Screening for Metallic Character in a Database of Novel Magnetic Metal-Organic Frameworks

Alexander C. Tyner,[1] Avinash Pathapati,[2] and Alexander V. Balatsky[1]

[1] *NORDITA, KTH Royal Institute of Technology and Stockholm University, Stockholm, Sweden*
[2] *Doublet labs, Stockholm, Sweden*

(Dated: August 3, 2026)

Realizing a metallic, magnetically ordered metal-organic framework (MOF) would combine itinerant carriers with rare-earth-free magnetism, a combination of interest for organic spintronics. A companion study generated a database of 149,420 candidate magnetic MOFs, M^2OF, by exhaustive site substitution of magnetic-ion positions in structural prototypes drawn from the QMOF database, ranking candidates purely by predicted total magnetic moment. That screen leaves the electronic character of the resulting structures, metallic or gapped, unexamined. Here we screen the full M^2OF database for metallic band structure using a SchNet regressor trained on 12,500 single-point DFT band gaps from the Organic Materials Database (OMDB-GAP1), applying a 0.1 eV predicted-gap cutoff to flag likely metals. Of 149,419 structures successfully screened, 8,800 (5.89%) fall below this threshold. The metallic fraction is not uniform across the substituted transition-metal site: Fe-substituted frameworks are metallic 9.6% of the time, versus a 5.4–6.3% range for Mn, V, Co, and Cr. Restricting to the 590 candidates with the largest total magnetic moment (mag. moment $\geq 60\,\mu_B$) identified for DFT follow-up in the companion study, none are predicted metallic, an apparent tension between the two target properties in this screen. As the band-gap checkpoint was trained on bulk inorganic crystals and applied out-of-distribution to porous organic frameworks, with a validation MAE (0.356 eV) larger than the metallic threshold itself, these results are intended as a prioritization tool for downstream DFT validation rather than a validated metal count.

I. INTRODUCTION

Metal-organic frameworks (MOFs) that combine long-range magnetic order with itinerant, metallic carriers would offer a synthetic route to rare-earth-free spintronic and molecular-conductor materials [1–3]. Identifying such compounds by first-principles search alone is intractable at scale: MOF unit cells routinely exceed 100 atoms, and density functional theory (DFT) scales as $O(N^3)$ in atom number, so exhaustively computing electronic structure for every candidate in a large generated database is not practical.

A companion study [4] addressed the magnetic half of this problem. Starting from magnetic MOF prototypes in the QMOF database [5, 6], that work fine-tuned CHGNet on a new OMDB-derived training set and used it to relax and rank 149,420 structures generated by substituting inequivalent magnetic sites (V, Cr, Mn, Fe, Co, Ni) in each prototype, producing the M^2OF database. Ranking there was driven entirely by predicted total magnetic moment; the electronic character of the generated structures, whether a given framework is a gapped insulator or has closed, metallic bands, was not part of that screen and is a materially different property from the magnetic moment used to rank candidates.

This work closes that gap. We apply an independently trained SchNet band-gap regressor, built on the OMDB-GAP1 band-gap dataset [7–9], directly to all 149,420 M^2OF structures and flag predicted-metallic candidates by a fixed band-gap threshold. The result is a ranked, searchable overlay on top of the existing magnetic-moment ranking, letting future DFT validation (of the kind performed in Ref. [4] via phonon stability and Quantum Espresso calculations) be prioritized toward candidates that are simultaneously magnetic *and* predicted metallic, rather than magnetic alone. This screening strategy follows the same OMDB-data-mining lineage as prior property-specific searches of the database, including a recent screen for superconducting MOFs [10].

II. METHODOLOGY

A. Band-gap regression model

The regressor is a SchNet architecture [11], implemented in SchNetPack [12], with 3 interaction blocks, an atomic embedding size of 64, a 5 Å interaction cutoff with cosine cutoff function, and a 20-basis Gaussian radial expansion, matching Table 3 of the SchNetPack reference [12]. Atomistic input graphs are constructed with the Atomic Simulation Environment [13].

