Supporting Information

Machine-learning-assisted Prediction of Long-term Performance Degradation on Solid Oxide Fuel Cell Cathodes Induced by Chromium Poisoning†

Kaichuang Yang,[a,b‡] Jiapeng Liu,[c,d‡] Yuhao Wang,[c,d‡] Xiangcheng Shi,[e] Jingle Wang,[f] Qiyang Lu,*[b,g] Francesco Ciucci,*[c,d,h,i] and Zhibin Yang,*[f]

[a.] School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310058, P. R. China

[b.] School of Engineering, Westlake University, Hangzhou, 310024, P. R. China

[c.] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, P. R. China

[d.] Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, P. R. China

[e.] Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore, 117543, Republic of Singapore

[f.] Research Center of Solid Oxide Fuel Cell, China University of Mining & Technology-Beijing, Beijing, 100083, P. R. China

[g.] Key Laboratory of 3D Micro/Nano Fabrication and Characterization of Zhejiang Province, Westlake University, Hangzhou 310024, P. R. China

[h.] Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, P. R. China

[i.] HKUST Energy Institute, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, P. R. China

† Electronic supplementary information (ESI) available.
‡ These authors contributed equally to this work.

*Corresponding authors: luqiyang@westlake.edu.cn; francesco.ciucci@ust.hk; yangzhibin0001@163.com
Keywords: Solid oxide fuel cells; Cathode degradation; Chromium poisoning; Machine learning

Implementation details of deep-DRT

Under the framework of distribution of relaxation times (DRT) analysis, impedance, Z_{DRT}, can be calculated using the equation below:

$$Z_{DRT}(\gamma, f) = i2\pi f L_0 + R_\infty + \int_{-\infty}^{\infty} \frac{\gamma(\ln \tau)}{1 + i2\pi f\tau} d\ln\tau \quad (1)$$

where f is the frequency, τ is a timescale variable, R_∞ is the ohmic resistance, γ is the distribution function of relaxation time, and L_0 is inductance. In the deep-DRT framework, two deep neural networks, i.e., the $\gamma-$net and the RL$-$net were constructed, as illustrated in Figure S4. Following the definition in deep-DRT, the experimental impedance at one frequency point, $Z_{exp}(f)$, can be re-written as $Z_{exp}(f,\psi)$, where ψ represents the state variables, e.g., temperature T, pressure p, materials (for fuel cell, ψ can be $\left[T, p_{O_2}\right]^\top$). It was assumed that both R_∞ and L_0 in (1) are only dependent on the state variables ψ, i.e., $R_\infty = R_\infty(\psi)$, $L_0 = L_0(\psi)$, whereas the DRT, γ, is a function of both timescale τ and ψ as $\gamma = \gamma(\log\tau,\psi)$, as demonstrated in Figure S4. As a result, the impedance by deep-DRT, $Z_{DRT}(f,\psi)$, in the vector form, can be written as

$$Z_{DRT}(f,\psi) = i2\pi L_0(\psi) f + R_\infty(\psi) 1 + A\,\gamma(\log\tau,\psi) \quad (2)$$

where $Z_{DRT}(f,\psi) = \left[Z_{DRT}(f_1,\psi), Z_{DRT}(f_2,\psi), \ldots, Z_{DRT}(f_M,\psi)\right]^\top$ is the impedance vector calculated by DRT formula at $f = [f_1, f_2, \ldots, f_M]^\top$ and ψ; $\gamma(\log\tau,\psi) = \left[\gamma(\log\tau_1,\psi), \gamma(\log\tau_2,\psi), \ldots, \gamma(\log\tau_N,\psi)\right]^\top$; $1 = [1, 1, \ldots, 1]^\top \in R^M$; and $A = A_{re} + iA_{im}$, with A_{re} and $A_{im} \in R^{M\times N}$ defined elsewhere[1, 2]. For a detailed derivation of the above equation, one is encouraged to read the published work[3]. Finally, the parameters of the deep-DRT network are optimized by minimizing the loss function, or the mean squared error (MSE), defined as

$$L(\theta_{RL},\theta_\gamma) = \frac{1}{K}\sum_{k=1}^{K} \left\| Z_{exp}(f,\psi_k) - i2\pi L_{0,\,DNN}(\psi_k,\theta_{RL})\, f - R_{\infty,\,DNN}(\psi_k,\theta_{RL}) 1 - A\,\gamma_{DNN}(\log\tau,\psi_k,\theta_\gamma) \right\|_2^2 \quad (3)$$

where ψ_k represents the state variables for k-th experiment, K is the total number of experiments for training, $L_{0,\ DNN}(\psi_k,\theta_{RL})$ and $R_{\infty,\ DNN}(\psi_k,\theta_{RL})$ are the outputs of RL-net with k-th experimental input, $\gamma_{DNN}(\log\tau,\psi_k,\theta_\gamma)$ is the output of $\gamma-$net with k-th experimental input, θ_{RL} and θ_γ are the network parameters of RL-net and $\gamma-$net, respectively.

To construct the state variables ψ for each experiment, we included three key components, i.e., temperature T, time t, and the feature of transition metal element substitute e. To be specific, the ψ is constructed as $\psi=[\psi_T,\psi_t,\psi_e]^\top$, where ψ_T, and ψ_t are normalized features of T and t, ψ_e is the elemental embedding. Here, we only use the feature of transition metal M to represent the cathode material. In particular, e is a vector with 16 entries from the MEGNet[4]. Besides, the T and t are normalized to be in the range of [0,1]. All other settings are consistent with the one implemented in deep-DRT[3] except for the number of hidden layers in this work is 3 for both $\gamma-$net and RL$-$net. The loss, or MSEs of the training dataset and validation dataset as a function of iteration numbers are shown in Figure S5. The MSEs of the training and validation datasets become stable after around 40,000 iterations. Additionally, no overfitting was observed. As a result, we will use the model after 50,000 iterations to predict the EIS spectrum as well as their corresponding DRT. The selected results for the test dataset are shown in Figure S6. Here we randomly show 2 cases to observe the model performance on unseen EIS spectra. The model can well predict the EIS response for different cathode materials at different temperatures and different measuring times. Considering that the EIS measurements are done for a variety of cathode materials, spanning a wide period, we believe such a model can work successfully to predict the EIS measurement under such a situation.

