

A New Paradigm in Tuning Learned Indexes: A Reinforcement Learning Enhanced Approach

TAIYI WANG, University of Cambridge, United Kingdom

LIANG LIANG*, EPFL, Switzerland

GUANG YANG, Imperial College London, United Kingdom

THOMAS HEINIS, Imperial College London, United Kingdom

EIKO YONEKI†, University of Cambridge, United Kingdom

Learned Index Structures (LIS) have significantly advanced data management by leveraging machine learning models to optimize data indexing. However, designing these structures often involves critical trade-offs, making it challenging for both designers and end-users to find an optimal balance tailored to specific workloads and scenarios. While some indexes offer adjustable parameters that demand intensive manual tuning, others rely on fixed configurations based on heuristic auto-tuners or expert knowledge, which may not consistently deliver optimal performance.

This paper introduces LITUNE, a novel framework for end-to-end automatic tuning of Learned Index Structures. LITUNE employs an adaptive training pipeline equipped with a tailor-made Deep Reinforcement Learning (DRL) approach to ensure stable and efficient tuning. To accommodate long-term dynamics arising from online tuning, we further enhance LITUNE with an on-the-fly updating mechanism termed the O2 system. These innovations allow LITUNE to effectively capture state transitions in online tuning scenarios and dynamically adjust to changing data distributions and workloads, marking a significant improvement over other tuning methods. Our experimental results demonstrate that LITUNE achieves up to a 98% reduction in runtime and a 17-fold increase in throughput compared to default parameter settings given a selected Learned Index instance. These findings highlight LITUNE's effectiveness and its potential to facilitate broader adoption of LIS in real-world applications.

CCS Concepts: • **Information systems** → **Data structures**.

Additional Key Words and Phrases: Learned Index, Reinforcement Learning, Parameter Tuning

ACM Reference Format:

Taiyi Wang, Liang Liang, Guang Yang, Thomas Heinis, and Eiko Yoneki. 2025. A New Paradigm in Tuning Learned Indexes: A Reinforcement Learning Enhanced Approach. *Proc. ACM Manag. Data* 3, 3 (SIGMOD), Article 120 (June 2025), 26 pages. <https://doi.org/10.1145/3725257>

1 Introduction

The intersection of data management and machine learning has given rise to learned index structures. These indexes integrate machine learning, replacing traditional algorithmic components, to capture

*Work has been partly performed at Imperial College London, London, UK.

†Corresponding author.

Authors' Contact Information: Taiyi Wang, University of Cambridge, Cambridge, United Kingdom, Taiyi.Wang@cl.cam.ac.uk; Liang Liang, EPFL, Lausanne, Switzerland, liang.liang@epfl.ch; Guang Yang, Imperial College London, London, United Kingdom, guang.yang15@imperial.ac.uk; Thomas Heinis, Imperial College London, London, United Kingdom, t.heinis@imperial.ac.uk; Eiko Yoneki, University of Cambridge, Cambridge, United Kingdom, eiko.yoneki@cl.cam.ac.uk.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 2836-6573/2025/6-ART120

<https://doi.org/10.1145/3725257>

data distributions and optimize search times. Notable examples include RMI [21], ALEX [10] and PGM [11], etc., which have become subjects of extensive research.

The effective design of a learned index involves deliberate trade-offs to achieve optimal performance for varying workloads. For instance, ALEX favors combined search and update performance by introducing gaps at the expense of space efficiency [10]. On the other hand, the dynamic PGM Index prioritizes update efficiency over search performance [11]. These design trade-offs also lead to more complex structures which generate configurable parameters. Tuning these parameters is the key to balancing the trade-offs that ensure higher performance over traditional indexes.

Beyond the primary parameters, learned indexes like ALEX have more subtle tunable factors that are often overlooked for simplicity. These parameters affect various aspects of the index performance, from operation cost (e.g., search and insertion cost) to the structure of the index (e.g., heights of the tree). For example, for ALEX, the *Max Node Size* parameter changes the size of the nodes, thereby affecting the height of the tree. On the other hand, *Split Policy* and *Gap Ratio* affect how insertion is carried out. These parameters are intertwined, and adjusting them in real-world scenarios can lead to substantial performance improvements, though it requires a more complex tuning process.

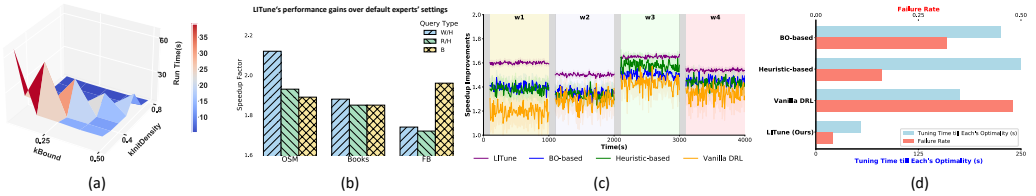


Fig. 1. (a) shows the performance surface of a learned index (ALEX) under a wild exploration of the parameter space. (b) highlights the optimal performance speedup achieved by LI-TUNE compared to default expert-selected parameters. (c) illustrates the continuous tuning performance of our system alongside other out-of-the-box methods under default configurations. (d) compares the tuning stability and costs across methods to reach their respective optimal performance levels.

Selecting the right tuning approach for learned indexes involves navigating a myriad of parameter configurations [16, 41]. For example, in practice, parameterized indexes can exhibit vastly different performance due to parameter choices. This is illustrated in Figure 1(a), where adjusting just two parameters leads to significant variability in runtime¹. This variability underscores the complexities and potential performance swings when considering the full spectrum of high-dimensional and continuous parameter configurations, which can scale into thousands or millions. This complexity is compounded by the fact that a misconfiguration can lead to drastic performance deviations, emphasizing the importance of precise tuning. Besides, our empirical experiments demonstrate that parameter interactions in learned indexes exhibit complex, workload-dependent relationships with no dominant parameters. As shown in Figure 2, where colors represent normalized parameter values and percentages show impact scores (ratio of individual-parameter to full-parameter tuning improvements), no parameter consistently exerts greater influence, with all impact scores falling between 10-25%. This heterogeneous distribution of parameter values across workloads, coupled with the balanced impact scores, indicates that performance optimization requires holistic parameter tuning rather than focusing on individual parameters. Moreover, unlike algorithmic indexes such as

¹In our experiments shown in Figure 1(a) and (d), we executed 16 million write and 16 million read queries on a 1 million SOSD dataset [18] using ALEX [10], selectively varying two parameters to validate the substantial performance gap among parameters.

B+trees that perform well out of the box, learned indexes are distribution-dependent. Furthermore, systems are not designed to automatically tune indexes and typically do not allocate substantial resources for this purpose. However, there is a critical requirement for indexes to perform optimally; this **necessitates that the tuner identifies high-quality solutions within a limited budget (Challenge C.1).**

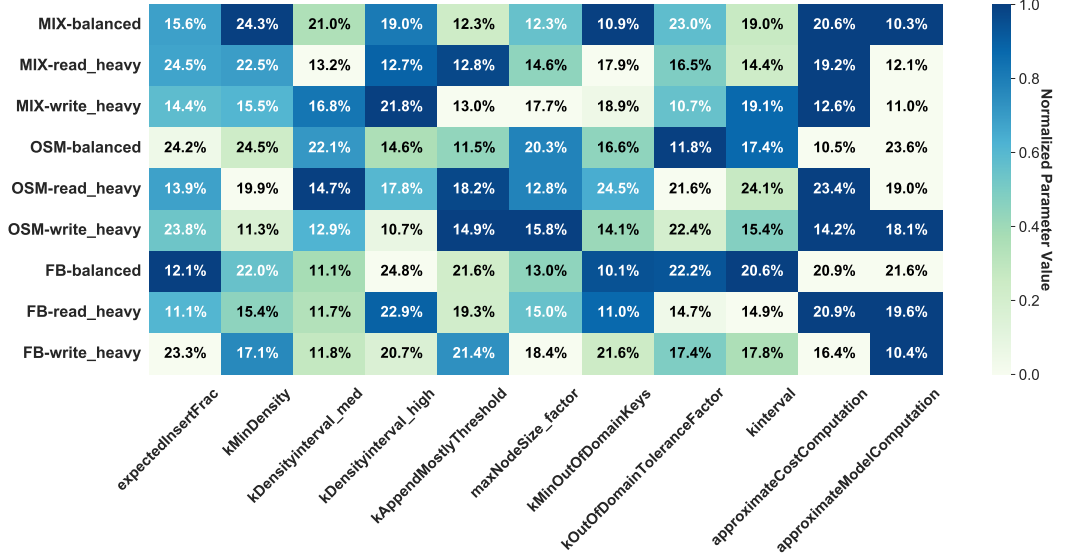


Fig. 2. Selected parameter value distributions and their impact scores across different workloads when tuning on ALEX. The heatmap colors represent normalized optimal parameter values, while the percentages indicate each parameter's individual tuning impact.

While learned indexes strive to be user-friendly by abstracting away underlying parameters, this abstraction can inadvertently lead to significant performance degradation, as hiding these critical parameters from the user may result in suboptimal default settings for various scenarios. By fine-tuning these parameters, we can achieve significant performance improvements. Figure 1(b) emphasizes the substantial performance gains (measured by query runtime speedup) achieved by our tuning system over SOSD [18] through comparisons between the optimal solution found by LITUNE and the default settings provided by experts. Importantly, unlike many DBMSs whose default configurations are often overly conservative [40, 49], the default settings in learned indexes are decided based on instance-optimized designs [9] which allow them to adapt to specific system environments given a "good" tuning algorithm. Currently, these algorithms are chosen by experts to optimize for common system environments. However, there hasn't been any end-to-end tuning system that is adaptive across data distributions. Thus, **tuning is necessary** and can lead to considerable enhancements.

Moreover, in real-world usage, data distributions and query types are not constant, which makes tuning more challenging. Figure 1(c) shows the degradation of tuning performance among existing out-of-box tuning methods during continuous online tuning when facing dynamic workloads, highlighting **the need for adaptive tuning to workloads and data distributions (Challenge C.2)**. We introduce four different workloads derived from mixture-distributed data from the SOSD dataset [18], involving various query types by adjusting the read-write ratio over time. To ensure

fair comparisons, each tuning method is provided with a preparation period, depicted as grey areas within the workload intervals. During this period, a preliminary smaller dataset reflecting future trends is used for warming up, initial tuning will happen during this period to make sure the system begins with reasonably optimized parameters.

Another challenge arises from the need for safe tuning, especially when dealing with a large parameter space and concurrent tuning demands. Recently, Reinforcement Learning enhanced by deep neural networks (DRL) has already been proved by many works [15, 49] as a good tuner when working within a large parameter space due to its intrinsic exploration abilities. Equipped with a learned module, an RL-based approach can be easily generalized and deployed to various data conditions. However, DRL, as a trial-and-error-based approach, normally presents aggressive tuning towards the optimal solutions, whose potential risks to the existing system were largely ignored. In such cases, **it is crucial to ensure tuning remains both safe and stable during exploration within the extensive parameter space (Challenge C.3)**. Figure 1(d) shows the failure rates caused by improper parameter settings from aggressive tuning approaches, particularly those resulting from exploratory vanilla DRL methods (PPO [31]). We also present the differences in tuning costs until each method reaches optimality, further emphasizing the motivation for introducing LITUNE. These observations underscore the necessity of a tailor-made tuning system rather than relying on out-of-box methods.