The model is trained from scratch on 12,500 single-point DFT band gaps from OMDB-GAP1: 9,000 structures for training and 1,000 for validation (drawn at random, seed 42, from the first 10,000 database entries) plus the dataset's official 2,500 structure test split. Training uses Adam with an initial learning rate of 1×10^{-3}, decayed by a factor of 0.96 every 10,000 steps, batch size 32, and an MSE loss with early stopping on validation MAE (patience 30 epochs, maximum 500 epochs). The best checkpoint, selected at epoch 37, reaches a validation MAE of 0.345 eV and a test-set MAE of 0.356 eV (Table I).

We emphasize that OMDB-GAP1 consists of bulk, relatively dense inorganic and organic crystal structures,

TABLE I. Training summary for the SchNet band-gap regressor.

Quantity	Value
Train / val / test structures	9,000 / 1,000 / 2,500
Best epoch (early stopping)	37
Validation MAE	0.345 eV
Test MAE	0.356 eV
Metallic threshold (this work)	0.1 eV

not porous MOF host lattices. Applying this checkpoint to the M^2OF database is therefore an out-of-distribution (OOD) prediction task, and the 0.356 eV test MAE is more than three times the 0.1 eV metallic cutoff used to label candidates; both points are treated as first-order caveats throughout Sec. IV.

B. Screening pipeline

Candidate metadata (framework identifier, chemical formula, and predicted total magnetic moment from Ref. [4]) is queried from a relational database; the corresponding relaxed structure is retrieved on demand as a compressed CIF file from object storage, decompressed, and parsed with ASE [13]. Structures are batched (16–64 per batch) and passed through the trained SchNet model on a single GPU; the predicted gap is clipped to non-negative values before thresholding, since gaps are physically bounded below by zero but the regression can produce small negative artifacts. A structure is labeled *metal* when its clipped predicted gap falls below 0.1 eV, otherwise it is left unlabeled (non-metal). The full pipeline is summarized in Fig. 1.

Structure retrieval, not model inference, is the throughput bottleneck: each prediction requires one small network request to fetch a compressed CIF file. We therefore overlap retrieval across many concurrent worker threads while inference proceeds on whichever batch of structures becomes available first. An initial attempt to screen the full database registered all ∼150,000 retrieval requests with the underlying futures executor up front; the completion-tracking overhead of that approach grows with the number of *outstanding* requests rather than staying constant, and the observed sustained rate degraded from a network-bound ∼11 structures/s to well under 1 structure/s over several hours. Replacing this with a bounded sliding window, in which only a small, fixed number of retrieval requests (worker count × 4) are ever in flight at once, restored throughput independent of total corpus size. With this fix, screening the full M^2OF database of 149,420 structures completed in 7 h 11 min on a single GH200 GPU node (average 5.8 structures/s), with a single structure skipped due to a retrieval failure.

III. RESULTS

Figure 2 shows the predicted band-gap distribution across the full corpus: mean 1.96 eV, standard deviation 1.01 eV, ranging from 0 to 5.10 eV after clipping. Of the 149,419 structures successfully screened, 8,800 (5.89%) fall below the 0.1 eV metallic threshold.

The metallic fraction is not uniform across the six substituted magnetic elements (Fig. 3). Fe-containing structures are predicted metallic in 9.6% of cases (7,030/73,041), a factor of ∼1.6 above the corpus mean, while Mn-containing structures are the least metallic-enriched at 5.4% (3,882/71,543). Cr, V, Co, and Ni fall between these extremes (5.7–7.5%).

Restricting to the 590 candidates with mag. moment $\geq 60\,\mu_B$, the subset identified in the companion study as the highest-priority targets for dynamical stability and DFT validation, none are predicted metallic (Fig. 4). The corresponding band-gap distribution for this subset, shown for reference in Fig. 5, is shifted toward larger gaps than the full corpus.

