AI-Assisted Descriptor Discovery for Electrochemical Interfacial States via Latent Organization of High-Throughput EIS

Anonymous Authors[1]

Abstract

Measurements do not automatically yield meaningful descriptors when existing theory cannot uniquely determine how complex observations should be reduced. Using high-throughput electrochemical impedance spectroscopy as a model case, we present a human-guided, AI-assisted workflow in which a conditional autoencoder organizes electrochemical responses into a latent space for descriptor construction and comparison, while experimental potential of zero charge (PZC) is used only as an external validation reference. The learned representation remained strongly ordered by electrochemical bias, enabling the construction of trajectory-derived descriptor candidates indexed by applied bias (V_b). Grounded comparison among these candidates identified an anchor-relative sine-like trajectory coordinate as the most credible scalar summary within the tested family. These results show that latent organization can support auditable descriptor discovery from complex electrochemical measurements in underdetermined settings.

1. Introduction

Recent discussion in AI for science has increasingly focused on autonomous or semi-autonomous systems that can plan experiments, navigate search spaces, and accelerate discovery through closed-loop decision making (Jumper et al., 2021; Merchant et al., 2023; Szymanski et al., 2023; Tom et al., 2024). That direction has been highly productive, but it does not capture the full range of scientific bottlenecks. In many domains, the limiting factor is not the lack of measurements or experimental throughput, but the difficulty of turning complex observations into scientifically meaningful variables. Rich data alone do not guarantee interpretable descriptors, especially when existing theory does not uniquely determine how those observations should be reduced. In such settings, the role of AI is not best framed as replacing scientific judgment with autonomy. A more useful role may be to reorganize complex measurements into a structured candidate space that scientists can inspect, compare, and validate against grounded external references (Tom et al., 2024; Szymanski et al., 2023).

Electrochemical impedance spectroscopy (Lazanas & Prodromidis, 2023; Chang & Park, 2010; Bard et al., 2022) provides a clear example of this problem. High-throughput impedance measurements can capture intertwined contributions from double-layer behavior, charge transfer, adsorption, transport, and other interfacial non-idealities. Yet physically plausible scalar summaries remain difficult to define without imposing strong modeling assumptions. Equivalent-circuit analysis remains valuable because it offers compact interpretability, but it also constrains the analysis to a predefined model class and can leave much of the measured response underused. As impedance measurements become denser and more systematically varied, the challenge shifts from merely fitting spectra to organizing them in a way that can support descriptor construction under scientific scrutiny (Ragoisha & Bondarenko, 2005; Pajkossy & Mészáros, 2020; Fortin et al., 2022).

This study presents a human-guided, AI-assisted workflow for that underdetermined setting. Rather than treating AI as an autonomous endpoint predictor, we use representation learning to organize high-throughput electrochemical impedance data into a latent space from which descriptor candidates can be derived and compared. Human judgment remains explicit in the problem formulation, acquisition logic, and validity criteria, while machine learning is used to expose structure that would be difficult to rank directly from raw spectra. Experimental potential of zero charge (PZC) is used only as an external validation reference, not as a supervised training target (Bard et al., 2022). Within this framework, we ask whether the learned representation can support auditable descriptor discovery from complex electrochemical responses and whether grounded comparison can distinguish better-supported one-dimensional summaries

---

[1] Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

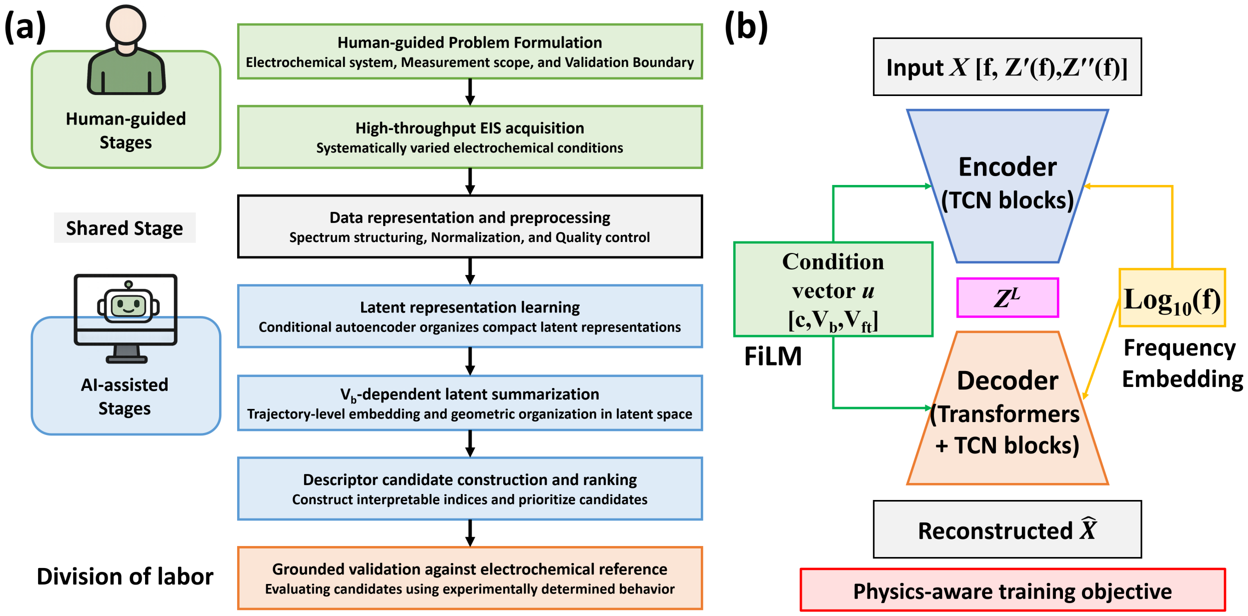


Figure 1. Human-guided workflow and model architecture for descriptor discovery from high-throughput electrochemical impedance data. (a) Schematic of the division of labor in the present study. Domain knowledge defines the problem formulation, acquisition logic, and validation boundary, while AI-assisted modeling organizes high-throughput impedance observations in latent space and supports descriptor construction, ranking, and reference-based validation. (b) Conditional autoencoder architecture (Hinton & Salakhutdinov, 2006; Sohn et al., 2015) for sequential impedance data. The model combines condition injection, frequency embedding, and an encoder-decoder backbone to learn compact latent structure and reconstruct the input spectrum under a physics-aware training objective.

from weaker alternatives. The contribution of this work is to show how AI can function as a scientifically useful organizing tool in measurement-rich yet interpretation-limited regimes.

2. Related Work

Electrochemical impedance spectroscopy (EIS) (Lazanas & Prodromidis, 2023; Chang & Park, 2010; Bard et al., 2022) has long been valued because the frequency dependence of impedance can encode interfacial kinetics, double-layer behavior, transport, adsorption, and other coupled non-ideal processes. Equivalent-circuit modeling became the dominant practical framework because it offers a compact route from complex spectra to interpretable parameters. At the same time, impedance analysis has always faced a persistent difficulty: model choice is rarely unique, physically distinct processes can generate similar spectral signatures, and richer measurements do not automatically yield clearer interpretation. EIS has therefore remained a powerful measurement modality with a substantial interpretation bottleneck, especially when information-rich spectra must be reduced to scientifically usable variables.

Recent AI-enabled electrochemistry has developed along two more visible directions (Chen et al., 2023; Bond et al., 2022). One direction has focused on data-driven interpretation of electrochemical signals, particularly in voltammetry and related mechanism-recognition tasks, where deep learning has been used for classification, regression, and automated mechanistic assignment (Chen et al., 2023; Bond et al., 2022; Kennedy et al., 2019; Hoar et al., 2022; Chen et al., 2022). Another direction has centered on automation, active learning, and self-driving experimental platforms, where electrochemical workflows are embedded within broader closed-loop systems for screening, optimization, or mechanistic investigation (Tom et al., 2024; Sheng et al., 2024; Pence et al., 2025). These developments have expanded the role of machine learning in electrochemistry, but they have mainly emphasized either direct prediction from structured signals or autonomous experimental execution.

