

ORIGINAL RESEARCH

Open access

Cross-Property Entanglement in Multi-Task Materials Learning Systems

Daniel Fischer^{1*}, Laura Meier¹, Thomas Braun², Stefan Koch²

Abstract

In the evolving landscape of computational and data-driven materials engineering, multi-task learning systems have emerged as pivotal infrastructures for accelerating discovery pipelines. These systems leverage shared representations across diverse material properties to enhance predictive accuracy and efficiency in high-dimensional spaces. However, a critical yet underexplored aspect is the entanglement of properties within these models, where interdependencies among physical, chemical, and structural attributes create emergent behaviors that influence overall system dynamics. This manuscript introduces a novel conceptual framework, the Property Entanglement Lattice (PEL), which interprets cross-property interactions as lattice-like structures facilitating integrated inference and discovery steering. By synthesizing recent advancements in materials informatics, machine learning architectures, and representation learning, we delineate how entanglement manifests in multimodal datasets and foundation models, impacting uncertainty quantification and simulation-experiment coupling. The framework elucidates computational workflows that harness entanglement for optimized resource allocation in autonomous systems, without relying on empirical validations. Implications extend to inverse design paradigms, where entangled representations enable more robust epistemic navigation in materials ecosystems. This work provides a systems-level lens for researchers to conceptualize trade-offs in multi-task setups, fostering infrastructural innovations in computational materials science. Ultimately, it positions cross-property entanglement as a core logic for advancing data-driven discovery, balancing technical depth with interpretive insights.

Keywords Materials informatics, Representation learning, Multi-task learning, Discovery pipelines, Property entanglement, Computational workflows

*Correspondence:

Daniel Fischer
daniel.fischer@gmail.com

¹ Department of Materials Simulation and Data Engineering, Faculty of Engineering, University of Freiburg, Freiburg, Germany

² Department of Computational Materials Systems, Faculty of Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany

Introduction

The paradigm shift in materials engineering

Materials engineering is undergoing a systemic transformation driven by the convergence of computational modeling, high-throughput infrastructures, and data-driven intelligence. Historically, materials discovery relied on empiricism—iterative synthesis, characterization, and refinement governed by domain intuition and incremental experimentation. While this paradigm yielded foundational advances, it constrained exploration to narrow compositional and structural neighborhoods due to cost, time, and instrumentation limitations. The emergence of computational materials science began to expand this search space, but it is the integration of artificial intelligence and informatics that has fundamentally reconfigured discovery logics.

This paradigm shift is catalyzed by exponential growth in computational capacity, cloud-scale storage, and algorithmic sophistication. High-throughput density functional theory, molecular dynamics simulations, and automated experimental platforms now generate vast multimodal datasets encompassing structural, electronic, thermodynamic, and kinetic descriptors. These infrastructures transform materials science from a data-scarce discipline into an information-rich ecosystem capable of supporting predictive analytics at scale.

At the center of this transition lies materials informatics, an interdisciplinary field that applies information science principles—data structuring, pattern recognition, and knowledge extraction—to materials discovery and design [1, 2]. By reconceptualizing materials datasets as structured epistemic repositories rather than isolated measurements, informatics frameworks enable the detection of latent structure–property relationships across chemical spaces. This shift supports accelerated innovation in energy storage systems, catalytic platforms, quantum materials, and advanced manufacturing alloys, where combinatorial complexity previously hindered systematic exploration.

Crucially, materials informatics does not merely accelerate existing workflows; it redefines the ontology of discovery itself. Computational pipelines now function as epistemic engines—systems that not only generate predictions but shape what is knowable within the materials design landscape.

Machine learning as an infrastructural driver

Machine learning (ML) has become the algorithmic backbone of this transformation, translating high-dimensional materials data into predictive and generative insights [3, 4]. However, ML deployment in materials engineering diverges from conventional data science contexts. Unlike consumer or financial datasets, materials data are sparse, noisy, multimodal, and constrained by physical laws governing atomic interactions.

Early machine learning applications focused on single-property prediction tasks—bandgap estimation, formation energy calculation, elastic modulus forecasting—demonstrating that statistical models could approximate quantum-mechanical outputs at reduced computational cost. While impactful, such single-task frameworks treated materials properties as independent targets, neglecting their intrinsic physicochemical coupling.

Recent advances foreground multi-task learning (MTL) as a more faithful modeling paradigm [5, 6]. By predicting multiple interdependent properties simultaneously, MTL architectures leverage shared representations that reflect the interconnected nature of materials behavior. Thermal conductivity, electrical resistivity, mechanical stiffness, and defect energetics, for example, often arise from shared electronic and structural determinants. Joint optimization therefore enhances both predictive performance and physical coherence.

Beyond efficiency gains, MTL introduces a conceptual reorientation: predictive systems begin to approximate the relational topology of materials properties rather than isolated observables. This relational modeling forms a precursor to the entanglement perspective advanced in this manuscript.

