSSA, partial equilibrium (PE)
   ⇒ coupled ODE and algebraic equations

2. Numerical representation of kinetic information:
   ILDM, ISAT, repromodelling
   ⇒ numerically stored kinetic equations:
   look-up tables, polynomials, ANNs

Removal of redundant reactions
the classic rate-of-production analysis

The traditional method: rate-of-production analysis (**ROPA**):

The percentage contribution of each reaction step to the formation and removal of each species is investigated in several time points.

A reaction step can be eliminated, if its contribution is less then (say) 5% to the formation/removal of any species.

Example:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -k_1xy + k_2ya - 2k_3x^2 + k_4xa - 0{,}5k_5xy$$

all producing terms — all consumption terms

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \quad +k_2ya + k_4xa \qquad -k_1xy - 2k_3x^2 - 0{,}5k_5xy$$

contribution of reaction step 4 to the producing terms

Easy to understand and easy to calculate.

BUT: large amount of data have to be considered
if there are many time poinst, reaction steps and species.

Table of contributions calculated by KINALC for a methane pyrolysis simulation

```
CH4                  Rate :   -1.082E-08
        No   Contribution           #  reaction

         1   -1.17547E-08   34.8 %C  152  C2H3+CH4  => C2H4+CH3
         2    9.13403E-09   39.8 %P  151  C2H4+CH3  => C2H3+CH4
         3   -7.16821E-09   21.2 %C  140  CH≡CCH2.+CH4  => CH≡CCH3+CH3
         4   -5.71162E-09   16.9 %C  136  CH2=CHCH2.+CH4  => CH2=CHCH3+CH3
         5    3.72978E-09   16.2 %P  139  CH≡CCH3+CH3  => CH≡CCH2.+CH4
         6    3.60756E-09   15.7 %P  135  CH2=CHCH3+CH3  => CH2=CHCH2.+CH4
         7    3.45436E-09   15.0 %P  143  C2H6+CH3  => C2H5+CH4

CH3                  Rate :    4.636E-10
        No   Contribution           #  reaction

         1   -1.69871E-08   22.7 %C    7  2CH3  => C2H6
         2    1.17547E-08   15.6 %P  152  C2H3+CH4  => C2H4+CH3
         3   -9.13403E-09   12.2 %C  151  C2H4+CH3  => C2H3+CH4
         4    7.84580E-09   10.4 %P    8  C2H6  => 2CH3

C2H2                 Rate :    1.996E-09
        No   Contribution           #  reaction

         1    6.06561E-09   44.1 %P   81  CH2=CHCH2.  => C2H2+CH3
         2   -5.12527E-09   43.6 %C   82  C2H2+CH3  => CH2=CHCH2.
         3    3.86101E-09   28.1 %P  190  C2H3(+M)  => H+C2H2(+M)
         4   -3.15073E-09   26.8 %C  189  H+C2H2(+M)  => C2H3(+M)
         5   -1.15098E-09    9.8 %C   67  C2H2+CH3  => CH≡CCH3+H
```

Principal component analysis of the sensitivity matrix: PCAS

Another approach for the identification of redundant reaction steps (= identification of redundant parameters)

It has been discussed in the local sensitivity analysis section.

In several time points the normalized local sensitivity matrix $\mathbf{S}_r$ is calculated, component matrix $\tilde{\mathbf{S}}$ is formed and the eigenvector-eigenvalue analysis of matrix $\tilde{\mathbf{S}}^{\mathrm{T}}\tilde{\mathbf{S}}$ is carried out.

$$\tilde{\mathbf{S}} = \begin{bmatrix} \tilde{\mathbf{S}}_1 \\ \tilde{\mathbf{S}}_2 \\ \vdots \\ \tilde{\mathbf{S}}_n \end{bmatrix}$$

$$\tilde{\mathbf{S}}_r = \{(p_k / Y_i)(\partial Y_i(t_r) / \partial p_k)\}$$

eigenvector j : $\mathbf{u}_j$ defines the parameters acting together for influencing the concentrations at time t_r

eigenvalue i : λ_i indicates the importance of the parameter group.

A reaction step is important, if all species concentrations are considered in the objective function and if the corresponding rate parameter belongs to an important parameter group (large λ_i) and it is important within the parameter group (large eigenvector element).

Principal component analysis of the rate sensitivity matrix: PCAF

the normed **F** matrix is calculated in several time points *r*

$$\tilde{\mathbf{F}}_r = \{(p_k / f_i)(\partial f_i(t_r)/\partial p_k)\}$$

eigenvector-eigenvalue analysis of matrix $\tilde{\mathbf{F}}_r^{\mathrm{T}}\tilde{\mathbf{F}}_r$

eigenvector *j* :	$\mathbf{u}_j$ defines the parameters acting together for influencing the production rates at time t_r
eigenvalue *i* :	λ_i indicates the importance of the parameter group.

A reaction step is important at time *r*, if all species production rate are considered in the objective function and if the corresponding rate parameter belongs to an important parameter group (large λ_i) and it is important within the parameter group (large eigenvector element).

A reaction step is important in a time interval, if it is important at least in a single time point

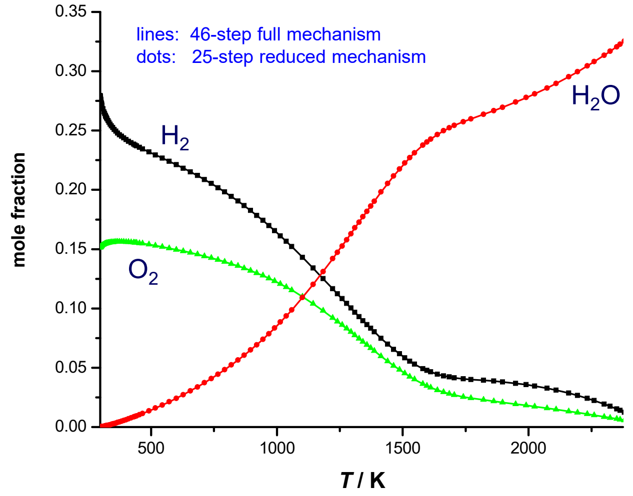
T. Turányi, T. Bérces, S. Vajda: Reaction rate analysis of complex kinetics systems *Int.J.Chem.Kinet.*, **21**, 83-99 (1989)

PCAS and PCAF are basically different

PCAF	**PCAS**
• investigation of reaction rates, which depend on the actual concentrations only	• investigation of sensitivity functions, which depend on the prehistory of the simulation
• analysis in time points	• analysis belongs to a time interval
• change of importance in time can be investigated	• direct inspection of the change of simulation results; close connection to parameter estimation
• **F** calculated analytically	• **S** calculated numerically

In all investigated cases PCAF and PCAS provided exactly the same reduced mechanisms.

Example: H_2-air flame mechanism reduction PCAS and PCAF provided identical reduced mechanisms



I. Gy. Zsély, T. Turányi: The influence of thermal coupling and diffusion on the importance of reactions: The case study of hydrogen-air combustion, *Phys.Chem.Chem.Phys.*, **5**, 3622-3631(2003)

Integer programming methods

$$\dot{\mathbf{y}} = \boldsymbol{\upsilon}\mathbf{R}(\mathbf{y})$$

d **y**/d *t* production rates belonging to the full mechanism
υ stoichiometric matrix
R rates of the reaction steps

$$\dot{\mathbf{z}} = \boldsymbol{\upsilon}\mathbf{D}\mathbf{R}(\mathbf{z})$$

d **z**/d *t* production rates belonging to the reduced mechanism
D diagonal matrix:
$d_{ii} = 0$ the reaction step is missing from the reduced mech
$d_{ii} = 1$ the reaction step is present in the reduced mech

The aim is finding the reduced mechanism having k reaction steps with minimal deviation between the production rates of the full and reduced mechanisms by varying vector **d**.

$$\begin{aligned}
&\min \| \mathbf{y} - \mathbf{z} \| \\
&\text{subject to} \\
&\dot{\mathbf{y}} = \boldsymbol{\upsilon}\mathbf{R}(\mathbf{y}), \qquad \mathbf{y}(0) = \mathbf{y}_o \\
&\dot{\mathbf{z}} = \boldsymbol{\upsilon}\mathbf{R}(\mathbf{z}), \qquad \mathbf{z}(0) = \mathbf{z}_o \quad 0 \le t \le b \\
&\sum_{i=1}^{N} d_{ii} = k, \qquad d_{ii} = 0 \text{ or } 1
\end{aligned}$$

L. Petzold, W. Zhu: Model reduction for chemical kinetics: An optimization approach. *AIChE Journal* **45**, 869-886 (1999)

Integer programming methods 2.

5-step full mechanism ⇒ 2^5-1 possible reduced mechanisms
⇒ trial of all combinations does not work
⇒ a classic integer programming problem

several approaches were suggested by Petzold and Zhu:
- greedy algorithm: elimination of the least important reactions first
- elimination of reactions first via the elimination of redundant species
- all variants require much computer time

Further improvements:

Androulakis: branch and bound algorithm, which splits the feasible region of input values into smaller subregions (branching) with the subregions forming a search tree.

Another approach: the simulation error
(deviation between the solutions of the reduced and full models)
is investigated instead of the deviations of the production rates.

I. P. Androulakis: Kinetic mechanism reduction based on an integer programming approach. *AIChE J.* **46**, 361-371 (2000)

Genetic algorithm-based methods

Again, the reduced mechanism is represented by a **d** vector with elements of either 1 or 0.

The conceptual background and the terminology is based on **evolution**

candidate reduced mechanisms (called **individuals**):

Individual 1	1100101010
Individual 2	1100111000
Individual 3	1100100011
Individual 4	1110001010

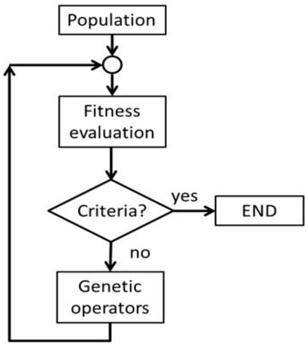
fitness criterion	agreement of the reduced mechanism with the full mech

genetic operators: cross over and mutation

cross over	exchange of sections of two individuals
mutation	changing 0 to 1 OR 1 to 0 in an individual

Genetic algorithm-based methods 2

1 random population of reduced mechanisms is produced.
2 the fitness is evalued (comparison of the full and reduced mechs)
3 if a good enough reduced mech is found ⇒ **END**
4 variants of the reduced mechanisms are produced by genetic operators
5 the newly produced mechs are added to the population;
the poor individuals are replaced by the new ones ⇒ **GOTO 2**



EVOLUTION: some of the the new individuals are at least not worse than the previous ones
⇒ **improvement is expected**

K. Edwards, T. F. Edgar, V. I. Manousiouthakis: Kinetic model reduction using genetic algorithms. *Comput. Chem. Eng.* **22**, 239-246 (1998)

N. Sikalo, O. Hasemann, C. Schulz, A. Kempf, I. Wlokas: A genetic algorithm-based method for the automatic reduction of reaction mechanisms. *Int. J. Chem. Kinet.* **46**, 41–59 (2014)

Topic 11: Reduction of reaction mechanisms 2: Elimination of redundant species

reaction rate and Jacobian-based methods for species removal,

species elimination via trial and error,
connectivity method (CM)

simulation error minimization connectivity method (SEM-CM)
directed relation graph (DRG) method
DRG-aided sensitivity analysis (DRG-SA),
DRG with error propagation (DRGEP)
path flux analysis method (PFA)

comparison of methods for species elimination

Identification of redundant species

important species and important features
- examples for important features: ignition delay time, period time
- alternative name: target species (DRG terminology)

The aim of the simulations is the accurate calculation of the important features and the concentrations of the important species

In different simulations the important species/features can be different even if the applied mechanism is identical.

necessary species
These are needed for the accurate calculation of the important features and the concentrations of the important species

redundant species
- not important and not necessary species
- can be eliminated from the mechanisms

T. Turányi: Reduction of large reaction mechanisms, *New J.Chem.*, **14**, 795-803 (1990)

Removal of redundant species 2

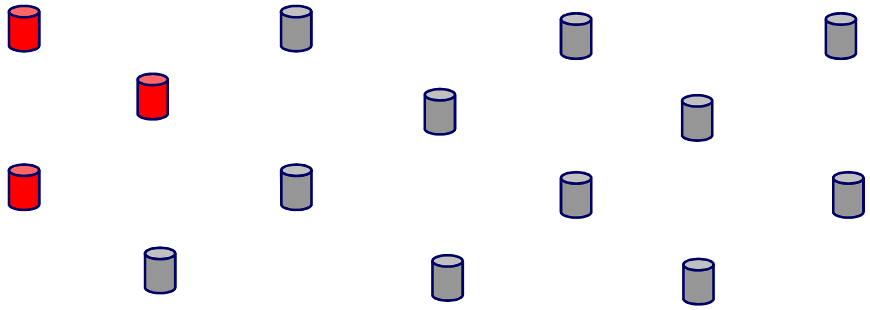
Approach 1: elimination of species one-by-one via trial calculations
The consuming reactions of each species is deleted one-by-one;
the effect on the important species/reactions is investigated.

Problem: not effective enough; some species can be eliminated only together (*e.g.* participating in fast preequilibrium reactions)
elimination of species groups ⇒ too many possibilities

Approach 2: Which species are directly linked to the important species?
(this approach cannot handle the reproduction of important features)

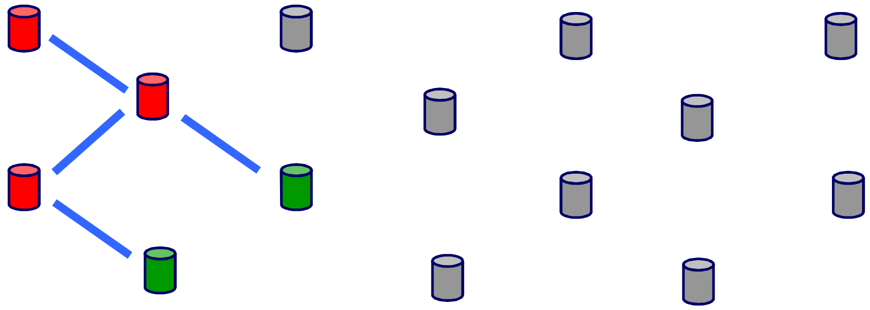
Step 1: using an appropriate method, the direct link of each species to the important species is investigated (detection of „species group 1”).
Step 2: assessment of the link of each species to the „species group 1”.
The newly identified closely lined species are added to the group, forming „species group 2”.

