ON BOUNDING THE UNION PROBABILITY USING PARTIAL WEIGHTED INFORMATION – SUPPLEMENTARY MATERIAL

JUN YANG, FADY ALAJAJI, AND GLEN TAKAHARA

1. **Relation to the Cohen-Merhav bound.** Let $f_i(B) > 0$ and $m_i(\omega_B)$ be non-negative real functions. Then by the Cauchy-Schwarz inequality,

$$\left[\sum_{B:i\in B} f_i(B)p_B\right]\left[\sum_{B:i\in B} \frac{p_B}{f_i(B)} m_i^2(\omega_B)\right] \geq \left[\sum_{B:i\in B} p_B m_i(\omega_B)\right]^2. \tag{1}$$

Thus, using

$$P\left(\bigcup_{i=1}^N A_i\right) = \sum_{B\in\mathscr{B}}\left(\sum_{i=1}^N f_i(B)\right)p_B = \sum_{i=1}^N \sum_{B\in\mathscr{B}:i\in B} f_i(B)p_B. \tag{2}$$

we have

$$P\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \sum_{B:i\in B} f_i(B)p_B \geq \sum_{i=1}^N \frac{\left[\sum_{B:i\in B} p_B m_i(\omega_B)\right]^2}{\sum_{B:i\in B} \frac{p_B}{f_i(B)} m_i^2(\omega_B)}. \tag{3}$$

If we define $f_i(B)$ by

$$f_i(B) = \begin{cases} \frac{1}{|B|} = \frac{1}{\deg(\omega_B)} & \text{if } i \in B \\ 0 & \text{if } i \notin B \end{cases} \tag{4}$$

so that

$$P\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \sum_{B\in\mathscr{B}:i\in B} \frac{p_B}{\deg(\omega_B)} = \sum_{i=1}^N \sum_{\omega\in A_i} \frac{p(\omega)}{\deg(\omega)}. \tag{5}$$

then the inequality reduces to

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_{i=1}^N \frac{\left[\sum_{B:i\in B} p_B m_i(\omega_B)\right]^2}{\sum_{B:i\in B} p_B m_i^2(\omega_B)|B|} = \sum_i \frac{\left[\sum_{\omega\in A_i} p(\omega)m_i(\omega)\right]^2}{\sum_j \sum_{\omega\in A_i\cap A_j} p(\omega)m_i^2(\omega)}, \tag{6}$$

where the equality holds when $m_i(\omega) = \frac{1}{\deg(\omega)}$ (i.e., $m_i(\omega_B) = \frac{1}{|B|}$), which was first shown by Cohen and Merhav [1, Theorem 2.1].

When $m_i(\omega) = c_i > 0$, (6) reduces to the DC bound

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_i \frac{[c_i P(A_i)]^2}{\sum_j c_i^2 P(A_i\cap A_j)} = \sum_i \frac{P(A_i)^2}{\sum_j P(A_i\cap A_j)} = \ell_{\mathrm{DC}}. \tag{7}$$

Note that as remarked in [2], the DC bound can be seen as a special case of the lower bound

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_j c_i^2 P(A_i\cap A_j)}, \tag{8}$$

when $c_i = \frac{P(A_i)}{\sum_j P(A_i \cap A_j)}$. This is because

$$\frac{\left[\sum_i \left(\frac{P(A_i)}{\sum_j P(A_i\cap A_j)}\right) P(A_i)\right]^2}{\sum_i \sum_j \left(\frac{P(A_i)}{\sum_j P(A_i\cap A_j)}\right)^2 P(A_i \cap A_j)} = \frac{\left(\sum_i \frac{P(A_i)^2}{\sum_j P(A_i\cap A_j)}\right)^2}{\sum_i \left\{\left(\frac{P(A_i)}{\sum_j P(A_i\cap A_j)}\right)^2 \sum_j P(A_i \cap A_j)\right\}} = \frac{\ell_{\mathrm{DC}}^2}{\ell_{\mathrm{DC}}} = \ell_{\mathrm{DC}}. \tag{9}$$

Note that although $c_i > 0$ is not assumed in (8), one can always replace c_i by $|c_i|$ in (8) if $c_i < 0$ to get a sharper bound.

However, the lower bound in (8) is looser than the following two (left-most) lower bounds (which we later derive in (16) and (18)):

$$\sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)} \geq \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_k c_i c_k P(A_i \cap A_k)} \geq \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_j c_i^2 P(A_i \cap A_j)}, \tag{10}$$

where $c_i > 0$ for all i and the last inequality can be proved using $2c_i c_j \leq c_i^2 + c_j^2$.