Challenges in data-driven ecosystems

Despite rapid methodological progress, data-driven materials engineering faces structural challenges rooted in ecosystem complexity [7, 8]. Contemporary discovery platforms integrate heterogeneous data modalities—crystallographic structures, spectroscopic signatures, microscopy images, thermodynamic simulations, and synthesis metadata—each governed by distinct noise regimes and measurement biases [9, 10].

High-throughput computational infrastructures generate data at unprecedented scale, yet quantity does not guarantee epistemic reliability. Variations in simulation parameters, convergence thresholds, and pseudopotential choices introduce latent biases that propagate into machine learning models [11, 12]. When such datasets train predictive systems, encoded assumptions may remain unarticulated, shaping inference trajectories in opaque ways.

Coupling computational predictions with experimental validation introduces additional friction. Discrepancies between simulated and observed properties—arising from idealized boundary conditions or measurement artifacts—can destabilize closed discovery loops if not systematically reconciled [13, 14]. These translation gaps highlight the infrastructural challenge of harmonizing virtual and physical knowledge systems.

Uncertainty and autonomous discovery pressures

Uncertainty quantification emerges as a critical stabilizing mechanism within these ecosystems [15, 16]. Predictive accuracy alone is insufficient when models guide costly synthesis or safety-critical materials deployment. Decision systems require calibrated confidence estimates that distinguish robust predictions from speculative extrapolations.

Traditional statistical uncertainty methods struggle to capture epistemic gaps intrinsic to underexplored chemical spaces [17, 18]. Advanced probabilistic architectures—Bayesian deep learning, ensemble modeling, conformal prediction—address this limitation by embedding uncertainty directly within inference pipelines.

The stakes of uncertainty intensify within autonomous discovery systems, where machine learning models iteratively propose candidates, trigger synthesis, assimilate results, and retrain in closed feedback loops [19, 20]. In such infrastructures, uncertainty is not merely diagnostic—it is operational. It directs exploration, allocates resources, and governs risk exposure.

As autonomy scales, so too does the need for interpretive frameworks capable of diagnosing how predictive signals propagate across interconnected properties and decision layers.

The role of representation learning

Representation learning provides the computational substrate enabling these predictive ecosystems [21, 22]. By transforming raw materials descriptors into structured latent embeddings, representation systems allow machine learning architectures to operate across atomic, molecular, and microstructural scales.

Graph neural networks (GNNs) exemplify this paradigm, encoding materials as relational graphs that preserve bonding topologies and geometric symmetries [23, 24]. Through message-passing operations, these models learn hierarchical feature abstractions transferable across materials classes.

Complementary architectures extend representational reach. Stoichiometry-based embeddings infer properties directly from elemental compositions, enabling rapid screening of hypothetical compounds [25, 26]. Meanwhile, scientific foundation models pretrained on multimodal corpora introduce few-shot adaptability, extending predictive capabilities into data-scarce regimes [27, 28].

Yet representation learning introduces an underexamined phenomenon: latent property entanglement. Shared embedding dimensions frequently encode multiple physicochemical attributes simultaneously—electronic, mechanical, thermodynamic—producing coupled inference pathways [29, 30]. While such entanglement enhances generalization and multi-objective optimization, it can also amplify correlated prediction errors or obscure mechanistic interpretability.

Current literature addresses these representational dynamics in fragmented ways, examining architecture performance or embedding transferability without a unified interpretive lens for cross-property coupling.

Positioning the Property Entanglement Lattice (PEL)

This manuscript addresses the conceptual gap by introducing the Property Entanglement Lattice (PEL)—a systems-level interpretive framework that reconceptualizes cross-property interactions as lattice structures embedded within multi-task learning architectures.

Rather than treating entanglement as an incidental byproduct of shared representations, PEL frames it as an organizing topology governing knowledge propagation across predictive systems. Within this lattice:

- Nodes represent property domains or task embeddings.
- Edges encode entanglement strength and directional influence.
- Lattice density reflects representational coupling intensity.
- Structural distortions signal epistemic risk or bias amplification.

By mapping these interactions, the framework provides an infrastructural lens for interpreting how computational workflows steer discovery trajectories. Importantly, PEL operates as a conceptual scaffold rather than an empirical model—designed to guide interpretive diagnostics, architectural design choices, and epistemic risk assessment within data-driven materials ecosystems shown in [table 1](#).

Table 1. Property entanglement regimes in multi-task materials learning and their infrastructural consequences

Entanglement regime (PEL view)	Lattice signature (conceptual)	Typical architectural driver	Primary benefit (epistemic gain)	Primary risk (epistemic cost)	Steering implication (what to do)
Sparse coupling	Few edges; fragmented clusters	Independent heads; weak sharing	Clean separability; localized interpretability	Missed cross-property signal; redundant learning	Add selective bridges between physically linked properties
Structured coupling	Modular lattice; communities	Soft sharing; task grouping	Transfer without collapse; stable propagation	Boundary errors between modules	Use module-aware losses and diagnostics on inter-module edges
Dense coupling	High edge density; short paths	Hard sharing; strong shared trunk	Efficiency; strong generalization in small data	Conflation; correlated error amplification	Introduce disentanglement constraints and uncertainty gating
Asymmetric coupling	Directed influence patterns	Imbalanced data; dominant tasks	Strong supervision from “anchor” tasks	Suppression of minority tasks; bias transfer	Reweight tasks; protect minority outputs via regularization
Dynamic coupling	Edge weights shift over loops	Closed-loop retraining; streaming data	Adaptation; continuous epistemic refinement	Drift-induced instability; oscillations	Add stability checks + loop damping via conservative updates
Pathological coupling	Orange dashed edges dominate	Dataset bias; modality dominance	Apparent multi-task “boost”	Spurious correlations; unsafe extrapolation	Trigger risk gates, recalibration, or lattice pruning

