REFRAMING LLM FINETUNING THROUGH BAYESIAN OPTIMIZATION LENSES

Anonymous authors
Paper under double-blind review

ABSTRACT

Chemical reaction optimization remains a critical bottleneck in drug discovery and materials science. While reaction procedures are naturally documented as text in research papers and protocols, converting these descriptions into structured features for machine learning poses significant challenges. We present a novel framework that leverages LLMs to directly process textual reaction descriptions, combined with deep kernel learning to accelerate optimization. Our approach adapts LLM embeddings through joint optimization with Gaussian processes, enabling dynamic reorganization of the latent space to reflect reaction performance. Unlike previous methods using static LLM embeddings, our approach induces a natural metric learning effect through the GP marginal likelihood, clustering successful reaction conditions while separating unsuccessful ones. We demonstrate that this embedding adaptation emerges independently of the initial LLM, suggesting broad applicability across different foundation models. Empirical evaluation on Buchwald-Hartwig reactions shows our method reduces the number of experiments needed to identify optimal conditions by 45% compared to static embeddings, while maintaining well-calibrated uncertainty estimates. Further experiments in drug discovery and catalyst design validate the framework's effectiveness across diverse chemical domains.

1 INTRODUCTION

Large Language Models (LLMs) have demonstrated remarkable capabilities in natural language understanding and generation Wei et al. (2022). Their success stems from an ability to learn rich representations of text that capture subtle patterns, relationships, and even domain-specific knowledge. This representational power has naturally led to growing interest in adapting LLMs beyond general language tasks to specialized domains - from scientific discovery to reasoning tasks Boiko et al. (2023); Bran et al. (2023)

The adaptation of LLMs to new domains presents both opportunities and challenges. While techniques like prompt engineering and fine-tuning can effectively guide LLM behavior, they often focus solely on improving model outputs without considering the reliability of those predictions. Yet in many real-world applications, understanding prediction uncertainty is as crucial as the predictions themselves. Whether in drug discovery, materials design, or automated reasoning, decisions based on model predictions can have significant consequences, making reliable uncertainty quantification essential

In this context, Bayesian optimization (BO) has emerged as a powerful strategy, capable of navigating sparse data landscapes by judiciously balancing exploration and exploitation Shields et al. (2021) The challenge of extracting reliable uncertainties from Large Language Models (LLMs) for efficient BO has also sparked diverse approaches. The methods explored range from prompt-based techniques that quantify uncertainty through averaging LLM responses Ramos et al. (2023), to PEFT-based finetuning that transforms LLMs into Bayesian neural networks via Laplace approximation Kristiadi et al. (2024). While the area of intersection between BO and LLMs inspire novel ideas, Gaussian Processes Williams & Rasmussen (2006) (GPs) stand out as a particularly compelling framework due to their principled approach to uncertainty quantification. GPs offer natural uncertainty estimates through their probabilistic modeling, provide interpretable confidence bounds, and maintain well-calibrated predictions even in regions of sparse data.

Recent works have demonstrated successful integration of LLMs with GPs for optimization tasks. However, these approaches primarily utilize LLMs as fixed feature extractors. Despite the rich information encoded in LLM representations, treating them statically creates an artificial barrier between the expressive power of LLMs and the probabilistic rigor of GPs. This gap is particularly significant in domains requiring sample-efficient optimization of complex objectives, where the quality of feature representations directly impacts exploration efficiency, such as chemistry.

Chemical synthesis optimization presents a compelling example of such challenges. As a critical aspect of drug discovery and materials science, it often serves as a bottleneck in research and development. Traditional approaches to synthesis optimization often rely on high-throughput experimentation or expert-guided exploration. While these methods have yielded success, they can be resource-intensive, time-consuming, and limited by human bias or the physical constraints of experimental setups. In recent years, machine learning techniques have emerged as powerful tools to accelerate this process, offering the potential to navigate vast chemical spaces more efficiently Coley et al. (2020); Jorner et al. (2021); Schwaller et al. (2022). Chemical reaction optimization presents a unique challenge in machine learning as well: finding optimal conditions in a vast design space with limited experimental data.

