Supplementary Material:

Spectral Approximate Inference

A. Proof of Claim 1

We first prove $\mathbf{f}(\Omega) \subset \mathcal{B}$. To this end we introduce the following inequalities for all $\mathbf{x} \in \{-1,1\}^n$:

$$|\langle \mathbf{u}_j, \mathbf{x} \rangle| \le \|\mathbf{u}_j\|_1, \qquad |\langle \mathbf{u}_j, \mathbf{x} \rangle - c \cdot f_j(\mathbf{x})| \le c \cdot \frac{n+1}{2} \tag{17}$$

which directly leads us to $|c \cdot f_j(\mathbf{x})| \le \|\mathbf{u}_j\|_1 + c \cdot (n+1)/2 \le c \cdot b_j$, and therefore $\mathbf{f}(\Omega) \subset \mathcal{B}$. Here, the first inequality of (17) is trivial. The second inequality of (17) is from the fact that the error between $c \cdot f_j(\mathbf{x})$ and $\langle \mathbf{u}_j, \mathbf{x} \rangle$ arises from a series of quantizations which is presented once in (8) and at most n times in (9). Since the quantization error is at most $c/2$ for each quantization, the second inequality of (17) holds.

Now we prove the bound of $|\mathcal{B}|$. From the definition of $\mathcal{B}$ and b_j, one can easily observe that the following bound on $|\mathcal{B}|$ holds:

$$\begin{aligned}
|\mathcal{B}| = \prod_{j=1}^{r} (2b_j + 1) &= 2^r \prod_{j=1}^{r} \left(\frac{1}{c} \|\mathbf{u}_j\|_1 + \frac{n}{2} + 1 \right) \\
&= 2^r \prod_{j=1}^{r} \left(\frac{1}{c} \sqrt{|\lambda_j|} \|\mathbf{v}_j\|_1 + \frac{n}{2} + 1 \right) \\
&\le 2^r \prod_{j=1}^{r} \left(\frac{1}{c} \sqrt{|\lambda_j| n} + \frac{n}{2} + 1 \right),
\end{aligned}$$

where the inequality is from $\|\mathbf{v}_j\|_1 \le \sqrt{n} \|\mathbf{v}_j\|_2 = \sqrt{n}$.

B. Proof of Claim 2

Claim 2 holds since

$$\begin{aligned}
t_i(\mathbf{k}) &= t_{i-1}(\mathbf{k}) + \sum_{\mathbf{x} \in \mathbf{f}^{-1}(\mathbf{k}) \cap (\mathcal{S}_i \setminus \mathcal{S}_{i-1})} \exp\big(\langle \boldsymbol{\theta}, \mathbf{x} \rangle\big) \\
&= t_{i-1}(\mathbf{k}) + \sum_{\mathbf{g}_i(\mathbf{x}) \in \mathbf{g}_i\left(\mathbf{f}^{-1}(\mathbf{k}) \cap (\mathcal{S}_i \setminus \mathcal{S}_{i-1})\right)} \exp\big(\langle \boldsymbol{\theta}, \mathbf{x} \rangle\big) \\
&= t_{i-1}(\mathbf{k}) + \sum_{\mathbf{x}' \in \mathbf{f}^{-1}(\mathbf{k} - [\widehat{u}_{ji}]_{j=1}^r) \cap \mathcal{S}_{i-1}} \exp\big(\langle \boldsymbol{\theta}, \mathbf{g}_i^{-1}(\mathbf{x}') \rangle\big) \\
&= t_{i-1}(\mathbf{k}) + \sum_{\mathbf{x}' \in \mathbf{f}^{-1}(\mathbf{k} - [\widehat{u}_{ji}]_{j=1}^r) \cap \mathcal{S}_{i-1}} \exp\big(2\theta_i + \langle \boldsymbol{\theta}, \mathbf{x}' \rangle\big) \\
&= t_{i-1}(\mathbf{k}) + \exp(2\theta_i) \cdot t_{i-1}(\mathbf{k} - [\widehat{u}_{ji}]_{j=1}^r).
\end{aligned} \tag{18}$$