In positioning entanglement as a structural property of computational discovery systems, this work advances a broader thesis: that the future of materials AI depends not only on predictive power but on interpretive architectures capable of revealing how knowledge is co-constructed across tasks, representations, and infrastructures.

Theoretical Background & Literature Synthesis

Foundations of multi-task learning in materials

Multi-task learning (MTL) in materials science emerges from the recognition that materials properties rarely exist in isolation; instead, they are co-manifestations of shared physicochemical drivers such as bonding environments, electronic structures, thermodynamic stability fields, and microstructural configurations. By learning these interdependencies simultaneously, MTL architectures leverage shared informational substrates to enhance predictive fidelity across multiple outputs [1, 5]. This paradigm departs from single-task modeling, where each property is predicted independently, often resulting in redundant feature learning and increased susceptibility to overfitting—particularly in small data regimes characteristic of experimental materials datasets [2, 6].

In neural implementations, parameter sharing across hidden layers enables the extraction of generalized structure–property embeddings, allowing correlated tasks—such as bandgap prediction, formation energy estimation, and elastic property forecasting—to mutually reinforce learning signals [3, 7]. Hard parameter sharing reduces model variance by constraining representational degrees of freedom, whereas soft sharing frameworks retain task-specific subnetworks linked through regularization penalties, balancing generalization with specialization [4, 8]. This architectural diversity reflects an epistemic trade-off between cross-task coherence and property-specific sensitivity.

Transfer learning further extends the MTL paradigm by enabling cross-domain knowledge migration. Pretrained models on large computational repositories—such as density functional theory (DFT) datasets—can be fine-tuned for scarce experimental tasks, effectively transferring learned physical priors into underrepresented discovery domains [4, 8]. This is particularly consequential for emerging materials classes where experimental data acquisition remains resource-intensive.

The small-data challenge remains central to materials informatics. Active learning strategies address this by iteratively selecting high-value data points for labeling or simulation, guided by uncertainty estimates or acquisition functions [9, 13]. When embedded within high-throughput pipelines, these adaptive sampling strategies prioritize regions of chemical space with maximal informational gain, accelerating discovery while minimizing computational expenditure [10, 14]. Importantly, such systems do not merely predict properties; they shape experimental trajectories, functioning as epistemic steering mechanisms.

Consequently, the literature increasingly frames MTL not only as a predictive tool but as an interpretive infrastructure capable of elucidating property co-dependencies. By modeling covariance structures across outputs, these systems offer insight into latent physicochemical couplings, thereby informing theoretical understanding alongside predictive performance [11, 15].

Representation and graph-based architectures

Representation learning constitutes the computational substrate upon which modern materials AI systems operate. Transforming atomic configurations, crystallographic symmetries, and compositional descriptors into machine-interpretable embeddings enables the deployment of deep learning across diverse materials domains [21, 23]. The representational challenge lies in encoding invariances—rotational, translational, and permutational—while preserving chemically meaningful interactions.

Graph neural networks (GNNs) have emerged as dominant architectures in this space, modeling materials as relational graphs in which atoms function as nodes and interatomic interactions as edges [22, 24]. Through iterative message-passing operations, local atomic environments propagate information across graph structures, yielding hierarchical embeddings that capture both short-range bonding and long-range structural order. This enables unified modeling across molecular crystals, extended solids, and amorphous systems.

Crucially, GNNs operationalize physics-consistent learning by embedding geometric constraints and interaction kernels into aggregation functions, facilitating property prediction across scales—from quantum observables to macroscopic mechanical responses [25, 27]. Their adaptability has catalyzed widespread adoption in high-throughput screening, catalysis discovery, and functional materials design.

Recent deep learning extensions integrate multimodal inputs, fusing spectroscopic data, microscopy imaging, simulation outputs, and compositional metadata into unified embedding spaces [26, 28]. Such multimodal fusion enhances representational richness, enabling models to reconcile discrepancies between experimental observations and computational approximations.

Stoichiometry-driven architectures offer an alternative representational paradigm by predicting properties directly from elemental compositions without explicit structural inputs [6, 12]. While computationally efficient, these models encode statistical correlations rather than mechanistic interactions, raising interpretability constraints.

A critical epistemic feature of these representation systems is property entanglement. Latent embedding dimensions frequently encode multiple physicochemical attributes simultaneously—e.g., electronic and thermodynamic signatures—rendering disentanglement non-trivial [16, 18]. While such entanglement supports multi-objective optimization, it complicates interpretive extraction, limiting targeted design strategies.

