Friction https://mc03.manuscriptcentral.com/friction
https://doi.org/10.26599/FRICT.2026.9441292

Research Article

A physics-informed active learning framework for the inverse design of vacuum coatings

Wanyu Wang[1,2], Bowen Xu[3], Zhangli Hou[1], Zhen Li[1], Zhiyuan Wang[3], Qifeng Cui[4], Zhinan Zhang[1]✉

[1] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

[2] SJTU Paris Elite Institute of Technology, Shanghai Jiao Tong University, Shanghai 200240, China

[3] School of Materials Science and Chemical Engineering, Harbin University of Science and Technology, Harbin 150080, China

[4] National Key Laboratory of Aerospace Mechanism, Institute of Aerospace System Engineering, Shanghai, Shanghai 201108, China

✉ Corresponding author. E-mail: zhinanz@sjtu.edu.cn

Received: March 26, 2026; Revised: June 25, 2026; Accepted: July 17, 2026

Abstract: The multi-objective optimization of Physical Vapor Deposition (PVD), as a predominant modality of vacuum coating, is indispensable for surface engineering, yet its optimization is bottlenecked by the conflict between experimental efficiency and data scale. Traditional trial-and-

error and purely data-driven machine learning suffer from low efficiency and severe overfitting under typical "small-sample" constraints, often recommending non-physical or unsafe parameters. To address these challenges, this study proposes a closed-loop Physics-Informed Active Learning (PIAL) framework integrating three core modules: physics-enhanced feature engineering, a Physics-Regularized Heteroscedastic Neural Network (PR-HNN), and a risk-aware dual-point batch acquisition strategy. By embedding fundamental sputtering kinetics into the PR-HNN via continuous partial derivative constraints, the model is softly regularized toward physically plausible monotonic and saturating trends within the investigated CoCrFeNiTi magnetron sputtering process window, while effectively quantifying non-stationary experimental noise. Experimental results demonstrate that physics-based regularization reduces the generalization error of the PR-HNN by 41% compared to purely data-driven baselines, effectively eliminating non-physical fluctuations in sparse data regions. Guided by the risk-aware acquisition function, the framework progressively expanded the initial Pareto front through three closed-loop iterations. The PIAL-optimized coatings achieved a peak hardness of approximately 13.4 GPa and a minimum wear depth of 0.25 μm, while maintaining robust corrosion resistance. This methodology establishes a highly efficient, physically interpretable, and safe paradigm for AI-driven surface engineering and high-performance coating development.

Keywords: Active Learning; Neural Network; Vacuum Coating; Physical Vapor Deposition; Tribo-informatics

1 Introduction

As a fundamental modality in surface engineering, vacuum deposition is indispensable for the longevity and reliability of core components in strategic sectors such as aerospace and maritime industries [1]. Among various vacuum-based technologies, Physical Vapor Deposition (PVD) stands out as a premier technique for fabricating protective films with superior chemical and mechanical

stability [2,3]. Driven by the escalating demands of modern industrial applications, PVD coating development is no longer limited to unilateral property enhancement. Instead, it now demands a synergistic optimization of multiple attributes, such as achieving a delicate balance between hardness and toughness while simultaneously maximizing wear resistance [4]. This paradigm shift requires research methodologies to move beyond simple functional testing toward complex, integrated process optimization.

The PVD process involves intricate plasma physics [5] and thin-film growth kinetics [6], which are governed by a combination of macroscopic process parameters, encompassing bias voltage [7,8], target power [9], gas pressure [10], substrate temperature, etc. [11]. The broad operating range of each parameter results in a vast, high-dimensional search space, where the influence on coating performance is highly non-linear and intricately coupled [12]. Meanwhile, the pursuit of multi-objective optimization is further complicated by the inherent antagonistic effects between different intrinsic properties [13]. While numerical simulations have been employed to explore PVD mechanisms [14–16], they remain secondary tools due to their prohibitive computational overhead and the inevitable "reality gap" regarding actual vacuum chamber conditions. Thus, physical experiments serve as the definitive validation method, though high operational costs and lengthy characterization cycles typically limit the resulting datasets to an extremely small scale [17].

Traditional experiment-intensive PVD development typically follows a static, open-loop sampling paradigm driven by empirical trial-and-error. Conventionally, the workflow initiates with discretizing process variables based on prior expertise, prescribing fixed experimental matrices via One-Factor-at-a-Time (OVAT) or Orthogonal Experimental Design (OED) schemes. Specifically, OVAT isolates individual variables along orthogonal axes, creating a restricted "cross-shaped" trajectory [18], while OED employs static arrays to establish a uniform grid of discrete points [19]. Both paradigms adhere to a "design-then-execute" logic, leading to a fundamental disconnection between sampling and performance feedback. Without the ability to dynamically integrate

intermediate results, these systems lack the flexibility to dynamically steer sampling toward high-performance regions [20]. Furthermore, traditional analysis relies heavily on human intuition to interpret discrete data, which struggles to fully resolve the complex synergistic couplings in multi-variable PVD processes [21]. Consequently, while providing a structured foundation for process exploration, these methods risk inefficient resource allocation in low-gradient regions and frequently overlook narrow-window global optima hidden within the gaps of the static grid.

To overcome the bottlenecks of conventional experimental methods, data-driven methodologies have been progressively leveraged to advance the coating research [22,23]. Fernández-López et al. [24] employed Random Forest (RF) and XGBoost to predict the thickness of plasma electrolytic oxidation coatings. Xu et al. [25] coupled Gaussian Mixture Models (GMM) with Support Vector Regression (SVR) for the process parameter design of plasma spraying. Zhang et al. [26] integrated feature engineering with multiple Machine Learning (ML) algorithms to discover superhard high-entropy nitride coatings within vast compositional spaces. Despite their strong fitting capabilities, the predictive reliability of conventional ML models depends strongly on the coverage and quality of the training data. When only limited experimental observations are available, flexible data-driven models may overfit the observed samples and produce poorly constrained extrapolations in unexplored regions of the PVD process space [27,28]. Furthermore, these "black-box" models frequently neglect the underlying physics of plasma processes [29], risking the blind recommendation of extreme parameters that could trigger catastrophic target poisoning or arc discharge. The emerging field of tribo-informatics [30,31] demonstrates that integrating machine learning with tribological principles and design methodology can effectively model complex tribo-behavior from small datasets while preserving physical interpretability [30–33], which is a capability highly relevant to advancing coating development and inverse design. Unlike traditional forward exploration that blindly maps inputs to outputs, inverse design represents a fundamentally objective-driven paradigm: it starts with the desired multi-dimensional performance targets and autonomously

deduces the optimal macroscopic process parameters required to achieve them. Consequently, an intelligent inverse design framework for PVD processes is urgently required to simultaneously ensure small-sample compatibility, strict experimental safety, and profound physical interpretability.

To address these requirements, this study proposes a closed-loop Physics-Informed Active Learning (PIAL) framework. The framework integrates three core modules: First, a physics-enhanced feature engineering module maps macroscopic parameters to fundamental sputtering kinetics. Second, a Physics-Regularized Heteroscedastic Neural Network (PR-HNN) is constructed, which embeds physical laws via partial derivative constraints to prevent irrational extrapolation under extreme data scarcity. Third, a risk-aware dual-point batch acquisition strategy is formulated to safely balance exploration and exploitation while strictly avoiding hazardous equipment regimes. Validated through the multi-objective inverse design of CoCrFeNiTi High-Entropy Alloy (HEA) multilayer films, this methodology establishes a highly efficient, physically interpretable, and safe paradigm for the accelerated discovery of advanced functional coatings.

2 Physics-Informed Active Learning Methodology

2.1 Problem Formulation and PIAL Framework

To bridge the gap between black-box modeling and physical causality under the small-sample constraints of PVD optimization, this study selects the CoCrFeNiTi HEA system as a representative benchmark. This system is critical for surface protection in extreme aerospace and marine environments, and its intricate multi-element synergies provide a rigorous high-dimensional space for the multi-objective formulation. The complex process optimization is thus decoupled into a Primary-Auxiliary constrained system.

Let the composite input space be defined as $\boldsymbol{\chi}$, the primary intrinsic objective as $\boldsymbol{y}_{\mathbf{pri}}$, and the auxiliary performance outputs as $\boldsymbol{y}_{\mathbf{aux}}$. The optimization problem is formulated as:

$$\max_{\boldsymbol{x}\epsilon\boldsymbol{\chi}} f_{\mathrm{pri}}(\boldsymbol{x}) + \epsilon(\boldsymbol{x}) \tag{1}$$

$$\text{subject to: } P(\boldsymbol{y}_{\mathbf{aux}}(\boldsymbol{x}) < \boldsymbol{y}_{\mathbf{th}}) \geq P_{\mathrm{th}} \tag{2}$$

$$g_j(\boldsymbol{x}) \leq 0, \qquad j = 1, \ldots, K \tag{3}$$

Where $f_{pri}(\boldsymbol{x})$ represents the underlying non-linear mapping of the primary objective approximated by the active learning surrogate; $\epsilon(\boldsymbol{x})$ denotes the input-dependent heteroscedastic noise characterizing dynamic measurement uncertainties; y_{th} represents the safety thresholds for auxiliary properties; and $g_j(\boldsymbol{x}) \leq 0$ defines the j-th explicit physical or equipment boundary constraint.

The PIAL framework addresses this constrained formulation through a closed-loop architecture. Traditional open-loop methods rely on static experimental designs followed by disjointed fabrication and characterization, which severely limits the optimization efficiency in high-dimensional spaces (Fig. 1a). In contrast, the proposed PIAL framework (Fig. 1b) realizes a fundamental paradigm shift from parameter-driven blind search to objective-driven inverse design. Rather than treating physical rules as mere post-processing filters, the PIAL framework embeds domain physical knowledge at the foundational algorithm level through three synergistic mechanisms:

Physics-Enhanced Feature Engineering: Mapping macroscopic process variables to intrinsic physical descriptors to simplify the high-dimensional mapping space.