2. **Relation to the Gallot-Kounias bound.** By the Cauchy-Schwarz inequality, or assuming $m_i(\omega) = 1$ in (1), we have

$$\left[\sum_{B:i\in B} f_i(B) p_B\right]\left[\sum_{B:i\in B} \frac{p_B}{f_i(B)}\right] \geq \left[\sum_{B:i\in B} p_B\right]^2 = P(A_i)^2. \tag{11}$$

Using $f_i(B)$ defined using $\boldsymbol{c}$ (note that $f_i(B) > 0$ is equivalent to $c_i > 0$ for all i), we have

$$\left[\sum_{B:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k}\right]\left[\sum_{B:i\in B} \left(\frac{\sum_{k\in B} c_k}{c_i}\right) p_B\right] \geq P(A_i)^2. \tag{12}$$

Note that

$$\sum_{B:i\in B} \left(\frac{\sum_{k\in B} c_k}{c_i}\right) p_B = \frac{1}{c_i} \sum_{k=1}^{N} \sum_{B:i\in B, k\in B} c_k p_B = \frac{\sum_k c_k P(A_i \cap A_k)}{c_i}. \tag{13}$$

Therefore, we have

$$\left[\sum_{B:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k}\right]\left[\frac{\sum_k c_k P(A_i \cap A_k)}{c_i}\right] \geq P(A_i)^2. \tag{14}$$

Then for all i,

$$\sum_{B:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k} \geq \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)} \tag{15}$$

By summing (15) over i, we get another new lower bound:

$$P\left(\bigcup_i A_i\right) \geq \sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)}. \tag{16}$$

Note that we can use Cauchy-Schwarz inequality again:

$$\left[\sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)}\right]\left[\sum_i c_i \sum_k c_k P(A_i \cap A_k)\right] \geq \left[\sum_i c_i P(A_i)\right]^2, \tag{17}$$

which yields

$$P\left(\bigcup_i A_i\right) \geq \sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)} \geq \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_k c_i c_k P(A_i \cap A_k)}. \tag{18}$$

Since the above inequality holds for any positive $\boldsymbol{c}$, we have

$$P\left(\bigcup_i A_i\right) \geq \max_{\boldsymbol{c}\in\mathbb{R}_+^N} \sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)} \geq \max_{\boldsymbol{c}\in\mathbb{R}_+^N} \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_k c_i c_k P(A_i \cap A_k)}. \tag{19}$$

One can show that by computing the partial derivative with respect to c_i and set it to zero that

$$\max_{\boldsymbol{c}\in\mathbb{R}^N} \sum_{i=1}^{N} \frac{c_i^2 P(A_i)^2}{c_i \sum_k c_k P(A_i \cap A_k)} = \max_{\boldsymbol{c}\in\mathbb{R}^N} \frac{\left[\sum_i c_i P(A_i)\right]^2}{\sum_i \sum_k c_i c_k P(A_i \cap A_k)} =: \ell_{\mathrm{GK}}, \tag{20}$$

where ℓ_{GK} is the Gallot-Kounias bound (see [2]), and the optimal $\tilde{\boldsymbol{c}}$ can be obtained from

$$\boldsymbol{\Sigma}\tilde{\boldsymbol{c}} = \boldsymbol{\alpha}, \tag{21}$$

where $\boldsymbol{\alpha} = (P(A_1), P(A_2), \ldots, P(A_N))^T$ and $\boldsymbol{\Sigma}$ is a $N \times N$ matrix whose (i,j)-th element equals to $P(A_i \cap A_j)$. Thus, we conclude that the lower bounds in (19) are equal to the GK bound as shown in [2] if $\tilde{\boldsymbol{c}} \in \mathbb{R}_+^N$; otherwise, the lower bounds in (19) are weaker than the GK bound.

3. Complete Results and Proof of Theorem 1.

Theorem. *For any given* $\boldsymbol{c}$ *that satisfies*

$$\sum_{k\in B} c_k \neq 0, \quad \textit{for all} \quad B \in \mathscr{B} \tag{22}$$

a new lower bound on the union probability is given by

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \sum_{i=1}^{N} \ell_i(\boldsymbol{c}) =: \ell_{\textit{NEW-I}}(\boldsymbol{c}), \tag{23}$$

where

$$\ell_i(\boldsymbol{c}) = P(A_i)\left(\frac{c_i}{\sum_{k\in B_1^{(i)}} c_k} + \frac{c_i}{\sum_{k\in B_2^{(i)}} c_k} - \frac{c_i \sum_k c_k P(A_i \cap A_k)}{P(A_i)\left(\sum_{k\in B_1^{(i)}} c_k\right)\left(\sum_{k\in B_2^{(i)}} c_k\right)}\right), \tag{24}$$

where $B_1^{(i)}$ *and* $B_2^{(i)}$ *are subsets of* $\{1,\ldots,N\}$ *that satisfy the following conditions.*