The literature thus reveals a trajectory toward transferable, foundation-scale embeddings pretrained across heterogeneous datasets. Yet, interpretive transparency lags behind representational sophistication, underscoring the need for frameworks capable of disentangling latent property couplings for domain-specific deployment [17, 19].

Uncertainty and epistemic structures

Uncertainty quantification (UQ) embeds probabilistic reasoning within materials AI systems, addressing both data variability and model incompleteness [15, 20]. Two principal uncertainty classes dominate: aleatoric uncertainty, arising from measurement noise and intrinsic stochasticity, and epistemic uncertainty, reflecting knowledge gaps due to limited training coverage.

Bayesian neural networks operationalize UQ by learning parameter distributions rather than deterministic weights, enabling posterior predictive intervals. Ensemble methods approximate similar effects through model diversity, aggregating predictions across independently trained networks [4, 10]. Both approaches inform risk-aware decision-making within discovery pipelines.

In multi-task environments, uncertainty assumes joint structural forms. Prediction confidence for one property may condition reliability in another, necessitating covariance-aware modeling of task distributions [7, 11]. Such joint uncertainty landscapes are critical for multi-objective optimization, where trade-offs between performance metrics must be navigated probabilistically.

Epistemic risks intensify when models extrapolate beyond training domains—common in exploratory materials discovery. Domain applicability assessments address this by mapping prediction reliability across chemical space, identifying out-of-distribution regimes [3, 9]. Techniques such as distance-aware embeddings and conformal prediction frameworks enhance reliability diagnostics.

From an infrastructural perspective, uncertainty functions as a steering signal rather than a passive metric. High-uncertainty regions attract targeted simulations or experiments, transforming UQ into an active exploration driver [13, 14]. Thus, uncertainty architectures operate as epistemic compasses within autonomous discovery ecosystems.

Autonomous and closed-loop systems

Autonomous materials discovery systems represent the convergence of machine learning, robotics, and high-throughput computation into self-optimizing scientific infrastructures [19, 20]. These platforms automate iterative cycles of hypothesis generation, materials synthesis, characterization, and model retraining.

Closed-loop experimentation operationalizes this paradigm by coupling predictive algorithms with robotic fabrication and in situ measurement technologies [16, 18]. Real-time feedback enables adaptive experimentation, where subsequent trials are dynamically informed by prior outcomes.

High-throughput infrastructures underpin this autonomy by scaling computational screening and experimental throughput. However, resource allocation becomes a critical optimization variable, requiring decision logics that balance exploration breadth with validation depth [12, 17].

Inverse design frameworks invert traditional forward prediction by optimizing input structures for target properties. Generative models—including variational autoencoders, diffusion systems, and generative adversarial networks—navigate latent design spaces to propose candidate materials [8, 26]. Multi-property constraints embedded within these systems rely heavily on entangled representations derived from MTL and GNN architectures.

While technologically advanced, these systems expose interpretive limitations. Entanglement that aids optimization may obscure mechanistic reasoning, constraining scientific explainability and trust calibration [23, 25]. This reveals a conceptual gap between autonomous capability and epistemic transparency.

Integration across ecosystems

At the ecosystem scale, materials informatics infrastructures increasingly coalesce around foundation models pretrained on heterogeneous datasets spanning simulations, experiments, and literature corpora [27, 28]. These models function as generalizable representation engines adaptable across tasks, accelerating downstream deployment.

Simulation–experiment coupling forms a critical integrative axis. Machine learning mediates discrepancies between computational approximations and empirical measurements, calibrating predictive pipelines across virtual and physical domains [29, 30]. This hybridization enhances model robustness while expanding applicability.

However, cross-property dynamics introduce systemic complexity. Feedback loops linking prediction, validation, and retraining can amplify both insight and error propagation depending on representational entanglement and uncertainty handling [31, 32]. Multi-task correlations, while beneficial, may inadvertently transmit biases or amplify epistemic blind spots.

Synthesizing these literatures reveals a convergent trajectory toward tightly coupled, self-steering discovery ecosystems. Yet interpretive frameworks capable of mapping cross-task entanglements, uncertainty propagation, and infrastructural feedbacks remain underdeveloped. This gap motivates the need for unified conceptual architectures that situate computational workflows not merely as predictive engines but as epistemic systems shaping the trajectory of materials discovery.

Proposed conceptual framework

The Property Entanglement Lattice (PEL) To address the interpretive needs in multi-task materials learning, we introduce the Property Entanglement Lattice (PEL), a novel conceptual structure that models cross-property interdependencies as a multidimensional lattice. In PEL, properties are nodes within a lattice graph, connected by edges representing entanglement strengths derived from shared representational subspaces. This framework conceptualizes multi-task systems as layered architectures where data ingestion, model inference, and discovery outputs interact through entanglement-mediated feedbacks.

Structurally, PEL comprises three layers: the input representation layer, the entanglement core, and the output steering layer. The input layer aggregates multimodal data streams—such as compositional vectors, graph embeddings, and uncertainty profiles—into a unified manifold. The entanglement core then weaves these into a lattice, where property nodes entangle via tensor products of latent features, capturing systemic correlations without empirical tuning. Finally, the output layer translates entangled states into discovery directives, prioritizing workflows based on lattice stability.

