Electronic Supplementary Information

Efficient Neural Network Models of Chemical Kinetics Using a Latent asinh Rate Transformation

Felix A. Döppel and Martin Votsmeier*
*martin.votsmeier@tu-darmstadt.de

June 12, 2023

1 Linearizing Rate Expressions

$$r_\mathrm{j} = k_{\mathrm{j},0} \cdot \exp\left(-\frac{E_A}{RT}\right) \cdot \prod_\mathrm{i} c_\mathrm{i}^{\nu_{\mathrm{i,j}}} \cdot \prod_\mathrm{k} \theta_\mathrm{k}^{\nu_{\mathrm{k,j}}} \tag{S1}$$

$$\ln r_\mathrm{j} = \ln k_{\mathrm{j},0} - \frac{E_A}{R} \cdot \frac{1}{T} + \sum_\mathrm{i} \nu_{\mathrm{i,j}} \cdot \ln c_\mathrm{i} + \sum_\mathrm{k} \nu_{\mathrm{k,j}} \cdot \ln \theta_\mathrm{k} \tag{S2}$$

$$y = b + m \cdot x \tag{S3}$$

2 Preferential Oxidation Mechanism

Table S1 Reactions and parameters for the PROX mechanism taken from [1] originating from [2]

No.	Reaction	s_0 or A_0		E_A	$\frac{\mathrm{d}E_\mathrm{A}}{\mathrm{d}\theta_\mathrm{H}}$	$\frac{\mathrm{d}E_\mathrm{A}}{\mathrm{d}\theta_\mathrm{O}}$	$\frac{\mathrm{d}E_\mathrm{A}}{\mathrm{d}\theta_\mathrm{OH}}$	$\frac{\mathrm{d}E_\mathrm{A}}{\mathrm{d}\theta_{\mathrm{H_2O}}}$	$\frac{\mathrm{d}E_\mathrm{A}}{\mathrm{d}\theta_\mathrm{CO}}$
		unitless or s^{-1}	unitless	kJ/mol	kJ/mol	kJ/mol	kJ/mol	kJ/mol	kJ/mol
H_2 oxidation on Pt									
R1	$H_2 + 2^* \rightleftarrows 2H^*$	1.287×10^{-1}	0.8584	-	-	-	-	-	-
		7.953×10^{12}	1.911	79.09	−25.10	0	0	0	0
R2	$O_2 + 2^* \rightleftarrows 2O^*$	5.423×10^{-1}	0.7656	-	-	-	-	-	-
		8.406×10^{12}	0.9275	208.9	0	−133.9	0	0	0
R3	$OH^* + ^* \rightleftarrows O^* + H^*$	1.950×10^{12}	13.286	111.2	67.321	−4.87	0	62.89	0
		6.325×10^{12}	10.812	32.48	−5.820	22.26	0	−41.71	0
R4	$H_2O^* + ^* \rightleftarrows OH^* + H^*$	9.358×10^{12}	−0.3949	74.57	4.795	50.33	52.30	−43.25	0
		9.989×10^{12}	−0.3664	53.43	−7.757	−87.75	−52.30	71.81	0
R5	$H_2O^* + O^* \rightleftarrows 2OH^*$	4.316×10^{10}	0.3262	36.74	0	70.10	52.30	−83.68	0
		1.700×10^{10}	0.5285	94.32	0	−139.12	−52.30	136.0	0
R6	$H_2O + ^* \rightleftarrows H_2O^*$	1.084×10^{-1}	11.624	-	-	-	-	-	-
		2.033×10^{12}	2.490	39.98	0	0	104.6	−10.46	0
CO oxidation on Pt									
R7	$CO + ^* \rightleftarrows CO^*$	1.000	0	-	-	-	-	-	-
		5.659×10^{15}	0.3946	165.9	0	0	0	0	−62.70
R8	$CO_2 + ^* \rightleftarrows CO_2{}^*$	1.950×10^{-1}	0.2500	-	-	-	-	-	-
		3.626×10^{12}	0.2459	11.60	0	0	0	0	0
R9	$CO_2{}^* + ^* \rightleftarrows CO^* + O^*$	4.178×10^{10}	−0.2778	110.4	0	4.27	0	0	24.21
		2.393×10^{11}	0.4558	85.42	0	−17.69	0	0	−38.53
Coupling reactions between CO- and H_2 oxidation									
R10	$CO_2{}^* + H^* \rightleftarrows CO^* + OH^*$	8.031×10^{8}	−0.3259	23.22	−6.276	60.25	0	−45.56	18.12
		1.245×10^{9}	0.8237	76.91	6.276	−77.83	0	59.04	−44.63
R11	$COOH^* + ^* \rightleftarrows CO^* + OH^*$	8.426×10^{8}	0.025 77	22.66	0	60.25	0	−45.56	18.13
		1.187×10^{9}	0.4719	77.46	0	−77.83	0	59.94	−44.63
R12	$COOH^* + ^* \rightleftarrows CO_2{}^* + H^*$	1.058×10^{11}	0.5812	4.975	6.276	0	0	0	0
		9.454×10^{10}	−0.1098	6.079	−62.844	0	0	0	0
R13	$CO^* + H_2O^* \rightleftarrows COOH^* + H^*$	1.103×10^{11}	0.4911	98.71	4.904	0	52.30	−5.230	−31.38
		9.070×10^{10}	−0.027 78	22.79	−7.657	0	−52.30	5.230	31.38
R14	$CO_2{}^* + OH^* \rightleftarrows COOH^* + O^*$	5.349×10^{10}	0.052 72	109.5	0	−41.13	0	52.30	0
		1.870×10^{11}	0.4515	29.70	0	30.00	0	−52.30	0
R15	$CO_2{}^* + H_2O^* \rightleftarrows COOH^* + OH^*$	8.642×10^{10}	−0.0481	71.31	0	52.45	52.30	−44.87	0
		1.157×10^{11}	0.5647	49.07	0	−85.62	−52.30	70.19	0
R16	$CO_2{}^* + H^* \rightleftarrows HCOO^* + ^*$	1.117×10^{11}	−0.075 25	73.49	−12.55	0	0	0	0
		8.957×10^{10}	0.4218	0.00	0	0	0	0	0
R17	$CO_2{}^* + OH^* \rightleftarrows HCOO^* + O^*$	6.168×10^{10}	−0.3443	154.8	0	−67.47	0	100.2	0
		1.621×10^{11}	−0.1617	2.577	0	3.657	0	−4.351	0
R18	$CO_2{}^* + H_2O^* \rightleftarrows HCOO^* + OH^*$	1.022×10^{11}	−0.3574	111.8	0	118.7	90.99	−98.17	0
		9.785×10^{10}	0.066 65	17.15	0	−19.40	−13.61	16.89	0