BO has shown particular promise in this domain, providing a framework for balancing exploration and exploitation in the search for optimal synthesis conditions. However, its effectiveness is heavily dependent on the quality of the underlying feature representations. While numerous representational schemes have been proposed—ranging from one-hot encodings Chuang & Keiser (2018) and molecular fingerprints Rogers & Hahn (2010); Schneider et al. (2015); Capecchi et al. (2020); Probst et al. (2022) to quantum mechanical descriptors Ahneman et al. (2018); Shields et al. (2021) and learned representations Schwaller et al. (2021)—each comes with its own set of trade-offs in terms of computational overhead, interpretability, and required domain expertise. Amidst this landscape, a surprisingly versatile and information-rich medium has been overlooked: natural language. Chemists have long documented reaction details using natural language in research papers and supplementary materials, creating a rich corpus of domain knowledge ripe for exploitation Vaucher et al. (2020); Guo et al. (2021); White (2023).

However, existing deep kernel methods have not been extended to leverage this rich chemical knowledge encoded in LLMs. This limitation is particularly significant for reaction optimization, where success depends on understanding and exploiting entire regions of favorable conditions rather than isolated optimal points. Previous approaches focusing on supervised fine-tuning with auxiliary prediction heads Kristiadi et al. (2024) can lead to overfitting and potentially miscalibrated uncertainty estimates, while failing to exploit the natural structure present in chemical spaces.

We address these limitations by introducing a framework that seamlessly integrates LLMs into the GP architecture through deep kernel learning. Our approach leverages the GP marginal likelihood as a training objective, enabling the model to automatically discover and exploit regions of high-performing reactions while maintaining well-calibrated uncertainties. This joint optimization induces an implicit contrastive learning effect, where embeddings of reactions with similar outcomes are pulled together while dissimilar reactions are pushed apart.

Critically, we observe that this embedding space organization emerges naturally from the optimization process, regardless of the choice of pretrained LLM. This result suggests that our approach effectively adapts even general-purpose LLMs into powerful chemical optimization tools, while maintaining the rigorous uncertainty quantification that makes GPs valuable for optimization tasks.

Our key contributions include:

1. A novel framework for LLM finetuning through GP optimization that simultaneously adapts embeddings and learns a probabilistic model of the chemical space
2. Demonstration that GP-based training naturally induces an implicit contrastive learning effect, organizing the latent space into regions of similar reaction outcomes
3. Theoretical analysis of how GP lengthscales interact with embedding space structure, providing insights into the effectiveness of LLM representations for Bayesian optimization
4. Empirical evidence of sample-efficient learning (from as few as 10 datapoints) and improved exploration of high-performing reaction conditions

2 METHODS

Our approach combines large language models with Gaussian processes through a deep kernel framework that enables joint optimization of the embedding space and the surrogate model. This design allows the model to discover and exploit regions of high-performing reactions while maintaining reliable uncertainty estimates.

2.1 DATA REPRESENTATION

Chemical reaction data presents unique challenges for machine learning due to its inherent complexity and heterogeneous nature. Reaction conditions typically comprise multiple parameter types: numerical values (temperature, concentration, time), categorical variables (catalyst type, solvent choice), and detailed procedural descriptions. This heterogeneity traditionally needs careful consideration of how to represent such diverse data types in a unified format suitable for machine learning models.

In chemistry, molecular representations have been extensively studied, leading to various approaches such as fingerprints, SMILES strings, and molecular graphs, each requiring specialized kernel functions to capture relevant similarity measures. However, reaction conditions and procedural descriptions present an additional layer of complexity beyond molecular representation, as they combine multiple data types with complex interdependencies.