Data → Model → Discovery Pipelines PEL delineates pipelines as entanglement-propagating flows. Data ingestion initiates with heterogeneous sources funneled into representational embeddings, where initial entanglements emerge from feature correlations. Model inference amplifies these through multi-task heads, optimizing joint losses that reinforce lattice connections. Discovery pipelines then leverage the lattice for inverse mappings, steering toward novel materials by navigating entangled property landscapes.

Feedback Loops and Computational Steering Integral to PEL are bidirectional feedback loops that dynamically adjust lattice configurations. Uncertainty signals from one property node propagate across edges, recalibrating adjacent nodes to maintain epistemic balance. Computational steering logics emerge from this, where entanglement metrics guide resource allocation—e.g., prioritizing simulations in highly entangled subspaces to maximize informational yield.

This can be conceptualized as an entanglement density function, expressed as $\tau = \sum_{i,j}^{(E)} w_{ij} \sigma(z_i \otimes z_j)$, where w_{ij} denotes edge weights between properties i and j , z are latent vectors, and σ a non-linear activation capturing interaction non-additivities. This formula interprets the aggregate entanglement as a weighted tensor interaction, highlighting how dense connections enhance pipeline robustness.

Another dynamic is the steering trade-off, may be expressed as $\tau = \underset{U}{\operatorname{argmin}} (\alpha U + \beta D(E))$, with U as uncertainty aggregate and $D(E)$ as disentanglement cost, balanced by scalars α, β . This captures the interaction between minimizing risks and preserving entanglements for efficient discovery.

A third aspect formalizes loop closure: $\Delta L = \int \nabla E f(p) dp$, where ∇E is the entanglement gradient over property path p , and f the feedback functional. This symbolizes iterative refinements as path integrals along lattice edges, steering toward converged states.

These elements integrate within PEL, as conceptualized in **Figure 1**, which depicts the lattice with layered nodes, pipeline arrows, and looping feedbacks, illustrating entanglement flows.

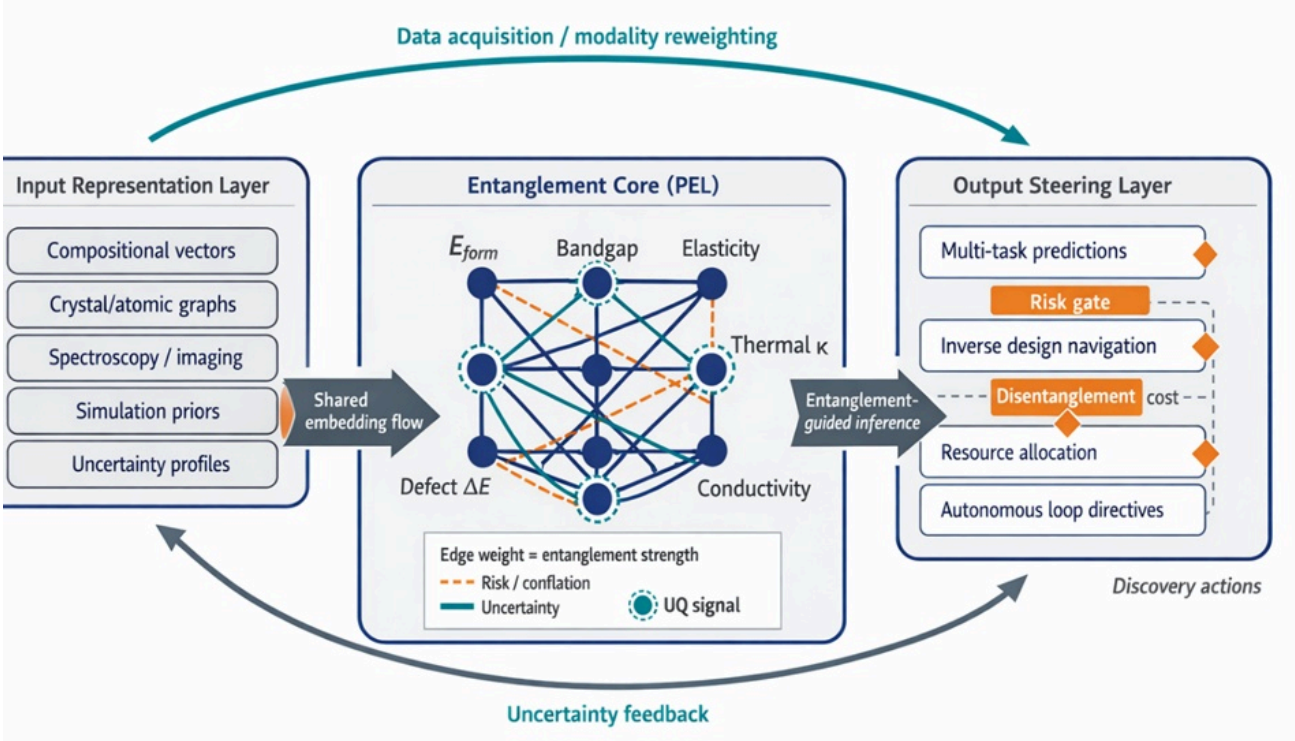


Figure 1. Property Entanglement Lattice (PEL) as an interprete architecture for multi-task materials learning.