Step *n*: The iteration is repeated, until no species is closely linked to „species group (*n*-1)”



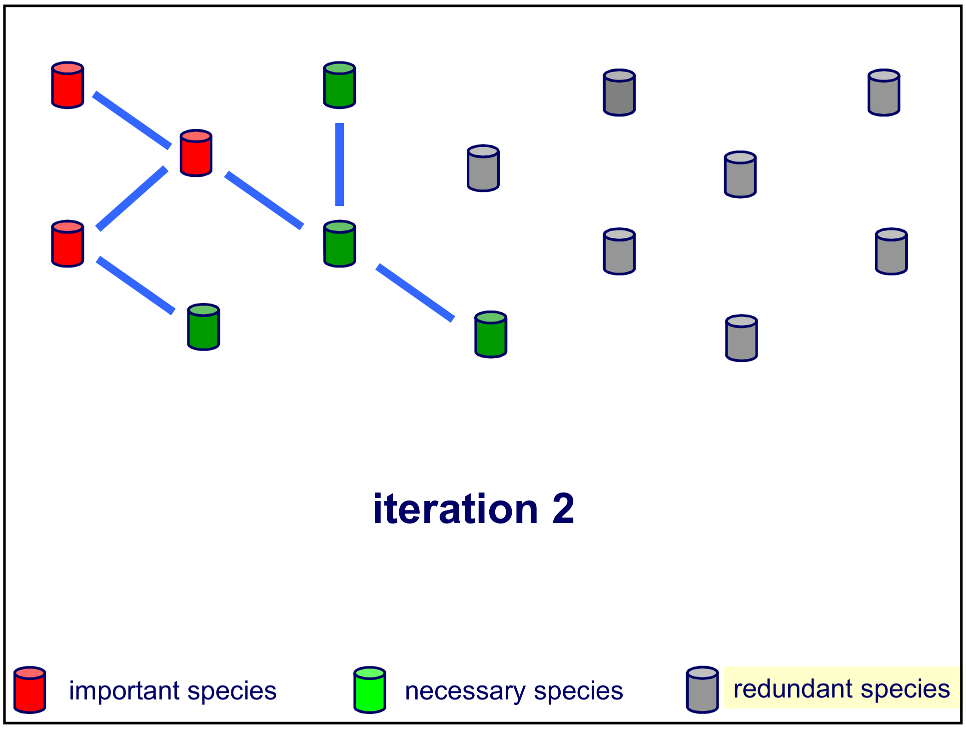
iteration 0

🟥 important species 🟩 necessary species ▨ redundant species



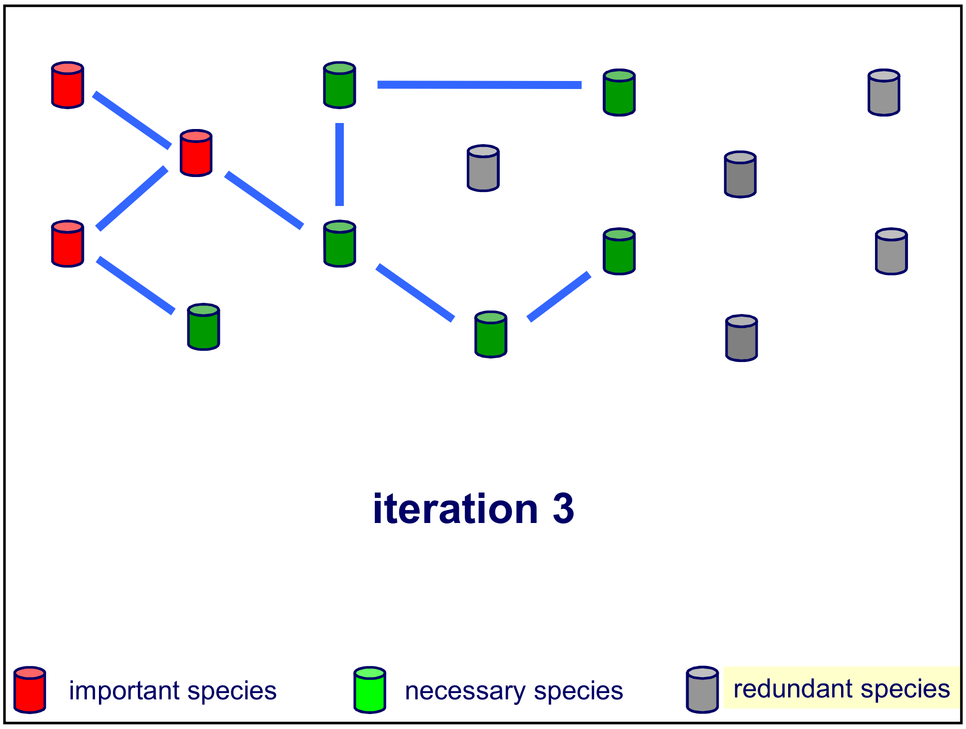
iteration 1

🟥 important species 🟩 necessary species ▨ redundant species



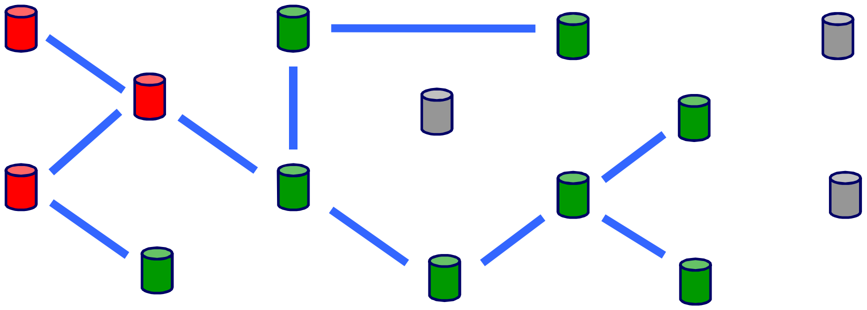
iteration 2

🟥 important species 🟩 necessary species ▨ redundant species



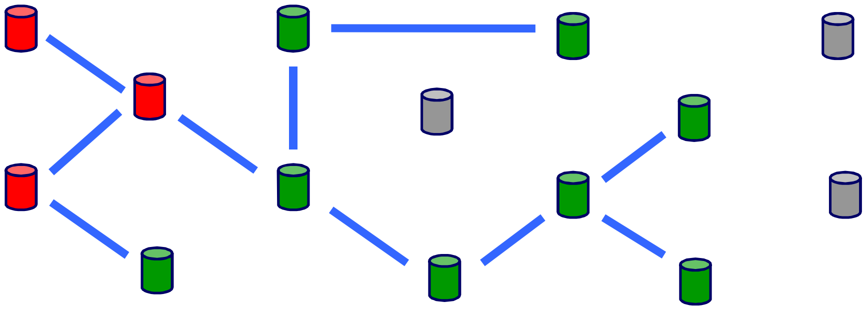
iteration 3

🟥 important species 🟩 necessary species ▨ redundant species



iteration 4

🟥 important species 🟩 necessary species ▨ redundant species



iteration 5

🟥 important species 🟩 necessary species ▨ redundant species

Connectivity Method

An element of the **normalized Jacobian** shows the effect of changing the concentration of species i on the production rate of species j:

$$\bar{\mathbf{J}}_{ij} = \left(c_i / f_j\right)\left(\partial f_j / \partial c_i\right)$$

B_i characterizes the strength of the direct link of species i to the group of important species:

$$B_i = \sum_{j \in \text{group}} \bar{\mathbf{J}}_{ij}^2$$

In an iterative process all species are identified that are strongly connected directly or through other species to the group of important species.

The procedure is repeated at each selected time.

T. Turányi: Reduction of large reaction mechanisms, *New J.Chem.*, **14**, 795-803 (1990)

Connectivity Method 2.

advantages and disadvanteges

Advantages:

- Simple, fast
- Available in KINALC
- Works well for small mechanisms

Disadvanteges:

The iteration is stopped, when a gap appers in the range of B_i values.
BUT
The best threshold is different from case to case.
The user has to select the threshold manually.

If there is a large number of species (500), there is no obvious gap!

The special role of important species diminishes after many iterations; this may lead to the selection of redundant species.

Directed Relation Graph (DRG) Method

The mechanism is represented by nodes and vertices.
Each species is a node.
The procedure starts from the important species (called here „target species”)
A vertex is drawn if the corresponding reaction has a significant contribution to the production rate of the species represented by the node.
This effect is characterized by the following measure:

$$R_{i \to j}^{(\mathrm{Lu})} = \frac{\sum_{\alpha \in C(i,j)} \left| \nu_{i\alpha} r_\alpha \right|}{\sum_{\alpha \in R(i)} \left| \nu_{i\alpha} r_\alpha \right|}$$

R(*i*)	set of reactions that are related to species *i*
C(*i*,*j*)	set of reactions in which both species *i* and *j* participate
$\upsilon_{i\alpha}$	stoichiometric coefficient of species *i* in reaction α
r_α	net reaction rate (the difference of the forward and backward rates)

T. Lu, C. K. Law: A directed relation graph method for mechanism reduction.
Proc. Comb. Inst., **30** (2005) 1333–1341.

DRG Method 2

Importance of species i: $I_i^{(\mathrm{DRG})} = \begin{cases} 1 & \text{if species } i \text{ is a target species} \\ \max\limits_{j \in S}\left(\min\left(R_{j \to i}, I_j^{(DRG)}\right)\right) & \text{otherwise} \end{cases}$

S	full set of chemical species
$R_{i \to j}$	connection weight
$R_{i \to j} = 0$	the two species are not connected

$I_i^{(\mathrm{DRG})}$ is calculated iteratively using a minimum-cost graph search algorithm

if $I_i^{(\mathrm{DRG})} < \varepsilon$ species *i* is considered to be redundant
for the simulation of the target species.

DRG+restart Method

tactic 1: getting a reduced mechanism in one step
by applying a large ε

tactic 2: getting a larger reduced mechanism by applying a smaller ε
- new simulation with the reduced mechanism
- new calculation of the r values
  (different r values since the mechanism is different!)
- new reduction with (the same or larger) smaller ε

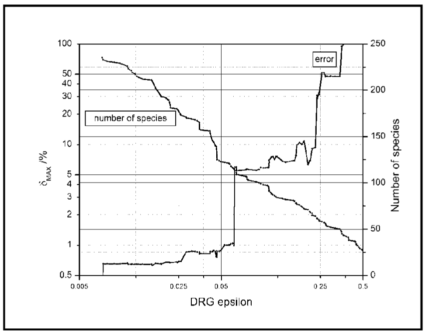
Lu and Law found that this two-stage approach (called DRG+restart) is more effective and allows the removal of further species at the second stage.

T. Lu, C. K. Law: Linear time reduction of large kinetic mechanisms with directed relation graph: *n*-heptane and *iso*-octane. *Combust. Flame*, **144** (2006) 24–36..

Directed Relation Graph (DRG) Methods

advantages

- simple, very fast



disadvanteges

- ε is not directly related to the error of the reduced mechanism
- decreasing ε does not always decrease the error of reduction
- very small fluxes of important species is also reproduced
  ⇒ selection of redundant species
- every selected species becomes equally important
  ⇒ selection of redundant species

DRGASA method

DRGASA: DRG-aided sensitivity analysis

misleading name:
- calculation of sensitivities is not included
- the DRG-estimation is checked using simulations

1 redundant species are selected by DRG using large ε

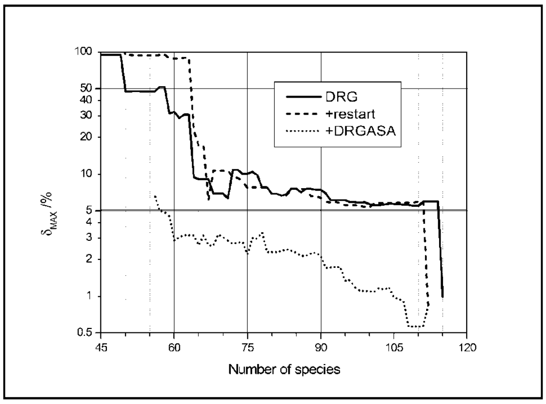
2 a second group of species is also identified using DRG with smaller ε

3 A series of simulations are carried out where the consequences of eliminating these species (one-by-one) are investigated using a series of simulations.

DRGASA method is more effective than the basic DRG approach, because it investigates the simulation error directly.
This simulation error belongs to the group of important species, and therefore the DRGASA indicates less species to be necessary than the original method for a prescribed error limit.

DRG vs. DRG-restart vs. DRGASA methods

Case study: reduction of a methane partial oxidation mechanism
full mechanism: 6874 reaction steps of 345 species
maximal simulation errors of the mechanisms as function of species number, for the original DRG DRG-restart and DRGASA methods



DRG+restart is slightly better (sometimes) than DRG

DRGASA is much better than DRG

T. Nagy, T. Turányi: Reduction of very large reaction mechanisms using methods based on simulation error minimization, *Combust. Flame*, **156**, 417–428 (2009)

Directed Relation Graph with Error Propagation: DRGEP method

Alternative definition of the weight of the vertices;
distinguishing reactions that create or destroy species *i*

$$R_{i\to j}^{(Pep)} = \frac{\left|\sum_{\alpha\in C(i,j)} \nu_{i\alpha} r_\alpha\right|}{\max\left(\sum_{\alpha\in R(i)} \left(\nu_{i\alpha} r_\alpha\right)^+, \sum_{\alpha\in R(i)} \left(\nu_{i\alpha} r_\alpha\right)^-\right)}$$

$(.)^+$ selects the positive terms
$(.)^-$ selects the negative terms

This equation calculates the sum of the rates belonging to the pair of species (*i*, *j*) to the total rate of formation or destruction of species *i*.

Importance index: $$I_i^{(\mathrm{DRGEP})} = \begin{cases} 1 & \text{if species } i \text{ is a target species} \\ \max\limits_{j\in S}\left(R_{j\to i} \cdot I_j^{(\mathrm{DRGEP})}\right) & \text{otherwise} \end{cases}$$

if $I_i^{(\mathrm{DRGEP})} < \varepsilon$ species *i* is considered to be redundant for the simulation of the target species.