The LITUNE system, utilizing a tailored Deep Reinforcement Learning (DRL) framework, significantly improves the tuning of learned index parameters, overcoming the constraints of traditional methods and optimizing performance without requiring the extensive training data needed by supervised models. Key contributions include:

(1) **Introduction of an Automatic Tuning System Using DRL (Addressing C.1)**: This system efficiently navigates large parameter spaces in real-time, dynamically adjusting to changing performance requirements, capturing the stateful transitions, thus rapidly finding optimal solutions in complex online environments.

(2) **Adaptive Design (Addressing C.2)**: During the training stage, we leverage Meta-RL to efficiently transfer knowledge from a smaller pre-training set to unseen tuning tasks through solid initialization and fast gradient descent. In the online tuning stage, LITUNE employs an on-the-fly updating system, termed the O2 system, to further boost adaptability over the longer term. This design enhances the tuner's ability to quickly adapt to new or evolving workloads and data distributions.

(3) **Operational Stability and Reliability (Addressing C.3)**: Featuring a Context-RL-based risk mitigation strategy (ET-MDP solver), LITUNE avoids dangerous configurations, ensuring the tuning process maintains the system's operational stability and reliability.

To our best knowledge, *LITUNE is the first system to enable stateful, online tuning of learned indexes using tailored Deep Reinforcement Learning methods, integrated with safety-aware mechanisms and O2 system to ensure reliable and adaptive performance.*

The paper is organized as follows: Section 2 reviews related work. Section 3 discusses our motivation and the design of LITUNE. Section 4 details our novel RL-based tuning methodologies. Section 5 presents our experimental analysis and results. Finally, Section 6 concludes with a summary and discussion.

2 Related Works

2.1 Parameter Tuning for System

Parameter tuning is a common practice to optimize systems. For example, knob tuning in database systems, where specific values or knobs can be adjusted to optimize for specific query operators

and improve data access efficiency [50]. Traditional search strategies include using random and grid search to find the optimal parameters for a given workload. Advanced frameworks such as GPTune [25], Spearmint [33], and Sequential Model-Based Optimization (SMBO) [29] attempt to refine this process through Bayesian Optimization, integrating multi-task and transfer learning. However, these methods struggle with accuracy and computational efficiency, as they require starting anew with each shift in workload or data distribution. The inefficiencies amplify when faced with real-time tuning requirements for learned indexes [12].

2.2 Deep Reinforcement Learning for Parameter Tuning

Recently, Deep reinforcement learning (DRL) has shown promising results in optimizing complex systems under dynamic conditions [6, 15]. Notably, CDBTune [49] uses deep deterministic policy gradients (DDPG [24]) to automatically navigate and tune the high-dimensional continuous space of database configurations. It has shown the adaptability and efficiency of DRL methods over traditional tuning tools and expert DBA interventions for handling dynamic conditions. However, CDBTune cannot easily adapt to the unique challenges in index tuning, not present database configuration tuning. Index tuning requires rapid and precise adjustments without system resets or prolonged downtime while answering queries and updating records within dynamic data distributions and workload patterns. Misconfiguration has severe consequences for performance and must be avoided. LITUNE, on the other hand, efficiently tackles these challenges while keeping the full advantages of DRL tuning.

2.3 Learned Index Tuning

Learned indexes [21] replace traditional indexing algorithms (like B+Trees) with models that predict the approximate location of a key using the Empirical Cumulative Distribution Function (CDF), potentially reducing search operations. Research on both static [11, 19, 21, 34] and updatable indexes [10, 11, 14, 22, 23, 38, 44, 46, 48] demonstrates that performance heavily depends on data distribution, and to achieve the best performance, the indexes must be "tuned" according to the data distribution.

Currently, there are two approaches for considering data distributions when designing learned indexes: (1) exposing parameters for adjustment by users or future research, as seen in [11, 21, 48], and (2) implementing self-tuning mechanisms through cost models, utilized by indexes such as [10, 23, 46]. In both approaches, the default parameter settings are crucial, with index designers asserting that tuning is unnecessary for satisfactory performance. However, this assumption only holds when the empirical cost of the default parameters is effective [22, 26, 38], and for updatable learned indexes, the insert cost model must also be valid [10, 23]. Early studies like [36] use grid search to find optimal default parameters but fail to capture the complexities of tuning learned indexes. Automating this tuning process remains under-explored and is fundamental to our work with LITUNE. This challenge is exacerbated for indexes supporting dynamic workloads, which must be reorganized to maintain model accuracy as data distributions shift. Unlike traditional indexes, learned index parameters depend not only on dataset size but also on rapidly changing data distributions. Additionally, updatable indexes require structural modifications—such as gaps [10, 23], hierarchical structures [11, 45], and buffers [14, 23]—to mitigate the impact of distribution shifts. These complexities create intricate dependencies between parameters (as shown in Figure 1(a)), making automated tuning particularly challenging.

CDFShop [27] automates learned index optimization by fine-tuning cumulative distribution function (CDF) models specifically for the RMI index, adjusting high-level hyperparameters (e.g., model type, branching factors). However, it is confined to RMI, lacks safety-aware mechanisms, and relies on iterative, game-theoretic, and heuristic exploration methods that can be time-consuming

and less scalable for complex indexes or larger datasets. Consequently, we exclude CDFShop from our baselines, comparing LITUNE instead with more general Bayesian Optimization and heuristic-based methods. AirIndex [4] employs a graph-based approach to automatically tune learned index structures but often has limited exploration by focusing on top- K heuristics, incurring high overhead despite parallelization and complicating optimization in high-dimensional parameter spaces. Recently, Reinforcement Learning (RL) has been applied to tuning learned indexes: RusKey [28] optimizes LSM-tree-based key-value stores [37], marking the first RL-driven LSM-tree transformations under dynamic workloads. Unlike white-box cost models centered on I/O complexities, RusKey's black-box RL approach offers a holistic view similar to LITUNE [8], yet it tunes only a single parameter and thus cannot be directly used for general learned index tuning.

To clarify our selection criteria and highlight the key tuning methods utilized in prior research, we have summarized relevant studies in Table 1. We observe that irrespective of the specific index type being targeted, the underlying tuning methods are generally transferable across different contexts. Consequently, we have extracted these transferable methods and adopted them as baselines for our experiments. This approach allows us to systematically evaluate the effectiveness of our proposed tuning system against established methodologies.

Method	Parameter Selection/Tuning Method
B-trees [7]	Expert selection, Heuristics
RMI [21]	Expert selection, Heuristics
ALEX [10]	Heuristic cost model, Grid search
SWIX [23]	Heuristic cost model, Grid search
CARMI [48]	Expert selection, Hardware-aware
CDFShop [27]	Heuristics; Pareto front
AirIndex [4]	Heuristics, Graph-based search
RusKey [28]	Vanilla DRL (DDPG)
Ours	Safe RL approach

Table 1. Summary of Existing Indexing and Tuning Works

2.4 Index Advisor

Index advisors are tools used in physical database design to help administrators select optimal indexes at the schema level based on query workloads [3, 39]. They focus on recommending which columns or combinations of columns should be indexed to improve query execution times and overall system efficiency. Recent RL-based index selection methods [20, 32, 43, 51] continue this approach by automating the index selection process but still operate at the schema level. In contrast, our work operates on a fundamentally different level by tuning the internal parameters of learned index structures themselves rather than selecting which indexes to create. Learned indexes leverage machine learning models to predict data positions within a dataset [21], and their performance heavily depends on hyperparameters and model configurations. Traditional index advisors do not address the challenges associated with optimizing these internal parameters. Therefore, our DRL-based tuning framework enhances the performance of learned indexes through dynamic internal optimization, offering adaptability and efficiency improvements that traditional index advising methods do not provide. This distinction underscores that we are addressing a different problem space, focusing on the internal mechanics of learned indexes rather than on schema-level index selection.

3 LITUNE System

3.1 Motivation

Unlike existing index tuning works that concentrate on a limited set of observable parameters directly reflected in data structures, our work tackles the more complex challenge of tuning within a vast parameter space where parameters interact in non-independent and intertwined ways. This complexity demands stable and efficient tuning strategies, particularly in the context of online and continuous learning tasks. Since parameter metrics are difficult to capture and do not readily lend themselves to the integration of strong heuristics or expert knowledge, we focus on capturing stateful transitions and propose tailored end-to-end tuners.

Furthermore, navigating complex parameter spaces poses significant challenges for traditional search strategies and advanced model-based approaches [49]. Traditional search strategies (random search, grid search, heuristics) fall short in navigating the extensive parameter space of learned indexes, while advanced out-of-box tuning frameworks (e.g., SMBO [29]) require starting a new navigation cycle with each shift in workload or data distribution, making them resource-intensive. Furthermore, establishing universal states for different index structures with different parameter sets creates complications. In this regard, LITUNE is designed to offer online and stateful index tuning using deep reinforcement learning for dynamic workloads and provides fast and safe configurations for multiple learned indexes. Our approach mitigates the instabilities and aimless explorations that often accompany the use of generic, out-of-the-box tuning methods.

3.2 System Overview

At the core of LITUNE is reinforcement learning (RL), which enhances adaptability to dynamic workloads beyond the capabilities of traditional cost models. Specifically, we design a Markov Decision Process framework that involves the RL agent interacting with an environment to maximize rewards. The decisions are made based on observed *states* and chosen *parameters* [37], which, in this case, are the structural responses to shifting data distributions and tuned parameters, respectively.

LITUNE operates in two primary phases: the Training Stage and the Online Tuning Stage. In the Training Stage, we implement an efficient adaptive training pipeline to generate a generalizable pre-trained model. Once this model is deployed, it undergoes continuous fine-tuning in the Online Tuning Stage, ensuring it remains current and effective under evolving operational conditions. To avoid misconfigurations during tuning, LITUNE adopts a context-aware RL system that prevents early terminations, ensuring stability and speed throughout the process. Additionally, to keep the RL agent updated during online use, we employ an O2 system for ongoing training and adjustments.

3.3 Training Stage

Part A of Figure 3 depicts the initial generation of a pre-trained RL agent. This agent is crucial as it forms the foundation of our adaptive RL-based tuner, designed to efficiently handle diverse tuning scenarios right from deployment.

3.3.1 Offline Training Preparation. The foundational step in our approach involves preparing a Deep Reinforcement Learning (DRL) agent for optimizing learned index configurations. This process commences with the generation of varied datasets and query sets, establishing a diverse training environment. Each training episode, defined by a unique dataset-query set combination, allows the agent to explore a spectrum of configurations through a trial-and-error strategy, thereby accumulating a rich set of initial training data.