Heterogeneous Surrogate Ensembles: Deploying an ensemble of PR-HNNs to anchor the topological consistency of the primary objective, alongside Standard Gaussian Processes (Standard GPs) to probabilistically monitor auxiliary responses.

Risk-Aware Constrained Batch Acquisition: Formulating the overarching PIAL acquisition function (α_{PIAL}) that multiplicative cascades statistical gains, performance feasibilities, and physical safety constraints. A dual-point batch strategy is further integrated to dynamically balance local exploitation and global exploration within the physically safe manifold.

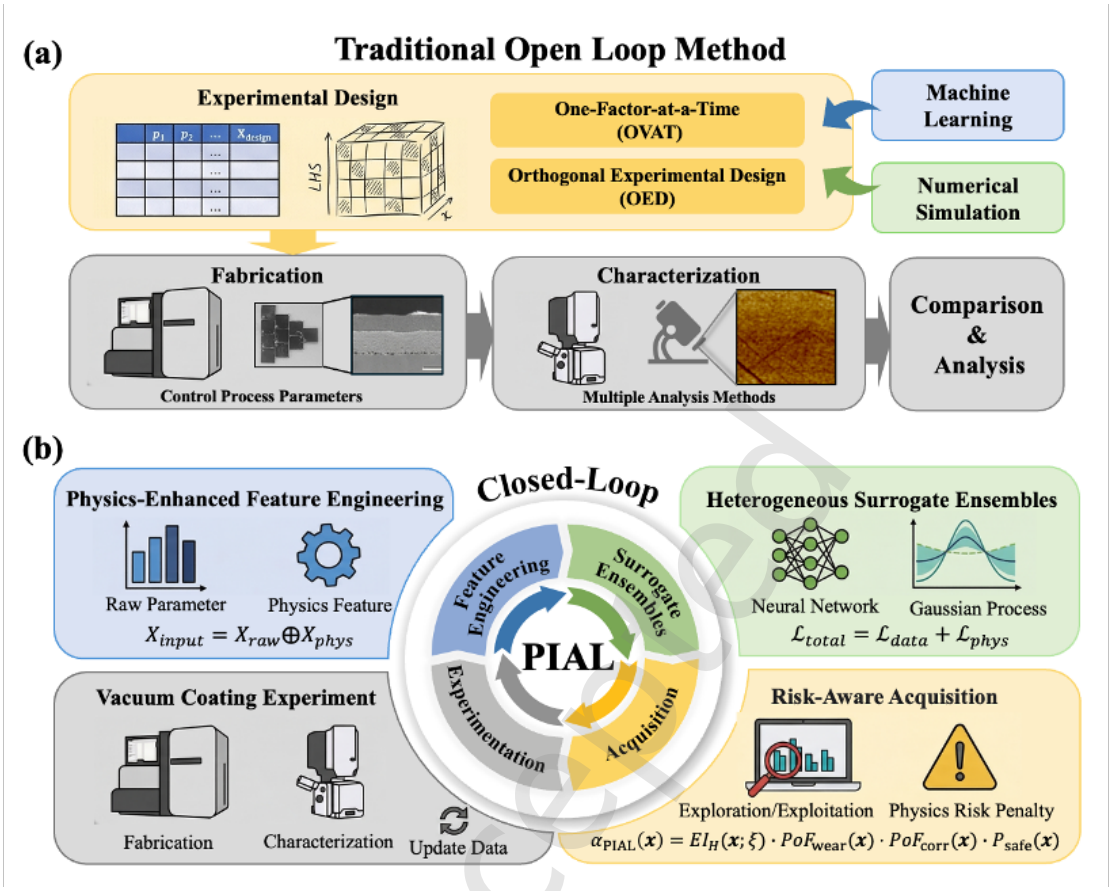

Fig. 1 Paradigm shift from traditional forward optimization to the closed-loop PIAL inverse design framework. (a) Traditional open-loop method (b)The closed-loop PIAL framework

2.2 Physics-Enhanced Feature Engineering

To explicitly capture the underlying plasma dynamics and film growth kinetics, the first layer of the PIAL framework transforms macroscopic operational variables into physically meaningful descriptors. For the multi-target co-sputtering system in this study, the raw parameter vector is defined as:

$$\boldsymbol{X}_{\mathbf{raw}} = \left[P_{\mathrm{Co}}, P_{\mathrm{Ti}}, V_{\mathrm{bias}}, t, N_{\mathrm{layer}}\right]^{\mathrm{T}} \tag{4}$$

Where P_{Co} and P_{Ti} represent the initially set sputtering powers for the respective metallic targets, V_{bias} is the substrate bias voltage, t is the total deposition time, and N_{layer} is the designated number of modulation layers. Under small-sample constraints, mapping this decoupled $\boldsymbol{X}_{\mathbf{raw}}$ directly to mechanical performance significantly increases the search complexity and the risk of overfitting.

To endow the surrogate model with *a priori* mechanistic awareness, a physics-enhanced feature vector $\boldsymbol{X}_{\mathbf{phys}}$ is constructed. As illustrated in Fig. 2a, this feature engineering process explicitly maps macroscopic control variables to microscopic physical phenomena, specifically target surface sputtering, substrate surface ion bombardment, and film cross-sectional growth kinetics. Based on glow discharge and film growth kinetics, three key derived features are formulated:

Total Sputtering Power: The macroscopic deposition rate and the overall plasma ionization level do not solely depend on individual target settings but are fundamentally governed by the aggregated energy input [34]. Therefore, this study introduces the total sputtering power as a primary derived feature to characterize the global intensity of the glow discharge. It is formulated as:

$$P_{\mathrm{total}} = P_{\mathrm{Co}} + P_{\mathrm{Ti}} \tag{5}$$

Bias Energy Flux: During the PVD process, the substrate bias accelerates argon and metal ions within the plasma, inducing an intense bombardment effect on the growing film. To accurately quantify the comprehensive intensity of this ion bombardment, the bias voltage is coupled with the total sputtering power to formulate the bias energy flux. This physical quantity strongly correlates with adatom mobility on the substrate, microstructural densification of the coating, and the evolution of residual compressive stress [34,35]. It is defined as:

$$f_2 = V_{\mathrm{bias}} \cdot P_{\mathrm{total}} \tag{6}$$

Deposition Time per Layer: In the design of multilayer nanostructures, the individual layer thickness is a critical variable dictating the ultimate mechanical properties of the material. To enable the model to explicitly perceive this micro-scale information, the deposition time per layer is extracted as a feature. From a kinetic perspective, this parameter governs the single-layer thickness and serves as the core metallurgical factor triggering the Hall-Petch strengthening effect, effectively hindering dislocation glide via heterogeneous interfaces within multilayer architectures [36,37]. It is calculated as:

$$f_3 = \frac{t}{N_{\mathrm{layer}}} \tag{7}$$

Consequently, the physical derived feature vector is established as $\boldsymbol{X}_{\mathbf{phys}} = [f_1, f_2, f_3]^{\mathrm{T}}$. To explicitly integrate macroscopic equipment control with microscopic growth kinetics, the ultimate input space is constructed by concatenating the raw parameters and the physics-enhanced features:

$$\boldsymbol{X}_{\mathbf{input}} = \boldsymbol{X}_{\mathbf{raw}} \oplus \boldsymbol{X}_{\mathbf{phys}} \tag{8}$$

Where $\oplus$ denotes the vector concatenation operator. Finally, to eliminate the adverse effects of dimensional heterogeneity across different physical quantities and accelerate the convergence of the surrogate model, all features undergo rigorous Z-Score standardization before being fed into the network:

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \tag{9}$$

Where z_i is the standardized value of the i-th feature dimension, and μ_i and σ_i denote the empirical mean and standard deviation of that specific feature across the training dataset, respectively. This physics-guided augmentation profoundly reduces the mapping complexity, bridging the gap between statistical learning and physical reality.

2.3 Heterogeneous Surrogate Ensembles

The surrogate model serves as the core engine driving the closed-loop optimization in active learning. To robustly navigate extreme sparse-data environments and accommodate the fundamentally different mechanistic behaviors between intrinsic mechanical properties and system-level responses, this study constructs a heterogeneous surrogate architecture (Fig. 2b). This specific design simultaneously improves physical topological fidelity and robust uncertainty quantification.

For the primary intrinsic objective (hardness), a PR-HNN is developed. In practical PVD characterizations, experimental measurement noise is rarely homoscedastic. Coatings fabricated under extreme parameters may exhibit internal micro-cracks or severe localized spalling, significantly amplifying the variance during nanoindentation or tribological testing. To account for this inherent heteroscedasticity, the base data-driven objective of the PR-HNN discards the standard

Mean Squared Error (MSE). Instead, it incorporates the true experimental variance ($\sigma^2_{\text{exp},i}$) of each sample to construct a variance-weighted empirical data loss ($\mathcal{L}_{\text{data}}$):

$$\mathcal{L}_{\text{data}} = \frac{1}{N}\sum_{i=1}^{N}\frac{(\hat{y}_i - y_i)^2}{\sigma^2_{\text{exp},i} + \epsilon} \tag{10}$$

Where y_i and $\hat{y}_i$ are the true and predicted hardness values, respectively. The constant ϵ is introduced to prevent excessively large sample weights when the measured variance approaches zero and thereby maintain numerical stability during optimization. A sensitivity analysis of ϵ is provided in Table S1. This mechanism adaptively down-weights low-confidence experimental data points during the backpropagation process, thereby significantly enhancing algorithmic robustness against anomalous measurement noise.

