NeurIPS 2024
The 38th Annual Conference on Neural Information Processing Systems

A Unified Confidence Sequence for Generalized Linear Models, with Applications to Bandits

Junghyun Lee (KAIST AI), Se-Young Yun (KAIST AI), Kwang-Sung Jun (Univ. of Arizona CS)

KAIST AI
Kim Jaechul Graduate School

OSI Optimization and Statistical Inference LAB
OptiML Optimization & Machine Learning

Computer Science

Generalized Linear Models

Problem Setting

Generalized Linear Models

Problem Setting

Consider the **Generalized Linear Model (GLM)**:

$$dp(r \mid x; \theta_\star) = \exp\left(\frac{r\langle x, \theta_\star \rangle - m(\langle x, \theta_\star \rangle)}{g(\tau)} + h(r, \tau)\right) d\nu,$$

with dispersion parameter $\tau > 0$, base measure ν, **context** $x \in X$, and **unknown parameter** $\theta_\star \in \Theta$**.**

Generalized Linear Models

Problem Setting

Consider the **Generalized Linear Model (GLM)**:

$$dp(r \,|\, x; \theta_\star) = \exp\left(\frac{r\langle x, \theta_\star\rangle - m(\langle x, \theta_\star\rangle)}{g(\tau)} + h(r, \tau)\right) d\nu,$$

with dispersion parameter $\tau > 0$, base measure ν, **context** $x \in X$, and **unknown parameter** $\theta_\star \in \Theta$.

Assumptions. $X \subseteq \mathbb{B}^d(1)$, $\emptyset \neq \Theta \subseteq \mathbb{B}^d(S)$, Θ compact & convex, $m(\cdot)$ is convex and three-times differentiable.

Properties. $\mathbb{E}[r \,|\, x, \theta_\star] = m'(\langle x, \theta_\star\rangle) =: \mu(\langle x, \theta_\star\rangle)$, $\mathrm{Var}[r \,|\, x, \theta_\star] = g(\tau)\dot{\mu}(\langle x, \theta_\star\rangle)$

Examples. $\mu(z) = z$: Gaussian, $\mu(z) = (1 + e^{-z})^{-1}$: **Bernoulli**, $\mu(z) = e^z$: Poisson

Generalized Linear Models

Problem Setting

Consider the **Generalized Linear Model (GLM)**:

$$dp(r \mid x; \theta_\star) = \exp\left(\frac{r\langle x, \theta_\star\rangle - m(\langle x, \theta_\star\rangle)}{g(\tau)} + h(r, \tau)\right) d\nu,$$

with dispersion parameter $\tau > 0$, base measure ν, **context** $x \in X$, and **unknown parameter** $\theta_\star \in \Theta$.

Assumptions. $X \subseteq \mathbb{B}^d(1)$, $\emptyset \neq \Theta \subseteq \mathbb{B}^d(S)$, Θ compact & convex, $m(\cdot)$ is convex and three-times differentiable.

Properties. $\mathbb{E}[r \mid x, \theta_\star] = m'(\langle x, \theta_\star\rangle) =: \mu(\langle x, \theta_\star\rangle)$, $\mathrm{Var}[r \mid x, \theta_\star] = g(\tau)\dot{\mu}(\langle x, \theta_\star\rangle)$

Examples. $\mu(z) = z$: Gaussian, $\mu(z) = (1 + e^{-z})^{-1}$: **Bernoulli**, $\mu(z) = e^z$: Poisson

Generalized Linear Bandits

Confidence Sequence (CS) for the Unknown Parameter

Generalized Linear Bandits

Confidence Sequence (CS) for the Unknown Parameter

Goal: For $\delta \in (0,1)$**, obtain** $\{\mathscr{C}_t(\delta)\}_{t \geq 1}$ **s.t.** $\mathbb{P}\left(\exists t \geq 1 : \theta_\star \notin \mathscr{C}_t(\delta) \right) \leq \delta$

Generalized Linear Bandits

Confidence Sequence (CS) for the Unknown Parameter

Goal: For $\delta \in (0,1)$**, obtain** $\{\mathscr{C}_t(\delta)\}_{t\geq 1}$ **s.t.** $\mathbb{P}\left(\exists t \geq 1 : \theta_\star \notin \mathscr{C}_t(\delta)\right) \leq \delta$

Setting. $\{(x_s, r_s)\}_{s\geq 1}$: adaptively collected observations satisfying $\mathbb{E}[r_s \,|\, \Sigma_s] = \mu(\langle x_s, \theta_\star \rangle)$, where $\Sigma_s := \sigma(\{x_1, r_1, \cdots, x_{s-1}, r_{s-1}, x_s\})$.