P. Pepiot-Desjardins, H. Pitsch: An efficient error-propagation-based reduction method for large chemical kinetic mechanisms, *Combust. Flame*, **154**, 67-81 (2008)

DRGEP method 2

advantages

- simple, very fast

disadvanteges

- ε is not directly related to the error of the reduced mechanism
- decreasing ε does not always decrease the error of reduction
- very small fluxes of important species is also reproduced
  ⇒ selection of redundant species
- ~~every selected species becomes equally important~~
  ~~⇒ selection of redundant species~~

Path Flux Analysis (PFA) method

The first generation production (P_A) and consumption (C_A) fluxes of species *A*:

$$P_{\mathrm{A}} = \sum_i \max\left(\nu_{\mathrm{A},i}\,\omega_i, 0\right)$$

$$C_{\mathrm{A}} = \sum_i \max\left(-\nu_{\mathrm{A},i}\,\omega_i, 0\right)$$

υ_{Ai} stoichiometric coefficient of species A in reaction *i*

ω_i net reaction rate (the difference of the forward and backward rates)

The production (P_{AB}) and consumption (C_{AB}) fluxes of species A via species B

$$P_{\mathrm{AB}} = \sum_i \max\left(\nu_{\mathrm{A},i}\,\omega_i\,\delta_{\mathrm{B}}^i, 0\right)$$

$$C_{\mathrm{AB}} = \sum_i \max\left(-\nu_{\mathrm{A},i}\,\omega_i\,\delta_{\mathrm{B}}^i, 0\right)$$

δ_{B}^i is unity if species B is involved in the i-th reaction and 0 otherwise

W. T. Sun, Z. Chen, X. L. Gou, Y. G. Ju: A path flux analysis method for the reduction of detailed chemical kinetic mechanisms *Combust. Flame* **157**, 1298-1307 (2010)

PFA method 2

Flux ratio: share of a particular production and consumption path via species B to the total production and consumption flux of species A.
The first generation flux ratios for the
production and consumption of species A via species B are defined as:

$$r_{\mathrm{AB}}^{pro-1st} = \frac{P_{\mathrm{AB}}}{\max\left(P_{\mathrm{A}}, C_A\right)}$$

$$r_{\mathrm{AB}}^{con-1st} = \frac{C_{\mathrm{AB}}}{\max\left(P_{\mathrm{A}}, C_A\right)}$$

$$r_{\mathrm{AB}} = \max\left(r_{\mathrm{AB}}^{pro-1st}, r_{\mathrm{AB}}^{con-1st}\right)$$

Starting from the set of important species, using the relation $r_{\mathrm{AB}} > \varepsilon$,
the set of other necessary species are identified.

Common features of the CM/DRG/DRGEP/PFA methods

The method is based on the investigation of the kinetic system of ODEs.

The size of the reduced mechanism is controlled by a threshold (B_i or ε);
this threshold is not directly related to the error of simulation.

The analysis is carried out at several reaction times (concentration sets)
The final reduced mechanism is the union of the
reduced mechanisms obtained at the different concentration sets

Only one or few reduced mechanisms are produced and
therefore the error of reduction is checked only once or few times.

New strategy for mechanism reduction:

Simulation error minimization (SEM)

Definition of error calculation

relative error ⇒ overemphasizes large relative deviations of small concentrations

absolute error ⇒ deviations of small concentrations are neglected

mixed error:

$$\delta_\mathrm{i}(t_\mathrm{j}) = 2\frac{c_\mathrm{i}^\mathrm{red}(t_\mathrm{j}) - c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})}{c_\mathrm{i}^\mathrm{full}(t_\mathrm{j}) - c_\mathrm{i,MAX}^\mathrm{full}} \approx \begin{cases} \dfrac{c_\mathrm{i}^\mathrm{red}(t_\mathrm{j}) - c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})}{c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})} & \text{if } c_\mathrm{i}^\mathrm{full}(t_\mathrm{j}) \sim c_\mathrm{i,MAX}^\mathrm{full} \quad \approx \text{relative error} \\ \dfrac{c_\mathrm{i}^\mathrm{red}(t_\mathrm{j}) - c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})}{c_\mathrm{i,MAX}^\mathrm{full}/2} & \text{if } c_\mathrm{i}^\mathrm{full}(t_\mathrm{j}) << c_\mathrm{i,MAX}^\mathrm{full} \quad \approx \text{scaled absolute error} \end{cases}$$

where

$c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})$ concentration from the full mechanism

$c_\mathrm{i}^\mathrm{red}(t_\mathrm{j})$ concentration from the reduced mechanism

$c_\mathrm{i,MAX}^\mathrm{full} = \max_\mathrm{j} c_\mathrm{i}^\mathrm{full}(t_\mathrm{j})$ maximal value of c_i^{full}

definition of worst case error:

$$\delta_\mathrm{i,MAX} = \max_\mathrm{j} \left|\delta_\mathrm{i}(t_\mathrm{j})\right|$$

$$\delta_\mathrm{MAX} = \max_\mathrm{i} \delta_\mathrm{i,MAX}$$

Consistency of a reduced mechanism

The aim is to produce a consistent reduced mechanism.

A mechanism is called consistent, if all species are living.

living species:
has nonzero concentration
or
produced from another species
or
has inflow term (e.g. in PSR)

complementary set of species:
species that are not yet selected, but
selection of these species
yields at least one additional selected reaction

Note: unions of complementary sets are also complementary sets

Looking for strongly connected species - the problem of selecting important reactions

$A + B \rightarrow C \qquad D \rightarrow B$

A is the only important species
non-zero initial concentrations: A, D
zero initial concentrations: B, C

Proper reduced mechanism: $A + B \rightarrow C \qquad D \rightarrow B$

Connectivity method selects B and D as necessary
(or B only at the initial time).

Rule 1: a reaction is selected, if all of its species are necessary
⇒ some important reactions are not selected,
including $A + B \rightarrow C$

or

Rule 2: a reaction is selected, if any of its species is necessary
⇒ redundant reactions are also selected

SEM-CM steps

1 initiation

- simulations using the full mechanism
- selection of the representative time points
- saving concentration sets and Jacobian matrices in these points
- identification of the important species by the modeller
  $\Rightarrow$ initially these species are the selected species

2 identification of complementary sets of species

Complementary sets of species are looked for
to the group of currently selected species
by going through all reaction steps one-by-one.

These sets may contain each other or overlap.

SEM-CM

3 ranking the complementary sets

The strength of the direct link of the complementary set *k*,
to the group of selected species is characterized by

$$C_\mathrm{k} = \frac{1}{n_\mathrm{k}} \sum_{\mathrm{j} \in \mathrm{set}} B_\mathrm{j} = \frac{1}{n_\mathrm{k}} \sum_{\mathrm{j} \in \mathrm{set}} \sum_{\mathrm{i} \in \mathrm{group}} \bar{\mathbf{J}}_{\mathrm{ij}}{}^2$$

Each complementary set is ranked according to their C_k values.

4 generation of extended sets of species

Several complementary sets exist
with similarly strong links to the group of selected species.

Building procedure to depth level *m* generates *m* extended sets of species
by adding each complementary set up to rank *m*
to the current group of selected species.
This procedure is repeated at each reaction time t_k.

SEM-CM

5 generation of consistent reduced mechanisms

Non-living species are identified and the corresponding complementary sets are determined. A living species at time t_k is formed previously from living species. Maximum values of the previous Jacobians are used for ranking the complementary sets:

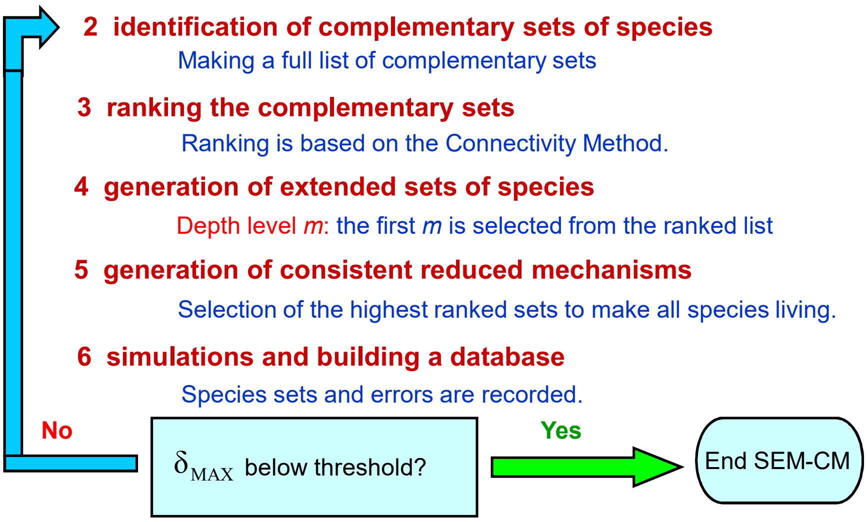
$$\overline{C}_k = \frac{1}{n_k}\sum_{j\in \text{set}}\overline{B}_j = \frac{1}{n_k}\sum_{j\in \text{set}}\ \sum_{\substack{i:\,\text{non-living}\\ \text{and selected}}}\overline{\mathbf{M}}_{ij}^{\,2} \qquad \overline{\mathbf{M}}_{ij}(t_k)^2 = \max_{l\le k}\mathbf{J}_{ij}(t_l)^2$$

Species belonging to the highest ranked complementary set are added to the group of selected species. This procedure is repeated, until all species become living.

6 simulations and building a database

Each generated reduced mechanism is investigated via a simulation. Reduction errors are recorded.

SEM-CM: The algorithm



1 initiation

Simulations with the full mechanism, saving **c** and **J**

2 identification of complementary sets of species

Making a full list of complementary sets

3 ranking the complementary sets

Ranking is based on the Connectivity Method.

4 generation of extended sets of species

Depth level *m*: the first *m* is selected from the ranked list

5 generation of consistent reduced mechanisms

Selection of the highest ranked sets to make all species living.

6 simulations and building a database

Species sets and errors are recorded.

SEM-PCAF: PCAF with simulation error minimization

Step 1: Identification of redundant reactions

Based on the PCAF method, many reduced mechanisms are generated by trying various thresholds for eigenvalues and eigenvector elements.

Step 2: Making the reduced mechanisms to be consistent

Some of the obtained reduced mechanisms contain not-living species.
Producing reactions of these species are restored.
The most important producing reactions (based on F-matrix analysis) are added.

Step 3: Finding the fastest reduced mechanism with small simulation error

Simulations are carried out; errors and required CPU time are recorded.
Many different reduced mechanisms may have similarly small error.

The reduced mechanism associated with the fastest simulation within a 2% margin of reduction error is selected as the best one.

Example:
gas-phase chemistry in solid-oxide fuel cells
the partial oxidation of methane

Solid-oxide fuel cells (SOFCs): power source for electric-driven vehicles.

Can be operated with hydrocarbon fuels.
Air is added to the hydrocarbon fuel to prevent deposit formation.
⇒ slow partial oxidation of the hydrocarbon before reaching the anode

Dean mechanism: homogeneous gas-phase chemistry in the anode channel of natural gas fuelled SOFCs.
Partial oxidation of methane up to high conversion.

Reduction is needed for **computer optimization** of fuel cell geometry and operating conditions.

Gas-phase chemistry in solid-oxide fuel cells: partial oxidation of methane 2

Full Dean mechanism: 345 species and 6874 irreversible reactions.

It was investigated at a typical set of SOFC conditions:

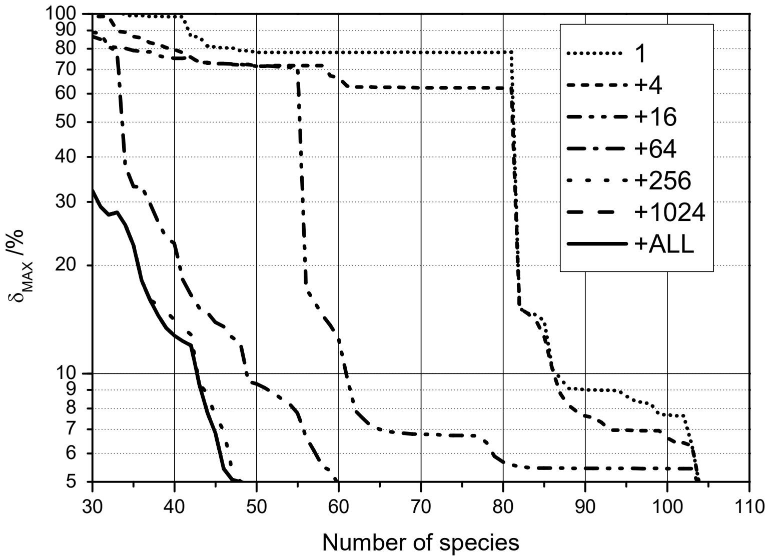
T = 900 ˚C (1173.15 K)
p = 1 atm (101325 Pa)

isothermal and isobaric conditions

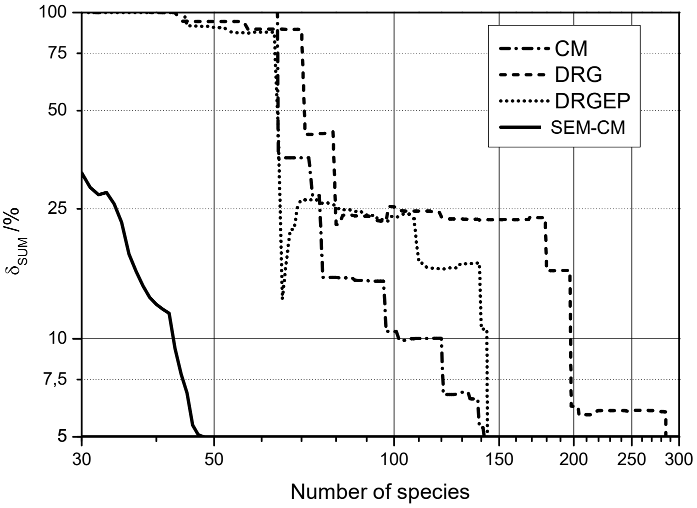
30.0 % v/v methane and 70.0 % v/v air

12 important species:
CH_4, N_2, O_2, H_2, H_2O, CH_2O, CO, CO_2, C_2H_2, C_2H_4, C_2H_6, C_6H_6.
(the mole fraction of these species exceed 0.001)

