Appendix: ERM for BayesOpt with STPs

1 DERIVATION

1.1 Expected Likelihood of Regret

Under the STP, the univariate PDF is [2]:

$$\mathcal{T}(\mu, \sigma, \nu) = \frac{C}{\sigma} \times \left(1 + \frac{[(y-\mu)/\sigma]^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}$$

with: mean μ, std. deviation σ and degrees-of-freedom ν. Define C as [2]:

$$C = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}$$

Probability of regret $r(\mathbf{x}) = f^* - f(\mathbf{x})$ [1] on a Student's-t posterior is:

$$\mathbb{P}[r(\mathbf{x})] = \frac{C}{\hat{\sigma}_{STP}(\mathbf{x})} \left(1 + \frac{[(\hat{\mu}_{STP}(\mathbf{x}) - f^* + r(\mathbf{x}))/\hat{\sigma}_{STP}(\mathbf{x})]^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}$$

$\mathbb{E}[r(\mathbf{x})]$ is expected regret [1] and for Student's-t is:

$$\int_0^{\infty} \frac{C \times r(\mathbf{x})}{\hat{\sigma}_{STP}(\mathbf{x})} \left(1 + \frac{[(\hat{\mu}_{STP}(\mathbf{x}) - f^* + r(\mathbf{x}))/\hat{\sigma}_{STP}(\mathbf{x})]^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dr(\mathbf{x})$$

1.2 Minimizing Expected Regret

The acquisition function minimizing STP expected regret is: $\alpha_{STP}^{ERM}(\mathbf{x}) = \mathbb{E}[r(\mathbf{x})]$ [1]:

$$\mathbb{E}[r(\mathbf{x})] = \int_t^{+\infty} \frac{C \times [\hat{\mu}_{STP}(\mathbf{x}) - f^* + r(\mathbf{x})]}{\hat{\sigma}_{STP}(\mathbf{x})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} \hat{\sigma}_{STP}(\mathbf{x}) dt$$

where: $t = \frac{\hat{\mu}_{STP}(\mathbf{x}) - f^* + r(\mathbf{x})}{\hat{\sigma}_{STP}(\mathbf{x})}$; $r(\mathbf{x}) = f^* - \hat{\mu}_{STP}(\mathbf{x}) + \hat{\sigma}_{STP}(\mathbf{x})t$; and $dr(\mathbf{x}) = \hat{\sigma}_{STP}(\mathbf{x})dt$ [1]:

$$= \int_{t=\frac{\hat{\mu}_{STP}(\mathbf{x}) - f^*}{\hat{\sigma}_{STP}(\mathbf{x})}}^{+\infty} C \times [f^* - \hat{\mu}_{STP}(\mathbf{x}) + \hat{\sigma}_{STP}(\mathbf{x})t] \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})] \int_{t=\frac{\hat{\mu}_{STP}(\mathbf{x}) - f^*}{\hat{\sigma}_{STP}(\mathbf{x})}}^{+\infty} C \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt$$

$$+ \hat{\sigma}_{STP}(\mathbf{x}) \int_{t=\frac{\hat{\mu}_{STP}(\mathbf{x}) - f^*}{\hat{\sigma}_{STP}(\mathbf{x})}}^{+\infty} Ct \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt \quad (1)$$

1.3 Eq. 1: 1st Term

Let: $z_s = \frac{f^* - \hat{\mu}_{STP}(\mathbf{x})}{\hat{\sigma}_{STP}(\mathbf{x})}$ [1]; By integration [1]:

$$[f^* - \hat{\mu}_{STP}(\mathbf{x})] \int_t^{+\infty} C \times \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})] \times \left(\int_{-\infty}^{+\infty} C \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt - \int_{-\infty}^{t} C \times \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt\right)$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})] \times \left(\int_{-\infty}^{+\infty} \phi_s(t)dt - \int_{-\infty}^{t} C \times \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt\right)$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})] \times [1 - \Phi_s\left(\frac{\hat{\mu}_{STP}(\mathbf{x}) - f^*}{\hat{\sigma}_{STP}(\mathbf{x})}\right)]$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})] \times \Phi_s\left(\frac{f^* - \hat{\mu}_{STP}(\mathbf{x})}{\hat{\sigma}_{STP}(\mathbf{x})}\right)$$

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})]\Phi_s(z_s) \quad (2)$$

1.4 Eq. 1: 2nd Term

Substitute: $s = 1 + \frac{t^2}{\nu}$; $\frac{ds}{dt} = \frac{2t}{\nu}$; $\frac{\nu}{2}ds = tdt$ [2]:

$$= \hat{\sigma}_{STP}(\mathbf{x}) \int_{1+\frac{z_s^2}{\nu}}^{+\infty} C \times \frac{\nu}{2} \times s^{-\frac{(\nu+1)}{2}} ds$$

$$= \hat{\sigma}_{STP}(\mathbf{x}) \times \frac{\nu}{2} \times \left(\frac{2}{1-\nu}\right) \times C \times s^{\frac{(1-\nu)}{2}} \Big|_{1+\frac{z_s^2}{\nu}}^{\infty}$$

$$= \hat{\sigma}_{STP}(\mathbf{x}) \times \frac{\nu}{2} \times \left(\frac{2}{1-\nu}\right) \times C \times s \times s^{-\frac{(\nu+1)}{2}} \Big|_{1+\frac{z_s^2}{\nu}}^{\infty}$$

$$= -\hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z_s^2}{\nu}\right) C \times s^{-\frac{(\nu+1)}{2}} \Big|_{1+\frac{z_s^2}{\nu}}^{\infty}$$

$$= -\hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z_s^2}{\nu}\right) \times C$$

$$\times \left(1 + \frac{[(\hat{\mu}_{STP}(\mathbf{x}) - f^* + r(\mathbf{x}))/\hat{\sigma}_{STP}(\mathbf{x})]^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} \Big|_{r(\mathbf{x})=0}^{r(\mathbf{x})=\infty}$$

$$= -\hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z_s^2}{\nu}\right) \times [\phi_s(\infty) - \phi_s(-z_s)]$$

$$= \hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z_s^2}{\nu}\right) [\phi_s(-z_s) - 0]$$

$$= \hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z_s^2}{\nu}\right) \phi_s(z_s) \quad (3)$$

where: $\phi_s(z_s) = \phi_s(-z_s)$ for symmetrical, standard Student's-t.

1.5 $\mathbb{E}[r(\mathbf{x})] = \alpha_{STP}^{ERM}(\mathbf{x})$

Adding Eq. 2 + Eq. 3:

$$= [f^* - \hat{\mu}_{STP}(\mathbf{x})]\Phi_s(z_s) + \hat{\sigma}_{STP}(\mathbf{x}) \left(\frac{\nu}{\nu-1}\right) \left(1 + \frac{z^2}{\nu}\right) \phi_s(z_s)$$

REFERENCES

[1] Vu Nguyen and Michael A. Osborne. 2020. Knowing The What But Not The Where in Bayesian Optimization. arXiv:arXiv:1905.02685

[2] Brendan Tracey and David Wolpert. 2018. Upgrading from Gaussian Processes to Student's-t Processes. 2018 AIAA Non-Deterministic Approaches Conference. https://doi.org/10.2514/6.2018-1659