1. *If* $\frac{\sum_k c_k P(A_i\cap A_k)}{c_i P(A_i)} \geq 0$ *and* $\min_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} < 0$, *then*

$$\begin{aligned} B_1^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \quad \textit{s.t.} \quad \frac{\sum_{k\in B} c_k}{c_i} < 0, \\ B_2^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i}. \end{aligned} \tag{25}$$

2. *If* $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} \geq 0$ *and* $\min_{\{B: i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \geq 0$, *then*

$$\begin{aligned} B_1^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \ \ s.t. \ \ \frac{\sum_{k\in B} c_k}{c_i} \leq \frac{\sum_k c_k P(A_i\cap A_k)}{c_i P(A_i)}, \\ B_2^{(i)} &= \arg\min_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \ \ s.t. \ \ \frac{\sum_{k\in B} c_k}{c_i} \geq \frac{\sum_k c_k P(A_i\cap A_k)}{c_i P(A_i)}. \end{aligned} \tag{26}$$

3. *If* $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} < 0$ *and* $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} < \left\{ \max_{\{B:i\in B, \frac{\sum_{k\in B} c_k}{c_i} < 0\}} \frac{\sum_{k\in B} c_k}{c_i}, \right\}$, *then*

$$\begin{aligned} B_1^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i}, \quad s.t. \ \ \frac{\sum_{k\in B} c_k}{c_i} < 0, \\ B_2^{(i)} &= \arg\min_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i}. \end{aligned} \tag{27}$$

4. *If* $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} < 0$ *and* $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} \geq \left\{ \max_{\{B:i\in B, \frac{\sum_{k\in B} c_k}{c_i} < 0\}} \frac{\sum_{k\in B} c_k}{c_i} \right\}$, *then*

$$\begin{aligned} B_1^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i}, \\ B_2^{(i)} &= \arg\max_{\{B:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \ \ s.t. \ \ \frac{\sum_{k\in B} c_k}{c_i} \leq \frac{\sum_k c_k P(A_i\cap A_k)}{c_i P(A_i)}. \end{aligned} \tag{28}$$

Proof. Note that for the third and fourth cases, under the condition $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} < 0$, the elements of $\boldsymbol{c}$ cannot be all positive or negative, so the set $\{B : i \in B, \frac{\sum_{k\in B} c_k}{c_i} < 0\}$ is not empty. Therefore, the solutions of $B_1^{(i)}$ and $B_2^{(i)}$ always exist.

We note that $\ell_i(\boldsymbol{c})$ is the solution of

$$\begin{aligned} \min_{\{p_B: i\in B\}} \sum_{B:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k} \quad \text{s.t.} \quad & \sum_{B:i\in B} p_B = P(A_i), \\ & \sum_{B:i\in B} \left(\frac{\sum_{k\in B} c_k}{c_i}\right) p_B = \frac{1}{c_i}\sum_k c_k P(A_i\cap A_k), \\ & p_B \geq 0, \quad \text{for all} \quad B\in\mathscr{B} \quad \text{such that} \quad i\in B. \end{aligned} \tag{29}$$

From (29) we have that

$$\sum_{B:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k} \geq \ell_i(\boldsymbol{c}). \tag{30}$$

Summing (30) over i and using

$$P\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \sum_{B\in\mathscr{B}:i\in B} \frac{c_i p_B}{\sum_{k\in B} c_k} = \sum_{i=1}^N \sum_{\omega\in A_i} \frac{c_i p(\omega)}{\sum_{\{k:\omega\in A_k\}} c_k}. \tag{31}$$

we directly obtain $P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_{i=1}^N \ell_i(\boldsymbol{c})$.

Note that we can solve (29) using the same technique used in [4, 5]. Consider two subsets B_1 and B_2 such that $p_{B_1} \geq 0$ and $p_{B_2} \geq 0$, then denoting

$$b := \frac{\sum_k c_k P(A_i\cap A_k)}{c_i P(A_i)}, b_1 := \frac{\sum_{k\in B_1} c_k}{c_i}, b_2 := \frac{\sum_{k\in B_2} c_k}{c_i}, \tag{32}$$

then problem (29) reduces to

$$\ell_i(\boldsymbol{c}) = \min_{\{p_{B_1}, p_{B_2}\}} \quad \frac{p_{B_1}}{b_1} + \frac{p_{B_2}}{b_2} \quad \text{s.t.} \;\; p_{B_1} + p_{B_2} = P(A_i), \\ b_1 p_{B_1} + b_2 p_{B_2} = bP(A_i), \\ p_{B_1} \geq 0, \quad p_{B_2} \geq 0. \tag{33}$$

