A.2 Bound on cumulant function with sticky hyperparameter.

Applying this bound to the case of the sticky HDP-HMM requires some further effort. Applying the above to this case gives an equation analogous to Eq. 37.

$$c_D(\alpha\beta_k + \delta_k\kappa) \geq K\log\alpha - \log(\alpha+\kappa) + \log(\alpha\beta_k+\kappa) + \sum_{\substack{m=1 \\ m\neq k}}^{K+1} \log(\beta_m) \tag{38}$$

Evaluating this term requires computing $\mathbb{E}_q[\log(\alpha\beta_k + \kappa)]$, which has no closed form in terms of elementary functions. Instead of calculating this directly, we use the concavity of logarithms to lower bound this term:

$$\begin{aligned} \log(\alpha\beta_k+\kappa) &\geq \beta_k\log(\alpha+\kappa) + (1-\beta_k)\log(\kappa) \\ &= \beta_k\big(\log(\alpha+\kappa) - \log(\kappa)\big) + \log\kappa \end{aligned} \tag{39}$$

We justify this bound by noting that for even moderate κ (say, $\kappa > 10$), this inequality is very tight, as shown in Figure 6 in the main body of this thesis. Empirically, we find that κ almost always needs to be either zero, in which case we do not apply the bound, or in the low hundreds, in which case the gap in the bound is completely negligible.

Plugging this bound in Eq. 38, we find

$$\begin{aligned} c_D(\alpha\beta_k+\delta_k\kappa) \geq c_{sur-\kappa}(\alpha,\kappa,\beta,k) \triangleq K\log\alpha + \log(\kappa) - \log(\alpha+\kappa) + \sum_{\substack{m=1 \\ m\neq k}}^{K+1}\log(\beta_m) \\ + \beta_k\big(\log(\alpha+\kappa) - \log(\kappa)\big) \end{aligned} \tag{40}$$

This equation gives us a surrogate bound on the cumulant function for a single sticky transition vector. We next need to compute the sum of K sticky cumulant functions, plus one non-sticky cumulant function for the starting state.

Using our surrogate functions, we have

$$\begin{aligned} c_{sur}(\alpha,\beta) + \textstyle\sum_{k=1}^K c_{sur-\kappa}(\alpha,\kappa,\beta,k) = {} & (K^2+K)\log\alpha \\ & + K(\log(\kappa) - \log(\alpha+\kappa)) \\ & + \big(\log(\alpha+\kappa) - \log(\kappa)\big)\textstyle\sum_{k=1}^K \beta_k \\ & + \textstyle\sum_{k=1}^{K+1}\log\beta_k + \sum_{k=1}^K \sum_{\substack{m=1 \\ m\neq k}}^{K+1}\log(\beta_m) \end{aligned} \tag{41}$$

In the last line, the first sum comes from the stating state's cumulant function, the second nested sum comes from the others. We can combine these two terms to find that

$$\begin{aligned} c_{sur}(\alpha,\beta) + \textstyle\sum_{k=1}^K c_{sur-\kappa}(\alpha,\kappa,\beta,k) = {} & (K^2+K)\log\alpha + K(\log(\kappa) - \log(\alpha+\kappa)) \\ & + \big(\log(\alpha+\kappa) - \log(\kappa)\big)\textstyle\sum_{k=1}^K\beta_k \\ & + \log\beta_{K+1} + K\textstyle\sum_{k=1}^{K+1}\log(\beta_k) \end{aligned} \tag{42}$$

Finally, we can rewrite these sums of surrogate cumulants in terms of u instead of β, since the transformation between them is deterministic. We find

$$\begin{aligned} c_{sur}(\alpha,\beta(u)) + \textstyle\sum_{k=1}^K c_{sur-\kappa}(\alpha,\kappa,\beta(u),k) = {} & (K^2+K)\log\alpha + K(\log(\kappa) - \log(\alpha+\kappa)) \\ & + \big(\log(\alpha+\kappa) - \log(\kappa)\big)\textstyle\sum_{k=1}^K\beta_k(u) \\ & + \textstyle\sum_{k=1}^K\big(K\log u_k + [K(K+1-k)+1]\log(1-u_k)\big) \end{aligned} \tag{43}$$

We can now easily compute expectations of Eq. 43, since $\mathbb{E}_q[\beta_k]$ and $\mathbb{E}_q[\log u_k]$ have known closed forms for $q(u_k) = \text{Beta}(\hat{\rho}_k\hat{\omega}_k, (1-\hat{\rho}_k)\hat{\omega}_k)$.