Speeding up high-throughput characterization of materials libraries by active learning: autonomous electrical resistance measurements

Supporting Information

Felix Thelen, Lars Banko, Rico Zehl, Sabrina Baha, and Alfred Ludwig*

Chair for Materials Discovery and Interfaces, Institute for Materials,
Ruhr University Bochum, Universitätsstraße 150, 44780 Bochum, Germany.
** corresponding author: alfred.ludwig@rub.de*

Contents:

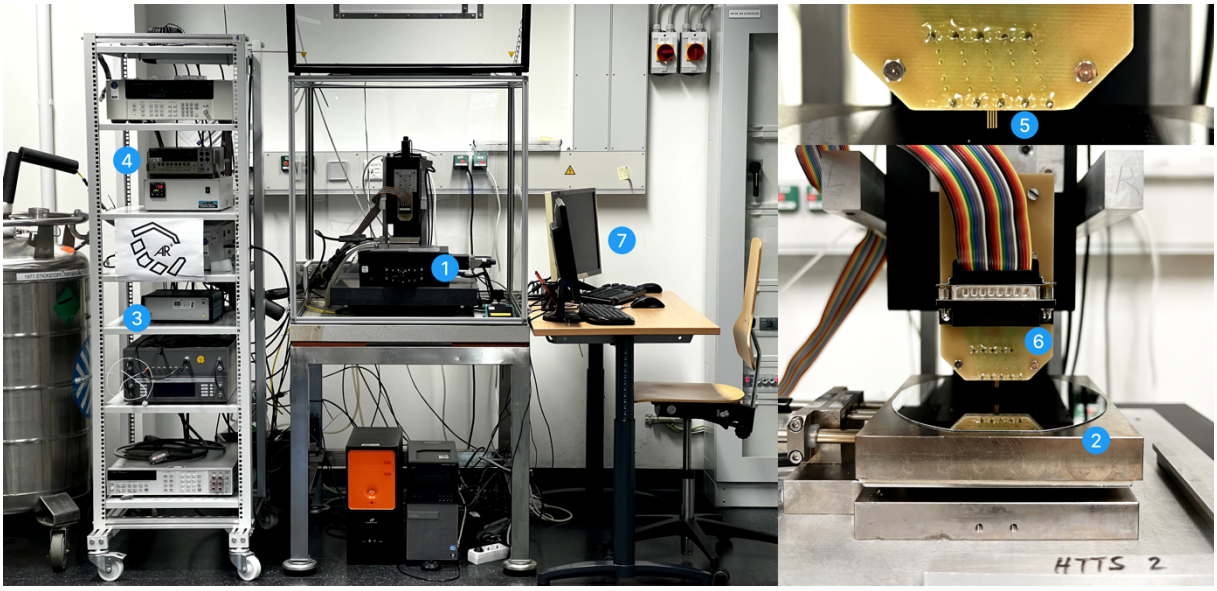
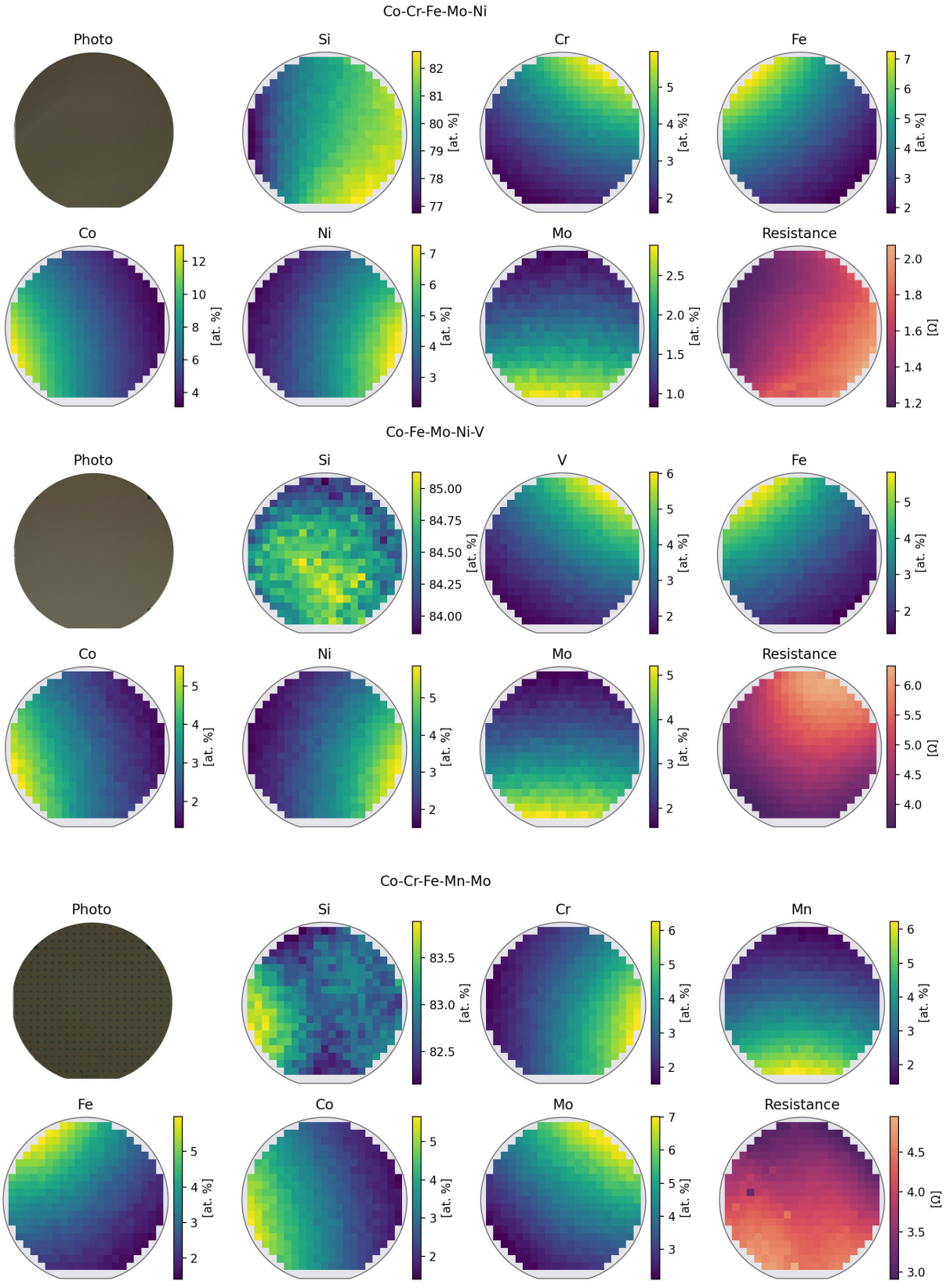
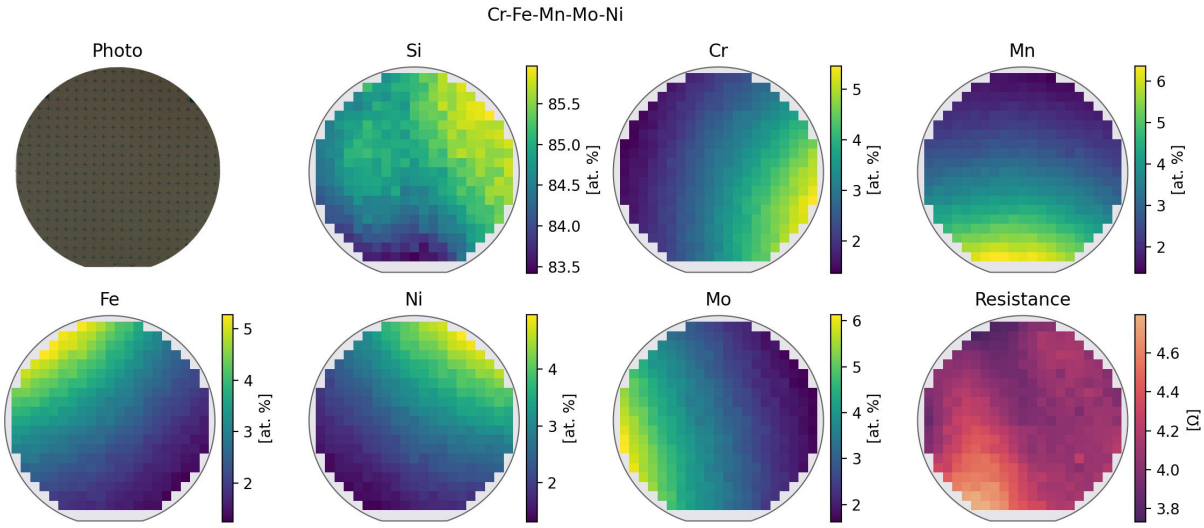
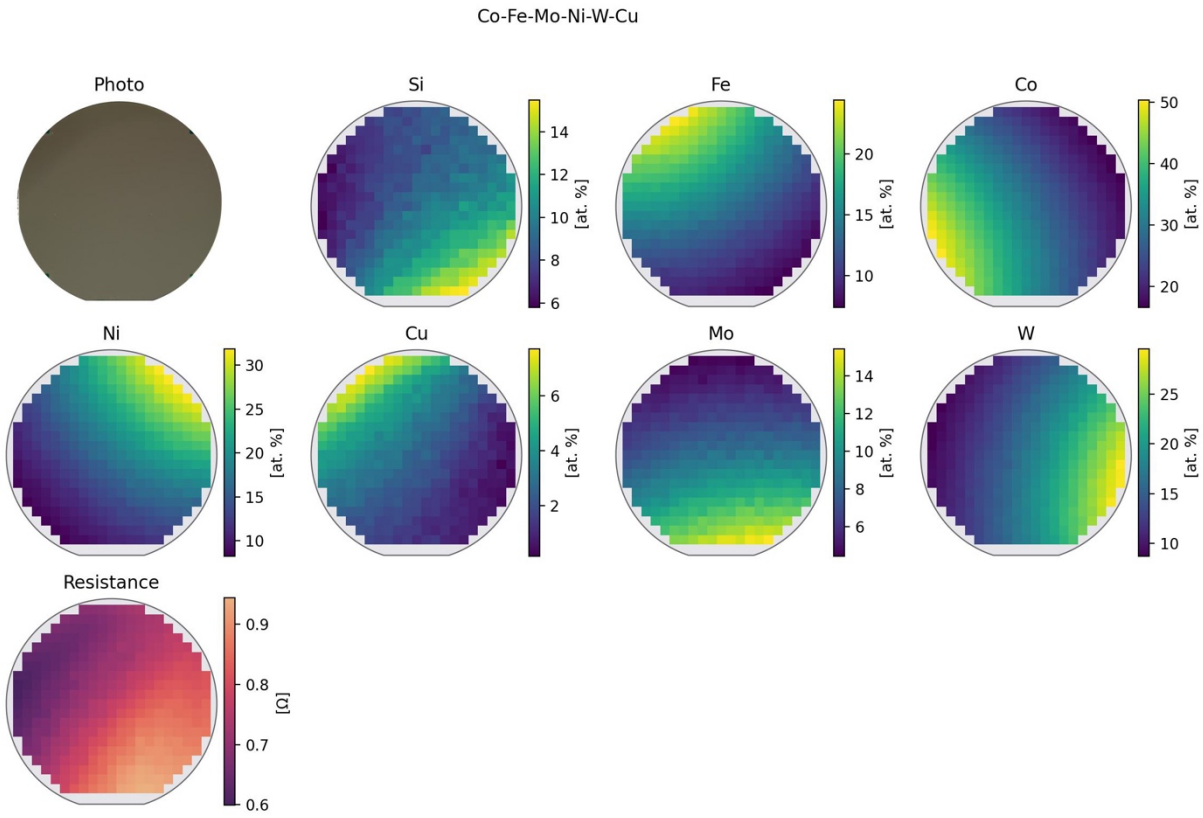


