An inner bound technique on the normalized set of entropy vectors

John M. Walsh and Steven Weber (Drexel University)

Entropy vectors

1. Let $\mathbf{X} = (X_1, \dots, X_N)$ be N discrete random variables with finite support.
2. Let $h(\mathbf{X}_\mathcal{A})$ be the entropy of the subset of rvs $\mathbf{X}_\mathcal{A} = (X_i, i \in \mathcal{A})$ for some non-empty subset $\mathcal{A} \subseteq \{1, \dots, N\} \equiv [N]$.
3. Let $\mathbf{h} = (h(\mathbf{X}_\mathcal{A}),\ \mathcal{A} \subseteq [N])$ be the vector of entropies of each non-empty subset $\mathcal{A} \subseteq [N]$. Note $\mathbf{h}$ has $2^N - 1$ entries.
   - Example: for $N = 3$, $\mathbf{h} = (h_1, h_2, h_3, h_{12}, h_{13}, h_{23}, h_{123})$.
4. A vector $\mathbf{h} \in \mathbb{R}^{2^N - 1}$ is called entropic if its elements are the entropies for some joint distribution $\mathbf{p}$ on the N rvs $\mathbf{X}$.
5. The entropy vector region (EVR) $\bar{\Gamma}^*_N$ is the set of all entropic vectors.
6. Normalize by the number of bits for the support m: $\tilde{\mathbf{h}} = \mathbf{h}/\log_2 m$, and define $\bar{\Omega}^*_N$ as the set of normalized entropy vectors (Hassibi and Shadbackt 2007).

Two representations of joint distributions on binary rvs

Lemma 1. *The joint distribution $\mathbf{p}$ on N binary rvs*

$$\mathbf{p} = (p_{[N]}(\mathbf{x}) = \mathbb{P}(\mathbf{X} = \mathbf{x}),\ \mathbf{x} \in \{0,1\}^N),$$

and the collection of probabilities of each subset of rvs each taking the value zero

$$\mathbf{q} = (p_\mathcal{A}(\mathbf{0}) = \mathbb{P}(\mathbf{X}_\mathcal{A} = \mathbf{0}),\ \mathcal{A} \subseteq [N])$$

are equivalent in that there is a bijection between them.

Specifying the zero probabilities for all *strict* subsets leaves one degree of freedom, the joint probability of zero.

Example: $N = 2$ rvs. There exists an invertible matrix such that $\mathbf{q} = \mathbf{M}\mathbf{p}$:

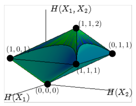
$$\begin{bmatrix} p_{(1)}(0) \\ p_{(2)}(0) \\ p_{(1,2)}(00) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & & \\ 1 & & 1 & \\ 1 & & & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p_{(1,2)}(00) \\ p_{(1,2)}(01) \\ p_{(1,2)}(10) \\ p_{(1,2)}(11) \end{bmatrix}.$$

A finite terminating algorithm to determine membership in the set of binary entropy vectors

1. The algorithm considers a sequence of subsets of increasing cardinality and finds the marginal distributions on these subsets consistent with the given entropy vector candidate.
2. The set of all possible binary distributions on a given set $\mathcal{B} \subset \mathcal{A}$ of size $|\mathcal{B}| = k - 1$ consistent with the given entropies is stored in the set $\mathcal{Q}_\mathcal{B}$.
3. Find candidate joint distributions $\mathbf{p}_\mathcal{A}$ on $\mathbf{X}_\mathcal{A}$ consistent with their marginals $\mathbf{p}_\mathcal{B}$ for each $\mathcal{B} \subseteq \mathcal{A}$.
4. For each candidate joint disbn $\mathbf{p}_\mathcal{A}$ set the remaining free variable using the specified entropy $h_\mathcal{A}$ (either 0,1,2 values).
5. Return all joint disbns $\mathbf{p}$ on $\mathbf{X}$ consistent with the given entropy vector $\mathbf{h}$ (if any).

An easy proof of Yeung's result that $\mathcal{P}_2 = \bar{\Omega}^*_2$ (and similarly $\mathcal{P}_3 = \bar{\Omega}^*_3$)

1. Let Φ_N be the collection of all entropy vectors for N *binary* rvs.
2. The Shannon-type inequalities Γ_N plus cardinality constraints ($\mathcal{B}_N = \{\mathbf{h} : h_i \leq 1, i = 1, \dots, N\}$), denoted $\mathcal{P}_N = \Gamma_N \cap \mathcal{B}_N$, form a polytope in entropy space that outer bounds $\mathrm{conv}(\Phi_N)$.
3. For $N = 2$, $\mathcal{P}_2$ is generated by the vertices below, each of which is in Φ_2 (using the algo.), as can be seen with the following constructions:
   - $(0,0,0)$: $(X_1, X_2) = (0,0)$ with probability one.
   - $(0,1,1)$: X_1 deterministic, X_2 uniform.
   - $(1,0,1)$: X_1 uniform, X_2 deterministic.
   - $(1,1,1)$: $X_1 = X_2$ with probability one, X_1 uniform.
   - $(1,1,2)$: X_1, X_2 independent and uniform.