According to [4, Appendix B], one can get that

$$\ell_i(\boldsymbol{c}) = \min_{\{b_1, b_2 : b_1 \leq b \leq b_2\}} \quad P(A_i) \left(\frac{1}{b_1} + \frac{1}{b_2} - \frac{b}{b_1 b_2} \right), \tag{34}$$

and the partial derivative of $P(A_i)\left(\frac{1}{b_1} + \frac{1}{b_2} - \frac{b}{b_1 b_2}\right)$ with respect to b_1 and b_2 are (see [4, Appendix B, Eq. (B.3)]):

$$\frac{\partial \left[P(A_i) \left(\frac{1}{b_1} + \frac{1}{b_2} - \frac{b}{b_1 b_2} \right) \right]}{\partial b_1} = \frac{P(A_i)}{b_1^2} \left(\frac{b - b_2}{b_2} \right), \\ \frac{\partial \left[P(A_i) \left(\frac{1}{b_1} + \frac{1}{b_2} - \frac{b}{b_1 b_2} \right) \right]}{\partial b_2} = \frac{P(A_i)}{b_2^2} \left(\frac{b - b_1}{b_1} \right). \tag{35}$$

Note that the partial derivatives are not continuous at $b_1 = 0$ and $b_2 = 0$. Therefore, the solution depends on the following different scenarios.

1. If $b \geq 0$ and $\min_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} < 0$, the solutions of (34) are given by

$$b_1 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} \quad \text{s.t.} \quad \frac{\sum_{k \in B} c_k}{c_i} < 0, \\ b_2 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}. \tag{36}$$

2. If $b \geq 0$ and $\min_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} \geq 0$, the solutions of (34) are given by

$$b_1 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} \quad \text{s.t.} \quad \frac{\sum_{k \in B} c_k}{c_i} \leq b, \\ b_2 = \min_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} \quad \text{s.t.} \quad \frac{\sum_{k \in B} c_k}{c_i} \geq b. \tag{37}$$

3. If $b < 0$ and $b < \left\{ \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}, \quad \text{s.t.} \; \frac{\sum_{k \in B} c_k}{c_i} < 0 \right\}$,, the solutions of (34) are given by

$$b_1 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}, \quad \text{s.t.} \; \frac{\sum_{k \in B} c_k}{c_i} < 0, \\ b_2 = \min_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}. \tag{38}$$

4. If $b < 0$ and $b \geq \left\{ \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}, \quad \text{s.t.} \; \frac{\sum_{k \in B} c_k}{c_i} < 0 \right\}$,, the solutions of (34) are given by

$$b_1 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i}, \\ b_2 = \max_{\{B : i \in B\}} \frac{\sum_{k \in B} c_k}{c_i} \quad \text{s.t.} \quad \frac{\sum_{k \in B} c_k}{c_i} \leq b. \tag{39}$$

□

4. **Proof of Lemma 2.**

Lemma. *When $\boldsymbol{c} \in \mathbb{R}_+^N$, the lower bound $\ell_{NEW\text{-}I}(\boldsymbol{c})$ can be computed in pseudo-polynomial time, and can be arbitrarily closely approximated by an algorithm running in polynomial time.*

Proof. The problems in (25) to (28) are exactly the 0/1 knapsack problem with mass equals to value (see [3], the corresponding decision problem is also called subset sum problem). Unfortunately, the 0/1 knapsack problem is NP-hard in general.

However, if $\boldsymbol{c} \in \mathbb{R}_+^N$, i.e, the case (26), there exists a dynamic programming solution which runs in pseudo-polynomial time, i.e., polynomial in N, but exponential in the number of bits required to represent $\frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)}$ (see [3]). Furthermore, there is a fully polynomial-time approximation scheme (FPTAS), which finds a solution that is correct within a factor of $(1-\epsilon)$ of the optimal solution (see [3]). The running time is bounded by a polynomial and $1/\epsilon$ where ϵ is a bound on the correctness of the solution.