Training Data: Unlike supervised learning, our methodology relies on the automatic collection of training quintuples $\langle d, q, a, s, r \rangle$ by the RL agent. d denotes the dataset, q a set of queries, a the parameters for index construction, s the index states, and r the performance metric. This approach

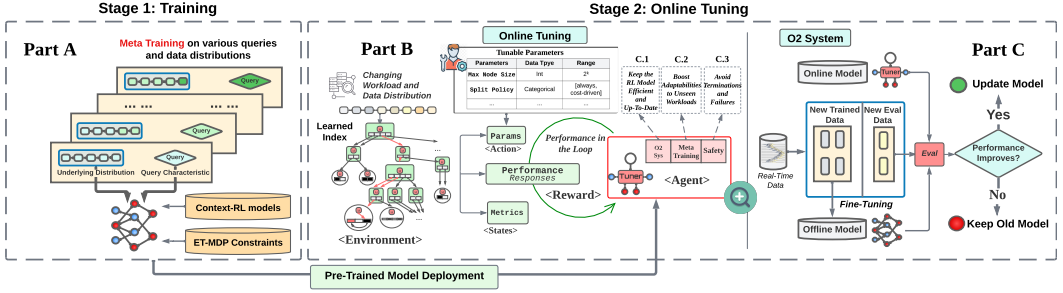


Fig. 3. The architecture of LITUNE. **Part A** illustrates the training phase, where RL-based models are trained. Once the training is complete, these models are deployed as online tuners in **Part B**. The operational details of the O2 system are explained in **Part C**.

ensures comprehensive feedback for the agent, which is pivotal for its learning process. Given the NP-hard nature of optimizing learned index configurations in a continuous parameter space, our model employs general DRL techniques. This choice enables the exploration of novel configurations and mitigates the risk of entrapment in local optima.

3.3.2 Adaptive Training (Meta Training) Design. As mentioned in Introduction, the **C.2** requires robust generalization of Deep Reinforcement Learning (DRL) agents in LITUNE to handle real-world scenarios with unseen queries and variable data distributions [47]. To achieve this, we introduce an adaptive training design based on Meta Reinforcement Learning (Meta-RL), which enhances the tuner’s adaptability to new situations and addresses.

Our method employs the Model-Agnostic Meta-Learning (MAML) approach [13], which trains agents to rapidly adapt with minimal updates. In the context of learned index tuning, the “tuning instances” represent specific scenarios with unique data distributions and query types. MAML integrates into the RL training process through a two-level training loop:

Inner Loop—Adaptation to Tuning Instances: Agents perform tuning-instance-specific updates to optimize performance on sampled tuning instances. This involves adjusting policy parameters based on interactions characterized by specific workload types and data distributions.

Outer Loop—Meta-Update for Generalization: The initial policy parameters are updated across all tuning instances to improve general performance. This enhances the agent’s ability to generalize to new, unseen tuning scenarios.

This dual-loop process enables the agent to handle individual tuning scenarios effectively while maintaining broad adaptability across diverse operational conditions, which is essential for efficiency in the dynamic landscape of learned index tuning. Here is a quick example on how it works in practice:

Example 3.1. Practice of Meta-RL

Consider two tuning instances for a learned index system:

- (1) *Instance A:* Primarily range queries on a uniformly distributed dataset.
- (2) *Instance B:* A mix of insert and range queries on a skewed, non-uniform dataset.

A traditional RL agent trained only on Instance A performs well for similar uniform range queries. However, when applied to Instance B, it struggles to maintain performance, requiring extensive retraining to adjust its policy. In contrast, a Meta-RL agent trained with the MAML approach has encountered diverse tuning instances during meta-training. When faced with Instance B, it can swiftly adapt its policy with just a few gradient updates, leveraging prior knowledge. This enables

the Meta-RL agent to maintain robust performance across varying scenarios without significant retraining. $\square$

3.4 Online Tuning Stage

After establishing a strong foundation in the training stage, we now transition to the online tuning stage, where the pre-trained model is put into practical use. This section describes how the LITUNE system operates and adapts to continuous tuning needs on the fly.

Parts B and C of Figure 3 illustrate the architecture of the online tuning system. The dashed box at the top represents the client and data storage system where end users send their queries to the LITUNE system below. This system is designed to handle the continuous adaptation required by varying data environments.

The operational flow is presented in Part B of Figure 3: Once the data storage setup and tuning requests are confirmed, the learned index and tuning system process the current data as the underlying distribution. The system executes queries on this data, generating performance data and states. Based on these observations, the tuning system automatically suggests adjustments to the learned index parameters to optimize future query handling and improve performance.

Furthermore, LITUNE leverages the pre-trained model to provide real-time recommendations for parameter settings during online tuning scenarios. It also continuously refines this model by incorporating feedback from each tuning request into its training process. This integration of Online Tuning and Offline Training (O2 System) ensures that LITUNE dynamically adapts to any changes in workload and data distribution, maintaining high efficiency and adaptability over time.

3.4.1 Online Tuning. End users can easily tune the target learned index by submitting a request to LITUNE. Upon receiving a request, the system gathers the necessary data and utilizes the pre-trained RL model for online tuning, ultimately recommending the best-performing parameters.

3.4.2 O2 System. As depicted in Part C of Figure 3, the O2 system integrates Online and Offline RL models to address [C.2](#), enhancing real-time adaptability and performance optimization. The system employs the online tuner with a pre-trained model for immediate index adjustments when no data changes occur. Conversely, significant data changes activate both the online and offline models: the offline model refines itself with new data, while the online model handles real-time optimizations.

The O2 system routinely assesses the necessity for model updates by comparing the online model's performance against new data and predefined criteria, including statistical divergence and user-defined thresholds. This assessment occurs at regular intervals or user-defined checkpoints, ensuring the O2 system remains responsive to changing data and workloads. During updates, the system carefully balances the current online model's performance with the insights gained from the offline model's continuous learning, thus maintaining optimal tuning efficacy. This efficiency also contributes to addressing [C.1](#).

It should be noted that the real-world queries encountered during online tuning might differ from those used in the offline training process. For this reason, this structure allows LITUNE to stay adaptive, leveraging the offline system's incremental fine-tuning for better alignment with real-world workloads, while the online system provides swift responses to immediate tuning needs. Here is a practical working example of O2 system:

Example 3.2. Running O2 System

Consider a learned index tuning scenario with two distinct phases of data workload:

- (1) *Stable Phase:* The dataset experiences minimal changes, and the workload consists mainly of read-heavy range queries.

- (2) *Dynamic Phase*: The dataset undergoes significant updates, including frequent insertions and deletions, and the workload shifts to a mix of read and write queries.

During the *Stable Phase*, the O2 system utilizes the online tuner with the pre-trained model to make immediate index adjustments, ensuring efficient query processing without the overhead of model retraining.

When transitioning to the *Dynamic Phase*, the significant data changes trigger the activation of both Online and Offline models. The offline model begins refining the tuning strategy by learning from the new data distribution and varied query patterns. Simultaneously, the online model continues to handle real-time optimizations to maintain query performance. $\square$

3.5 LITUNE Working Process

Summarizing the learned index parameter tuning in LITUNE, the learned index (tuning target) serves as the RL environment, while the deep RL model acts as the agent, recommending configurations based on the learned index state. As the index is constructed and queries are executed, the learned index state alters, reflected in the metrics. These metrics evaluate learned index performance, calculating the corresponding RL reward value, and the agent updates its policy accordingly, continuing until the tuning time budget is exhausted, ultimately revealing the most fitting parameter settings. Thus, the RL-based tuner can easily capture and memorize the state responses to the data distribution and workload shifts.

To illustrate how our RL-based tuner is specifically designed for learned index scenarios, Figure 4(a) demonstrates a tuning example using ALEX. As the workload transitions from a balanced to a write-heavy workload, LITUNE detects changes in ALEX's states and performance metrics. Initially, it increases the maximum node size from the default 16MB and reduces the threshold for out-of-domain inserts before triggering node expansion. This adjustment leads to a notable decrease in the "no_expand_and_retrain" metric. With positive feedback from the training environment, LITUNE stores new data and asynchronously fine-tunes the model using the O2 system, evaluating potential updates based on pre-defined criteria. This enables continuous learning with minimal disruption, allowing swift adaptation to workload changes. Eventually, LITUNE recommends further increasing the maximum node size to 64MB and raising the minimum number of out-of-domain inserts required before expansion for future trials. The Safe RL module ensures system safety by automatically preventing the selection of dangerous states when configuring more aggressive values for the maximum node size and minimum number of out-of-domain inserts. We avoid these aggressive settings, which may yield immediate rewards but could lead to system failure in the long run. While some learned indexes (like ALEX) require full reconstruction for structural parameter changes due to their codebase constraints, LITUNE provides flexible on-the-fly reconfiguration mechanisms across different index implementations. For cases where reconstruction is unavoidable, LITUNE minimizes overhead through efficient sampling: maintaining a small reservoir ($\approx 1\%$ of dataset) and proportionally selecting queries across workload types (read/write), only applying final configurations on the full dataset. This sampling strategy enables rapid performance estimation while preserving workload characteristics, with detailed cost analysis provided in Section 5.4.4.

4 Methodology

In this section, we explore the integration of novel Reinforcement Learning (RL) models to address the unique challenges of tuning learned indexes. While vanilla RL frameworks like Deep Deterministic Policy Gradient (DDPG) methods [24, 49] are adept at managing high-dimensional spaces and continuous actions, they often fall short in ensuring safety for tuning systems. To address these limitations, we have augmented the standard DDPG framework by incorporating context-aware

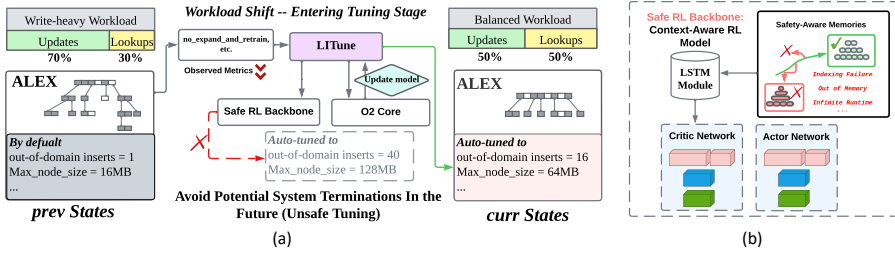


Fig. 4. (a) Running example of LITUNE. This example demonstrates how LITUNE’s tuner components respond to changes in performance metrics and adjust to workload shifts. (b) The safe-RL approach prevents aggressive tuning by learning from instabilities encountered during training.

learning. Notably, our online tuner components are designed with flexibility in mind and can be replaced by any existing vanilla DRL methods, showcasing our framework’s adaptability to accept similar DRL approaches. Detailed discussions on these DDPG framework enhancements are presented in the subsequent sections.

4.1 RL formalization for LITUNE

Using RL in LITUNE requires a nuanced formalization of learned index-tuning scenarios. Part B in Figure 3 illustrates the interaction diagram of LITUNE components and the functional workflow of LITUNE in practice.

Agent: Viewed as the tuning system, the agent receives rewards and states from the learned index, updating the policy to steer parameter adjustments towards higher rewards.

Environment: Representing the tuning target, the environment is an execution instance of the learned index.

State: In LITUNE, the state of the agent, denoted as s_t , represents the current condition of the learned index after applying recommended parameter settings. We categorize states into two main categories: structural and operational. Structural metrics might include features such as the number of internal nodes or tree height that represent the structure and measurable mechanisms of the index. However, not all features can be captured with just structural features. Therefore, we define a new category of operational metrics that captures which parameters affect the actual operation of the index. For each of the key operations (search, insert, update, and delete), we define a set of features that capture the cost of these operations using non-index-specific metrics. For example, we use search distance to capture the search cost, which is universal to all indexes regardless of the exact search algorithm used (linear, binary, or exponential). These metrics act as empirical proxies due to the complex inter-dependencies among features, offering a detailed snapshot of the index’s internal states necessary for effective tuning across various scenarios.