Impedance-based measurements occupy a somewhat different position within this broader landscape. Several studies (Chang, 2023; Doonyapisut et al., 2023b;a; Schaeffer et al., 2023; Jones et al., 2022; Lv et al., 2024) have started to apply machine learning directly to EIS data, including analysis of potential-dependent spectral sequences, forecasting from full impedance responses, and supervised use of frequency-domain electrochemical data for sensing and state estimation. These efforts are important because they show that impedance spectra can support learning beyond handcrafted circuit parameters alone. Much of this literature, however, has still been oriented toward prediction,

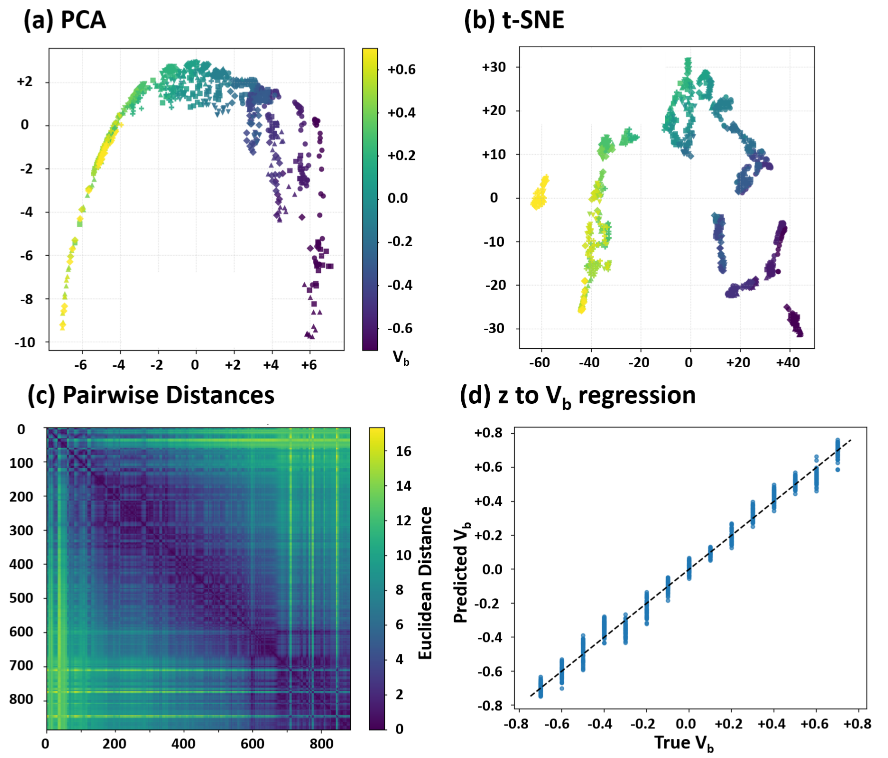


Figure 2. Latent space organization is strongly ordered by electrochemical bias. (a) PCA projection of latent representations, colored by V_b and marked by quantiles of $V_b + V_p$, showing a smooth global ordering primarily along the bias coordinate. (b) t-SNE projection of the same latent representations, indicating that local neighborhoods remain ordered by V_b despite nonlinear embedding. (c) Pairwise latent-distance matrix after sorting samples by V_b, revealing banded structure rather than random mixing. (d) Regression from latent coordinates to V_b, showing that electrochemical bias remains the dominant global ordering variable in the learned representation ($R^2 = 0.989$).

classification, or parameter inference from spectral data.

Comparatively less attention has been given to the intermediate problem addressed here: using representation learning to reorganize high-throughput, impedance-rich observations into a structured latent space that supports descriptor construction and candidate ranking under explicit experimental grounding.

3. Methods and Model Architecture

3.1. Overall framework

The overall workflow contains four linked stages: measurement, latent representation learning, descriptor construction, and external validation (Figure 1). Electrochemical knowledge determines what is measured, which variables are treated as relevant conditions, and what kind of external check is scientifically admissible. Machine learning then operates within those boundaries by restructuring the high-dimensional observations into a latent space in which systematic variation becomes easier to compare. This division of labor is central to the framework. Scientific judgment remains explicit at the level of task definition and interpretive accountability, while AI is used to structure the representation rather than to act as an autonomous decision-maker.

3.2. Data acquisition and representation

The dataset consists of high-throughput electrochemical impedance measurements collected under systematically varied electrochemical conditions. In the present implementation, the measurements are obtained through a high-throughput EIS protocol, but the methodological argument is framed at the level of high-throughput impedance data more broadly (Ragoisha & Bondarenko, 2005; Pajkossy & Mészáros, 2020; Fortin et al., 2022). Each sample includes the relevant experimental conditions, including base potential V_b, perturbation potential V_p, and the frequency-dependent real and imaginary impedance components; the present dataset was acquired at a fixed concentration condition. The spectra are arranged as sequential observations so that each input retains both its conditioning variables and its frequency-resolved response profile. Robust input normalization based on median centering and interquartile-range scaling was applied for stable training, while evaluation was carried out with reference to both normalized and physi-

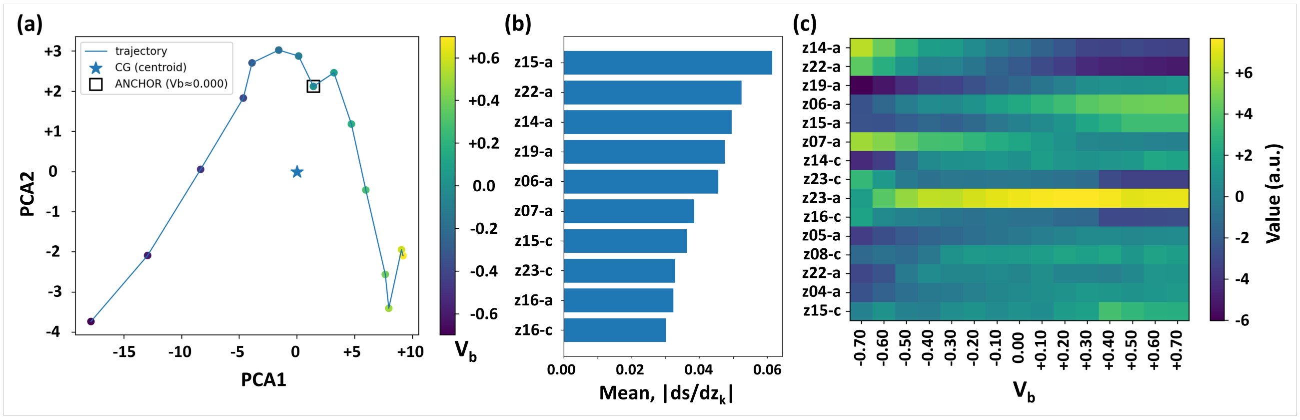


Figure 3. Trajectory-based descriptor-candidate construction from V_b-indexed latent summaries. (a) Trajectory of per-V_b latent summaries in the PCA1-PCA2 subspace, with the global centroid and an anchor near $V_b \approx 0$ used to define oriented trajectory coordinates. (b) Top latent-summary features ranked by mean sensitivity for the anchor-relative sine-like descriptor candidate within the focus region. (c) Heatmap of representative high-variance latent-summary features across V_b. Together, these panels show that descriptor construction is supported by coordinated latent reorganization across multiple features rather than by a single dominant coordinate.

cal space so that reconstruction quality and downstream constraints remained tied to the original electrochemical quantities. Detailed preprocessing, data splitting, and reference matching are provided in the Appendix.

3.3. Model architecture

A conditional autoencoder (Hinton & Salakhutdinov, 2006; Sohn et al., 2015) is used to learn a compact latent representation of the impedance observations (Figure 1b). The encoder is built from temporal convolutional blocks, and the decoder combines Transformer and temporal convolutional components (Bai et al., 2018; Vaswani et al., 2017). Experimental conditions are injected explicitly rather than being left for the model to infer indirectly, and frequency information is supplied through a dedicated embedding pathway (Perez et al., 2018). That architecture follows the structure of the data itself: neighboring frequency points must remain locally coherent, longer-range spectral dependencies must remain accessible, and electrochemical condition variables should remain explicit rather than being absorbed into an unconditional latent code.

3.4. Loss design

Training is guided by a composite objective rather than by reconstruction loss alone. Reconstruction preserves the primary spectral content of the input, but electrochemical impedance data also impose structural demands that a purely generic objective need not respect.

Band-based penalties discourage reconstructions that drift outside acceptable spectral behavior, including token-level deviation and magnitude-phase consistency. A physical positivity term suppresses outputs that violate basic electrochemical plausibility in the reconstructed components, while a total-variation term promotes local smoothness along the frequency axis. A latent regularization term further discourages unstable or unnecessarily diffuse latent representations. Rather than encoding a complete electrochemical theory into the network, this objective biases learning toward reconstructions and latent geometries that remain more compatible with smooth, bounded, and physically plausible spectra. Detailed loss-design ablation is reported in Appendix.

3.5. Latent summarization and descriptor construction

After training, each impedance sample is mapped to a latent vector and exported for downstream analysis. These sample-level latent codes are then reorganized by electrochemical condition so that descriptor construction is performed on condition-level structure rather than on isolated observations. In particular, samples sharing the same base potential V_b are grouped, and the dependence of each latent coordinate on perturbation potential V_p within that V_b slice is summarized parametrically.

In the reference analysis, this within-slice dependence is represented by quadratic fitting in V_p. The resulting coefficients are then concatenated to form a V_b-level latent summary, yielding one organized latent summary per base-potential condition. This summary rule was adopted as the reference parameterization because, among the tested summary modes, it provided the strongest overall grounded screening performance for the selected descriptor family (Appendix F).

Descriptor construction is carried out on these organized V_b-indexed summaries rather than on raw spectra or on single latent coordinates. Low-dimensional projections are used only as geometric views of the organized summaries

(Figure 3a), from which multiple trajectory-derived indices can be defined. These include angular coordinates, anchor-relative forms, radial or mixed forms, and local-shape quantities derived from variation along the latent trajectory. The purpose of this step is not to privilege one coordinate *a priori*, but to construct a physically plausible candidate family from the learned latent organization.