Therefore, if $\boldsymbol{c} \in \mathbb{R}_+^N$, one can get a lower bound for $\ell_i(\boldsymbol{c})$ in polynomial time which can be arbitrarily close to $\ell_i(\boldsymbol{c})$ by setting ϵ small enough, i.e.,

$$\ell_i(\boldsymbol{c}) \geq \ell_i^L(\boldsymbol{c}, \epsilon), \quad \lim_{\epsilon \to 0^+} \ell_i^L(\boldsymbol{c}, \epsilon) = \ell_i(\boldsymbol{c}). \tag{40}$$

The details are as follows. First, assume $\hat{B}_1$ and $\hat{B}_2$ are obtained by the FPTAS which satisfy

$$(1-\epsilon) \sum_{k \in B_1^{(i)}} c_k \leq \sum_{k \in \hat{B}_1^{(i)}} c_k \leq \sum_{k \in B_1^{(i)}} c_k, \quad \sum_{k \in B_2^{(i)}} c_k \leq \sum_{k \in \hat{B}_2^{(i)}} c_k \leq (1+\epsilon) \sum_{k \in B_2^{(i)}} c_k. \tag{41}$$

Then we have

$$\begin{aligned} \sum_{k \in B_1^{(i)}} c_k &\leq \min\left\{ \frac{\sum_{k \in \hat{B}_1^{(i)}} c_k}{1-\epsilon}, \frac{\sum_k c_k P(A_i \cap A_k)}{P(A_i)} \right\} =: b_1^{(i)}, \\ \sum_{k \in B_2^{(i)}} c_k &\geq \max\left\{ \frac{\sum_{k \in B_2^{(i)}} c_k}{1+\epsilon}, \frac{\sum_k c_k P(A_i \cap A_k)}{P(A_i)} \right\} =: b_2^{(i)}. \end{aligned} \tag{42}$$

Then one can get the arbitrarily close lower bound for $\ell_i(\boldsymbol{c})$ as

$$\ell_i(\boldsymbol{c}) \geq \ell_i^L(\boldsymbol{c}, \epsilon) := P(A_i) \left(\frac{c_i}{b_1^{(i)}} + \frac{c_i}{b_2^{(i)}} - \frac{c_i \sum_k c_k P(A_i \cap A_k)}{P(A_i) b_1^{(i)} b_2^{(i)}} \right). \tag{43}$$

Therefore, we can get a lower bound for $P\left(\bigcup_{i=1}^N A_i\right)$ that is arbitrarily close to $\ell_{\text{NEW-I}}(\boldsymbol{c})$ in polynomial time: $P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_i \ell_i(\boldsymbol{c}) \geq \sum_i \ell_i^L(\boldsymbol{c}, \epsilon)$. □

5. **Proof of Corollary 1.**

Corollary. *(New class of upper bounds $\hbar_{NEW\text{-}I}(\boldsymbol{c})$): We can derive an upper bound for any given $\boldsymbol{c} \in \mathbb{R}_+^N$ by*

$$\begin{aligned} P\left(\bigcup_i A_i\right) \leq & \left(\frac{1}{\min_k c_k} + \frac{1}{\sum_k c_k} \right) \sum_i c_i P(A_i) \\ & - \frac{1}{(\min_k c_k) \sum_k c_k} \sum_i \sum_k c_i c_k P(A_i \cap A_k) =: \hbar_{NEW\text{-}I}(\boldsymbol{c}). \end{aligned} \tag{44}$$

Proof. We get the upper bound by maximizing, instead of minimizing, the objective function of (29). More specifically, for any given $\boldsymbol{c} \in \mathbb{R}^+$, a upper bound can be obtained by

$$\hbar(\boldsymbol{c}) = \sum_{i=1}^{N} \hbar_i(\boldsymbol{c}), \tag{45}$$

where $\hbar_i(\boldsymbol{c})$ is defined by

$$\begin{aligned}\hbar_i(\boldsymbol{c}) := \max_{\{p_B : i \in B\}} \sum_{B: i \in B} \frac{c_i p_B}{\sum_{k \in B} c_k} \quad \text{s.t.} \quad & \sum_{B: i \in B} p_B = P(A_i), \\ & \sum_{B: i \in B} \left(\frac{\sum_{k \in B} c_k}{c_i} \right) p_B = \frac{1}{c_i} \sum_k c_k P(A_i \cap A_k), \\ & p_B \geq 0, \quad \text{for all} \quad B \in \mathscr{B} \quad \text{such that} \quad i \in B.\end{aligned} \tag{46}$$

The resulting upper bound is given by

$$\begin{aligned} P\left(\bigcup_i A_i\right) \quad & \leq \quad \sum_i \Bigg\{ P(A_i) \Bigg[\frac{c_i}{\min_k c_k} + \frac{c_i}{\sum_k c_k} \\ & \qquad - \frac{c_i^2}{(\min_k c_k) \sum_k c_k} \frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)} \Bigg] \Bigg\} \\ & = \quad \left(\frac{1}{\min_k c_k} + \frac{1}{\sum_k c_k} \right) \sum_i c_i P(A_i) \\ & \qquad - \frac{1}{(\min_k c_k) \sum_k c_k} \sum_i \sum_k c_i c_k P(A_i \cap A_k). \end{aligned} \tag{47}$$