We notice that the predicted DRT for each spectrum consists of only one main peak, which may deviate from reality. To further improve the model performance, we would like to use the deep-DRT to predict the relative DRT rather than the true one. In specific, we will first obtain the DRT using DRTTools[5], and use the deep-DRT to predict the coefficient with respect to the one

from DRTTools. In detail, the pre-DRT is obtained by ridge regression (RR) as implemented in DRTTools using piece-wise linear basis functions together with the 2nd derivative. Later, the deep-DRT outputs the relative values. Hence, the final impedance by DRT is expressed as

$$Z_{DRT}(f,\psi) = i2\pi L_{0,\,DNN}(\psi,\theta_{RL})L_{0,\,RR}f - R_{\infty,\,DNN}(\psi,\theta_{RL})R_{\infty,\,RR}1 - A\,\gamma_{DNN}(\log\tau,\psi,\theta_{\gamma}) \otimes \gamma_{RR} \quad (4)$$

where $\otimes$ denotes the element-wise multiplication for two vectors, $L_{0,\,RR}$, $R_{\infty,\,RR}$, and γ_{RR} are obtained from DRTTools. One should note that the $L_{0,\,DNN}$, $R_{\infty,\,DNN}$, and γ_{DNN} now represent the relative values with respect to the ones obtained by RR. Based on the pre-trained DRT, the losses of both the training dataset and validation dataset drop significantly. The MSEs as a function of iterations are shown in Figure S7. As we can notice, the absolute values of MSEs at the initial iteration are much smaller in comparison to the model from scratch. This is mainly attributed to the well-behaved DRT by ridge regression. Furthermore, we notice that the MSEs become stable only after 10,000 iterations, a shorter period than the 50,000 iterations needed for the previous model. Finally, the stable MSEs are much smaller than the ones for the model from scratch, implying that the pre-training with ridge-regression helps improve the model performance further. This suggests that the ridge-regression can benefit the model convergence as well as model performance.

The code was implemented using the Pytorch [6] The activation functions were chosen to be non-saturating exponential linear (ELU) units[7] for the first three layers and softplus for the last layer. The weights and biases were initialized using the Xavier uniform method[8] and to zero, respectively. Furthermore, the parameters of the network θ were optimized using the Adam algorithm[9] with a learning rate of 10^{-5} and a maximum of 50,000 iterations. The structure of deep-DRT in this work is slightly different from the one published[3]. The number of hidden layers in this work was set to 3 rather than 5. This was verified by comparing the final MSE with different hidden layers after fixed iterations (50,000) with identical settings.

As explained in the main text, we used only the transition metal M to represent the cathode material. The transition metal element was assigned with a vector with 16 entries obtained from the MEGNet.[4] It was demonstrated that the elemental embeddings show good behavior in differentiating the elements in the periodic table. For temperature and time encoding, the real values were normalized into the range of $[0,1]$ following

$$\psi_T = \frac{T_{max} - T}{T_{max} - T_{min}} \tag{5}$$

$$\psi_t = \frac{t}{t_{last} - t_0} \tag{6}$$

For one example, the state variable ψ for $SrFe_{0.75}M_{0.25}O_{3-\delta}$ (M = Co) at 700 °C, 24 h is represented as $\psi = [\psi_T, \psi_t, \psi_e]^\top = [1.0, 0.25, \psi_{Co}]^\top$, where ψ_{Co} is the embedding array of Co with 16 entries.

For the DRT calculations, the γ vector is evaluated between the 10^{-7} s and 10^3 s with 301 equispaced data points in the log-scale, i.e., the τ vector has a length of 301. To keep consistent with the other input features, the $\log \tau_k$ was normalized between $[-1,1]$ by

$$\log \tau_k := -1 + 2\frac{\log \tau_k - \min(\log \tau_k)}{\max(\log \tau_k) - \min(\log \tau_k)} \tag{7}$$

Specifically, we predict the latent γ vector by predicting each γ_k at the timescale of $\log \tau_k$.

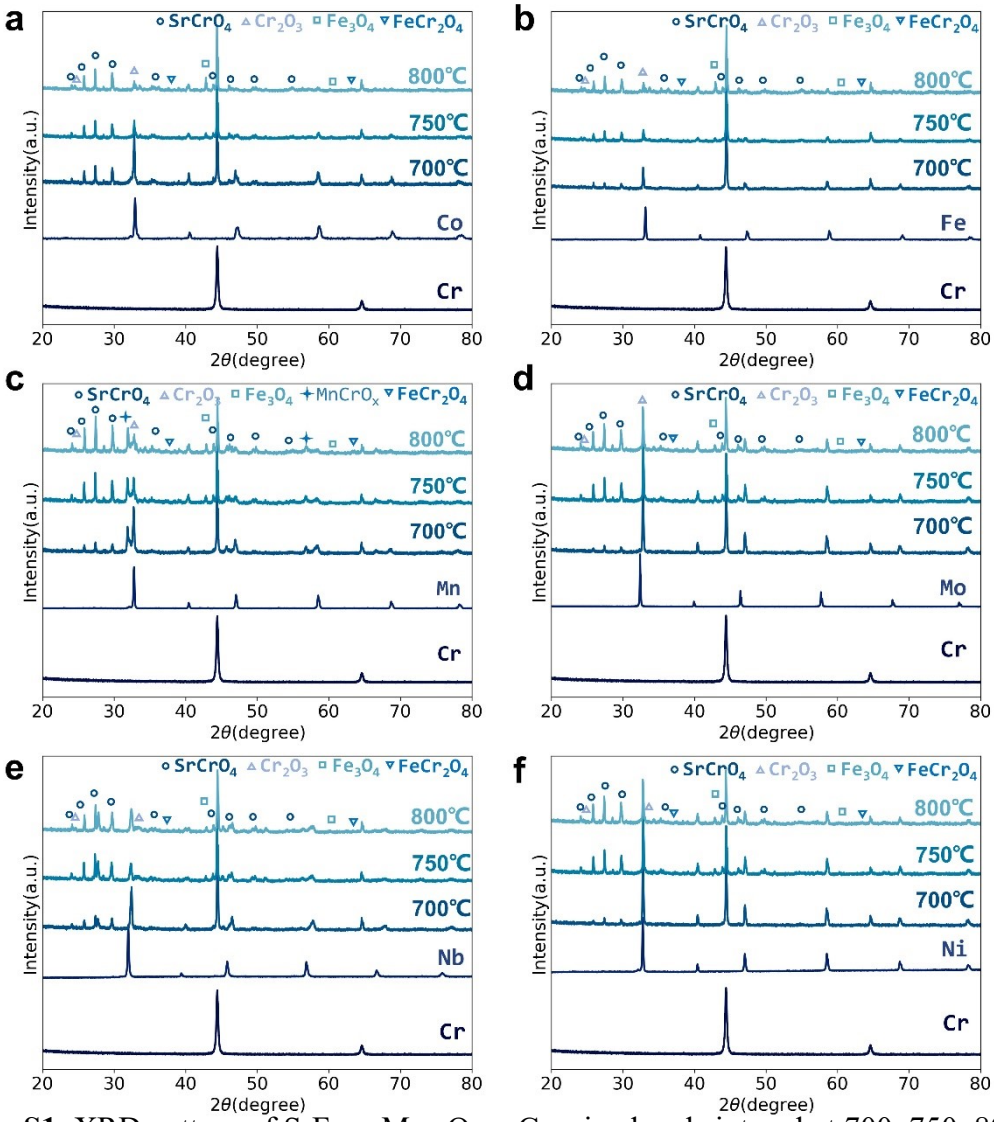

Figure S1. XRD pattern of $SrFe_{0.75}M_{0.25}O_{3-\delta}$ -Cr mixed and sintered at 700, 750, 800 °C for 5 hours. M= (a) Co, (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni.