In the above, $\mathbf{g}_i : \mathcal{S}_i \setminus \mathcal{S}_{i-1} \to \mathcal{S}_{i-1}$ is a bijection defined by $\mathbf{g}_i(\mathbf{x}) = \mathbf{x}'$ such that $x'_\ell = x_\ell$ except for $\ell = i$. The second equality of (18) is from replacing the summation over $\mathbf{f}^{-1}(\mathbf{k}) \cap (\mathcal{S}_i \setminus \mathcal{S}_{i-1})$ by that over $\mathbf{g}_i\big(\mathbf{f}^{-1}(\mathbf{k}) \cap (\mathcal{S}_i \setminus \mathcal{S}_{i-1})\big)$. The third equality of (18) is based on (9) which implies that for all $\mathbf{x} \in \mathcal{S}_i \setminus \mathcal{S}_{i-1}$, $\mathbf{x}' = \mathbf{g}_i(x)$ satisfies

$$\mathbf{f}(\mathbf{x}) - [\widehat{u}_{ji}]_{j=1}^r = \mathbf{f}(\mathbf{x}'). \tag{19}$$

Hence, (19) leads us to

$$\begin{aligned}\mathbf{g}_i\big(\mathbf{f}^{-1}(\mathbf{k}) \cap (\mathcal{S}_i \setminus \mathcal{S}_{i-1})\big) &= \mathbf{g}_i\big(\{\mathbf{x} \in \mathcal{S}_i \setminus \mathcal{S}_{i-1} : \mathbf{f}(\mathbf{x}) = \mathbf{k}\}\big) \\ &= \{\mathbf{x}' \in \mathcal{S}_{i-1} : \mathbf{f}(\mathbf{x}') = \mathbf{k} - [\widehat{u}_{ji}]_{j=1}^r\} \\ &= \mathbf{f}^{-1}(\mathbf{k} - [\widehat{u}_{ji}]_{j=1}^r) \cap \mathcal{S}_{i-1}\end{aligned}$$

and the third equality of (18) follows. The fourth equality of (18) directly follows from the definition of $\mathbf{g}_i$ that $x_i' = -1$ and $\big(\mathbf{g}_i^{-1}(\mathbf{x}')\big)_i = x_i = 1$.

C. Proof of Theorem 3

We first prove the computational complexity of Algorithm 1. Since each $t(\mathbf{k}), t'(\mathbf{k})$ possesses a memory of $O(|\mathcal{B}|)$ and $|\mathcal{B}| \leq 2^r \prod_{j=1}^r (\sqrt{|\lambda_j| n}/c + n/2 + 1)$ from Claim 1, the space complexity of Algorithm 1 is $O\big(2^r \prod_{j=1}^r (\sqrt{|\lambda_j| n}/c + n/2 + 1)\big)$. In addition, as the algorithm iterates n times while each iteration accesses to $t(\mathbf{k})$ and $t'(\mathbf{k})$, Algorithm 1 has $O\big(n 2^r \prod_{j=1}^r (\sqrt{|\lambda_j| n}/c + n/2 + 1)\big)$ computational complexity.

Now we provide the bound on the partition function approximation. First, we refer the following error bound introduced in the proof of Claim 1.

$$|\langle \mathbf{u}_j, \mathbf{x} \rangle - c \cdot f_j(\mathbf{x})| \leq c \cdot \frac{n+1}{2}. \tag{20}$$