LLMs offer a straightforward solution to this representation challenge. By structuring reaction conditions as natural language descriptions, we can leverage LLMs' ability to process and embed mixed data types into a continuous vector space. We construct these representations through a two-step process:

1. Template Construction: We define each reaction r through a standardized template: $r = \text{template}(\{\text{parameters}, \text{values}\})$ where the template converts various parameter types into a structured text format:

```
The reaction was prepared with:
temperature: {numerical_value}°C
solvent: {solvent_smile}
ligand: {ligand_smile}
```

2. LLM Embedding: We process the templated description through the LLM to obtain a fixed-dimensional embedding: $\mathbf{x} = \text{LLM}(r) \in \mathbb{R}^d$ which unifies representation of heterogeneous parameters and offers natural handling of categorical and numerical values. Moreover, it provides compatibility with standard continuous kernels (e.g., Matérn) and scalability to varying numbers of parameters while preserving the relationships between parameters.

The resulting embedding vectors $\mathbf{x}$ capture both the individual parameter values and their interactions, providing a rich representation space for subsequent Gaussian process modeling. This representation strategy removes the need for specialized kernel functions for different parameter types, as the LLM embedding space is already equipped with meaningful distance metrics suitable for standard continuous kernels.

2.2 VANILLA GAUSSIAN PROCESS

Given a chemical reaction template r, we obtain initial embeddings $\mathbf{x}$ and use them as fixed input features to a Gaussian process model with a Matérn-3/2 kernel:

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \left(1 + \frac{\sqrt{3}||\mathbf{x} - \mathbf{x}'||_2}{\ell}\right) \exp\left(-\frac{\sqrt{3}||\mathbf{x} - \mathbf{x}'||_2}{\ell}\right)$$

where ℓ is the lengthscale and σ^2 is the signal variance.

The hyperparameters $\{\ell, \sigma^2, \sigma_n^2\}$ are optimized by maximizing the log marginal likelihood:

$$\log p(\mathbf{y}|\mathbf{X}) = -\frac{1}{2}\mathbf{y}^\top(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{K} + \sigma_n^2\mathbf{I}| - \frac{n}{2}\log(2\pi)$$

Using the vanilla GP approach, the LLM embeddings remain fixed throughout the optimization process. The model's adaptability comes solely from tuning the GP hyperparameters to fit the observed data. This setup preserves the underlying geometric relationships between reactions in the embedding space, relying on the pretrained LLM's inherent understanding of chemical similarity.

2.3 Deep Kernel Gaussian process

However, general pretrained LLMs do not explicitly come with the chemical understanding of the data at hand. By using fixed features from LLMs we depend on the predefined knowledge incorporated in their weights. The success of the Bayesian optimization procedure builds upon the ability of the GP to model the fixed input space that hopefully contains enough information.

Deep kernel Gaussian processes, on the other hand, combine the flexibility of deep neural networks with the principled uncertainty quantification of Gaussian processes. In this approach, the kernel function is composed with a learned feature transformation:

$$k_\theta(\mathbf{x}, \mathbf{x}') = k(f_\theta(\mathbf{x}), f_\theta(\mathbf{x}'))$$

where f_θ is a parameterized feature extractor with parameters θ. This composition allows the model to learn task-specific feature representations while maintaining the probabilistic properties of the GP framework. The learned transformation and GP parameters are jointly optimized through the marginal likelihood where $\mathbf{K}_\theta$ is the kernel matrix computed using the transformed features.

2.4 LLM-based Deep Kernel

In our framework, we explore different approaches to constructing the feature transformation f_θ.

1. Projection Layer: A learned non-linear transformation $\mathbf{P} \in \mathbb{R}^{m \times d}$ applied to fixed LLM embeddings: $f_\theta(\mathbf{x}) = \mathbf{P}\text{LLM}(r)$ where m is the projection dimension.

Using a combination of linear layer network with an activation function on top of the fixed LLM features most closely aligns with standard deep kernel learning methods applied to GPs. This approach is particularly beneficial when the LLM model weights are not available as in the case of closed-source models from OpenAI. The projection layer can learn to emphasize or suppress different aspects of the fixed embeddings, effectively creating a task-specific view of the chemical space.