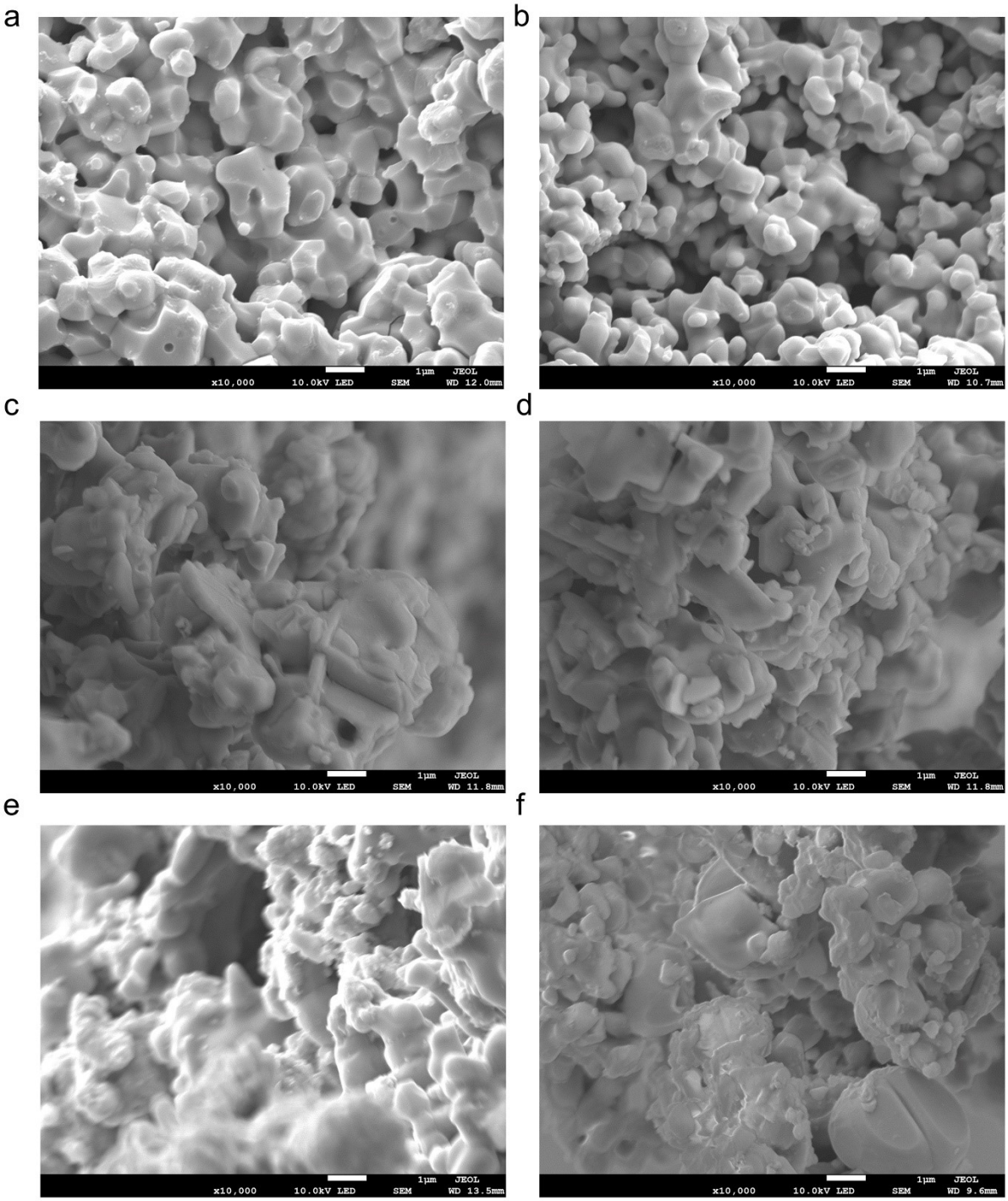

Figure S2. SEM images of $SrFe_{0.75}M_{0.25}O_{3-\delta}$, M= (a) Co, (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni.