Testing the “depth level” of SEM-CM



Comparison of methods CM, DRG, DRGEP, SEM-CM simulation error *vs.* number of species

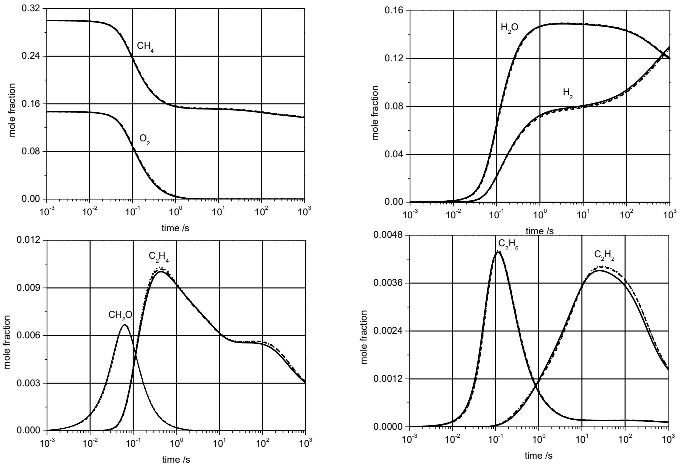


SOFC chemistry example: comparison of the obtained reduced mechanisms at 5% maximal error

original mechanism:	345 species	6874 reactions	
DRG reduction	286 species	5637 reactions	1.34 times faster
DRGEP reduction	144 species	2482 reactions	6.18 times faster
CM reduction	139 species	2494 reactions	5.57 times faster
SEM-CM reduction	47 species	613 reactions	58.4 times faster
SEM-CM + SEM-PCAF	47 species	297 reactions	**103.0 times faster**

CPU for the generation of the reduced mechanism (Athlon XP 2500+ PC):
(hh:mm:ss) DRG: 00:15:00 DRGEP: 00:04:40 CM: 00:01:30
SEM-CM(256): 9:29:00 SEM-CM(all): 18:10:00

Testing the obtained reduced mechanisms



_____	full mechanism	**6,874 reactions of 345 species**
- - - - -	result of SEM-CM	**613 reactions of 47 species**
…………	result of SEM-CM+ SEM-PCAF	**297 reactions of 47 species**

Almost identical concentrations can be obtained with a much smaller mechanism!

Summary

1. In all current methods for the elimination of redundant species and redundant reactions from a large reaction mechanism one or few reduced mechanisms are generated. The controlling parameter of the method is not directly related to the error of reduction.
2. New reduction philosophy: **SIMULATION ERROR MINIMIZATION** Thousands of candidate reduced mechanisms are generated in a guided way. The best mechanism (smallest reduction error and/or fastest simulation) is accepted.
3. **SEM-CM:** guided building up of a series of consistent reduced mechanisms, starting from the important species.
4. **SEM-PCAF:** optimized PCAF method for the elimination of redundant reactions
5. **SEM-CM** and **SEM-PCAF** together are very effective for the reduction of large reaction mechanisms

T. Nagy, T. Turányi: Reduction of very large reaction mechanisms using methods based on simulation error minimization, *Combust. Flame*, **156**, 417–428 (2009)

Topic 12: Reduction of reaction mechanisms 4: Lumping

reaction lumping;

species lumping:
- linear lumping,
- general nonlinear methods,
- chemical lumping,
- continuous lumping

Reaction lumping

Lumping paralel reaction pathways

A + B → C + D	$4k_1$
A + B → E + F	$6k_1$

Lumped reaction: A + B → 0,4 C + 0,4 D + 0,6 E + 0,6 F $10k_1$

It generates exactly the same kinetic system of ODEs
Computer time is not saved

Reaction lumping based on the rate limiting step

$$\begin{array}{lll} & \mathrm{A + B \rightarrow C + D} & v_1 = k_1ab \quad \text{slow} \Leftarrow \text{rate limiting step} \\ & \mathrm{D + E \rightarrow F} & \mathrm{v}_2 = k_2de \quad \text{fast} \\ + & \underline{\hphantom{\mathrm{A + B \rightarrow C + F - E}}} & \\ & \mathrm{A + B \rightarrow C + F - E} & v = k_1ab \end{array}$$

Less stiff ODEs, computer time is saved

Species lumping

The matematical definitions:

$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = \mathbf{f}(\mathbf{Y})$	Original kinetic system of ODEs, dimension of **Y** is *n*
$\frac{\mathrm{d}\hat{\mathbf{Y}}}{\mathrm{d}t} = \hat{\mathbf{f}}(\hat{\mathbf{Y}})$	ODE for the lumped variables dimension $n' \leq n$
$\hat{\mathbf{Y}} = \mathbf{h}(\mathbf{Y})$	Definition of lumped variables

h linear function ⇒ linear lumping
h nonlinear function ⇒ nonlinear lumping

identical solutions ⇒ exact lumping
almost identical solutions ⇒ approximate lumping

$\mathbf{Y}, \hat{\mathbf{Y}}$ some elements of these two vectors are identical
⇒ constrained lumping

Linear species lumping

$\hat{\mathbf{Y}} = \mathbf{M}\mathbf{Y}$	New variables are obtained by multiplying the original variable vector with a matrix
$\mathbf{Y} = \mathbf{M}^{-1}\hat{\mathbf{Y}}$	Regaining the original concentrations
$\mathbf{M}$	Lumping matrix (*n* x *n*')

Jacobian is constant
(= first order reactions only) ⇒ exact lumping is possible

Jacobian is not constant ⇒ no exact lumping
methods for approximate lumping
do not work well

Species lumping – a chemical approach

Lumping is frequently used in an intuitive way
„chemical approach"

If species with similar reactions and reactivity are present:

⇒ these species are lumped
⇒ concentration of the lumped species
= sum of the concentrations of the member species

called as the „*family method*" in atmospheric chemistry

A more involved approach:

the reactivity of the member species are different
concentration of the lumped species
= weighted sum of the concentrations of the member species

Species lumping – a chemical approach

R_1: $A_1 \rightarrow L_1$ $r_1 = k_1 [A_1]$
R_2: $A_2 \rightarrow L_2$ $r_2 = k_2 [A_2]$
R_3: $L_1 \rightarrow B_1$ $r_3 = k_3 [L_1]$
R_4: $L_2 \rightarrow B_2$ $r_4 = k_4 [L_2]$

Species lumping: $[L] = [L_1] + [L_2]$

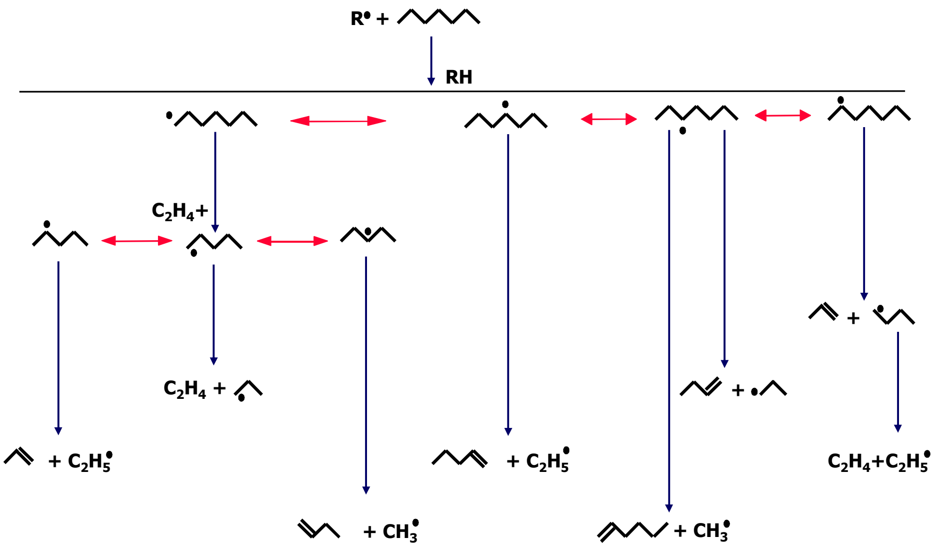
The mechanism after species lumping:

R_1: $A_1 \rightarrow L$ $r_1 = k_1 [A_1]$
R_2: $A_2 \rightarrow L$ $r_2 = k_2 [A_2]$
R_3: $L \rightarrow B_1$ $r_3 = k_{3m} [L]$
R_4: $L \rightarrow B_2$ $r_4 = k_{4m} [L]$

New rate coefficients:

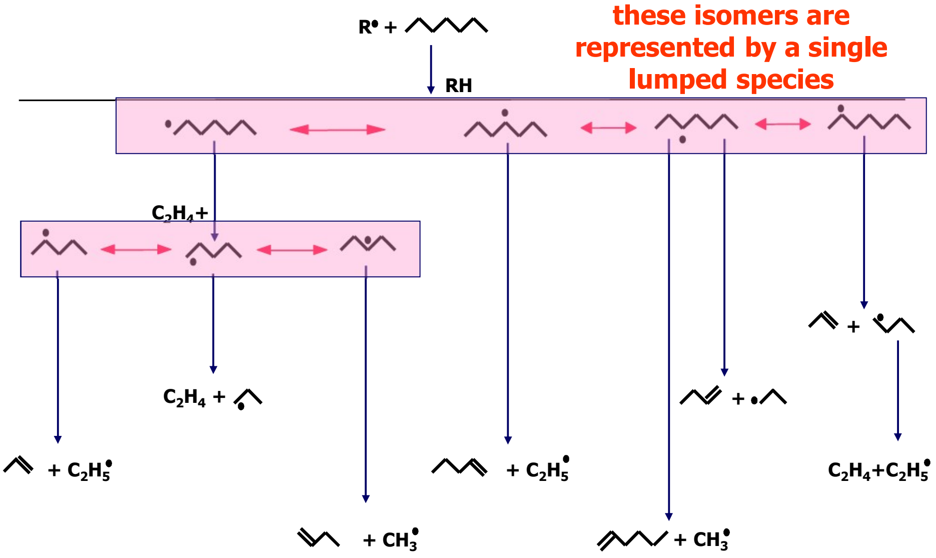
$k_{3m} = k_3 * [L_1]/[L]$
$k_{4m} = k_4 * [L_2]/[L]$

H abstraction reactions of *n*-heptane



Rice-Kossakoff decomposition mechanism

„Horizontal” species lumping: species with identical number of C-atoms are lumped



these isomers are represented by a single lumped species

The lumped reaction obtained

reaction rates of the parallel reactions can be calculated at a given temperature ⇒ stoichiometric coefficients of the lumped reaction

The obtained lumped reaction at 1000 K:

$$C_7H_{15} \rightarrow 0.17\ C_2H_4 + 0.17\ C_5H_{11} + 0.43\ C_3H_6 + 0.43\ C_4H_9 + 0.20\ C_4H_8 + 0.20\ C_3H_7 + 0.16\ C_5H_{10} + 0.16\ C_2H_5 + 0.04\ C_6H_{12} + 0.04\ CH_3$$

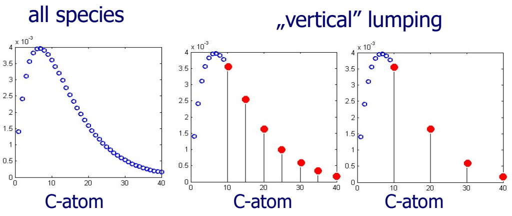
The stoichiometric coefficients of the lumped reaction change very little with changing temperature

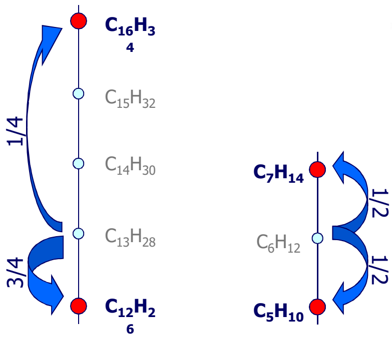
	800	**1000**	**1200**	**1500**
CH_3	0.03	0.04	0.044	0.045
C_2H_5	0.21	0.16	0.13	0.11
C_3H_7	0.18	0.20	0.21	0.23
C_4H_9	0.43	0.43	0.42	0.41
C_5H_{11}	0.15	0.17	0.196	0.205

E. Ranzi, M. Dente, A. Goldaniga, G. Bozzano, T. Faravelli: Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures. *Prog. Energy Combust. Sci.* **27**, 99-139 (2001)

„Vertical” species lumping

Similar types of species are identified (e.g. *n*-alkane, iso-alkane, alcene, …), and these are represented by surrogate species having similar ranges of carbon atoms.





n-heptane primary oxidation reactions

Detailed Scheme	Lumped Scheme
135 Primary reactions	15 Primary lumped reactions
38 Intermediate radicals	4 Intermediate lumped radicals
30 Primary products (retaining nC_7 structure)	4 Primary lumped products
3 n-heptenes	1 lumped n-heptene
8 cyclic-ethers	1 lumped cyclic-ether
4 hydroperoxides	1 lumped hydroperoxide
15 keto-hydroperoxides	1 lumped keto-hydroperoxide

Continous lumping

Continous species

species in a petroleum feedstocks, polimerisation systems
- very large numbers of species (several hundred thousands)
- can be ordered according to a chemical or physical feature (e.g. molecular weight)
- a feature is a continous function of the ordering variable (e.g. the melting point and the reactivity of the oligomers is a smooth function of the number of monomer units in the oligomer species.)