Generalized Linear Bandits

Confidence Sequence (CS) for the Unknown Parameter

Goal: For $\delta \in (0,1)$, obtain $\{\mathscr{C}_t(\delta)\}_{t\geq 1}$ s.t. $\mathbb{P}\left(\exists t \geq 1 : \theta_\star \notin \mathscr{C}_t(\delta)\right) \leq \delta$

Setting. $\{(x_s, r_s)\}_{s\geq 1}$: adaptively collected observations satisfying $\mathbb{E}[r_s \mid \Sigma_s] = \mu(\langle x_s, \theta_\star \rangle)$, where $\Sigma_s := \sigma(\{x_1, r_1, \cdots, x_{s-1}, r_{s-1}, x_s\})$.

We consider **CS** of the form $\mathscr{C}_t(\delta) := \left\{\theta \in \Theta : \mathscr{L}_t(\theta) - \mathscr{L}_t(\widehat{\theta}_t) \leq \beta_t(\delta)^2\right\}$, where

$$\mathscr{L}_t(\theta) := \sum_{s=1}^{t-1} \left\{ \ell_s(\theta) \triangleq \frac{-r_s\langle x_s, \theta\rangle + m(\langle x_s, \theta\rangle)}{g(\tau)} \right\}, \quad \widehat{\theta}_t := \mathrm{argmin}_{\theta\in\Theta} \mathscr{L}_t(\theta) .$$

where $\mathscr{L}_t(\theta)$ is the cumulative log-likelihood loss til time $t-1$, with **Lipschitz constant** L_t.

New, State-of-the-Art CS for GLMs!

Contribution #1

Proof via PAC-Bayes

Theorem 3.1. We have $\mathbb{P}\left(\exists t \geq 1 : \theta_\star \notin \mathcal{C}_t(\delta)\right) \leq \delta$, where

$$\mathcal{C}_t(\delta) := \left\{\theta \in \Theta : \mathcal{L}_t(\theta) - \mathcal{L}_t(\widehat{\theta}_t) \leq \beta_t(\delta)^2\right\}$$

$$\beta_t(\delta)^2 := \log\frac{1}{\delta} + d\log\left(e \vee \frac{2eSL_t}{d}\right)$$

Bernoulli: $\beta_t(\delta)^2 \lesssim_\delta d\log\frac{St}{d}$ => poly(S)-free for **Bernoulli**!!!

<=> prior work [Lee et al., AISTATS'24]: $\mathcal{O}_\delta\left(S + d\log\frac{St}{d}\right)$

Rmk. For self-concordant GLMs, one can have an *ellipsoidal form* of the CS.

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Lemma 3.3. For any data-independent "prior" $\mathbb{Q}$ and any sequence of adapted "posterior" distributions (possibly learned from the data) $\{\mathbb{P}_t\}$, the following holds:

$$\mathbb{P}\left(\exists t \geq 1 : \mathscr{L}_t(\theta_\star) - \mathbb{E}_{\theta \sim \mathbb{P}_t}[\mathscr{L}_t(\theta)] \geq \log \frac{1}{\delta} + D_{KL}(\mathbb{P}_t \| \mathbb{Q}) \right) \leq \delta$$

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Lemma 3.3. For any data-independent "prior" $\mathbb{Q}$ and any sequence of adapted "posterior" distributions (possibly learned from the data) $\{\mathbb{P}_t\}$, the following holds:

$$\mathbb{P}\left(\exists t \geq 1 : \mathscr{L}_t(\theta_\star) - \mathbb{E}_{\theta \sim \mathbb{P}_t}[\mathscr{L}_t(\theta)] \geq \log \frac{1}{\delta} + D_{KL}(\mathbb{P}_t \| \mathbb{Q}) \right) \leq \delta$$

pf. Consider the likelihood ratio $M_t(\theta) = \exp(\mathscr{L}_t(\theta_\star) - \mathscr{L}_t(\theta))$.

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Lemma 3.3. For any data-independent “prior” $\mathbb{Q}$ and any sequence of adapted “posterior” distributions (possibly learned from the data) $\{\mathbb{P}_t\}$, the following holds:

$$\mathbb{P}\left(\exists t \geq 1 : \mathscr{L}_t(\theta_\star) - \mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathscr{L}_t(\theta)] \geq \log\frac{1}{\delta} + D_{KL}(\mathbb{P}_t\|\mathbb{Q})\right) \leq \delta$$

pf. Consider the likelihood ratio $M_t(\theta) = \exp(\mathscr{L}_t(\theta_\star) - \mathscr{L}_t(\theta))$.

1. $M_t(\theta)$ is a nonnegative martingale, and so is $\mathbb{E}_{\theta\sim\mathbb{Q}}[M_t(\theta)]$ by Tonelli's theorem

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Lemma 3.3. For any data-independent "prior" $\mathbb{Q}$ and any sequence of adapted "posterior" distributions (possibly learned from the data) $\{\mathbb{P}_t\}$, the following holds:

$$\mathbb{P}\left(\exists t \geq 1 : \mathcal{L}_t(\theta_\star) - \mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathcal{L}_t(\theta)] \geq \log\frac{1}{\delta} + D_{KL}(\mathbb{P}_t \| \mathbb{Q})\right) \leq \delta$$

pf. Consider the likelihood ratio $M_t(\theta) = \exp(\mathcal{L}_t(\theta_\star) - \mathcal{L}_t(\theta))$.