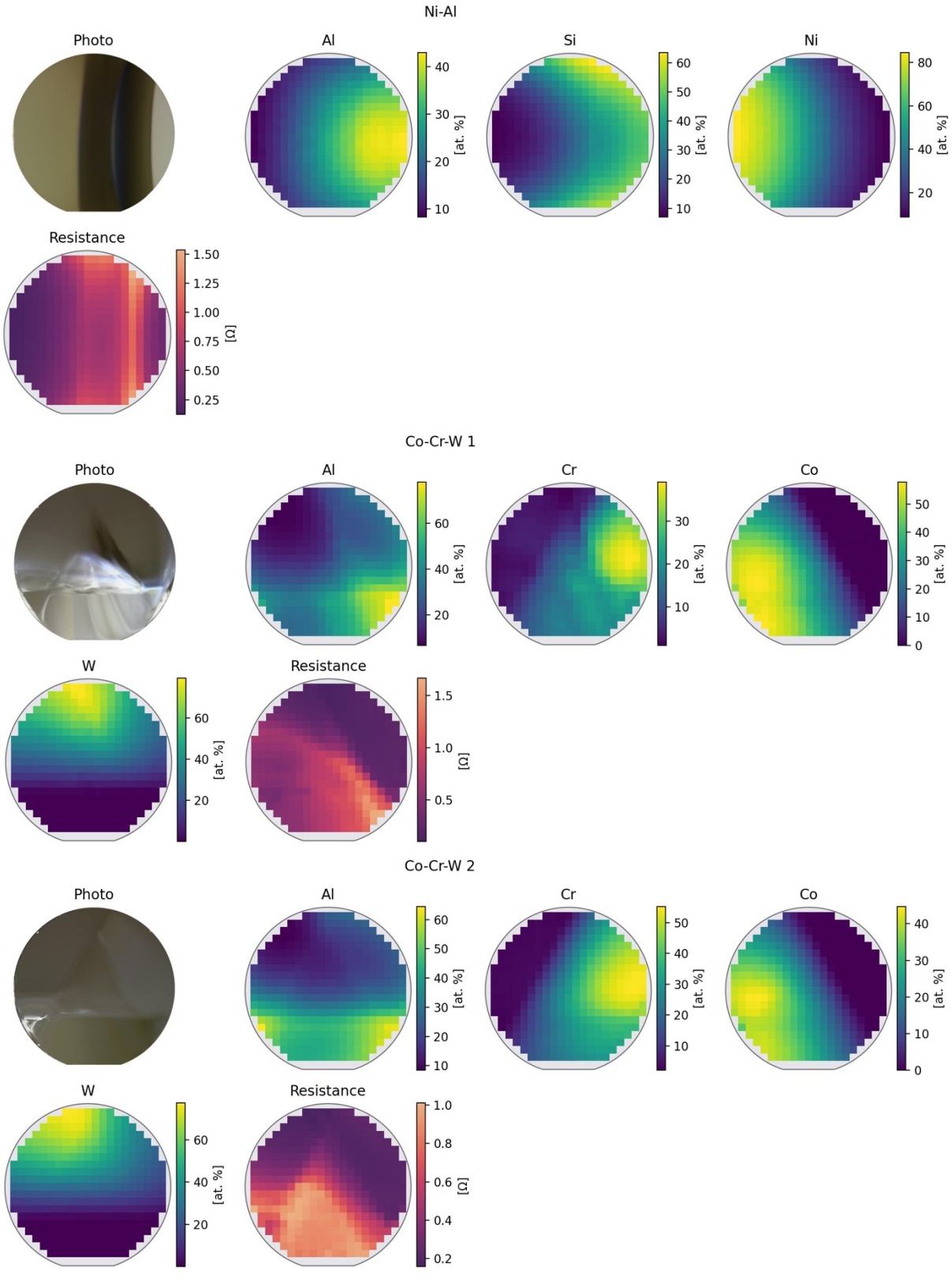
Figure S-1: Photo of the resistance measurement setup. The materials library is secured on the x-y-stage (1) (UHL GT8-NSNA) by a suction plate (2). The plate also accommodates a heating and nitrogen cooling system allowing temperature dependent measurements. However, this system was not used for the findings of this work. The stages are controlled by UHL F9S-3-M positioning controller (3). The resistance measurements are conducted using a Keithley 2400 source meter (4). The contact pins (5) (Feinmetall F238) are mounted to the stage by a custom spring-loaded circuit board (6). The setup is controlled via an Intel Core i7 8 GB RAM Windows 10 PC (7).