3 Ammonia Oxidation Mechanism

Table S2 shows all 15 reversible reactions of the ammonia oxidation mechanism considered in this work, the activation energy E_A associated with the forward direction and the energy difference ΔE of the reaction. Table S3 lists all non-imaginary frequencies ν of the initial- and transition states. Values have to be multiplied by the elementary charge e and divided by Planck's constant h to convert to the unit Hz. Sticking coefficients of all adsorption reactions are assumed to be one.

Table S2 All 15 reversible ammonia oxidation reactions, the activation energy E_A associated with the forward direction and the energy difference ΔE of the reaction

Reaction	E_A / J mol^{-1}	ΔE / J mol^{-1}
$O_2 + 2^* \rightleftharpoons 2O^*$	0	−238 319
$NH_3 + ^* \rightleftharpoons NH_3{}^*$	0	−66 575
$NH_3{}^* + O^* \rightleftharpoons NH_2{}^* + OH^*$	67 540	39 559
$NH_2{}^* + O^* \rightleftharpoons NH^* + OH^*$	78 153	−14 473
$NH^* + O^* \rightleftharpoons N^* + OH^*$	154 377	31 840
$NH_3{}^* + OH^* \rightleftharpoons NH_2{}^* + H_2O^*$	33 770	−13 508
$NH_2{}^* + OH^* \rightleftharpoons NH^* + H_2O^*$	965	−67 540
$NH^* + OH^* \rightleftharpoons N^* + H_2O^*$	39 559	−22 192
$2OH^* \rightleftharpoons O^* + H_2O^*$	0	−53 067
$H_2O^* \rightleftharpoons H_2O + ^*$	18 332	18 332
$2N^* \rightleftharpoons N_2 + 2^*$	244 108	−52 102
$N^* + O^* \rightleftharpoons NO^* + ^*$	213 233	3859
$NO^* \rightleftharpoons NO + ^*$	184 287	184 287
$N^* + NO^* \rightleftharpoons N_2O^* + ^*$	164 990	89 731
$N_2O^* \rightleftharpoons N_2O + ^*$	0	−2895

Table S3 A list of all non-imaginary frequencies ν of the initial states IS and transition states TS