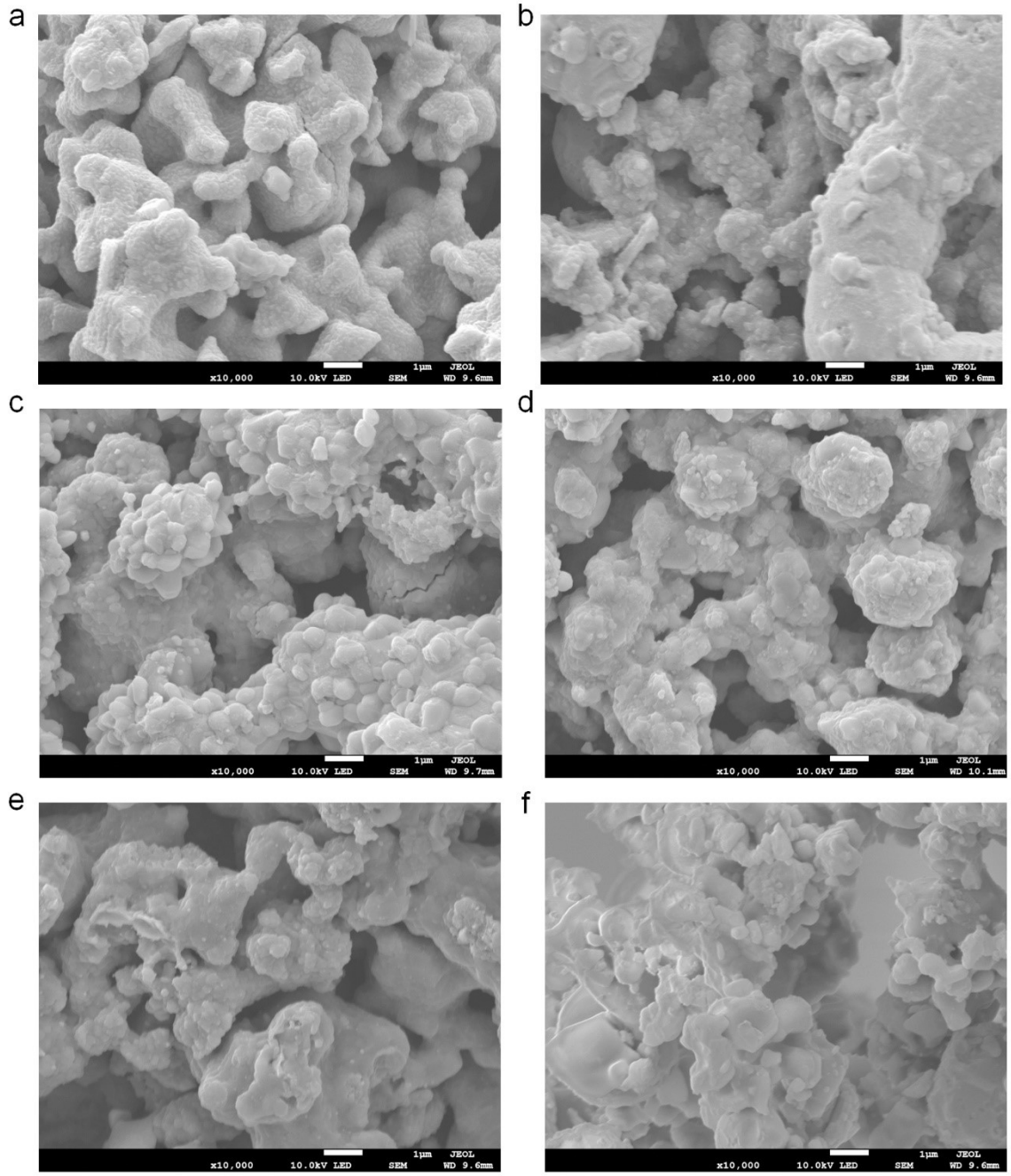

Figure S3. SEM images of $SrFe_{0.75}M_{0.25}O_{3-\delta}$ after 96 h Cr deposition and poisoning test, M= (a) Co (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni.

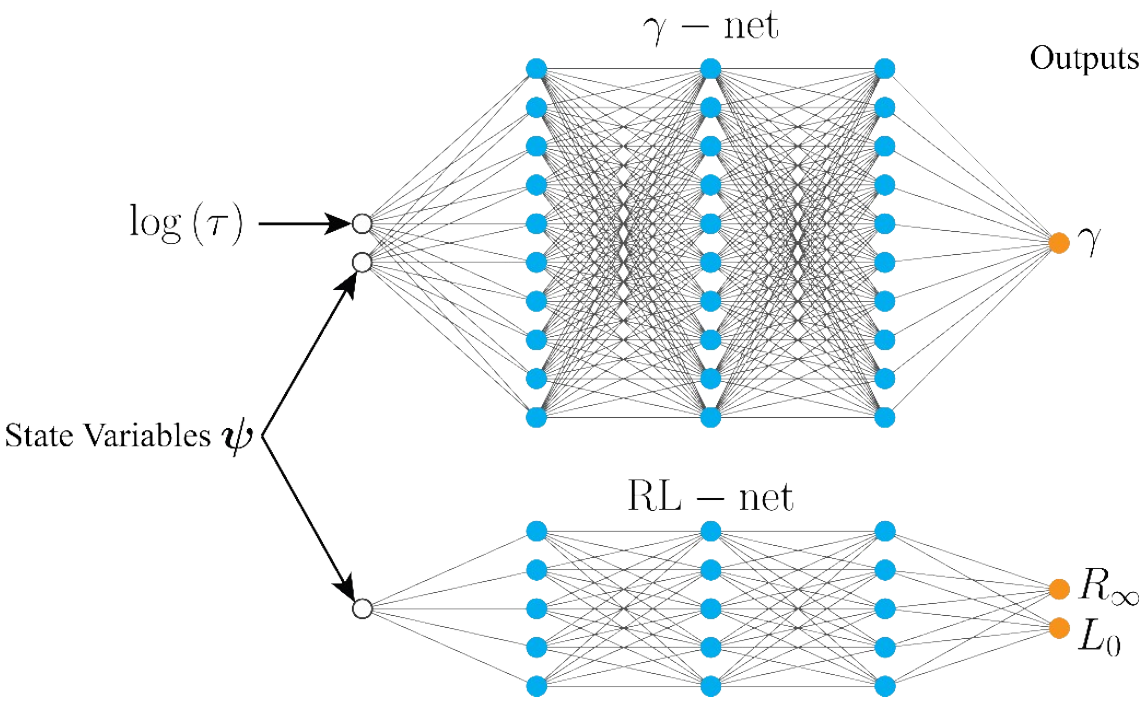

Figure S4. Schematic illustration of deep-DRT network structure. Two deep neural networks are constructed: the upper one is $\gamma-$net, and the lower one is RL$-$net. The $\gamma-$net takes the inputs of parameters combining $\log(\tau)$ and state variables ψ and outputs the DRT at the timescale τ. The RL$-$net takes only the state variables ψ as input and outputs the R_∞ and L_0.

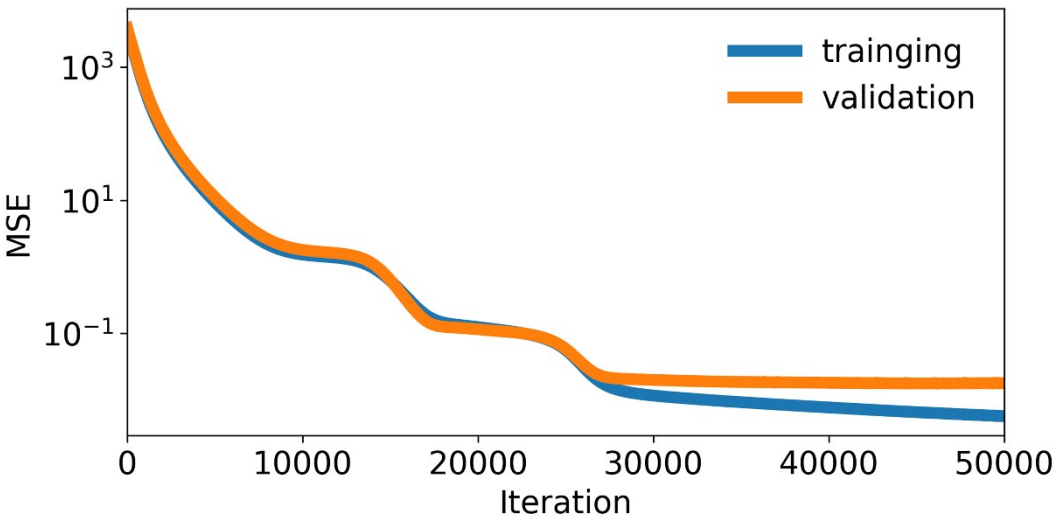

Figure S5. MSE of training and validation datasets as a function of the iteration numbers.