Using probability density function (*pdf*) of the feature
instead of the concentrations of the individual species

= the many discrete species are represented by a continuum

The chemical reactions modify this *pdf.*

R. Aris, G. R. Gavalas: On the theory of reactions in continuous mixtures *Philos. Trans. R. Soc.* **A260**, 351-393 (1966)

M.S. Okino, M.L. Mavrovouniotis: Simplification of mathematical models of chemical reaction systems. *Chem. Rev.* **98**, 391–408 (1998)

Topic 13: Reduction of reaction mechanisms 4: Time scales

history of the quasi steady-state approximation,

calculation of the local QSSA error,

interpretation of QSSA,

Computational Singular Perturbation (CSP)

reduction of models in reaction kinetics with direct calculation of slow manifolds (ILDM),
reaction diffusion manifolds (REDIM)

repro-modelling

History of the quasi-steady state approximation

1913 – 1960: analytical solution of the kinetic ODEs

Bodenstein (1913): analytical solution to the equations of the H_2/Br_2 reaction system
Szemjonov (1939): applied the QSSA for a part of the intermediates only

1960 – 1971: conversion of the stiff ODEs to non-stiff ODEs

Edelson (1973) claimed that the stiff ODEs can already be solved directly (no need for the QSSA); nobody can assess the error introduced by the application of the QSSA ⇒ QSSA should be banned

1971 – : production of skeleton models from detailed mechanims for solving PDEs

theoretical study of the QSSA method: (about 60 articles)
main lines:
- justification of the application of QSSA for small mechanisms
- singular perturbation: analytical investigation of small systems

An early general article and its commented English translation.
We developed this line further, see the following coming pages.

D. A. Frank-Kamenetskii: Условия примениности метода Боденштейна в химической кинетике (Conditions for the applicability of the Bodenstein method in chemical kinetics)
Ж. Физ. Хим. (*Zh. Fiz. Him.*) **14**, 695-700 (1940)

T. Turányi, J. Tóth: Comments to an article of Frank-Kamenetskii on the quasi-steady-state approximation. *ACH Models In Chemistry* **129**, 903-907 (1992)

Quasi-steady-state approximation (QSSA)

The original kinetic system of differential equations:

$$\mathrm{d}\mathbf{c}/\mathrm{d}t = \mathrm{f}(\mathrm{c},\mathrm{k}), \qquad \mathbf{c}(0) = \mathbf{c}_0$$

The concentration vector is divided to two parts:

$\mathbf{c}^{(1)}$ concentration vector of non-QSSA species

$\mathbf{c}^{(2)}$ concentration vector of QSSA species

The Jacobian:

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}^{(11)} & \mathbf{J}^{(12)} \\ \mathbf{J}^{(21)} & \mathbf{J}^{(22)} \end{pmatrix} = \begin{pmatrix} \partial\mathbf{f}^{(1)}/\partial\mathbf{c}^{(1)} & \partial\mathbf{f}^{(1)}/\partial\mathbf{c}^{(2)} \\ \partial\mathbf{f}^{(2)}/\partial\mathbf{c}^{(1)} & \partial\mathbf{f}^{(2)}/\partial\mathbf{c}^{(2)} \end{pmatrix}$$

Quasi-steady-state approximaton:

$$\mathrm{d}\mathbf{c}^{(1)}/\mathrm{d}t = \mathrm{f}^{(1)}(\mathrm{c},\mathrm{k})$$

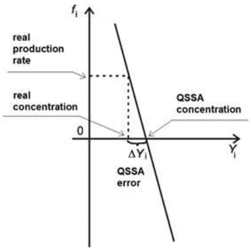
$$\mathbf{0} = \mathrm{f}^{(2)}(\mathrm{c},\mathrm{k})$$

⇐ denote $\mathbf{C}^{(2)}$ the concentration vector calculated from the algebraic equation

$$\mathbf{c}(0) = \mathbf{c}_0$$

The local error of the QSSA

$\Delta\mathbf{c}^{(2)} = \mathbf{c}^{(2)} - \mathbf{C}^{(2)}$ Local error of the QSSA



Taylor expansion of the production rate
of the QSSA species at the QSSA concentrations:

$$\frac{\mathrm{d}\mathbf{c}^{(2)}}{\mathrm{d}t} = \underbrace{\mathbf{f}^{(2)}(\mathbf{c},\mathbf{k})\big|_{\mathbf{c}=\mathbf{C}}}_{0} + \mathbf{J}^{(22)}\,\Delta\mathbf{c}^{(2)}$$

Calculation of the local error for several QSSA species:

$$\mathrm{d}\mathbf{c}^{(2)}/\mathrm{d}t = \mathbf{J}^{(22)}\,\Delta\mathbf{c}^{(2)}$$

Calculation of the local error for a single QSSA species:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = J_{ii}\Delta c_i \qquad \Rightarrow \qquad -\Delta c_i = \left(\frac{-1}{J_{ii}}\right)\frac{\mathrm{d}c_i}{\mathrm{d}t}$$

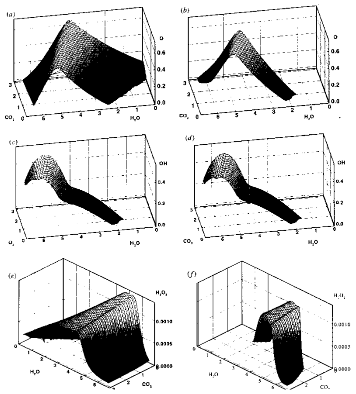
error of QSSA approximation = lifetime × production rate of the species

T. Turányi, A. S. Tomlin, M. J. Pilling: On the error of the quasi-steady-state approximation. J. Phys. Chem. **97**, 163-172 (1993)

Example: QSSA error at methane pyrolysis

```
The QSSA local error of each species at 50 s

                         relative error      absolute error

   19. CH3               -6.200E-01 %      -2.631E-13 mole/cm**3
   20. CH3CH2CH3         -5.284E-01 %      -4.568E-14 mole/cm**3
   21. CH.=CHCH2CH3       5.050E-01 %       3.022E-21 mole/cm**3
   22. (CH3)2C=CH.        4.889E-01 %       1.708E-19 mole/cm**3
   23. C2H5               4.013E-01 %       2.989E-16 mole/cm**3
   24. CH≡CCH2.          -3.385E-01 %      -4.109E-14 mole/cm**3
   25. CH2=CHCH=CH2       3.062E-01 %       8.733E-13 mole/cm**3
   26. CH.=CHCH3          3.030E-01 %       4.148E-18 mole/cm**3
   27. C2H3               2.612E-01 %       1.241E-16 mole/cm**3
   28. CH2=CHCH2.         2.134E-01 %       1.153E-14 mole/cm**3
   29. CH≡CCH3           -2.040E-01 %      -1.283E-11 mole/cm**3
   30. CH2                1.591E-01 %       1.041E-20 mole/cm**3
   31. C                 -1.366E-01 %      -8.633E-30 mole/cm**3
   32. CH2=C.CH3          1.364E-01 %       5.094E-18 mole/cm**3
   33. CH2S               1.025E-01 %       5.811E-23 mole/cm**3
   34. H                  7.083E-02 %       1.699E-16 mole/cm**3
   35. CH2=C=CH2         -6.929E-02 %      -1.209E-12 mole/cm**3
```

Direct calculation of slow manifolds



$\mathbf{y}^M$ a point „on the surface” of the manifold

$\mathbf{f}(\mathbf{y}^M)$ velocity of the point
„on the surface” of the manifold

$\mathbf{W}_s$ tangent plane of the manifold
in this point (includes vector $\mathbf{f}$)

$\mathbf{W}_f$ vector of fast directions
linearly independent of $\mathbf{W}_s$

$\mathbf{f}(\mathbf{y}^M)$ this vector is orthogonal
to the $\mathbf{w}_f$ vectors:

$$\mathbf{W}_f\ \mathbf{f}\left(\mathbf{y}^M\right) = \mathbf{0}$$

1. Selection of the dimension (denote n_D)
2. Selection of n_D variables used for parameterization (vector $\mathbf{x}$).
   The equation above defines all other concentrations that
   belong to the manifold as a function of vector $\mathbf{x}$.

PROBLEM: numerically ill-conditioned task, because vectors $\mathbf{W}_f$ have frequently almost identical direction.

Numerical determination of slow manifolds

Numerically more stable is the
Schur decomposition of the Jacobian:

$$\mathbf{Q}^{\mathrm{T}}\mathbf{J}\mathbf{Q} = \begin{pmatrix} \mathbf{J}'^{(11)} & \mathbf{J}'^{(12)} \\ 0 & \mathbf{J}'^{(22)} \end{pmatrix}$$

eigenvalues of matrix $\mathbf{N}_s$ $\lambda_1, \lambda_2, \ldots, \lambda_m$
eigenvalues of matrix $\mathbf{N}_f$ $\lambda_{m+1}, \ldots, \lambda_n$ and $\mathrm{Re}(\lambda_i) \geq \mathrm{Re}(\lambda_j)$ if $i < j$,
thus $\mathbf{Q}_f$ belongs to the fast modes.

The following equation is solved: $\mathbf{Q}_f(\mathbf{Y})\,\mathbf{f}(\mathbf{Y}) = 0$

ILDM: Intrinsic Low-Dimensional Manifolds

Using the ILDM in simulations: a look-up table contains
- the change of parameterizing variables in time ($\mathrm{d}\,\mathbf{Y}_{\mathrm{p}}/\mathrm{d}t$)
- the related values of all other concentrations

simulations =
solution of the small ODE + search in the table

REDIM: reaction diffusion manifolds

In most of these slides the reduction of
spatially homogeneous chemical kinetic systems is discussed.

However, in some systems the chemical reactions and
the effect of chemical flows have to be simulated together:

$$\frac{\partial\boldsymbol{\psi}}{\partial t} = \mathbf{F}(\boldsymbol{\psi}) - \vec{\boldsymbol{\upsilon}}\cdot\mathrm{grad}\boldsymbol{\psi} + \frac{1}{\rho}\mathrm{div}\mathbf{D}\,\mathrm{grad}\boldsymbol{\psi}$$

$\boldsymbol{\psi}$	is the vector of thermokinetic state with elements specific enthalpy h, the pressure P, mass fractions w_i
F	chemical source term
$\vec{\boldsymbol{\upsilon}}$	flow velocity
ρ	density
D	matrix of transport coefficients

V. Bykov, U. Maas: The extension of the ILDM concept to reaction-diffusion manifolds. *Combust. Theory Model.* **11**, 839-862 (2007)

REDIM: reaction diffusion manifolds 2

Possible cases:

1 fast chemistry – slow transport:
the concentration changes are dictated by the chemistry only

2 fast chemistry – moderately fast transport:
the manifolds are dictated by the chemistry
with perturbation by the diffusion

U. Maas, S. B. Pope: Laminar flame calculations using simplified chemical kinetics based on intrinsic low-dimensional manifolds. *Proc. Combust. Inst.* **25**, 1349-1356 (1994)

3 fast chemistry – fast transport:
a joint manifold has to determined,
dictated by both the chemistry and the diffusion
⇒ REDIM: reaction-diffusion manifolds

U. Maas, V. Bykov: The extension of the reaction/diffusion manifold concept to systems with detailed transport models. *Proc. Combust. Inst.* **33**, 1253-1259 (2011)

Computational Singular Perturbation (CSP)

Assume that there are M fast modes in an n-variate dynamical system
Denote Ω an $(n\text{-}m)$-dimensional manifold

vectors $\mathbf{a}_i$, $i = 1, \dots, M$ span the fast subspace
vectors $\mathbf{a}_i$, $i = M+1, \dots, n$ span the manifold
$\mathbf{A}_r = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m]$
$\mathbf{A}_s = [\mathbf{a}_{m+1}, \mathbf{a}_{m+2}, \dots, \mathbf{a}_n]$
production rates: $\mathbf{f} = \mathbf{f}_{\text{fast}} + \mathbf{f}_{\text{slow}}$

fast directions of production rates: $\mathbf{f}_{\text{fast}} = \mathbf{A}_r\ \mathbf{z}$
production rates along the surface: $\mathbf{f}_{\text{slow}} = \mathbf{A}_s\ \mathbf{z}$

fast amplitudes $\mathbf{z}_{\text{fast}} = \mathbf{B}_r\ \mathbf{f}$
slow amplitudes $\mathbf{z}_{\text{slow}} = \mathbf{B}_s\ \mathbf{f}$
Matrices **A** and **B** are orthogonal

S. H. Lam, D. A. Goussis: Understanding complex chemical kinetics with computational singular perturbation. *Proc. Combust. Inst.* **22**, 931-941 (1988)

CSP 2

Matrices **A** and **B** are obtained by an iterative „refinement procedure” from the Jacobian

The value of the amplitude allows sorting the modes:

z always zero — conserved property

in classic kinetics: *e.g.* element conservation

z was large ⇒ now *z* almost zero — exhausted mode

in classic kinetics: QSSA, partial equilibrium

z almost zero ⇒ may become large — dormant mode

in classic kinetics: pool component on the actual time scale

z large now — active mode

movement of the trajectory along the manifold

CSP 3

The CSP analysis is based on the application of a series of pointers:

1 CSP Participation Index

z_k^m denotes the contribution of the *k*-th reaction step to the *m*-th fast amplitude:

$$z^m = z_1^m + z_2^m + \ldots + z_{N_R}^m \approx 0$$

$$z^m = \mathbf{b}^m \, \mathbf{f}, \qquad m = 1, \ldots, M$$ amplitude of the *m*-th fast mode

$$z_k^m = \left(\mathbf{b}^m \, \mathbf{v}_k\right) r_k$$

ν_k stoichiometric vector of reaction *k*
r_k rate of reaction *k*

$$P_k^m = \frac{z_k^m}{\sum_j^{N_r} \left| z_j^m \right|}$$ **CSP Participation Index**

The sum of the absolute values of P_k^m is equal to unity. A relatively large P_k^m value indicates that the *k*-th reaction step is a significant participant in the *m*-th equilibrium.

D. A. Goussis, P. D. Kourdis: Glycolysis in *saccharomyces cerevisiae*:
Algorithmic exploration of robustness and origin of oscillations. *Math. Biosci.* **243**, 190–214 (2013)

CSP 4

2 CSP Importance Index

contribution of the *k*-th reaction step to the evolution of the *n*-th variable on the manifold

$$f^{n}_{slow} = f^{n,1}_{slow} + f^{n,2}_{slow} + \ldots + f^{n,N_r}_{slow}$$

$$f^{n,k}_{slow} = \sum_{j=M+1}^{N_S} a^{n}_{j} \left(\mathbf{b}^{j} \mathbf{v}_{k}\right) r_k, \quad k = 1, \ldots, N_r$$

a^n_j denotes the *n*-th element of column vector $\mathbf{a}_j$ in matrix $\mathbf{A}_s$

$$I^{n}_{k} = \frac{f^{n,k}_{slow}}{\sum_{j}^{N_r} \left| f^{n,k}_{slow} \right|}$$

CSP Importance Index

The sum of the absolute values of I_k^n is equal to unity. A relatively large I_k^n value indicates that the *k*-th reaction step has a significant contribution to the change of the *n*-th variable on the manifold.