The resulting workflow therefore proceeds in five linked steps: sample-level encoding, condition-wise grouping, parametric summarization of V_p-dependence, trajectory-level candidate construction, and grounded candidate screening. This structure makes explicit what is otherwise easy to conflate: the latent code is learned from spectra, the summary is constructed from organized condition slices, and the final descriptor candidates are derived only after this intermediate reorganization step.

Candidate indices are then compared through internal consistency and external grounded validation against experimentally determined PZC. The aim is not to claim that the latent space contains one uniquely correct scalar by construction, but to organize and rank a family of plausible one-dimensional summaries under explicit scientific criteria.

4. Results and Discussion

4.1. Bias-ordered latent organization

The first result is that the learned latent space is not an arbitrary compression of the impedance data, but an organization in which electrochemical bias remains the dominant ordering principle. Across linear projection, nonlinear embedding, pairwise distance structure, and latent-to-V_b regression, the same pattern recurs: samples recorded under nearby bias conditions remain nearby after compression, and the global manifold follows a coherent progression with V_b (Figure 2). The significance of this result lies less in recovering a known experimental variable than in showing that compression has preserved a physically relevant ordering rather than obscuring it.

That point matters because descriptor construction in latent space is only scientifically useful when the learned organization remains stably linked to electrochemically meaningful variation. A latent map that reconstructs spectra accurately but destroys such ordering could still be numerically convenient, yet it would not provide a credible basis for downstream scalarization. Here, the latent-to-V_b regression reaches $R^2 = 0.989$, and the same bias-ordered structure is visible across multiple views of the representation. Taken together, these results indicate that the learned space has retained a coherent electrochemical backbone on which candidate summaries can be constructed. Against that structured backdrop, V_p-related variation appears as a secondary modulation rather than as the primary axis of latent structure.

4.2. Trajectory geometry supports descriptor construction

Once that ordered manifold is established, the next question is not simply whether the latent space is structured, but at what level that structure should be summarized for descriptor construction. The key observation is that the relevant organization is not captured by a single dominant latent coordinate. Instead, the V_b-indexed latent summaries trace an asymmetric trajectory around the global centroid (Figure 3a), indicating that the electrochemically relevant variation is expressed as coordinated movement through latent space rather than as monotonic change along one axis.

The supporting analyses reinforce that interpretation. The sensitivity ranking of the candidate indices (Figure 3b) is supported by a distributed set of latent-summary features rather than by one exceptional term, and the V_b-wise heatmap likewise shows coordinated drift, saturation, sign reversal, and weakly non-monotonic behavior across multiple features (Figure 3c). This pattern argues against reducing descriptor construction to one latent coordinate, one coefficient family, or one distance measure in isolation. A trajectory-level description is therefore not a cosmetic geometric choice, but the level at which the latent organization remains scientifically interpretable. In that sense, a useful scalarization should preserve coordinated latent reorganization rather than flatten it prematurely into a single naive index. This is why the subsequent comparison focuses on trajectory-derived indices that preserve orientation, anchor dependence, or local shape.

4.3. Grounded validation and interpretation

Grounded validation serves a stronger role here than simply selecting the coordinate with the highest numerical association. In an underdetermined latent space, many one-dimensional readouts can be constructed after compression, so numerical agreement alone does not establish scientific usefulness. The more important question is which form of scalarization preserves electrochemically ordered variation in a way that remains credible under an external check. In the present study, experimental PZC is used for that purpose only as a validation reference, not as a supervised training target.

Within the screened candidate family, the anchor-relative sine-like trajectory coordinate, implemented here as $\sin(\theta)$, emerges as the most credible descriptor candidate (Figure 4). Its advantage does not lie only in strong agreement with experimental PZC, but in the fact that it preserves directionally organized trajectory information relative to a physically motivated anchor while remaining stable across the focus re-

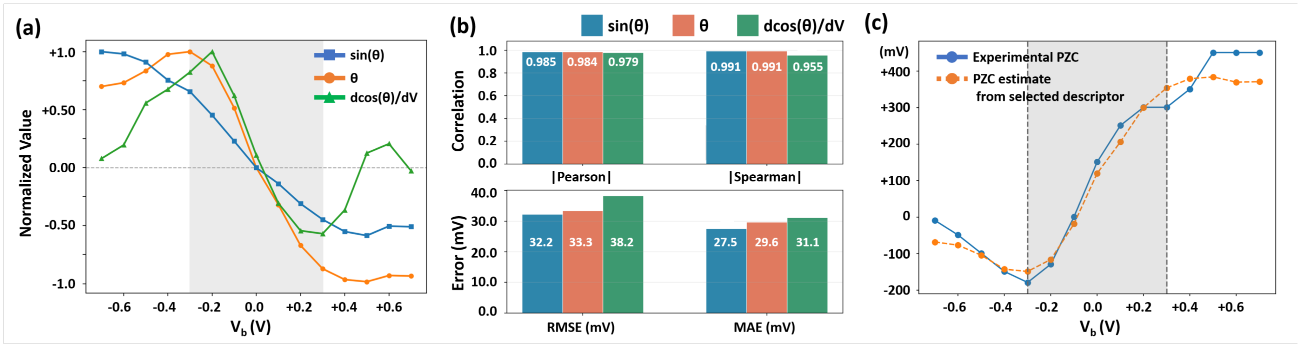


Figure 4. Comparison, selection, and grounded validation of trajectory-derived descriptor candidates. (a) Normalized V_b-dependent profiles of the top candidates, $\sin(\theta)$, θ, and $d\cos(\theta)/dV_b$, with the focus region shaded. (b) Quantitative comparison of these candidates within the focus region using absolute Pearson and Spearman correlation and reconstruction error (RMSE and MAE) against experimental PZC. (c) Experimental PZC and the PZC estimate derived from the selected candidate after one-dimensional calibration. The selected anchor-relative sine-like descriptor reproduces the major nonlinear features of the PZC–V_b relation, including the negative-bias minimum, zero crossing, and high-bias saturation.

gion. The top-screened alternatives, the unmodified angular coordinate θ and the local-shape coordinate $d\cos(\theta)/dV_b$, also retain electrochemical signal, indicating that relevant structure is not confined to a single construction. Screening within the focus region first narrows the candidate family on grounded quantitative criteria, after which one-dimensional calibration is used only to place the selected descriptor on the PZC scale for external comparison.

The comparison among these candidates sharpens the interpretation. The unmodified angular coordinate tracks the overall trajectory ordering and performs competitively, but the anchor-relative sine-like form provides a slightly better overall balance of correlation and calibrated reconstruction. The derivative-based coordinate $d\cos(\theta)/dV_b$ remains informative, especially as a measure of local variation along the latent path, yet its weaker grounded metrics indicate that local change alone is not the best-supported one-dimensional summary for the present validation target. The main contribution is therefore not a claim of universal optimality, but a grounded distinction between better-supported scalarizations and weaker alternatives within a structured latent candidate space. Latent organization becomes scientifically useful here not because it automatically yields a unique descriptor, but because it makes candidate construction, comparison, and external validation more disciplined and auditable. A further qualification is important for interpreting this result. Part of the dominant PZC-relevant structure is already linearly accessible, and a PCA baseline can remain competitive on top-level scalar agreement metrics (Appendix G). The value of the learned nonlinear representation should therefore not be framed as universal superiority in every summary statistic. Its practical advantage is that it supports a broader scientific workflow in which reconstruction, condition-aware latent organization, trajectory-based candidate construction, and grounded screening are all carried out within a single controllable representation. In that sense, the contribution of the present approach lies less in winning a benchmark-style comparison than in making descriptor discovery more structured, auditable, and scientifically navigable.

4.4. Limitations and scope

Three limits define the scope of the present claim. First, alignment of the latent space with V_b does not by itself demonstrate disentanglement of underlying interfacial physics. Because V_b is a known and injected condition, strong latent-to-V_b alignment shows coherent organization of the compressed representation rather than mechanistic separation in a strict causal sense. Second, PZC is used here only as a grounded validation reference, not as a supervised training endpoint and not as proof that the extracted descriptor is uniquely correct. Agreement with PZC therefore functions as an external plausibility check on descriptor quality rather than as a complete theory of the interface (Bard et al., 2022). Third, the proposed descriptor should not be interpreted as universal across chemistries, electrochemical systems, or acquisition protocols. The evidence presented here shows only that, within the present high-throughput impedance framework, a stable latent trajectory can be learned and a physically plausible scalar summary can be prioritized from that trajectory. In this sense, the significance of the nonlinear representation here lies primarily in providing a unified space for reconstruction, condition-aware organization, and descriptor screening under explicit experimental grounding.

5. Conclusion

This study shows that descriptor discovery from impedance-rich electrochemical measurements can be made more systematic when representation learning is used to organize observations into a structured latent space rather than to predict an endpoint directly. Within the present framework, the learned representation preserves electrochemically ordered variation, condition-level summarization converts that organization into a trajectory that can be inspected explicitly, and grounded comparison against independently measured PZC helps distinguish better-supported scalar summaries from weaker alternatives, prioritizing an anchor-relative sine-like trajectory coordinate as the most credible scalar summary within the tested candidate family.