Anytime-valid *Markov's inequality* for supermartingales

1. $M_t(\theta)$ is a nonnegative martingale, and so is $\mathbb{E}_{\theta\sim\mathbb{Q}}[M_t(\theta)]$ by Tonelli's theorem
2. By Ville's inequality [Ville, 1939], we have $\mathbb{P}\left(\exists t \geq 1 : \mathbb{E}_{\theta\sim\mathbb{Q}}[M_t(\theta)] \geq \frac{1}{\delta}\right) \leq \delta$

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Lemma 3.3. For any data-independent "prior" $\mathbb{Q}$ and any sequence of adapted "posterior" distributions (possibly learned from the data) $\{\mathbb{P}_t\}$, the following holds:

$$\mathbb{P}\left(\exists t \geq 1 : \mathscr{L}_t(\theta_\star) - \mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathscr{L}_t(\theta)] \geq \log\frac{1}{\delta} + D_{KL}(\mathbb{P}_t \| \mathbb{Q})\right) \leq \delta$$

pf. Consider the likelihood ratio $M_t(\theta) = \exp(\mathscr{L}_t(\theta_\star) - \mathscr{L}_t(\theta))$.

Anytime-valid *Markov's inequality* for supermartingales

1. $M_t(\theta)$ is a nonnegative martingale, and so is $\mathbb{E}_{\theta\sim\mathbb{Q}}[M_t(\theta)]$ by Tonelli's theorem

2. By Ville's inequality [Ville, 1939], we have $\mathbb{P}\left(\exists t \geq 1 : \mathbb{E}_{\theta\sim\mathbb{Q}}[M_t(\theta)] \geq \frac{1}{\delta}\right) \leq \delta$

3. "Change" $\mathbb{Q}$ to $\mathbb{P}_t$ via **Donsker-Varadhan variational representation of KL** [Donsker & Varadhan, 1983].

$$\mathrm{KL}(\mathbb{P}_t || \mathbb{Q}) = \sup_{g:\Theta\to\mathbb{R}} \mathbb{E}_{\theta\sim\mathbb{P}_t}[g(\theta)] - \log \mathbb{E}_{\theta\sim\mathbb{Q}}[e^{g(\theta)}]$$

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Journal of Machine Learning Research 24 (2023) 1-61 Submitted 3/23; Revised 10/23; Published 12/23

A Unified Recipe for Deriving (Time-Uniform) PAC-Bayes Bounds

Ben Chugg BENCHUGG@CMU.EDU
Hongjian Wang HJNWANG@CMU.EDU
Aaditya Ramdas ARAMDAS@STAT.CMU.EDU
Departments of Statistics and Machine Learning
Carnegie Mellon University

Proof of Theorem 3.1

Step 1. Time-Uniform PAC-Bayes Bound

Journal of Machine Learning Research 24 (2023) 1-61 Submitted 3/23; Revised 10/23; Published 12/23

