I. Theorem and Proof

Given a real number $\phi > 0$, we define

$$\xi(\gamma) \triangleq e^{(-(1-p_{max})\cdot(1-\gamma)\cdot C)} + \frac{\left(\frac{2^\gamma}{2^\gamma-1}\right)^{2^\gamma} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma\cdot C}}{e\cdot 2^{\gamma-C}\cdot\Big(\gamma\cdot\ln 2-\ln(2^\gamma-1)\Big)\cdot(2^C\cdot\phi+1)}$$

where $p_{max} \triangleq \max_{j=1}^{r} p_j$. The following theorem establishes the optimality of $\hat{\mathbf{k}}$.

Theorem I.1. *For any ϵ and γ satisfying $\xi(\gamma) \le 1-\epsilon$, we can guarantee, with probability ϵ, $\hat{\mathbf{k}}$ is a feasible solution of $\mathbf{P}$ and $\Phi(\hat{\mathbf{k}}) \le \Phi(\mathbf{k}^*) + \phi\cdot\left(\frac{1}{2}\right)^{D_{p||q}}$.*

Proof. We first state an auxiliary lemma.

Lemma I.2. $\forall\delta\in(0,1)$*, it holds that*

$$\Pr\left[\sum_{j=1}^{r} p_j\cdot\hat{k}_j > (1+\delta)\cdot\gamma\cdot C\right] < \left(\frac{e^{p_{max}\delta}}{(1+\delta)^{1+\delta}}\right)^{\gamma\cdot C},$$

where $p_{max} \triangleq \max_{1\le j\le r}\{p_j\}$.

Proof. It follows from Markov inequality that

$$\begin{aligned}\Pr\left[\sum_{j=1}^{r} p_j\hat{k}_j > (1+\delta)\gamma C\right] &= \Pr\left[\exp\left(t\sum_{j=1}^{r} p_j\hat{k}_j\right) > \exp[t(1+\delta)\gamma C]\right]\\ &< \exp[-t(1+\delta)\gamma C]\cdot\mathbb{E}\left[\exp\left(t\sum_{j=1}^{r} p_j\hat{k}_j\right)\right].\end{aligned}$$

Denote $y_j \triangleq k_j^0 - \lfloor k_j^0\rfloor$. It follows from the definition of $\hat{k}_j$ that

$$\begin{aligned}\mathbb{E}\left[\exp\left(t\sum_{j=1}^{r} p_j\hat{k}_j\right)\right] &= \prod_{j=1}^{r} y_j\exp\left[p_j t\left(\lfloor k_j^0\rfloor+1\right)\right] + (1-y_j)\exp\left[p_j t\lfloor k_j^0\rfloor\right]\\ &= \prod_{j=1}^{r}\exp\left(p_j t\lfloor k_j^0\rfloor\right)[y_j\exp(p_j t)+1-y_j]\\ &< \prod_{j=1}^{r}\exp\left(p_j t\lfloor k_j^0\rfloor\right)\cdot y_j\cdot[\exp(p_j t)-1].\end{aligned}$$

Imposing $t = \ln(1+\delta)$, we have

$$\begin{aligned}
(1+\delta)^{-(1+\delta)\gamma C}\prod_{j=1}^{r}\exp\left(p_j t\left\lfloor k_j^0\right\rfloor\right)\cdot y_j\cdot[\exp(p_j t)-1] &= (1+\delta)^{-(1+\delta)\gamma C}\prod_{j=1}^{r}\exp\left(p_j t\left\lfloor k_j^0\right\rfloor\right)\cdot y_j\cdot[(1+\delta)^{p_j}-1]\\
&\le (1+\delta)^{-(1+\delta)\gamma C}\prod_{j=1}^{r}\exp\left(p_j t\left\lfloor k_j^0\right\rfloor\right)\cdot y_j\cdot\delta p_j\\
&\le (1+\delta)^{-(1+\delta)\gamma C}\prod_{j=1}^{r}\exp\left(p_j t\left\lfloor k_j^0\right\rfloor\right)\cdot\delta p_{max}\\
&\le (1+\delta)^{-(1+\delta)\gamma C}\prod_{j=1}^{r}\exp(p_j t k_j^0\delta p_{max})\\
&= (1+\delta)^{-(1+\delta)\gamma C}\cdot(1+\delta)^{\sum_j p_j k_j^0\delta p_{max}}\\
&= \left[\frac{e^{p_{max}\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\gamma C}.
\end{aligned}$$

The lemma is thus proved. □

Lemma I.2 leads to the following corollary.

Corollary I.3. *The probability that* $\hat{\mathbf{k}}$ *is a feasible solution of* $\mathbf{P}$ *is at least* $1-\exp\big(-(1-p_{max})\cdot(1-\gamma)\cdot C\big)$.

Proof. Imposing $(1+\delta)\cdot\gamma = 1$ and applying lemma I.2, we have

$$\begin{aligned}
\Pr\left[\sum_{j=1}^{r}p_j\hat{k}_j > C\right] &= \Pr\left[\sum_{j=1}^{r}p_j\hat{k}_j > (1+\delta)\gamma C\right]\\
&< \left[\frac{e^{p_{max}\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\gamma C}\\
&= \exp\Big(\big(p_{max}\cdot\delta-(1+\delta)\ln(1+\delta)\big)\cdot\gamma\cdot C\Big)\\
&= \exp\Big(\big(p_{max}\cdot\delta-\delta+o(\delta)\big)\cdot\gamma\cdot C\Big)\\
&= \exp\Big(-(1-p_{max})\cdot\delta\cdot\gamma\cdot C\Big)\\
&= \exp\Big(-(1-p_{max})\cdot(1-\gamma)\cdot C\Big).
\end{aligned}$$

The corollary is proved. □

We proceed to prove the main theorem.