2. PEFT-Adapted LLM: Low-rank adaptation of LLM parameters: $f_\theta(\mathbf{x}) = \text{LLM}_\theta(r)$ where θ represents the trainable adapter parameters. Parameter-Efficient Fine-Tuning (PEFT) provides a clever solution to the challenge of adapting large language models. Instead of fine-tuning the entire model, which would be computationally prohibitive and potentially catastrophic to the model's learned representations, PEFT introduces small, trainable adapter modules within the LLM. Specifically, we employ Low-Rank Adaptation (LoRA) **?**, which decomposes weight updates into low-rank matrices. Following this approach allows us to preserve the potential chemical knowledge captured during pretraining, avoiding catastrophic forgetting. Through reducing the number of trainable parameters by several orders of magnitude compared to full fine-tuning, the low-rank updates can capture task-specific patterns while keeping the model's general capabilities.

3. Combined Approach: Sequential application of PEFT and projection: $f_\theta(\mathbf{x}) = \mathbf{P}\text{LLM}_\theta(r)$, thus combining the benefits of both worlds. The PEFT adapters allow the LLM to adapt its internal representations to the optimization task, while the projection layer provides an additional degree of freedom to reshape the embedding space.

With any of these methods, we optimize the parameters θ (projection matrix and/or PEFT parameters) jointly with the GP hyperparameters through the marginal likelihood. In other words, we are finetuning the LLM through the GP loss which allows the model to learn transformations that both preserve relevant chemical information, organize the latent space to better reflect the structure of the optimization objective and provide well-calibrated uncertainty measures.

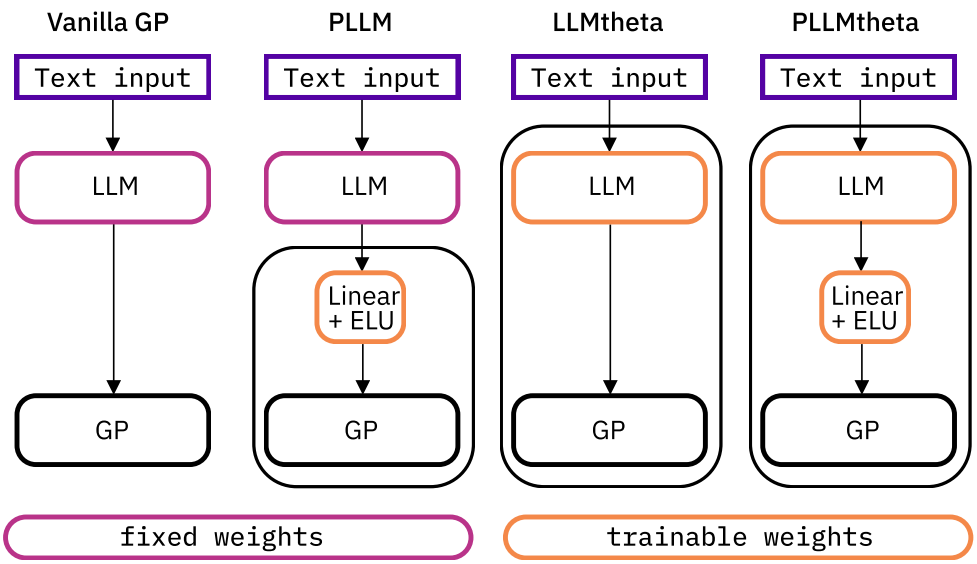

Figure 1: GP+LLMs architectures. From left to right: Vanilla GP (LLM serves as a fixed feature extractor); LLM+Projection layer (PLLM) - GP learns a mappping to a more compact latent space from the LLM embedding; LLM$_\theta$ - LLM weights are trainable through GP loss; PLLM$_\theta$ - both the projection and LLM weights are trained jointly with the GP.