Tuning-oriented Reward Design. We define the reward r_t as a scalar that accounts for performance changes from both the initial baseline (D_0) and the immediately preceding step ($t - 1$). The idea is to capture the sort of incremental, forward-looking decisions that human experts make when tuning complex systems, balancing short-term gains with a broader trajectory of improvement. Concretely, the RL agent starts from an initial performance D_0 and seeks an improved state D_n . Once tuning begins, the system transitions to D_1 and the RL-agent computes $\Delta(D_1, D_0)$. From there, each iteration aims to surpass its predecessor, reflecting the principle that D_i should exceed D_{i-1} for all $i < n$. This design effectively encodes both immediate and foundational improvements in a single reward function.

Our focus metric for performance, denoted as R , signifies the end-to-end runtime, which is paramount for understanding and optimizing query performance. To track optimization progress, we define two key differential metrics: $\Delta = \begin{cases} \Delta_{t \rightarrow 0} = \frac{-R_t + R_0}{R_0} \\ \Delta_{t \rightarrow t-1} = \frac{-R_t + R_{t-1}}{R_{t-1}} \end{cases}$ Inspired by [49], the reward, r , is articulated as:

$$r = \begin{cases} ((1 + \Delta_{t \rightarrow 0})^2 - 1)^\omega (1 + \Delta_{t \rightarrow t-1})^\kappa, & \text{if } \Delta_{t \rightarrow 0} > 0 \\ -((1 - \Delta_{t \rightarrow 0})^2 - 1)^\omega (1 - \Delta_{t \rightarrow t-1})^\kappa, & \text{if } \Delta_{t \rightarrow 0} \leq 0 \end{cases}$$

Here, $\Delta_{t \rightarrow 0}$ and $\Delta_{t \rightarrow t-1}$ measure performance changes relative to the initial setting and the previous step. The scalar parameters ω (odd) and κ (even) control how strongly the reward emphasizes near-term versus longer-term improvements: ω governs the significance of changes from the initial baseline, and κ dictates the impact of recent performance gains or losses. By adjusting these scalars, practitioners can configure how aggressively the system pursues performance gains, how quickly it penalizes regressions, and how it balances near-term improvements against long-term objectives. In our practice, $\omega = 1$ and $\kappa = 2$ often strike a useful balance between achieving notable gains and maintaining stability.

Multi-objective trade-offs and achieving the tuning goal via RL. The reward function underpins our RL-based tuning process by connecting the search mechanism to diverse performance objectives. By maximizing the expected cumulative discounted reward, $\max_{\pi} \mathbb{E}_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t r_t \right]$, where π is the policy, τ the trajectory, γ the discount factor, and r_t the immediate reward, the RL agent explores parameter configurations aligned with user priorities. It refines its policy via trial-and-error (temporal-difference methods) to approximate expected returns across diverse states. Practitioners can adjust the performance metric R to emphasize or de-emphasize latency or throughput (e.g., $R = 0.8 \cdot \text{latency} + 0.2 \cdot \text{throughput}^{-1}$), steering the tuner's optimization without altering the underlying RL framework, ensuring that the tuner can effectively meet the final goal—be it latency-sensitive, throughput-oriented, or a balanced mix of both.

Action: Denoted as a_t , actions derive from the parameter configurations space and correspond to parameter tuning operations, influencing all tunable parameters concurrently.

Policy: Policy, mapping from state to action, maintains state transitions and is represented by a deep neural network. RL aims to learn the optimal policy.

4.2 Backbone: Safe RL approach for LITUNE

As mentioned in the introduction, we face the challenge of optimizing performance while ensuring safety in the context of tuning learned index [\(C.3\)](#), i.e., avoiding configurations that lead to system instability or failures such as out-of-memory errors or endless runtime. To address this, we propose a minimalist approach by employing an *Early Terminated Markov Decision Process* (ET-MDP) solver that triggers an *early termination* whenever the learning policy violates predefined constraints. Early termination has been previously used to improve sample efficiency in solving regular MDPs [42], as it accelerates learning by reducing the policy search space and shortening the time horizon. Moreover, an *ideal* policy should *never* violate the constraints, eliminating the need to learn to proceed or recover after violations.

To effectively handle the constraints in our tuning problem, we model it as a *Constrained Markov Decision Process* (CMDP) and transform it into its early terminated counterpart, the ET-MDP [35]. This transformation allows us to apply standard RL algorithms while ensuring safety through the early termination mechanism.

Definition 4.1 (Constrained Markov Decision Process). A *Constrained Markov Decision Process* (CMDP) is a deterministic MDP with a fixed horizon $H \in \mathbb{N}^+$, defined by the tuple $(\mathcal{S}, \mathcal{A}, H, r, c, \mathbb{T})$,

where $\mathcal{S}$ and $\mathcal{A}$ represent the state and action spaces; $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the reward function; $c : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the cost function representing constraints; $C \in \mathbb{R}^+$ is the upper bound on the permitted expected cumulative cost; and $\mathbb{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ is the transition function. The policy class Π consists of stationary policies $\pi : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$ such that $\sum_a \pi(a|s) = 1$ for all $s \in \mathcal{S}$.

In our learned index tuning problem, constraints such as out-of-memory errors and endless runtime define dangerous or constrained states, as illustrated in Figure 4(b). We incorporate these constraints into the CMDP framework by defining appropriate cost functions. Specifically, we assign costs (e.g., c_m for memory violations and c_r for runtime violations) which are set to 1 upon violation, ensuring that each type of violation contributes equally to the cumulative cost. This approach penalizes the policy for entering unsafe areas and guides it toward safer trajectories.

REMARK. *By incorporating system constraints into the cost function, we can effectively model the safe learned index tuning problem as a CMDP.*

To handle the constraints and ensure safety during the learning process, we transform the CMDP into an *Early Terminated MDP* (ET-MDP), which introduces an absorbing termination state whenever the cumulative cost exceeds a predefined threshold.

Definition 4.2 (Early Terminated MDP). For any CMDP, we define its *Early Terminated MDP* (ET-MDP) as a new unconstrained MDP $(\mathcal{S} \cup \{s_e\}, \mathcal{A}, H, r', \mathbb{T}')$, where s_e is the absorbing state after termination. The transition function and reward function in the ET-MDP are adjusted to handle terminations:

$$\begin{aligned}\mathbb{T}'(s, a) &= \mathbb{T}(s, a) \mathbb{1}(b_t \leq C) + s_e \mathbb{1}(b_t > C), \\ r'(s, a) &= r(s, a) \mathbb{1}(b_t \leq C) + r_e \mathbb{1}(b_t > C).\end{aligned}$$

where $b_t = \sum_{\tau=1}^t c_\tau^m + c_\tau^r$ records the cumulative costs up to time t , and $r_e \in \mathbb{R}$ is a small termination reward. The parameter C represents the total tolerated failures we can accept during training.

These adjustments integrate the costs into the ET-MDP framework to ensure that once a constraint is violated, the associated cost is counted and added to the cumulative costs. This guides the agent to avoid actions leading to high-penalty states, effectively steering the policy towards safer and more optimal trajectories.

REMARK. *By converting the CMDP into an ET-MDP and solving it using an appropriate RL algorithm, we can effectively ensure that the learned policy respects the constraints by avoiding actions that lead to early termination.*

Solving the ET-MDP: The transformation of the CMDP into an ET-MDP simplifies the problem by converting it into an unconstrained MDP where constraints are implicitly handled via early termination. The goal is to find an optimal policy π that maximizes the expected cumulative reward while respecting the constraints:

$$\max_{\pi \in \Pi} \mathbb{E}_{\tau \sim \pi, \mathbb{T}} \left[\sum_{t=1}^H r_t \right], \quad \text{s.t.} \quad \mathbb{E}_{\tau \sim \pi, \mathbb{T}} \left[\sum_{t=1}^H c_t \right] \leq C,$$

where $\tau = (s_1, a_1, r_1, \dots, s_H, a_H, r_H)$ represents the trajectory generated by policy π .

To solve this constrained optimization problem, the Lagrangian method relaxes it to an unconstrained one with a penalty term:

$$\pi^* = \arg \max_{\pi \in \Pi} \min_{\lambda \geq 0} \mathbb{E}_{\tau \sim \pi, \mathbb{T}} \left[\sum_{t=1}^H r_t - \lambda \sum_{t=1}^H c_t \right] + \lambda C, \quad (1)$$

Where $\lambda \geq 0$ is the Lagrangian multiplier. In practice, if the policy π is parameterized by θ , i.e., $\pi = \pi_\theta$, the optimization over θ and λ can be conducted iteratively through policy gradient ascent for θ and stochastic gradient descent for λ according to Eqn. (1).

However, as pointed out by [5], one possible defect of the Lagrangian methods is the violation of constraints during training, as the method may not strictly enforce the constraints at every iteration. This issue can be mitigated by incorporating *context models* [35].

Context models in an ET-MDP solver learn generalizable representations across similar tasks. In our setting, each state corresponds to a different task within the same distribution, allowing context models to transfer policies to unseen states and avoid constraint violations. By modeling tuning as a CMDP and transforming it into an ET-MDP, we naturally incorporate safe RL constraints into learned index tuning.

Implementation in LITUNE: In LITUNE, we handle constraints such as memory limits and runtime bounds by triggering early termination when these constraints are violated. The normal CMDP solver, integrated with the Deep Deterministic Policy Gradient (DDPG) algorithm enhanced with Long Short-Term Memory (LSTM), allows the system to manage large state and action spaces effectively. The LSTM module maintains context from past explorations, enabling the RL agent to adapt to dynamic data environments and explore safely while maximizing reward collection.

Figure 4 summarizes how the ET-MDP solver in LITUNE addresses safe tuning issues. The core strategy involves linking dangerous system *metrics* to the learned index *states* (understood as the structural information of the constructed index). The RL agent learns to avoid these dangerous states and maximize rewards within the given constraints during training. Once deployed, with memory units embedded in LSTM, the RL agent can autonomously identify safe tuning areas and achieve more reliable tuning performance. This strategy is significantly effective to handle the

C.3.

5 Experimental Study

In this section, we evaluate the performance of LITUNE in comparison to existing tuning approaches, focusing on two specific learned index instances across various workloads and datasets.