Species	Type	ν / V
$NH_3{}^*$	IS	0.4320 0.4319 0.4152 0.1936 0.1934 0.1342 0.0820 0.0819 0.0458 0.0150 0.0149
$NH_2{}^*$	IS	0.4327 0.4189 0.1800 0.0969 0.0954 0.0825 0.0595 0.0436 0.0269
NH^*	IS	0.4242 0.0994 0.0992 0.0665 0.0588 0.0587
N^*	IS	0.0612 0.0612 0.0612
N_2O^*	IS	0.2980 0.1614 0.0541 0.0540 0.0225 0.0150 0.0149 0.0033
NO^*	IS	0.1924 0.0512 0.0511 0.0388 0.0187 0.0184
H_2O^*	IS	0.4578 0.4448 0.1919 0.0635 0.0573 0.0167 0.0122 0.0114
OH^*	IS	0.4391 0.0930 0.0896 0.0466 0.0244 0.0225
O^*	IS	0.0558 0.0454 0.0453
*	IS	–
$NH_3{-}O$	TS	0.4379 0.4262 0.3916 0.1802 0.1211 0.1125 0.0810 0.0685 0.0641 0.0588 0.0418 0.0316 0.0201 0.0137
$NH_2{-}O$	TS	0.4129 0.1870 0.1610 0.1155 0.1045 0.0786 0.0669 0.0576 0.0458 0.0332 0.0272
$NH{-}O$	TS	0.2167 0.1402 0.0776 0.0656 0.0569 0.0523 0.0355 0.0298
$NH_3{-}OH$	TS	0.4564 0.4366 0.4309 0.4057 0.1952 0.1826 0.1033 0.0825 0.0671 0.0638 0.0426 0.0333 0.0270 0.0201 0.0143 0.0088 0.0065
$NH_2{-}OH$	TS	0.4623 0.4009 0.3930 0.1915 0.1032 0.0995 0.0750 0.0607 0.0569 0.0502 0.0460 0.0197 0.0178 0.0113
$NH{-}OH$	TS	0.4531 0.2032 0.1503 0.1021 0.0783 0.0637 0.0626 0.0563 0.0448 0.0306 0.0198
$OH{-}OH$	TS	0.4549 0.1881 0.1673 0.1026 0.0799 0.0709 0.0632 0.0534 0.0368 0.0231 0.0180
$N{-}N$	TS	0.0716 0.0702 0.0537 0.0474 0.0228
$N{-}O$	TS	0.0722 0.0613 0.0525 0.0299 0.0287
$N{-}NO$	TS	0.2166 0.0731 0.0574 0.0515 0.0367 0.0339 0.0159 0.0060

4 Training Times

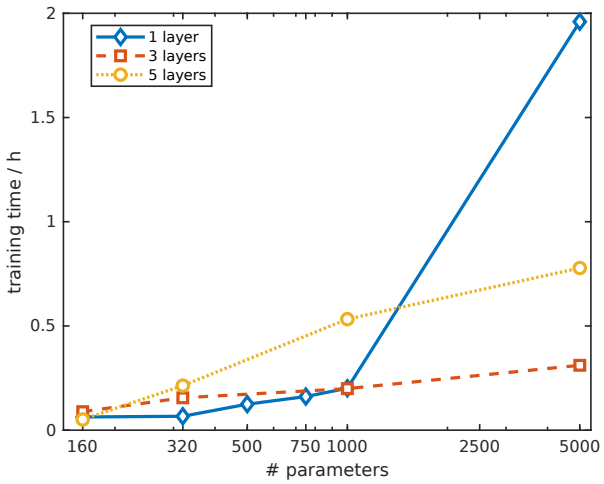


Figure S1 Median training times of the neural networks predicting CO source terms with the latent hyperbolic sine transformation strategy. Measured using three threads of a Intel® Xeon® Platinum 9242 Processor each.

5 Comparing Relative Error and MATE

So far, we applied the inverse hyperbolic sine transformation in a latent way. To investigate the effects of the latent approach, we also applied the inverse hyperbolic sine transformation in the conventional way. This means data are transformed in a preprocessing step and the transformed values are used as targets to be learned by a conventional neural network. As shown in table S4, the conventional approach leads to relative prediction errors above 1000 %, while the latent approach achieves 15 %. This can be attributed to the fact that instead of the relative error, the conventional approach minimizes an error measure defined in terms of the transformed values $\operatorname{asinh}(\dot{s})$ which we call MATE. This error measure, however, is not relevant for reactor simulations.

