A Proof of Proposition 1

Proposition 2. *For any π on Ω as in (1), and any $\epsilon > 0$, there are positive constants $w_i = w_i(\epsilon) > 0$, and normalized modular functions $m_i = m_i(\epsilon)$, $i \in \{1, \ldots, r\}$, such that, if we define $q(S) := \sum_{i=1}^r w_i \exp(m_i(S))$, for all $S \in \Omega$, then $d_{TV}(\pi, q) \le \epsilon$.*

Proof. Let $r = |\Omega|$, and let $(S_i)_{i=1}^r$ be an enumeration of all sets in Ω. For any $i \in \{1, \ldots, r\}$, and any $v \in V$, we define

$$m_{iv} = \begin{cases} \beta_i\,, & \text{if } v \in S_i \\ -\beta_i\,, & \text{otherwise} \end{cases},$$

and $m_i(S) = \sum_{v \in S} m_{iv}$, for all $S \in \Omega$. We also define

$$w_i = \frac{\pi(S_i)}{Z_i} = \frac{\pi(S_i)}{\left(1 + e^{\beta_i}\right)^{|S_i|} \left(1 + e^{-\beta_i}\right)^{|V \setminus S_i|}}.$$

Then, for all $i \in \{1, \ldots, r\}$, we have

$$\begin{aligned} &d_i(\beta_1, \ldots, \beta_r) := |\pi(S_i) - q(S_i)| \\ &= \left| \pi(S_i) - \sum_{j=1}^r \pi(S_j) \frac{e^{\beta_j |S_j|}}{\left(1 + e^{\beta_j |S_j|}\right)\left(1 + e^{-\beta_j |V \setminus S_j|}\right)} \right| \\ &\le \pi(S_i) \left(1 - \frac{e^{\beta_i |S_i|}}{\left(1 + e^{\beta_i |S_i|}\right)\left(1 + e^{-\beta_i |V \setminus S_i|}\right)} \right) + \\ &\quad \sum_{j : S_j \neq S_i} \pi(S_j) \frac{e^{\beta_j |S_i|}}{\left(1 + e^{\beta_j |S_j|}\right)\left(1 + e^{-\beta_j |V \setminus S_j|}\right)}. \end{aligned}$$

Note that both terms vanish if we let all $\beta_j \to \infty$. Therefore, for any $\delta > 0$, there are $\beta_{ij} = \beta_{ij}(\delta)$, for all $j \in \{1, \ldots, r\}$, such that $d_i(\beta_{i1}, \ldots, \beta_{ir}) \le \delta$.

Finally, choosing $\hat{\beta}_j := \max_{i \in \{1, \ldots, r\}} \beta_{ij}$, for all $j \in \{1, \ldots, r\}$, we get

$$d_{\mathrm{TV}}(\pi, q) = \frac{1}{2} \sum_{i=0}^r d_i(\hat{\beta}_1, \ldots, \hat{\beta}_r) \le 2^{n-1}\delta.$$

The result follows by choosing $\delta = \epsilon / 2^{n-1}$. □

B Ising Model on the Complete Graph

B.1 Bounds on Gibbs mixing

Theorem B1 (Theorem 15.3 in (Levin et al., 2008b))**.** *If $\beta > 1$, then the Gibbs sampler on $\textsc{Ising}_\beta$ has a bottleneck ratio $\Phi_* = \mathcal{O}\left(e^{-c(\beta)n}\right)$, where $c(\beta)$ is a non-decreasing function of β.*

Corollary 1 (cf. Theorem 15.3 in (Levin et al., 2008b))**.** *For $n \ge 3$, the Gibbs sampler on $\textsc{Ising}$ has spectral gap $\gamma^{\mathrm{G}} = \mathcal{O}\left(e^{-cn}\right)$, where $c > 0$ is a constant.*

Corollary 2 (cf. Theorem 2 in (Ding et al., 2009))**.** *For all $n \ge 3$, the restriction chains P_i^{G}, $i = 0, 1$, of the Gibbs sampler on $\textsc{Ising}$ have spectral gap $\gamma_i^{\mathrm{G}} = \Theta\left(\frac{2\ln(n) - 1}{n}\right)$.*

B.2 Bounds on M^3 mixing

M^3 sampler. The proposal distribution can be written as follows,

$$q(S) = \frac{1}{2}\left(\frac{\exp(-d_n(n-1)|S|)}{Z_1} + \frac{\exp(d_n(n-1)|S|)}{Z_2} \right), \tag{4}$$

where $Z_1 = (1 + \exp(-d_n(n-1)))^n$, and $Z_2 = (1 + \exp(d_n(n-1)))^n$.

Lemma B1 (Fact 6 in (Anari et al., 2016))**.** *The spectral gap of any reversible two-state chain P with stationary distribution π that satisfies $P(0,1) = c\,\pi(1)$ is c.*

Lemma 1. *For all $n \ge 10$, the projection chain $\bar{P}^{\mathrm{M}}$ of the M^3 sampler on $\textsc{Ising}$ has spectral gap $\bar{\gamma}^{\mathrm{M}} = \Omega(1)$.*

Proof. We define $\pi_k = \sum_{S \in \Omega, |S| = k} \pi(S)$, and $q_k = \sum_{S \in \Omega, |S| = k} q(S)$.