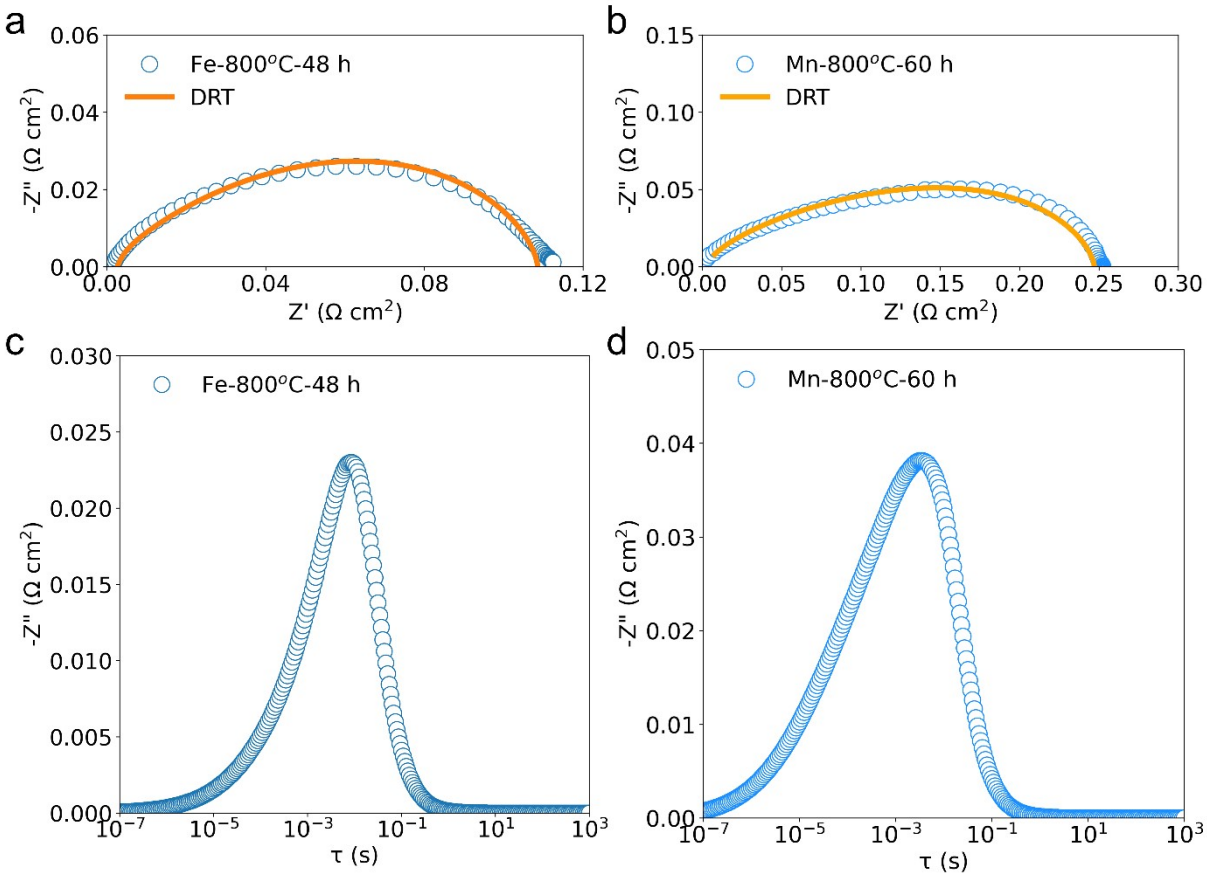

Figure S6. Nyquist plot of the EIS response of (a) Fe-800-48 and (b) Mn-800-60. The experimental results are labeled with dots while the predicted values by DRT are labeled in lines. (c) and (d) correspond to the predicted DRT using the gamma-net for (b) Fe-800-48, and (d) Mn-800-60, respectively.

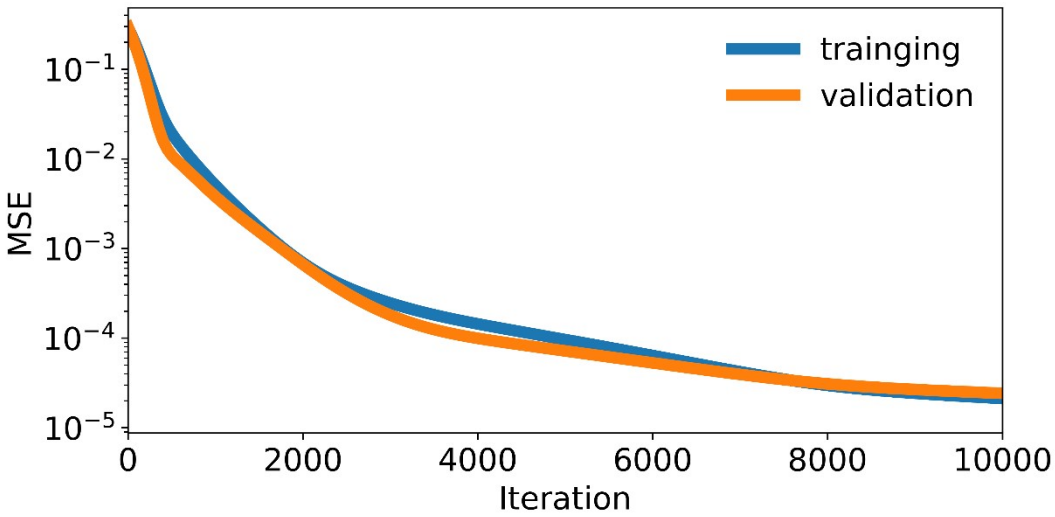

Figure S7. MSE of training and validation datasets as a function of the iteration numbers for the model with pretrained DRT by ridge regression.