Using (20), we provide a bound for $|\langle \mathbf{u}_j, \mathbf{x} \rangle^2 - (c \cdot f_j(\mathbf{x}))^2|$ as follows

$$\begin{aligned}|\langle \mathbf{u}_j, \mathbf{x} \rangle^2 - (c \cdot f_j(\mathbf{x}))^2| &= |\langle \mathbf{u}_j, \mathbf{x} \rangle - c \cdot f_j(\mathbf{x})||\langle \mathbf{u}_j, \mathbf{x} \rangle + c \cdot f_j(\mathbf{x})| \\ &\leq c \cdot \frac{n+1}{2} \left(|2\langle \mathbf{u}_j, \mathbf{x} \rangle| + c \cdot \frac{n+1}{2} \right) \\ &\leq \frac{1}{4} c^2 (n+1)^2 + c\sqrt{|\lambda_j| n}(n+1)\end{aligned} \tag{21}$$

where the first inequality is from (20) and the second inequality is from $|\langle \mathbf{u}_j, \mathbf{x} \rangle| \leq \|\mathbf{u}_j\|_1 \leq \sqrt{|\lambda_j| n}$. From (21), the error bound can be derived as

$$\begin{aligned}\frac{Z}{\widehat{Z}} &= \frac{\sum_{\mathbf{x} \in \Omega} \exp\left(\langle \boldsymbol{\theta}, \mathbf{x} \rangle + \sum_{j=1}^r \operatorname{sign}(\lambda_j) \langle \mathbf{u}_j, \mathbf{x} \rangle^2 \right)}{\sum_{\mathbf{x} \in \Omega} \exp\left(\langle \boldsymbol{\theta}, \mathbf{x} \rangle + \sum_{j=1}^r \operatorname{sign}(\lambda_j) (c \cdot f_j(\mathbf{x}))^2 \right)} \\ &\leq \max_{\mathbf{x} \in \Omega} \frac{\exp\left(\langle \boldsymbol{\theta}, \mathbf{x} \rangle + \sum_{j=1}^r \operatorname{sign}(\lambda_j) \langle \mathbf{u}_j, \mathbf{x} \rangle^2 \right)}{\exp\left(\langle \boldsymbol{\theta}, \mathbf{x} \rangle + \sum_{j=1}^r \operatorname{sign}(\lambda_j) (c \cdot f_j(\mathbf{x}))^2 \right)} \\ &\leq \max_{\mathbf{x} \in \Omega} \exp\left(\sum_{j=1}^r |\langle \mathbf{u}_j, \mathbf{x} \rangle^2 - (c \cdot f_j(x))^2| \right) \\ &\leq \exp\left(\frac{1}{4} r c^2 (n+1)^2 + c\sqrt{n}(n+1) \sum_{j=1}^r \sqrt{|\lambda_j|} \right)\end{aligned}$$

where the last inequality follows from (21). One can obtain a same bound for $\widehat{Z}/Z$ and this completes the proof of Theorem 3.

D. Proof of Claim 4

The result of Claim 4 directly follows from the following inequality:

$$
\begin{aligned}
\mathrm{KL}\Big(P_{\mathcal{Y}}(\mathbf{y}) || \prod_{j=1}^{r} q_j(y_j)\Big) &= -\sum_{\mathbf{y}\in\mathcal{Y}} P_{\mathcal{Y}}(\mathbf{y}) \log \prod_{j=1}^{r} q_j(y_j) - H(P_{\mathcal{Y}}(\mathbf{y})) \\
&= -\sum_{\mathbf{y}\in\mathcal{Y}} P_{\mathcal{Y}}(\mathbf{y}) \sum_{j=1}^{r} \log q_j(y_j) - H(P_{\mathcal{Y}}(\mathbf{y})) \\
&= -\sum_{j=1}^{r} \sum_{y_j:\mathbf{y}\in\mathcal{Y}} P_{\mathcal{Y}}(y_j) \log q_j(y_j) - H(P_{\mathcal{Y}}(\mathbf{y})) \\
&\geq -\sum_{j=1}^{r} \sum_{y_j:\mathbf{y}\in\mathcal{Y}} P_{\mathcal{Y}}(y_j) \log P_{\mathcal{Y}}(y_j) - H(P_{\mathcal{Y}}(\mathbf{y}))
\end{aligned}
$$

where the last inequality follows from the source coding theorem (Shannon, 1948).