5.1 Key Insights

Before presenting our experimental results, we highlight several essential insights from our study:

(a) Tuning is essential: Effective default parameters are hard to set due to the complexity and variability of systems and problems. Unlike claims in [10, 48], default settings often fail to achieve optimal performance because they can't handle parameter inter-dependencies and dynamic data conditions. This highlights the crucial need for tuning to enhance learned index performance by synchronously adjusting parameters. (details in section 5.4.1)

(b) LITUNE is efficient: Contrary to the belief that deep models sacrifice tuning efficiency for quality, LITUNE uses online tuning methods to significantly improve efficiency. Experiments demonstrate that LITUNE outperforms other methods in performance gains, regardless of the tuning budget. (details in section 5.4.3, corresponding to C.1)

(c) LITUNE is adaptive: LITUNE effectively handles online and continuous tuning scenarios, adapting to different data distributions and workload dynamics. Its ability to adjust to varying tuning budgets showcases its robustness and applicability across diverse index types, proving its effectiveness under changing data conditions. (details in section 5.4.5, corresponding to C.2)

(d) LITUNE is safe: While setting safe ranges for individual parameters is simple, adjusting multiple parameters simultaneously poses safety challenges. LITUNE employs a context-aware

Table 2. Parameter space characteristics of learned indexes

Index	# Dims	Parameter Types
ALEX	14	• 5 Continuous [0,1] (density, workload ratios, etc.) • 3 Boolean (computation, splitting controls, etc.) • 4 Integer (node sizes, buffer limits, etc.) • 2 Discrete choice (tree policies, etc.)
CARMI	13	• 10 Continuous (operation timings, etc.) • 2 Integer (node sizes, etc.) • 1 Hybrid continuous/discrete (lambda, etc.)

strategy that links failure situations with states, using constraints and penalties to maintain parameters within safe zones. This approach ensures system stability and safety during online tuning, as demonstrated by our experimental results. (details in section 5.5.2, corresponding to [C.3](#))

5.2 Experimental Settings

5.2.1 Platform. Experiments are conducted on a machine equipped with an NVIDIA Quadro RTX 8000 GPU, an Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz, 8 vCPUs, and 64GB RAM. Since most of the parameter tuning methods have a certain degree of randomness, we repeated each experiment 5 times with different seeds.

5.2.2 Tuned Indexes and parameters. **ALEX** [10] is a dynamic, updatable learned index that allows tuning to optimize workload performance while not requiring parameters for basic operation. It offers statistics such as the number of model and data nodes, as well as expansions and splits, which serve as states for tuning parameters like read-write ratios, node sizes, and gap settings.

CARMI [48], based on RMI, enhances cache-awareness in learned indexes. It does not expose statistics directly but allows for state identification, including structural elements like leaf node counts and operational factors such as query visits and keys scanned. These states are tuned for parameters like leaf size and search operation weights, showcasing our method's capability to adapt tuning strategies to various learned index configurations.

We selected these two as our tuned instances because they exemplify two distinct parameterization mechanisms, each utilizing different optimization methods. Specifically, ALEX employs a heuristic-based cost model to automatically tune its parameters, whereas CARMI optimizes its parameters involving hardware-specific considerations.

Table 2 presents an overview of the parameter space in learned indexes, highlighting the marked differences in tuning complexity across existing methods. While other recent index tuning works typically handle only a few parameters—CDFShop [27] focuses on 2–4 RMI-specific parameters, RusKey [28] primarily tunes the compaction policy K in FLSM-Tree (fewer than three parameters), and AirIndex [10] employs layer-wise parameters which, despite their seemingly large number, are governed by strong hierarchical constraints, pruned to a top- K set, predominantly focused on I/O-aware optimizations, and limited to static workload conditions—our system manages a substantially larger optimization space of around 10-15 parameters.

The complexity of our parameter space, with its high dimensionality and mixed discrete-continuous variables, poses significant challenges for traditional optimization methods under limited tuning budgets. RL-based methods excel by learning parameter interactions through experience, enabling efficient global exploration while avoiding local optima.

5.2.3 Evaluation and Training Datasets. We evaluate LITUNE using the Search On Sorted Data (SOSD) benchmark suite [18], which includes datasets with up to 200 million 64-bit keys from various domains: Amazon books, OpenStreetMap (OSM), Facebook user IDs, and MIX (a combination of uniform, FB, books, and OSM distributions).

To prevent overfitting and ensure the tuner encounters unseen distributions during evaluation, we adopt strategies from prior RL-based training [28, 49]. Instead of training directly on SOSD, we generate synthetic 1-D key-value pairs with diverse distributions (e.g., uniform, beta, normal) and vary the Write-Read Ratio (W/R Ratio) between 1:10 and 10:1. These experiments test the generalization ability of the methods across diverse query patterns in various tuning scenarios.

5.2.4 Workloads. LITUNE evaluation highlights its versatility across static and dynamic datasets, establishing it as a robust tuning solution for learned indexes. All runtime performances depicted in the section's figures represent the average runtime for basic operations (INSERT, SEARCH, and DELETE) based on the Write-Read Ratio within workloads.

(a) Static Workload: We evaluate LITUNE on static datasets (OSM, books, Facebook, MIX) to demonstrate its adaptability to varied data distributions. Our setup uses 90M records from a 100M dataset², reserving 10M for a fixed distribution and 80M for INSERT/DELETE operations. Workloads differ by Write-Read Ratio (W/R): Balanced (1), Read-Heavy (1/3), and Write-Heavy (3). We evaluate tuning efficiency over varying steps (Figure 5) and also conduct "**extensive tuning**" (up to 50 steps or 1000 seconds) to gauge near-optimal performance (Figure 6).

(b) Data-shifting Workload: We divide 100M records into 30 tumbling windows [30], each with 1M base data and 8M updates. LITUNE frequently adjusts parameters to handle rapid data evolution, with at most 5 tuning steps or 100 seconds (whichever first), as shown in Figures 9 and 10³.

5.2.5 Evaluation Metrics. The key metric used for evaluating the different tuning methods is End-to-end runtime performance. We crafted a dual-faceted experimental setup to encompass both static and data-shifting cases. In addition to query runtime performance shown in figure 6, we also evaluate throughput in static settings, as shown in figure 7. It's worth noting that these throughput metrics were obtained under continuous system tuning requests, which introduces unique considerations that may account for the observed discrepancies with the runtime speed-up results. We also documented the effects and insights related to tuning the structural parameters of the learned index, which are detailed in section 5.4.1.

5.3 Baseline Methodologies

Our evaluation sought to underscore the efficacy of LITUNE by contrasting it with various established tuning methodologies. Initially, the *Default* method, utilizing unaltered system parameter settings from index designers or experts, set a fundamental baseline for comparative analysis. This is the adaptive configuration that responds to dynamic workloads and is designed by the authors of the corresponding indexes. *Random Search* scanned the parameter space indiscriminately, while *Grid Search* exhaustively tested predefined parameter combinations. Notably, *Grid Search* does not rely on expert defaults but instead uses a fixed parameter grid determined at the outset. *Heuristic Search*, implemented via a simulated annealing kernel from OpenTuner [1], pursued more focused exploration by leveraging existing expert heuristics. *Sequential Model-Based Optimization (SMBO)* employed the TPE method [2, 29], effectively managing complex parameter spaces (an approach similar to OtterTune [40]). Lastly, we incorporated a *vanilla DDPG-based tuner* [24], referred to as

²A 90M subset aligns with the *NIPS mlforays'19* fb distribution in the SOSD benchmark.

³These workload combinations reflect their worst performance under streaming constraints.

"DDPG", pretrained and fine-tuned with the same data as LITUNE, to demonstrate that direct RL pipelines from other domains (e.g., DBMS [49]) may offer moderate gains when embedded in our framework, but are less effective overall than LITUNE.

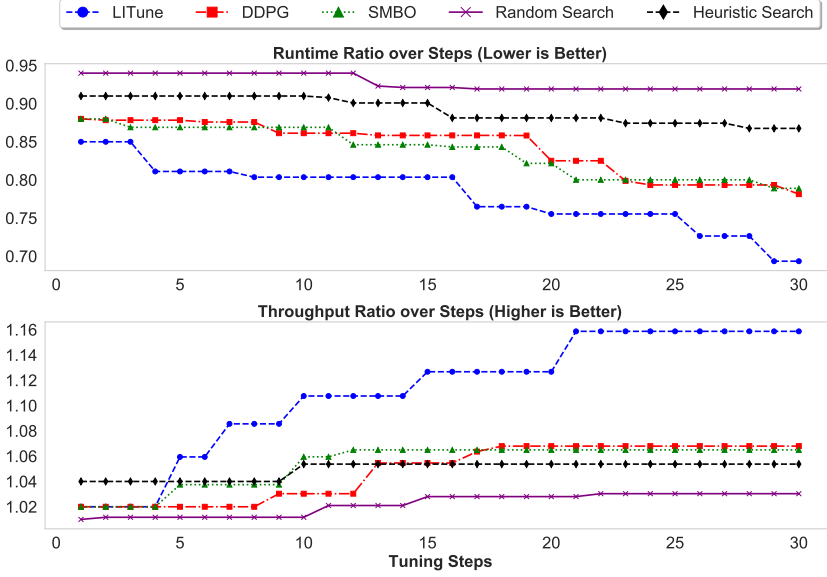


Fig. 5. Tuning efficiency–Performance as tuning steps increase. **Above:** runtime ratio (best found vs. default settings). **Below:** throughput ratio (best found vs. default settings).

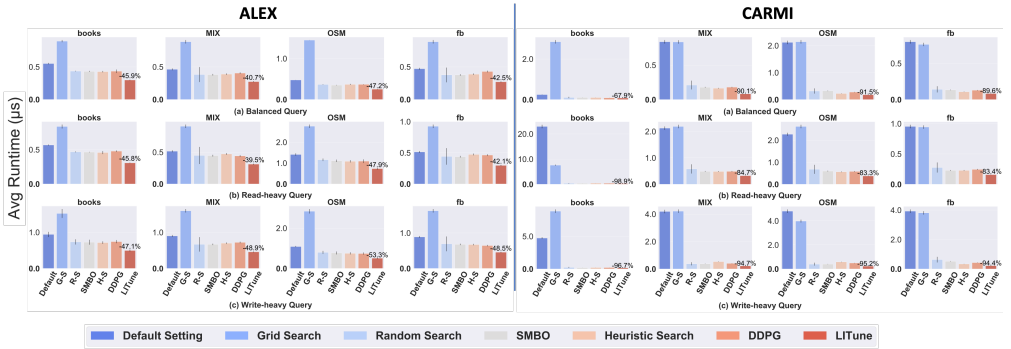


Fig. 6. Runtime performance (average on operation tuples) with extensive tuning. Performance improvements relative to default settings are marked.

5.4 Results Overview

Our comprehensive experimentation yielded results strongly in favor of LITUNE, with the method outperforming all baseline methods under varying conditions. Notably, LITUNE's advantage becomes especially pronounced whether the available tuning steps are restricted or extended, emphasizing its efficacy and efficiency in optimizing learned index types.

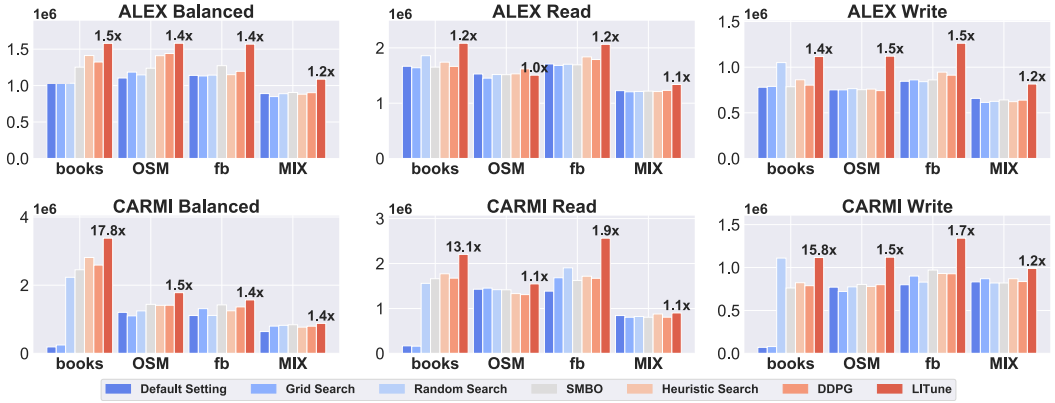


Fig. 7. Throughput performance (ops/sec) with extensive tuning. Performance improvements relative to default settings are marked.