The resulting contribution is methodological as much as empirical. AI is used here not to declare a final descriptor autonomously, but to make descriptor construction, candidate comparison, and external checking more disciplined in an underdetermined measurement setting. That role remains meaningful even when part of the relevant structure is linearly accessible, because the value of the learned representation lies in supporting a broader and more controllable scientific workflow rather than in maximizing a single benchmark metric.

More broadly, the present results suggest that, for measurement-rich but interpretation-limited problems, AI may be most useful when it operates as an organizing layer under explicit human control. A natural next step is to test how well the same workflow transfers across chemistries, acquisition protocols, and other underdetermined electrochemical measurement settings in which scientifically useful descriptors are not uniquely specified in advance.

References

Bai, S., Kolter, J. Z., and Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018. doi: 10.48550/arXiv.1803.01271.

Bard, A. J., Faulkner, L. R., and White, H. S. *Electrochemical Methods: Fundamentals and Applications*. Wiley, 3 edition, 2022. ISBN 9781119334064.

Bond, A. M., Zhang, J., Gundry, L., and Kennedy, G. F. Opportunities and challenges in applying machine learning to voltammetric mechanistic studies. *Current Opinion in Electrochemistry*, 34:101009, 2022. doi: 10.1016/j.coelec.2022.101009.

Chang, B.-Y. A novel analysis method for electrochemical impedance spectra using deep learning. *Electrochimica Acta*, 462:142741, 2023. doi: 10.1016/j.electacta.2023.142741.

Chang, B.-Y. and Park, S.-M. Electrochemical impedance spectroscopy. *Annual Review of Analytical Chemistry*, 3(1):207–229, 2010. doi: 10.1146/annurev.anchem.012809.102211.

Chen, H., Kätelhön, E., and Compton, R. G. Predicting voltammetry using physics-informed neural networks. *The Journal of Physical Chemistry Letters*, 13(2):536–543, 2022. doi: 10.1021/acs.jpclett.1c04054.

Chen, H., Kätelhön, E., and Compton, R. G. Machine learning in fundamental electrochemistry: Recent advances and future opportunities. *Current Opinion in Electrochemistry*, 38:101214, 2023. doi: 10.1016/j.coelec.2023.101214.

Doonyapisut, D., Kannan, P.-K., Kim, B., Kim, J. K., Lee, E., and Chung, C.-H. Analysis of electrochemical impedance data: Use of deep neural networks. *Advanced Intelligent Systems*, 5(8):2300085, 2023a. doi: 10.1002/aisy.202300085.

Doonyapisut, D., Kim, B., Kim, J. K., Lee, E., and Chung, C.-H. Deep generative learning for exploration in large electrochemical impedance dataset. *Engineering Applications of Artificial Intelligence*, 126:107027, 2023b. doi: 10.1016/j.engappai.2023.107027.

Fortin, P., Gerhardt, M. R., Ulleberg, Ø., Zenith, F., and Holm, T. Multi-sine EIS for early detection of PEMFC failure modes. *Frontiers in Energy Research*, 10:855985, 2022. doi: 10.3389/fenrg.2022.855985.

Hinton, G. E. and Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science*, 313 (5786):504–507, 2006. doi: 10.1126/science.1127647.

Hoar, B. B., Zhang, W., Xu, S., et al. Electrochemical mechanistic analysis from cyclic voltammograms based on deep learning. *ACS Measurement Science Au*, 2(6): 595–604, 2022. doi: 10.1021/acsmeasuresciau.2c00045.

Jones, P. K., Stimming, U., and Lee, A. A. Impedance-based forecasting of lithium-ion battery performance amid uneven usage. *Nature Communications*, 13(1):4806, 2022. doi: 10.1038/s41467-022-32422-w.

Jumper, J., Evans, R., Pritzel, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021. doi: 10.1038/s41586-021-03819-2.

Kennedy, G. F., Zhang, J., and Bond, A. M. Automatically identifying electrode reaction mechanisms using deep neural networks. *Analytical Chemistry*, 91(19):12220–12227, 2019. doi: 10.1021/acs.analchem.9b01891.

Lazanas, A. C. and Prodromidis, M. I. Electrochemical impedance spectroscopy—a tutorial. *ACS Measurement Science Au*, 3(3):162–193, 2023. doi: 10.1021/acsmeasuresciau.2c00070.

Lv, J., Yu, Z., Sun, G., and Liu, J. Deep learning-based fault diagnosis and electrochemical impedance spectroscopy frequency selection method for proton exchange membrane fuel cell. *Journal of Power Sources*, 591:233815, 2024. doi: 10.1016/j.jpowsour.2023.233815.

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., and Cubuk, E. D. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023. doi: 10.1038/s41586-023-06735-9.

Pajkossy, T. and Mészáros, G. Connection of CVs and impedance spectra of reversible redox systems, as used for the validation of a dynamic electrochemical impedance spectrum measurement system. *Journal of Solid State Electrochemistry*, 24:2883–2889, 2020. doi: 10.1007/s10008-020-04661-8.

Pence, M. A., Hazen, G., and Rodríguez-López, J. Closed-loop navigation of a kinetic zone diagram for redox-mediated electrocatalysis using bayesian optimization, a digital twin, and automated electrochemistry. *Analytical Chemistry*, 97(12):6771–6779, 2025. doi: 10.1021/acs.analchem.5c00099.

Perez, E., Strub, F., de Vries, H., Dumoulin, V., and Courville, A. FiLM: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018. doi: 10.1609/aaai.v32i1.11671.

Ragoisha, G. A. and Bondarenko, A. S. Potentiodynamic electrochemical impedance spectroscopy. *Electrochimica Acta*, 50(7–8):1553–1563, 2005. doi: 10.1016/j.electacta.2004.10.047.

Schaeffer, J., Gasper, P., Garcia-Tamayo, E., Gasper, R., Adachi, M., Gaviria-Cardona, J. P., Montoya-Bedoya, S., Bhutani, A., Schiek, A., Goodall, R., Findeisen, R., Braatz, R. D., and Engelke, S. Machine learning benchmarks for the classification of equivalent circuit models from electrochemical impedance spectra. *Journal of The Electrochemical Society*, 170(6):060512, 2023. doi: 10.1149/1945-7111/acd8fb.

Sheng, H., Sun, J., Rodríguez, O., Hoar, B. B., Zhang, W., Xiang, D., Tang, T., Hazra, A., Min, D. S., Doyle, A. G., Sigman, M. S., Costentin, C., Gu, Q., Rodríguez-López, J., and Liu, C. Autonomous closed-loop mechanistic investigation of molecular electrochemistry via automation. *Nature Communications*, 15(1):2781, 2024. doi: 10.1038/s41467-024-47210-x.

Sohn, K., Lee, H., and Yan, X. Learning structured output representation using deep conditional generative models. In *Advances in Neural Information Processing Systems 28 (NeurIPS 2015)*, 2015. doi: 10.5555/2969442.2969628.

Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., Milsted, D., McDermott, M. J., Gallant, M., Cubuk, E. D., Merchant, A., Kim, H., Jain, A., Bartel, C. J., Persson, K. A., Zeng, Y., and Ceder, G. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 624(7990):86–91, 2023. doi: 10.1038/s41586-023-06734-w.

Tom, G., Schmid, S. P., Baird, S. G., Cao, Y., Darvish, K., Hao, H., Lo, S., Pablo-García, S., Rajaonson, E. M., Skreta, M., Yoshikawa, N., Corapi, S., Akkoc, G. D., Strieth-Kalthoff, F., Seifrid, M., and Aspuru-Guzik, A. Self-driving laboratories for chemistry and materials science. *Chemical Reviews*, 124(16):9633–9732, 2024. doi: 10.1021/acs.chemrev.4c00055.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, 2017. doi: 10.5555/3295222.3295349.

Appendix

This appendix provides methodological details and supplementary analyses that support the main text while keeping the main narrative focused on descriptor discovery from latent organization. The appendix is organized around the same logic as the paper itself: external grounding, data preparation, representation learning, candidate construction, candidate screening, summary construction, and additional validation analyses.

A. Experimental PZC Determination and Reference Matching

A.1. Independent experimental determination of PZC

Experimental PZC values were used only as an external validation reference and were never supplied as supervised targets during model training. This distinction is central to the interpretation of the present workflow. The latent representation and the descriptor candidates were learned entirely from the impedance data and explicit experimental conditions, and PZC entered the analysis only after candidate construction as an independently measured reference.

In the present study, PZC was determined experimentally by sweeping both the base potential, V_b, and the perturbation potential, V_p, while acquiring high-throughput EIS measurements over the resulting condition grid. At each condition, the interfacial capacitance was estimated on the basis of the Gouy–Chapman–Stern (GCS) model, and the potential corresponding to the capacitance minimum was identified as the experimental PZC. This operational definition was applied consistently across the measured conditions and provided the grounded reference used for descriptor evaluation in the main text.

Because the experimental PZC values were obtained independently of the representation-learning procedure, the reported descriptor–PZC agreement should be interpreted as external validation rather than endpoint supervision. This separation is important for the scientific interpretation of the results, as it clarifies that the selected descriptors were screened against an experimentally defined electrochemical reference rather than optimized directly to reproduce it.