However, relying solely on minimizing the empirical risk in a high-dimensional space inevitably leads to gradient evolutions that violate materials science common sense. Inspired by Physics-Informed Neural Networks (PINNs) [38,39], the PR-HNN leverages the automatic differentiation capabilities of deep learning frameworks to explicitly embed partial-derivative physical laws into the total loss function:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{phys}} \tag{11}$$

Where the aggregate physics-regularized term, $\mathcal{L}_{phys}$, is defined as the weighted sum of individual constraint penalties:

$$\mathcal{L}_{\text{phys}} = \sum_{k=1}^{3}\lambda_k \mathcal{L}_{\Phi_k} \tag{12}$$

Where λ_k represents the penalization weight for the k-th physical constraint. As explicitly illustrated in Fig. 2b, the relative contributions of the empirical and physics-based losses are adjusted according to data availability. During the initial sparse-data stage, the physics loss ($\mathcal{L}_{\text{phys}}$) is assigned a relatively larger weight, allowing the derivative constraints to provide stronger inductive bias and suppress nonphysical gradients and curvatures in poorly sampled regions. As additional experimental

data are incorporated through the active-learning iterations, the relative contribution of the empirical data-fitting loss ($\mathcal{L}_{\mathrm{data}}$) is progressively increased, enabling the model to capture more detailed local process–property relationships. Importantly, the physical constraints are implemented as soft inequality penalties with finite weights rather than as substitutes for experimental information. The empirical loss, therefore, remains active throughout training and can override or correct the prescribed prior trends when they are inconsistent with the observations. Accordingly, the PR-HNN is designed as a physics-regularized surrogate initialized from baseline samples and progressively refined through experimental feedback, thereby improving predictive stability and sample efficiency under the investigated small-sample conditions.

To ensure the robust generalization reliability of the model when facing non-ideal experimental perturbations such as target poisoning or sudden arc discharge, the aforementioned derivative constraints are integrated as soft inequality penalties. The finite magnitude of the penalty weights allows the empirical data loss ($\mathcal{L}_{\mathrm{data}}$) to dynamically compete with the physics loss ($\mathcal{L}_{\mathrm{phys}}$) during optimization, thereby effectively avoiding the risk of introducing systematic biases caused by rigid physical hard boundaries locking the model. Furthermore, as a post-training diagnostic mechanism to systematically evaluate the compatibility between physical assumptions and sparse empirical data, this study introduces the Physical Violation Rate (PVR_{k}) to quantitatively monitor the collocation points after model convergence:

$$PVR_k = \frac{1}{N_{grid}} \sum_{i=1}^{N_{grid}} I(\Phi_k(x_i) > 0) \tag{13}$$

Where $I(\cdot)$ is the indicator function. If a high PVR or large violation residual persists after model convergence, or if further increasing λ_k leads to a significant degradation in validation error, it indicates a potential mismatch between the designated physical prior and the actual experimental state, which instructs the framework to relax, re-weight, or deactivate the corresponding constraint.

Instead of imposing rigid piecewise boundaries, this study defines three continuous derivative-level regularization operators (Φ_k) based on specific metallurgical and kinetic mechanisms:

Monotonic Kinetic Constraint (Φ_1): According to sputtering kinetics, an increase in the total sputtering power (P_{total}) fundamentally enhances the plasma density and deposition rate, leading to a denser microstructure and higher hardness. Assuming the process operates within a safe regime devoid of catastrophic macroscopic cracking, this kinetic densification dictates that the first-order partial derivative is expected to be non-negative within the investigated safe window [34]. A ReLU-based penalty is triggered if the model predicts a spurious negative slope:

$$\mathcal{L}_{\Phi_1} = \frac{1}{N_{\text{grid}}} \sum \left[\text{ReLU}\left(-\frac{\partial H}{\partial P_{\text{total}}} \right) \right]^2 \tag{14}$$

Where H denotes the predicted intrinsic hardness of the coating, and N_{grid} represents the total number of collocation points uniformly sampled across the input parameter space to evaluate the physical penalty.

Ion Bombardment Concavity Constraint (Φ_2): The evolution of coating hardness with respect to substrate bias (V_{bias}) exhibits a complex saturation effect. Moderate bias promotes surface adatom migration and film densification; however, excessive ion bombardment triggers severe residual compressive stress accumulation and re-sputtering, causing the hardness growth to stall or degrade [34,35,40]. To elegantly capture this dynamic without enforcing rigid inflection points, a second-order concavity constraint is introduced:

$$\mathcal{L}_{\Phi_2} = \frac{1}{N_{\text{grid}}} \sum \left[\text{ReLU}\left(\frac{\partial^2 H}{\partial V_{\text{bias}}^2} \right) \right]^2 \tag{15}$$

By penalizing positive second derivatives, this constraint is primarily utilized to prevent the model from predicting unreasonable exponential hardness growth in the high-bias region, guiding the model via the soft physics prior to learn a plausible evolutionary trend from ion-bombardment enhancement and film densification to hardness gain saturation.

Hall-Petch Strengthening Constraint (Φ_3): In multilayer architectures, reducing the single-layer deposition time (t_{layer}) decreases the modulation period, which strengthens the coating by hindering dislocation glide via dense interfaces [41]. This inverse monotonic relationship is guided by penalizing any positive gradient predictions:

$$\mathcal{L}_{\Phi_3} = \frac{1}{N_{\text{grid}}} \sum \left[\text{ReLU}\left(\frac{\partial H}{\partial t_{\text{layer}}} \right) \right]^2 \tag{16}$$

While the PR-HNN explicitly improves physical consistency, standard neural networks only output deterministic mean values, lacking the uncertainty estimation strictly required by Bayesian active learning. To support uncertainty-aware active learning, we therefore employ a deep-ensemble strategy. Multiple independently initialized PR-HNN models are trained under the same architecture and loss formulation. For an unseen candidate, the ensemble mean is used as the hardness prediction, whereas the dispersion among ensemble outputs is used as an estimate of epistemic uncertainty. This uncertainty estimate is used as a diagnostic indicator of prediction stability rather than as an independent claim of improved predictive accuracy. An ensemble of M independent PR-HNNs is trained in parallel, each initialized with different random seed weights. For any unseen candidate parameter $\boldsymbol{x}$, the predictive mean $\mu_{\text{pri}}(\boldsymbol{x})$ and the epistemic uncertainty σ_{pri}^2 are quantified through the statistical distribution of the ensemble's outputs:

$$\mu_{\text{pri}}(\boldsymbol{x}) = \frac{1}{M} \sum_{m=1}^{M} \hat{y}_m(\boldsymbol{x}) \tag{17}$$

$$\sigma_{\text{pri}}^2 = \frac{1}{M} \sum_{m=1}^{M} \left(\hat{y}_m(\boldsymbol{x}) - \mu_{\text{pri}}(\boldsymbol{x}) \right)^2 \tag{18}$$

This architecture elegantly resolves the exploration and exploitation dilemma for the primary objective, enabling the algorithm to confidently explore highly uncertain regions while keeping its patterns regularized within plausible physical boundaries.

Conversely, the situation is fundamentally different for the auxiliary system-level responses, which encompass wear depth, Coefficient of Friction (COF), and corrosion current. Friction and wear are dynamic system behaviors involving contact mechanics, microscopic fracture, and tribochemical reactions, while corrosion is a multivariable coupled interfacial electrochemical process. At the macroscopic process parameter level, these phenomena strictly lack universal analytical formulas that can be directly translated into explicit partial derivative constraints.

Forcing a physical neural network to fit these highly complex and unconstrained dimensions under extreme data sparsity would fail to leverage the inherent advantages of physics-informed models. Instead, it would introduce severe epistemic noise and significantly exacerbate the risk of overfitting.

Therefore, conforming to the primary and auxiliary decoupling strategy, this study deploys Standard GP regressors equipped with Radial Basis Function (RBF) and White Kernels as lightweight surrogate models for these auxiliary properties. Operating in a purely data-driven Bayesian inference mode, these GPs robustly output the predictive distributions (μ_{aux},σ^2_{aux}) for both tribological and electrochemical responses. Here, σ_{aux} denotes the posterior predictive standard deviation of the corresponding auxiliary GP rather than a sample-specific standard deviation derived from repeated physical measurements.

Ultimately, this heterogeneous architecture maximizes the utilization of diverse domain knowledge. By combining physics-regularized deep networks for primary intrinsic traits with non-parametric GPs for auxiliary system traits to formulate the final concatenated predictive output, the framework effectively avoids conflicting gradient optimizations caused by imposing unreasonable physical constraints.

2.4 Risk-Aware Constrained Batch Acquisition

In active learning, striking an optimal balance between exploring unknown regions and exploiting known high-yield areas is critical. For PVD coating optimization, the objectives inherently

possess a strict engineering hierarchy: mechanical hardness dictates the primary performance ceiling, tribological and electrochemical responses act as necessary survival thresholds, and macroscopic equipment limits form absolute physical safety boundaries. Traditional unconstrained acquisition functions struggle to encapsulate this asymmetric hierarchy and are prone to recommending extreme parameters that could trigger equipment arcing. To explicitly address this, this study formalizes a custom PIAL acquisition function, α_{PIAL}, structured as a multiplicative cascade of the primary driving force, auxiliary performance feasibilities, and physical safety barriers:

$$\alpha_{\mathrm{PIAL}}(\boldsymbol{x}) = EI_H(\boldsymbol{x};\xi) \cdot PoF_{\mathrm{wear}}(\boldsymbol{x}) \cdot PoF_{\mathrm{corr}}(\boldsymbol{x}) \cdot P_{\mathrm{safe}}(\boldsymbol{x}) \tag{19}$$

The primary optimization engine is driven by the Expected Improvement (EI) of hardness (EI_H) [42]. Leveraging the predictive mean $\mu_H(\boldsymbol{x})$and epistemic uncertainty $\sigma_H(x)$ derived from the PR-HNN ensemble, EI_H quantifies the statistical expectation of surpassing the current maximum observed hardness (f∗). It is mathematically defined as

$$EI_H(\boldsymbol{x};\xi) = \mu_H(\boldsymbol{x}) - f^* - \xi)\Phi(Z) + \sigma_H(\boldsymbol{x})\phi(Z) \tag{20}$$

Where $Z = (\mu_H(\boldsymbol{x}) - f^* - \xi)/\sigma_H(\boldsymbol{x})$, and Φ, ϕ represent the standard normal cumulative distribution and probability density functions, respectively. The trade-off parameter ξ dynamically dictates the algorithm's greediness.