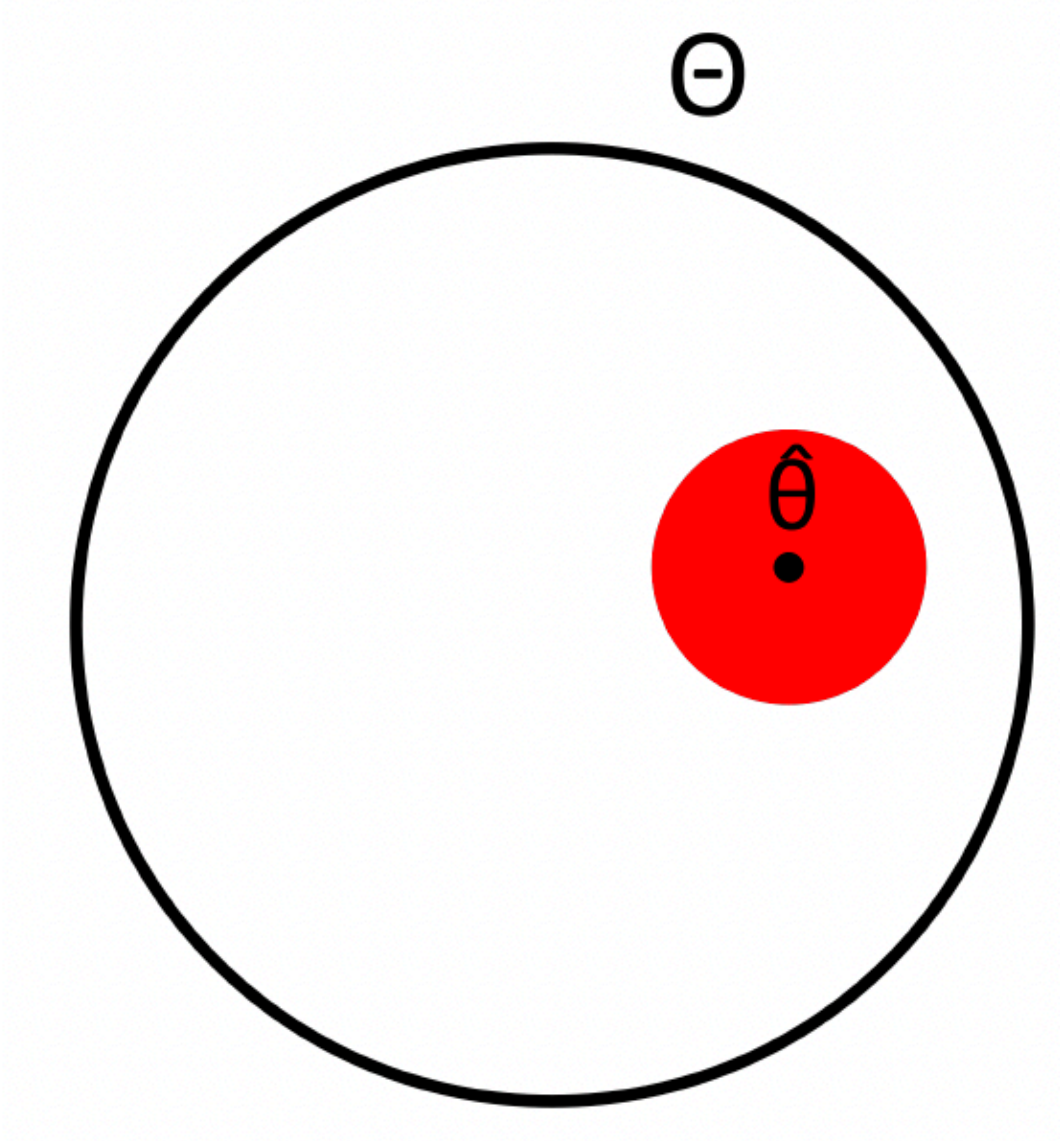
SURVEY

A Unified Recipe for Deriving (Time-Uniform) PAC-Bayes Bounds

Ben Chugg BENCHUGG@CMU.EDU
Hongjian Wang HJNWANG@CMU.EDU
Aaditya Ramdas ARAMDAS@STAT.CMU.EDU
Departments of Statistics and Machine Learning
Carnegie Mellon University