□

6. **Proof of Theorem 2.**

Theorem. *Defining* $\mathscr{B}^- = \mathscr{B} \setminus \{1, \ldots, N\}$, $\tilde{\gamma}_i := \sum_k c_k P(A_i \cap A_k)$, $\tilde{\alpha}_i := P(A_i)$ *and*

$$\tilde{\delta} := \max_i \left[\frac{\tilde{\gamma}_i - \left(\sum_k c_k - \min_k c_k\right) \tilde{\alpha}_i}{\min_k c_k} \right]^+, \tag{48}$$

where $\boldsymbol{c} \in \mathbb{R}_+^N$, *another class of lower bounds is given by*

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \tilde{\delta} + \sum_{i=1}^N \ell_i'(\boldsymbol{c}, \tilde{\delta}) =: \ell_{NEW\text{-}II}(\boldsymbol{c}), \tag{49}$$

where

$$\begin{aligned}\ell_i'(\boldsymbol{c}, x) = & [P(A_i) - x] \cdot \\ & \left(\frac{c_i}{\sum_{k \in B_1^{(i)}} c_k} + \frac{c_i}{\sum_{k \in B_2^{(i)}} c_k} - \frac{c_i \sum_k c_k \left[P(A_i \cap A_k) - x \right]}{[P(A_i) - x] \left(\sum_{k \in B_1^{(i)}} c_k \right) \left(\sum_{k \in B_2^{(i)}} c_k \right)} \right),\end{aligned} \tag{50}$$

and

$$\begin{aligned} B_1^{(i)} &= \arg \max_{\{B\in\mathscr{B}^-:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \; s.t. \; \frac{\sum_{k\in B} c_k}{c_i} \leq \frac{\sum_k c_k\left[P(A_i\cap A_k)-x\right]}{c_i\left[P(A_i)-x\right]}, \\ B_2^{(i)} &= \arg \min_{\{B\in\mathscr{B}^-:i\in B\}} \frac{\sum_{k\in B} c_k}{c_i} \; s.t. \; \frac{\sum_{k\in B} c_k}{c_i} \geq \frac{\sum_k c_k\left[P(A_i\cap A_k)-x\right]}{c_i\left[P(A_i)-x\right]}. \end{aligned} \tag{51}$$

Proof. Let $x = p_{\{1,2,\dots,N\}}$ and consider $\sum_i \ell_i'(\boldsymbol{c},x)+x$ as a new lower bound where where $\ell_i'(\boldsymbol{c},x)$ equals to the objective value of the problem

$$\begin{aligned} &\min_{\{p_B:i\in B, B\in\mathscr{B}^-\}} \sum_{B:i\in B,B\in\mathscr{B}^-} \frac{c_i p_B}{\sum_{k\in B} c_k} \\ &\text{s.t.} \quad \sum_{B:i\in B,B\in\mathscr{B}^-} p_B = P(A_i)-x, \\ &\sum_{B:i\in B,B\in\mathscr{B}^-} \left(\frac{\sum_{k\in B} c_k}{c_i}\right) p_B = \frac{1}{c_i}\sum_k c_k\left[P(A_i\cap A_k)-x\right], \\ &p_B \geq 0, \quad \text{for all} \quad B\in\mathscr{B}^- \quad \text{such that} \quad i\in B. \end{aligned} \tag{52}$$

The solution of (52) exists if and only if $\min_k c_k \leq \frac{\tilde{\gamma}_i-(\sum_k c_k)x}{\tilde{\alpha}_i - x} \leq \sum_k c_k - \min_k c_k$, which gives $\max_i \left[\frac{\tilde{\gamma}_i-(\sum_k c_k-\min_k c_k)\tilde{\alpha}_i}{\min_k c_k}\right]^+ \leq x \leq \min_i \left[\frac{\tilde{\gamma}_i-(\min_k c_k)\tilde{\alpha}_i}{\sum_k c_k - \min_k c_k}\right]$. Therefore, the new lower bound can be written as

$$\min_x \left[x+\sum_{i=1}^N \ell_i'(\boldsymbol{c},x)\right] \text{ s.t. } \left[\frac{\tilde{\gamma}_i-(\sum_k c_k-\min_k c_k)\,\tilde{\alpha}_i}{\min_k c_k}\right]^+ \leq x \leq \frac{\tilde{\gamma}_i-(\min_k c_k)\tilde{\alpha}_i}{\sum_k c_k-\min_k c_k}, \forall i. \tag{53}$$