CSP 5

3 CSP Pointer (earlier name: „radical pointer”)

Identification of variables (*i.e.* species concentrations) that have a large contribution to the exhausted modes.

$$\mathbf{D}_m = \mathrm{diag}\left[\mathbf{a}_m \ \mathbf{b}^m\right]$$

CSP Pointer

A value of D^i_m close to unity indicates that the *i*-th variable is strongly connected to the *m*-th mode and its corresponding time-scale.

D^i_m identifies the QSS-species and
the non-QSS-species participating in fast equilibria

Application of CSP for mechanism reduction: a non-stiff reduced model can be obtained that well describes the change of modes belonging to the characteristic time-scale of the system.

M. Valorani, F. Creta, D. A. Goussis, J. Lee, H. Najm: An automatic procedure for the simplification of chemical kinetic mechanisms based on CSP. *Combust. Flame* **146**, 29–51 (2006)

Repromodelling

Stages of reduction of a detailed reaction mechanism

original stiff ODE, many variables **SLOW SIMULATION**

→ making a skeleton mechanism
(by elimination of redundant species and reactions, lumping)
FASTER

skeleton mechanism → manifold based mathematical model (ODE)

skeleton mechanism → difference equations obtained by the repromodelling method **FASTEST**

Principle of repromodelling:
The chemical kinetic model is simulated several thousand times at different conditions. A polynomial is fitted to the simulation results. In the further simulations the fitted polynomial is used, instead of solving again the differential equations.

Repromodel from skeleton models

We have a skeleton mechanism having few species.
The kinetic simulation has a natural timescale and natural time step Δt.
The solutions of the ODE (calculated concentrations) are printed at every Δt times:
$\mathbf{c}(t)$, $\mathbf{c}(t+\Delta t)$, $\mathbf{c}(t+2\Delta t)$, $\mathbf{c}(t+3\Delta t)$, *etc.*

Building a database from the simulation results:
$\mathbf{c}(t) \rightarrow \mathbf{c}(t+\Delta t)$
$\mathbf{c}(t+\Delta t) \rightarrow \mathbf{c}(t+2\Delta t)$
$\mathbf{c}(t+2\Delta t) \rightarrow \mathbf{c}(t+3\Delta t)$

Fitting function **G** to the data: $\mathbf{c}(t+\Delta t)= \mathbf{G}(\mathbf{c}(t))$
This is a system of difference equations.
Recursive application of **G** results in concentartion-time functions with resolution Δt

number of variables of function **G** = number of variables of the skeleton model
Fast calculations, because the integrated solution is stored.

Repromodel from detailed mechanisms

Presence of a slow manifold:
The trajectory quickly approaches a n_1 dimensional slow manifold
The location of the points of the n_1-dimensional manifold in the n-dimensional space of concentrations can be parameterized by vector $\mathbf{x}$ of n_1 elements:

function $\mathbf{G}_2$: $\mathbf{c}(t) = \mathbf{G}_2(\ \mathbf{x}(\mathrm{t})\ )$

Movement on the manifold during time Δt is characterized by function $\mathbf{G}_1$:

$$\mathbf{x}(t+\Delta t) = \mathbf{G}_1\ (\mathbf{x}(t))$$

functions $\mathbf{G}_1$ and $\mathbf{G}_2$ can be determined by repro-modelling:

- Selection of Δt and n_1
- Building a database from the simulation results of the detailed mechanism:

$$\mathbf{c}(t) \rightarrow \mathbf{c}(t+\Delta t)$$
$$\mathbf{c}(t+\Delta t) \rightarrow \mathbf{c}(t+2\Delta t)$$
$$\mathbf{c}(t+2\Delta t) \rightarrow \mathbf{c}(t+3\Delta t)$$

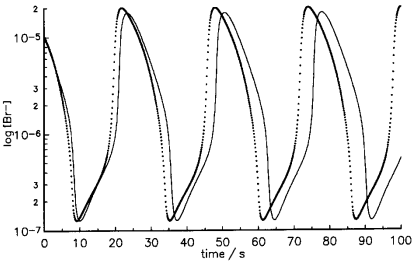
- Fitting functions $\mathbf{G}_1$ and $\mathbf{G}_2$ to the data.

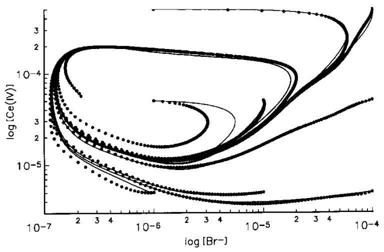
Original detailed mechanim: n variables; $\mathbf{G}_2$ repromodel: n_1<n variables
The calculation is much faster, because the repromodel has fewer variables AND because of the integrated results are fitted (*e.g.:* $n = 50$, $n_1 = 3$)

Example 1: the Oregonator model

3-variable skeleton model of the Belousov-Zhabotinskii reaction
200 simulations started from the neighborhood of the limit cycle
Result: 20 thousand $\mathbf{c}(t)$, $\mathbf{c}(t+\Delta t)$ datasets

Up to 8th order, 3-variate polynomials were fitted using the Gram−Schmidt orthonormalization process





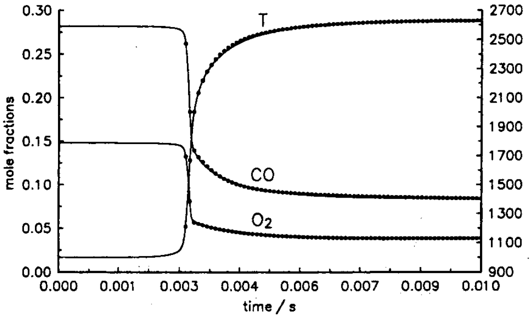
The repromodel can be calculated 60 times faster than the solution of the ODE of the Oregonator model

Example 2: ignition of CO/H_2/air mixtures

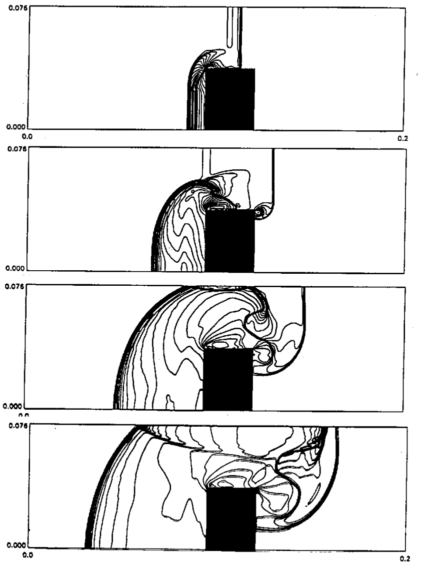
Detailed mechanism: 67 reactions of 13 species
The ignition can be characterized by a 3-dimensional manifold (Maas and Pope (1992))

φ=0.5–1.5, H:C=1:10, T=990–1010 K; 300 random initial compositions
$\Delta t = 10^{-4}$ s time step, $t_{final} = 0.01$ s
The database contains 30000 entries.
A 4th order polynomials were fitted to the change of temperature and the mole fractions of CO and O_2.

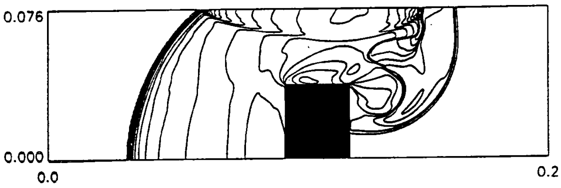
The repromodel calculated 11700 times faster
the change of T, X_{CO} and X_{O2} during the ignition of CO/H_2/air mixtures



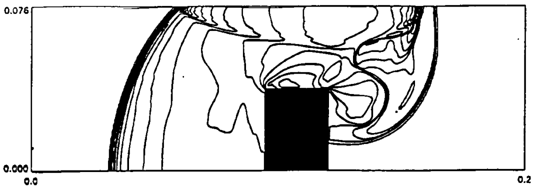
Example 3: Spread of detonation wave in a H_2/O_2/Ar mixture



Spread of detonation wave in $H_2/O_2/Ar$ mixture



Calculated density map based on a detailed hydrogen combustion mechanism (9 variables)



Calculated density map based on a repromodel (polynomial approximation with 3 variables)
100 times faster

Overview of mechanism reduction methods

I. without time scale analysis

1. determination of a skeleton mechanism (a part of the original one)
   elimination of redundant species and reactions
   ⇒ smaller kinetic system of ODEs

2. species lumping and reaction lumping
   ⇒ smaller kinetic system of ODEs

II. using time scale analysis

1. classic methods: QSSA and partial equilibrium
   ⇒ smaller kinetic system of ODEs or
   coupled differerential/algebraic equations
2. slow manifolds
   ⇒ smaller kinetic system of ODEs
3. repromodelling
   ⇒ difference equations

Topic 14: Computer codes for the study of complex reaction systems

general simulation codes in reaction kinetics,

simulation of gas kinetics systems,

analysis of reaction mechanisms,

investigation of biological reaction kinetic systems,

codes for global uncertainty analysis,

ReSpecTh Information Site

Reaction kinetics simulation codes

WINPP/XPP Windows simulation code
solving systems of ODEs, DAEs and PDEs.
The user has to provide the rate equations ⇒ applicable for small systems only
http://www.math.pitt.edu/~bard/classes/wppdoc/readme.htm

KPP: Kinetic Preprocessor http://people.cs.vt.edu/~asandu/Software/Kpp/
production of the kinetic ODE from the reaction mechanism
numerical solution of stiff ODEs; sparse matrix routines

V. Damian, A. Sandu, M. Damian, F. Potra, G. R. Carmichael:
The Kinetic PreProcessor KPP - A software environment for solving chemical kinetics.
Comp. Chem. Eng. **26**, 1567-1579 (2002)

SUNDIALS: SUite of Nonlinear and DIfferential/ALgebraic equation Solvers
https://computation.llnl.gov/casc/sundials/main.html

	MATLAB interface to the following solvers:
CVODE	solves initial value problems for ordinary differential equation (ODE) systems
CVODES	solves ODE systems and includes sensitivity analysis capabilities
ARKODE	solves initial value ODE problems with additive Runge-Kutta methods
IDA	solves initial value problems for differential-algebraic equation (DAE) systems
IDAS	solves DAE systems and includes sensitivity analysis capabilities
KINSOL	solves nonlinear algebraic systems.

CHEMKIN

Developed at the SANDIA National Laboratories, Livermore, CA, USA
CHEMKIN-I (1975-1985)
CHEMKIN-II (1985-1995)
Simulation codes: SENKIN, PSR, PREMIX, SHOCK, EQLIB
+ utility programs, data bases
FORTRAN codes, controlled by the input files

> Kee R. J., Rupley F. M., Miller J. A.
> CHEMKIN-II: A FORTRAN *Chemical Kinetics Package for the Analysis of Gas-Phase Chemical Kinetics*
> SANDIA report No. SAND79-8009B

Reaction Design www.reactiondesign.com (1995-)
Commertial codes; source code is not provided

Chemkin 3.x,
Graphical User Interface (GUI) to CHEMKIN-II
Chemkin 4.x
really new solvers, graphical interface, versatile
Chemkin Pro
Chemkin + additional utility codes (*e.g.* pathway plotting)

CHEMKIN simulation codes

www.reactiondesign.com

CHEMKIN → CHEMKIN -II → CHEMKIN 3 → CHEMKIN 4 → CHEMKIN PRO

CHEMKIN (1975–)	classified code
CHEMKIN-II (1986–)	classified code, then freeware
since CHEMKIN 3 (1996–)	commercial code

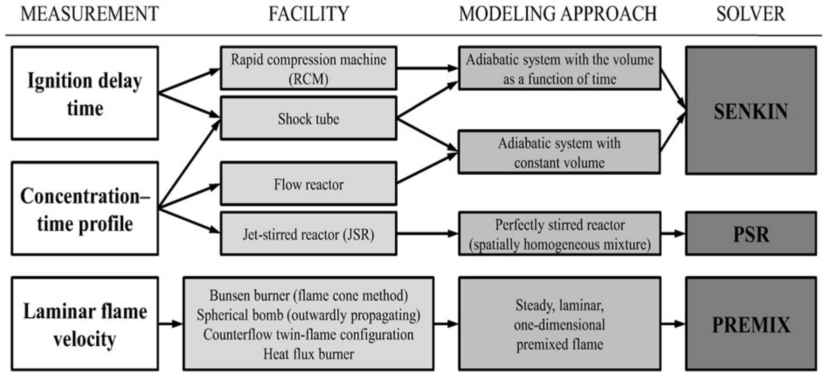
CHEMKIN-II simulation codes:

SENKIN	spatially homogeneous reactions
PREMIX	laminar premixed flames
SHOCK	shock tube simulations
PSR	perfectly stirred reactor simulations

Options of SENKIN:

- adiabatic system, constant p pressure
- adiabatic system, constant V volume
- adiabatic system, $V(t)$ function
- closed system, constant p, T
- closed system, constant V, T
- closed system, $p(t)$ and $T(t)$ function

CHEMKIN-II simulation codes



MEASUREMENT | FACILITY | MODELING APPROACH | SOLVER

Ignition delay time

Concentration–time profile

Laminar flame velocity

Rapid compression machine (RCM)

Shock tube

Flow reactor

Jet-stirred reactor (JSR)

Bunsen burner (flame cone method)
Spherical bomb (outwardly propagating)
Counterflow twin-flame configuration
Heat flux burner

Adiabatic system with the volume as a function of time

Adiabatic system with constant volume

Perfectly stirred reactor (spatially homogeneous mixture)

Steady, laminar, one-dimensional premixed flame

SENKIN

PSR

PREMIX

Relation between the
- measured quantity
- experimental setup
- modelling approach
- the corresponding CHEMKIN-II simulation code

C. Olm, I. Gy. Zsély, R. Pálvölgyi, T. Varga, T. Nagy, H. J. Curran, T. Turányi: Comparison of the performance of several recent hydrogen combustion mechanisms. *Combust. Flame*, **161**, 2219-2234 (2014).