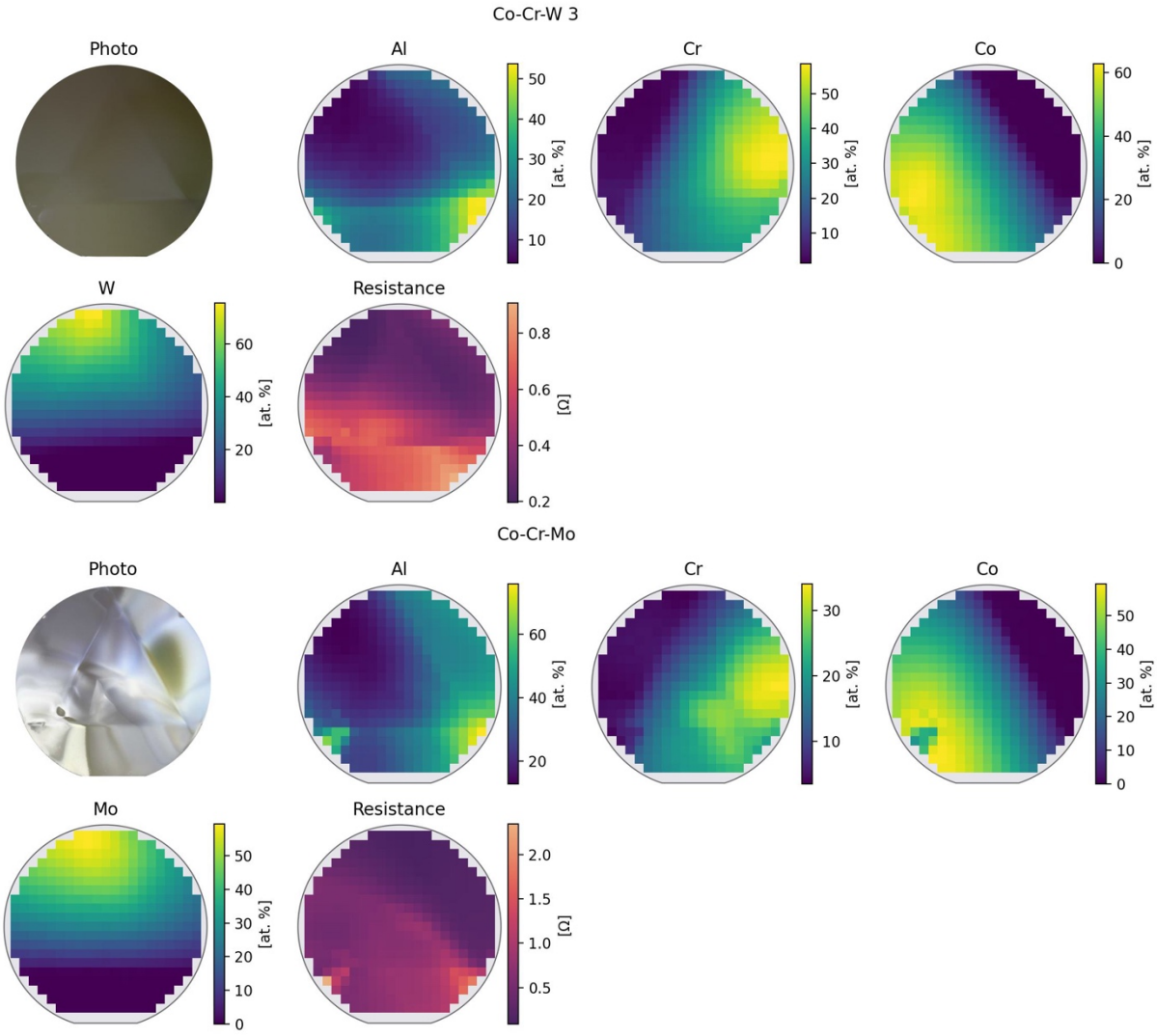
Figure S-2: Visualization of each of the ten test libraries containing a photo of the library, compositional data, and the measured electrical resistance.