PEL conceptualizes cross-property coupling as a lattice core mediating flows from multimodal inputs to discovery steering. Property nodes (tasks) are connected by weighted entanglement edges that encode shared latent subspaces, enabling information propagation and coordinated uncertainty handling across outputs. The framework integrates three layers—representation ingestion, entanglement core, and steering outputs—linked by bidirectional feedback loops where uncertainty and mismatch signals reconfigure lattice connectivity to support resource allocation, inverse design navigation, and epistemic risk management within autonomous discovery pipelines.

Analytical implications

The Property Entanglement Lattice (PEL) offers a lens for interpreting systemic behaviors in multi-task materials learning, revealing implications for computational infrastructures and discovery workflows. At the infrastructure level, PEL underscores how entanglement structures influence data processing pipelines, where densely connected property nodes facilitate efficient information propagation across tasks. This interpretation suggests that in high-throughput systems, entanglement can optimize resource distribution by prioritizing computations in subspaces with high interconnection densities, thereby reducing redundant simulations while enhancing coverage of correlated material attributes. Operational design levers for entanglement-aware pipelines are consolidated in Table 2.

Table 2. PEL design checklist: mapping pipeline decisions to entanglement-aware controls

Pipeline locus	Design decision (what changes)	Entanglement effect (PEL interpretation)	Control / diagnostic (conceptual)	Success criterion (interpretive)
Data ingestion	Modality balance (sim vs exp vs literature)	Prevents dominance-driven lattice skew	Modality reweighting; coverage audits	No single modality dictates edge formation
Representation learning	Shared embedding depth vs task-specific	Sets baseline coupling capacity	Adapter layers; modular encoders	Structured coupling without task collapse

	adapters			
Multi-task objective	Loss weighting across properties	Alters edge strengths and directionality	Task reweighting; gradient conflict checks	Stable lattice; reduced negative transfer
Uncertainty layer	Joint vs independent UQ	Governs propagation across edges	Covariance-aware UQ; ensemble diversity	Uncertainty signals steer without over-spreading
Risk gating	Thresholds for extrapolation / conflation	Converts distortions into steerable constraints	Applicability checks; orange “risk gates”	High-risk edges trigger conservative actions
Discovery steering	Exploration policy (where to sample next)	Chooses paths through lattice	Entanglement-guided acquisition	Sampling increases coverage of weak/critical edges
Closed-loop updates	Retraining cadence + damping	Determines coupling drift over time	Loop damping; stability monitoring	Edge weights converge rather than oscillate
Inverse design	Multi-objective constraints handling	Uses lattice as feasible manifold	Constraint prioritization; Pareto path tracing	Designs follow physically coherent edge paths

In terms of representation-inference interactions, PEL highlights the role of lattice edges as conduits for feature sharing, where entanglements emerge from overlapping latent dimensions in deep architectures. This dynamic implies that model designs incorporating modular lattice components could better handle multimodal inputs, integrating disparate data modalities—such as spectroscopic signals and structural graphs—into cohesive embeddings. Consequently, inference processes become more resilient to data heterogeneity, as entangled nodes buffer against isolated property variances, fostering smoother transitions in simulation-experiment couplings.

Discovery steering logics within PEL emphasize adaptive navigation through entangled spaces, where feedback loops modulate lattice configurations to align with epistemic priorities. For instance, in inverse design scenarios, steering might favor paths along strongly entangled edges to explore property combinations that reflect physical interdependencies, such as mechanical-thermal couplings in alloys. This interpretive approach reveals trade-offs in workflow efficiency: while entanglement amplifies exploratory breadth, excessive density could introduce computational overheads in uncertainty propagation, necessitating balanced lattice pruning strategies.

Epistemic risk structures are another key implication, as PEL frames risks as distortions in lattice geometry. Weak entanglements may signal underrepresented domains, prompting targeted data acquisition to reinforce connections, whereas over-entanglements risk conflating distinct properties, leading to blurred inference boundaries. This perspective informs uncertainty quantification by viewing variances as edge perturbations, enabling systems to calibrate confidence intervals based on lattice stability metrics.

Overall, these implications position PEL as a tool for conceptualizing infrastructure trade-offs, where entanglement serves as a metric for evaluating multi-task viability. By interpreting interactions at this level, researchers can devise workflows that leverage entanglement for epistemic gains, such as in autonomous systems where closed-loop decisions hinge on lattice-derived insights.

Results and Discussion

Integrating PEL into broader computational materials paradigms invites reflection on its alignment with existing ecosystems. In materials informatics, the lattice structure complements graph-based representations by extending them to property-level networks, potentially enhancing scalability in foundation models that span diverse scientific domains [1, 2]. This alignment suggests that entanglement-aware designs could mitigate biases in pretrained models, where shared subspaces inadvertently amplify dataset artifacts across tasks [3, 4].