Next, we can prove that the objective function of (53) is non-decreasing with x. First, we prove $\ell_i'(\boldsymbol{c},x)$ is continuous when $\exists B'\in\mathscr{B}^-$ such that $\frac{\sum_k c_k[P(A_i\cap A_k)-x]}{c_i[P(A_i)-x]} = \frac{\sum_{k\in B'} c_k}{c_i}$. This can be proved by $\lim_{h\to 0^+} \ell_i'(\boldsymbol{c},x+h) = \lim_{h\to 0^+} \ell_i'(\boldsymbol{c},x-h) = \frac{c_i}{\sum_{k\in B'} c_k}$. Then one can prove that when $\frac{\sum_{k\in B_2^{(i)}} c_k}{c_i} < \frac{\sum_k c_k[P(A_i\cap A_k)-x]}{c_i[P(A_i)-x]} < \frac{\sum_{k\in B_1^{(i)}} c_k}{c_i}$, the partial derivative of $\ell_i'(\boldsymbol{c},x)+\frac{c_i}{\sum_k c_k}x$ w.r.t. x is non-negative:

$$\begin{aligned} \frac{\partial\left(\ell_i'(\boldsymbol{c},x)+\frac{c_i}{\sum_k c_k}x\right)}{\partial x} &= \frac{c_i}{\sum_k c_k} - \frac{c_i}{\sum_{k\in B_1^{(i)}} c_k} - \frac{c_i}{\sum_{k\in B_2^{(i)}} c_k} \\ &\quad + \frac{c_i\sum_k c_k}{\left(\sum_{k\in B_1^{(i)}} c_k\right)\left(\sum_{k\in B_2^{(i)}} c_k\right)} \\ &= \frac{c_i\left(\sum_k c_k - \sum_{k\in B_1^{(i)}} c_k\right)\left(\sum_k c_k - \sum_{k\in B_2^{(i)}} c_k\right)}{\left(\sum_k c_k\right)\left(\sum_{k\in B_1^{(i)}} c_k\right)\left(\sum_{k\in B_2^{(i)}} c_k\right)} \\ &= \frac{c_i\left(\sum_{k\notin B_1^{(i)}} c_k\right)\left(\sum_{k\notin B_2^{(i)}} c_k\right)}{\left(\sum_k c_k\right)\left(\sum_{k\in B_1^{(i)}} c_k\right)\left(\sum_{k\in B_2^{(i)}} c_k\right)} \geq 0. \end{aligned} \tag{54}$$

Therefore, the objective function of (53), $\sum_i \ell_i'(\boldsymbol{c},x)+x = \sum_i\left(\ell_i'(\boldsymbol{c},x)+\frac{c_i}{\sum_k c_k}x\right)$, is non-decreasing with x.

Finally, defining $\tilde{\delta}$ as in (48), the new lower bound can be written as $P\left(\bigcup_{i=1}^N A_i\right) \geq \tilde{\delta} + \sum_{i=1}^N \ell_i'(\boldsymbol{c}, \tilde{\delta})$, where $\ell_i'(\boldsymbol{c}, \tilde{\delta})$ can be obtained using the solution for $\ell_i(\boldsymbol{c})$ with $b = \frac{\sum_k c_k P(A_i \cap A_k)}{c_i P(A_i)}$ replaced by $\tilde{b} = \frac{\sum_k c_k \left[P(A_i \cap A_k) - \tilde{\delta}\right]}{c_i \left[P(A_i) - \tilde{\delta}\right]}$. □

7. **Proof of Corollary 2.**

Corollary. *(Improved class of upper bounds $\hbar_{NEW\text{-}II}(\boldsymbol{c})$): We can improve the upper bound $\hbar_{NEW\text{-}I}(\boldsymbol{c})$ in (44) by*

$$\begin{aligned} P\left(\bigcup_i A_i\right) \leq \min_i & \left\{ \frac{\sum_k c_k P(A_i \cap A_k) - (\min_k c_k) P(A_i)}{\sum_k c_k - \min_k c_k} \right\} \\ & + \left(\frac{1}{\min_k c_k} + \frac{1}{\sum_k c_k - \min_k c_k} \right) \sum_i c_i P(A_i) \\ & - \frac{1}{(\min_k c_k)(\sum_k c_k - \min_k c_k)} \sum_i \sum_k c_i c_k P(A_i \cap A_k), \\ & =: \hbar_{NEW\text{-}II}(\boldsymbol{c}). \end{aligned} \tag{55}$$

Note that the upper bound $\hbar_{NEW\text{-}II}(\boldsymbol{c})$ in (55) is always sharper than $\hbar_{NEW\text{-}I}$ in (44).