Cr-Fe-Mn-Mo-Ni
Photo
Si
Cr
Mn
Fe
Ni
Mo
Resistance
[at. %]
[Ω]


Co-Fe-Mo-Ni-W-Cu
Photo
Si
Fe
Co
Ni
Cu
Mo
W
Resistance
[at. %]
[Ω]


Ni-Al
Photo
Al
Si
Ni
[at. %]
Resistance
[Ω]
Co-Cr-W 1
Photo
Al
Cr
Co
[at. %]
W
Resistance
[Ω]
Co-Cr-W 2
Photo
Al
Cr
Co
[at. %]
W
Resistance
[Ω]


Co-Cr-W 3
Photo
Al
Cr
Co
W
Resistance
[at. %]
[Ω]
Co-Cr-Mo
Photo
Al
Cr
Co
Mo
Resistance
[at. %]
[Ω]

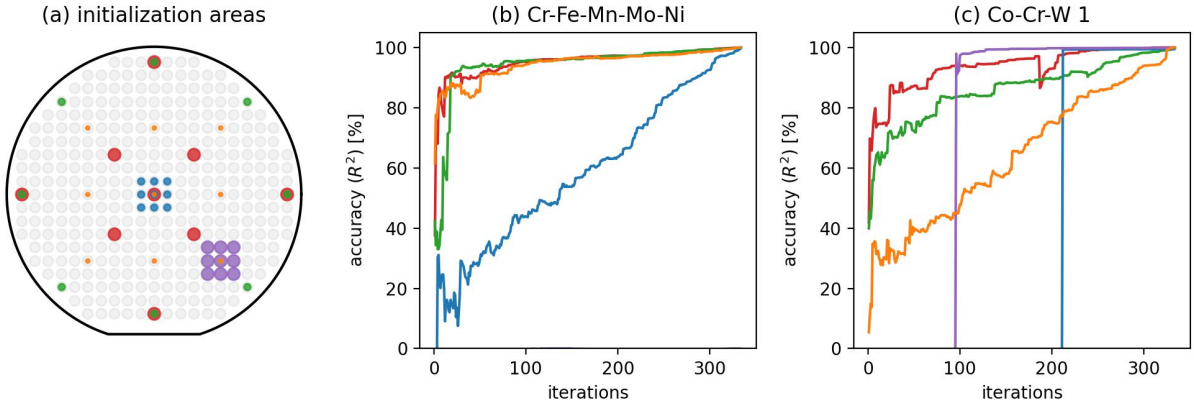


Figure S-3: Test of different initialization area arrangements. One co- (b) and one multilayer sputtered materials library (c) were examined. For nine initialization areas, a total of $n = \binom{349}{9} \approx 1.58 \cdot 10^{17}$ combinations are possible, that is why a random selection of arrangements is not feasible. The red one is used for all other tests of the algorithm, as it generally performed reliably across all libraries. Generally, the choice of initialization areas matters most for libraries showing a resistance distribution of high variance. Since the multilayer sputtered libraries cover more area on their respective composition space, the choice is in this case more important. The blue and purple areas show that for a narrow selection of initialization areas, the algorithm is converging late or not at all. The arrangements shown in green and orange show that the edges of the libraries are most important to achieve a sufficient initial fit. The edges often feature the areas with the highest uncertainty of the Gaussian process, therefore the performance improves when adding these areas early on. In the arrangement shown in green, the area in the center of the library is neglected. The performance of (b) with this arrangement shows that the center point is generally not needed for the prediction of uniform co-sputtered libraries. However, the center region is important in case of the multilayer sputtered libraries. Since they were designed to cover the complete ternary composition space with the composition space visible in form of a triangle in the center of the library, the center region is of great importance. In order to improve the prediction for these libraries, 5 initialization areas were placed in the center region, resulting in the arrangement shown in red.

