Optimization of Plasma-Surface Kinetics via Machine Learning Methods

José Afonso, Vasco Guerra and Pedro Viegas
Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
josefafonso@tecnico.ulisboa.pt
pedro.a.viegas@tecnico.ulisboa.pt

Introduction

Optimizing plasma-surface kinetics is a critical challenge in material science, often hindered by computationally expensive simulations. Moreover, a key challenge is the uncertainty of model parameters, such as energy barriers, for some chemical reactions, which are not well-defined in the literature [1,2]. This work addresses this challenge by employing a data-driven optimization approach, minimizing the discrepancy between the physical model's predictions and experimental data.
We also analyze how well the data constrains each optimized parameter, which provides further confidence in the model's predictions and helps guide future experiments to target the most important sources of uncertainty.

Physical Model: Surface Kinetics on Pyrex

This work uses a mesoscopic model to describe the interaction between a O_2/CO_2 glow discharge and the Pyrex reactor wall. The model focuses on the adsorption and recombination of reactive oxygen and carbon species [1,2,3]. The experimental data comes from [2,4] and it corresponds to 225 data points. They include discharges with different O_2 and CO_2 mixtures, pressures within the interval 0.2-10 Torr, discharge currents between 10 and 40 mA and wall temperatures between -20°C and 50°C.

Surface Sites and Mechanisms

The Pyrex surface is described by three types of adsorption sites:

- **Physisorption (F_v)**: Weak, reversible van der Waals bonds covering the entire surface, allowing for particle diffusion and desorption [1,2].
- **Chemisorption (S_v)**: Strong, irreversible chemical bonds at surface defects, with a much lower density than physisorption sites [1,2].
- **Metastable (S_v^*)**: Highly reactive, reversible sites created by ion bombardment at low pressures (<1 Torr) [3].

Two main recombination mechanisms [1] occur:

- **Eley-Rideal (E-R)**: A gas-phase particle reacts directly with an adsorbed particle:

$$A(g) + A_{(F/S)} \rightarrow A_2 + (F/S)_v$$

- **Langmuir-Hinshelwood (L-H)**: A diffusing physiosorbed particle reacts with another adsorbed particle:

$$A_F + A_{(F/S)} \rightarrow A_2(g) + (F/S)_v + F_v$$

Reaction Scheme

It is composed of 62 chemical reactions, where the following components are highlighted:

Oxygen species (O/O_2) kinetics

The model includes the recombination of atomic oxygen (O) into O_2 or ozone (O_3) [2].

- At low wall temperatures ($T_w < 25$°C), the most important mechanisms are:

$$\text{L-H: } O_F + O_F \rightarrow O_2(g) + 2\,F_v \text{ and E-R: } O(g) + O_F \rightarrow O_2(g) + F_v)$$

- At higher temperatures ($T_w > 25$°C), reactions involving chemisorbed oxygen become dominant.

$$O(g) + O_S \rightarrow O_2(g) + S_v \text{ and } O_F + O_S \rightarrow O_2(g) + S_v + F_v$$

Carbon species (CO/CO_2) kinetics

The oxygen scheme is extended to include carbon species, which are critical in CO_2 plasmas [4].

- CO can both physisorb and chemisorb, while the stable CO_2 molecule only physisorbs.
- Main effect: CO competes with O for reactive chemisorption sites. This passivation of the surface by CO significantly reduces the overall recombination rate of atomic oxygen.

Optimization of Surface Kinetics Parameters

Parameters for Optimization

The goal is to refine the less-known parameters in the Arrhenius rate equations, $k_i = A_i \cdot e^{-E_{a_i}/k_B T}$. These parameters form a vector, θ, to be optimized. It includes:

- **Energy Barriers (E_{a_i})**: For creation of the S_v^* sites and its destructions.
- **Steric Factors (A_i)**: For the chemical reactions involving CO and the metastable sites [4].
- **Desorption frequency (ν_d)**: For the desorption of physisorbed species (O_F, O_{2F}, CO_F), parametrized as $\nu_d = A + B \cdot e^{E/k_B T}$ [2].

In total, we consider an optimization problem with 29 parameters.

The Optimization Problem

We define the optimal parameter vector θ^* that minimizes the objective function $\Phi(\theta)$, given by:

$$\theta^* = \arg\min_{\theta \in \Theta} \Phi(\theta)$$

with Θ the domain of physically admissible solutions and $\Phi(\theta) = \frac{1}{2}\sum_i^N r_i^2(\theta)$ and the residual $r_i(\theta) = (\gamma_i - \hat{\gamma}_i(\theta))/\gamma_i$. The recombination probability γ_i corresponds to the measured observables and $\hat{\gamma}_i(\theta)$ to the predicted ones. i iterates over the experimental set of conditions [2,4].