However, interpretive challenges arise in applying PEL to real-world pipelines, particularly in handling dynamic data environments. High-throughput computations generate evolving datasets, and PEL's static lattice metaphor may require extensions to accommodate temporal evolutions, such as adaptive edge weights responsive to new experimental inputs [5, 6]. This points to opportunities for hybrid frameworks that blend PEL with active learning logics, where uncertainty-driven sampling strengthens entanglements in under-explored regions [7, 8].

Trade-offs in representation learning are evident, as PEL illuminates how deep architectures foster entanglements that boost generalization but complicate explainability [9, 10]. For graph neural networks, lattice interpretations could guide layer designs to explicitly encode property interlinks, improving transferability across material classes without sacrificing interpretability [11, 12]. Yet, in multimodal contexts, fusing datasets risks uneven entanglement distributions, where dominant modalities overshadow others, necessitating equitable weighting schemes [13, 14].

In autonomous discovery, PEL's feedback loops resonate with closed-loop systems, offering a conceptual basis for steering optimizations that prioritize epistemic closure over brute-force exploration [15, 16]. This could refine inverse design by mapping desired outputs onto lattice paths, balancing multi-objective constraints inherent to materials optimization [17, 18].

Broader field implications include fostering interdisciplinary bridges, as PEL's systems-level view applies to analogous domains like chemical engineering, where property entanglements mirror reaction networks [19, 20]. Limitations stem from its conceptual nature, lacking direct mapping to quantifiable metrics, though it provides a scaffold for future formalizations [21, 22]. Ultimately, PEL encourages a shift toward entanglement-centric thinking, enriching data-driven materials engineering with nuanced interpretive tools.

Conclusion

The Property Entanglement Lattice (PEL) advances a conceptual understanding of cross-property dynamics in multi-task materials learning, interpreting entanglements as foundational elements of computational workflows and discovery infrastructures. By framing these interactions through lattice structures, pipelines, and steering logics, PEL elucidates systemic trade-offs and epistemic structures, offering interpretive insights for optimizing multi-task systems. This framework integrates seamlessly with advancements in materials informatics, representation learning, and autonomous discovery, positioning entanglement as a pivotal logic for navigating complex material spaces. As data-driven paradigms evolve, PEL serves as a conceptual anchor, guiding infrastructural innovations toward more cohesive and efficient materials engineering ecosystems.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 12 Sep 2022

Revised: 16 Oct 2022

Accepted: 17 Nov 2022

Published online: 18 March 2023

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Xu P, Ji X, Li M, Lu W. Small data machine learning in materials science. *npj Comput Mater*. 2023;9(1):42.
<https://doi.org/10.1038/s41524-023-01000-z>.
- Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. *npj Comput Mater*. 2018;4(1):25.
<https://doi.org/10.1038/s41524-018-0081-z>.
- Umehara M, Stein HS, Guevarra D, Newhouse PF, Boyd DA, Gregoire JM. Analyzing machine learning models to accelerate generation of fundamental materials insights. *npj Comput Mater*. 2019;5(1):34.
<https://doi.org/10.1038/s41524-019-0172-5>.
- Sutton C, Boley M, Ghiringhelli LM, Rupp M, Vreeken J, Scheffler M. Identifying domains of applicability of machine learning models for materials science. *Nat Commun*. 2020;11(1):4428.
<https://doi.org/10.1038/s41467-020-17112-9>.
- Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater*. 2019;5(1):21.
<https://doi.org/10.1038/s41524-019-0153-8>.
- Goodall REA, Lee AA. Predicting materials properties without crystal structure: deep representation learning from stoichiometry. *Nat Commun*. 2020;11(1):6280.
<https://doi.org/10.1038/s41467-020-19964-7>.
- Gupta V, Choudhary K, Tavazza F, Campbell C, Liao W-k, Choudhary A, et al. Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data. *Nat Commun*. 2021;12(1):6595.
<https://doi.org/10.1038/s41467-021-26921-5>.
- Antunes LM, Grau-Crespo R, Butler KT. Distributed representations of atoms and materials for machine learning. *npj Comput Mater*. 2022;8(1):44.

<https://doi.org/10.1038/s41524-022-00729-3>.

Tao Q, Xu P, Li M, Lu W. Machine learning for perovskite materials design and discovery. *npj Comput Mater.* 2021;7(1):23.
<https://doi.org/10.1038/s41524-021-00495-8>.

Kailkhura B, Gallagher B, Kim S, Hiszpanski A, Han TY. Reliable and explainable machine-learning methods for accelerated material discovery. *npj Comput Mater.* 2019;5(1):108.
<https://doi.org/10.1038/s41524-019-0248-2>.

Feng S, Fu H, Zhou H, Wu Y, Lu Z, Dong H. A general and transferable deep learning framework for predicting phase formation in materials. *npj Comput Mater.* 2021;7(1):10.
<https://doi.org/10.1038/s41524-020-00488-z>.

