Meta-Learning via PAC-Bayesian with Data-Dependent Prior: Generalization Bounds from Local Entropy

Shiyu Liu[1,4], **Wei Shi**[1], **Zenglin Xu**[2,4,5*], **Shaogao Lv**[3], **Yehong Zhang**[4] and **Hui Wang**[4]
[1]University of Electronic Science and Technology of China
[2]Artificial Intelligence Innovation and Incubation (AI[3]) Institute, Fudan University
[3]Nanjing audit university
[4]Peng Cheng Laboratory
[5]Harbin Institute of Technology, Shenzhen
shyu.liu@foxmail.com, shiwei9624@gmail.com, zenglin@gmail.com
lvsg716@nau.edu.cn, {zhangyh02, wangh06}@pcl.ac.cn,

Abstract

Meta-learning accelerates the learning process on unseen learning tasks by acquiring prior knowledge through previous related tasks. The PAC-Bayesian theory provides a theoretical framework to analyze the generalization of meta-learning to unseen tasks. However, previous works still encounter two notable limitations: (1) they merely focus on the data-free priors, which often result in inappropriate regularization and loose generalization bounds; (2) more importantly, their optimization process usually involves nested optimization problems, incurring significant computational costs. To address these issues, we derive new generalization bounds and introduce a novel PAC-Bayesian framework for meta-learning that integrates data-dependent priors. This framework enables the extraction of optimal posteriors for each task in closed form, thereby allowing us to minimize generalization bounds incorporated data-dependent priors with only a simple local entropy. The resulting algorithm, which employs SGLD for sampling from the optimal posteriors, is stable, efficient, and computationally lightweight, eliminating the need for nested optimization. Extensive experimental results demonstrate that our proposed method outperforms the other baselines.

1 Introduction

Meta-learning [Thrun and Pratt, 1998; Baxter, 2000], or learning-to-learn, extracts prior knowledge from past training tasks to accelerate future learning. This process enables knowledge transfer across tasks, leveraging previous experience to eliminate the need for training models from scratch. Recent advances in meta-learning have achieved impressive success across various fields, including computer vision [Hospedales *et al.*, 2021], natural language processing [Lee *et al.*, 2022], and reinforcement learning [Zhao *et al.*, 2022]. Most current meta-learning methods [Finn *et al.*, 2017; Zhu *et al.*, 2023] presume access to a large pool of tasks, but this contrasts with the limited tasks typically available in real-world settings. This gap raises the risk of overfitting to the few tasks used during meta-training, potentially impairing performance on unseen target tasks. Hence, a critical challenge is how to effectively regularize the meta-learner to generalize well to unseen tasks.

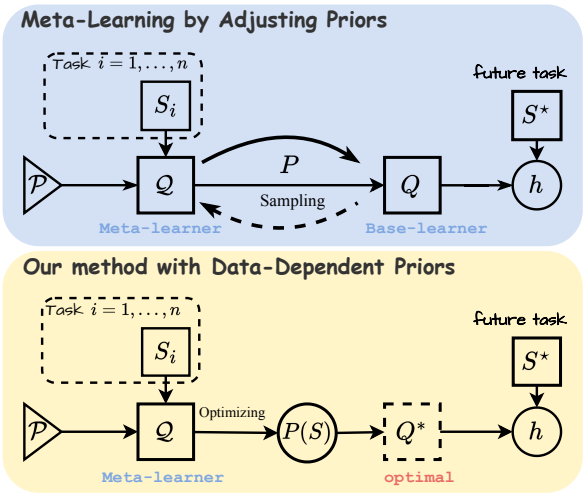

Figure 1: Existing methods typically learn posterior for each task and hyper-posterior simultaneously using a complex nested procedure. In contrast, our method directly optimizes for priors best satisfying the generalization needs with a simple local entropy objective, whose corresponding optimal posteriors can be extracted in closed form.

Fortunately, the PAC-Bayesian theory [McAllester, 1999] provides a rigorous framework for understanding the generalization performance of learners, which incorporates prior knowledge and implicitly regularizes the learners. Recently, there has been a surge in studying meta-learning with PAC-Bayesian theory [Amit and Meir, 2018; Guan *et al.*, 2022; Rezazadeh, 2022]. However, these approaches still face several limitations: (1) They primarily focus on data-free priors, which tends to restrict achieving tighter bounds. Typically, generalization bounds balance empirical risk against model

*Corresponding author

complexity, dominated by the gap between the prior and posterior. Priors that closely match the underlying task distribution can lead to tighter bounds. (2) Their optimization process is a complex and usually unstable procedure, fine-tuning posteriors that reflect task preferences without drifting too far from repetitive priors, and fitting the hyper-posterior to minimize the corresponding bound. Hence, the resulting algorithms rely on a complex nested optimization problem, incurring expensive computational costs. Here a question arises: *are there priors to regularize meta-learning, achieving a tighter generalization guarantee and simplifying the optimization process?*

To this end, we study the generalization property of meta-learning through the PAC-Bayesian perspective with *data-dependent priors*. First, we derive the relationship between the local entropy and the results in PAC-Bayesian learning theory, which helps us directly optimize the priors to adhere to a tighter upper bound without explicit posteriors. Although it is possible to acquire data-dependent priors by optimizing local entropy [Chaudhari *et al.*, 2017; Chaudhari *et al.*, 2019], such an approach is usually ineffective due to the restriction in PAC-Bayesian theory, which states that the prior cannot depend on the training samples. An intuitive solution to relax this restriction is the differential privacy (DP) mechanism [Dziugaite and Roy, 2018b]. DP enables the data-dependent priors to weakly depend on the data while ensuring they suffice for the valid generalized bounds. To this end, we extend the PAC-Bayesian analysis framework for meta-learning to the data-dependent priors under privacy constraints. We further refine our PAC-Bayesian bounds to directly relate to the local entropy of Bayesian learners, rather than explicit posteriors, thus getting around the need for complex nested optimization. Besides, we develop a novel meta-learning algorithm utilizing Stochastic Gradient Langevin Dynamics (SGLD) [Welling and Teh, 2011], which is straightforward to implement and simple to train. Experiments show that our approach yields competitive results compared to previous PAC-Bayesian meta-learning methods.

Contributions. The main contributions of this paper can be summarized as follows:

- We study the generalization of meta-learning through the PAC-Bayesian perspective incorporating *data-dependent priors* (Section 3).
- We provide a new framework for PAC-Bayesian meta-learning incorporating *data-dependent priors*. Our PAC-Bayesian generalization bound is defined as a function of the local entropy, without the reliance on explicit posteriors for each task, which ultimately attains fast convergence ability (Section 3.2).
- With our bounds, we develop an efficient meta-learning algorithm grounded on SGLD, namely PAC-MLE. This algorithm eliminates the necessity of nested optimization, by effectively and implicitly sampling from the Gibbs optimal poster (Section 4).
- Numerical experiments on both classification and regression tasks demonstrate that our method outperforms previous theoretical work in terms of fast adaptation and generalization guarantee, and effectively impedes over-fitting (Section 5).

1.1 Related Work

In this subsection, we briefly overview the main related work. We provide a more detailed literature review in Appendix B.