2.5 Implicit Metric Learning

A key feature of our approach is that the GP marginal likelihood objective naturally induces a contrastive effect in the embedding space. The kernel function $k(\mathbf{x}, \mathbf{x}')$ measures similarity between points, and optimizing the marginal likelihood encourages:

$$\text{If } |y_i - y_j| \text{ is small: } ||f_\theta(\mathbf{x}_i) - f_\theta(\mathbf{x}_j)||_2 \downarrow$$
$$\text{If } |y_i - y_j| \text{ is large: } ||f_\theta(\mathbf{x}_i) - f_\theta(\mathbf{x}_j)||_2 \uparrow$$

In other words, the joint GP optimization induces high kernel values (small distances) between points with similar outputs and low kernel values (large distances) between points with different outputs. This reorganization of the embedding space happens automatically through the optimization of the deep kernel parameters, adapting the feature space to better align with the reaction outcomes without requiring explicit contrastive loss terms.

2.6 Bayesian optimization

We use expected improvement as our acquisition function to select the next reaction to evaluate from the held-out set. In our framework, the acquisition function plays a dual role: not only does it guide exploration-exploitation trade-off in the traditional BO sense, but it also influences the evolution of the embedding space itself through the sequential optimization process.

This creates an interesting dynamic system where each component influences the others: the acquisition function's suggestions affect which regions of the space are explored, which in turn impacts how the LLM adapts its embeddings through PEFT, which then shapes how the GP models the space and generates uncertainties. This interconnected nature shares conceptual similarities with reinforcement learning, where actions (acquisition suggestions) influence both the state space (embedding organization) and the value estimates (GP predictions). However, unlike traditional RL, our system maintains explicit uncertainty quantification through the GP and operates directly in the learned representation space.

The result is a well-integrated joint system where the acquisition function, embedding adaptation, and GP modeling work in concert to discover and exploit high-performing regions. This synergy is particularly evident in how the embedding space progressively organizes itself around regions of similar reaction outcomes, as shown in our empirical analysis.