To account for the auxiliary performance requirements, wear depth (D_{wear}) and corrosion current (I_{corr}) were treated as minimization-type feasibility constraints. The corresponding thresholds were anchored to the best auxiliary performance observed in the initial training set. These values serve as baseline feasibility criteria requiring a candidate to have a finite probability of matching or exceeding the best initial auxiliary performance. A sensitivity analysis over nine combinations of the wear-depth and corrosion-current thresholds confirmed that the top-ranked recommendation remained unchanged; the detailed results are provided in Table S2.

Utilizing the predictive Gaussian distributions generated by the lightweight GPs, the Probabilities of Feasibility (PoF) computes the exact statistical probability that a candidate parameter

will strictly satisfy the predefined survival thresholds ($D_{\mathrm{wear}}^{\mathrm{th}}$ for wear and $I_{\mathrm{corr}}^{\mathrm{th}}$ for corrosion). The wear depth feasibility probability is calculated as

$$PoF_{\mathrm{wear}}(\boldsymbol{x}) = \Phi\left(\frac{D_{\mathrm{wear}}^{\mathrm{th}} - \mu D_{\mathrm{wear}}(\boldsymbol{x})}{\sigma D_{\mathrm{wear}}(\boldsymbol{x})}\right) \tag{21}$$

Where $\mu D_{\mathrm{wear}}(\boldsymbol{x})$ and $\sigma D_{\mathrm{wear}}(\boldsymbol{x})$ denote the posterior predictive mean and standard deviation of the wear-depth GP. The corrosion-current feasibility probability is defined analogously. A higher PoF value indicates a greater predicted probability of satisfying the corresponding auxiliary-performance requirement. This mechanism elegantly transforms complex antagonistic trade-offs into probabilistic gating thresholds.

Simultaneously, the absolute physical boundaries of the PVD equipment are encoded into an equipment-safety indicator (P_{safe}). Based on the explicit physical constraints $g_j(x) \leq 0$ defined in Section 3.1, the safety penalty is formulated as:

$$P_{\mathrm{safe}}(x) = \prod_{j=1}^{J} \mathbb{I}\left[g_j(x) \leq 0\right] \tag{22}$$

Where $\mathbb{I}$ is the indicator function. Accordingly, $P_{\mathrm{safe}}(x) = 1$ only when the candidate satisfies all predefined equipment-safety constraints; violation of any constraint results in $P_{\mathrm{safe}}(x) = 0$. This multiplicative cascade mechanism endows the PIAL algorithm with a mathematical "zero-out" property. Even if an extreme parameter combination possesses an exceptionally high expected hardness gain (EI_H), violation of an equipment-safety boundary causes its overall acquisition value to become zero, strictly confining the active learning trajectory within the physically feasible manifold. This "zero-out" effect is visually evident in the contour plot of Fig. 2c, where the entire high-risk region defined by the substrate bias and CoCrFeNi power coupling is strictly masked out, securely confining the active learning trajectory within the physically feasible manifold.

In practical PVD process development, the cost of a single vacuum pumping and deposition cycle is prohibitively high. Validating only a single parameter per iteration leads to extremely low

experimental throughput. To maximize validation efficiency while explicitly breaking the exploration-exploitation dilemma, this study designs a Dual-Point Sequential Batch Strategy (Fig. 2c). Instead of naively selecting the top two highest-scoring points, which inevitably causes severe information redundancy through spatial clustering, the batch is sequentially generated by dynamically modulating the trade-off parameter ξ alongside a Local Penalization mechanism:

$$\boldsymbol{x_1} = \arg\max_{\boldsymbol{x}}[\alpha_{\mathrm{PIAL}}(\boldsymbol{x}|\xi = 0.01)] \tag{23}$$

$$\boldsymbol{x_2} = \arg\max_{\boldsymbol{x}}[\alpha_{\mathrm{PIAL}}(\boldsymbol{x}|\xi = 0.1) \cdot \varphi(\boldsymbol{x}, \boldsymbol{x_1})] \tag{24}$$

The first recommended point ($\boldsymbol{x_1}$) acts as the Exploitation Point. By setting ξ to a near-zero value, the algorithm becomes highly greedy, pinpointing the Pareto optimal peak with the absolute highest expected hardness within the safe zone. The second point ($\boldsymbol{x_2}$) acts as the Exploration Point. By significantly increasing ξ, the algorithm prioritizes regions with high epistemic uncertainty. To completely eradicate redundancy, the Gaussian local repulsion function is defined as

$$\varphi(\boldsymbol{x}, \boldsymbol{x_1}) = \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x_1}\|^2}{2l^2}\right) \tag{25}$$

Where l controls the spatial extent of the penalized neighborhood. When a candidate approaches the previously selected point $\boldsymbol{x_1}$, $\varphi(\boldsymbol{x}, \boldsymbol{x_1})$ approaches zero, thereby suppressing redundant recommendations. Through this dual-point combinatorial strategy, the system simultaneously dispatches one point for immediate yield maximization and one point for high-risk boundary charting, perfectly reconciling algorithmic foresight with the economic viability of experimental validation.

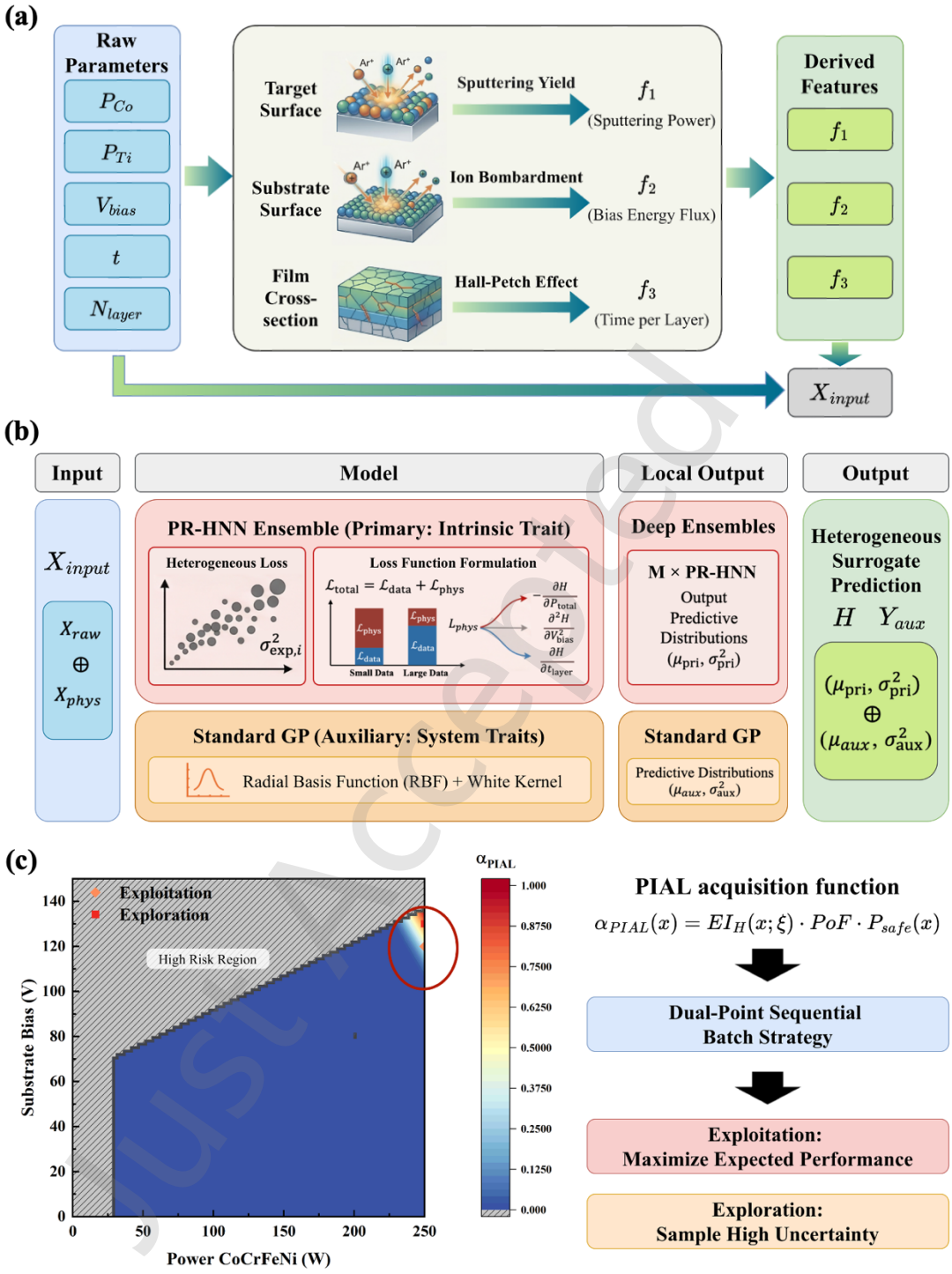

Fig. 2 Core Components of the PIAL Framework. (a) Physics-enhanced feature engineering, illustrating the transformation of raw macroscopic process parameters into physically meaningful derived features. (b) Architectural decomposition of the heterogeneous surrogate ensembles. The primary PR-HNN pipeline integrates a variance-weighted heteroscedastic loss and embeds continuous partial derivative physical constraints, coupled with Deep Ensembles for epistemic-uncertainty quantification. The auxiliary pipeline utilizes Standard Gaussian

Processes for unconstrained system traits. (c) Risk-aware constrained batch acquisition. Balance dynamically local exploitation and global exploration within the physically safe manifold.

2.5 Workflow of PIAL Inverse Design

Traditional materials development and vacuum coating process optimization predominantly adhere to a forward trial-and-error paradigm, wherein parameters are empirically set prior to observing and testing the ultimate performance. Navigating a vast, high-dimensional parameter space in this manner is highly inefficient. The PIAL framework proposed in this study fundamentally shifts this paradigm, realizing a goal-driven "inverse design" approach. Guided by the ultimate objective of achieving Pareto optimality, the algorithm proactively and inversely deduces the optimal process parameter combinations required to attain such performance, strictly bounded by physical laws.