A.2. Matching the PZC reference to the impedance dataset

The experimental PZC table and the impedance-derived descriptor table were aligned through their shared base-potential values V_b after experimental PZC extraction. Because the PZC reference was available only at the condition level used for descriptor comparison, grounded evaluation was performed on matched V_b points rather than on the full impedance dataset at the individual-sample level.

The matching procedure consisted of four steps: (1) extracting the experimental PZC reference from the independently analyzed electrochemical data, (2) constructing the descriptor table from the learned latent summaries, (3) aligning the two tables over shared V_b values, and (4) separating the matched points into the full region and the prespecified focus region used for primary screening. In the present dataset, this yielded 15 matched points in the full region and 7 matched points in the focus region.

This explicit matching step is scientifically important because the reported descriptor–PZC agreement reflects sparse external grounding rather than dense supervision. The validation results should therefore be interpreted as agreement on independently matched reference conditions, not as endpoint fitting over the entire dataset.

B. Data Preprocessing and Dataset Split

B.1. Input construction and normalization

Each learning sample consists of a frequency-ordered impedance sequence with 17 frequency points and two response channels, Z_r and Z_i, accompanied by explicit conditioning variables including V_b and V_p. The full dataset contains 6840 samples spanning 15 base-potential values and 59 perturbation-potential values under a fixed concentration condition.

Inputs were normalized using median centering and interquartile-range scaling for optimization stability, whereas reconstruction quality, pass metrics, and downstream descriptor screening were interpreted with reference to both normalized and physical space. This separation was retained intentionally so that latent-space quality would not be judged only by abstract compression performance, but also by whether reconstructed spectra and downstream summaries remained meaningful after inversion to the original electrochemical scale.

Table A1. Dataset composition, preprocessing, split design, and reference matching used in the main analysis.

Section	Item	Value
Dataset	Total impedance samples	6840
Dataset	Base-potential values V_b	15
Dataset	Perturbation-potential values V_p	59
Dataset	Frequency points per spectrum	17
Dataset	Concentration condition	Fixed single concentration
Input	Spectral channels	Z_r, Z_i
Input	Explicit condition variables	V_b, V_p
Preprocessing	Input normalization	Median centering + interquartile-range scaling
Split	Train / validation / test	0.80 / 0.15 / 0.05
Split	Grouping key	Shared (V_b, V_p) condition sets
Split	Group shuffling	True
Evaluation	Domains	Normalized and physical space
Reference	Matched PZC points (all region)	15
Reference	Matched PZC points (focus region)	7
Screening	Focus region	$-0.3 \leq V_b \leq 0.3$

B.2. Grouped split and leakage prevention

Train, validation, and test partitions were constructed using grouped splitting rather than random sample-level partitioning. The split ratio in the reference experiments was 0.80/0.15/0.05. Grouping was defined at the level of shared (V_b, V_p) condition sets so that samples belonging to the same electrochemical condition context were kept in the same partition.

This design was adopted because neighboring measurements within a high-throughput impedance grid can remain strongly correlated even when treated as distinct samples. Preventing such leakage is necessary if latent organization is to be interpreted scientifically rather than as an artifact of train–test overlap. Group shuffling was applied before partitioning, and all downstream analyses reported in the main text were based on models trained under this grouped-split protocol.

C. Reference Architecture and Training Hyperparameters

C.1. Reference model configuration

The reference model used throughout the main text is a conditional autoencoder that combines a temporal-convolution encoder with a hybrid decoder containing both attention-based and temporal-convolution components. Experimental conditions are injected explicitly rather than inferred indirectly, and frequency information is supplied through a dedicated embedding pathway. This design reflects the structure of the input data: neighboring frequency points should remain locally coherent, longer-range spectral dependencies should remain accessible, and condition variables should remain explicit rather than being absorbed into an unconditional latent code.

The reference configuration used in the main analysis corresponds to a latent dimension of 24, a model width of 512, an encoder depth of 5 blocks, a decoder temporal-convolution depth of 3 blocks, 2 decoder attention layers, 8 attention heads, a feed-forward multiplier of 2, dropout of 0.4, and condition dropout of 0.2. Optimization was performed with a learning rate of 5×10^{-4}, a batch size of 64, and early stopping with patience 50. The complete reference configuration is summarized in Table A2.

C.2. Architecture and hyperparameter ablation

The architecture ablation is best interpreted as a robustness analysis rather than as a blind search for a numerically optimal network. The sweep covers latent dimension, model width, encoder depth, decoder depth, number of attention layers, number of attention heads, feed-forward multiplier, dropout, and condition-vector dropout. Taken together, these experiments show that the main conclusions do not depend on one arbitrarily chosen architecture.

The sweep is therefore best interpreted as a robustness study around the main analysis configuration rather than as a search for a single universally optimal architecture. Several variants remain competitive, and some improve particular downstream

Table A2. Reference architecture and optimization hyperparameters used in the main text.

Group	Parameter	Value
Latent	z_{dim}	24
Latent	d_{model}	512
Encoder	`enc_blocks`	5
Encoder	`enc_dilations`	[1, 2, 4, 8, 16]
Decoder	`dec_tcn_blocks`	3
Decoder	`dec_dilations`	[1, 2, 4, 8, 16]
Decoder	`dec_mhsa_layers`	2
Decoder	`attn_heads`	8
Decoder	`ff_mult`	2
Regularization	`dropout`	0.4
Regularization	`condition_dropout`	0.2
Optimization	learning rate	0.0005
Optimization	batch size	64
Optimization	epochs	200
Optimization	early-stopping patience	50
Loss	$\lambda_{\text{band,start}}$	1.0
Loss	$\lambda_{\text{band,end}}$	1.0
Loss	$\lambda_{\text{magphase}}$	1.0
Loss	λ_{phys}	0.001
Loss	λ_{tv}	0.001
Loss	$\lambda_{z,\text{L2}}$	0.0001

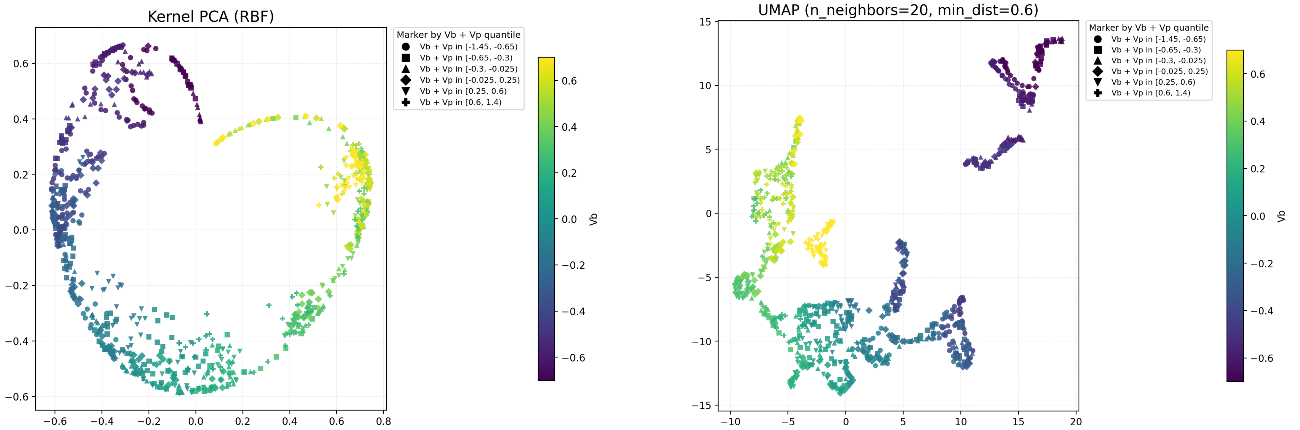
metrics. For that reason, the reference model used in the main text should be interpreted as a stable and interpretable analysis configuration rather than as the unique numerical optimum. Its role is to provide a consistent basis for reconstruction, latent organization, and descriptor screening across the full workflow.

Table A3. Architecture and hyperparameter ablation results relative to the reference model. Only ablated architectural settings and downstream metrics are shown. Loss weights were fixed at the reference values. Metrics are rounded to two decimal places for readability.