Proof. Letting $x = p_{\{1,\ldots,N\}}$. Defining $\mathscr{B}^- = \mathscr{B} \setminus \{1, \ldots, N\}$, then

$$\hbar'(\boldsymbol{c}) = \max_x \left[x + \sum_{i=1}^N \hbar_i'(\boldsymbol{c}, x) \right], \tag{56}$$

where $\hbar_i'(\boldsymbol{c}, x)$ is defined by

$$\begin{aligned} \hbar_i'(\boldsymbol{c}, x) := \max_{\{p_B : i \in B, B \in \mathscr{B}^-} & \sum_{B : i \in B, B \in \mathscr{B}^-} \frac{c_i p_B}{\sum_{k \in B} c_k} \\ \text{s.t.} \quad & \sum_{B : i \in B, B \in \mathscr{B}^-} p_B = P(A_i) - x, \\ & \sum_{B : i \in B, B \in \mathscr{B}^-} \left(\frac{\sum_{k \in B} c_k}{c_i} \right) p_B = \frac{1}{c_i} \sum_k c_k \left[P(A_i \cap A_k) - x \right], \\ & p_B \geq 0, \quad \text{for all} \quad B \in \mathscr{B}^- \quad \text{such that} \quad i \in B. \end{aligned} \tag{57}$$

The solution of $\hbar_i'(\boldsymbol{c}, x)$ is independent with x:

$$\begin{aligned} \hbar_i'(\boldsymbol{c}, x) &= (P(A_i) - x) \left(\frac{c_i}{\min_k c_k} + \frac{c_i}{\sum_k c_k - \min_k c_k} \right) \\ & \quad - \frac{c_i}{(\min_k c_k)(\sum_k c_k - \min_k c_k)} \sum_k c_k \left(P(A_i \cap A_k) - x \right), \\ &= P(A_i) \left(\frac{c_i}{\min_k c_k} + \frac{c_i}{\sum_k c_k - \min_k c_k} \right) \\ & \quad - \frac{c_i}{(\min_k c_k)(\sum_k c_k - \min_k c_k)} \sum_k c_k P(A_i \cap A_k), \end{aligned} \tag{58}$$

and the solution exists if and only if for all i

$$\min_k c_k \leq \frac{\sum_k c_k P(A_i \cap A_k) - (\sum_k c_k)x}{P(A_i) - x} \leq \sum_k c_k - \min_k c_k. \tag{59}$$

Thus, we get

$$\begin{aligned} &\left\{\max_i \frac{\sum_k c_k P(A_i \cap A_k) - \left(\sum_k c_k - \min_k c_k\right) P(A_i)}{\min_k c_k}\right\}^+ \\ &\leq x \leq \min_i \frac{\sum_k c_k P(A_i \cap A_k) - (\min_k c_k) P(A_i)}{\sum_k c_k - \min_k c_k} \end{aligned} \tag{60}$$

Therefore, we get the upper bound

$$\begin{aligned} P\left(\bigcup_i A_i\right) \leq \min_i &\left\{\frac{\sum_k c_k P(A_i \cap A_k) - (\min_k c_k) P(A_i)}{\sum_k c_k - \min_k c_k}\right\} \\ &+ \left(\frac{1}{\min_k c_k} + \frac{1}{\sum_k c_k - \min_k c_k}\right) \sum_i c_i P(A_i) \\ &- \frac{1}{(\min_k c_k)(\sum_k c_k - \min_k c_k)} \sum_i \sum_k c_i c_k P(A_i \cap A_k). \end{aligned} \tag{61}$$

□

REFERENCES

[1] A. Cohen and N. Merhav, Lower bounds on the error probability of block codes based on improvements on de Caen's inequality, *IEEE Transactions on Information Theory*, **50** (2004), 290–310.

[2] C. Feng, L. Li and J. Shen, Some inequalities in functional analysis, combinatorics, and probability theory, *The Electronic Journal of Combinatorics*, **17** (2010), 1.

[3] V. V. Vazirani, *Approximation Algorithms*, Springer-Verlag New York, Inc., New York, NY, USA, 2001.

[4] J. Yang, F. Alajaji and G. Takahara, Lower bounds on the probability of a finite union of events, URL http://arxiv.org/abs/1401.5543, Submitted, 2014.

[5] J. Yang, F. Alajaji and G. Takahara, New bounds on the probability of a finite union of events, in *2014 IEEE International Symposium on Information Theory (ISIT)*, 2014, 1271–1275.

E-mail address: jun@utstat.toronto.edu
E-mail address: fady@mast.queensu.ca
E-mail address: takahara@mast.queensu.ca