Table S4 Prediction errors of lightweight neural networks with 40 nodes in a single hidden layer (≈320 parameters) modeling the CO source terms $\dot{s}$ with two different approaches: Latent transformation minimizes the relative error during training resulting in an average accuracy of 15 %. Conventional transformation minimizes the error of transformed values MATE instead.[1] Therefore, its predictions are two orders of magnitude less accurate, as measured by the relative error. Because MATE is not a relevant measure for the application in reactor simulations, the slightly better MATE score of the conventional approach poses no considerable advantage over the latent approach. The equations show how the errors are computed using the neural network predictions h

error measure	equation	latent (this work)	conventional
relative error	$\left\lvert\frac{\dot{s}-h(\dot{s})}{\dot{s}}\right\rvert$	15 %	>1000 %
MATE	$\lvert\operatorname{asinh}(\dot{s})-h(\operatorname{asinh}(\dot{s}))\rvert$	120 %	110 %

Figure S2 shows the relative error and the mean absolute transformed error MATE as a function of the predicted value. The exact value is arbitrarily chosen to be one. Both error measures share the same minimum and show similar asymptotic behavior for underestimations. Because of the logarithmic x-axis, the relative error grows exponentially for overestimations while MATE grows linearly. This leads to much more significant overestimations occuring in models trained with the conventional transformation approach as opposed to the latent approach.

[1] Actually, both approaches minimize the root mean square of the term shown in the respective equation, not the mean absolute which is reported in table S4. As the root mean square is more sensitive to outliers it is well suited for model training but not an intuitive indicator of the model quality.

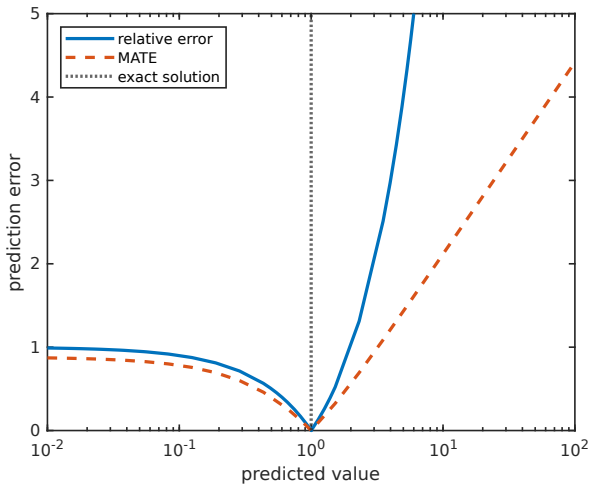


Figure S2 Comparison between the relative error and the mean absolute transformed error MATE (y-axis) for different predicted values (x-axis) assuming an exact solution of one.

6 Comparing Conventional and Latent Transformation Results

Figure S3 shows the prediction accuracy gain when using latent inverse hyperbolic sine transformation instead of the conventional approach modeling steady state CO source terms from test case one. The numbers shown are computed by dividing the relative prediction error obtained with the conventional approach by the error of the latent approach.

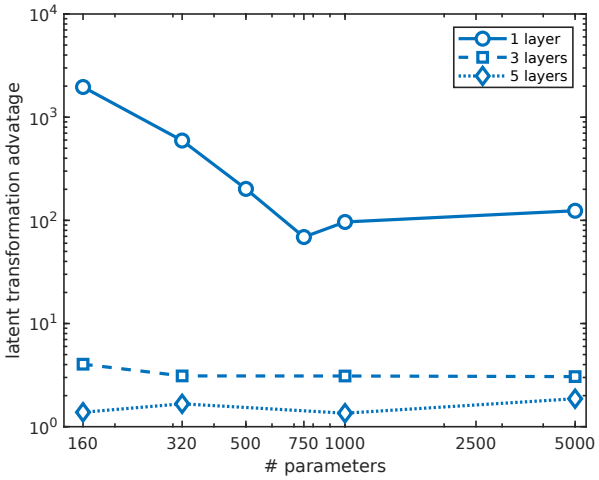


Figure S3 The accuracy gain of latent over conventional data transformation is shown as the quotient between relative prediction errors of models with identical complexity modeling steady state CO source terms. Results are shown for different numbers of hidden layers as a function of the number of model parameters.