3 RESULTS

4.1 Bayesian Optimization with Fixed LLM Features

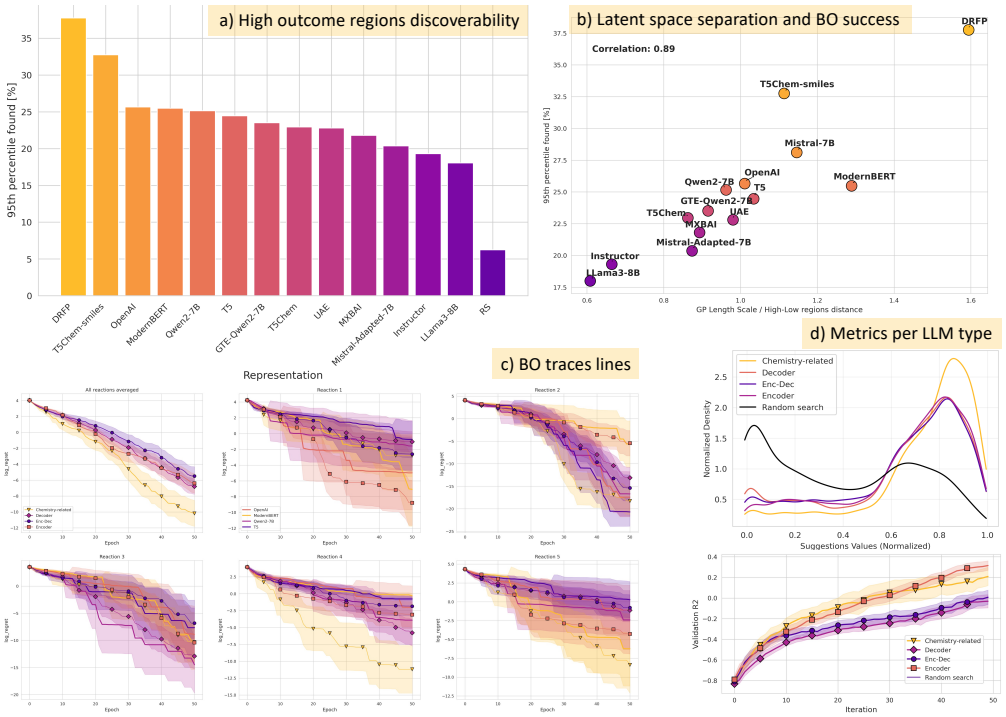

Figure 2: BO performance with fixed LLM features as inputs to a vanilla GP. a) On the top left we demonstrate the average discovery of high-impact regions of the design space (reactions with high yield). The y-axis represents the percentage of the top 5% reactions found during the optimization process, across all five Buchwald-Hartwig reactions. b) We analyze the relation between different LLM embeddings and their success rates in BO. By uncovering high correlation between the ratio of the GP learned length scales and the average separability of the latent space, we show that for a successful optimization process the embeddings should be differentiable enough to learn important (di)similarities. c) Optimization paths for 1 - all reactions averaged across different LLM types (Decocer only, Encoder only, Encoder-Decoder); 2-6 - individual reactions and LLM models' performance. d) Top: Distribution of suggestions generated throughout the entire optimization process (50 iterations and 20 seed runs). Bottom: R2 score over different LLM types, averaged across all reactions.

We first evaluate the effectiveness of LLM embeddings as fixed feature extractors for Bayesian optimization across five Buchwald-Hartwig reactions. As shown in Figure 1a, while chemistry-specialized representations (DRFP, T5Chem-smiles) excel at discovering high-yield reactions (95th percentile), general-purpose LLM embeddings also provide competitive performance. However, we observe substantial variation in performance across different LLM embeddings, prompting an analysis of the underlying factors.

To understand this variability, we examine the relationship between embedding space structure and BO success. Figure 1b reveals a strong positive correlation (0.89) between the ratio of GP length-scale to high-low region distances and the model's ability to discover high-yield reactions. This ratio

effectively captures how well the GP's understanding of similarity (through its lengthscale) aligns with the actual separation between different performance regions in the embedding space. Models with higher ratios, like DRFP and T5Chem-smiles, allow the GP to maintain meaningful correlations across performance regions, leading to smoother fits that better capture the underlying structure of the objective function. This smoother GP behavior enables more effective exploration-exploitation trade-offs during optimization, as the acquisition function can better assess uncertainties across both high and low-performing regions.

The BO traces in Figure 1c demonstrate that performance varies across different reactions, with no representation consistently outperforming others across all tasks - including chemistry-specialized ones. While all LLM types show similar distributions of suggested point evaluations (Figure 1d, top), encoder-based models tend to achieve higher R^2 values during training (Figure 1d, bottom). However, this improved function approximation does not necessarily translate to better BO performance, highlighting that the primary objective of BO is discovering optimal values rather than complete function mapping.

Interestingly, among LLM models, T5Chem only performs well when using input representations similar to its pretraining data (reaction SMILES), reinforcing previous findings about the limited generality of domain-specialized LLMs. This observation suggests that input representation influences BO success through two mechanisms: (1) contextual alignment with pretraining helps models better leverage their learned weights, and (2) the resulting embedding space organization affects the GP's ability to learn appropriate lengthscales for modeling the objective function.

3.1 LLM-based Deep Kernel

This analysis motivates exploring approaches where embeddings and GP parameters can be jointly optimized. Such co-adaptation could enable embeddings to maintain meaningful neighbor relationships at distances that match the GP's lengthscale, while allowing the GP to adjust its similarity assumptions to match patterns in the embedding space. In the following section, we investigate this direction through deep kernel LLM-GP models.

This joint training approach leads to significant improvements in performance, increasing the discovery rate of high-performing reactions (95th percentile) compared to using fixed LLM embeddings by more than 50%. This substantial improvement validates our earlier analysis about the importance of alignment between GP lengthscales and embedding space structure.

The evolution of the embedding space (Figure 3) reveals how the GP's marginal log likelihood objective guides the LLM's representations through PEFT. Starting from the initial embedding distribution where high and low-performing points are mixed (bottom left), the space gradually reorganizes to create clearer separations between different performance regions. This reorganization happens without explicit contrastive learning objectives, emerging naturally from the GP's need to model the objective function effectively.

The pairwise distance distributions (right) track this evolution through the optimization process, showing how the initially overlapped distributions of high-high, high-low, and low-low distances gradually separate. This separation both mathematical and semantical as it reflects the model learning meaningful chemical relationships that help guide the optimization process. The GP predictions (top left) closely match the ground truth (middle left), suggesting the model has learned a reliable mapping between molecular structure and performance.

