Supplementary Information: Gradient-enhanced neural networks for model parameter estimation applied to flow chemistry automated platforms

Francisco Bolaños-García[1,*], Jean-Marc Commenge[1], and Laurent Falk[1]

[1]Université de Lorraine, CNRS, LRGP, Nancy F-54000, France

[*]Corresponding author: francisco.bolanos-garcia@univ-lorraine.fr

S.1 Gaussian Process comparison

To provide a baseline comparison to the artificial neural network (ANN) surrogates, a standard Gaussian Process (GP) regression model was trained using the same dataset and data splits employed for ANN training. The GP model was implemented using an exact regression formulation with a constant mean function and a radial basis function (RBF) kernel. Model hyperparameters, including kernel lengthscales and noise variance, were optimized by maximizing the marginal log-likelihood using the Adam optimizer.

Predictions were evaluated on the test set, and gradients with respect to the kinetic parameters (k_1, k_2) were obtained via automatic differentiation of the GP predictive mean. To ensure consistency with the ANN evaluation, both outputs and gradients were scaled identically before computing performance metrics. The GP model achieved a total mean squared error (MSE) of 1.89×10^{-2}.

Parity plots for output predictions and parameter sensitivities are shown in Figure S.1. While the GP provides accurate output predictions, it exhibits similar limitations in reproducing parameter sensitivities.

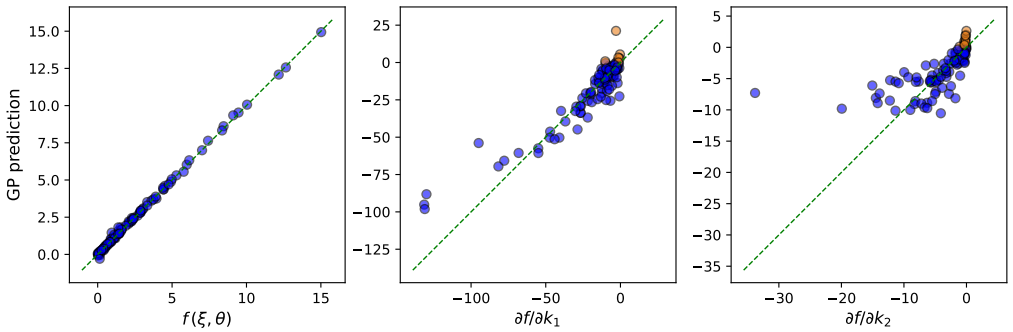
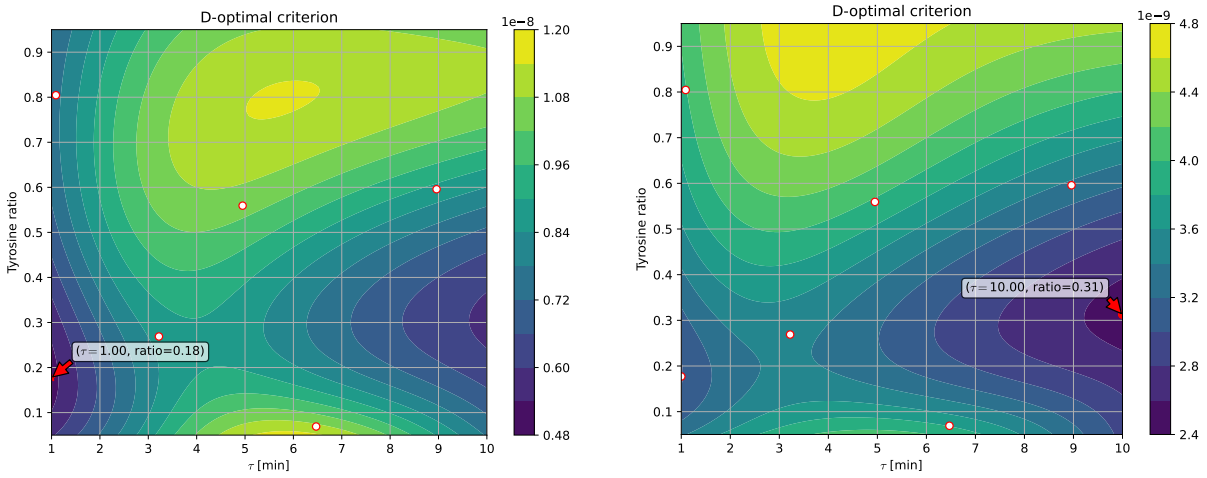


Figure S.1: Parity plots comparing the GP predicted output and gradients with respect to k_1 and k_2 under standard training against the true values obtained from the mechanistic model.