Nyshadham C, Rupp M, Bekker B, Shapeev AV, Mueller T, Rosenbrock CW, et al. Machine-learned multi-system surrogate models for materials prediction. *npj Comput Mater.* 2019;5(1):51.
<https://doi.org/10.1038/s41524-019-0189-9>.

Kim Y, Kim Y, Yang C, Park K, Gu GX, Ryu S. Deep learning framework for material design space exploration using active transfer learning and data augmentation. *npj Comput Mater.* 2021;7(1):140.
<https://doi.org/10.1038/s41524-021-00609-2>.

Jennings PC, Lysgaard S, Hummelshøj JS, Vegge T, Bligaard T. Genetic algorithms for computational materials discovery accelerated by machine learning. *npj Comput Mater.* 2019;5(1):46.
<https://doi.org/10.1038/s41524-019-0181-4>.

Bartók AP, De S, Poelking C, Csányi G, Ceriotti M, Kermode JR, et al. Machine learning unifies the modeling of materials and molecules. *Sci Adv.* 2017;3(12):e1701816.
<https://doi.org/10.1126/sciadv.1701816>.

Allen AEA, Tkatchenko A. Machine learning of material properties: Predictive and interpretable multilinear models. *Sci Adv.* 2022;8(18).
<https://doi.org/10.1126/sciadv.abm7185>.

He T, Huo H, Bartel CJ, Wang Z, Chen K, Ceder G. Precursor recommendation for inorganic synthesis by machine learning materials similarity from scientific literature. *Sci Adv.* 2023;9(23).
<https://doi.org/10.1126/sciadv.adg8180>.

Sun W, Zheng Y, Yang K, Zhang Q, Shah AA, Wu Z, et al. Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Sci Adv.* 2019;5(11).
<https://doi.org/10.1126/sciadv.aay4275>.

Cagliani A, Coli GM, Baroni E, Dijkstra M, Frenkel L. Inverse design of soft materials via a deep learning–based evolutionary strategy. *Sc Adv.* 2022;8(3):eabj6731.
<https://doi.org/10.1126/sciadv.abj6731>.

Pilania G. Machine learning in materials science: From explainable predictions to autonomous design. *Comput Mater Sci.* 2021;193(193):110360.
<https://doi.org/10.1016/j.commatsci.2021.110360>.

Picklum M, Beetz M. MatCALO: Knowledge-enabled machine learning in materials science. *Comput Mater Sci.* 2019;163:50-62.

Deringer VL, Bartók AP, Proserpio L, Day GM, Csányi G, Pickard CJ. Machine learning interatomic potentials as emerging tools for materials science. *Adv Mater.* 2019;31(49):1902765.
<https://doi.org/10.1002/adma.201902765>.

Lv C, Zhou X, Zhong L, Yan C, Srinivasan M, Seh ZW, et al. Machine learning: An advanced platform for materials development and state prediction in lithium-ion batteries. *Adv Mater.* 2022;34(25):2101474.

Dinic J, Sharma A, Rekar Y, Veeravalli V, Bediako B, Jones R, et al. Applied machine learning for developing next-generation functional materials. *Adv Funct Mater.* 2021;31(51):2104195.
<https://doi.org/10.1002/adfm.202104195>.

Suwardi A, Lim DH, Li Y, Lan Z, Chen Y, Wee ATS. Machine-Learning-Driven biomaterials evolution. *Adv Mater.* 2022;34(1):2102703.
<https://doi.org/10.1002/adma.202102703>.

Wang AY-T, Murdock RJ, Kauwe SK, Oliynyk AO, Gurlo A, Brgoch J, et al. Machine learning for materials scientists: An introductory guide toward best practices. *Chem Mater.* 2020;32(12):4954-65.
<https://doi.org/10.1021/acs.chemmater.0c01907>.

Vasudevan RK, Choudhary K, Mehta A, Smith R, Kusne G, Tavazza F, et al. Materials science in the artificial intelligence age: high-throughput library generation, machine learning, and a pathway to equity. *Nat Mater.* 2021;20(9):1173-81.

Danilov DL, Tseng S, Nordlund K, Djurabekova F, Zhukovskii YF, Sokolov PS. Machine-Learning-Assisted determination of the global zero-temperature phase diagram of materials. *Adv Mater.* 2023;35(14):2107888.
<https://doi.org/10.1002/adma.202210788>.

Suwardi A, Cao J, Hu L, Wei F, Wu M, Zhao Y. Innovative materials science via machine learning. *Adv Funct Mater.* 2022;32(10):2108044.
<https://doi.org/10.1002/adfm.202108044>.

Zhang X, Wang D, Chen Z, Malik V, Wang H, Zhu S, et al. Finding the next superhard material through ensemble learning. *Adv Mater.* 2021;33(7):2005112.
<https://doi.org/10.1002/adma.202005112>.

Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. *Nature.* 2018;559(7715):547-55.
<https://doi.org/10.1038/s41586-018-0337-2>.

Ramprasad R, Batra R, Pilania G, Mannodi-Kanakithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput Mater.* 2017;3(1):54.
<https://doi.org/10.1038/s41524-017-0056-5>.