The closed-loop algorithmic data flow of this physics-informed inverse design strictly mirrors the logic of the deployed computational architecture, which can be deconstructed into the following four core iterative steps:

Initialization and Physics-Enhanced Mapping: In the algorithm initiation phase, a minimal empirical baseline dataset is established based on preliminary physical coating experiments. The corresponding multi-objective performances are acquired, while a sample-specific empirical variance is estimated from repeated nanoindentation measurements for the primary hardness objective. The uncertainties of the auxiliary wear and corrosion responses are subsequently represented by the posterior predictive distributions of their respective GP surrogates. Subsequently, the algorithm frontend automatically extracts the derived physics-enhanced features based on PVD kinetic formulas. The raw parameters and derived features undergo rigorous Z-Score standardization, thereby pre-loading physical priors into the system environment.

Heterogeneous Surrogate Ensembles Training: Based on the current dataset, the system trains the heterogeneous surrogate models. For the primary objective, an ensemble of M PR-HNNs is trained in parallel. The backpropagation jointly minimizes the variance-weighted data loss and the

soft physical-derivative penalties. Simultaneously, standard GPs are fitted to the auxiliary responses. This step ensures that the predictive landscape under sparse data conditions adheres to known sputtering and metallurgical kinetics while quantifying epistemic uncertainty.

Risk-Aware Dual-Point Batch Recommendation: Once model training converges, the algorithm enters the inverse deduction phase. Within the high-dimensional space containing thousands of discretized candidate points, the system conducts a global search. The PIAL acquisition function (α_{PIAL}) mathematically cascades the expected improvement of hardness, the PoF for auxiliary traits, and the physical safety survival rate. By dynamically modulating the trade-off parameter and applying a local penalization operator, the algorithm outputs a sequential batch of two mutually independent process parameters: one for aggressive exploitation and one for physically safe exploration.

Experimental Feedback and Closed-Loop Update: Based on the batch inversely recommended by the algorithm, researchers execute coating and characterization experiments on actual magnetron sputtering equipment. The acquired performance results, together with the replicate-derived hardness variance, are appended to the database. The auxiliary GP models are then refitted to update the posterior predictive distributions of wear depth and corrosion current. The aforementioned four steps iterate cyclically until the expected performance of the recommended parameters converges or the pre-set experimental budget is exhausted, whereupon the system ultimately locks in the global optimal process parameter window.

3 Result and Discussion

3.1 Experimental Materials and Methods

To rigorously validate the proposed PIAL framework, the CoCrFeNiTi HEA system was strategically selected as the benchmark material. All CoCrFeNiTi HEA films were deposited using a high-vacuum magnetron sputtering system (JGP600, China, Fig. 3a-I) equipped with equiatomic

CoCrFeNi and pure Ti targets (99.99% purity). The deposition process was conducted under a high-purity argon atmosphere to fabricate the alternating multilayer architecture (Fig. 3b).

As defined in Fig. 3c, the active learning closed-loop primarily regulated five key process parameters: t, P_{Co}, P_{Ti}, V_{bias}, and N_{layer}. The initial training set consisted of five coatings selected, based on prior experimental experience, as an expert-seeded sparse baseline within the equipment-safe process window, thereby establishing the empirical starting point for the active learning closed-loop. Three subsequent dual-point active-learning rounds introduced six additional coatings, resulting in a total experimental dataset of 11 samples.

To comprehensively evaluate the service performance, specifically the mechanical hardness, tribological behavior, and electrochemical corrosion resistance (Fig. 3c), and construct the high-fidelity dataset for the PIAL inverse design, the specialized testing equipment shown in Fig. 3a was employed.

For micro-mechanical properties, the intrinsic hardness and elastic modulus were measured using a nanoindenter (Hysitron TI950, Fig. 3a-II) operating in continuous stiffness measurement (CSM) mode with a maximum load of 5 mN. To explicitly quantify the heteroscedastic measurement noise required by the surrogate model, six independent indentations were randomly performed on each sample to extract the statistical mean and variance from the load-displacement curves.

Tribological performance was assessed via reciprocating sliding tests at room temperature using a multifunctional tribometer (Rtec-5000S, USA, Fig. 3a-III). Tests were conducted under a normal load of 1 N with durations of 30 and 60 minutes. Following the tests, 3D surface profiles of the wear tracks were acquired using a white light interferometer to accurately quantify the wear depth.

Finally, the chemical stability was evaluated using an electrochemical workstation (Bio-Logic VMP-3, Fig. 3a-IV) in a 35 g/L NaCl solution. A standard three-electrode system was employed, with a 1 cm² exposed film area as the working electrode, a saturated calomel electrode as the reference, and a platinum sheet as the counter electrode. Potentiodynamic polarization curves were

recorded at a scanning rate of 0.2 mV/s ranging from −1.4 to +0.2 V to determine the corrosion potential (E_{corr}) and corrosion current density (I_{corr}). Furthermore, electrochemical impedance spectroscopy (EIS) measurements were conducted from 0.01 to 100 000 Hz with a 5 mV RMS amplitude, and the data (two sets per sample) were analyzed using ZView equivalent circuits.

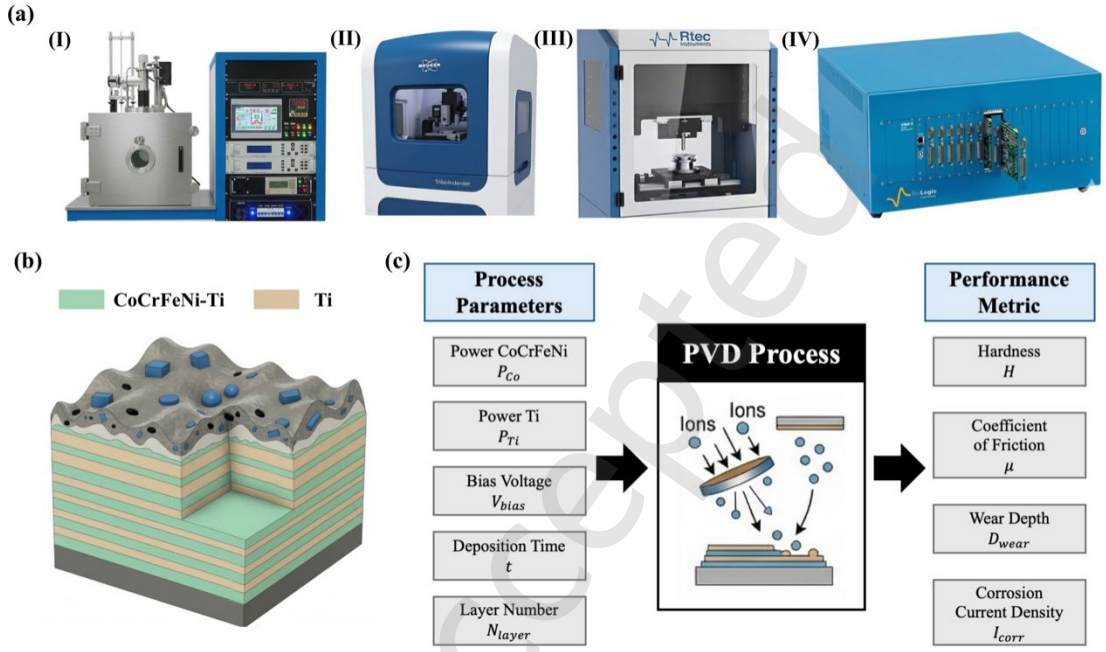

Fig. 3 Experimental Setup, Multilayer Coating Architecture, and Parameter-Performance Mapping Framework. (a) Photographs of the specialized equipment utilized for coating fabrication and multi-objective property characterization: (I) High-vacuum magnetron sputtering system; (II) Nanoindenter; (III) Multifunctional tribometer; (IV) Electrochemical workstation. (b) CoCrFeNi-Ti/Ti HEA multilayer architecture. (c) Block diagram of the PVD process mapping, demonstrating the explicit correlation between the macroscopic input process parameters and the targeted multi-objective performance metrics.

3.2 Physics Consistency and Predictive Capability Validation

To verify that the proposed PR-HNN successfully captures the underlying PVD physics robustly navigates the highly coupled 5D feature space under extreme data scarcity ($N = 5$), three rounds of progressive validation were conducted along the actual experimental trajectory, as summarized in Table S3.

The validation strictly followed the order of sample acquisition. At each stage, all samples available at that point were used for model training, while the subsequently acquired Samples 6–7, 8–9, and 10–11 were sequentially used for testing. Each model was trained exclusively on data obtained before the corresponding experimental round, thereby preventing future-data leakage and faithfully reflecting the continuous model-updating process within the active-learning closed loop.

To evaluate the predictive performance of the PR-HNN, a Standard GP using the original five-dimensional process parameters (Standard GP, raw 5D) and a GP using eight-dimensional physics-enhanced features (GP + physics features, 8D) were selected as baseline models. As summarized in Table S4 and Fig. 4a, the PR-HNN achieved RMSE values of 1.272, 0.927, and 0.812 GPa in the three validation rounds, respectively. These values were consistently lower than those obtained by the Standard GP (1.709, 1.547, and 1.924 GPa) and the GP with physics-enhanced features (1.708, 1.198, and 1.673 GPa).

Across the three rounds, the PR-HNN achieved a mean RMSE of 1.004 GPa and a pooled RMSE of 1.022 GPa. These values correspond to reductions of 41.9% and 41.0%, respectively, relative to the raw-5D Standard GP, and reductions of 34.2% and 33.8%, respectively, relative to the physics-feature-enhanced GP. Moreover, the RMSE of the PR-HNN decreased monotonically from 1.272 GPa in Round 1 to 0.812 GPa in Round 3, indicating progressive improvement and stabilization as additional experimental data were incorporated. In contrast, neither GP baseline exhibited a comparable reduction over the successive rounds. These results demonstrate that the predictive advantage of the PR-HNN was maintained throughout the model-updating process rather than being confined to a single data split.