S.2 D-optimal experimental conditions

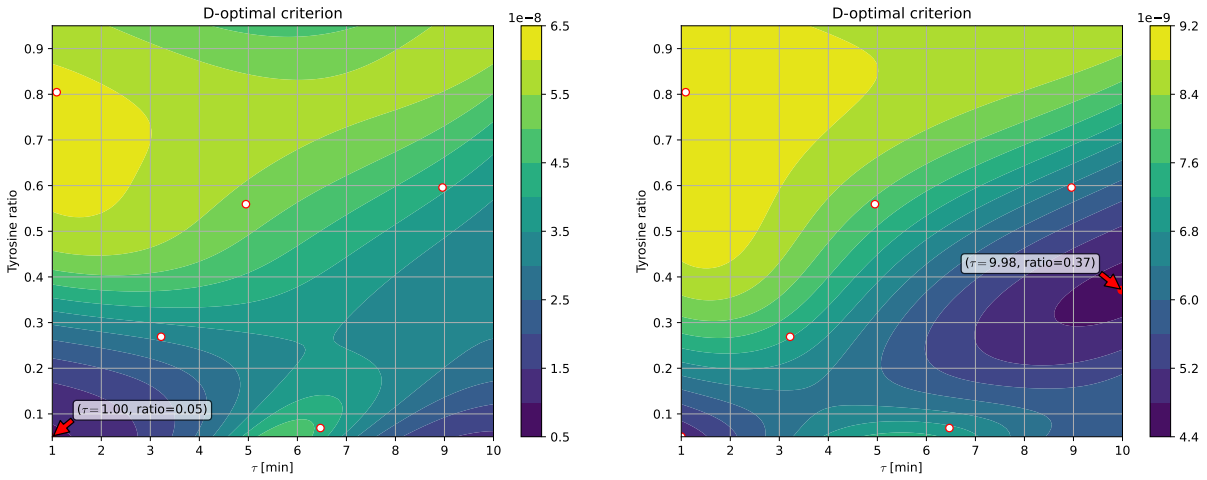
Figures S.2 to S.4 present the experimental conditions proposed after optimizing the D-optimal criterion using three different ANNs: the standard ANN with the lowest $\text{RMSE}_{\text{total}}$ ($\text{RMSE}_{\text{total}} = \text{RMSE}_{\text{output}} + \text{RMSE}_{\text{grad}}$), the standard ANN with the lowest $\text{RMSE}_{\text{output}}$, and the gradient-enhanced ANN with the lowest $\text{RMSE}_{\text{total}}$. Although all three models were trained on the same dataset and the experimental planning incorporated the same five preliminary experiments, the D-optimal experimental conditions are not the same.



(a) 6^{th} experiment proposed

(b) 7^{th} experiment proposed

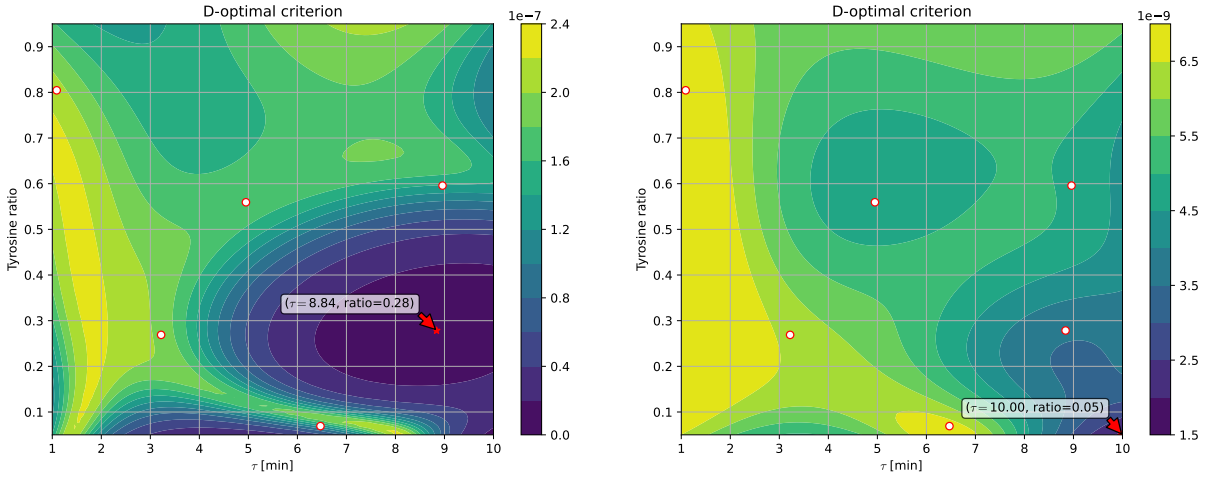
Figure S.2: First two experimental conditions proposed when optimizing the D-optimal criterion for the standard trained ANN with the lowest $\text{RMSE}_{\text{total}}$. White dots represent previous experiments