5.4.1 Tuning results and takeaways. The values marked in Figure 6 & 7 show the performance gains compared against the default parameters as evidence of the effectiveness and need for a tuning system. We dive into some insight from the parameters of ALEX. Across all workloads, the thresholds for out-of-domain insertions exhibit significant increases compared to default parameters. Specifically, the minimum threshold shows an 80–100× increase and the maximum threshold ranges from 3–5×, which suggests that ALEX can achieve better by "buffering" out-of-domain keys in each node prior to expansion. Moreover, LITUNE adjusts workload-specific parameters such as the expected insertion fraction (otherwise fixed to a write-only workload) and toggles between approximate or exact model/cost computations. Specifically, ALEX benefits more from exact computations under balanced workloads, whereas read-heavy or write-heavy scenarios may favor approximate approaches. Additionally, LITUNE minimizes expansions and retrains, reducing retrains for the OSM dataset under a balanced workload from 7196 to nearly zero. This outcome aligns with recent findings [17, 36] emphasizing the overhead of retraining in learned indexes.

5.4.2 Radar Chart: Overall Evaluations among Tuning Methods. The showcase in Figure 8, based on 200 tuning trials using the MIX data distribution (the most complex) and a balanced workload on CARMi, quantitatively demonstrates the advantages of LITUNE. It provides a quick comparison of various tuning approaches across five attributes: Adaptability, Solution Quality, Stability, Tuning Efficiency, and Preparation Time. Scores are normalized on a scale from 0 to 9 to enable direct comparison. The results are illustrated in a radar chart, emphasizing the strengths and weaknesses of each method. Metrics are defined as: **Adaptability:** Lower runtime variances across scenarios indicates better adaptability. **Solution Quality:** Higher average runtime performance signifies superior results. **Stability:** More successful trials indicate higher stability. **Tuning Efficiency:** Higher performance-to-tuning budget ratios show better efficiency. **Preparation Time:** Shorter setup and pre-to-tuning time are preferred.

5.4.3 Comparative E2E Performance of LITUNE. LITUNE demonstrates strong End-to-End (E2E) performance across diverse datasets and workloads, leveraging Meta-RL, the ET-MDP solver, and O2 System components. As shown in Figure 5, on the MIX dataset with a Balanced workload on ALEX, it rapidly achieves over 60% of the optimal result within 10 steps, while others require at least twice as many. Its tuning efficiency—query runtime improvements per step—is about 2× that of

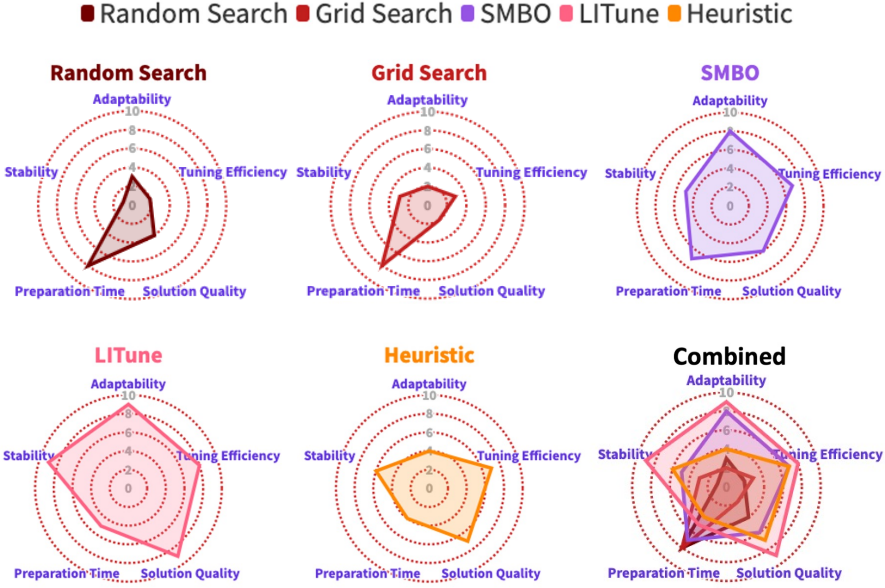


Fig. 8. Navigating the Parameter Tuning Landscape: LITUNE vs Other Tuning Methods

SMBO and DDPG, and $3\times$ that of random search, surpassing baselines from the very first recorded step. Only inference is required online, consuming just seconds per step on modern GPUs. Figure 6 highlights CARMI's notable optimization headroom, with over 90% runtime reduction (compared to 30–40% for ALEX). Complex datasets like MIX and OSM pose additional challenges, yet LITUNE consistently excels.

When looking into tuning methods, *Grid Search* and *Random Search* display contrasting outcomes. *Grid Search* consistently under-performs across all tuning budget scenarios, illustrating the challenges of exhaustively navigating extensive parameter spaces, and thus is not included in Figure 5. In contrast, *Random Search* occasionally reaches optimal or near-optimal configurations under more generous budgets despite its inherent performance variability. *SMBO*, while generally effective, can under-perform due to its reliance on historical evaluations, which may lead it into sub-optimal search areas influenced by noise or model discrepancies. Moreover, *SMBO* risks venturing into system risk areas, potentially wasting the tuning budget without efficient exploration. *Heuristic Search* maintains stable and reasonable performance but requires specific heuristic designs, which may not be universally applicable across all systems or scenarios. The performance of *Heuristic Search* is notable in the ALEX scenario due to effective heuristic discovery. Yet, it faces limitations in the CARMI scenario where appropriate insights into structures and parameter impacts are lacking.

Failure of DDPG-Tuner: As shown in Figure 6 and Figure 7, vanilla DDPG only matches the performance levels of SMBO and heuristic searches, lagging 10–15% behind LITUNE. The under-performance of a vanilla RL tuner using DDPG, compared to LITUNE, can be attributed to several key factors. Firstly, DDPG-Tuner often fails to capture the complex dependencies and dynamics within the tuning environment due to its simpler reinforcement learning model. This leads to underfitting, where the tuner is unable to generalize well from its training data to new or unseen scenarios. In contrast, LITUNE integrates advanced Meta-RL and the O2 system, which not only accelerates learning adaptations but also enhances its predictive accuracy by utilizing historical

Table 3. Training and tuning cost comparison across different methods. LITune-X denotes X% sampling ratio, and Col 3-Col 6 indicate tuning time required to achieve target performance improvements.

Method	Training	Tuning Time				Best Perf. (default 403s)
		-5%	-10%	-20%	-45%	
Grid Search	-	32m	1.0h	-	-	→ 360s
Heuristic Search	-	15s	35s	5m	-	→ 312s
SMBO	-	12s	25s	3m	-	→ 314s
DDPG ([24, 28])	12h	28s	35s	45s	-	→ 326s
LITune-0.1%	5h	6s	12s	18s	25s	→ 288s
LITune-1%(ours)	6h	8s	15s	22s	28s	→ 212s
LITune-10%	7h	12s	20s	26s	32s	→ 211s
LITune-Full	12h	18s	25s	32s	38s	→ 208s

tuning experiences. Furthermore, DDPG-Tuner lacks the advanced memory and learning mechanisms of LITUNE, making it less effective in handling diverse tuning tasks, leading to sub-optimal decisions and reduced performance in complex environments.

5.4.4 Tuning and Training Costs of LITUNE. While LITUNE offers significant performance improvements, managing training and tuning costs is essential for practical deployment. Under a scaling workload with 400M balanced queries on ALEX using OSM data, Table 3 presents a comprehensive comparison of tuning overhead across different methods. We evaluate LITUNE under different sampling rates to demonstrate the effectiveness of our sampling strategy. For instance, to achieve a 20% runtime reduction, LITUNE with 1% sampling requires only 22 seconds of tuning time, compared to 25 seconds for vanilla DDPG and several minutes to hours for traditional approaches. We choose the 1% sampling rate as it achieves nearly identical performance (212s vs 208s) to LITune-Full while having training and tuning overhead close to LITune-0.1%, which sacrifices too much performance (288s vs 212s). Our results also showcase the highest attainable performance for various tuning methods when given substantial tuning budgets (1000s), while noting *Grid Search*'s limitation of becoming computationally infeasible due to its expansive search domain.

The training phase represents a one-time investment in computational resources, where LITUNE develops generic tuning strategies across various index scenarios. As shown in Table 3, LITUNE's training time with 1% sampling (6 hours) is half that of DDPG used in RusKey [28] (12 hours), while providing superior tuning capabilities - achieving 47% runtime reduction compared to DDPG's 19%. Furthermore, our O2 System (detailed in Section 3.4.2) allows instant model updates for real-time applications without additional training overhead.

5.4.5 Adaptability of LITUNE. LITUNE consistently exhibits superior performance and adaptability across varied queries, data distributions, and data shifts, as demonstrated in Figures 6, 7, and 9.

This robust performance amid diverse and shifting scenarios can be ascribed to the foundational Deep Reinforcement Learning (DRL) framework, which enables LITUNE to continually refine its policies and swiftly adapt its parameter settings, ensuring optimal performance amidst varied query contexts and data scenarios. The LITUNE not only promotes intelligent and dynamic exploration and exploitation of the parameter space but also facilitates quick convergence to optimal or near-optimal configurations, making it especially adept at navigating complex, heterogeneous, and dynamically evolving data and query environments, thereby addressing [C.2](#). Specifically, LITUNE demonstrates:

1. *Consistency across Query Types*: Different rows in Figure 6 correspond to diverse query types (B, RH, WH) when read vertically. Notably, LITUNE's framework efficiently navigate through different query types, adapting its parameter settings in real time to ensure optimal performance amidst shifting query contexts.

2. *Versatility across Data Distributions*: Different columns in Figure 6 correspond to varied data distributions (OSM, books, fb, MIX) when read horizontally. LITUNE also showcases adaptability to varied data distributions, adeptly managing complex scenarios like the MIX distribution by autonomously identifying and applying optimal parameter configurations.

3. *Robustness amidst Data Shifts*: As evident in Figure 9, LITUNE sustains high performance across continuous data chunks in online tuning, quickly adapting to evolving data distributions without necessitating re-initialization. It is important to note that we did not include DDPG or vanilla-DRL methods for comparison here, as their extreme instability in handling dynamic scenarios without tailored designs made them unsuitable for this evaluation.