Derivative-Free Optimization Methods

Our parameter estimation task is defined as a black-box, derivation-free optimization problem, as the LoKI-C [5] simulator does not provide analytical gradient information. Thus, we must use algorithms that do not rely on derivatives.
As a baseline, we employ the global search algorithms Differential Evolution (DE) [6] and Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [7] to explore the entire parameter space Θ, and Powell's Method [8], which is more focused on local refinement.
We propose a Hierarchical Optimization Algorithm based on Sloppy Model Theory [9,10].

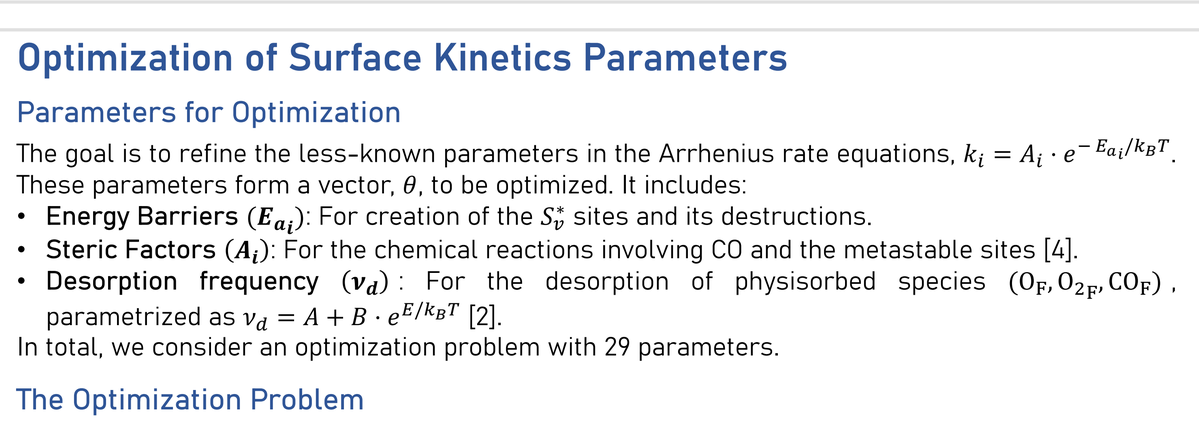
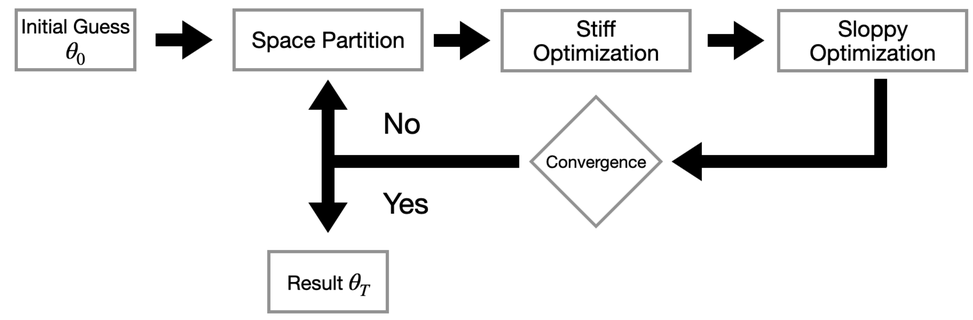


Figure 1(a) shows the eigenvalue spectrum ($|\lambda_i|$) of the Gauss-Newton Hessian, $H_{GN} \approx \nabla^2\tilde{\Phi}(\theta)$, as a function of the eigendirections V_i at the minimum θ^* found. The H_{GN} is represented as $H_{GN} = V\Lambda V^T$, with $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Figure 1(b) compares the convergence of the loss function (Φ) versus iterations for our hierarchical method (red) against the baseline methods, showing that our approach converges faster and yields better results.

Hierarchical Optimization Algorithm

In Figure 1(a), the eigenvalues span many orders of magnitude, confirming the model's sloppy nature. The large eigenvalues correspond to stiff directions, while the tiny eigenvalues correspond to sloppy directions that are challenging for standard optimization methods to resolve.
To solve this, our hierarchical algorithm (see scheme) iteratively partitions parameters into stiff and sloppy subspaces, based on the eigen decomposition of H_{GN}. It then optimizes each subspace in sequence, efficiently navigating the complex landscape to find the final parameters θ_T, as shown in Figure 1(b).



Results

We now present the results of the optimized parameters obtained using the hierarchical optimization algorithm, as shown in Figure 1(b). It reduced the loss Φ from the value of 0.1038 using default parameters [2,3,4] to the optimized value of Φ = 0.0672.