(a) 6^{th} experiment proposed

(b) 7^{th} experiment proposed

Figure S.3: First two experimental conditions proposed when optimizing the D-optimal criterion for the standard trained ANN with the lowest $\text{RMSE}_{\text{output}}$. White dots represent previous experiments



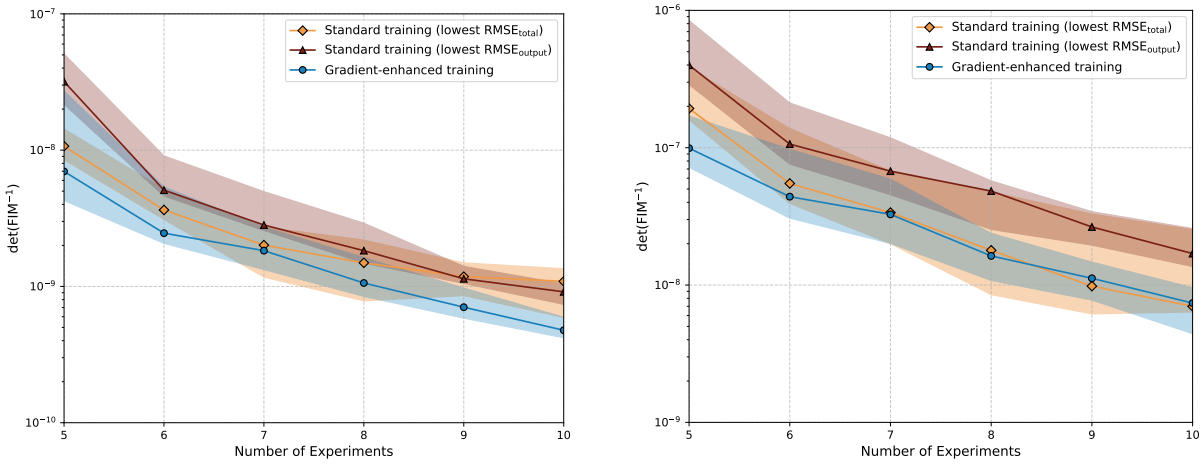
(a) 6^{th} experiment proposed

(b) 7^{th} experiment proposed

Figure S.4: First two experimental conditions proposed when optimizing the D-optimal criterion for the gradient-enhanced trained ANN. White dots represent previous experiments

S.3 Behaviour of confidence region based on the FIM

Figure S.5 reports the evolution of the determinant of the inverse Fisher Information Matrix (FIM^{-1}), which is proportional to the volume of the estimated confidence region. This quantity is directly related to the D-optimality criterion used for experimental design, where maximizing det(FIM) is equivalent to minimizing $\det(\mathrm{FIM}^{-1})$.