Proof of Theorem 3.1

Step 2. Novel choice of of “prior” and “posterior” & Lipschitzness



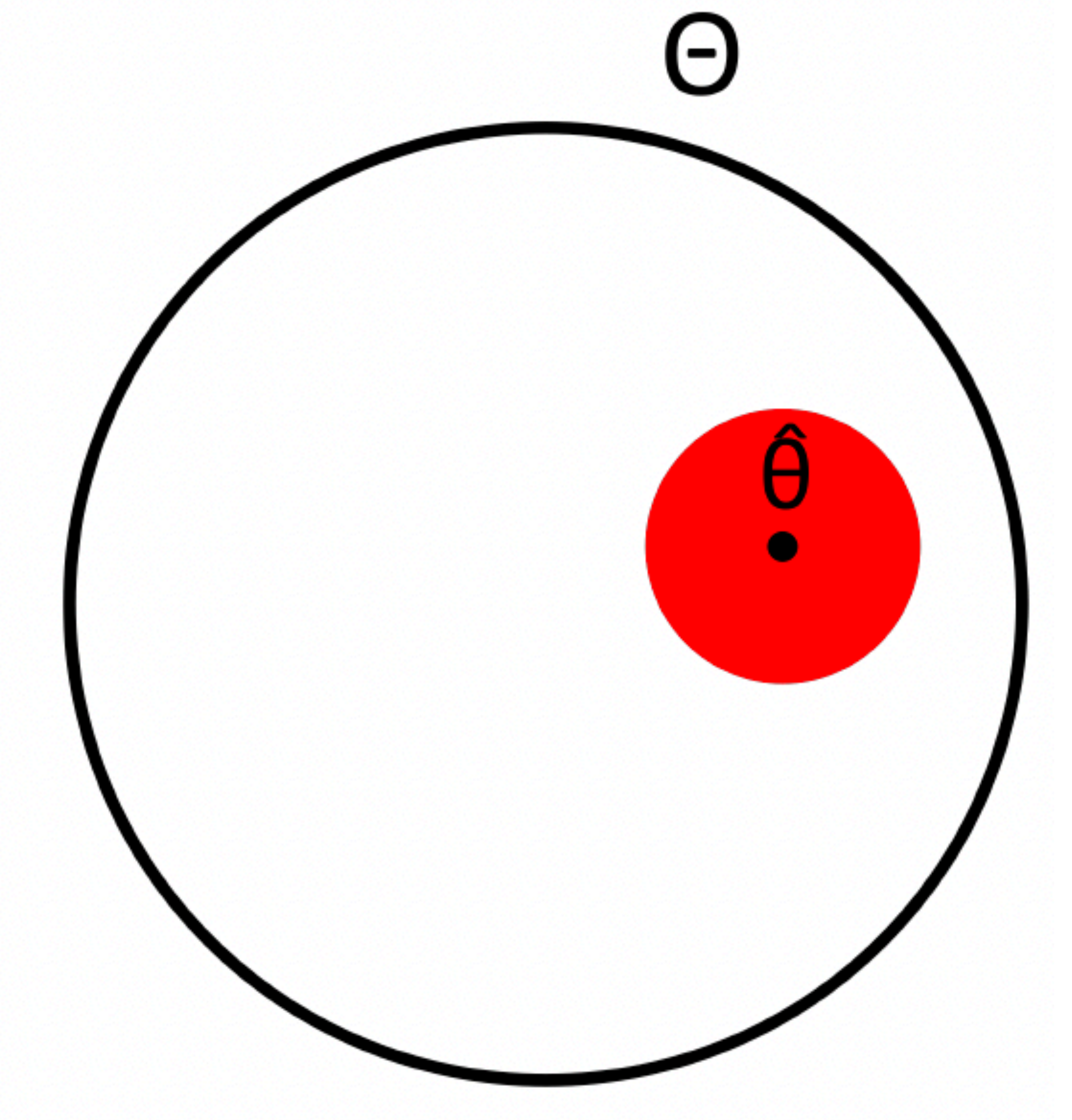
From P. Alquier’s MLSS lecture slides

Proof of Theorem 3.1

Step 2. Novel choice of of “prior” and “posterior” & Lipschitzness

$$\mathbb{Q} = \mathrm{Unif}(\Theta), \quad \mathbb{P}_t = \mathrm{Unif}\left(\widetilde{\Theta}_t \triangleq (1-c)\hat{\theta}_t + c\Theta\right)$$

Remark. Originally considered in portfolio optimization [Blum and Kalai, 1999] and fast rates in online learning [Hazan et al., 2007; Foster et al., COLT'18].



From P. Alquier's MLSS lecture slides

Proof of Theorem 3.1

Step 2. Novel choice of of “prior” and “posterior” & Lipschitzness

$$\mathbb{Q} = \mathrm{Unif}(\Theta), \quad \mathbb{P}_t = \mathrm{Unif}\left(\widetilde{\Theta}_t \triangleq (1-c)\hat{\theta}_t + c\Theta\right)$$

Remark. Originally considered in portfolio optimization [Blum and Kalai, 1999] and fast rates in online learning [Hazan et al., 2007; Foster et al., COLT’18].

Proof of Theorem 3.1

Step 2. Novel choice of of "prior" and "posterior" & Lipschitzness

$$\mathbb{Q} = \mathrm{Unif}(\Theta), \quad \mathbb{P}_t = \mathrm{Unif}\left(\widetilde{\Theta}_t \triangleq (1-c)\widehat{\theta}_t + c\Theta\right)$$

Remark. Originally considered in portfolio optimization [Blum and Kalai, 1999] and fast rates in online learning [Hazan et al., 2007; Foster et al., COLT'18].

=> $D_{KL}(\mathbb{P}_t || \mathbb{Q}) = \log \dfrac{\mathrm{vol}(\Theta)}{\mathrm{vol}(\widetilde{\Theta})} = \log \dfrac{\mathrm{vol}(\Theta)}{\mathrm{vol}(c\Theta)} = d \log \dfrac{1}{c}$

Also, $\mathbb{E}_{\theta \sim \mathbb{P}_t}[\mathscr{L}_t(\theta)] = \mathscr{L}_t(\widehat{\theta}_t) + \mathbb{E}_{\theta \sim \mathbb{P}_t}[\mathscr{L}_t(\theta) - \mathscr{L}_t(\widehat{\theta}_t)] \leq \mathscr{L}_t(\widehat{\theta}_t) + 2SL_t c$,

Proof of Theorem 3.1

Step 2. Novel choice of of "prior" and "posterior" & Lipschitzness

$$\mathbb{Q} = \mathrm{Unif}(\Theta), \quad \mathbb{P}_t = \mathrm{Unif}\left(\widetilde{\Theta}_t \triangleq (1-c)\widehat{\theta}_t + c\Theta\right)$$

Remark. Originally considered in portfolio optimization [Blum and Kalai, 1999] and fast rates in online learning [Hazan et al., 2007; Foster et al., COLT'18].