Meta-learning [Thrun and Pratt, 1998] enables knowledge transfer across tasks, leveraging prior experience to eliminate the need for training models from scratch. Various methods include developing a shared embedding space across tasks [Vinyals *et al.*, 2016; Xu *et al.*, 2020], learning the initialization for quick adaptation [Finn *et al.*, 2017; Rothfuss *et al.*, 2019], or incorporating probabilistic modeling for uncertainty quantification [Finn *et al.*, 2018; Chen and Chen, 2022]. Despite their ability to learn complex patterns, these methods often require many meta-training tasks and lack performance guarantees. To address this issue, we study the generalization of meta-learning through the PAC-Bayesian perspective and introduce a novel meta-learning method with generalization guarantees.

The PAC-Bayesian learning theory [McAllester, 1999] provides upper bounds on the generalization error that holds with an arbitrarily high probability. Further works [Pentina and Lampert, 2014; Amit and Meir, 2018; Liu *et al.*, 2021; Guan *et al.*, 2022; Rezazadeh, 2022] extend previous bounds to the scenario where priors are meta-learned. Liu *et al.* [2021] propose to use held-out data to construct data-dependent priors. However, these methods are considerably complex as they leave both the hyper-posterior and posterior unspecified. Rothfuss *et al.* [2021; 2023] derive the PAC-optimal hyper-posterior but overlook data-dependent priors, limiting the potential for achieving tighter bounds. In contrast, we directly optimize data-dependent priors to best meet generalization preferences using a simple local entropy, which greatly simplifies the learning process.

2 Preliminaries: PAC-Bayesian Framework

Background. In the common setting, a learning task τ is characterized by an unknown distribution $\mathcal{D}$ over space $\mathcal{Z}$. We are given a set of m observations, $S = \{z_j\}_{j=1}^m$, with each $z_j \sim \mathcal{D}$. In supervised learning, the observations $z_j = (x_j, y_j)$ comprise input features $x_j \in \mathcal{X}$ and target labels $y_j \in \mathcal{Y}$. Given S, our goal is to find a hypothesis $h \in \mathcal{H}$, which can minimize the *expected error* over $\mathcal{D}$, $\mathcal{L}(h, \mathcal{D}) = \mathbb{E}_{z\sim\mathcal{D}}[\ell(h, z)]$, where $\ell : \mathcal{H} \times \mathcal{Z} \to \mathbb{R}^+$ is a loss function. Since $\mathcal{D}$ is unknown in general, we typically evaluates the *empirical error* instead, $\widehat{\mathcal{L}}(h, S) = \frac{1}{m}\sum_{j=1}^m \ell(h, z_j)$.

In the PAC-Bayesian framework, we are concerned with *randomized predictors* i.e., probability measures on the hypothesis space $\mathcal{H}$. This allows us to reason about the uncertainty of h, resulting from the fact that only a finite number of data points are available for training. Our goal is to output a posterior $Q := Q(S, P) \in \mathcal{M}(\mathcal{H})$, where $P \in \mathcal{M}(\mathcal{H})$ represents the prior knowledge and $\mathcal{M}(\mathcal{H})$ denotes the set of probability measures over $\mathcal{H}$. Then, we can define the *expected error* of posterior Q as $\mathcal{L}(Q, \mathcal{D}) = \mathbb{E}_{h\sim Q}\mathcal{L}(h, \mathcal{D})$ and its *empirical* counterpart as $\hat{\mathcal{L}}(Q, S) = \mathbb{E}_{h\sim Q}\hat{\mathcal{L}}(h, S)$.

PAC-Bayesian bounds. In practice, $\mathcal{L}(Q, \mathcal{D})$ is typically unknown as well. Thus, it is common to minimize $\hat{\mathcal{L}}(Q, S)$ instead. If there are no structural assumptions with additional

constraints, this often results in overfitting and limited generalization performance. Hence, it is important to analyze the difference between the empirical estimate and its expected one. PAC-Bayesian theory helps us do so by bounding the unknown generalization error based on its empirical estimate:

Theorem 1 (McAllester's single-task PAC-Bayesian bound [McAllester, 1999])**.** *Given a data distribution $\mathcal{D}$, prior $P \in \mathcal{M}(\mathcal{H})$, for any confidence level $\delta \in (0, 1]$, the following inequality holds uniformly for all posteriors $Q \in \mathcal{M}$ over samples $S \sim \mathcal{D}^m$, with probability at least $1 - \delta$,*

$$\mathcal{L}(Q, \mathcal{D}) \leq \widehat{\mathcal{L}}(Q, S) + \sqrt{\frac{\mathrm{KL}(Q\|P) + \log(m/\delta)}{2(m-1)}}. \quad (1)$$

Generally, PAC-Bayesian bounds attempt to balance the trade-off between empirical error (fitting error) and the complexity term (deviation from prior), which promotes the choice of a simple hypothesis. The choice of prior distribution P not only impacts the tightness of PAC-Bayesian bounds, but also notably affects its performance. Priors that closely match the true data-generating process can lead to better bounds.

3 PAC-Bayesian Bounds for Meta-Learning

In this section, we present our main theoretical results. We begin with a general PAC-Bayesian bound for meta-learning and discuss its implications. We then introduce the local entropy and show its connection with the PAC-Bayesian framework. Based on this, we extend the general PAC-Bayesian meta-learning bound to incorporate data-dependent priors. The overview of our proposed framework is illustrated in Fig. 2. The corresponding proofs can be found in Appendix F.

3.1 General PAC-Bayesian Meta-Learning Bounds

In meta-learning settings, we aim to acquire such a prior P in a data-driven manner through a set of related learning task $\{\tau_i\}_{i=1}^n$, which can speed up the learning process on new target tasks $\tau \sim \mathcal{T}$. All of these tasks $\tau_i := (\mathcal{D}_i, S_i)$ share the same sample space $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, hypothesis space $\mathcal{H}$ and employ the same loss function $\ell(h, z)$. However, they may differ regarding the (unknown) data distributions $\mathcal{D}_i$. For simplicity, we can assume that m_i is constant across tasks, $m = m_i, \forall i$. Moreover, every task $\tau_i \sim \mathcal{T}$ is considered to be drawn i.i.d. from an environment $\mathcal{T}$.