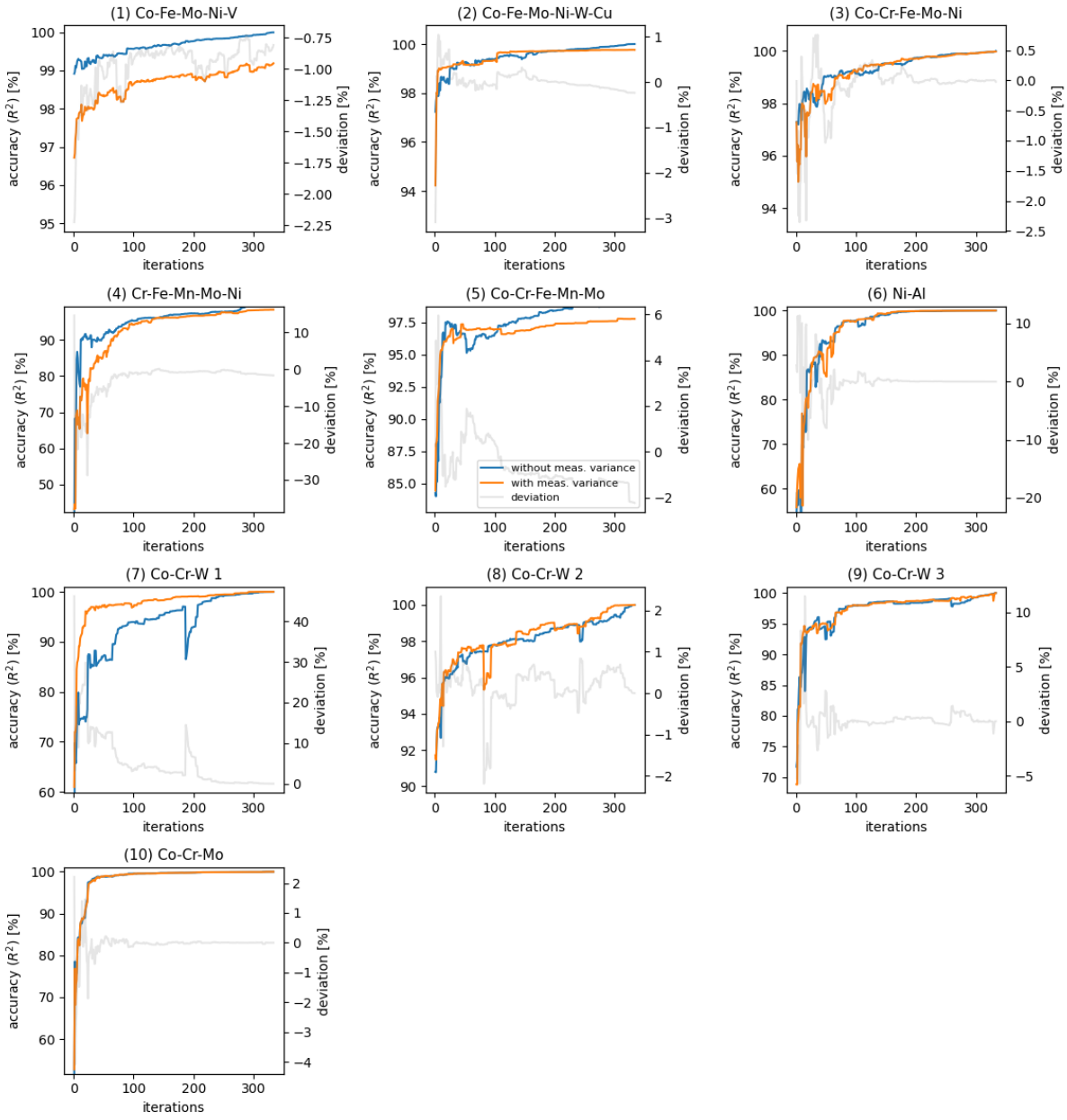


Figure S-4: Comparison of the GP performance on the dataset shown in Figure S-2 with and without including the measurement variance.

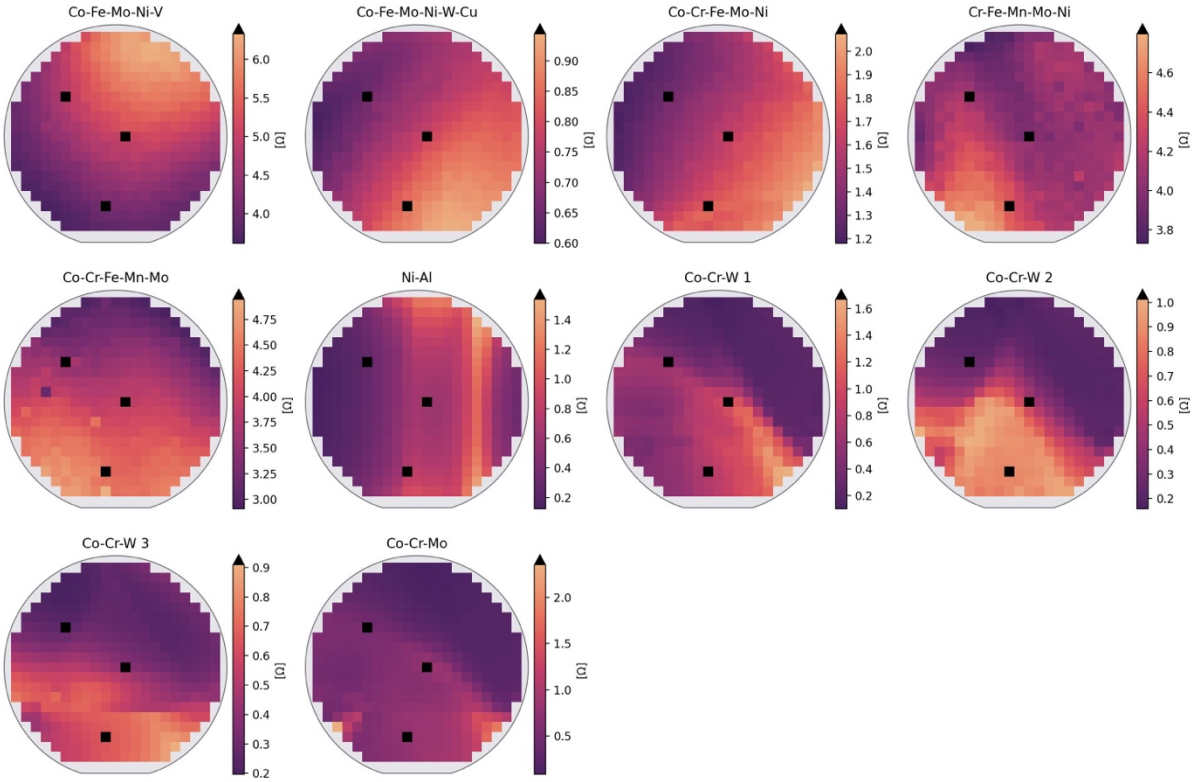


Figure S-5: Dataset with simulated outliers. As a single failed touchdown of the pins should be simulated, ten out of 30 resistance measurements were exchanged by outliers chosen randomly between $0.8 - 1.2\ \mathrm{M\Omega}$. The position of the outliers was fixed to ensure comparability. Each plot shows the mean resistance at each measurement area.