IV. DISCUSSION

Two results merit caution before being read as physical trends rather than screening artifacts. First, the apparent Fe enrichment in Fig. 3, while a factor of ∼1.6 above the corpus mean, is based on a model applied out-of-distribution: the training checkpoint never saw a MOF host lattice, only bulk inorganic and dense organic crystals from OMDB-GAP1, and its 0.356 eV test MAE is substantially larger than the 0.1 eV cutoff separating a metal label from a non-metal label in this work. A shift of even one MAE in the predicted gap for a structure near the threshold is enough to flip its label. The 5.89% corpus-wide metallic fraction, and the elemental breakdown built on it, should therefore be read as an order-of-magnitude prioritization signal, not a validated count of metallic MOFs.

Second, the absence of any predicted metal among the 590 highest-magnetic-moment candidates (Fig. 4) is consistent with, but does not establish, a genuine tension between maximizing total magnetic moment and retaining a closed band gap in this generation scheme: fully spin-polarized, high-moment configurations from aggressive magnetic-site substitution may systematically favor wider gaps in this model's OOD regime, or the effect may simply reflect the small size of that subset (590 of 149,420 structures). Distinguishing between these explanations requires first-principles verification of the kind already applied to the top magnetic-moment candidates in Ref. [4], extended to compute the actual electronic band structure rather than relying on the trained SchNet gap alone.

Given both caveats, we suggest the practical use of this screen is as a *prioritization filter* layered on top of the existing magnetic-moment ranking: of the 8,800 structures

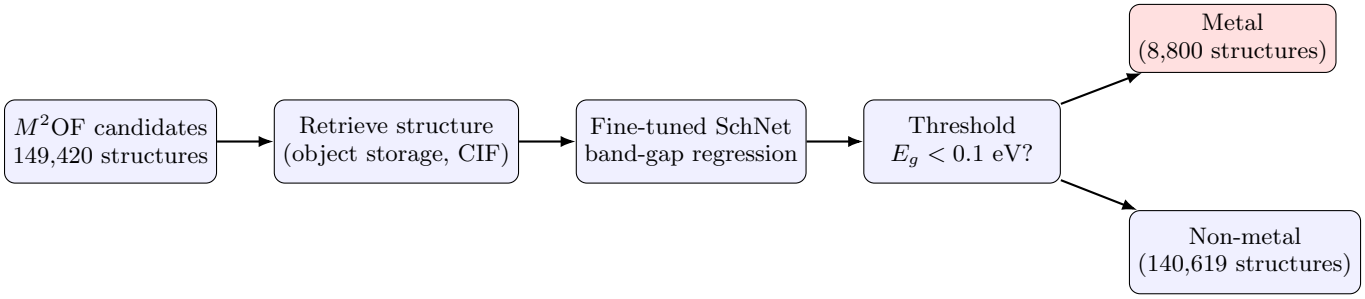

FIG. 1. Metallic-character screening pipeline applied to the full M^2OF candidate database. Each structure is retrieved on demand, converted to an atomistic graph, and passed through the trained SchNet regressor; a predicted band gap below 0.1 eV labels the structure metallic.

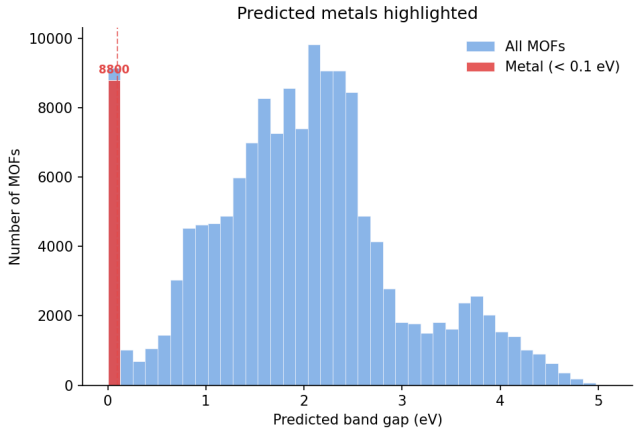

FIG. 2. Distribution of predicted band gaps over all 149,419 successfully screened M^2OF structures (blue), with the metallic subset ($E_g < 0.1$ eV, dashed line) overlaid in red and directly labeled by bin count.