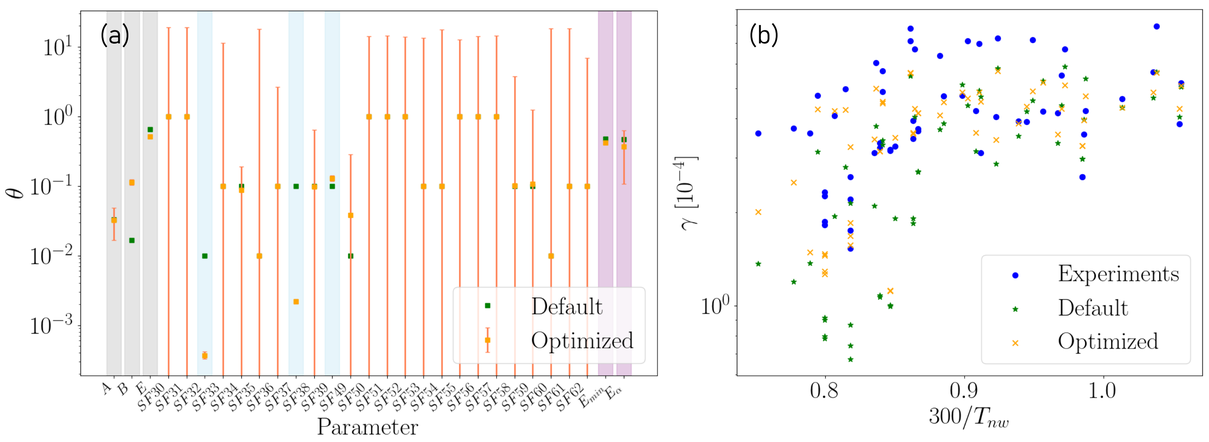


Figure 2(a): Comparison between the default (green) values and the final optimized parameters (orange). Error bars represent the uncertainty calculated from the Hessian of the loss, indicating how well each parameter is constrained by the data. Figure 2(b): Comparison of the experimentally measured γ with the model simulations with the optimized parameters (green) and the default ones (orange) for the set of points with $p < 1$ Torr.

Analysis of the optimized parameters, as shown in Fig. 2(a), reveals that the experimental data most strongly constrain a specific subset of reactions. These critical (stiff) parameters include:

- the desorption frequency of O_F, O_{2F}, CO_F (A, B, E) (gray bars);
- The steric factors of key reactions governing the lifecycle of chemisorbed CO: its formation ($CO + S_V \rightarrow CO_S$) and its subsequent recombination with O ($O + CO_S \rightarrow CO_2 + S_V$ and $O_F + CO_S \rightarrow CO_2 + S_V$) (blue bars).
- the energy for creating and destroying metastable states: E_α and E_{min} (pink bars).

The resulting model, validated in Figure 2(b), demonstrates an improvement in predictive power when compared to the default one.

Conclusions:

- Given the sloppy nature of the underlying physical model, the proposed hierarchical optimization algorithm provides a significant performance improvement when compared to the baseline methods like DE, CMA-ES, and Powell's method.
- A data-driven uncertainty estimates for the final parameters is derived from the Hessian of the loss function. Its analysis confirms which parameters are well-constrained by the experiment and which are less sensitive.
- The methods approached here are model-agnostic. Hence, it serves as a framework that can be applied to other complex, multi-parameter calibration problems in plasma physics and other scientific subjects.

Perspectives:

- Validate the generality and robustness of the hierarchical optimization method by applying it to other kinetics schemes.
- Determine the most important O recombination mechanisms in CO_2 plasmas in Pyrex using the optimized reaction scheme.

References:

[1] Guerra V. (2007). Analytical model of heterogeneous atomic recombination on silicalike surfaces, *IEEE Trans. Plasma Sci.* **35**, 1397-412
[2] Viegas P. et al. (2024). Surface recombination in Pyrex in oxygen DC glow discharges: mesoscopic modelling and comparison with experiments, *Plasma Sources Sci. Technol.* **33** (5), 05503
[3] Afonso J. et al. (2024). Plasma-induced reversible surface modification and its impact on oxygen heterogeneous recombination, *J. Phys. D. Appl. Phys.* **57**, 04LT01
[4] Berdugo B. (2024). Surface Recombination on Pyrex in CO_2 Glow Discharges, Université Paris-Saclay
[5] https://nprime.tecnico.ulisboa.pt/loki
[6] Das S. and Suganthan P. N. (2011). Differential Evolution: A survey of the state-of-the-art, *IEEE Transactions on Evolutionary Computation* **15** (1), 4-31
[7] Hansen N. (2016). The CMA Evolution Strategy: A Tutorial, arxiv.org/abs/1604.00772
[8] Powell M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating dervatives, *The Computer Journal* **7** (2), 155-162
[9] Transtrum M. K., Machta B. B. and Sethna J. P. (2010). Why are nonlinear fits so challenging?, *Physical Review Letters* **104**, 060201
[10] Transtrum M. K., Machta B. B. and Sethna J. P. (2011). An efficient, geometrically motivated algorithm for fitting sloppy models, *Physical Review E* **83** (3), 036701

Acknowledgements:

This work was supported by FCT (Fundação para a Ciência e a Tecnologia), under funding to IPFN: UIDB/50010/2020 (https://doi.org/10.54499/UIDB/50010/2020), UIDP/50010/2020 (https://doi.org/10.54499/UIDP/50010/2020), LA/P/0061/2020 (https://doi.org/109.54499/LA/P/0061/2020); and by European Union under Horizon Europe project CANMILK (DOI:10.3030/101069491).