This bidirectional optimization - where the LLM adapts its embeddings to facilitate GP modeling while the GP learns to better navigate the evolving latent space - offers a more principled approach to optimization than using fixed representations. The emergence of clear high-performance clusters in the embedding space (visualized in 2D) suggests the model is learning to recognize and exploit patterns in chemical structures that correlate with desired properties.

4 Conclusion

This works presents a novel framework that reframes LLM finetuning through the lens of Bayesian optimization, demonstrating how joint training with Gaussian processes can dramatically improve the utility of LLM embeddings for optimization tasks.

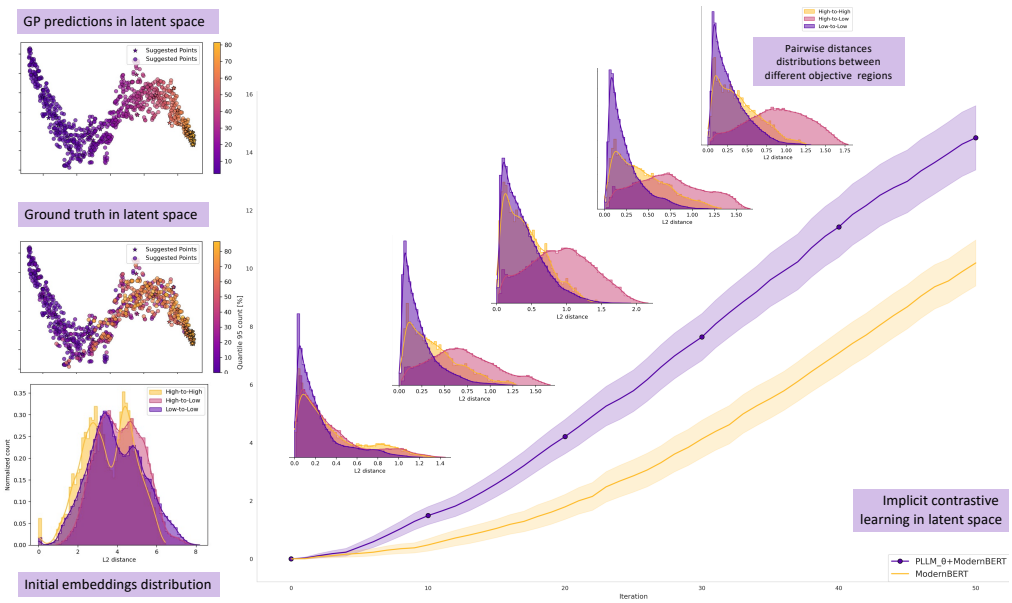

Figure 3: Deep Kernel LLM Finetuning through GP. We present the metric of the quantile coverage, while introducing the relation between the embeddings and their inner structure throughout the optimization process. Starting from a very densely populated and overlapped latent space (bottom left), we progressively improve the latent space structure throught the implicit contrastive learning patern that occurrs as a natural consequence of training a GP distance-based similarity loss. The lines show performance of the $PLLM_\theta$ architecture in comparison to Vanilla GP. We also visualize the latent space of embeddings and mark regions where the most sampling happens in the space. Accordingly to the latent space organization, the highly concentrated region of well performing chemical reactions gets sampled the most. Moreover, the GP aligns the predictions close to the ground truth, as observed in the top left plot.

The key innovation of our approach lies in leveraging the GP marginal likelihood to naturally induce organization in the embedding space, creating representations that better support sample-efficient optimization. This process occurs without explicit contrastive learning objectives, emerging instead from the GP's need to model the objective function effectively. The consistent improvement across different LLM architectures suggests we have identified a fundamental principle for adapting pretrained models to optimization tasks.

Our results on chemical reaction optimization demonstrate practical benefits compared to static embeddings. This improvement, combined with maintained uncertainty calibration, suggests promising applications beyond chemistry in domains where sample efficiency is crucial and data collection is expensive.