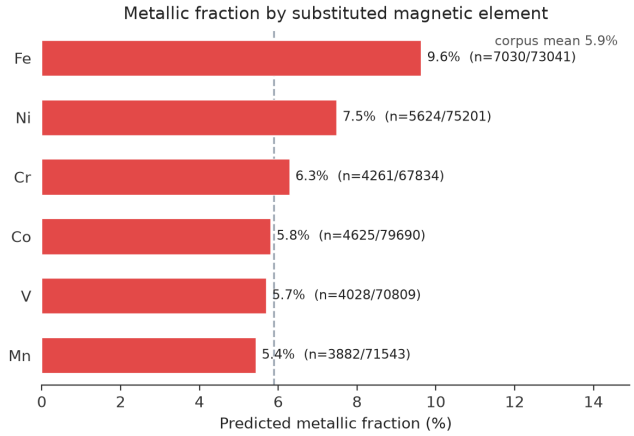

FIG. 3. Predicted metallic fraction among M^2OF structures containing each substituted magnetic transition-metal element, compared against the corpus-wide mean (dashed line, 5.89%). Counts in parentheses give (metal, total) structures containing that element; elements co-occur within a single structure, so the six subsets are not mutually exclusive.

flagged here as likely metallic, those with a large but not extreme predicted magnetic moment, so as to avoid the apparent depletion seen at the $\geq 60\,\mu_B$ tail, are natural next candidates for the phonon-stability and DFT validation pipeline already established for this database.

V. CONCLUSION

We have screened the full 149,420-structure M^2OF database of site-substituted magnetic MOF candidates for metallic band character, using a SchNet regressor trained on OMDB-GAP1 band gaps and a fixed 0.1 eV threshold. 8,800 structures (5.89%) are flagged as likely metallic, with Fe-substituted frameworks enriched relative to the corpus mean. Restricting to the highest-magnetic-moment candidates identified for DFT follow-up in the companion generative study [4], none are predicted metallic, motivating a joint magnetism-and-conductivity objective for future rounds of candidate selection rather than ranking by magnetic moment alone. As with the underlying generative workflow, this screen is intended to reduce the search space ahead of first-principles validation, not to replace it.

DATA AVAILABILITY

Predicted band gaps and metal/non-metal labels for all 149,419 screened structures, together with the trained SchNet checkpoint, are available upon request from the authors, consistent with the data-availability terms of Ref. [4].

ACKNOWLEDGMENTS

NORDITA is supported in part by NordForsk. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS).

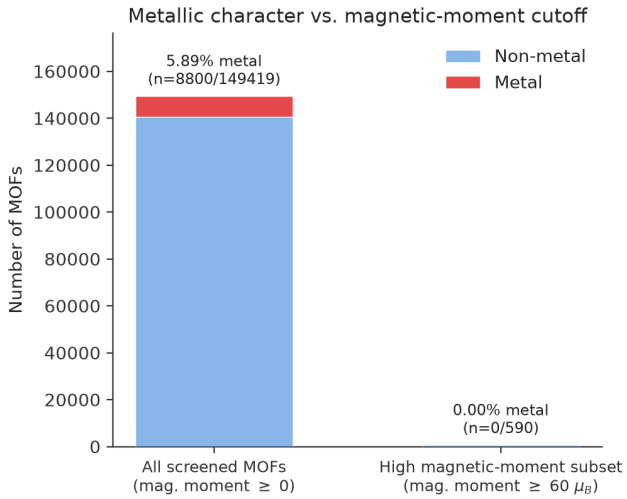

FIG. 4. Predicted metal vs. non-metal counts for the full screened corpus compared against the 590-structure subset with the largest predicted total magnetic moment (mag. moment $\geq 60\,\mu_B$) selected for DFT follow-up in Ref. [4].