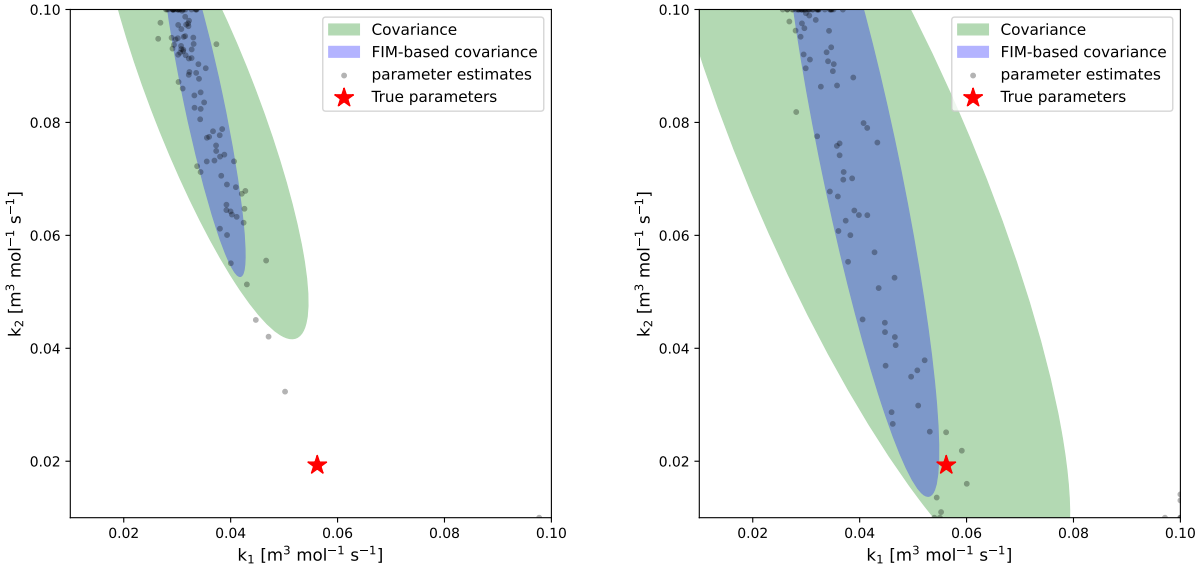
(a) $\sigma_e = 0.2\,\mathrm{mol\ s/m^3}$ (b) $\sigma_e = 0.4\,\mathrm{mol\ s/m^3}$

Figure S.5: Median determinant of the covariance matrix after each experiment. The shaded area shows the interquartile range of the distribution of values obtained for 10 different initial LHS designs.

Figures S.6–S.8 compare the empirical distribution of parameter estimates with the confidence regions predicted from the FIM. For each surrogate model, 100 independent MBDoE runs were performed under the same experimental budget. Each point in the scatter plot corresponds to the final parameter estimates obtained from one run.

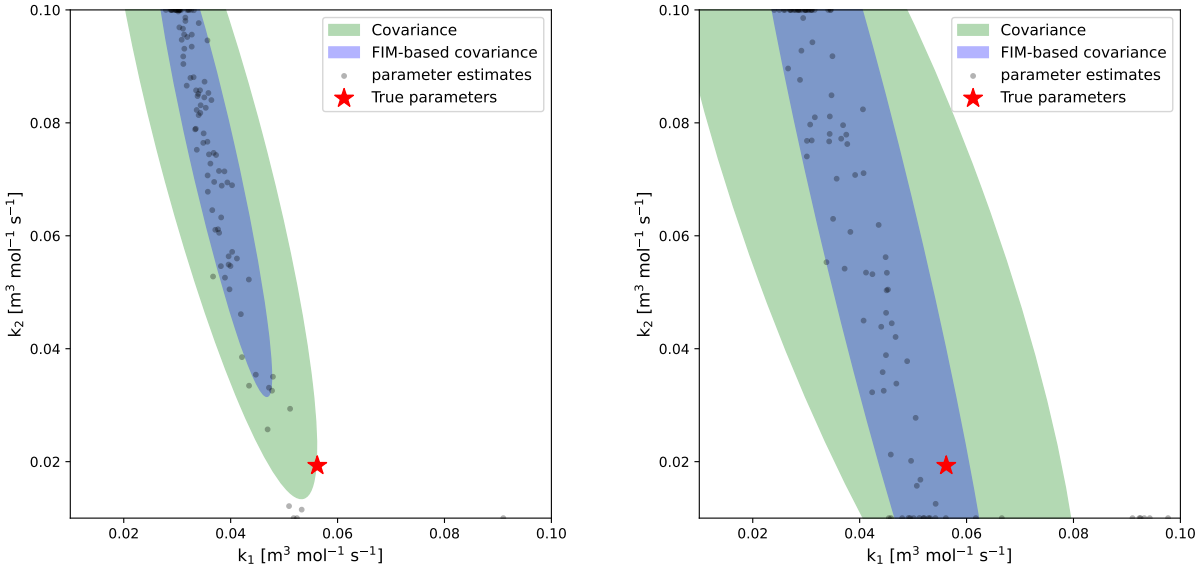
The empirical covariance ellipse (green region) was constructed from the sample covariance matrix of these 100 parameter estimates. The FIM-based covariance ellipse (blue region) was obtained from the inverse of the FIM corresponding to the run whose volume ratio is closest to the median value reported in the main text, thereby providing a representative realization of the FIM-based uncertainty.