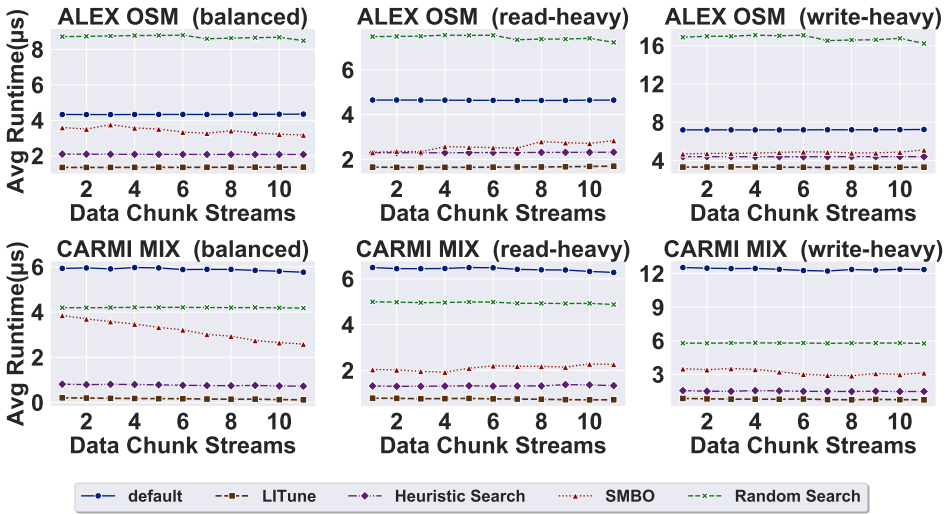


Fig. 9. Online and continuous tuning performance (averaged over operation tuples) in data streams (ALEX+OSM and CARMIX+MIX)

5.5 Ablation Study

5.5.1 *Effects of O2 System. Resilience to Unseen Data*: Figure 10 shows that the O2 system's online component quickly adapts to new trends using a sliding window of recent queries, enhancing resilience to unforeseen data and ensuring adaptability.

Handling Data Distribution Shifts: Demonstrated in the MIX dataset with ALEX (See Figure 10), the O2 system adeptly handles data distribution shifts. The offline tuner, enriched with diverse training data, robustly supports varied distributions. When the online model detects significant shifts, a threshold in the divergence measure initiates a model swap, ensuring continuous adaptation.

Consistency in Performance: As demonstrated in Figure 10, LITUNE equipped with the O2 system consistently outperforms versions without it. This superior performance is attributed to its hybrid approach: the online model rapidly adapts to new conditions, while the offline model, trained on a wider array of data scenarios, seamlessly maintains performance. This synergy ensures dependable and stable operation across diverse data environments.

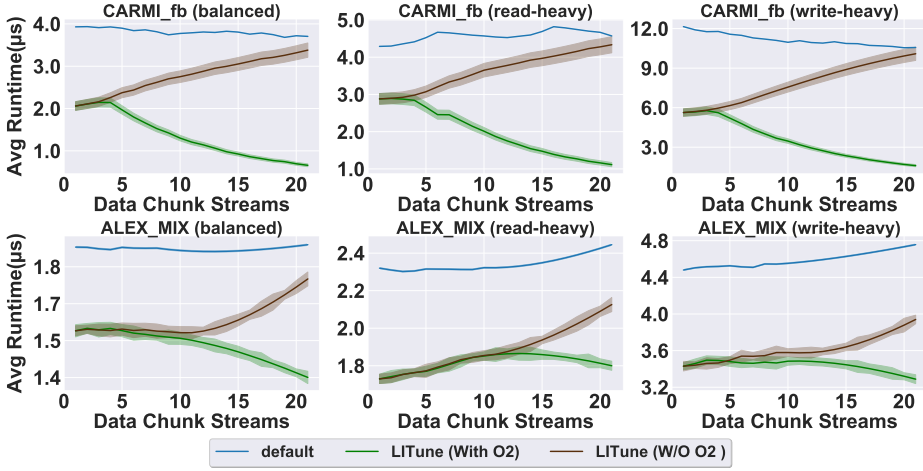


Fig. 10. Benefits of the O2 system in online and continuous tuning, comparing with pre-trained models W/O O2 system (CARMI+fb and ALEX+MIX), averaged over operation tuples.

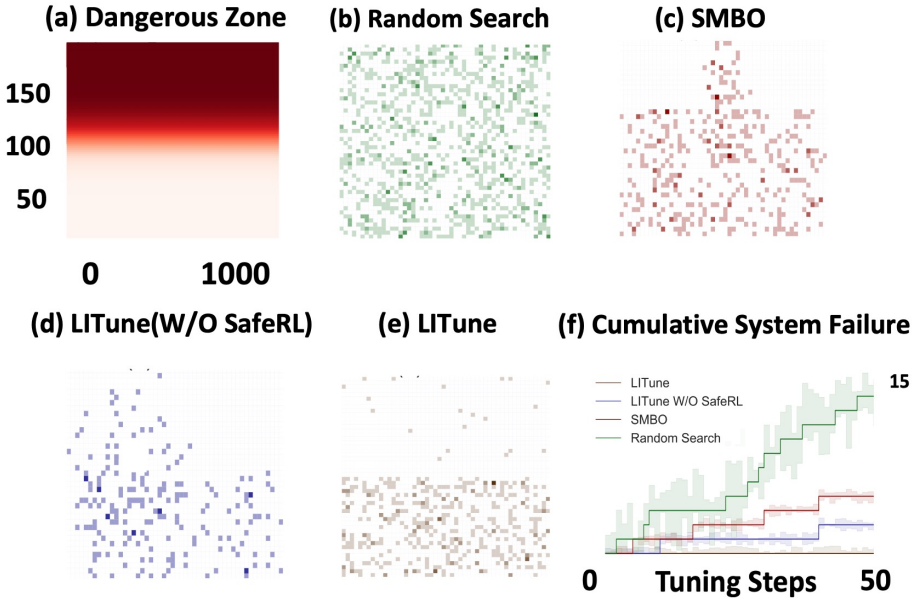


Fig. 11. Exploring Parameter Spaces Across Tuning Methods(ALEX + OSM + Balanced).

5.5.2 Safe Tuning. As mentioned in the Introduction, exploring parameter spaces can enhance system performance but also introduces risks, especially when parameters significantly impact system reliability and stability during tuning and operation (C.3).

Figure 12(a) demonstrates the importance of our Safe-RL module through training performance comparison. Without Safe-RL, the reward signals exhibit high volatility during the latter stages of training due to severe penalty signals triggered by system terminations (e.g., infinite loops, memory

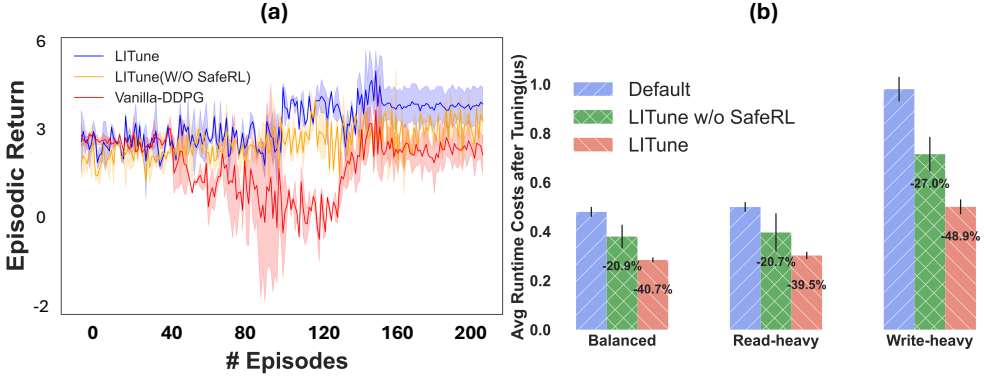


Fig. 12. (a) Training stability comparison between LITune variants with and without Safe-RL (ALEX). (b) End-to-end runtime performance of LITune variants after tuning on the MIX dataset for ALEX.

crashes) from aggressive parameter exploration. In contrast, LITUNE with Safe-RL maintains stable reward improvement throughout training, achieving both better final performance and more consistent learning progress compared to vanilla DDPG. Besides, within 200 epochs, we found that the vanilla DDPG cannot learn a converged policy and requires further training. This stability difference is particularly pronounced after the 100th training episode, where LITUNE without Safe-RL shows large reward fluctuations from frequent system failures, while the Safe-RL version maintains steady improvement by proactively avoiding parameter combinations that could lead to system termination. As shown in Figure 12(b), this training stability translates directly to better end-to-end performance from trained policies: In average, LITUNE with Safe-RL achieves 30% lower runtime with 60% less variance compared to the version without Safe-RL, demonstrating that safer exploration leads to both better and more reliable tuning outcomes.

Figure 11 further illustrates the benefits of our safety-aware design. Figures 11(a)–(e) show how four methods explore two critical parameters: *kMaxOutOfDomainKeys* and *kOutOfDomainToleranceFactor*, which are crucial under specific configurations (*fanoutSelectionMethod* = 1, *splittingPolicyMethod* = 1, *allowSplittingUpwards* = True). The red-highlighted **Dangerous Zone** in subfigure (a) marks parameter regions prone to causing system instabilities and degraded performance, emphasizing the need for a tuner like LITUNE. subfigure (f) presents the cumulative number of index system failures (e.g., infinite loops, memory issues) during tuning over five trials, highlighting the effectiveness of our safety design.

Random Search, lacking predictive or memory capabilities, explores the parameter space indiscriminately, entering risky zones and introducing instability into the learned index structures. *SMBO* incorporates a model to partially understand the parameter space, offering semi-protected exploration but still occasionally ventures into risky regions due to limited learning capacity. It tends to focus sampling in specific areas, sometimes within dangerous zones, increasing the risk of system instability. *LITUNE (without Safe-RL)*, lacking a context-aware design, also struggles with risky areas but performs better than other baselines because its reward function penalizes long runtimes, raising awareness of safe tuning regions.

In contrast, LITUNE embeds stability into exploration by leveraging historical knowledge to avoid the **Dangerous Zone**, ensuring reliable and stable index structures.

6 Discussion and conclusion

LITUNE is effective across diverse learned indexes, though its impact varies with parameter space size and interdependencies. By treating tuning as a black-box problem, it first identifies optimal configurations, then maps them back for analysis. Our studies indicate RL tolerates certain redundant parameters without harming performance, provided dimensionality stays manageable. While we currently rely on domain expertise to pinpoint key parameters, a systematic approach could automate this. Techniques like checkpointing and caching successful configurations further reduce retraining overhead as workloads evolve. In future work, preliminary sensitivity analysis can prune superfluous parameters for new or evolving indexes.

Furthermore, while LITUNE shows considerable promise in learned index optimization, several areas warrant further exploration. Like other RL-based methods, it occasionally faces convergence and stability issues during training, highlighting the need for future research to address these challenges to improve system reliability. Additionally, accelerating offline training (e.g., using parallelization) could further enhance the efficiency of LITUNE's deployment. Despite these challenges, LITUNE's innovative approach, combining Meta-RL and context-aware strategies, sets a new standard by improving system efficiency and query performance while ensuring stability through proactive risk management.