To extract knowledge from the observed datasets, the goal of learning focus shifts from a singular hypothesis $h \in \mathcal{H}$ to the prior $P \in \mathcal{M}(\mathcal{H})$. Here, we assumes a hyper-prior $\mathcal{P} \in \mathcal{M}(\mathcal{M}(\mathcal{H}))$, i.e., a distribution over priors P. We combine the hyper-prior $\mathcal{P}$ with the datasets $\{S_i\}_{i=1}^n$ from multiple tasks, and consequently output a hyper-posterior $\mathcal{Q}$ over priors which can then be used for new tasks. Accordingly, the quality of the hyper-posterior can be measured by the expected loss of learning new tasks using priors P drawn from $\mathcal{Q}$, the so-called *transfer error*:

$$\mathcal{L}(\mathcal{Q}, \mathcal{T}) := \mathbb{E}_{P\sim\mathcal{Q}}\left[\mathbb{E}_{(\mathcal{D},m)\sim\mathcal{T}}[\mathbb{E}_{S\sim\mathcal{D}^m}[\mathcal{L}(Q(S, P), \mathcal{D})]]\right].$$

While the environment τ is unknown in practice, we can estimate it using the *empirical multi-task error*,

$$\widehat{\mathcal{L}}(\mathcal{Q}, S_{1:n}) := \mathbb{E}_{P\sim\mathcal{Q}}\left[1/n \sum\nolimits_{i=1}^{n} \widehat{\mathcal{L}}(Q(S_i, P), S_i)\right].$$

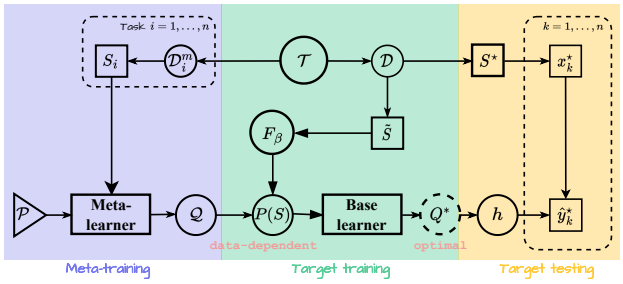

Figure 2: Overview of our framework for PAC-Bayesian meta-learning with data-dependent prior.

Building upon existing bounds, which affine-transformation steps at both the meta and task levels, we present a fast-rate bound for meta-learning,

Theorem 2 (Fast-rate PAC-Bayesian meta-learning bound)**.** *Let $Q : \mathcal{Z}^m \times \mathcal{M} \to \mathcal{M}(\mathcal{H})$ be a base learner, $\mathcal{P}$ be some fixed hyper-prior. Given any confidence level $\delta \in (0, 1]$, and $\lambda_t, \lambda_e \in (0, 2)$, the following inequality holds uniformly for all hyper-posteriors $\mathcal{Q}$, with probability $1 - \delta$,*

$$\mathcal{L}(\mathcal{Q}, \mathcal{T}) \leq \frac{\widehat{\mathcal{L}}(\mathcal{Q}, S_{1:n})}{(1-\lambda_e/2)(1-\lambda_t/2)} + \frac{\mathrm{KL}(\mathcal{Q}\|\mathcal{P}) + \ln(2/\delta)}{n\lambda_e(1-\lambda_e/2)}$$
$$+ \frac{1}{n}\sum_{i=1}^{n} \frac{\mathrm{KL}(\mathcal{Q}\|\mathcal{P}) + \mathbb{E}_{\mathcal{Q}}[\mathrm{KL}(Q_i\|P)] + \ln(2n/\delta)}{m\lambda_t(1-\lambda_t/2)(1-\lambda_e/2)}, \quad (2)$$

where $Q_i := Q(S_i, P)$.

Remark. *Theorem 2 indicates that the transfer error is bounded by the empirical multi-task error under $\mathcal{Q}$, plus the KL-divergence terms which serve as regularizers on the meta-level and task-level. This implies that regularization should be strong with limited data, but can vanish asymptotically as $n, m \to \infty$. Our fast-rate bound exhibits a convergence rate $\mathcal{O}(1/m + 1/n)$.*

The PAC-Bayesian theory provides a framework for understanding meta-learning generalization, leading to algorithms with generalization guarantees through corresponding upper bounds. However, previous methods [Amit and Meir, 2018; Rezazadeh, 2022] often struggle with data-free priors. In practice, the mismatch between prior and posteriors often results in inefficiency estimates and slow convergence.

3.2 PAC-Bayesian Meta-Learning Bounds with Data-dependent Priors

To overcome the drawbacks of data-free priors, a straightforward approach is to derive data-dependent priors $P(S_i)$ by reserving part of the training data specifically for this purpose [Thiemann *et al.*, 2017; Liu *et al.*, 2021; Dziugaite *et al.*, 2021]. Although it is possible to apply priors that are data-dependent but independent from the training set, such approaches are ineffective due to the intrinsic slow convergence and potential risks of overfitting. An alternative is to employ differentially private (DP), [Dwork *et al.*, 2015] mechanisms, which offer a compromise by ensuring priors are only weakly dependent on the data, thereby sufficient to the PAC-Bayesian bounds

under the relaxed restrictions [Dziugaite and Roy, 2018a; Dziugaite and Roy, 2018b; Rivasplata *et al.*, 2020].

In this way, the prior captures the underlying data distribution instead of being limited to the specific dataset it is trained on, maintaining resilience against data perturbations. This facilitates the somewhat loose but reliable bounds. Hence, we establish the PAC-Bayesian analysis framework for meta-learning, wherein a data-dependent prior is incorporated.

Theorem 3 (PAC-Bayesian meta-learning bound with ϵ-DP data-dependent priors)**.** *Under the same settings of Theorem 2, the following inequality holds with probability at least* $1-\delta$*,*

$$\mathcal{L}(\mathcal{Q},\mathcal{T}) \leq \frac{\widehat{\mathcal{L}}(\mathcal{Q},S_{1:n})}{(1-\lambda_e/2)(1-\lambda_t/2)} + \frac{\mathrm{KL}(\mathcal{Q}\|\mathcal{P})+\ln(2/\delta)}{n\lambda_e(1-\lambda_e/2)} + \frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{KL}(\mathcal{Q}\|\mathcal{P})+\mathbb{E}_{\mathcal{Q}}[\mathrm{KL}(Q_i\|P(S_i))]+C(m,\epsilon,\delta)}{m\lambda_t(1-\lambda_t/2)(1-\lambda_e/2)}, \tag{3}$$

where $C(m,\epsilon,\delta)=\ln 2\sqrt{m}+2\max\{\ln(3/\delta), m\epsilon^2\}$.

Remark. *Compared with the result in Theorem 2, the third term incurs an additional loss regarding private mechanisms. The bound Eq.* (3) *could significantly outperform the previous ones when* $\epsilon \in \mathcal{O}(m^{-1/2})$ *are relatively mild and matches the rate at which the KL term decays [Dziugaite and Roy, 2018b].*

3.3 PAC-Bayesian Meta-Learning Bounds with Data-dependent Priors from Local Entropy

We will show that the priors that achieve a tighter generalization bound can be optimized exactly with a simple local entropy objective, greatly simplifying the learning pipeline. Before we look into the specific solution to data-dependent priors, let us first explore the posterior Q which can help us to minimize the generalization bound in Eq. (3).

Intuitively, Catoni [2007] indicates the closed-form solutions for such a minimization problem $\beta\widehat{\mathcal{L}}(Q,S_i)+\mathrm{KL}(Q||P)$,

$$\begin{aligned} Q_i^\star(h) :&= \arg\min_{Q\in\mathcal{M}(\mathcal{H})}\ \beta\widehat{\mathcal{L}}(Q,S_i)+\mathrm{KL}(Q||P) \\ &= P(h)\exp\big(-\beta\widehat{\mathcal{L}}(h,S_i)\big)/Z_\beta(S_i,P), \end{aligned} \tag{4}$$

where $Z_\beta(S_i,P)=\int_{\mathcal{H}}P(h)\exp(-\beta\widehat{\mathcal{L}}(h,S_i))\mathrm{d}h$ is the partition function, and P, S_i, m, δ are defined same in Eq. (3). The corresponding distribution $Q_i^\star$, denoted as $Q_i^\star = P_{\exp(-\beta\widehat{\mathcal{L}}(h,S_i))}$[1], is also known as the *optimal Gibbs posterior* [Catoni, 2007; Lever *et al.*, 2013].

