Optimality of Thompson Sampling for Gaussian Bandits Depends on Priors: Supplementary Material

We prove Lemmas 3, 4 and 9 in this supplementary material.

Before the proof of these lemmas we give a simple inequality to evaluate the ratio of gamma functions to evaluate densities of normal, chi-squared and t-distributions.

Lemma 10. *For $z \geq 1/2$*

$$e^{-2/3} \leq \frac{\Gamma(z+\frac{1}{2})}{\Gamma(z)} \leq e^{1/6}\sqrt{z}\,.$$

Proof of Lemma 10. Since

$$\sqrt{2\pi}z^{z-1/2}e^{-z} \leq \Gamma(z) \leq \sqrt{2\pi}e^{1/6}z^{z-1/2}e^{-z}$$

for $z \geq 1/2$ from Stirling's formula (Olver et al., 2010, Sect. 5.6(i)), we have

$$\begin{aligned}\frac{\Gamma(z+\frac{1}{2})}{\Gamma(z)} &\geq e^{-2/3}\sqrt{z}\left(1+\frac{1}{2z}\right)^{z}\\ &\geq e^{-2/3}\sqrt{1/2}\left(1+\frac{1}{2\cdot 1/2}\right)^{1/2}\\ &= e^{-2/3}\,.\end{aligned}$$

Similarly we have

$$\begin{aligned}\frac{\Gamma(z+\frac{1}{2})}{\Gamma(z)} &\leq e^{-1/3}\sqrt{z}\left(1+\frac{1}{2z}\right)^{z}\\ &\leq e^{1/6}\sqrt{z}\,,\end{aligned}$$

which completes the proof. □

We prove Lemma 3 based on Cramér's theorem given below.

Proposition 11 (Dembo & Zeitouni, 1998, Ex. 2.2.38)**.** *Let $Z_1, Z_2, \cdots$ be i.i.d. random variables on $\mathbb{R}^d$. Then, for $\bar{Z} = n^{-1}\sum_{m=1}^{n} Z_m \in \mathbb{R}^d$ and any convex set $C \in \mathbb{R}^d$,*

$$\Pr[\bar{Z} \in C] \leq \exp\left(-n \inf_{z\in C} \Lambda^*(z)\right)\,,$$

where

$$\Lambda^*(z) = \sup_{\lambda\in\mathbb{R}^d}\{\lambda\cdot z - \log \mathrm{E}[e^{\lambda\cdot Z_1}]\}\,.$$

Proof of Lemma 3. Eq. (8) is straightforward from Cramér's theorem with $Z_m := X_{i,m}$ (see also e.g. Dembo & Zeitouni, 1998, Ex. 2.2.23).

Now we show (9). Let $Z_m = (Z_m^{(1)}, Z_m^{(2)}) := (X_{i,m},\, X_{i,m}^2) \in \mathbb{R}^2$. Then it is easy to see that the Fenchel-Legendre transform of the cumulant generating function of Z_i is given by

$$\begin{aligned}&\Lambda^*(z^{(1)}, z^{(2)})\\ &= \begin{cases} h\left(\frac{z^{(2)}-(z^{(1)})^2}{\sigma_i^2}\right) + \frac{(z^{(1)}-\mu_i)^2}{2\sigma_i^2}, & z^{(2)} > (z^{(1)})^2,\\ +\infty, & z^{(2)} \leq (z^{(1)})^2.\end{cases}\end{aligned}$$

Eq. (9) follows from

$$\begin{aligned}&\Pr[S_{i,n} \geq n\sigma^2]\\ &= \Pr[\bar{Z}^{(2)} - (\bar{Z}^{(1)})^2 \geq \sigma^2]\\ &\leq \exp\left(-n \inf_{(z^{(1)},z^{(2)}):\, z^{(2)}-(z^{(1)})^2\geq\sigma^2} \Lambda^*(z^{(1)}, z^{(2)})\right)\\ &\leq \exp\left(-nh\left(\frac{\sigma^2}{\sigma_i^2}\right)\right)\,,\end{aligned}$$

where the first and the second inequalities follow because $\{(z^{(1)}, z^{(2)}) : z^{(2)} - (z^{(1)})^2 \geq \sigma^2\}$ is a convex set and $h(x)$ is increasing in $x \geq 1$, respectively. □

Next we prove Lemma 4 based on Lemma 10.