Bounding π_k. By definition, we can write $\pi_k = \hat{\pi}_k / Z$, where $\hat{\pi}_0 = 1$, and for $k > 0$ we have

$$\begin{aligned} \hat{\pi}_k &:= \binom{n}{k} \exp\left(-\frac{2\ln(n)}{n} k(n-k) \right) \\ &= \binom{n}{k} n^{-\frac{2k}{n}(n-k)} \\ &\le \left(\frac{en}{k}\right)^k n^{-\frac{2k}{n}(n-k)} \\ &= \left(\frac{e}{k}\right)^k n^{-k + \frac{2k^2}{n}}. \end{aligned}$$

It follows that

$$\ln(\hat{\pi}_k) \le -k \ln\left(\frac{k}{e}\right) + \left(\frac{2k^2}{n} - k\right)\ln(n). \tag{5}$$

It is easy to verify that for all $n \ge 10$ and $3 \le k \le \lfloor n/2 \rfloor$, it holds that $(2k - n)\ln(n) \le 0.5n \ln(k/e)$. Substituting this into (5), we get

$$\begin{aligned} &\ln(\hat{\pi}_k) \le -0.5k \ln\left(\frac{k}{e}\right) \\ \Rightarrow\ \ &\hat{\pi}_k \le \exp(-0.5k \ln(k/e)). \end{aligned}$$

Noting that, for all k, $\hat{\pi}_k \leq 1$, and using the fact that $\hat{\pi}_{n-k} = \hat{\pi}_k$, we get

$$\begin{aligned}
Z &= \sum_{k=0}^{n} \hat{\pi}_k \\
&\leq 2 \sum_{k=0}^{\lfloor n/2 \rfloor} \hat{\pi}_k \\
&= 2(\hat{\pi}_0 + \hat{\pi}_1 + \hat{\pi}_2 + \sum_{k=3}^{\lfloor n/2 \rfloor} \hat{\pi}_k) \\
&\leq 3 + \sum_{k=3}^{\lfloor n/2 \rfloor} \exp(-0.5k \ln(k/e)) \\
&\leq c_1, \qquad (6)
\end{aligned}$$

where c_1 is a constant.

Bounding q_k. First, it is easy to see that, for all $n \geq 1$, $Z_1 \leq 3$.

$$\begin{aligned}
q_k &= \sum_{S \in \Omega, |S|=k} q(S) \\
&\geq \sum_{S \in \Omega, |S|=k} \frac{1}{2} \frac{\exp(-d_n(n-1)|S|)}{Z_1} \qquad \text{(by (4))} \\
&\geq \frac{1}{6} \binom{n}{k} \exp(-d_n(n-1)|S|)
\end{aligned}$$

Bounding the spectral gap. For the projection chain $\bar{P}^{\mathrm{M}}$, we have

$$\begin{aligned}
\bar{P}^{\mathrm{M}}(0,1) &= \frac{1}{\bar{\pi}(0)} \sum_{\substack{S \in \Omega_i \\ R \in \Omega_j}} \pi(S) P^{\mathrm{M}}(S,R) \\
&\geq 2\pi_0 q_n \qquad (\bar{\pi}(0) = 1/2 \text{ by symmetry of } \pi) \\
&= 2\pi_0 q_0 \qquad \text{(by symmetry of } q) \\
&\geq 2 \frac{\hat{\pi}_0}{Z} \frac{1}{6} \qquad (q_0 \geq \tfrac{1}{6}) \\
&\geq 2 \frac{1}{c_1} \frac{1}{6} \qquad (\hat{\pi}_0 = 1) \\
&= c\bar{\pi}(1),
\end{aligned}$$

where $c = (2/3)c_1$.

Finally, it follows from Lemma B1 that the spectral gap of $\bar{P}^{\mathrm{M}}$ is c.

□

B.3 Bounds on combined sampler mixing

Lemma B2. *For all $n \geq 10$, the projection chain $\bar{P}^{\mathrm{C}}$ of the combined chain on* Ising *has spectral gap $\bar{\gamma}^{\mathrm{C}} = \Omega(1)$.*

Proof. By definition, $\bar{P}^{\mathrm{C}}(S,R) \geq \alpha \bar{P}^{\mathrm{M}}(S,R)$, therefore a simple comparison argument (e.g., Lemma 13.22 in (Levin et al., 2008b)) combined with the result of Lemma 1 gives us $\bar{\gamma}^{\mathrm{C}} \geq \alpha \bar{\gamma}^{\mathrm{M}} = \Omega(1)$. □

Lemma B3. *For all $n \geq 3$, each of the restriction chains P_i^{C} of the combined chain on* Ising *has spectral gap $\gamma_i^{\mathrm{C}} = \Theta\left(\frac{2\ln(n)-1}{2n}\right)$.*

Proof. By definition, $P_i^{\mathrm{C}}(S,R) \geq \alpha P_i^{\mathrm{G}}(S,R)$, therefore, using a comparison argument like above together with Lemma 2 gives us $\gamma_i^{\mathrm{C}} \geq \alpha \gamma_i^{\mathrm{G}} = \Theta\left(\frac{2\ln(n)-1}{2n}\right)$. □

Theorem B2 (Theorem 1 in (Jerrum et al., 2004))**.** *Given a reversible Markov chain P, if the spectral gap of its projection chain $\bar{P}$ is bounded below by $\bar{\gamma}$, and the spectral gaps of its restriction chains P_i are uniformly bounded below by γ_{min}, then the spectral gap of P is bounded below by*

$$\gamma = \min\left\{\frac{\bar{\gamma}}{3}, \frac{\bar{\gamma}\gamma_{min}}{3P_{max} + \bar{\gamma}}\right\},$$

where $p_{max} := \max_{i \in \{0,1\}} \max_{S \in \Omega_i} \sum_{R \in \Omega \setminus \Omega_i} P(S,R)$.

Theorem 2. *For all $n \geq 10$, the combined chain P^{C} on* Ising *has spectral gap*

$$\gamma^{\mathrm{C}} = \Omega\left(\frac{2\ln(n)-1}{2n}\right).$$

Proof. The result follows directly by combining the spectral gap bounds of Lemmas B2 and B3 in Theorem B2, and noting that $P_{\max} \leq 1$. □