Model	Variant	z_{dim}	d_{model}	Enc.	Dec.-TCN	MHSA	Heads	FF	Dropout / Cond.	Both-pass	Pearson / RMSE
M0	Reference	24	512	5	3	2	8	2	0.40 / 0.20	0.84	0.98 / 36.40
M1	Latent dim = 8	8	512	5	3	2	8	2	0.40 / 0.20	0.84	0.97 / 47.68
M2	Latent dim = 16	16	512	5	3	2	8	2	0.40 / 0.20	0.83	0.97 / 43.93
M3	Latent dim = 32	32	512	5	3	2	8	2	0.40 / 0.20	0.84	0.98 / 38.84
M4	Latent dim = 48	48	512	5	3	2	8	2	0.40 / 0.20	0.83	0.97 / 43.72
M5	Model width = 256	24	256	5	3	2	8	2	0.40 / 0.20	0.84	0.96 / 51.62
M6	Model width = 384	24	384	5	3	2	8	2	0.40 / 0.20	0.84	0.97 / 42.19
M7	Encoder depth = 3	24	512	3	3	2	8	2	0.40 / 0.20	0.83	0.97 / 45.03
M8	Encoder depth = 7	24	512	7	3	2	8	2	0.40 / 0.20	0.83	0.99 / 32.28
M9	Decoder TCN depth = 1	24	512	5	1	2	8	2	0.40 / 0.20	0.83	0.97 / 44.46
M10	Decoder TCN depth = 5	24	512	5	5	2	8	2	0.40 / 0.20	0.83	0.97 / 47.23
M11	Attention layers = 1	24	512	5	3	1	8	2	0.40 / 0.20	0.83	0.97 / 46.23
M12	Attention layers = 3	24	512	5	3	3	8	2	0.40 / 0.20	0.84	0.98 / 39.05
M13	Heads = 4	24	512	5	3	2	4	2	0.40 / 0.20	0.83	0.98 / 38.97
M14	Heads = 16	24	512	5	3	2	16	2	0.40 / 0.20	0.84	0.97 / 44.49
M15	FF multiplier = 1	24	512	5	3	2	8	1	0.40 / 0.20	0.84	0.97 / 45.14
M16	FF multiplier = 4	24	512	5	3	2	8	4	0.40 / 0.20	0.83	0.98 / 38.18
M17	Dropout = 0.2	24	512	5	3	2	8	2	0.20 / 0.20	0.85	0.97 / 43.27
M18	Dropout = 0.6	24	512	5	3	2	8	2	0.60 / 0.20	0.80	0.97 / 47.27
M19	Cond. dropout = 0.0	24	512	5	3	2	8	2	0.40 / 0.00	0.84	0.98 / 37.45
M20	Cond. dropout = 0.4	24	512	5	3	2	8	2	0.40 / 0.40	0.84	0.96 / 50.93

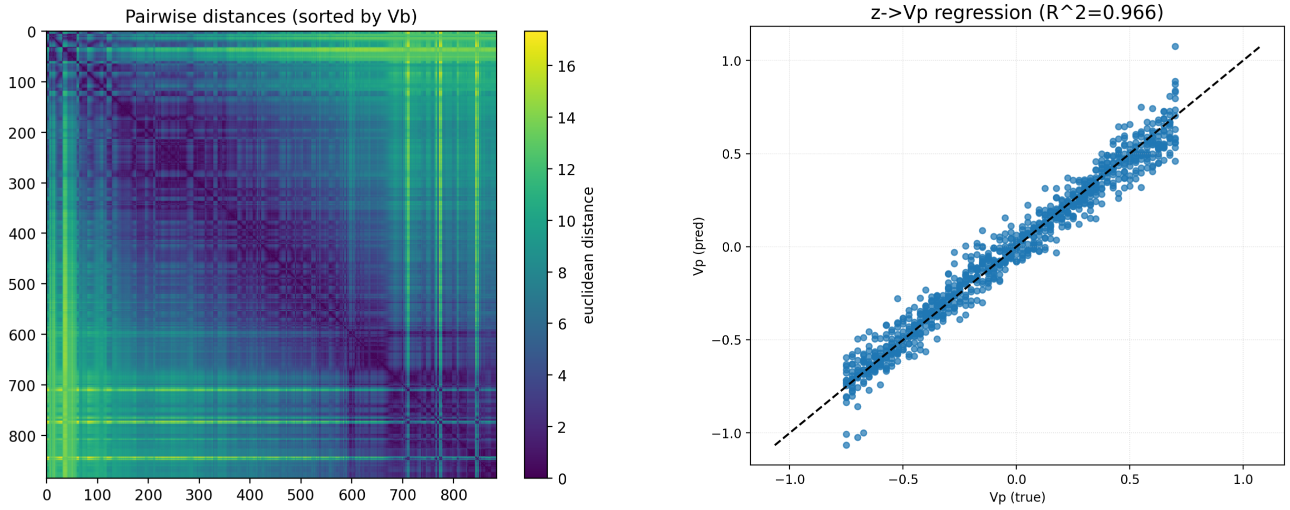
D. Loss-Design Ablation

Unless otherwise noted, descriptor selection in the main text is based on the abc-only summary mode. The auxiliary loss-ablation comparison shown here is retained in the same setting as the plotted panel so that the relative effect of each loss component can be read directly from the displayed results.



(a) Kernel PCA view of the reference latent space.

(b) UMAP view of the reference latent space.



(c) Pairwise latent-distance matrix after sorting by V_b.

(d) Regression from latent code to V_p.

Figure A1. Additional latent-space organization views for the reference model. The latent geometry is shown using kernel PCA and UMAP, with points colored by V_b and marker styles indicating quantile bins of $V_b + V_p$. A regression scatter from the latent code to V_p is included to show that the learned representation retains substantial information about the perturbation-potential axis. Together with the pairwise-distance view, these panels provide a complementary summary of how experimental conditions are organized in the learned latent space.

The loss-design ablation addresses a distinct question from the architecture sweep. The relevant issue is not only whether the model reconstructs spectra, but whether the learned latent representation remains useful for grounded descriptor construction after training. In the present results, reconstruction-only training provides the weakest baseline because it lacks explicit constraints that encourage physically plausible spectral structure and stable latent organization. Adding spectral-band constraints already recovers much of the downstream improvement, indicating that physically informed structure is important for descriptor discovery rather than merely for cosmetic reconstruction quality.

The full composite objective provides additional stabilization through magnitude–phase consistency, physical plausibility penalties, total-variation smoothing, and latent regularization. More broadly, these results support the view that, in underdetermined descriptor-discovery settings, physically informed losses are valuable not because they encode a complete electrochemical theory into the network, but because they bias the learned representation toward reconstructions and latent geometry that remain scientifically useful downstream. Figure A2 summarizes the corresponding ablation results.

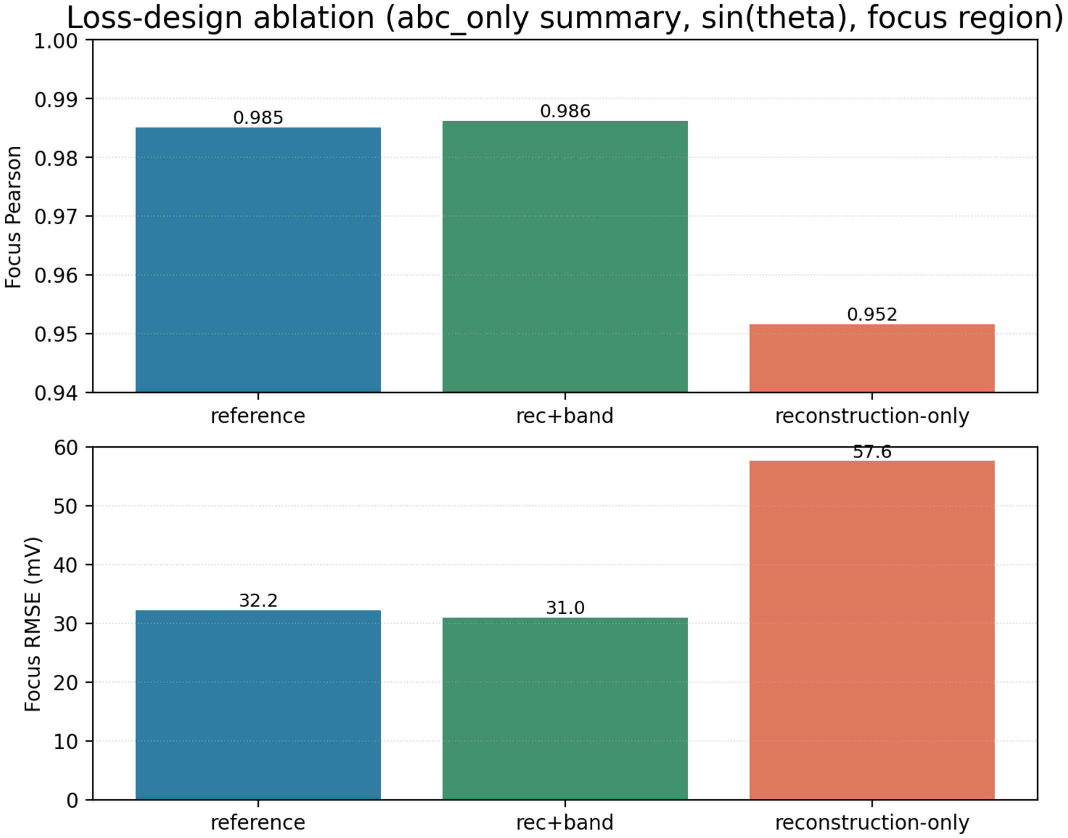


Figure A2. Loss-design ablation in the focus region. The reference objective, the reconstruction-plus-band variant, and the reconstruction-only baseline are compared using the focus-region Pearson correlation and RMSE for $\sin(\theta)$. The results show that physically informed auxiliary terms improve downstream descriptor utility beyond reconstruction loss alone.