Then, plugging $Q_i^\star$ back to Eq. (4), we have

$$\min_Q \beta\widehat{\mathcal{L}}(Q,S_i)+\mathrm{KL}(Q\|P) = -\ln\int_{\mathcal{H}}P(h)e^{-\beta\widehat{\mathcal{L}}(h,S_i)}\mathrm{d}h. \tag{5}$$

Note that Eq. (5) establishes the connections between the PAC-Bayesian framework and the partition function w.r.t. the prior P. This suggests that the minimum of upper bounds essentially relies on the extent to which the prior can reflect the data. Recall that the hypothesis h is parameterized by $\mathbf{w}$ such that $\widehat{\mathcal{L}}(h,S_i):=\widehat{\mathcal{L}}(\mathbf{w},S_i)$. Like previous works [Amit and Meir, 2018], assume that $\mathbf{w}$ has prior $P(\mathbf{w})=\mathcal{N}(\mathbf{w}|\mathbf{v},\sigma^2)$ and rewrite the right hand side (RHS) of Eq. (5),

$$F_\beta(P,S_i) = \ln\int\mathcal{N}(\mathbf{w}|\mathbf{v},\sigma^2)e^{-\beta\widehat{\mathcal{L}}(\mathbf{w},S_i)}\,\mathrm{d}\mathbf{w} \tag{6}$$

$$= \ln\int e^{-\beta\widehat{\mathcal{L}}(\mathbf{w},S_i)-\|\mathbf{w}-\mathbf{v}\|^2/2\sigma^2}\,\mathrm{d}\mathbf{w}. \tag{7}$$

It is worth noting that Eq. (7), known as *local entropy* [Chaudhari *et al.*, 2017], is frequently used to measure flatness since it can measure the amount of "good" parameters $\mathbf{w}$ around given $\mathbf{v}$. The resulting algorithm [Chaudhari *et al.*, 2019; Pittorino *et al.*, 2021] prioritizes exploring flat regions, minimizing negative local entropy rather than the original loss function $\widehat{\mathcal{L}}(\mathbf{w},S_i)$, to enhance generalization performance. Thus, the connection between the generalization bound and local entropy can be reformulated as,

Lemma 4 ([Dziugaite and Roy, 2018a])**.** *Assume some constant* $\sigma^2>0$*. Then, the weights* $\mathbf{v}$ *that maximize the local entropy* $F_\beta(P,S_i)$ *in Eq.* (7)*, also minimize* $\beta\widehat{\mathcal{L}}(Q_i^\star,S_i)+\mathrm{KL}(Q_i^\star\|P)$*, where* $Q_i^\star = P_{\exp(-\beta\widehat{\mathcal{L}}(\mathbf{w},S_i))}$ *and* $P=\mathcal{N}(\mathbf{v},\sigma^2)$.

Lemma 4 indicates that maximizing local entropy allows us to optimize the prior within the PAC-Bayesian bound. In other words, the acquired $\mathbf{v}$ can serve as the mean of the data-dependent prior $P(S_i)$. Recall that Theorem 3 holds for any base learner $Q(S_i,P)$ with data-dependent priors for optimal performance. By Eq. (5), we can reformulate Theorem 3 in terms of local entropy, which yields

Corollary 5. *Under the same settings of Theorem 3, assume an optimal Gibbs posterior* $Q_i^\star$ *as the base learner, the following inequality holds with probability at least* $1-\delta$*,*

$$\begin{aligned} \mathcal{L}(\mathcal{Q},\mathcal{T}) \leq \frac{1}{n}\sum\nolimits_{i=1}^{n}\frac{1}{\lambda\beta}\mathbb{E}_{P\sim\mathcal{Q}}\big[-F_\beta(P,S_i)\big] \\ +\Big(\frac{1}{n\alpha}+\frac{1}{\lambda\beta}\Big)\mathrm{KL}(\mathcal{Q}\|\mathcal{P})+C(\delta,\lambda,\alpha,\beta,\epsilon), \end{aligned} \tag{8}$$

where $\lambda=(1-\lambda_t/2)(1-\lambda_e/2), \alpha=\lambda_e(1-\lambda_e/2), \beta=m\lambda_t$*, constant* $C(\delta,\lambda,\alpha,\beta,\epsilon)$ *is deferred to proofs in Appendix.*

Remark. *Corollary 5 provides a tighter bound than previous works, where bounds hold for any* $Q\in\mathcal{M}(\mathcal{H})$*. It is worth mentioning that Eq.* (8) *eliminates the explicit dependence on* $Q(S_i,P)$*, and turns the nested optimization problem into a standard stochastic optimization problem. This bound consists of the expected local entropy under the hyper-posterior* $\mathcal{Q}$ *and the KL-divergence term, serving as meta-level regularization. As the number of training tasks* n *increases, the emphasis on the KL term diminishes, aligning with the general understanding that more data reduces the need for regularization, eventually becoming negligible as* $n,m\to\infty$.

In practice, achieving strong DP can be computationally intractable. However, Dziugaite and Roy [2018b] relax the privacy requirement and show that convergence in distribution to a DP mechanism is adequate for ensuring generalization.

[1]Formally, a distribution like $P_{\exp(-\beta g)}$ is called a Gibbs distribution with energy function g and inverse temperature β.

This insight permits the use of SGLD [Welling and Teh, 2011]), which under certain conditions, is known to weakly converge to its target distribution. Consequently, we can formulate an optimization method for Meta-learning with PAC-Bayesian guarantee that incorporates data-dependent priors derived from local entropy.

4 PAC-Bayesian Meta-Learning Algorithm

In this section, we discuss how to translate our theoretical findings into a practical algorithm. We begin with the objective function from the upper bound in Corollary 5 and proceed to detail the derivation of low-variance gradient estimators.

For simplicity, we focus on priors $P_{\mathbf{v}_i}(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\mathbf{v}_i, \sigma_i^2)$, where σ_i^2 is a predefined constant. Then, we limit the space of hyper-prior and hyper-posteriors as a family of isotropic Gaussian distributions: $\mathcal{P}(\mathbf{v}) := \mathcal{N}(0, \sigma_{\mathcal{P}}^2)$ and $\mathcal{Q}_\nu(\mathbf{v}) = \mathcal{N}(\mathbf{v}|\nu, \sigma_{\mathcal{Q}}^2)$, where $\sigma_{\mathcal{P}}^2, \sigma_{\mathcal{Q}}^2 > 0$ are some predefined constants.