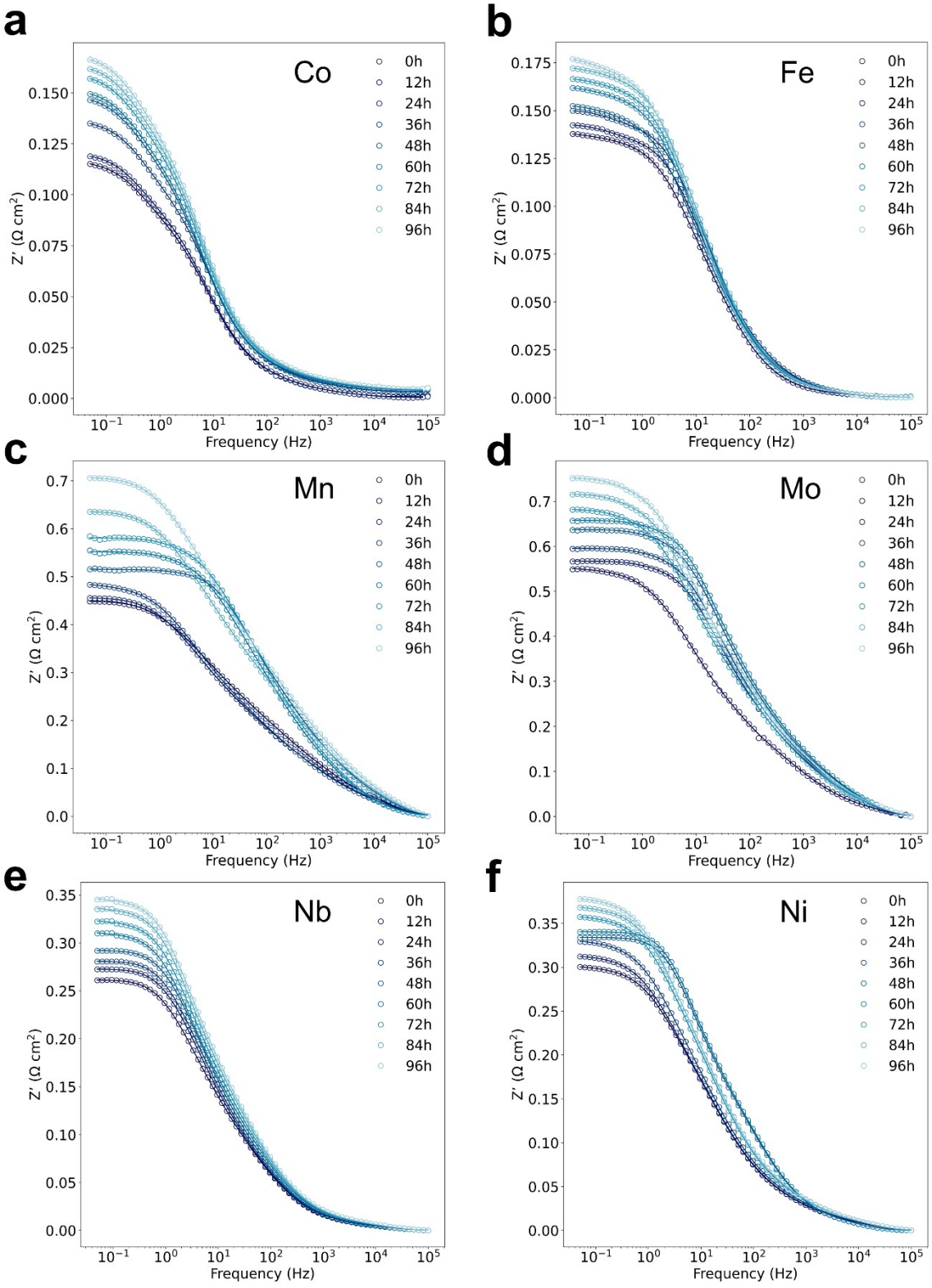

Figure S8. EIS response of SFMO/LSGM/SFMO (M=(a) Co, (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni) at 700°C. The experimental results are labeled with dots while the simulated values by deep-DRT are labeled in color lines.

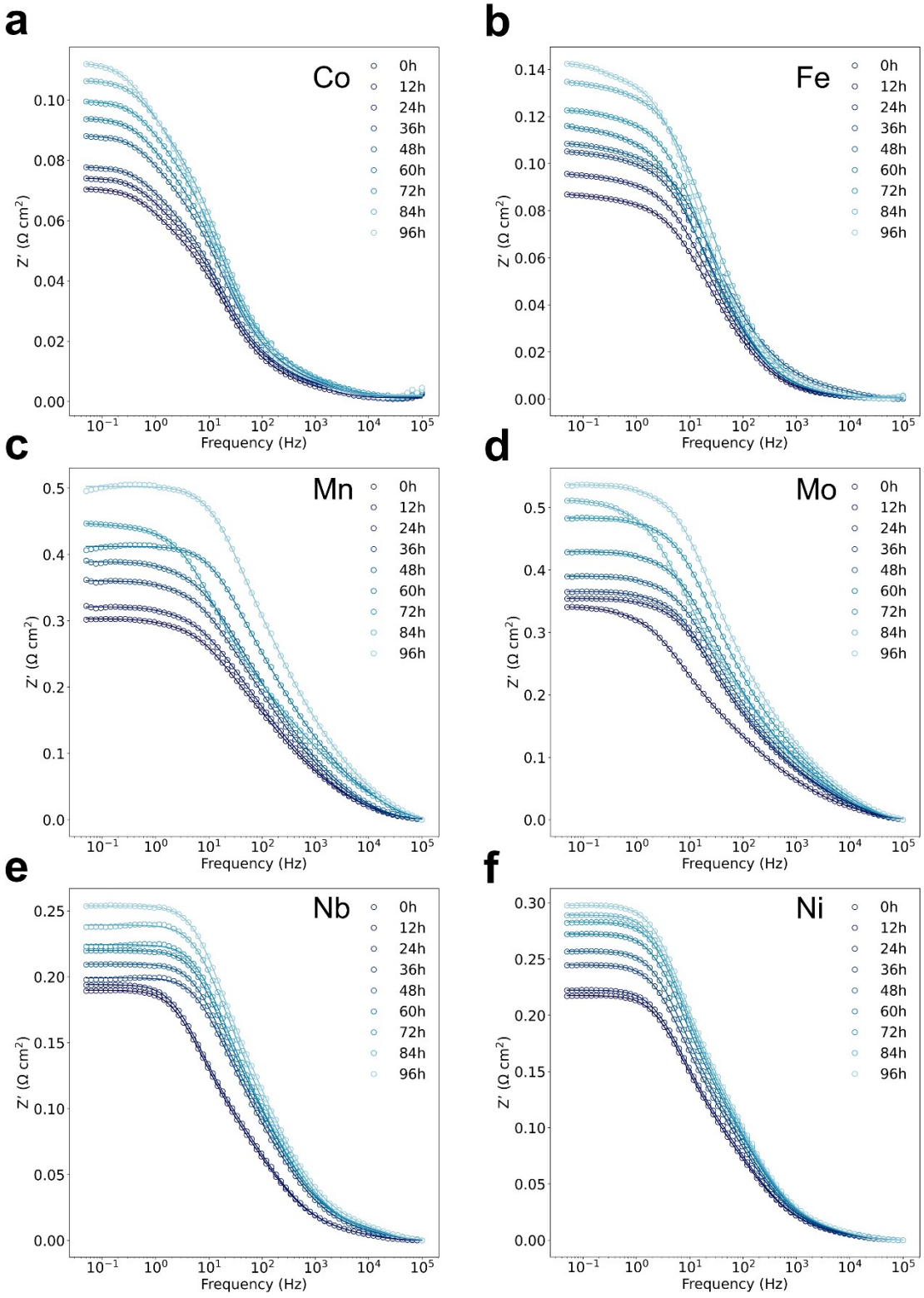

Figure S9. EIS response of SFMO/LSGM/SFMO (M=(a) Co, (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni) at 750°C. The experimental results are labeled with dots while the simulated values by deep-DRT are labeled in color lines.