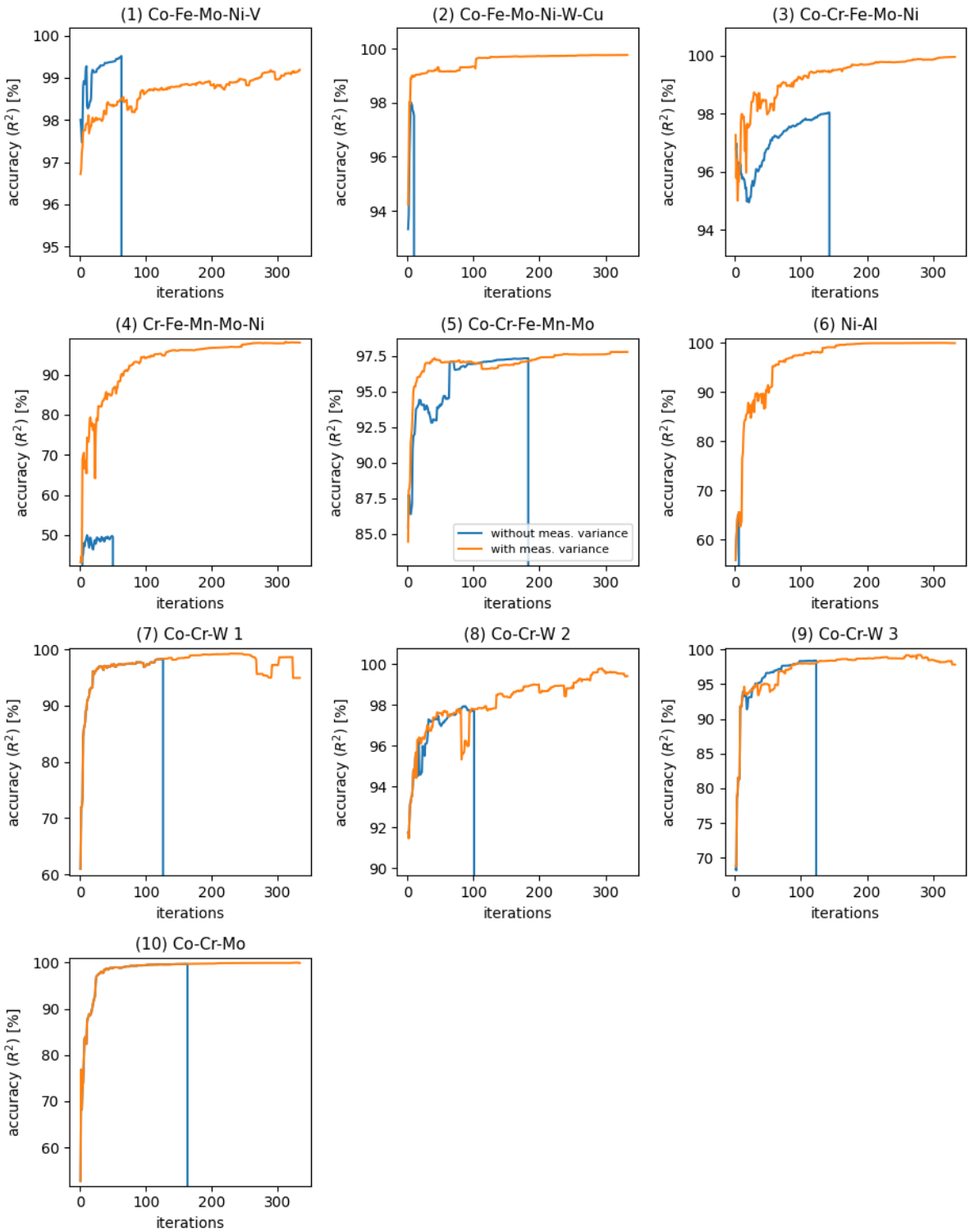


Figure S-6: Visualization of the resistance distributions with randomly added measurement noise. As soon as the vanilla GP encounters an outlier, the prediction fails, while with incorporating the measurement noise into the model, the GP is able to skip the outliers.