As shown in the parity plots, the Standard GP (Fig. 4b-I) exhibits a loose scatter indicative of poor correlation. Incorporating physics-enhanced features improves the overall agreement to some extent, but noticeable prediction errors remain (Fig. 4b-II). By comparison, the PR-HNN predictions (Fig. 4b-III) align closely with the ideal $y = x$ dashed line.

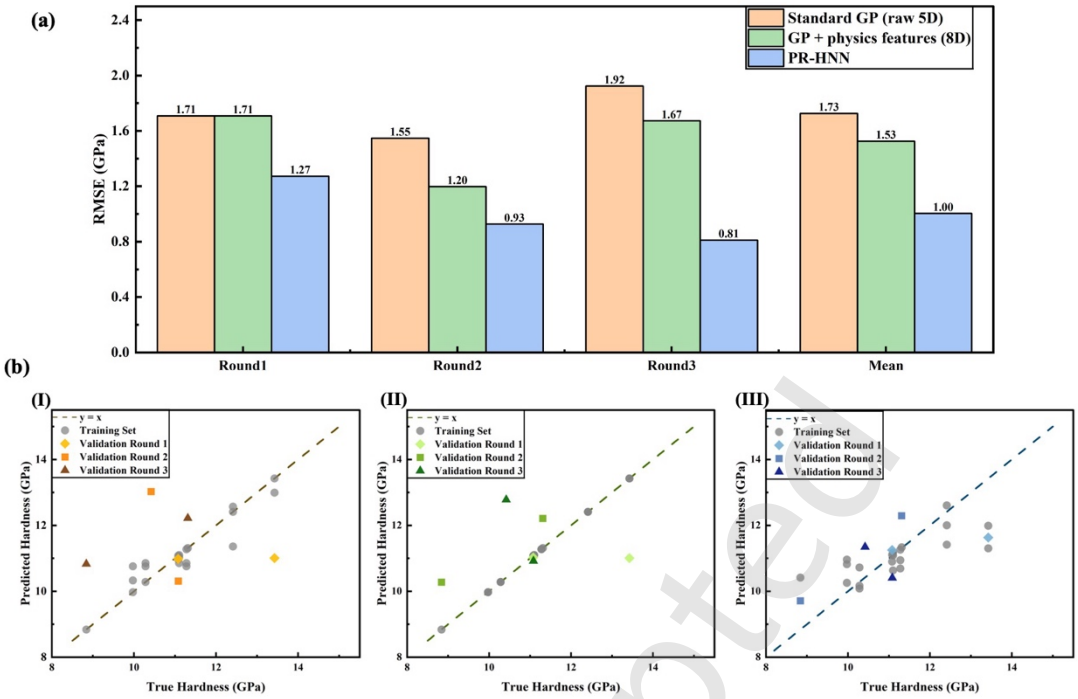

Fig. 4 Chronological Validation and Predictive Performance Comparison of Surrogate Models. (a) RMSE comparison of the Standard GP with raw five-dimensional inputs, the GP with eight-dimensional physics-enhanced features, and the PR-HNN across three successive validation rounds and their mean performance. (b) Parity plots comparing the predicted and experimental hardness values for: (I) Standard GP with raw five-dimensional inputs; (II) GP with physics-enhanced features; and (III) PR-HNN.

To further assess the contributions of the individual model components, ablation analyses were conducted for the physics-enhanced features, physical derivative constraints, and heteroscedastic noise modeling, with the detailed results provided in Table S4. As shown in Fig. 5, the Standard GP produces an almost flat response surface under the sparse Round-1 initialization, indicating severe underfitting and limited sensitivity to changes in target power and substrate bias (Fig. 5a). The unconstrained neural network avoids complete surface collapse but generates a response topology that is not consistently aligned with the expected physical dependence of hardness on the process variables, illustrating that a flexible neural network alone cannot reliably recover physically meaningful trends from only five training samples (Fig. 5c).

Introducing the physical derivative constraints produces a smoother response surface with a more physically reasonable dependence on target power and substrate bias, whereas heteroscedastic modeling improves robustness to sample-dependent measurement noise. Fig. 5f demonstrates the profound impact of physics-informed inductive biases. The PR-HNN model, regularized by the full-dimensional partial derivative penalties ($\mathcal{L}_{phys}$), reconstructs a smooth, highly rational, and physically consistent topology even in data-scarce regions. Along the CoCrFeNi target power axis, the surface adheres to the monotonic increase dictated by the Monotonic Kinetic Constraint (Φ_1), accurately reflecting the densification driven by enhanced plasma ionization and sputtering yield.

The results show that both the physical derivative constraints and heteroscedastic modeling independently improved the predictive performance of the neural network under small-sample conditions. When jointly incorporated into the complete PR-HNN, these components yielded the lowest mean and pooled RMSE values across the three rounds. This finding indicates that physics-based regularization and data-driven uncertainty modeling play complementary roles in improving predictive accuracy, stability, and physical consistency.

Overall, the validations demonstrate that the PR-HNN maintained a consistent predictive advantage throughout successive data augmentation and model updates. The ablation results further indicate that the synergistic integration of physical derivative constraints and heteroscedastic noise modeling enables the model to generate more accurate, stable, and physically consistent predictions within an extremely small-sample and highly coupled process space. This capability provides a reliable surrogate-model foundation for efficient and risk-aware parameter exploration in the subsequent active-learning closed loop.

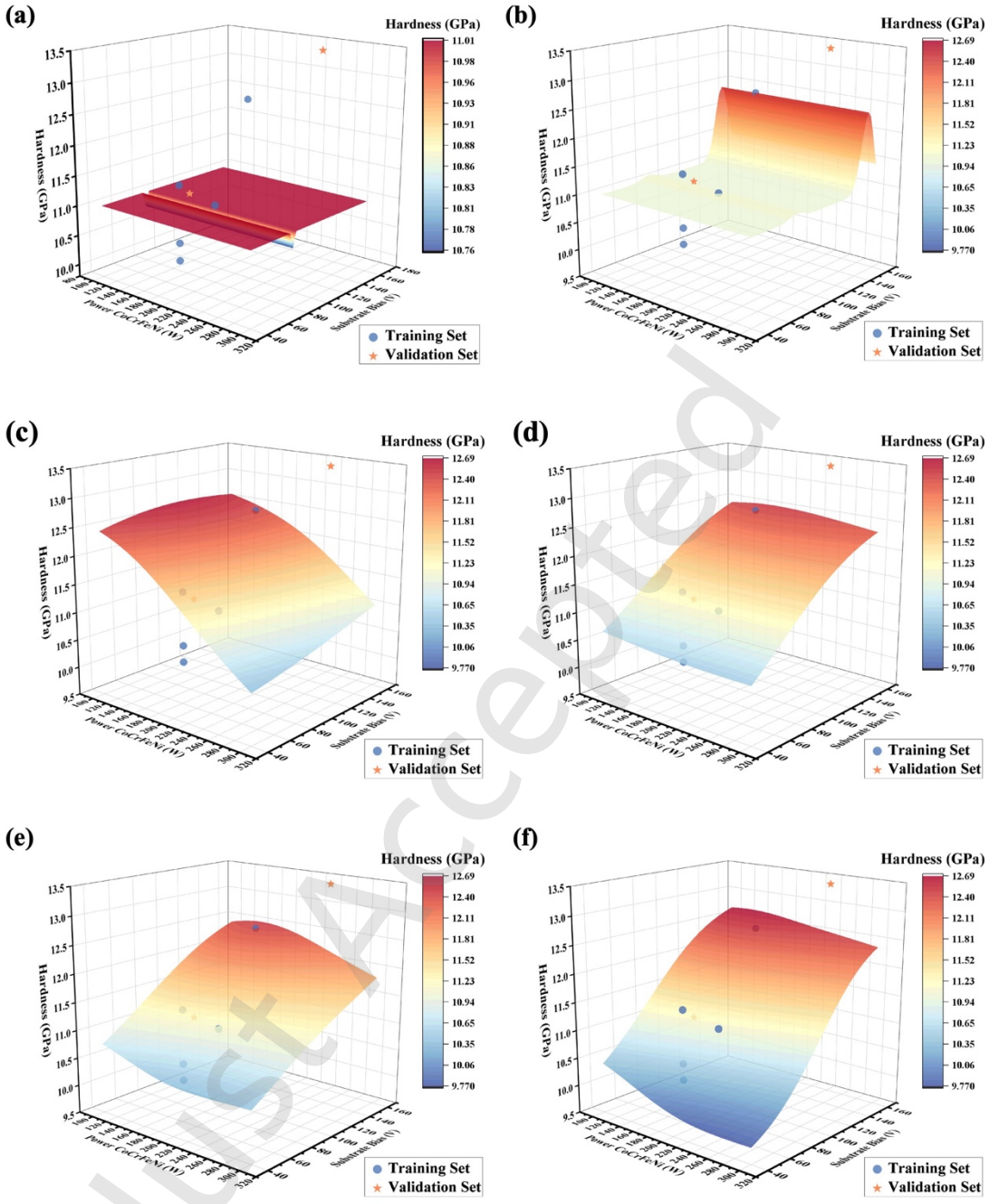

Fig. 5 Ablation Analysis of Hardness Prediction Surfaces under Sparse-Data Conditions. Three-dimensional response surfaces of coating hardness with respect to CoCrFeNi target power and substrate bias predicted by: (a) Standard GP with raw five-dimensional inputs; (b) GP with eight-dimensional physics-enhanced features; (c) neural network without physical constraints or heteroscedastic modeling; (d) neural network with physical derivative constraints; (e) neural network with heteroscedastic modeling; and (f) PR-HNN

3.3 Multi-Objective Optimization and Performance Enhancement

The ultimate goal of the PIAL framework is to transcend the inherent trade-offs between mechanical, tribological, and electrochemical properties in HEA multilayer films. By analyzing the experimental feedback from three PIAL iterations, a clear and highly efficient transition from heuristic space exploration to targeted Pareto optimization is observed.