Proof of Lemma 4. Letting

$$\begin{aligned}\tilde{A} &= \frac{\Gamma(\frac{n}{2}+\alpha)}{\sqrt{\pi(n+2\alpha-1)}\Gamma(\frac{n-1}{2}+\alpha)}\,,\\ x_0 &= \sqrt{\frac{n(n+2\alpha-1)}{S_{i,n}}}(\mu - \bar{x}_{i,n})\,,\end{aligned}$$

we can express $p_n(\mu|\hat{\theta}_{i,n})$ from (4) and (5) as

$$p_n(\mu|\hat{\theta}_{i,n}) = \tilde{A}\int_{x_0}^{\infty}\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n}{2}-\alpha}\mathrm{d}x. \quad (25)$$

This integral is bounded from above by

$$\begin{aligned}
&p_n(\mu|\hat{\theta}_{i,n})\\
&=\tilde{A}\int_{x_0}^{\infty}\frac{1}{x}\cdot x\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n}{2}-\alpha}\mathrm{d}x\\
&=\tilde{A}\left[\frac{1}{x}\cdot\frac{\frac{n-1}{2}+\alpha}{\frac{n-2}{2}+\alpha}\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n-2}{2}-\alpha}\right]_{\infty}^{x_0}\\
&\quad-\tilde{A}\int_{x_0}^{\infty}\frac{2}{x^2}\frac{\frac{n-1}{2}+\alpha}{\frac{n-2}{2}+\alpha}\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n-2}{2}-\alpha}\mathrm{d}x\\
&\le\frac{\tilde{A}}{x_0}\frac{\frac{n-1}{2}+\alpha}{\frac{n-2}{2}+\alpha}\left(1+\frac{x_0^2}{n+2\alpha-1}\right)^{-\frac{n-2}{2}-\alpha}.
\end{aligned}$$

From Lemma 10

$$\begin{aligned}
\frac{\tilde{A}}{x_0}\frac{\frac{n-1}{2}+\alpha}{\frac{n-2}{2}+\alpha}&=\frac{\Gamma(\frac{n-2}{2}+\alpha)}{2\sqrt{\pi n}\Gamma(\frac{n-1}{2}+\alpha)}\frac{\sqrt{S_{i,n}}}{\mu-\bar{x}_{i,n}}\\
&\le\frac{1}{2\sqrt{\pi n}\mathrm{e}^{-2/3}}\frac{\sqrt{S_{i,n}}}{\mu-\bar{x}_{i,n}}\\
&\le\frac{1}{\sqrt{n}}\frac{\sqrt{S_{i,n}}}{\mu-\bar{x}_{i,n}}
\end{aligned}$$

and we obtain (10).

On the other hand, the integral (25) is bounded from below by

$$\begin{aligned}
&p_n(\mu|\hat{\theta}_{i,n})\\
&=\tilde{A}\int_{x_0}^{\infty}\left(1+\frac{x^2}{n+2\alpha-1}\right)^{\frac{1}{2}}\\
&\qquad\cdot\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n+1}{2}-\alpha}\mathrm{d}x\\
&\ge\tilde{A}\int_{x_0}^{\infty}\frac{x}{\sqrt{n+2\alpha-1}}\left(1+\frac{x^2}{n+2\alpha-1}\right)^{-\frac{n+1}{2}-\alpha}\mathrm{d}x\\
&\qquad\qquad(\text{by }\sqrt{1+u^2}\ge u)\\
&=\frac{\tilde{A}}{\sqrt{n+2\alpha-1}}\left(1+\frac{x_0^2}{n+2\alpha-1}\right)^{-\frac{n-1}{2}-\alpha}.
\end{aligned}$$

From Lemma 10

$$\begin{aligned}
\frac{\tilde{A}}{\sqrt{n+2\alpha-1}}&=\frac{\Gamma(\frac{n}{2}+\alpha)}{2\sqrt{\pi}\Gamma(\frac{n+1}{2}+\alpha)}\\
&\ge\frac{1}{2\mathrm{e}^{1/6}\sqrt{\pi(\frac{n}{2}+\alpha)}}
\end{aligned}$$

and we obtain (11). □

Finally we prove Lemmas 8 and 9 on the regret bound of Thompson sampling.