The objective function for the RHS of Eq. (8) reads as

$$J(\nu) = \frac{1}{n}\sum\nolimits_{i=1}^{n} \frac{1}{\lambda\beta}\mathbb{E}_{\mathbf{v}_i\sim\mathcal{Q}_\nu}\big[-F_\beta(P_{\mathbf{v}_i}, S_i)\big] + \big(\frac{1}{n\alpha} + \frac{1}{\lambda\beta}\big)\mathrm{KL}(\mathcal{Q}_\nu\|\mathcal{P}). \tag{9}$$

To reduce variance, we use the reparameterization trick [Kingma and Welling, 2013], and rewrite $\mathbf{v}_i := f(\nu, \varepsilon) = \nu + \sigma_i\varepsilon$ with $\varepsilon \sim \mathcal{N}(0, 1)$. Hence, the corresponding gradients follow:

$$\nabla J(\nu) \approx \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\lambda\beta}\big[-\nabla F_\beta(P_{\mathbf{v}_i}, S_i)\big] + \Big(\frac{1}{n\alpha} + \frac{1}{\lambda\beta}\Big)\frac{1}{\sigma_{\mathcal{P}}^2}\nu. \tag{10}$$

Gradient Estimation. Then we can compute the derivative of $F_\beta(P_{\mathbf{v}_i}, S_i)$ w.r.t. $\mathbf{v}_i$ by using methods such as REINFORCE [Williams, 1992]:

$$\nabla F_\beta(P_{\mathbf{v}_i}, S_i) = \mathbb{E}_{Q(S_i, P_{\mathbf{v}_i})}\big[(\mathbf{v}_i - \mathbf{w})/\sigma_i^2\big], \tag{11}$$

where $Q(S_i, P_{\mathbf{v}_i}) \propto \mathcal{N}(\mathbf{w}|\mathbf{v}_i, \sigma_i^2)\exp(-\beta\widehat{\mathcal{L}}(\mathbf{w}, S_i))$ is the optimal Gibbs posterior.

Eq. (11) does not have closed-form solutions when $\widehat{\mathcal{L}}(\mathbf{w}, S_i)$ is not a squared loss, due to the high-dimensional integrations for the neural network. In practice, we draw samples $\mathbf{w}_i \sim Q(S_i, P_{\mathbf{v}_i})$ through Monte Carlo sampling methods, i.e., SGLD, which gives

$$\hat{\nabla} F_\beta(P_{\mathbf{v}_i}, S_i) \approx \big(\mathbf{v}_i - \mu^K\big)/\sigma_i^2, \tag{12}$$

where $\mu_i^k = \alpha\mathbf{w}_i^k + (1-\alpha){\mu_i}^{k-1}$, $\mu_i^1 = \mathbf{w}_i^1$, and $\{\mathbf{w}_i^k\}_{k=1}^K$ are the (approximately) i.i.d. samples from $Q(S_i, P_{\mathbf{v}_i})$, $\alpha \in (0, 1)$ denotes the average weight. With the help of injected noise in each update step, SGLD can explore regions beyond local optima and converge asymptotically to the global minimum for non-convex objectives [Raginsky *et al.*, 2017]. It is also worth noting that the sequence of $\{\mathbf{w}_i^k\}_{k>K}$ converges to Gibbs distribution $Q(S_i P_{\mathbf{v}_i})$ with sufficiently large K, under the conditions that $\sum_k^\infty \eta_k' \to \infty$ and $\sum_k^\infty {\eta_k'}^2 \to 0$.

Algorithm 1 PAC-MLE algorithm: meta-training phase

Require: hyper-prior $\mathcal{P}$, datasets $\{S_i\}_{i=1}^n$, learning rate η, η', weight average α, initialize $\nu^0 \sim \mathcal{P}$
for $t = 1, \ldots, T$ **do**
  sample $|I_t|$ tasks uniformly at random, $I_t \subseteq [n]$
  $\eta_t \leftarrow \varphi_\eta(\eta_{t-1})$ ▷ learning rate decay
  for $i = 1, \ldots, |I_t|$ **do**
    $\mathbf{v}_i \sim \mathcal{Q}_{\nu^t}$, $\mathbf{w}_i^0, \mu_i^0 \leftarrow \mathbf{v}_i$
    for $k = 1, \ldots, K$ **do**
      Sample a mini-batch $S_i' \subset S_i$
      $\eta_k' \leftarrow \varphi_{\eta'}(\eta_{k-1}')$ ▷ learning rate decay
      $\nabla F(\mathbf{w}_i^{k-1}) \leftarrow -\beta\frac{1}{|S_i|}\sum\nabla\widehat{\mathcal{L}}(\mathbf{w}_i^{k-1}) - (\mathbf{w}_i^{k-1} - \mathbf{v}_i)/\sigma_i^2$
      $\mathbf{w}_i^k \leftarrow \mathbf{w}_i^{k-1} + \eta_k'\nabla F(\mathbf{w}_i^{k-1}) + \sqrt{2\eta_k'}\zeta^{k,i}$
      $\mu_i^k \leftarrow \alpha\mu_i^{k-1} + (1-\alpha)\mathbf{w}_i^k$ ▷ moving average
    end for
  end for
  $\nu^t \leftarrow \nu^{t-1} + \eta_t\frac{1}{|I_t|\lambda\beta}\sum_i(\mathbf{v}_i^{t-1} - \mu_i^K)/\sigma_i^2 - \eta_t\frac{1}{\sigma_{\mathcal{P}}^2}(\frac{1}{n\alpha} + \frac{1}{\lambda\beta})\nu^{t-1} + \sqrt{2\eta_t}\xi^t$
end for

It is computationally expensive to learn meta-information from all the tasks when the number n of tasks is large. Thus, we can even use mini-batching on the task level. For each iteration t, we can sample a mini-batch of tasks, which are indexed by $I_t \subseteq [n]$. Subsequently, the update rule w.r.t. ν is expressed as follows:

$$\nu^t = \nu^{t-1} - \eta_t\Big(\frac{1}{\lambda\beta|I_t|}\sum\nolimits_i^{|I_t|}\big[-\hat{\nabla}F_\beta(P_{\mathbf{v}_i}, S_i)\big] + \big(\frac{1}{n\alpha} + \frac{1}{\lambda\beta}\big)\frac{1}{\sigma_{\mathcal{P}}^2}\nu^{t-1}\Big) + \sqrt{2\eta_t}\xi^t. \tag{13}$$

where η_t is learning rate at t, and $\xi^t \sim \mathcal{N}(0, 1)$ is isotropic Gaussian noise injected during the update.

Algorithm. The Pseudocode of the proposed meta-learning algorithm is summarized in Algorithm 1. At the beginning of each round t, we sample a subset of tasks I_t from the task sets. During the task update phase, each sampled task runs SGLD for K steps using its data, followed by updating its state concerning priors with learning rate η_t'. In the subsequent meta-parameter update phase, we aggregate the updates of the sampled tasks and then update the meta-information by Eq. (13) with learning rate η_t. Here, $\varphi_\eta(\cdot)$ and $\varphi_{\eta'}(\cdot)$ denote the learning rate decay function (e.g., cosine decay).