Fig. 6a illustrates the evolution of the multi-objective Pareto fronts for Hardness versus Wear Depth (Fig. 6a-I) and Hardness versus Corrosion Resistance (Fig. 6a-II). The initial baseline samples generated are widely dispersed and generally exhibit sub-optimal performance, reflecting the high epistemic uncertainty of the unoptimized 5D parameter space. In contrast, the three successive dual-point batches recommended by the PIAL acquisition function progressively expanded and refined the initial performance boundaries. The first-round candidates identified a high-hardness and corrosion-resistant region, while the subsequent rounds further explored alternative mechanical–tribological trade-offs and extended the Pareto front toward lower wear depth.

Specifically, Sample 7, located at the zenith of the new Pareto front, achieved a peak hardness of approximately 13.4 GPa, which is a significant improvement over the best initial sample (Fig. S1). Crucially, as explicitly constrained by the PoF in the algorithm, this profound hardening did not induce catastrophic embrittlement. As shown in Fig. 6a-I, Sample 7 exhibited a wear depth of 0.30 μm, corresponding to a 25% reduction relative to the best initial wear depth of 0.40 μm. Moreover, the third-round Sample 11 further reduced the wear depth to 0.25 μm, extending the mechanical Pareto front toward the minimum-wear region while retaining a hardness of approximately 11.1 GPa. This significant tribological improvement is visually corroborated by the optical micrographs in Fig. 6b, which reveal a noticeably narrower and smoother wear track for the PIAL-optimized coating.

Simultaneously, the sample maintained a robust corrosion resistance (Fig. 6a-II). The multi-round results therefore demonstrate that the PIAL framework can progressively navigate the highly coupled parameter space and identify multiple Pareto-efficient solutions with distinct performance trade-offs, rather than converging to a single property-specific candidate (Fig. S2). This confirms that

the physics-informed constraints successfully steered the model away from recommending mathematically extreme but physically fragile parameters, ultimately yielding coatings with improved and balanced mechanical, tribological, and electrochemical performance.

Finally, to evaluate the comprehensive multi-functional enhancement, a normalized radar chart was constructed comparing the initial best performer with the PIAL-optimized sample 7 Fig. 6c. The axes encompass Hardness, Lubricity, Wear Resistance, and Corrosion Resistance. The radar profile of the PIAL-optimized sample (red area) entirely encapsulates that of the initial best (blue area), indicating a unidirectional expansion of the performance envelope. By maintaining excellent electrochemical stability while drastically boosting mechanical and tribological traits, the PIAL framework achieved a comprehensive high-performance HEA coating within three closed-loop iterations, while the additional Pareto-efficient samples identified in the subsequent rounds further demonstrated the framework's capacity to explore diverse performance trade-offs under a limited experimental budget.

Table S5 further compares the PIAL-optimized coating with representative CoCrFeNiTi-based coatings reported in the literature, demonstrating its favorable overall balance of hardness, friction, wear, and corrosion performance [43–47].

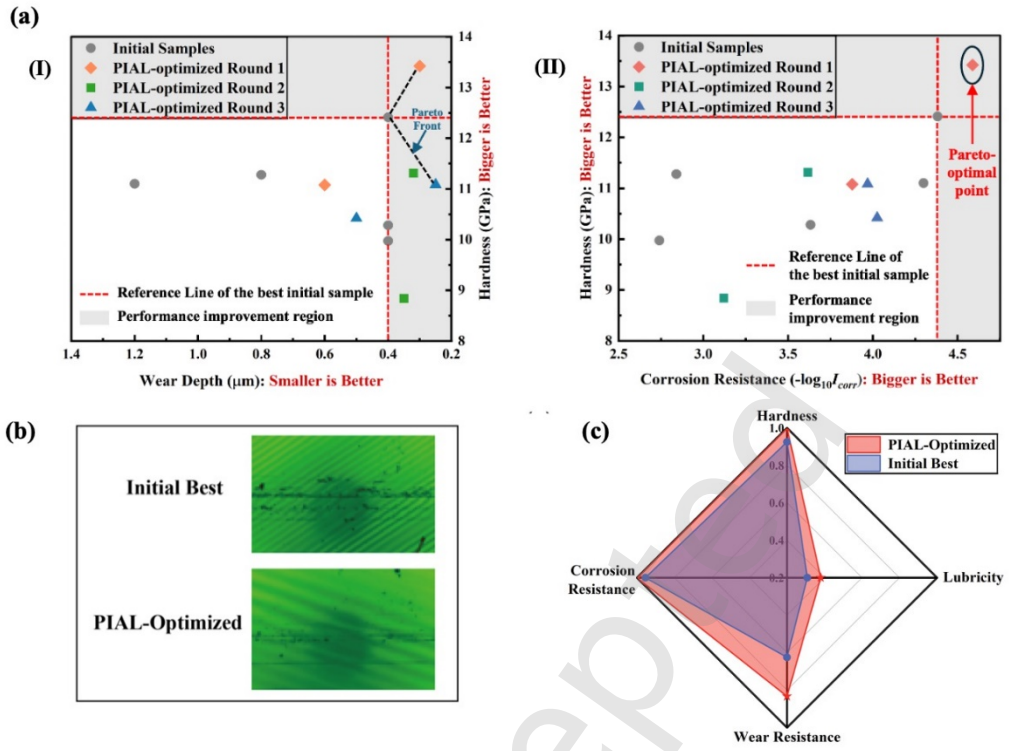

Fig. 6 Multi-Objective Optimization and Comprehensive Performance Enhancement of PIAL-Engineered Coatings. (a) Evolution of the multi-objective Pareto fronts: (I) hardness versus wear depth; (II) hardness versus corrosion resistance, demonstrating the superior performance of the PIAL-optimized samples and highlighting the optimal Pareto-efficient solution. (b) Optical micrographs of the wear tracks after reciprocating sliding tests, visually comparing the tribological damage morphology between the initial best sample and the PIAL-optimized sample. (c) Normalized radar chart illustrating the comprehensive multi-functional enhancement of the PIAL-optimized coating over the initial best performer across hardness, lubricity, wear resistance, and corrosion resistance.

4 Conclusion and future work

To address the high-dimensional, multi-objective optimization challenges and inherent physical trade-offs in the PVD of CoCrFeNiTi HEA multilayer films under extremely small-sample constraints, this study proposes a closed-loop PIAL framework. The framework synergistically integrates physics-enhanced feature engineering, a PR-HNN surrogate, and a risk-aware dual-point batch acquisition strategy. By embedding fundamental sputtering kinetics and metallurgical principles into the machine learning objective function via partial derivative penalties, the PR-HNN

effectively suppresses nonphysical fluctuations and consistently outperforms both the Standard GP and its physics-feature-enhanced counterpart across three iterative validation rounds, demonstrating that the model maintains stable predictive performance as the experimental dataset is progressively expanded.

Guided by the progressively updated surrogate models, the PIAL framework expanded the initial Pareto front over three closed-loop iterations and identified multiple Pareto-efficient coatings with different performance trade-offs. The best observed hardness reached approximately 13.4 GPa, while another candidate achieved a minimum wear depth of 0.25 μm. These results provide an initial experimental demonstration that integrating soft physical priors with active learning can improve the efficiency of process exploration within the investigated material system and equipment-safe process window. Fundamentally, the proposed PIAL architecture is established as a generalized, modular framework. Its underlying derivative-level regularization engine allows the governing equations to be flexibly substituted, modified, or expanded to accommodate a broader range of material systems and advanced manufacturing scenarios.

Building upon this algorithmic cornerstone, our future work aims to further scale up the physical experimental workflow and accelerate coating development. Specifically, research will focus on two practical directions. First, as the present study primarily provides an initial validation of the feasibility and effectiveness of the proposed framework, future work will involve additional rounds of active-learning experiments and extend the investigation to broader process windows and more complex coating architectures. Second, future work will explore a simulation-to-reality transfer optimization paradigm. By incorporating computational simulations to generate large-scale virtual datasets, the algorithm can efficiently map the global physical trends of the parameter space. Subsequently, guided by active learning, targeted physical experiments can be used to calibrate the surrogate model, reduce the simulation-to-reality discrepancy, and progressively identify high-performance process regions. Ultimately, this methodology provides a promising, interpretable, and

sample-efficient foundation for the accelerated inverse design of next-generation advanced functional coatings.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (5257050242), Natural Science Foundation of Shanghai(25ZR1402227) and the Industry-University-Research Cooperation Fund of the Eighth Academy of China Aerospace Science and Technology Corporation (USCAST2023-20).

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Electronic Supplementary Material (ESM)

Supplementary material (including the systematic summary of PVD process variables for the initial training and active learning phases, the comparison of nanoindentation hardness between empirical and PIAL optimized coatings, and the evaluation of corrosion resistance via Tafel extrapolation analysis) is provided.