Recall the proof of Theorem 3.1. We can solve $\mathbf{k^0}$ similarly as follows.

$$k_j^0 = \gamma\cdot C + D_{p||q}\cdot\log_{1-\frac{1}{2^\gamma}}\frac{1}{2}+\log_{1-\frac{1}{2^\gamma}}\frac{p_j}{q_j}.$$

We then have

$$\sum_{j=1}^{r}q_j\cdot\left(1-\frac{1}{2^\gamma}\right)^{k_j^0} = \left(1-\frac{1}{2^\gamma}\right)^{\gamma\cdot C}\cdot\left(\frac{1}{2}\right)^{D_{p||q}}.$$

We further get

$$\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{\hat{k}_j}\right] &= \sum_{j=1}^{r} q_j \cdot \mathbb{E}\left[\left(1-\frac{1}{2^\gamma}\right)^{\hat{k}_j}\right] \\
&= \sum_{j=1}^{r} q_j \cdot \left(\left(1-\frac{1}{2^\gamma}\right)^{\lfloor k_j^0 \rfloor + 1} \cdot (k_j^0 - \lfloor k_j^0 \rfloor) + \left(1-\frac{1}{2^\gamma}\right)^{\lfloor k_j^0 \rfloor} \cdot (1 - (k_j^0 - \lfloor k_j^0 \rfloor))\right) \\
&= \sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{\lfloor k_j^0 \rfloor} \cdot \left(1 - \frac{1}{2^a}(k_j^0 - \lfloor k_j^0 \rfloor)\right) \\
&= \sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{k_j^0} \cdot \left(1-\frac{1}{2^\gamma}\right)^{-(k_j^0 - \lfloor k_j^0 \rfloor)} \cdot \left(1 - \frac{1}{2^\gamma}(k_j^0 - \lfloor k_j^0 \rfloor)\right) \\
&\overset{x \triangleq k_j^0 - \lfloor k_j^0 \rfloor}{=} \sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{k_j^0} \cdot \left(1-\frac{1}{2^\gamma}\right)^{-x} \cdot \left(1 - \frac{x}{2^\gamma}\right) \\
&\leq A_{max} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma \cdot C} \cdot \left(\frac{1}{2}\right)^{D_{p||q}},
\end{aligned}$$

where

$$A_{max} = \max_{x \in [0,1)} \left(1-\frac{1}{2^\gamma}\right)^{-x} \cdot \left(1 - \frac{x}{2^\gamma}\right) = \frac{\left(\frac{2^\gamma}{2^\gamma - 1}\right)^{2^\gamma}}{e \cdot 2^\gamma \cdot \left(\gamma \cdot \ln 2 - \ln(2^\gamma - 1)\right)}.$$

We study the probability that $\Phi(\hat{\mathbf{k}}) > \Phi(\mathbf{k}^*) + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}$ by applying Markov inequality as follows.

$$\begin{aligned}
&\Pr\left[\Phi(\hat{\mathbf{k}}) > \Phi(\mathbf{k}^*) + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}\right] \\
&= \Pr\left[\sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{\hat{k}_j} > \sum_{j=1}^{r} q_j \cdot \left(\frac{1}{2}\right)^{k_j^*} + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}\right] \\
&< \frac{\mathbb{E}\left[\sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{\hat{k}_j}\right]}{\sum_{j=1}^{r} q_j \cdot \left(\frac{1}{2}\right)^{k_j^*} + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}} \\
&\leq \frac{A_{max} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma \cdot C} \cdot \left(\frac{1}{2}\right)^{D_{p||q}}}{\left(\frac{1}{2}\right)^{C} \cdot \left(\frac{1}{2}\right)^{D_{p||q}} + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}} \\
&= \frac{A_{max} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma \cdot C} \cdot 2^C}{2^C \cdot \phi + 1} \\
&= \frac{\left(\frac{2^\gamma}{2^\gamma - 1}\right)^{2^\gamma} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma \cdot C}}{e \cdot 2^{\gamma - C} \cdot \left(\gamma \cdot \ln 2 - \ln(2^\gamma - 1)\right) \cdot (2^C \cdot \phi + 1)}
\end{aligned}$$

Therefore, the probability that $\hat{\mathbf{k}}$ is a feasible solution of $\mathbf{P}$ and $\Phi(\hat{\mathbf{k}}) > \Phi(\mathbf{k}^*) + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}$ can be lower-bounded as follows.

$$
\begin{aligned}
&1-\Pr\left[\sum_{j=1}^{r} p_j \hat{k}_j > C\right] - \Pr\left[\sum_{j=1}^{r} q_j \cdot \left(1-\frac{1}{2^\gamma}\right)^{\hat{k}_j} > \sum_{j=1}^{r} q_j \cdot \left(\frac{1}{2}\right)^{k_j^*} + \phi \cdot \left(\frac{1}{2}\right)^{D_{p||q}}\right] \\
&> 1 - e^{(-(1-p_{max})\cdot(1-\gamma)\cdot C)} - \frac{\left(\frac{2^\gamma}{2^\gamma-1}\right)^{2^\gamma} \cdot \left(1-\frac{1}{2^\gamma}\right)^{\gamma \cdot C}}{e \cdot 2^{\gamma - C} \cdot \Big(\gamma \cdot \ln 2 - \ln(2^\gamma - 1)\Big) \cdot (2^C \cdot \phi + 1)} \\
&= 1 - \xi(\gamma) \geq 0.
\end{aligned}
$$

Theorem I.1 is thus proved. □