7 Inference Time

Source term prediction time t_{predict} for 100 000 reaction conditions is averaged over 1000 runs with the models used for the preferential oxidation plug-flow reactor simulations. Results are shown in table S5 using an AMD Ryzen 7 5800X CPU and a GEFORCE RTX 3070 GPU. Computing the exact solution on the same machine takes 800 s resulting in a speed-up of 45 700 on the CPU and 95 200 on the GPU.

Table S5 Neural network prediction times t_{predict} for 100 000 reaction conditions averaged over 1000 measurements with the same models as used for the preferential oxidation plug-flow reactor simulations

Hardware	$t_{\text{predict}}(\text{O}_2)$ / ms	$t_{\text{predict}}(\text{CO})$ / ms
CPU: AMD Ryzen 7 5800X	5.5	12
GPU: GEFORCE RTX 3070	3.0	5.4

8 Comparing to our Previous Work Using Mechanistic Insights

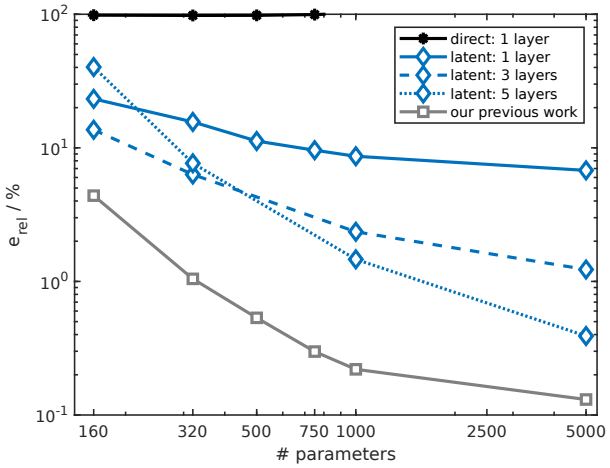
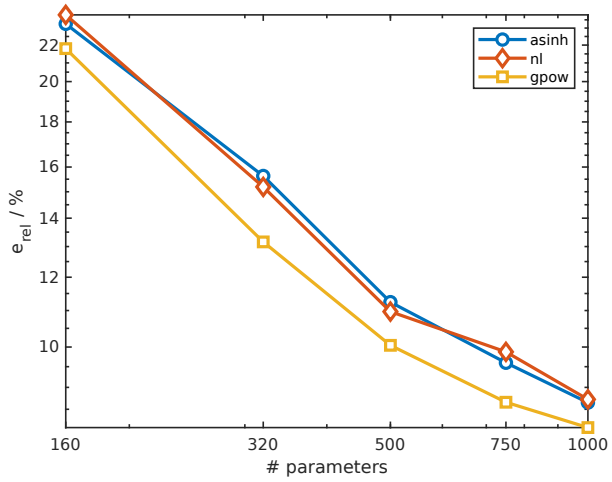
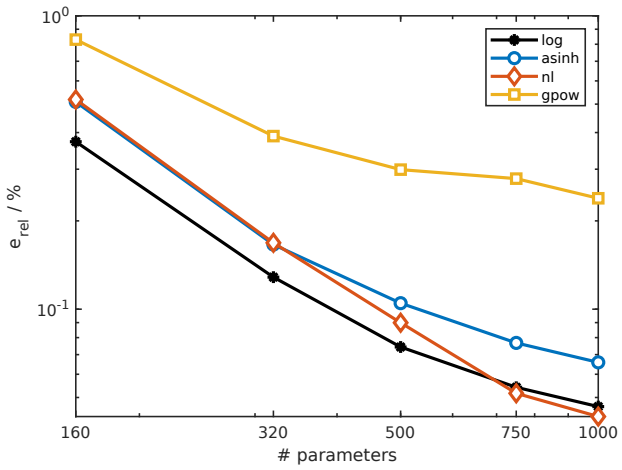


Figure S4 Relative prediction error of CO source terms dependent of the total number of learnable parameters in a neural network compared between different modeling strategies. Our previous work [1] performs best but requires a reaction path analysis and therefore cannot be used with experimental data or highly complex computational models.

9 Alternatives to the Hyperbolic Sine



(a) . . . CO source terms.



(b) . . . O_2 source terms.

Figure S5 Comparing the prediction accuracy of different transformation functions as a function of trainable model parameters in a single hidden layer for . . .

References

[1] F. A. Döppel and M. Votsmeier, *Chemical Engineering Science*, 2022, **262**, 117964.

[2] W. Hauptmann, M. Votsmeier, H. Vogel and D. G. Vlachos, *Applied Catalysis A: General*, 2011, **397**, 174–182.