# Empirical PAC-Bayes Bounds for Markov Chains

Vahe Karagulyan Pierre Alquier

ESSEC BUSINESS SCHOOL

AISTATS

# Problem Setup

- We observe sample trajectory: $\mathcal{S} = ((U_1, Y_1), \ldots, (U_n, Y_n))$.
- **Objects:** $(U_t)_{t \geq 1}$ form a **stationary Markov chain**.
- **Labels:** $Y_t$ depend only on $U_t$.
- **Prediction:** $\mathcal{F} = \{f_\theta : \mathcal{U} \to \mathcal{Y} \mid \theta \in \Theta\}$, and loss $\ell(.,.) \leq c$.

$$R(\theta) = \mathbb{E}_{\mathcal{S}}\left[\frac{1}{n}\sum_{t=1}^{n} \ell(f_\theta(U_t), Y_t)\right], \quad r(\theta) = \frac{1}{n}\sum_{t=1}^{n} \ell(f_\theta(U_t), Y_t).$$

# PAC-Bayes bound for Markov Chains

**Theorem 2.1:** Assume $\{U_t\}_{t=1}^{n}$ is a stationary Markov chain with pseudo-spectral gap $\gamma_{ps} > 0$. Then for any constants $0 < \lambda < \frac{n}{10}$, $\delta \in (0, 1)$, and prior $\mu \in \mathcal{P}(\Theta)$,

With probability $\geq 1 - \delta$, for all posteriors $\rho$:

$$\mathbb{E}_{\theta\sim\rho}[R(\theta)] \leq \mathbb{E}_{\theta\sim\rho}[r(\theta)] + \frac{2\lambda c^2\left(1 + \frac{1}{n\gamma_{ps}}\right)}{n - 10\lambda} + \frac{\mathrm{KL}(\rho||\mu) + \log\frac{1}{\delta}}{\lambda\gamma_{ps}}.$$

**Key insight:** Dependency penalty appears as a $\dfrac{1}{\gamma_{ps}}$ factor.

# From Non-Empirical to Empirical

Relative Deviation Bounds

$$\mathbb{P}_{\mathcal{S}}\left(\left|\frac{\widehat{\gamma}_{ps}}{\gamma_{ps}} - 1\right| \geq \varepsilon\right) \leq C_1 e^{-n\varepsilon^2 C_2} \tag{1}$$

# Examples of such $\widehat{\gamma}_{ps}$

**Example 1:** Finite state-space ergodic Markov chains: [1]

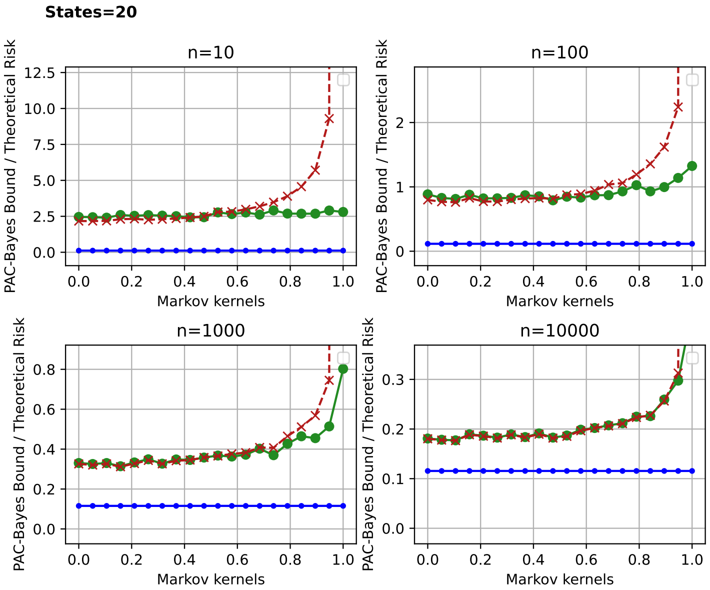
$$\widehat{\gamma}_{ps,[K]} := \max_{k \in [K]} \left\{ \frac{\gamma\left( \left( \hat{P}^{\dagger} \right)^{k} \hat{P}^{k} \right)}{k} \right\}. \tag{2}$$

**Example 2:** Stationary AR(1) process $U_t = aU_{t-1} + \zeta_t$ where $-1 < a < 1$ and the $\zeta_t$ are i.i.d. from $\mathcal{N}(0,1)$.

$$\widehat{\gamma}_{ps} := \min \left\{ \frac{1}{\frac{1}{n}\sum_{t=1}^{n} U_t^2}, 1 \right\}. \tag{3}$$

[1](Wolfer and Kontorovich, 2024)

# Experiments: Bound Tightness




**Green:** Emp. PAC-Bayes bound

**Red:** PAC-Bayes bound

**Blue:** True risk $\mathbb{E}_{\theta\sim\rho}[R(\theta)]$

Bernoulli labels:

$$Y_t \mid U_t = u \sim \mathrm{Ber}(p_u)$$

Predictors:

$$f_\theta(u) = \mathbf{1}(u \geq \theta)$$

# Contribution

- First empirical PAC-Bayes bounds for Markov chains.
- The empirical bound is just as tight as the non-empirical bound.

# Thank You!

(See you on the poster session ...)

**Vahe Karagulyan (website)**

**Pierre Alquier (website)**

**Full paper**