It is worth mentioning that previous works [Amit and Meir, 2018; Liu *et al.*, 2021; Rezazadeh, 2022] adopt a similar variational form of our theoretical bound, in which they set factorized Gaussian distribution $\mathcal{Q}_\theta = \mathcal{N}(\theta, I_{2d})$ and $Q_{\phi_i} = \mathcal{N}(\mu_i, \sigma_i^2)$ as the hyper-posterior and the posterior of task τ_i, respectively. In addition, they utilize $\mathcal{P} = \mathcal{N}(0, I_{2d})$ for the hyper-prior distribution and $P_\theta = \mathcal{N}(\mu_P, \sigma_P^2)$ for the prior distribution, where $(\mu_P, \sigma_P^2) \sim Q_\theta$. They then optimize the empirical multi-task error in addition to the regularization terms (e.g., Theorem 2) w.r.t. parameters $\theta \in \mathbb{R}^{2d}$ and $\phi_i \in \mathbb{R}^{2d}$ using SGD. In contrast, our algorithm avoids such

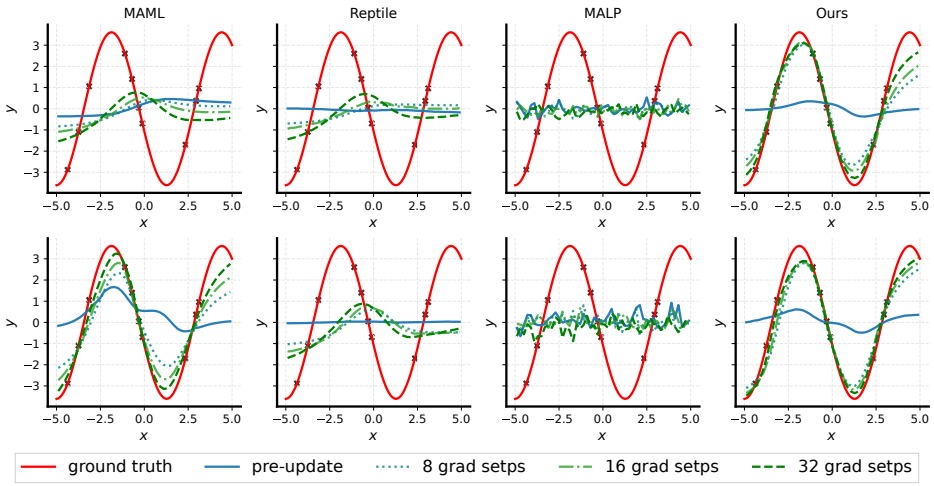

Figure 3: Results on few-shot adaptation for the simple regression task in the Sinusoids environment after 300 (top). vs 3000 (bottom) rounds of training.

a complex optimization process, reducing the required optimization parameters by a quarter. Instead, we employ SGLD to simulate the distribution.

Hence, one key advantage of our approach, compared to previous works, is its ability to simplify the nested optimization problem into a simpler stochastic optimization problem. This simplification not only enhances the stability of meta-learning, but also improves its scalability. Specifically, explicit computation of the task posteriors $Q(S_i, P)$ is unnecessary; instead, mini-batching over tasks can be employed. On the other hand, our approach incorporates principled meta-level regularization in the form of the hyper-prior $\mathcal{P}$. This regularization effectively addresses overfitting to the meta-training tasks. Consequently, our approach enables successful meta-learning with a limited number of tasks, contrasting with the majority of popular meta-learners [Finn *et al.*, 2017; Finn *et al.*, 2018; Nichol *et al.*, 2018] that heavily rely on numerous tasks to achieve generalization at the meta-level.

5 Numerical Experiments

In this section, we validate our theoretical results and demonstrate the efficiency of the algorithm through numerical experiments, including tasks related to few-shot regression and image classification. We defer complete setup details and further experiments to Appendix C.

5.1 Toy Example

Setup. This subsection considers a simple regression problem on a synthetic dataset. Each task involves regressing from the input to the output of a sine wave, with varying amplitude and phase [Finn *et al.*, 2017; Finn *et al.*, 2018]. To solve each regression task, we employ a fully connected network with 2 hidden layers, each with 64 hidden units. The goal is to recover the underlying periodic structure of unseen sine waves from a few observations and previous tasks. We set the same experiment setup as that in [Finn *et al.*, 2017]. We can evaluate the quality of the structure recovered by measuring the root mean squared errors (RMSE) over the novel tasks.

Results. The Fig. 3 illustrates the qualitative performance of some baselines. MAML nearly recovers the underlying function but needs more rounds r and is slower than our method. MLAP, hindered by a small m, fails to capture underlying patterns, highlighting a gap between theory and practice The meta-knowledge extracted by our method captures more comprehensive task environment information and can be easily adapted to new tasks.

5.2 Few-shot Regression

Setup. In this subsection, we consider regression experiments in one synthetic and four real-world meta-learning environments. The synthetic environment involves regression on 2-dimensional mixture of Cauchy distributions plus a random Gaussian Processes function. We utilize datasets corresponding to various calibration sessions of the Swiss Free Electron Laser (SwissFEL) [Milne *et al.*, 2017]. Additionally, we employ data from the PhysioNet 2012 challenge [Silva *et al.*, 2012], consisting of time series of electronic health measurements from patients in intensive care. Specifically, this includes the Glasgow Coma Scale (GCS) and the hematocrit value (HCT). In this context, each task corresponds to a different patient. Finally, we make use of the Intel Berkeley Research Lab temperature sensor dataset (Berkeley-Sensor) [Madden, 2004], where the tasks involve auto-regressive prediction of temperature measurements from sensors placed in various locations throughout the building.

Results. Table 1 shows that our method achieves comparable results across five regression environments. Specifically, it can obtain the best test errors in two regression environments and competitive errors in the remaining three. This demonstrates that the introduced meta-learning framework is not only theoretical sound, but also yields competitive empirical performance in practice.

Method	Cauchy	SwissFel	Physionet-GCS	Physionet-HCT	Berkeley-Sensor
BNN [Liu and Wang, 2016]	0.327(0.008)	0.529(0.022)	2.664(0.274)	3.938(0.869)	0.109(0.004)
MAML [Finn *et al.*, 2017]	0.219(0.004)	0.730(0.057)	1.895(0.141)	2.413(0.113)	0.045(0.003)
BMAML [Yoon *et al.*, 2018]	0.225(0.004)	0.577(0.044)	1.894(0.062)	2.500(0.002)	0.073(0.014)
MLAP [Amit and Meir, 2018]	0.219(0.004)	0.486(0.026)	2.009(0.248)	2.470(0.039)	0.050(0.005)
PACOH [Rothfuss *et al.*, 2021]	**0.195(0.001)**	0.372(0.002)	1.561(0.061)	**2.405(0.017)**	**0.043(0.001)**
λ-bound [Liu *et al.*, 2021]	0.227(0.002)	0.490(0.036)	1.575(0.203)	2.435(0.043)	0.057(0.003)
classic bound [Rezazadeh, 2022]	0.230(0.002)	0.593(0.036)	1.629(0.136)	2.487(0.098)	0.069(0.011)
PAC-MLE (ours)	0.203(0.014)	**0.365(0.017)**	**1.513(0.073)**	2.429(0.006)	**0.044(0.003)**

Table 1: Comparison of meta-learning algorithms in terms of test RMSE in 5 regression environments. Reported are mean and standard deviation across 5 seeds.