E. Full Candidate Pool, Screening Policy, and Rejected Alternatives

Candidate notation. Here, r denotes radial distance from the chosen reference point, θ denotes the oriented angular coordinate, and arc length denotes the cumulative trajectory coordinate along the organized V_b-indexed latent path. Derivative and mixed forms are defined from these same trajectory coordinates.

The main text reports $\sin(\theta)$, θ, and $\frac{d\cos(\theta)}{dV_b}$ as the top-screened candidates, but these descriptors were selected from a broader family rather than chosen *ad hoc*. The full candidate pool includes angular, radial, anchor-relative, and local-derivative forms defined on the organized latent trajectory. Reporting this broader pool is important because it makes the selection logic transparent and avoids the appearance of post hoc cherry-picking.

Table A4. Candidate-screening results for the abc-only summary mode. Rows are grouped by candidate family for readability. Metrics are rounded to two decimal places.

Candidate	**Family**	**Focus n**	**All n**	**Focus Pearson / RMSE**	**All Pearson / RMSE**
$\sin(\theta)$	Main-text	7	15	0.99 / 32.25	0.99 / 44.61
θ	Main-text	7	15	0.98 / 33.33	0.92 / 140.76
$d\cos(\theta)/dV_b$	Main-text	7	15	0.98 / 38.21	0.73 / 164.66
$r\sin(\theta)$	Radial / proxy	7	15	0.96 / 55.26	0.88 / 208.35
arc length	Radial / proxy	7	15	0.96 / 52.60	0.82 / 318.19
r	Radial / proxy	7	15	0.37 / 174.60	0.07 / 361.19
$r\theta$	Radial / proxy	7	15	0.95 / 60.79	0.70 / 553.66
$d^2\theta/dV_b^2$	Angular	7	15	0.86 / 96.14	0.57 / 215.29
$d\theta/dV_b$	Angular	7	15	0.54 / 157.40	0.54 / 279.26
$\cos(\theta)$	Angular	7	15	0.58 / 152.54	0.44 / 287.01
$d(r\cos\theta)/dV_b$	Derivative / mixed	7	15	0.86 / 94.21	0.65 / 353.74
dr/dV_b	Derivative / mixed	7	15	0.97 / 45.21	0.66 / 403.73
$d(r\sin\theta)/dV_b$	Derivative / mixed	7	15	0.76 / 122.78	-0.46 / 476.03
ds/dV_b	Derivative / mixed	7	15	0.46 / 167.02	-0.57 / 656.24
$r\theta'$	Other	7	15	0.13 / 185.99	0.66 / 181.48
$d(r\theta)/dV_b$	Other	7	15	0.03 / 187.46	0.58 / 216.15
$d\sin(\theta)/dV_b$	Other	7	15	0.26 / 181.01	0.06 / 256.68
$r^2\,d\theta/dV_b$	Other	7	15	0.26 / 180.90	-0.59 / 341.84
$r\cos(\theta)$	Other	7	15	0.90 / 83.22	0.53 / 717.23
$d(r^2\theta)/dV_b$	Other	7	15	0.26 / 181.22	-0.41 / 806.54
$r^2\theta$	Other	7	15	0.90 / 83.22	0.48 / 1803.00

Candidate screening was carried out using grounded quantitative criteria, including Pearson correlation, Spearman correlation, RMSE, and MAE relative to experimental PZC after one-dimensional calibration. The focus region was used as the primary screening regime, while all-region behavior was also recorded to identify candidates that appeared strong locally but generalized poorly across the broader bias range. In the current results, $\sin(\theta)$ provides the strongest overall balance, θ remains competitive but less stable outside the focus region, and $\frac{d\cos(\theta)}{dV_b}$ emerges as a credible but weaker local-shape alternative. By contrast, purely radial or strongly proxy-like candidates are less convincing as general descriptor choices despite occasional partial alignment.

This broader comparison strengthens the interpretation of the selected descriptor. The preferred candidate is not simply the best among a few handpicked curves; it is the most credible member of a more diverse screened family. Figures A3–A5 and Table A4 present the broader screening landscape, the appendix-only candidate families, and representative rejected alternatives.

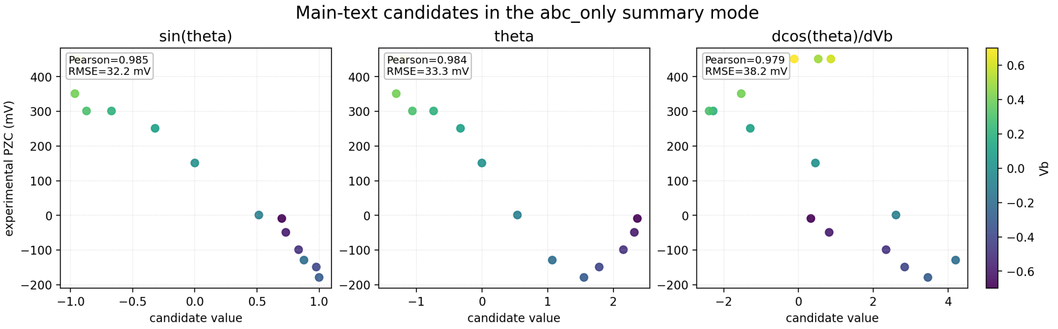


Figure A3. Candidate-pool screening in the `abc-only` summary mode: combined scatter view for the three main-text candidates against experimental PZC.

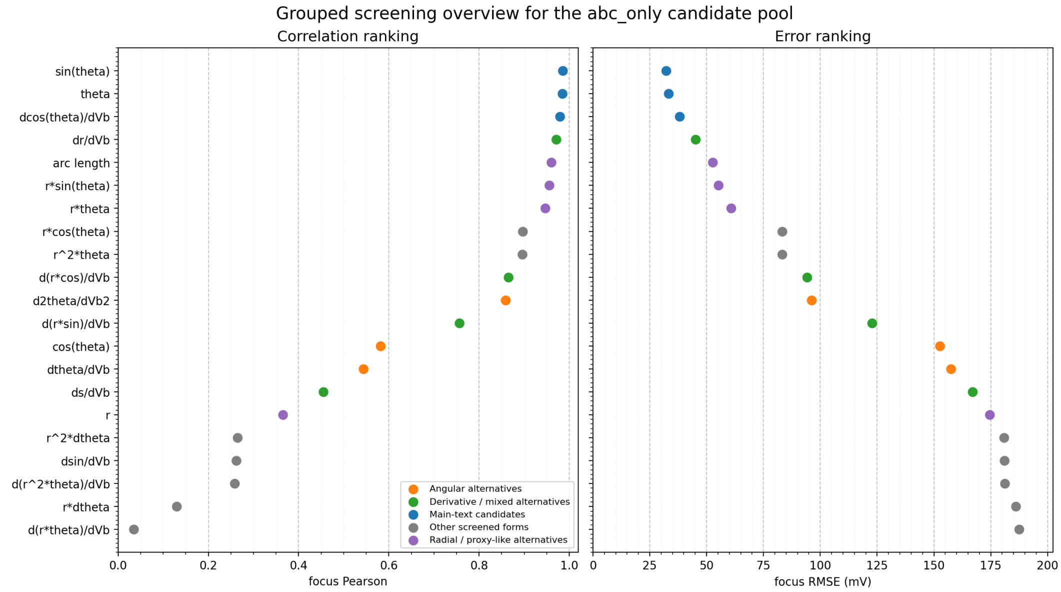


Figure A4. Candidate-pool screening in the `abc-only` summary mode: grouped screening overview with candidate-family labels.

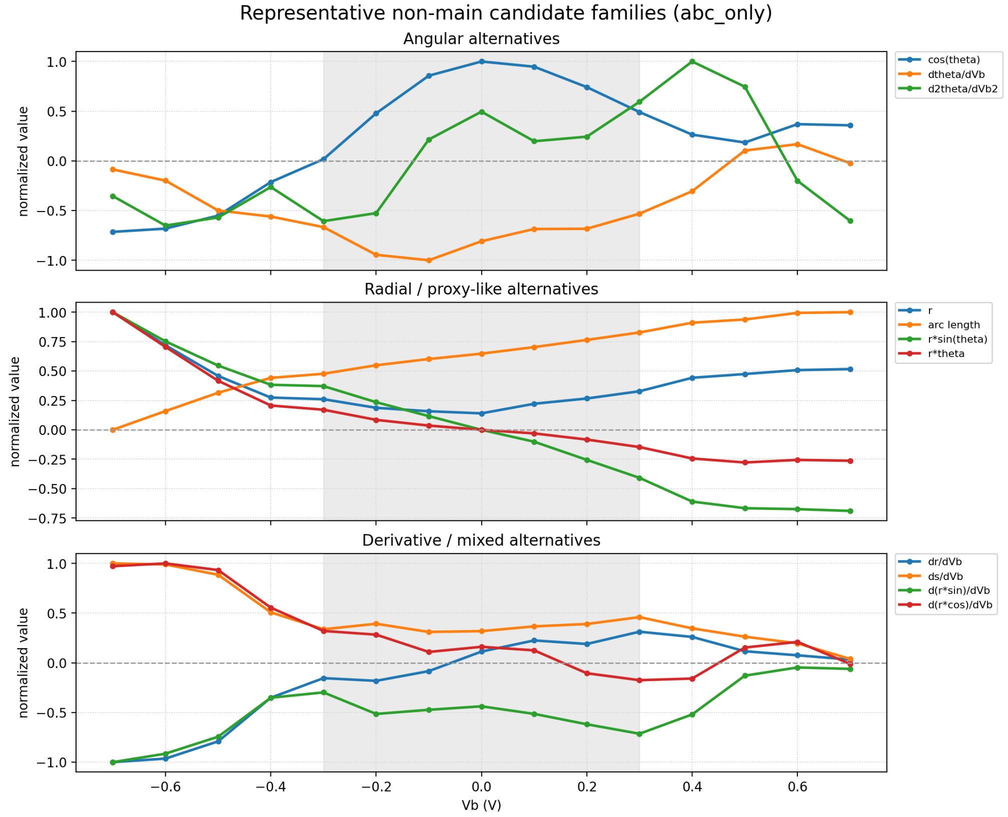


Figure A5. Representative non-main candidate families in the `abc-only` summary mode. The plotted profiles show angular alternatives, radial/proxy-like alternatives, and derivative/mixed alternatives across V_b.