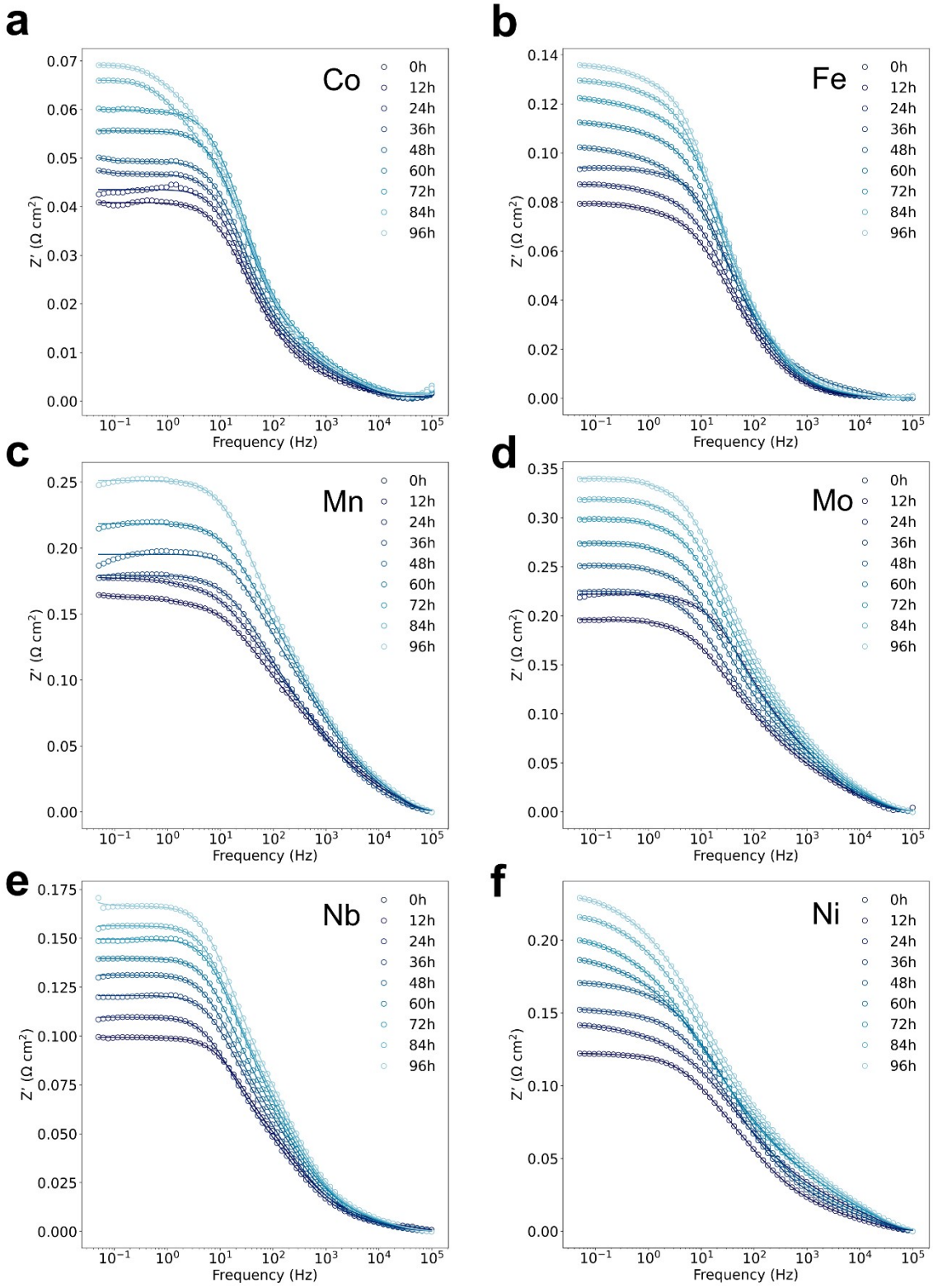

Figure S10. EIS response of SFMO/LSGM/SFMO (M=(a) Co, (b) Fe, (c) Mn, (d) Mo, (e) Nb and (f) Ni) at 800°C. The experimental results are labeled with dots while the simulated values by deep-DRT are labeled in color lines.