Alternatives to CHEMKIN

Cantera www.cantera.org
Open source code, available from SourceForge.net
chemical equilibrium, homogeneous and heterogeneous kinetics
reactor networks, 1D flames

Kintecus www.kintecus.com
Excel workbook; free for academic use
Simulation of combustion, atmospheric chemical and biological systems

COSILAB www.softpredict.com
commertial combustion simulation and mechanism analysis code
- visualization of reaction pathways
- reduction of kinetic mechanisms
- simulation of reactor networks
- two-dimensional reactors and flames
- spray and dust flames

Alternatives to CHEMKIN 2

OpenSmoke++ (Milano; http://www.opensmoke.polimi.it/)
free for academic use
numerical modelling of laminar reacting flows
built on the OpenFOAM framework
homogeneous reactions, heterogeneous reactions on catalytic surfaces

FlameMaster (Aachen; http://www.itv.rwthaachen.de/downloads/flamemaster/)
free for academic use

- homogeneous reactor and perfectly stirred reactor calculations
- freely propagating premixed flames
- steady counter-flow diffusion flames

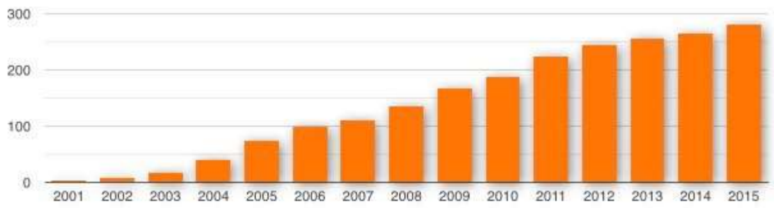
ZeroRK (LLNL, Livermore, USA, Ca; https://github.com/LLNL/zero-rk)
Zero-RK (Zero-Order Reaction Kinetics v3.5) is a software package that simulates chemically reacting systems using sparse, preconditioned, adaptive matrix methods to achieve orders-of-magnitude reduction in simulation time while maintaining accurate results.

SBML

SBML: Systems Biology Markup Language http://sbml.org/

SBML model definition format was created
to promote the exchange of systems biology models
(similar to the role of CHEMKIN format in gas kinetics)

281 SBML-compatible software packages are available (January 2016)
The list of these simulation codes can be looked at http://sbml.org/



Increase of the number of SBML-based computer codes
(these include both academic and commertial codes)

Copasi

COPASI (COmplex PAthway SImulator)
http://copasi.org/

S*imulation* and ***analysis*** of biochemical network ***models***.
Free, support, but source code is not provided.

Homogeneous kinetic systems in interacting compartments
Import and export of models in the SBML format (levels 1 to 3).
Export of models in many format (XPP, C code, Latex).

- ODE-based and stochastic simulatons
- stoichiometric analysis of the reaction networks
- optimization of models; parameter estimation
- local sensitivity analysis.
- time scale separation analysis
- characterization of non-linear dynamics properties (oscillations and chaos)

S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, U. Kummer: COPASI — a COmplex PAthway SImulator. *Bioinformatics* **22**, 3067-3074 (2006)

Global uncertainty analysis codes

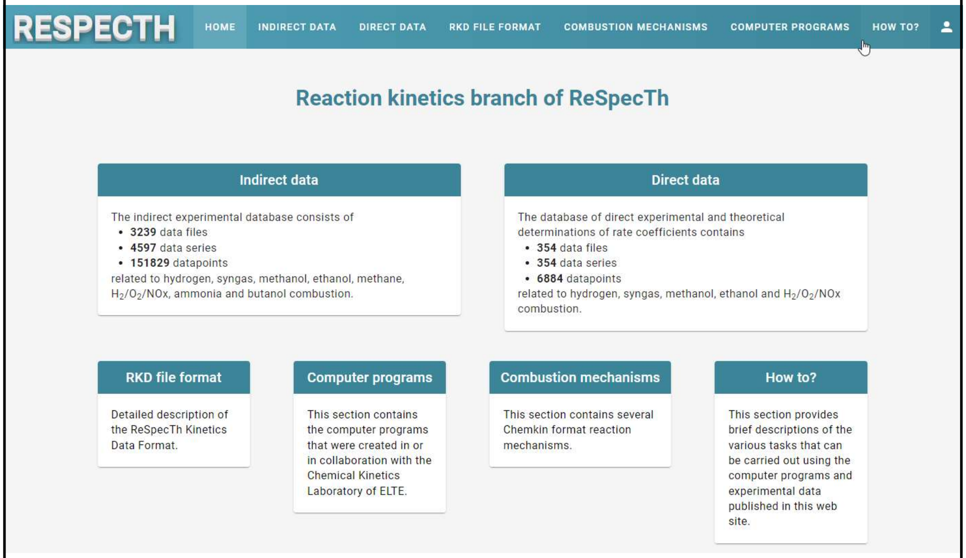
GUI-HDMR http://www.gui-hdmr.de
The GUI-HDMR software is based on the RS-HDMR approach, where all component functions are approximated by orthonormal polynomials using random (or quasi-random) samples. Calculation of up to second-order global sensitivity indices based on user supplied sets of input/output data. The component functions are approximated by up to 10th order orthonormal polynomials.

T. Ziehn, A. S. Tomlin: GUI-HDMR - A software tool for global sensitivity analysis of complex models *Environmental Modelling & Software*, **24**, 775-785 (2009)

SimLab https://ec.europa.eu/jrc/en/samo/simlab
Developed at the EC Joint Research Centre (EC-JRC) in Ispra, Italy.
Versions up to 2.2: GUI based nice education tool
(1) generation of random or quasi-random parameter sets
(2) running the models (within SimLab or externally)
(3) processing of the simulation results (FAST, Morris' and Sobol methods)
visualisation of the outcome of uncertainty/sensitivity analyses.

SimLab versions from 3.0:
subroutine can be called from Fortran, Python, C++, or Matlab

Reaction kinetics branch of the ReSpecTh web site https://respecth.hu



Reaction kinetics branch of ReSpecTh

Indirect data

The indirect experimental database consists of
- **3239** data files
- **4597** data series
- **151829** datapoints

related to hydrogen, syngas, methanol, ethanol, methane, $H_2/O_2/NOx$, ammonia and butanol combustion.

Direct data

The database of direct experimental and theoretical determinations of rate coefficients contains
- **354** data files
- **354** data series
- **6884** datapoints

related to hydrogen, syngas, methanol, ethanol and $H_2/O_2/NOx$ combustion.

RKD file format

Detailed description of the ReSpecTh Kinetics Data Format.

Computer programs

This section contains the computer programs that were created in or in collaboration with the Chemical Kinetics Laboratory of ELTE.

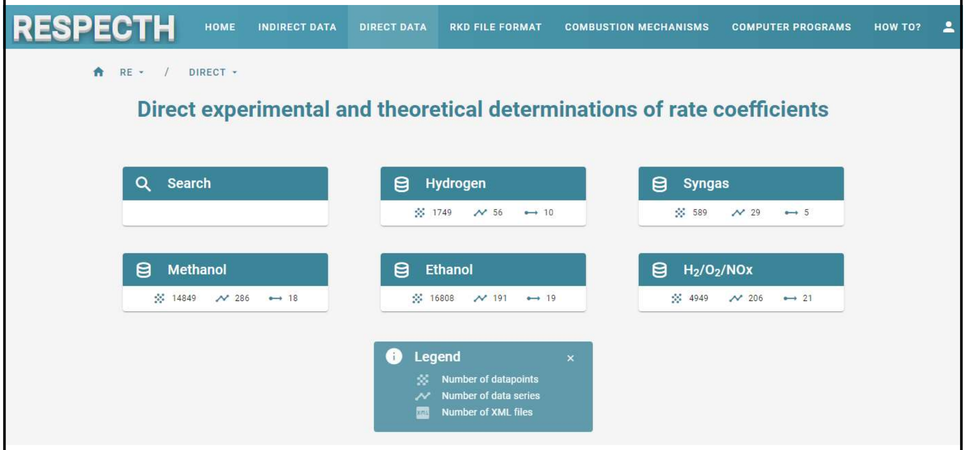
Combustion mechanisms

This section contains several Chemkin format reaction mechanisms.

How to?

This section provides brief descriptions of the various tasks that can be carried out using the computer programs and experimental data published in this web site.

Experimental and theoretical rate determinations

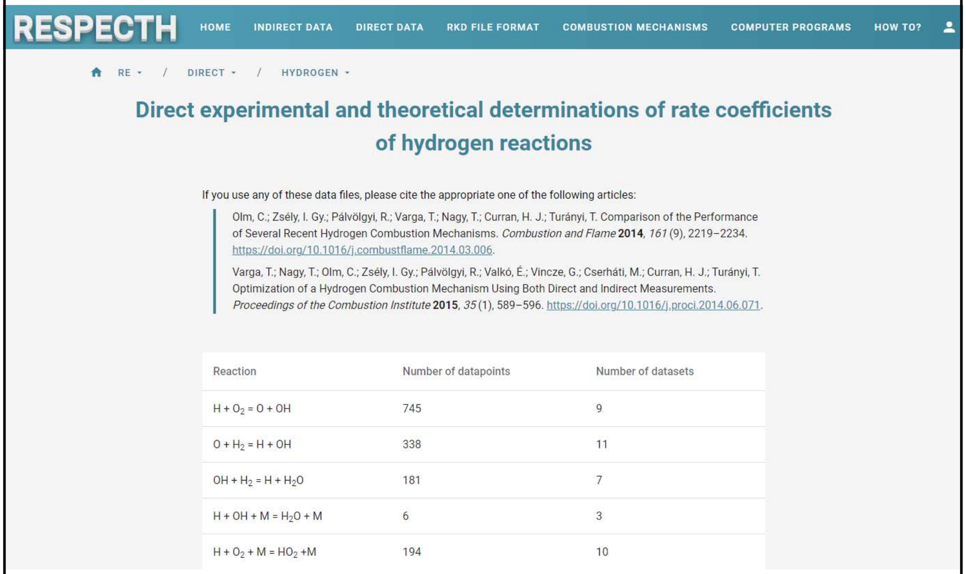


Direct experimental and theoretical determinations of rate coefficients

Hydrogen	1749	56	10
Syngas	589	29	5
Methanol	14849	286	18
Ethanol	16808	191	19
$H_2/O_2/NOx$	4949	206	21

Legend
Number of datapoints
Number of data series
Number of XML files

Experimental and theoretical rate determinations – Hydrogen



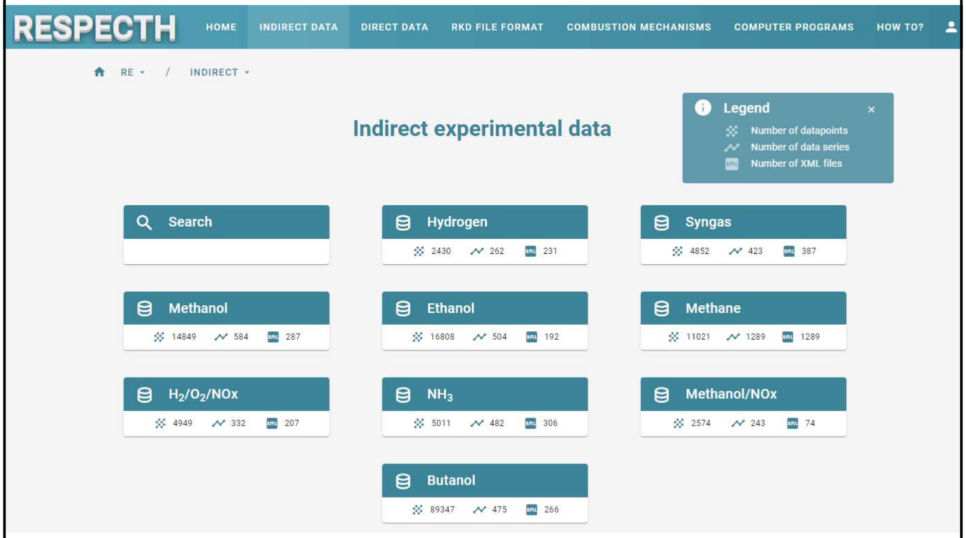
Direct experimental and theoretical determinations of rate coefficients of hydrogen reactions

If you use any of these data files, please cite the appropriate one of the following articles:

> Olm, C.; Zsély, I. Gy.; Pálvölgyi, R.; Varga, T.; Nagy, T.; Curran, H. J.; Turányi, T. Comparison of the Performance of Several Recent Hydrogen Combustion Mechanisms. *Combustion and Flame* **2014**, *161* (9), 2219–2234. https://doi.org/10.1016/j.combustflame.2014.03.006.
>
> Varga, T.; Nagy, T.; Olm, C.; Zsély, I. Gy.; Pálvölgyi, R.; Valkó, É.; Vincze, G.; Cserháti, M.; Curran, H. J.; Turányi, T. Optimization of a Hydrogen Combustion Mechanism Using Both Direct and Indirect Measurements. *Proceedings of the Combustion Institute* **2015**, *35* (1), 589–596. https://doi.org/10.1016/j.proci.2014.06.071.