Method	100 Pixels Swaps		200 Pixels Swaps		Permuted labels	
	Test Bound	Test Error (%)	Test Bound	Test Error (%)	Test Bound	Test Error (%)
MAML [Finn *et al.*, 2017]	N/A	1.876(0.001)	N/A	2.241(0.002)	N/A	26.50(0.018)
[Pentina and Lampert, 2014]	0.190(0.022)	1.939(0.001)	0.240(0.030)	2.631(0.002)	6.026(0.436)	15.660(0.063)
MLAP [Amit and Meir, 2018]	0.126(0.012)	1.587(0.001)	0.197(0.019)	1.948(0.001)	2.834(0.075)	8.571(0.004)
PACOH [Rothfuss *et al.*, 2021]	0.174(0.023)	1.921(0.001)	0.224(0.030)	2.634(0.001)	5.434(0.416)	12.520(0.061)
λ-bound [Liu *et al.*, 2021]	0.067(0.015)	1.630(0.001)	0.151(0.015)	2.097(0.001)	3.830(0.181)	11.340(0.017)
quadratic bound [Liu *et al.*, 2021]	0.085(0.033)	1.590(0.001)	0.150(0.030)	1.944(0.001)	5.280(0.778)	12.580(0.013)
λ bound [Guan *et al.*, 2022]	0.055(0.006)	1.643(0.001)	0.114(0.029)	1.937(0.001)	3.502(0.237)	12.930(0.022)
quadratic bound [Guan *et al.*, 2022]	0.081(0.018)	1.624(0.001)	0.163(0.020)	2.000(0.001)	5.306(0.338)	12.750(0.023)
kl-bound [Guan and Lu, 2022]	0.075(0.013)	1.610(0.001)	0.1341(0.022)	1.977(0.001)	4.363(0.262)	**6.205(0.011)**
Catoni-bound [Guan and Lu, 2022]	0.107(0.018)	1.880(0.001)	0.244(0.033)	2.375(0.001)	**1.849(0.138)**	8.014(0.001)
classic bound [Rezazadeh, 2022]	0.260(0.026)	2.088(0.001)	0.267(0.028)	2.832(0.001)	5.362(0.374)	10.950(0.015)
PAC-MLE (ours)	**0.053(0.011)**	**1.570(0.001)**	**0.091(0.017)**	**1.925(0.001)**	2.662(0.134)	**6.806(0.003)**

Table 2: Comparisons of different PAC-Bayesian meta-learning methods. The average test bounds and test errors are reported over 20 test tasks in three pixel-shuffled environments.

5.3 Image Classification

Setup. In this subsection, we conduct classification experiments on augmented MNIST dataset, and compare our bounds with previous works. Each task from the same environment is constructed by a limited number of pixel swaps or a random permutation of the labels to ensure task-relatedness. We followed previous works [Amit and Meir, 2018; Liu *et al.*, 2021] on all experimental details During the meta-training phase, we choose 10 training tasks, each consisting of 60,000 training examples; while in the meta-test phase, each task is constructed with reduced training samples, specifically 2,000. We utilize a four-layer fully-connected network for shuffled pixels experiments, and a four-layer convolutional network for permuted labels experiments.

Results. Table 2 shows the comparison of various bounds for both shuffled pixels and permuted labels experiments. The performance of our classic bounds is significantly better than the existing bounds, in terms of test bounds and test errors over the novel tasks. This demonstrates that our framework is not only theoretically sound, but also achieves a competitive generalization performance guarantee over all tasks from different environments. Meanwhile, the prediction performance of different methods gets worse with the increase in the number of pixel swaps. This also indicates the success of meta-learning is significantly influenced by the task-relatedness of the environment.

Next, we investigate the impact of different numbers of training tasks n and sample sizes per task m on the quality of the learned prior. The Fig. 5 (see Appendix C.3) plots the average test error of learning new tasks based on the number of training tasks and sample sizes per task in two distinct environments: the permuted labels environment and the permuted pixels environment where 100, 200, and 300 pixel swaps. The results demonstrate that the increased number of tasks (n) and sample sizes (m) used for learning the prior leads to improved performance on novel tasks. As anticipated, this effect is particularly discernible when the number of meta-training tasks is limited and diminishes as the value of n increases.

6 Conclusion

In this paper, we have introduced a simple framework for meta-learning with generalization guarantees, grounded on a PAC-Bayesian bound that incorporates data-dependent priors. Rather than coercing the data-dependent priors into standard meta-learning settings, we identify a connection between local entropy and PAC-Bayesian theory that allows us to optimize priors for tighter bounds, with a simple local entropy, without the explicit reliance on posteriors. The resulting algorithm is stable, performant, and computationally lightweight, eliminating the need for complex nest optimization. Experimental results demonstrate that our method outperforms previous approaches in terms of fast adaptation, generalization performance, and combating meta-overfitting.

Acknowledgments

Shaogao Lv's research was partially supported by the National Natural Science Foundation of China (No. 12371291, No. 72342019). This work was partially supported by the National Key Research and Development Program of China (No. 2022ZD0115301, No. 2018AAA0100204, No. 2021ZD0112900), the National Natural Science Foundation of China (No. 62206139), the Major Key Project of PCL (PCL2023A09), a key program of fundamental research from Shenzhen Science and Technology Innovation Commission (No. JCYJ20200109113403826), and an Open Research Project of Zhejiang Lab (NO. 2022RC0AB04).

References

[Amit and Meir, 2018] Ron Amit and Ron Meir. Meta-learning by adjusting priors based on extended pac-bayes theory. In *International Conference on Machine Learning*, 2018.

[Baxter, 2000] Jonathan Baxter. A model of inductive bias learning. *Journal of artificial intelligence research*, 2000.

[Catoni, 2007] Olivier Catoni. Pac-bayesian supervised classification: the thermodynamics of statistical learning. *arXiv preprint arXiv:0712.0248*, 2007.

[Chaudhari *et al.*, 2017] Pratik Chaudhari, Anna Choromanska, Stefano Soatto, Yann LeCun, Carlo Baldassi, Christian Borgs, Jennifer Chayes, Levent Sagun, and Riccardo Zecchina. Entropy-sgd: Biasing gradient descent into wide valleys. In *International Conference on Learning Representations*, 2017.

[Chaudhari *et al.*, 2019] Pratik Chaudhari, Anna Choromanska, Stefano Soatto, Yann LeCun, Carlo Baldassi, Christian Borgs, Jennifer Chayes, Levent Sagun, and Riccardo Zecchina. Entropy-sgd: Biasing gradient descent into wide valleys. *Journal of Statistical Mechanics: Theory and Experiment*, 2019.

[Chen and Chen, 2022] Lisha Chen and Tianyi Chen. Is bayesian model-agnostic meta learning better than model-agnostic meta learning, provably? In *International Conference on Artificial Intelligence and Statistics*, 2022.

[Dwork *et al.*, 2015] Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Leon Roth. Preserving statistical validity in adaptive data analysis. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 2015.