=> $D_{KL}(\mathbb{P}_t \| \mathbb{Q}) = \log \dfrac{\mathrm{vol}(\Theta)}{\mathrm{vol}(\widetilde{\Theta})} = \log \dfrac{\mathrm{vol}(\Theta)}{\mathrm{vol}(c\Theta)} = d \log \dfrac{1}{c}$

Also, $\mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathscr{L}_t(\theta)] = \mathscr{L}_t(\widehat{\theta}_t) + \mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathscr{L}_t(\theta) - \mathscr{L}_t(\widehat{\theta}_t)] \leq \mathscr{L}_t(\widehat{\theta}_t) + 2SL_t c$,

All in all, with probability at most δ, there exists a $t \geq 1$ such that

$$\mathscr{L}_t(\theta_\star) - \mathscr{L}_t(\widehat{\theta}_t) \geq \log\frac{1}{\delta} + d\log\frac{1}{c} + \mathbb{E}_{\theta\sim\mathbb{P}_t}[\mathscr{L}_t(\theta)] - \mathscr{L}_t(\widehat{\theta}_t) \geq \log\frac{1}{\delta} + d\log\frac{1}{c} + 2SL_t c$$

Choose $c = \min\left\{1, d/(2SL_t)\right\}$ and we are done.

Generalized Linear Bandits

Problem Setting

For $t \in [T]$:

1. The learner observes a potentially infinite (contextual) arm-set $\mathcal{X}_t \subset X$
2. The learner chooses $x_t \in \mathcal{X}_t$ according to some policy
3. Receive a reward $r_t \sim GLM(x_t, \theta_\star; \mu(\cdot))$
    - $\theta_\star$ is unknown to the learner

Goal: Minimize the regret

$$\mathrm{Reg}^B(T) := \sum_{t=1}^{T} \left\{ \mu(\langle x_{t,\star}, \theta_\star \rangle) - \mu(\langle x_t, \theta_\star \rangle) \right\} \text{ where } x_{t,\star} := \mathrm{argmax}_{x \in \mathcal{X}_t} \mu(\langle x, \theta_\star \rangle).$$

Generalized Linear Bandits

Contribution #2

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

1. Compute $\widehat{\theta}_t$ and $\mathscr{C}_t(\delta)$ - **tighter confidence sequence** (Theorem 3.1)**!**
2. $(x_t, \theta_t) = \mathrm{argmax}_{x \in \mathscr{X}_t, \theta \in \mathscr{C}_t(\delta)} \mu(\langle x, \theta \rangle)$
3. Play x_t and observe/receive a reward $r_t \sim GLM(x_t, \theta_\star; \mu(\cdot))$

Theorem 4.1. OFUGLB attains the following regret bound for self-concordant generalized linear bandits w.p. at least $1 - \delta$:

Nontrivial proof!!

$$\mathrm{Reg}(T) \lesssim \underbrace{d\sqrt{\frac{g(\tau)T}{\kappa_\star(T)} \log \frac{SL_T}{d} \log \frac{R_{\dot{\mu}}ST}{d}}}_{\text{permanent term}} + \underbrace{d^2 R_s R_{\dot{\mu}} \sqrt{g(\tau)} \kappa(T)}_{\text{transient term}}$$

Generalized Linear Bandits

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

- **Linear Bandits:** $\tilde{\mathcal{O}}\left(\sigma d\sqrt{T}\right)$
  - => matches state-of-the-art [Flynn et al., NeurIPS'23]
- **Logistic Bandits:** $\tilde{\mathcal{O}}\left(d\sqrt{T/\kappa_\star(T)} + d^2\kappa(T)\right)$
  - => *first* poly(S)-free regret with **computationally tractable, purely optimistic approach**!!
  - => improves upon prior state-of-the-art [Lee et al., AISTATS'24]
  - => similar guarantee in a *concurrent* work [Sawarni et al., arXiv'24], but is intractable and involves explicit warmup + their guarantees only apply to *bounded* GLBs.
- **Poisson Bandits:** $\tilde{\mathcal{O}}\left(dS\sqrt{T/\kappa_\star(T)} + d^2e^{2S}\kappa(T)\right)$
  - => *state-of-the-art* regret guarantee

Brief Proof Sketch of Theorem 4.1

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

Brief Proof Sketch of Theorem 4.1

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

Previously: use self-concordance control lemma to obtain

$\|\theta_\star - \hat{\theta}_t\|_{H_t(\hat{\theta}_t)} = \mathcal{O}(S\beta_T(\delta))$

Brief Proof Sketch of Theorem 4.1

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

Previously: use self-concordance control lemma to obtain
$\|\theta_\star - \hat{\theta}_t\|_{H_t(\hat{\theta}_t)} = \mathcal{O}(S\beta_T(\delta))$

Here: maximally avoid self-concordance control => use "exact" Taylor expansion,
$\|\theta_\star - \hat{\theta}_t\|_{\tilde{G}_t(\hat{\theta}_t,\nu_t)} = \mathcal{O}(\beta_T(\delta))$, where $\tilde{G}_t(\hat{\theta}_t, \nu_t) = \lambda \mathbf{I} + \frac{1}{g(\tau)} \sum_{s=1}^{t-1} \tilde{\alpha}_s(\hat{\theta}_t, \nu_t) x_s x_s^\top$ and
$\tilde{\alpha}_s(\theta, \nu) = \int_0^1 (1 - v)\dot{\mu}_t(\theta + v(\nu - \theta))dv.$