Both ellipses correspond to the same confidence level of 95% and are centered at the mean of the estimated parameters. The true parameter values are shown for reference. This comparison highlights whether the FIM-based approximation is statistically consistent with the observed variability of the parameter estimates.



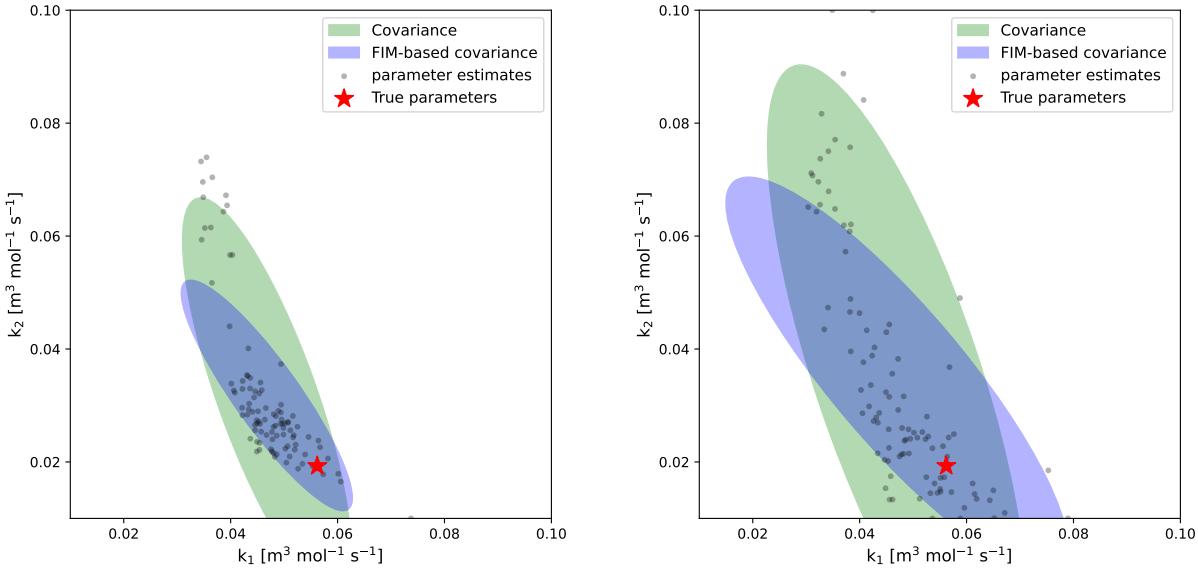
(a) $\sigma_e = 0.2\,\mathrm{mol\ s/m^3}$ (b) $\sigma_e = 0.4\,\mathrm{mol\ s/m^3}$

Figure S.6: Comparison between empirical parameter distribution (100 independent runs) and FIM-based confidence region for the standard ANN with the lowest $\mathrm{RMSE}_{\mathrm{total}}$.



(a) $\sigma_e = 0.2\,\mathrm{mol\ s/m^3}$ (b) $\sigma_e = 0.4\,\mathrm{mol\ s/m^3}$

Figure S.7: Comparison between empirical parameter distribution (100 independent runs) and FIM-based confidence region for the standard ANN with the lowest $\mathrm{RMSE}_{\mathrm{output}}$.



(a) $\sigma_e = 0.2\,\mathrm{mol\ s/m^3}$ (b) $\sigma_e = 0.4\,\mathrm{mol\ s/m^3}$

Figure S.8: Comparison between empirical parameter distribution (100 independent runs) and FIM-based confidence region for the gradient-enhanced ANN.