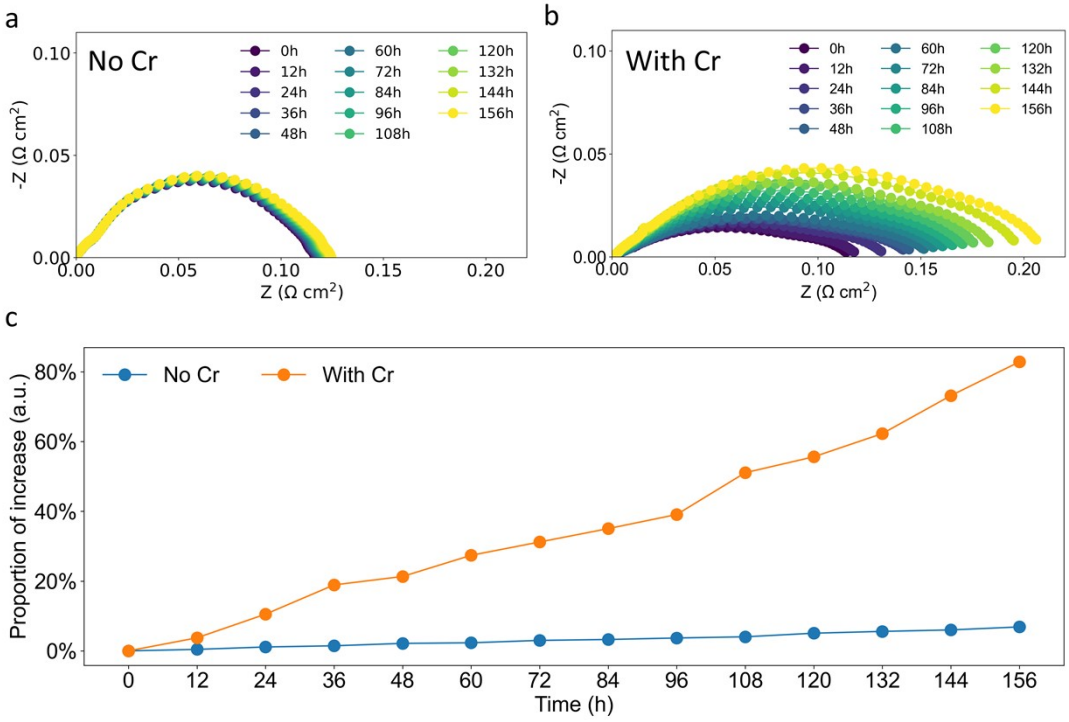

Figure S11. (a) Nyquist plots of the EIS response of $SrFe_{0.75}Co_{0.25}O_{3-\delta}$ (SFCO)/LSGM/SFCO tested at 700°C for 156 h without Cr source and (b) with Cr source; (c) The proportion of the R_p increase calculated by the data from (a) and (b).

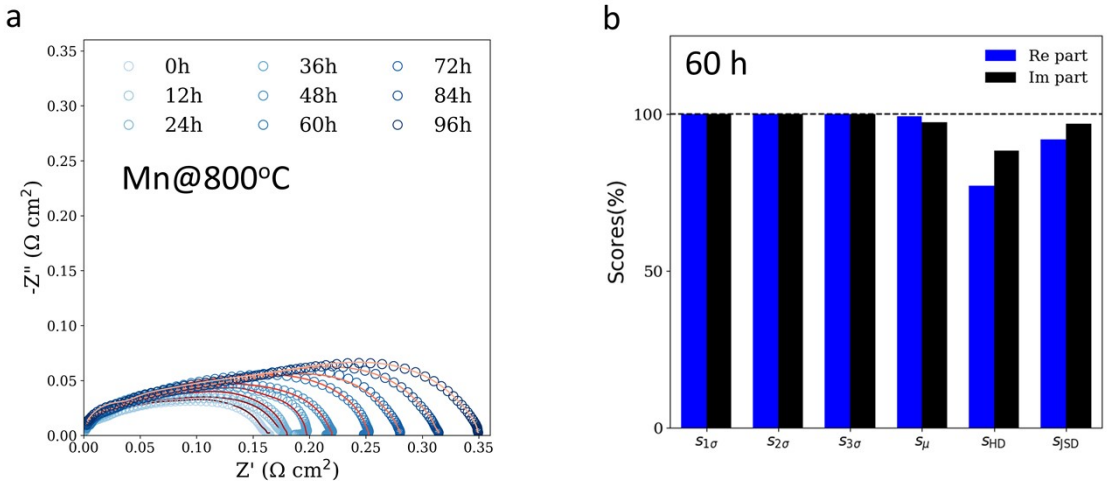

Figure S12 (a) The EIS of $SrFe_{0.75}Mn_{0.25}O_{3-\delta}$/LSGM/$SrFe_{0.75}Mn_{0.25}O_{3-\delta}$ at 800°C (calculated by Kramers-Kronig (dots) and the experimental data (solid lines)), (b) Kramers-Kronig test analysis of $SrFe_{0.75}Mn_{0.25}O_{3-\delta}$/LSGM/$SrFe_{0.75}Mn_{0.25}O_{3-\delta}$ at 800°C for 60 h

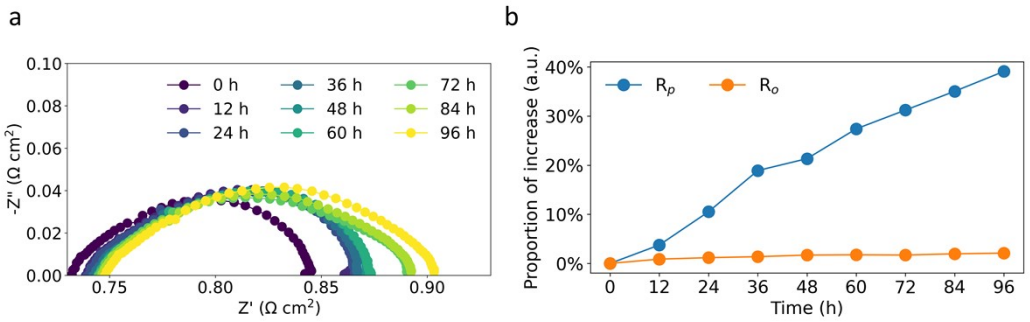

Figure S13. (a) Nyquist plot of the EIS response of SFCO/LSGM/SFCO tested at 750°C for 96 hours. (b) the relative changes (in percentage) of R_p and R_o as a function of elapsed testing time.

Table S1. Common chromium containing interconnect materials and their chromium weight percentages[10].

Material Name	Weight % Cr
Crofer22 APU	23
Crofer22 H	20-24
SUS430 Nisshin Steel Co. Ltd.	16.03
430SS	17
SUS420 Nippon Steel Corp., Japan	16-18
Ducrolloy ($Cr_5Fe_1Y_2O_3$)	94
Inconel 600	16
RA446	23-27
ZMG232 Hitachi Metals Co. Ltd.	22
17-4 PH SS	15.21
E-Brite	26
RA600	14-17
FeCM	16.65
ITM/ITM-14	26

References

1. M. Saccoccio, T. H. Wan, C. Chen and F. Ciucci, *Electrochimica Acta*, 2014, **147**, 470-482.
2. J. Liu, T. H. Wan and F. Ciucci, *Electrochimica Acta*, 2020, **357**, 136864.
3. E. Quattrocchi, T. H. Wan, A. Belotti, D. Kim, S. Pepe, S. V. Kalinin, M. Ahmadi and F. Ciucci, *Electrochimica Acta*, 2021.
4. C. Chen, W. Ye, Y. Zuo, C. Zheng and S. P. Ong, *Chemistry of Materials*, 2019, **31**, 3564-3572.
5. T. H. Wan, M. Saccoccio, C. Chen and F. Ciucci, *Electrochimica Acta*, 2015, **184**, 483-499.
6. A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga and A. Lerer, 2017.
7. D.-A. Clevert, T. Unterthiner and S. Hochreiter, *eprint arXiv:1511.07289*, 2015, arXiv:1511.07289.
8. X. Glorot and Y. Bengio, 2010/03/31, 2010.
9. D. P. Kingma and J. Ba, *eprint arXiv:1412.6980*, 2014, arXiv:1412.6980.
10. L. Zhou, J. H. Mason, W. Li and X. Liu, *Renewable and Sustainable Energy Reviews*, 2020, **134**, 110320.