Brief Proof Sketch of Theorem 4.1

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

Brief Proof Sketch of Theorem 4.1

OFUGLB: Optimism in the Face of Uncertainty for Generalized Linear Bandits

BUT, the remaining term of Cauchy-Schwartz, $\sum_t \|x_t\|^2_{\tilde{G}_t(\hat{\theta}_t, \nu_t)^{-1}}$, how to apply ***elliptical potential lemma?***

$$\tilde{G}_t(\hat{\theta}_t, \nu_t) = \lambda \mathbf{I} + \frac{1}{g(\tau)} \sum_{s=1}^{t-1} \tilde{\alpha}_s(\hat{\theta}_t, \nu_t) x_s x_s^\top$$

Lemma B.2 (Elliptical Potential Lemma; EPL[5]). *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_T \in \mathcal{B}^d(X)$ be a sequence of vectors and $\mathbf{V}_t := \lambda \boldsymbol{I} + \sum_{s=1}^{t-1} \boldsymbol{x}_s \boldsymbol{x}_s^\intercal$. Then, we have that*

$$\sum_{t=1}^{T} \min \left\{1, \|\boldsymbol{x}_t\|_{\mathbf{V}_t^{-1}}^2 \right\} \leq 2d \log \left(1 + \frac{X^2 T}{d\lambda}\right). \tag{23}$$

BUT, the remaining term of Cauchy-Schwartz, $\sum_t \|x_t\|^2_{\tilde{G}_t(\hat{\theta}_t, \nu_t)^{-1}}$, how to apply ***elliptical potential lemma?***

$$\tilde{G}_t(\hat{\theta}_t, \nu_t) = \lambda \mathbf{I} + \frac{1}{g(\tau)} \sum_{s=1}^{t-1} \tilde{\alpha}_s(\hat{\theta}_t, \nu_t) x_s x_s^\top$$

Lemma B.2 (Elliptical Potential Lemma; EPL[5]). *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_T \in \mathcal{B}^d(X)$ be a sequence of vectors and $\mathbf{V}_t := \lambda \boldsymbol{I} + \sum_{s=1}^{t-1} \boldsymbol{x}_s \boldsymbol{x}_s^\intercal$. Then, we have that*

$$\sum_{t=1}^{T} \min\left\{1, \|\boldsymbol{x}_t\|_{\mathbf{V}_t^{-1}}^2\right\} \leq 2d \log\left(1 + \frac{X^2 T}{d\lambda}\right). \tag{23}$$

BUT, the remaining term of Cauchy-Schwartz, $\sum_t \|x_t\|^2_{\tilde{G}_t(\hat{\theta}_t, \nu_t)^{-1}}$, how to apply ***elliptical potential lemma?***

$$\tilde{G}_t(\hat{\theta}_t, \nu_t) = \lambda \mathbf{I} + \frac{1}{g(\tau)} \sum_{s=1}^{t-1} \tilde{\alpha}_s(\hat{\theta}_t, \nu_t) x_s x_s^\top$$