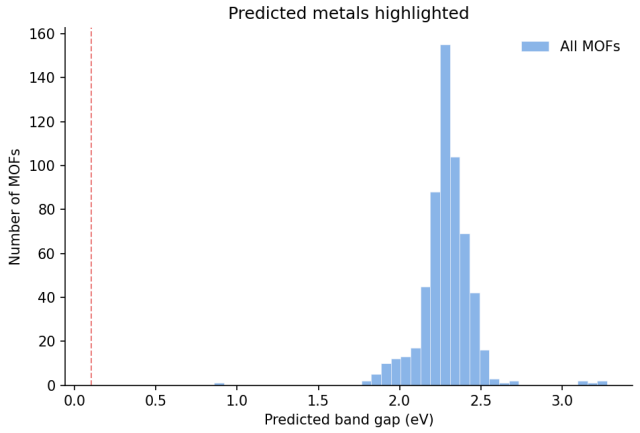

FIG. 5. Predicted band-gap distribution restricted to the 590 candidates with mag. moment $\geq 60\,\mu_B$. No structure in this subset falls below the 0.1 eV metallic threshold (dashed line).

[1] M. Kurmoo, Magnetic metal-organic frameworks, Chemical Society Reviews **38**, 1353 (2009).

[2] G. M. Espallargas and E. Coronado, Magnetic functionalities in MOFs: from the framework to the pore, Chemical Society Reviews **47**, 533 (2018).

[3] A. E. Thorarinsdottir and T. D. Harris, Metal-organic framework magnets, Chemical Reviews **120**, 8716 (2020).

[4] A. C. Tyner, A. Pathapati, and A. V. Balatsky, Generation of magnetic metal-organic frameworks, arXiv preprint arXiv:2604.27879 (2026).

[5] A. S. Rosen, V. Fung, P. Huck, C. T. O'Donnell, M. K. Horton, D. G. Truhlar, K. A. Persson, J. M. Notestein, and R. Q. Snurr, High-throughput predictions of metal-organic framework electronic properties: theoretical challenges, graph neural networks, and data exploration, npj Computational Materials **8**, 112 (2022).

[6] A. S. Rosen, S. M. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J. M. Notestein, and R. Q. Snurr, Machine learning the quantum-chemical properties of metal-organic frameworks for accelerated materials discovery, Matter **4**, 1578 (2021).

[7] B. Olsthoorn, R. M. Geilhufe, S. S. Borysov, and A. V. Balatsky, Band gap prediction for large organic crystal structures with machine learning, Advanced Quantum Technologies **2**, 1900023 (2019).

[8] S. S. Borysov, R. M. Geilhufe, and A. V. Balatsky, Organic materials database: An open-access online database for data mining, PLoS ONE **12**, e0171501 (2017).

[9] R. M. Geilhufe, B. Olsthoorn, and A. V. Balatsky, Shifting computational boundaries for complex organic materials, Nature Physics **17**, 152 (2021).

[10] A. Tyner and A. V. Balatsky, Screening the organic materials database for superconducting metal-organic frameworks, PLOS ONE **19**, 1 (2024).

[11] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, SchNet – a deep learning architecture for molecules and materials, The Journal of Chemical Physics **148**, 241722 (2018).

[12] K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller, SchNetPack: A deep learning toolbox for atomistic systems, Journal of Chemical Theory and Computation **15**, 448 (2019).

[13] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, *et al.*, The atomic simulation environment – a Python library for working with atoms, Journal of Physics: Condensed Matter **29**, 273002 (2017).