REFERENCES

Derek T Ahneman, Jesús G Estrada, Shishi Lin, Spencer D Dreher, and Abigail G Doyle. Predicting reaction performance in C–N cross-coupling using machine learning. *Science*, 360(6385):186–190, 2018.

Daniil A Boiko, Robert MacKnight, and Gabe Gomes. Emergent autonomous scientific research capabilities of large language models. *arXiv preprint arXiv:2304.05332*, 2023.

Andres M Bran, Sam Cox, Andrew D White, and Philippe Schwaller. Chemcrow: Augmenting large-language models with chemistry tools. *arXiv preprint arXiv:2304.05376*, 2023.

Alice Capecchi, Daniel Probst, and Jean-Louis Reymond. One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome. *Journal of cheminformatics*, 12(1):1–15, 2020.

Kangway V Chuang and Michael J Keiser. Comment on "predicting reaction performance in C–N cross-coupling using machine learning". *Science*, 362(6416):eaat8603, 2018.

Connor W Coley, Natalie S Eyke, and Klavs F Jensen. Autonomous discovery in the chemical sciences part i: Progress. *Angewandte Chemie International Edition*, 59(51):22858–22893, 2020.

Jiang Guo, A Santiago Ibanez-Lopez, Hanyu Gao, Victor Quach, Connor W Coley, Klavs F Jensen, and Regina Barzilay. Automated chemical reaction extraction from scientific literature. *Journal of Chemical Information and Modeling*, 62(9):2035–2045, 2021.

Kjell Jorner, Anna Tomberg, Christoph Bauer, Christian Sköld, and Per-Ola Norrby. Organic reactivity from mechanism to machine learning. *Nature Reviews Chemistry*, 5(4):240–255, 2021.

Agustinus Kristiadi, Felix Strieth-Kalthoff, Marta Skreta, Pascal Poupart, Alán Aspuru-Guzik, and Geoff Pleiss. A sober look at LLMs for material discovery: Are they actually good for Bayesian optimization over molecules? In *ICML*, 2024.

Daniel Probst, Philippe Schwaller, and Jean-Louis Reymond. Reaction classification and yield prediction using the differential reaction fingerprint drfp. *Digital Discovery*, 1(2):91–97, 2022.

Mayk Caldas Ramos, Shane S. Michtavy, Marc D. Porosoff, and Andrew D. White. Bayesian optimization of catalysts with in-context learning, 2023.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, 2010.

Nadine Schneider, Daniel M Lowe, Roger A Sayle, and Gregory A Landrum. Development of a novel fingerprint for chemical reactions and its application to large-scale reaction classification and similarity. *Journal of Chemical Information and Modeling*, 55(1):39–53, 2015.

Philippe Schwaller, Daniel Probst, Alain C Vaucher, Vishnu H Nair, David Kreutter, Teodoro Laino, and Jean-Louis Reymond. Mapping the space of chemical reactions using attention-based neural networks. *Nature Machine Intelligence*, 3(2):144–152, 2021.

Philippe Schwaller, Alain C Vaucher, Ruben Laplaza, Charlotte Bunne, Andreas Krause, Clemence Corminboeuf, and Teodoro Laino. Machine intelligence for chemical reaction space. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, pp. e1604, 2022.

Benjamin J Shields, Jason Stevens, Jun Li, Marvin Parasram, Farhan Damani, Jesus I Martinez Alvarado, Jacob M Janey, Ryan P Adams, and Abigail G Doyle. Bayesian reaction optimization as a tool for chemical synthesis. *Nature*, 590(7844):89–96, 2021.

Alain C Vaucher, Federico Zipoli, Joppe Geluykens, Vishnu H Nair, Philippe Schwaller, and Teodoro Laino. Automated extraction of chemical synthesis actions from experimental procedures. *Nature communications*, 11(1):3601, 2020.

Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.

Andrew D White. The future of chemistry is language. *Nature Reviews Chemistry*, pp. 1–2, 2023.

Christopher K Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.

A Appendix