Main proof novelty: designate the "worst-case" $\bar{\theta}_t$'s such that

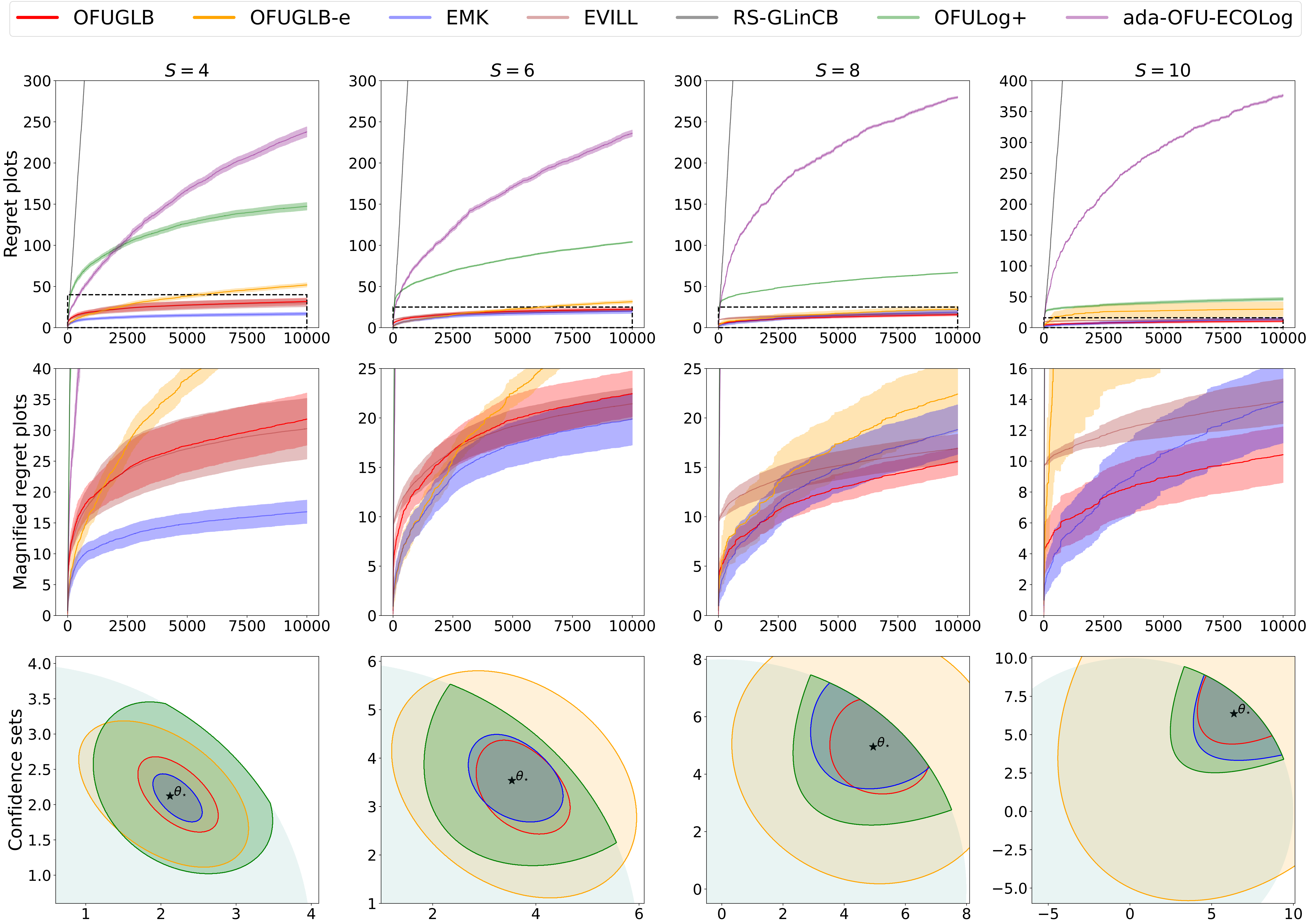
$$\sum_t \|x_t\|^2_{\tilde{G}_t(\hat{\theta}_t, \nu_t)^{-1}} \leq \sum_t \min\left\{1, \dot{\mu}(\bar{\theta}_s) \|x_t\|^2_{\bar{H}_t^{-1}}\right\}, \text{ where } \bar{H}_t = 2g(\tau)\lambda I + \sum_{s=1}^{t-1} \dot{\mu}_s(\bar{\theta}_s) x_s x_s^\top$$

Experiments for Logistic Bandits

Better than most of existing approaches

- One may wonder, does shaving off dependencies on S really help in practice?
- Synthetic experiments show that this is indeed beneficial, by a large margin!!

- On
- Sy



OFUGLB
OFUGLB-e
EMK
EVILL
RS-GLinCB
OFULog+
ada-OFU-ECOLog
$S = 4$
$S = 6$
$S = 8$
$S = 10$
Regret plots
Magnified regret plots
Confidence sets
$\theta_\star$

Thank you for your attention!

Poster Session 3 (Dec. 12, 11AM ~ 2PM)

1. A *unified,* state-of-the-art construction of likelihood ratio-based CS for any convex GLMs, with explicit constants!
2. **OFUGLB:** A new computationally tractable, optimistic algorithm that achieves state-of-the-art regrets for self-concordant GLBs.
3. For logistic bandits, its efficacy is shown numerically.

arXiv

References

J. Lee, S.-Y. Yun, and K.-S. Jun. Improved Regret Analysis of (Multinomial) Logistic Bandits via Regret-to-Confidence-Set Conversion. In *AISTATS* 2024.

A. Sawarni, N. Das, S. Barman, and G. Sinha. Generalized Linear Bandits with Limited Adaptivity. In *NeurIPS* 2024.

H. Flynn, D. Reeb, M. Kandemir, and J. R. Peters. Improved Algorithms for Stochastic Linear Bandits Using Tail Bounds for Martingale Mixtures. In *NeurIPS* 2023.

J. Ville. Étude critique de la notion de collectif. *Monographies des Probabilités*. Paris: Gauthier-Villars, 1939.

M. D. Donsker and S. R. S. Varadhan. Asymptotic Evaluation of Certain Markov Process Expectations for Large Times. IV. *Communications on Pure and Applied Mathematics*, 36(2):183-212, 1983.

A. Blum and A. Kalai. Universal Portfolios With and Without Transaction Costs. *Machine Learning*, 35(3):193-205, 1999.

E. Hazan, Z. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169-192, 2007.

D. J. Foster, S. Kale, H. Luo, M. Mohri, and K. Sridharan. Logistic Regression: The Importance of Being Improper. In *COLT 2018*.