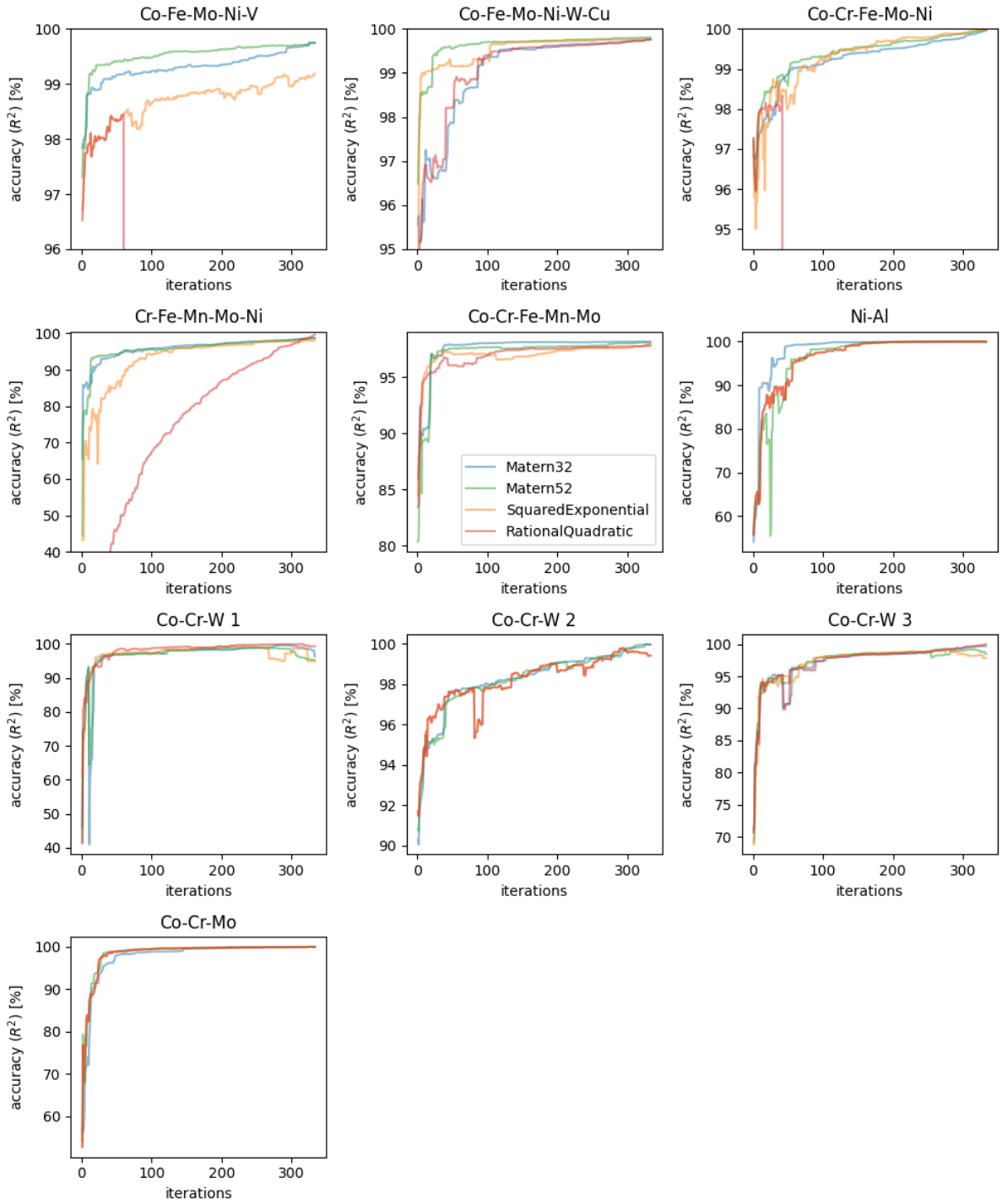
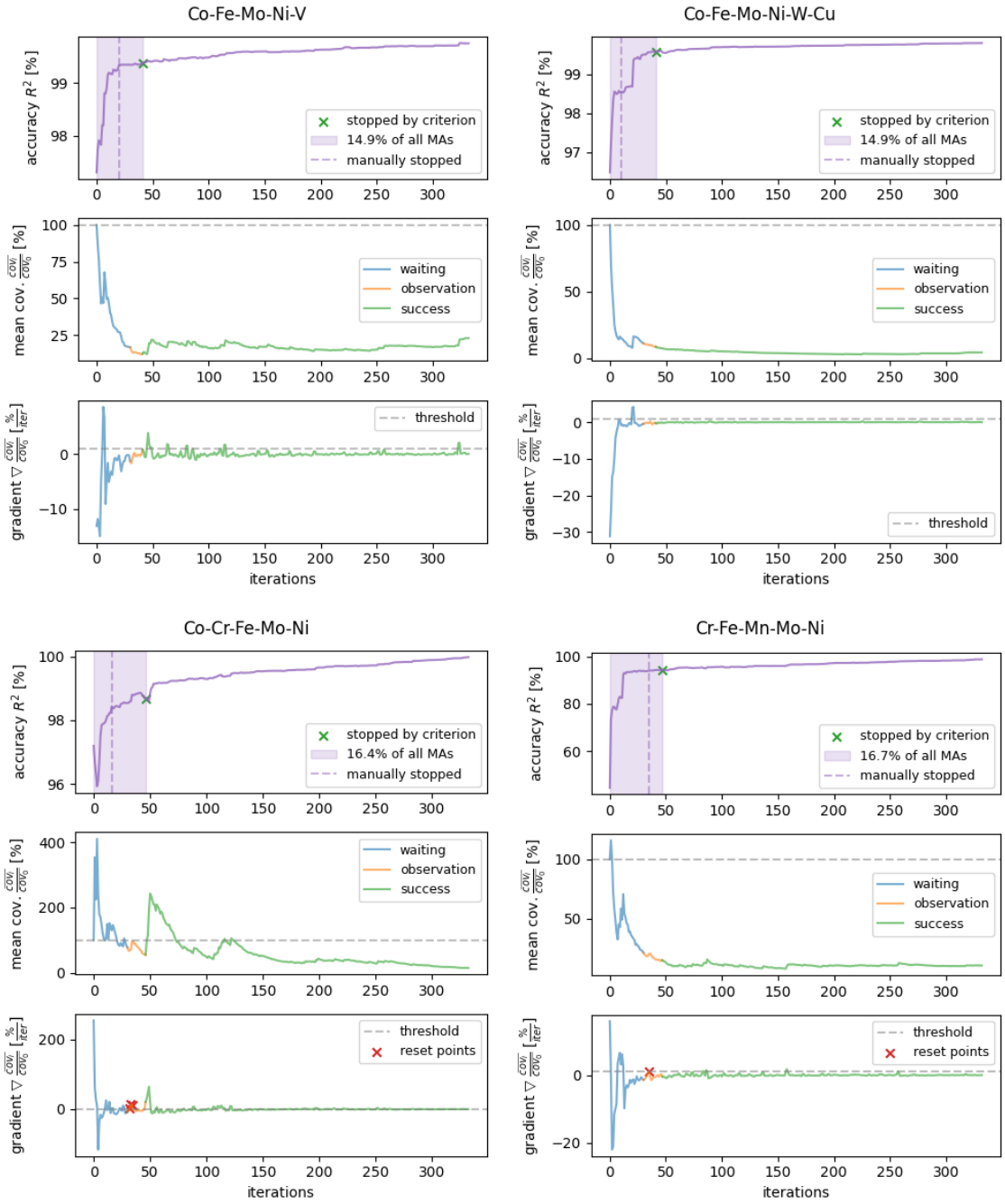
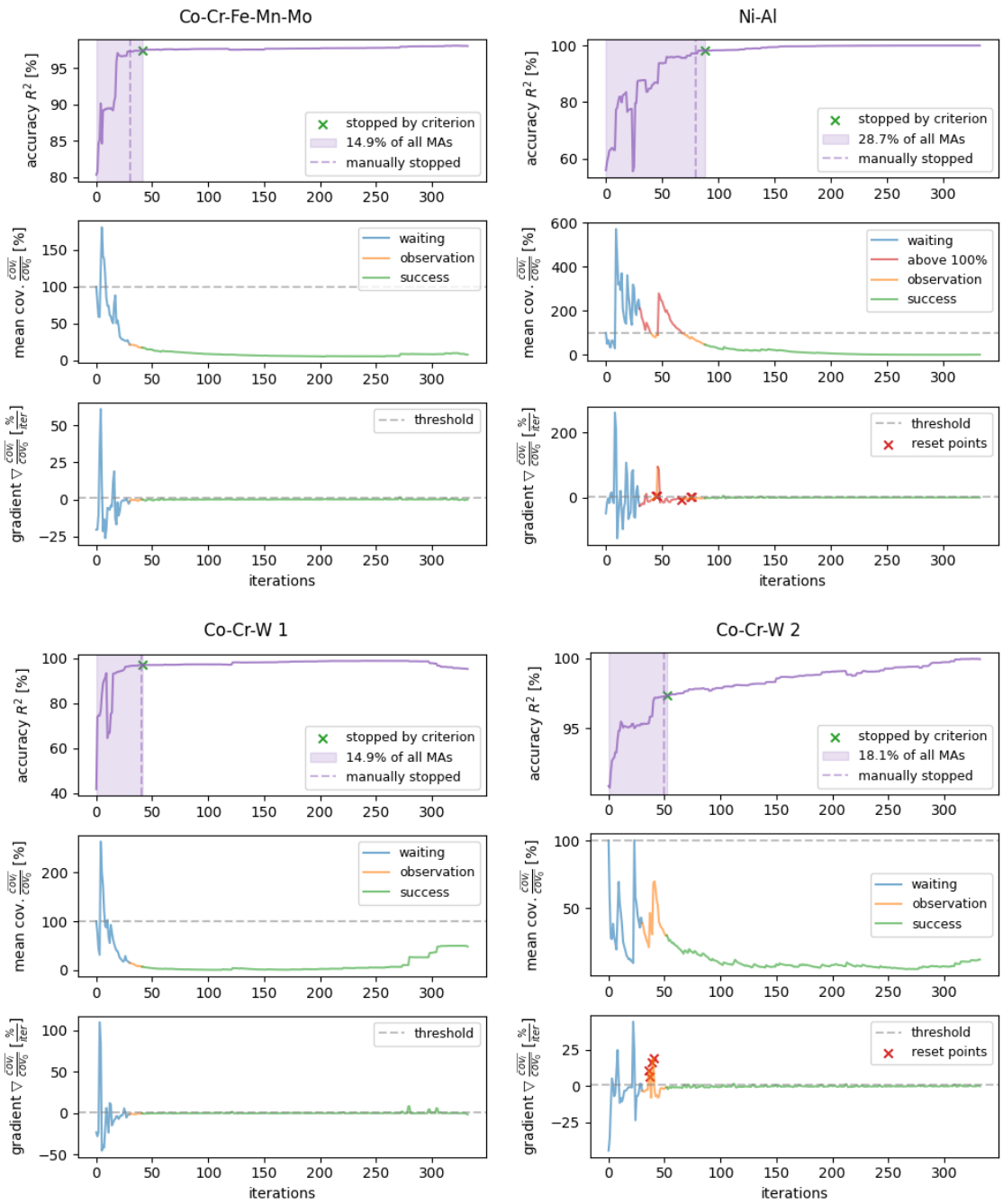
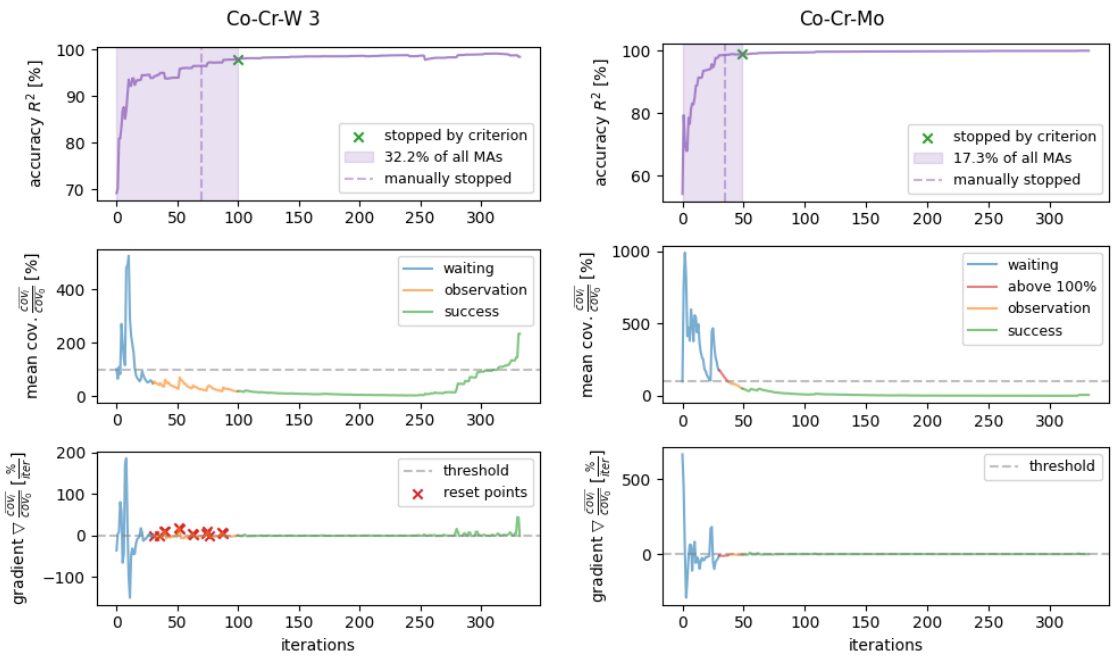