[Dziugaite and Roy, 2018a] Gintare Karolina Dziugaite and Daniel Roy. Entropy-sgd optimizes the prior of a pac-bayes bound: Generalization properties of entropy-sgd and data-dependent priors. In *International Conference on Machine Learning*, 2018.

[Dziugaite and Roy, 2018b] Gintare Karolina Dziugaite and Daniel M Roy. Data-dependent pac-bayes priors via differential privacy. In *Advances in neural information processing systems*, 2018.

[Dziugaite *et al.*, 2021] Gintare Karolina Dziugaite, Kyle Hsu, Waseem Gharbieh, Gabriel Arpino, and Daniel Roy. On the role of data in pac-bayes bounds. In *International Conference on Artificial Intelligence and Statistics*, 2021.

[Finn *et al.*, 2017] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning*, 2017.

[Finn *et al.*, 2018] Chelsea Finn, Kelvin Xu, and Sergey Levine. Probabilistic model-agnostic meta-learning. In *Advances in neural information processing systems*, 2018.

[Guan and Lu, 2022] Jiechao Guan and Zhiwu Lu. Fast-rate pac-bayesian generalization bounds for meta-learning. In *International Conference on Machine Learning*, 2022.

[Guan *et al.*, 2022] Jiechao Guan, Zhiwu Lu, and Yong Liu. Improved generalization risk bounds for meta-learning with pac-bayes-kl analysis. 2022.

[Hospedales *et al.*, 2021] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 2021.

[Kingma and Welling, 2013] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[Lee *et al.*, 2022] Hung-yi Lee, Shang-Wen Li, and Ngoc Thang Vu. Meta learning for natural language processing: A survey. *arXiv preprint arXiv:2205.01500*, 2022.

[Lever *et al.*, 2013] Guy Lever, François Laviolette, and John Shawe-Taylor. Tighter pac-bayes bounds through distribution-dependent priors. *Theoretical Computer Science*, 2013.

[Liu and Wang, 2016] Qiang Liu and Dilin Wang. Stein variational gradient descent: A general purpose bayesian inference algorithm. In *Advances in neural information processing systems*, 2016.

[Liu *et al.*, 2021] Tianyu Liu, Jie Lu, Zheng Yan, and Guangquan Zhang. Statistical generalization performance guarantee for meta-learning with data dependent prior. *Neurocomputing*, 2021.

[Madden, 2004] Samuel Madden. Intel lab data. http://db.csail.mit.edu/labdata/labdata.html, 2004. Accessed: Sep 8, 2020.

[McAllester, 1999] David A McAllester. Pac-bayesian model averaging. In *Proceedings of the twelfth annual Conference on Computational learning theory*, 1999.

[Milne *et al.*, 2017] Christopher J Milne, Thomas Schietinger, Masamitsu Aiba, Arturo Alarcon, Jürgen Alex, Alexander Anghel, Vladimir Arsov, Carl Beard, Paul Beaud, Simona Bettoni, et al. Swissfel: the swiss x-ray free electron laser. *Applied Sciences*, 2017.

[Nichol *et al.*, 2018] Alex Nichol, Joshua Achiam, and John Schulman. On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999*, 2018.

[Pentina and Lampert, 2014] Anastasia Pentina and Christoph Lampert. A pac-bayesian bound for lifelong learning. In *International Conference on Machine Learning*, 2014.

[Pittorino *et al.*, 2021] Fabrizio Pittorino, Carlo Lucibello, Christoph Feinauer, Gabriele Perugini, Carlo Baldassi, Elizaveta Demyanenko, and Riccardo Zecchina. Entropic gradient descent algorithms and wide flat minima. *Journal of Statistical Mechanics: Theory and Experiment*, 2021.

[Raginsky *et al.*, 2017] Maxim Raginsky, Alexander Rakhlin, and Matus Telgarsky. Non-convex learning via stochastic gradient langevin dynamics: a nonasymptotic analysis. In *Conference on Learning Theory*, 2017.

[Rezazadeh, 2022] Arezou Rezazadeh. A unified view on pac-bayes bounds for meta-learning. In *International Conference on Machine Learning*, 2022.

[Rivasplata *et al.*, 2020] Omar Rivasplata, Ilja Kuzborskij, Csaba Szepesvári, and John Shawe-Taylor. Pac-bayes analysis beyond the usual bounds. In *Advances in Neural Information Processing Systems*, 2020.

[Rothfuss *et al.*, 2019] Jonas Rothfuss, Dennis Lee, Ignasi Clavera, Tamim Asfour, and Pieter Abbeel. Promp: Proximal meta-policy search. In *International Conference on Machine Learning*, 2019.

[Rothfuss *et al.*, 2021] Jonas Rothfuss, Vincent Fortuin, Martin Josifoski, and Andreas Krause. Pacoh: Bayes-optimal meta-learning with pac-guarantees. In *International Conference on Machine Learning*, 2021.

[Rothfuss *et al.*, 2023] Jonas Rothfuss, Martin Josifoski, Vincent Fortuin, and Andreas Krause. Scalable pac-bayesian meta-learning via the pac-optimal hyper-posterior: From theory to practice. *arXiv preprint arXiv:2211.07206*, 2023.

[Silva *et al.*, 2012] Ikaro Silva, George Moody, Daniel J Scott, Leo A Celi, and Roger G Mark. Predicting in-hospital mortality of icu patients: The physionet/computing in cardiology challenge 2012. In *Computing in Cardiology*, 2012.

[Thiemann *et al.*, 2017] Niklas Thiemann, Christian Igel, Olivier Wintenberger, and Yevgeny Seldin. A strongly quasiconvex pac-bayesian bound. In *International Conference on Algorithmic Learning Theory*, 2017.

[Thrun and Pratt, 1998] Sebastian Thrun and Lorien Pratt. Learning to learn: Introduction and overview. *Learning to learn*, 1998.

[Vinyals *et al.*, 2016] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. In *Advances in neural information processing systems*, 2016.

[Welling and Teh, 2011] Max Welling and Yee W Teh. Bayesian learning via stochastic gradient langevin dynamics. In *International conference on machine learning*, 2011.

[Williams, 1992] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 1992.

[Xu *et al.*, 2020] Jin Xu, Jean-Francois Ton, Hyunjik Kim, Adam Kosiorek, and Yee Whye Teh. Metafun: Meta-learning with iterative functional updates. In *International Conference on Machine Learning*, 2020.

[Yoon *et al.*, 2018] Jaesik Yoon, Taesup Kim, Ousmane Dia, Sungwoong Kim, Yoshua Bengio, and Sungjin Ahn. Bayesian model-agnostic meta-learning. In *Advances in neural information processing systems*, 2018.

[Zhao *et al.*, 2022] Mandi Zhao, Pieter Abbeel, and Stephen James. On the effectiveness of fine-tuning versus meta-reinforcement learning. In *Advances in Neural Information Processing Systems*, 2022.

[Zhu *et al.*, 2023] Zhuangdi Zhu, Kaixiang Lin, Anil K Jain, and Jiayu Zhou. Transfer learning in deep reinforcement learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.