References

[1] Chen R, Zhang H, Ma X, Chai Y, Bi H, Ma H, Zhu N, Ji J, Xue M. Two-dimensional reduced graphene oxide/polypyrrloe-based coating enable superior corrosion protection and photothermal-induced in-situ internal environmental regulation. *Chemical Engineering Journal* **458**: 141481 (2023)

[2] Fotovvati B, Namdari N, Dehghanghadikolaei A. On coating techniques for surface protection: A review. *Journal of Manufacturing and Materials processing* **3**(1): 28 (2019)

[3] Kalscheuer C, Bobzin K, Liu X. Determination of mechanical properties of physical vapor deposition tool coatings using machine learning. *Wear* 206296 (2025)

[4] Le T C, Nguyen D-N, Kaminski D, Kolver T, Subramanian P, Bateman S. Harnessing Machine Learning for the Design of Surface Coatings: Challenges and Opportunities. *ACS Applied Materials & Interfaces* **17**(28): 39795–39808 (2025)

[5] Bobzin K, Kalscheuer C, Carlet M, Janowitz J. Machine learning based model for plasma prediction in high power pulsed magnetron sputtering processes. *Thin Solid Films* **777**: 139903 (2023)
[6] Santo D, Castro J D, Cruz S, Carvalho I, Cavaleiro A, Carvalho S. Customisation of PVD coatings for biomedical devices. *Surface and Coatings Technology* **512**: 132277 (2025)
[7] Bait L, Azzouz L, Madaoui N, Saoula N. Influence of substrate bias voltage on the properties of TiO2 deposited by radio-frequency magnetron sputtering on 304L for biomaterials applications. *Applied Surface Science* **395**: 72–77 (2017)
[8] Seynstahl A, Köbrich M, Rosnitschek T, Göken M, Tremmel S. Enhancing the lifetime and vacuum tribological performance of PVD-MoS2 coatings by nitrogen modification. *Surface and Coatings Technology* **477**: 130343 (2024)
[9] El-Hossary F, El-Rahman A A, Raaif M, El-Kassem M A. Physical, electrochemical, and biocompatibility characteristics of Ti-Al-N thin film synthesized by DC pulsed magnetron sputtering. *Journal of the Australian Ceramic Society* **56**(3): 1155–1165 (2020)
[10] Che J, Yi P, Peng L, Lai X. Impact of pressure on carbon films by PECVD toward high deposition rates and high stability as metallic bipolar plate for PEMFCs. *International Journal of Hydrogen Energy* **45**(32): 16277–16286 (2020)
[11] Wang J, Liang M, Pan Y, Sun S, Shen T, Wei X, Zhu Y, Liu J, Huang Q. Control of surface composition and microstructure of nano super-hydrophilic TiO2-CuOy coatings through reactive sputtering to improve antibacterial ability, corrosion resistance, and biocompatibility. *Applied Surface Science* **578**: 151893 (2022)
[12] Bobzin K, Brögelmann T, Kruppe N, Arghavani M, Engels M. Correlation of HPPMS plasma and coating properties using artificial neural networks. *Surface and Coatings Technology* **349**: 1130–1136 (2018)
[13] Lee J A, Park J, Sagong M J, Ahn S Y, Cho J-W, Lee S, Kim H S. Active learning framework to optimize process parameters for additive-manufactured Ti-6Al-4V with high strength and ductility. *Nature Communications* **16**(1): 931 (2025)
[14] Zhang D, Yi P, Lai X, Peng L, Li H. Active machine learning model for the dynamic simulation and growth mechanisms of carbon on metal surface. *Nature Communications* **15**(1): 344 (2024)
[15] Xuan W, Ye Z, Han D, Shi J, Chen J, Kang J. Multiscale simulation of physical vapor deposition. *Materials Science and Engineering: B* **295**: 116596 (2023)
[16] Mahne N, Čekada M, Panjan M. Total and differential sputtering yields explored by SRIM simulations. *Coatings* **12**(10): 1541 (2022)
[17] Ohkubo I, Hou Z, Lee J, Aizawa T, Lippmaa M, Chikyow T, Tsuda K, Mori T. Realization of closed-loop optimization of epitaxial titanium nitride thin-film growth via machine learning. *Materials Today Physics* **16**: 100296 (2021)
[18] Yang Y, Pei L, Zhang H, Feng K, Ju P, Duan W, Ji L, Li H, Liu X, Zhou H, others. Magnetron sputtering NbSe2 film as lubricant for space current-carrying sliding contact. *Friction* **11**(3): 383–394 (2023)
[19] Ayyagari S V G, Daware S V, Rai B. Data-driven approach to optimize composition and process parameters of hydrophobic coating formulation. *Surfaces and Interfaces* **23**: 100972 (2021)
[20] Han X-Q, Wang X-D, Xu M-Y, Feng Z, Yao B-W, Guo P-J, Gao Z-F, Lu Z-Y. AI-driven inverse design of materials: Past, present, and future. *Chinese Physics Letters* **42**(2): 027403 (2025)
[21] Lee J, Park D, Lee M, Lee H, Park K, Lee I, Ryu S. Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: a review. *Materials horizons* **10**(12): 5436–5456 (2023)
[22] Cao B, Adutwum L A, Oliynyk A O, Luber E J, Olsen B C, Mar A, Buriak J M. How to optimize materials and devices via design of experiments and machine learning: Demonstration using organic photovoltaics. *ACS nano* **12**(8): 7434–7444 (2018)
[23] Ning Z, Zhao X, Fan L, Peng Z, Ma F, Jin Z, Deng J, Duan J, Hou B. Application of artificial

neural networks throughout the entire life cycle of coatings: A comprehensive review. *Progress in Organic Coatings* **189**: 108279 (2024)
[24] Fernández-López P, Alves S A, Rogov A, Yerokhin A, Quintana I, Duo A, Aguirre-Ortuzar A. Data-driven optimization of plasma electrolytic oxidation (PEO) coatings with explainable artificial intelligence insights. *Coatings* **14**(8): 979 (2024)
[25] Xu P, Chen C, Chen S, Lu W, Qian Q, Zeng Y. Machine learning-assisted design of yttria-stabilized zirconia thermal barrier coatings with high bonding strength. *ACS omega* **7**(24): 21052–21061 (2022)
[26] Zhang X, Jia B, Zeng Z, Zeng X, Wan Q, Pogrebnjak A, Zhang J, Pelenovich V, Yang B. Machine learning-based design of superhard high-entropy nitride coatings. *ACS Applied Materials & Interfaces* **16**(28): 36911–36922 (2024)
[27] Pruksawan S, Lambard G, Samitsu S, Sodeyama K, Naito M. Prediction and optimization of epoxy adhesive strength from a small dataset through active learning. *Science and technology of advanced materials* **20**(1): 1010–1021 (2019)
[28] Sheng Y, Wu Y, Yang J, Lu W, Villars P, Zhang W. Active learning for the power factor prediction in diamond-like thermoelectric materials. *npj Computational Materials* **6**(1): 171 (2020)
[29] Pogue E A, New A, McElroy K, Le N Q, Pekala M J, McCue I, Gienger E, Domenico J, Hedrick E, McQueen T M, others. Closed-loop superconducting materials discovery. *npj Computational Materials* **9**(1): 181 (2023)
[30] Zhang Z, Yin N, Chen S, Liu C. Tribo-informatics: Concept, architecture, and case study. *Friction* **9**(3): 642–655 (2021)
[31] Zhang Z, Yin N. *Tribo-informatics: the systematic fusion of AI and tribology*. CRC Press. 2024 (2024)
[32] Yin N, Xing Z, He K, Zhang Z. Tribo-informatics approaches in tribology research: A review. *Friction* **11**(1): 1–22 (2023)
[33] Yin N, Hou Z, Wang X, Zhang S, Zhang Z. Tribo-informatics empowered research on triboelectrification. *Friction* **13**(2): 9441045 (2025)
[34] Vaz F, Machado P, Rebouta L, Cerqueira P, Goudeau P, Riviere J, Alves E, Pischow K, De Rijk J. Mechanical characterization of reactively magnetron-sputtered TiN films. *Surface and Coatings Technology* **174**: 375–382 (2003)
[35] Thornton J A. High rate thick film growth. *Annual review of materials science* **7**(1): 239–260 (1977)
[36] Hansen N. Hall–Petch relation and boundary strengthening. *Scripta materialia* **51**(8): 801–806 (2004)
[37] Ebrahimi M, Luo B, Wang Q, Attarilar S. High-performance nanoscale metallic multilayer composites: Techniques, mechanical properties and applications. *Materials* **17**(9): 2124 (2024)
[38] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics* **378**: 686–707 (2019)
[39] Hou Z, Wu Z, Li Z, Zhang Z. Bridging Physics-Informed Neural Networks and the Stribeck Curve: A Tribo-informatics Approach to Multi-Regime Lubrication Analysis. *Tribology International* 111702 (2026)
[40] Windischmann H. Intrinsic stress in sputtered thin films. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* **9**(4): 2431–2436 (1991)
[41] Clemens B, Kung H, Barnett S. Structure and strength of multilayers. *Mrs Bulletin* **24**(2): 20–26 (1999)
[42] Zhan D, Xing H. Expected improvement for expensive optimization: a review. *Journal of Global Optimization* **78**(3): 507–544 (2020)
[43] Zhao L, Wang Z, Liu Y, Jiang W, Bao F, Xing Z, Yang S. Construction and protection

mechanism of CoCrFeNiTi/Ti amorphous–nanocrystalline multilayer coatings fabricated by magnetron sputtering. *Journal of Alloys and Compounds* 183135 (2025)
[44] Manea C A, Geambazu L E, Tălpeanu D, Marinescu V, Sbârcea G B, Pătroi D, Udrea R M, Lungu M V, Lucaci M. CoxCrFeNiTi High-Entropy Alloys Prepared via Mechanical Alloying and Spark Plasma Sintering for Magnetron Sputtering Coatings. *Materials* **16**(19): 6386 (2023)
[45] Liu H, Gao W, Liu J, Du X, Li X, Yang H. Microstructure and properties of CoCrFeNiTi high-entropy alloy coating fabricated by laser cladding. *Journal of Materials Engineering and Performance* **29**(11): 7170–7178 (2020)
[46] Wang J, Man J, Wang Q, Liu G, Xiao S, Dong N. Effect of Nano WC on wear and corrosion resistances of CoCrFeNiTi high entropy alloy coating. *Journal of Materials Engineering and Performance* **34**(2): 1078–1087 (2025)
[47] Liu H, Xu Q, Wang L, Chen P, Liu X, Gao Q, Hao J, Yang H. Microstructure and wide temperature range tribological properties of CoCrFeNiTi high-entropy alloy coating by laser cladding. *Journal of Materials Engineering and Performance* **34**(2): 1515–1525 (2025)

Wanyu Wang. She received her B.S. degrees in French and Mechanical Engineering from Shanghai Jiao Tong University. She is currently pursuing her M.S. degree in Mechanical Engineering at Shanghai Jiao Tong University. Her research focuses on the inverse design of materials and tribo-informatics.

Zhinan Zhang. He received his Ph.D. degree in 2011 from Shanghai Jiao Tong University. He is a professor at the School of Mechanical Engineering, Shanghai Jiao Tong University. His research interests include computational design and analysis of tribosystems, tribo-informatics, and theory and methods of design engineering and innovation.