Proof of Lemma 8. Let $n_i>0$ be arbitrary. Then

$$\begin{aligned}
&\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[J(t)=i,\ \mathcal{A}(t),\ \mathcal{B}_i(t)]\\
&\le\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[\tilde{\mu}_i(t)\ge-\epsilon,\ B_i(l)]\\
&\le n_i+\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[\tilde{\mu}_i(t)\ge-\epsilon,\ \mathcal{B}_i(t),\ N_i(t)\ge n_i]\,.
\end{aligned}\tag{26}$$

Under the condition $\{\mathcal{B}_i(t),\ N_i(t)=n\}$, the probability of the event $\tilde{\mu}_i(t)\ge-\epsilon=\mu^*-\epsilon$ is bounded from Lemma 4 as

$$\begin{aligned}
p_n(-\epsilon|\hat{\theta}_{i,n})&\le\frac{\sqrt{\sigma_i^2+\epsilon}}{\Delta_i-2\epsilon}\left(1+\frac{(\Delta_i-2\epsilon)^2}{\sigma_i^2+\epsilon}\right)^{-\frac{n}{2}-\alpha+1}\\
&=\frac{\sqrt{\sigma_i^2+\epsilon}}{\Delta_i-2\epsilon}\mathrm{e}^{-(n+2\alpha-2)D_{\mathrm{inf}}(\Delta_i-2\epsilon,\sigma_i^2+\epsilon)}.
\end{aligned}$$

Therefore the expectation of (26) is bounded as

$$\begin{aligned}
&\mathrm{E}\left[\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[J(t)=i,\ \mathcal{A}(t),\ \mathcal{B}_i(t)]\right]\\
&\le n_i+\sum_{t=K\bar{n}+1}^{T}\Pr\left[\tilde{\mu}_i(t)\ge-\epsilon,\ \mathcal{B}_i(t),\ N_i(t)\ge n_i\right]\\
&\le n_i+T\frac{\sqrt{\sigma_i^2+\epsilon}}{\Delta_i-2\epsilon}\mathrm{e}^{-(n_i+2\alpha-2)D_{\mathrm{inf}}(\Delta_i-2\epsilon,\sigma_i^2+\epsilon)}
\end{aligned}$$

and we complete the proof by letting $n_i=(\log T)/D_{\mathrm{inf}}(\Delta_i-2\epsilon,\sigma_i^2+\epsilon)+2-2\alpha$. □

Proof of Lemma 9. First we have

$$\begin{aligned}
&\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[J(t)=i,\ \mathcal{B}_i^c(t)]\\
&=\sum_{n=\bar{n}}^{T}\mathbb{1}\left[\bigcup_{t=K\bar{n}+1}^{T}\{J(t)=i,\ \mathcal{B}_i^c(t),\ N_i(t)=n\}\right]\\
&\le\sum_{n=\bar{n}}^{T}\mathbb{1}\left[\bar{x}_{i,n}\ge\mu_i+\delta\text{ or }S_{i,n}\ge n(\sigma_i^2+\epsilon)\right].
\end{aligned}$$

Therefore, from Lemma 3,

$$\begin{aligned}
&\mathrm{E}\left[\sum_{t=K\bar{n}+1}^{T}\mathbb{1}[J(t)=i,\ \mathcal{B}_i^c(t)]\right]\\
&\le\sum_{n=\bar{n}}^{T}\left(\mathrm{e}^{-n\frac{\epsilon^2}{2\sigma_i^2}}+\mathrm{e}^{-nh\left(1+\frac{\epsilon}{\sigma_i^2}\right)}\right)\\
&\le\frac{1}{1-\mathrm{e}^{-\frac{\epsilon^2}{2\sigma_i^2}}}+\frac{1}{1-\mathrm{e}^{-h\left(1+\frac{\epsilon}{\sigma_i^2}\right)}}\\
&=\mathrm{O}(\epsilon^{-2})+\mathrm{O}(\epsilon^{-2})=\mathrm{O}(\epsilon^{-2}).
\end{aligned}$$

□