Reaction	Number of datapoints	Number of datasets
$H + O_2 = O + OH$	745	9
$O + H_2 = H + OH$	338	11
$OH + H_2 = H + H_2O$	181	7
$H + OH + M = H_2O + M$	6	3
$H + O_2 + M = HO_2 + M$	194	10

Indirect combustion experimental data



Indirect experimental data

Legend
- ⁘ Number of datapoints
- ⟋ Number of data series
- ▣ Number of XML files

Hydrogen ⁘ 2430 ⟋ 262 ▣ 231

Syngas ⁘ 4852 ⟋ 423 ▣ 387

Methanol ⁘ 14849 ⟋ 584 ▣ 287

Ethanol ⁘ 16808 ⟋ 504 ▣ 192

Methane ⁘ 11021 ⟋ 1289 ▣ 1289

$H_2/O_2/NOx$ ⁘ 4949 ⟋ 332 ▣ 207

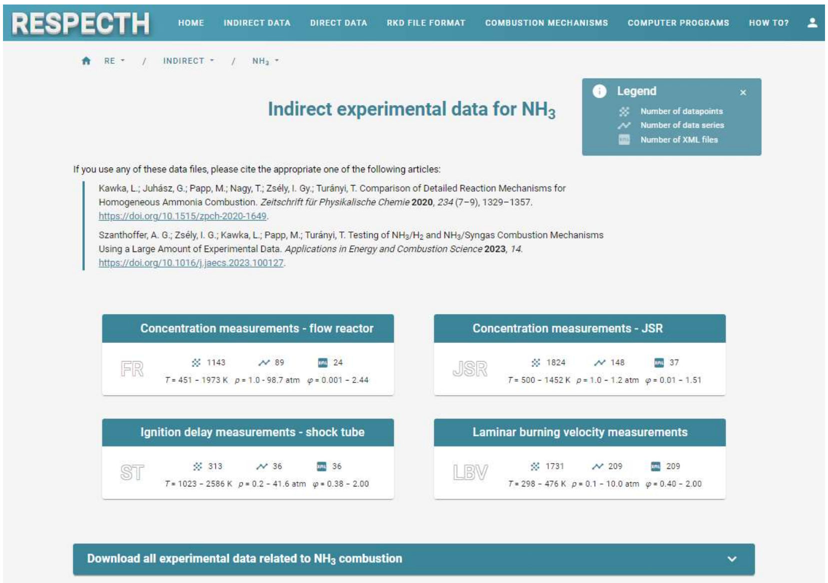
NH_3 ⁘ 5011 ⟋ 482 ▣ 306

Methanol/NOx ⁘ 2574 ⟋ 243 ▣ 74

Butanol ⁘ 89347 ⟋ 475 ▣ 266

Indirect combustion experimental data

Experimantal types – Ammonia



Indirect experimental data for NH_3

If you use any of these data files, please cite the appropriate one of the following articles:

> Kawka, L.; Juhász, G.; Papp, M.; Nagy, T.; Zsély, I. Gy.; Turányi, T. Comparison of Detailed Reaction Mechanisms for Homogeneous Ammonia Combustion. *Zeitschrift für Physikalische Chemie* **2020**, *234* (7−9), 1329−1357. https://doi.org/10.1515/zpch-2020-1649.
>
> Szanthoffer, A. G.; Zsély, I. G.; Kawka, L.; Papp, M.; Turányi, T. Testing of NH_3/H_2 and NH_3/Syngas Combustion Mechanisms Using a Large Amount of Experimental Data. *Applications in Energy and Combustion Science* **2023**, *14*. https://doi.org/10.1016/j.jaecs.2023.100127.

Concentration measurements - flow reactor
1143 89 24
$T = 451 - 1973$ K $p = 1.0 - 98.7$ atm $\varphi = 0.001 - 2.44$

Concentration measurements - JSR
1824 148 37
$T = 500 - 1452$ K $p = 1.0 - 1.2$ atm $\varphi = 0.01 - 1.51$

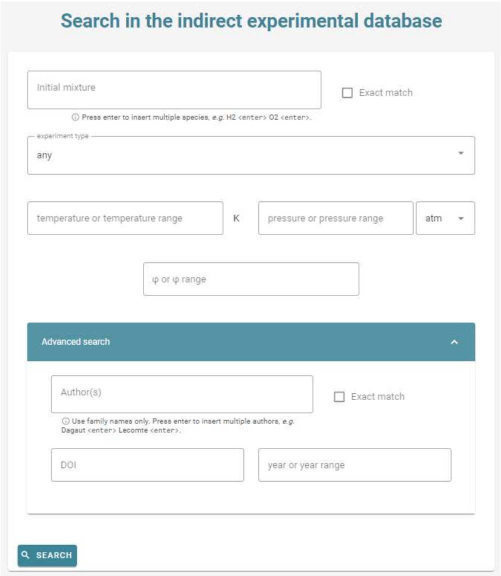
Ignition delay measurements - shock tube
313 36 36
$T = 1023 - 2586$ K $p = 0.2 - 41.6$ atm $\varphi = 0.38 - 2.00$

Laminar burning velocity measurements
1731 209 209
$T = 298 - 476$ K $p = 0.1 - 10.0$ atm $\varphi = 0.40 - 2.00$

Download all experimental data related to NH_3 combustion

Indirect combustion experimental data

Search engine



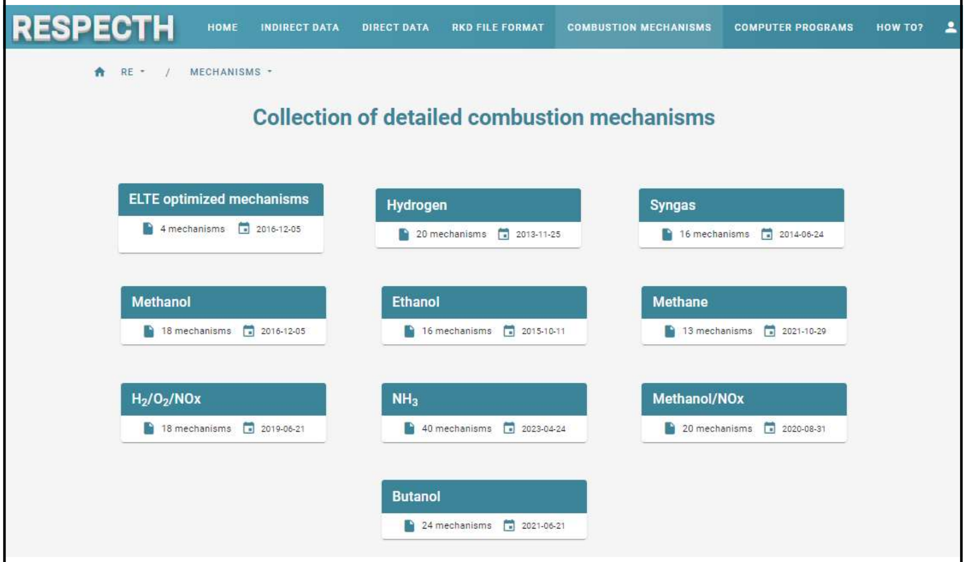
Search in the indirect experimental database

Press enter to insert multiple species, *e.g.* H2 <enter> O2 <enter>.

experiment type: any

Use family names only. Press enter to insert multiple authors, *e.g.* Dagaut <enter> Lecomte <enter>.

Combustion mechanisms



Collection of detailed combustion mechanisms

ELTE optimized mechanisms
4 mechanisms 2016-12-05

Hydrogen
20 mechanisms 2013-11-25

Syngas
16 mechanisms 2014-06-24

Methanol
18 mechanisms 2016-12-05

Ethanol
16 mechanisms 2015-10-11

Methane
13 mechanisms 2021-10-29

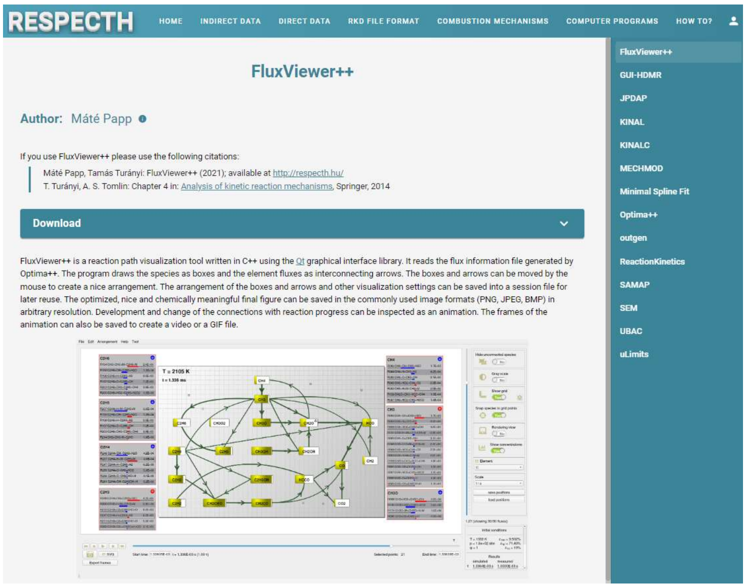
H_2/O_2/NOx
18 mechanisms 2019-06-21

NH_3
40 mechanisms 2023-04-24

Methanol/NOx
20 mechanisms 2020-08-31

Butanol
24 mechanisms 2021-06-21

Computer programs



FluxViewer++

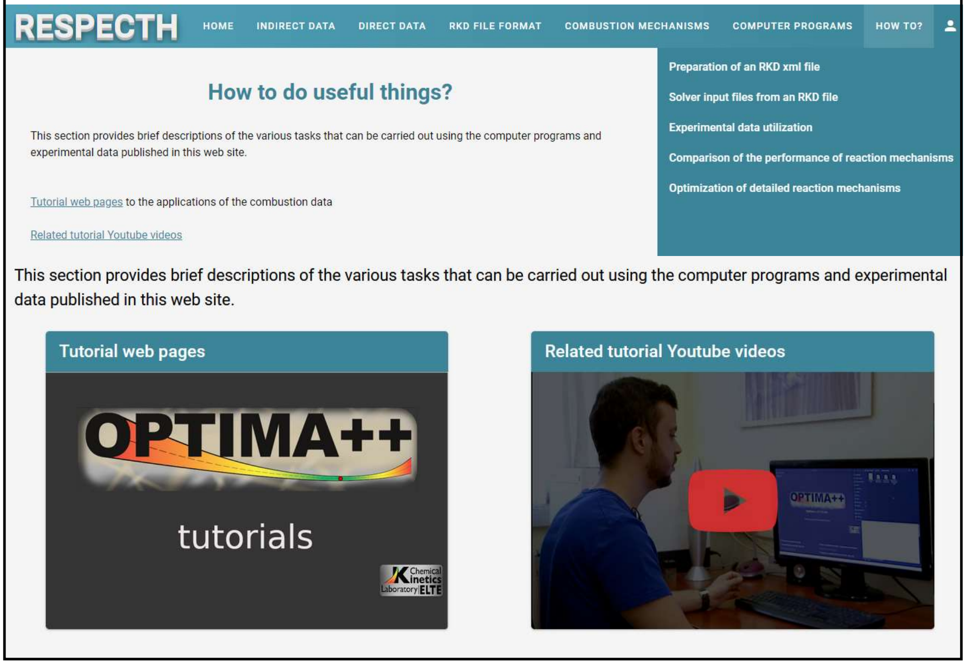
Author: Máté Papp

If you use FluxViewer++ please use the following citations:

> Máté Papp, Tamás Turányi: FluxViewer++ (2021); available at http://respecth.hu/
> T. Turányi, A. S. Tomlin: Chapter 4 in: Analysis of kinetic reaction mechanisms, Springer, 2014

FluxViewer++ is a reaction path visualization tool written in C++ using the Qt graphical interface library. It reads the flux information file generated by Optima++. The program draws the species as boxes and the element fluxes as interconnecting arrows. The boxes and arrows can be moved by the mouse to create a nice arrangement. The arrangement of the boxes and arrows and other visualization settings can be saved into a session file for later reuse. The optimized, nice and chemically meaningful final figure can be saved in the commonly used image formats (PNG, JPEG, BMP) in arbitrary resolution. Development and change of the connections with reaction progress can be inspected as an animation. The frames of the animation can also be saved to create a video or a GIF file.

How to ...



How to do useful things?

This section provides brief descriptions of the various tasks that can be carried out using the computer programs and experimental data published in this web site.

Tutorial web pages to the applications of the combustion data

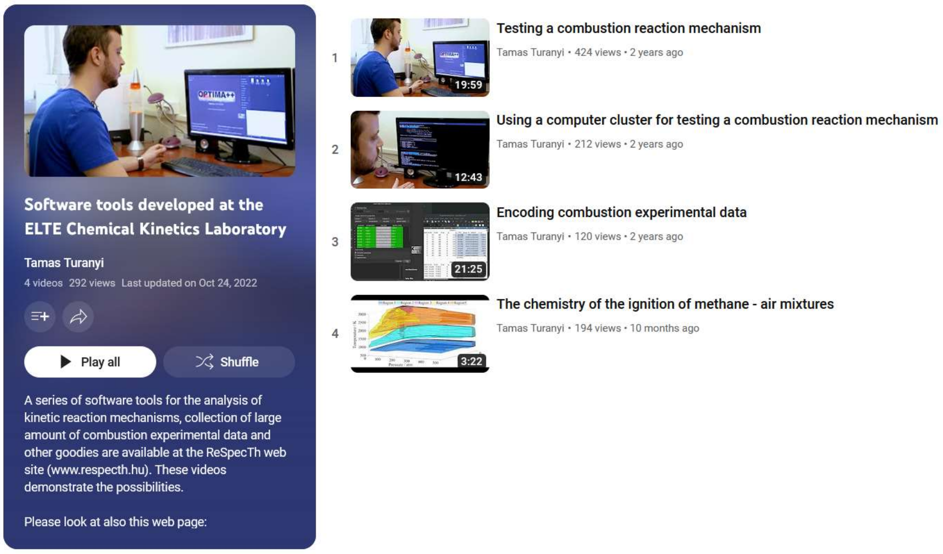
Related tutorial Youtube videos

This section provides brief descriptions of the various tasks that can be carried out using the computer programs and experimental data published in this web site.

Tutorial web pages

Related tutorial Youtube videos

Tutorial videos



Software tools developed at the ELTE Chemical Kinetics Laboratory

Tamas Turanyi

4 videos 292 views Last updated on Oct 24, 2022

A series of software tools for the analysis of kinetic reaction mechanisms, collection of large amount of combustion experimental data and other goodies are available at the ReSpecTh web site (www.respecth.hu). These videos demonstrate the possibilities.

Please look at also this web page:

1 19:59 Testing a combustion reaction mechanism
Tamas Turanyi • 424 views • 2 years ago

2 12:43 Using a computer cluster for testing a combustion reaction mechanism
Tamas Turanyi • 212 views • 2 years ago

3 21:25 Encoding combustion experimental data
Tamas Turanyi • 120 views • 2 years ago

4 3:22 The chemistry of the ignition of methane - air mixtures
Tamas Turanyi • 194 views • 10 months ago

Thank you for
your attention!

Supported by the Higher Education Restructuring Fund allocated to ELTE
by the Hungarian Government