1 Approximate Negative Correlation

Let \( D_{k \rightarrow l} \in \mathbb{R}^{\binom{n}{k} \times \binom{n}{l}} \) be the first half (down part) of the random walk between sets of size \( k \) and sets of size \( l \), defined by

\[
D_{k \rightarrow l}(S, T) = \begin{cases} 
\frac{1}{\binom{l}{i}} & \text{if } T \subseteq S, \\
0 & \text{otherwise}.
\end{cases}
\]

For any distribution \( v : \binom{n}{l} \rightarrow \mathbb{R}_{\geq 0} \), applying \( D_{k \rightarrow l} \) gives us a distribution \( v D_{k \rightarrow l} : \binom{n}{l} \rightarrow \mathbb{R}_{\geq 0} \) defined by the vector-matrix product

\[
v D_{k \rightarrow l}(T) = \sum_{S \in \binom{n}{l}} v(S) D_{k \rightarrow l}(S, T) = \frac{\sum_{S \supseteq T} v(S)}{\binom{k}{l}}.
\]

In class we showed that if \( v, \mu : \binom{n}{l} \rightarrow \mathbb{R}_{\geq 0} \) are distributions where \( v \) is arbitrary and \( \mu \) has a log-concave polynomial \( g_{\mu}(z_1, \ldots, z_n) = \sum_{S} \mu(S) \prod_{i \in S} z_i \), then \( D_{k \rightarrow l} \) shrinks their KL-divergence by a factor of \( l/k \). In other words

\[
D_{KL}(v D_{k \rightarrow l} \| \mu D_{k \rightarrow l}) \leq \frac{l}{k} D_{KL}(v \| \mu).
\]

As a reminder, the KL-divergence is defined as the \( f \)-divergence with \( f(x) = x \log x \):

\[
D_f(v \| \mu) := \mathbb{E}_{S \sim \mu} [f(v(S)/\mu(S))] - f(\mathbb{E}_{S \sim v} [v(S)/\mu(S)]).
\]

In the following, let \( \mu : \binom{n}{k} \rightarrow \mathbb{R}_{\geq 0} \) be a distribution with a log-concave polynomial.

1. Let \( T \in \binom{n}{l} \) be a fixed set and let \( v = \mathbb{I}_T \) be the distribution whose entire probability mass is on \( T \). Use the shrinkage of the KL-divergence for the operator \( D_{k \rightarrow l} \) to prove the following inequality relating \( \mu(T) \) and marginals of \( \mu \):

\[
\mathbb{P}_{S \sim \mu}[S = T] \leq \prod_{i \in T} \mathbb{P}_{S \sim \mu}[i \in S].
\]

2. Now let \( l \in \{0, 1, \ldots, k\} \) be some integer and let \( T \in \binom{n}{l} \) be a fixed set. Define \( v = \mathbb{I}_T \) to be the distribution whose entire probability mass is on the set \( T \). Let \( \mu' = \mu D_{k \rightarrow l} \). Note that \( v \) and \( \mu' \) are defined on the same ground set, and that \( \mu' \) has a log-concave polynomial. What is the value of \( \mu'(T) \)? How about the marginals of \( \mu' \)? Use the results of the previous part to obtain

\[
\mathbb{P}_{S \sim \mu}[T \subseteq S] \leq c(k, l) \prod_{i \in T} \mathbb{P}_{S \sim \mu}[i \in S],
\]

for some function \( c(k, l) \).

3. Show that we can take \( c(k, l) \leq l^n/l! \). Conclude that \( c(k, 2) \leq 2 \) and \( c(k, l) \leq e^l \).