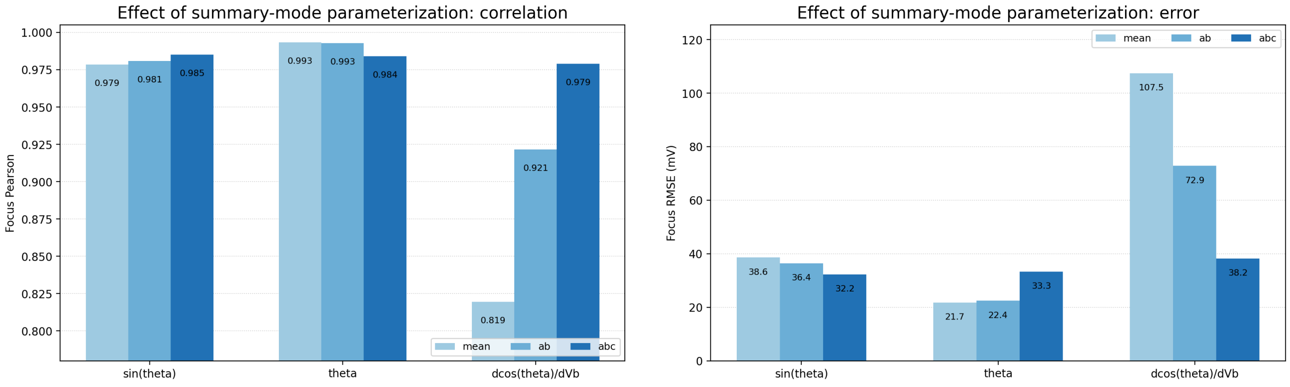
Table A5. Effect of summary parameterization on representative candidates in the focus region.

Summary mode	Candidate	Focus Pearson	Focus RMSE (mV)
mean-only	$\sin(\theta)$	0.98	38.59
mean-only	θ	0.99	21.72
mean-only	$d\cos(\theta)/dV_b$	0.82	107.50
ab-only	$\sin(\theta)$	0.98	36.40
ab-only	θ	0.99	22.42
ab-only	$d\cos(\theta)/dV_b$	0.92	72.85
abc-only	$\sin(\theta)$	0.99	32.25
abc-only	θ	0.98	33.33
abc-only	$d\cos(\theta)/dV_b$	0.98	38.21

F. Effect of Latent-Summary Parameterization and Choice of Polynomial Order

The main text adopts the `abc-only` summary mode, and that choice should be justified directly. The three summary modes correspond to different ways of compressing the V_p dependence within each V_b slice before descriptor screening: `mean-only` uses direct averaging, `ab-only` uses a first-order summary, and `abc-only` uses a quadratic summary. The purpose of this comparison is therefore not to claim invariance across all summary constructions, but to explain why the quadratic parameterization was chosen for the main analysis.

The comparison shows that `abc-only` provides the strongest overall alignment for the selected descriptor $\sin(\theta)$. In the focus region, the calibrated Pearson correlation for $\sin(\theta)$ improves from 0.979 in `mean-only` to 0.981 in `ab-only` and 0.985 in `abc-only`, while the calibrated RMSE decreases from 38.6 mV to 36.4 mV and then to 32.2 mV. This monotonic improvement provides a direct justification for using the quadratic summary in the main text. At the same time, θ remains competitive in all three settings, and $\frac{d\cos(\theta)}{dV_b}$ becomes substantially more informative under `abc-only`, indicating that the richer parameterization exposes additional candidate structure that is largely absent under simpler summaries. Figure A6 and Table A5 summarize this comparison.



(a) Correlation comparison across summary modes.

(b) Error comparison across summary modes.

Figure A6. Effect of latent-summary parameterization on candidate screening. The grouped bar charts compare representative candidates under the `mean-only`, `ab-only`, and `abc-only` summary rules. The results provide a direct rationale for using the quadratic summary in the main text.

G. Baseline PCA Comparison

A linear baseline is useful not because the paper is framed as a benchmark contest, but because it clarifies what part of the relevant structure is already linearly accessible. In the present results, the PCA baseline recovers part of the dominant organization and can achieve competitive alignment for some top descriptors. This observation should be stated plainly, as it indicates that the strongest PZC-relevant structure contains a substantial linear component.

At the same time, the learned representation remains valuable because it supports a richer workflow that includes spectrum reconstruction, condition-aware latent organization, geometry-based candidate construction, and more flexible screening beyond a fixed linear projection. The most accurate interpretation is therefore not that the nonlinear model dominates every scalar metric, but that it supports a more structured and controllable scientific workflow. Figure A7 and Table A6 document this comparison.

Table A6. Quantitative comparison between the reference autoencoder and the PCA baseline.

Method	**Latent type**	**Candidate**	**Region**	**Pearson**	**Spearman**	**RMSE (mV)**
AE reference	conditional autoencoder	$\sin(\theta)$	all	0.98	0.98	47.20
AE reference	conditional autoencoder	$\sin(\theta)$	focus	0.98	0.95	36.40
PCA baseline	linear PCA	$\sin(\theta)$	all	0.98	0.99	44.22
PCA baseline	linear PCA	$\sin(\theta)$	focus	0.98	0.99	35.26

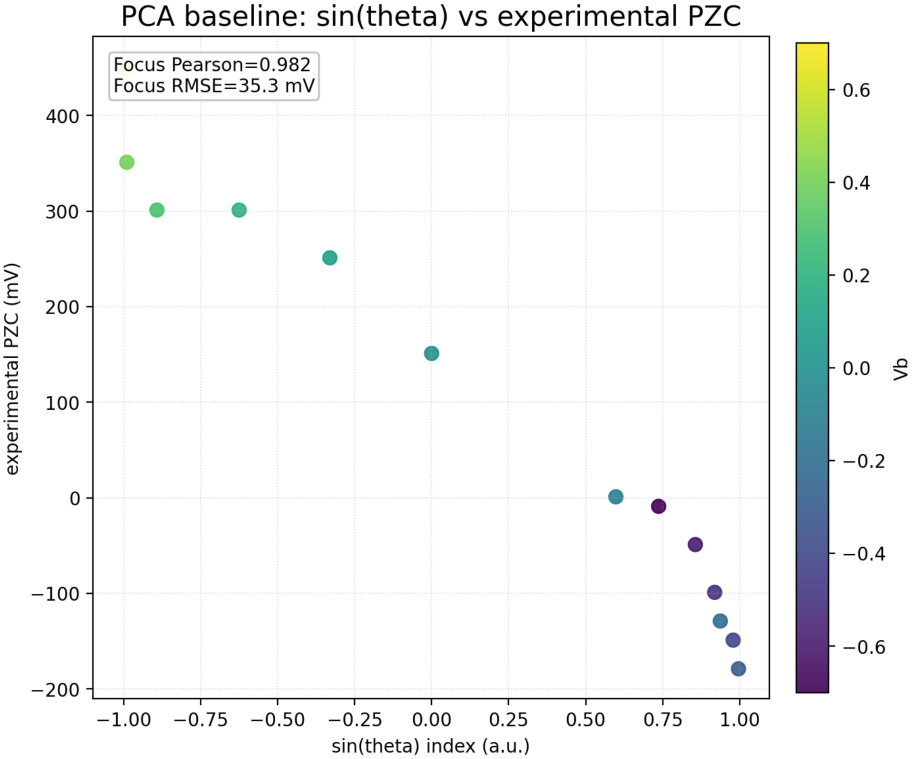


Figure A7. Representative descriptor–PZC comparison for the PCA baseline. This panel shows that part of the dominant PZC-relevant organization is already linearly accessible. At the same time, the learned nonlinear representation remains valuable because it supports a broader workflow for reconstruction, condition-aware organization, and geometry-based descriptor construction.