References

- [1] Jason Ansel, Shoaib Kamil, Kalyan Veeramachaneni, Jonathan Ragan-Kelley, Jeffrey Bosboom, Una-May O'Reilly, and Saman Amarasinghe. 2014. Opentuner: An extensible framework for program autotuning. In *Proceedings of the 23rd international conference on Parallel architectures and compilation*. 303–316.
- [2] James Bergstra, Dan Yamins, David D Cox, et al. 2013. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, Vol. 13. Citeseer, 20.
- [3] Surajit Chaudhuri and Vivek Narasayya. 1997. An Efficient Cost-Driven Index Selection Tool for Microsoft SQL Server. In *Proceedings of the 23rd International Conference on Very Large Data Bases (VLDB)*. 146–155.
- [4] Supawit Chockchowwat, Wenjie Liu, and Yongjoo Park. 2023. Airindex: versatile index tuning through data and storage. *Proceedings of the ACM on Management of Data* 1, 3 (2023), 1–26.
- [5] Yinlam Chow, Ofir Nachum, Edgar Dueñez-Guzmán, and Mohammad Ghavamzadeh. 2018. A lyapunov-based approach to safe reinforcement learning. *Advances in neural information processing systems* 31 (2018).
- [6] Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. 2018. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. *Advances in neural information processing systems* 31 (2018).
- [7] Douglas Comer. 1979. Ubiquitous B-tree. *ACM Computing Surveys (CSUR)* 11, 2 (1979), 121–137.
- [8] Niv Dayan and Stratos Idreos. 2018. Dostoevsky: Better space-time trade-offs for LSM-tree based key-value stores via adaptive removal of superfluous merging. In *Proceedings of the 2018 International Conference on Management of Data*. 505–520.
- [9] Jialin Ding, Ryan Marcus, Andreas Kipf, Vikram Nathan, Aniruddha Nrusimha, Kapil Vaidya, Alexander van Renen, and Tim Kraska. 2022. Sagedb: An instance-optimized data analytics system. *Proceedings of the VLDB Endowment* 15, 13 (2022).
- [10] Jialin Ding, Umar Farooq Minhas, Jia Yu, Chi Wang, Jaeyoung Do, Yinan Li, Hantian Zhang, Badrish Chandramouli, Johannes Gehrke, Donald Kossmann, et al. 2020. ALEX: an updatable adaptive learned index. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 969–984.
- [11] Paolo Ferragina and Giorgio Vinciguerra. 2020. The PGM-index: a fully-dynamic compressed learned index with provable worst-case bounds. *Proceedings of the VLDB Endowment* 13, 8 (2020), 1162–1175.
- [12] Matthias Feurer, Jost Springenberg, and Frank Hutter. 2015. Initializing bayesian hyperparameter optimization via meta-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.
- [13] Qiming Fu, Zhechao Wang, Nengwei Fang, Bin Xing, Xiao Zhang, and Jianping Chen. 2023. MAML2: meta reinforcement learning via meta-learning for task categories. *Frontiers of Computer Science* 17, 4 (2023), 174325.
- [14] Alex Galakatos, Michael Markovitch, Carsten Binnig, Rodrigo Fonseca, and Tim Kraska. 2019. Fiting-tree: A data-aware index structure. In *Proceedings of the 2019 international conference on management of data*. 1189–1206.
- [15] Sanket Kamthe and Marc Deisenroth. 2018. Data-efficient reinforcement learning with probabilistic model predictive control. In *International conference on artificial intelligence and statistics*. PMLR, 1701–1710.

- [16] Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, and Le Song. 2017. Learning combinatorial optimization algorithms over graphs. *Advances in neural information processing systems* 30 (2017).
- [17] Minsu Kim, Jinwoo Hwang, Guseul Heo, Seiyoon Cho, Divya Mahajan, and Jongse Park. 2024. Accelerating String-key Learned Index Structures via Memoization-based Incremental Training. *arXiv preprint arXiv:2403.11472* (2024).
- [18] Andreas Kipf, Ryan Marcus, Alexander van Renen, Mihail Stoian, Alfons Kemper, Tim Kraska, and Thomas Neumann. 2019. SOSD: A benchmark for learned indexes. *arXiv preprint arXiv:1911.13014* (2019).
- [19] Andreas Kipf, Ryan Marcus, Alexander van Renen, Mihail Stoian, Alfons Kemper, Tim Kraska, and Thomas Neumann. 2020. RadixSpline: a single-pass learned index. In *Proceedings of the third international workshop on exploiting artificial intelligence techniques for data management*. 1–5.
- [20] Jan Kossmann, Onur Mutlu, Scott Posner, and Felix Nöth. 2022. SWIRL: Selection of Workload-aware Indexes via Reinforcement Learning. In *Proceedings of the 2022 International Conference on Management of Data (SIGMOD)*. 1570–1583.
- [21] Tim Kraska, Alex Beutel, Ed H Chi, Jeffrey Dean, and Neoklis Polyzotis. 2018. The case for learned index structures. In *Proceedings of the 2018 international conference on management of data*. 489–504.
- [22] Pengfei Li, Yu Hua, Jingnan Jia, and Pengfei Zuo. 2021. FINEdex: a fine-grained learned index scheme for scalable and concurrent memory systems. *Proceedings of the VLDB Endowment* 15, 2 (2021), 321–334.
- [23] Liang Liang, Guang Yang, Ali Hadian, Luis Alberto Croquevielle, and Thomas Heinis. 2024. SWIX: A Memory-efficient Sliding Window Learned Index. *Proceedings of the ACM on Management of Data* 2, 1 (2024), 1–26.
- [24] TP Lillicrap. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
- [25] Yang Liu, Wissam M Sid-Lakhdar, Osni Marques, Xinran Zhu, Chang Meng, James W Demmel, and Xiaoye S Li. 2021. Gptune: Multitask learning for autotuning exascale applications. In *Proceedings of the 26th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. 234–246.
- [26] Baotong Lu, Jialin Ding, Eric Lo, Umar Farooq Minhas, and Tianzheng Wang. 2021. APEX: a high-performance learned index on persistent memory. *arXiv preprint arXiv:2105.00683* (2021).
- [27] Ryan Marcus, Emily Zhang, and Tim Kraska. 2020. Cdfshop: Exploring and optimizing learned index structures. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 2789–2792.
- [28] Dingheng Mo, Fanchao Chen, Siqiang Luo, and Caihua Shan. 2023. Learning to Optimize LSM-trees: Towards A Reinforcement Learning based Key-Value Store for Dynamic Workloads. *arXiv preprint arXiv:2308.07013* (2023).
- [29] Yoshihiko Ozaki, Yuki Tanigaki, Shuhei Watanabe, and Masaki Onishi. 2020. Multiobjective tree-structured parzen estimator for computationally expensive optimization problems. In *Proceedings of the 2020 genetic and evolutionary computation conference*. 533–541.
- [30] Kostas Patroumpas and Timos Sellis. 2006. Window specification over data streams. In *International Conference on Extending Database Technology*. Springer, 445–464.
- [31] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- [32] Tarique Siddiqui and Wentao Wu. 2024. ML-Powered Index Tuning: An Overview of Recent Progress and Open Challenges. *ACM SIGMOD Record* 52, 4 (2024), 19–30.
- [33] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* 25 (2012).
- [34] Mihail Stoian, Andreas Kipf, Ryan Marcus, and Tim Kraska. 2021. Towards Practical Learned Indexing. *arXiv preprint arXiv:2108.05117* (2021).
- [35] Hao Sun, Ziping Xu, Zhenghao Peng, Meng Fang, Taiyi Wang, Bo Dai, and Bolei Zhou. 2022. Constrained MDPs can be Solved by Eearly-Termination with Recurrent Models. In *NeurIPS 2022 Foundation Models for Decision Making Workshop*.
- [36] Zhaoyan Sun, Xuanhe Zhou, and Guoliang Li. 2023. Learned Index: A Comprehensive Experimental Evaluation. *Proceedings of the VLDB Endowment* 16, 8 (2023), 1992–2004.
- [37] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.
- [38] Chuzhe Tang, Youyun Wang, Zhiyuan Dong, Gansen Hu, Zhaoguo Wang, Minjie Wang, and Haibo Chen. 2020. XIndex: a scalable learned index for multicore data storage. In *Proceedings of the 25th ACM SIGPLAN symposium on principles and practice of parallel programming*. 308–320.
- [39] Grubb Valentin, Michael Zuliani, Diego Zilio, Guy M. Lohman, and Alan Skelley. 2000. DB2 Advisor: An Optimizer Smart Enough to Recommend Its Own Indexes. In *Proceedings of the 16th International Conference on Data Engineering (ICDE)*. 101–110.
- [40] Dana Van Aken, Andrew Pavlo, Geoffrey J Gordon, and Bohan Zhang. 2017. Automatic database management system tuning through large-scale machine learning. In *Proceedings of the 2017 ACM international conference on management of data*. 1009–1024.

- [41] Linnan Wang, Yiyang Zhao, Yu Jinnai, Yuandong Tian, and Rodrigo Fonseca. 2020. Neural architecture search using deep neural networks and monte carlo tree search. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 9983–9991.
- [42] Tingwu Wang, Xuchan Bao, Ignasi Clavera, Jerick Hoang, Yeming Wen, Eric Langlois, Shunshi Zhang, Guodong Zhang, Pieter Abbeel, and Jimmy Ba. 2019. Benchmarking Model-Based Reinforcement Learning. *arXiv preprint arXiv:1907.02057* (2019).
- [43] Zijia Wang, Haoran Liu, Chen Lin, Zhifeng Bao, Guoliang Li, and Tianqing Wang. 2024. Leveraging Dynamic and Heterogeneous Workload Knowledge to Boost the Performance of Index Advisors. *Proceedings of the VLDB Endowment* 17, 7 (2024), 1642–1654.
- [44] Jiacheng Wu, Yong Zhang, Shimin Chen, Jin Wang, Yu Chen, and Chunxiao Xing. 2021. Updatable learned index with precise positions. *arXiv preprint arXiv:2104.05520* (2021).
- [45] Jiacheng Wu, Yong Zhang, Shimin Chen, Jin Wang, Yu Chen, and Chunxiao Xing. 2021. Updatable learned index with precise positions. *arXiv preprint arXiv:2104.05520* (2021).
- [46] Guang Yang, Liang Liang, Ali Hadian, and Thomas Heinis. 2023. FLIRT: A Fast Learned Index for Rolling Time frames.. In *EDBT*. 234–246.
- [47] Wei Ying, Yu Zhang, Junzhou Huang, and Qiang Yang. 2018. Transfer learning via learning to transfer. In *International Conference on Machine Learning*. PMLR, 5085–5094.
- [48] Jiaoyi Zhang and Yihan Gao. 2021. Carmi: A cache-aware learned index with a cost-based construction algorithm. *arXiv preprint arXiv:2103.00858* (2021).
- [49] Ji Zhang, Yu Liu, Ke Zhou, Guoliang Li, Zhili Xiao, Bin Cheng, Jiashu Xing, Yangtao Wang, Tianheng Cheng, Li Liu, et al. 2019. An end-to-end automatic cloud database tuning system using deep reinforcement learning. In *Proceedings of the 2019 International Conference on Management of Data*. 415–432.
- [50] Xinyang Zhao, Xuanhe Zhou, and Guoliang Li. 2023. Automatic database knob tuning: a survey. *IEEE Transactions on Knowledge and Data Engineering* 35, 12 (2023), 12470–12490.
- [51] Wei Zhou, Chen Lin, Xuanhe Zhou, and Guoliang Li. 2024. Breaking It Down: An In-Depth Study of Index Advisors. *Proceedings of the VLDB Endowment* 17, 10 (2024), 2405–2418.

Received October 2024; revised January 2025; accepted February 2025