Figure S-7: Performance of the autonomous measurement with different Gaussian process kernels. Two kernels of the Matérn kernel class, the squared exponential as well as the rational quadratic kernel are compared.

Figure S-8: Visualization of the developed stopping criterion for all tested materials libraries. The accuracy of the optimization process, the mean covariance and well as the gradient of the mean covariance is shown over the iterations until all MAs are measured. The stopping iteration is marked in green, while the purple dashed lined shows the stopping iteration determined by observing the accuracy as well as a visual representation of the prediction (the optimal stopping opportunity). The purple range show the percentage of measured values compared to the entire library. The different colors of the mean covariance plot signal the different phases of the stopping criterion.




Co-Cr-Fe-Mn-Mo
Ni-Al
Co-Cr-W 1
Co-Cr-W 2
accuracy R^2 [%]
mean cov. $\frac{\overline{cov}_i}{\overline{cov}_0}$ [%]
gradient $\nabla \frac{\overline{cov}_i}{\overline{cov}_0}$ [$\frac{\%}{iter}$]
iterations
stopped by criterion
14.9% of all MAs
28.7% of all MAs
18.1% of all MAs
manually stopped
waiting
above 100%
observation
success
threshold
reset points


Co-Cr-W 3
Co-Cr-Mo
accuracy R^2 [%]
stopped by criterion
32.2% of all MAs
manually stopped
17.3% of all MAs
mean cov. $\frac{cov}{cov_0}$ [%]
waiting
observation
success
above 100%
gradient $\nabla \frac{cov}{cov_0}$ [$\frac{\%}{iter}$]
threshold
reset points
iterations