   Therefore $\mathcal{P}_2 = \mathrm{conv}(\Phi_2)$.
4. Further, since $\mathrm{conv}(\Phi_N) \subseteq \bar{\Omega}^*_N \subseteq \mathcal{P}_N$, we have $\mathcal{P}_2 = \bar{\Omega}^*_2$.
5. Similar result and proof for $N = 3$ (both results originally due to Yeung).

But, not all normalized entropy vectors are binary: $\mathrm{conv}(\Phi_N) \neq \bar{\Omega}^*_N$ for $N \geq 4$

1. Enumerating the vertices of $\mathcal{P}_4$ gives a vertex at

h_1	h_2	h_3	h_4	h_{12}	h_{13}	h_{14}	h_{23}	h_{24}	h_{34}	h_{123}	h_{124}	h_{134}	h_{234}	h_{1234}
1	1/2	1/2	1/2	1	1	1	1	1	1	1	1	1	1	1

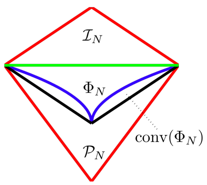
2. Evaluation of this point $\mathbf{h}$ using the membership algorithm asserts this point is not in Φ_4.
3. But, this point is in Ω^*_4 using the following construction: fix Z_1, Z_2 to be independent uniform bits, then take

$$X_1 = (Z_1, Z_2),\quad X_2 = (Z_1 \oplus Z_2, 0),\quad X_3 = (Z_1, 0),\quad X_4 = (Z_2, 0).$$

4. Can argue this point is an extreme point of $\bar{\Omega}^*_4$, and from here we can argue the result.

An algorithm to generate an inner bound for a given outer bound

1. Enumerate the extreme points of the polytope $\mathcal{P}_N$, by using a double description algorithm to convert the linear inequality representation into the generating vertices representation.
2. For each of these vertices, determine if they lie in Φ_N, using the membership algorithm. Keep only those vertices lying in Φ_N.
3. Take the convex hull of these vertices to get the polytope $\mathcal{I}_N$. $\mathcal{I}_N$ can be expressed in normal linear inequality form by using the double description method again.

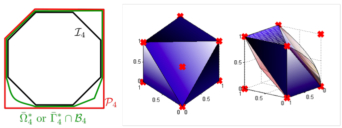


Every Shannon type inequality is tight for $N = 4$

Running the above algorithm using the $\mathcal{P}_4$ outer bound gives an inner bound with the following property.

Theorem 1. *For $N = 4$ every Shannon type inequality is tight on a full ($2^N - 2 = 14$) dimensional exposed face of conv(Φ_4) and hence $\bar{\Omega}^*_4$. Furthermore, for every such full dimensional exposed face $\mathcal{F}_\mathcal{I}$ of conv(Φ_4) that is contained in an exposed face $\mathcal{F}_\mathcal{P}$ of $\mathcal{P}_4$, the containment is strict ($\mathcal{F}_\mathcal{I} \subsetneq \mathcal{F}_\mathcal{P}$).*

Thus all Shannon inequalities are necessary for $N \geq 4$, but each such Shannon inequality can be improved!



Key properties of the inner bound algorithm

Theorem 2. *Given any polytopic outer bound $\mathcal{O}_N$ to $\bar{\Omega}^*_N$, the inner bound algorithm provides after a finite number of computations a polytopic inner bound $\mathcal{I}_N(\mathcal{O}_N)$ to conv(Φ_N) and hence $\bar{\Omega}^*_N$ and $\bar{\Gamma}^*_N$. Every exposed face of $\mathcal{O}_N$ which is also an exposed face of conv(Φ_N) will also be an exposed face of $\mathcal{I}_N(\mathcal{O}_N)$. Such an exposed face will also necessarily be an exposed face of $\bar{\Omega}^*_N$.*

The inner bound algorithm applies to *any* outer bound, i.e., not just $\mathcal{P}_N$. In particular, augmenting the Shannon outer bound with the recently discovered non-Shannon-type inequalities yields a better outer bound.

The quality of the inner bound improves with the quality of the outer bound. That is, given a sequence of increasingly tight outer bounds, our algorithm generates a corresponding sequence of increasingly tight inner bounds, in the sense that any face of the outer bound that is tight on an exposed face of Φ_N will generate a face of the inner bound that is also tight.

Extensions and future work

1. We further applied this general technique to the Shannon outer bound augmented with the non-Shannon type inequality due to Zhang and Yeung (1998), and established that none of the non-Shannon exposed faces are tight on a full face of $\mathrm{conv}(\Phi_4)$.
2. We have also extended the inner bound algorithm to the set of entropy vectors under one or more distribution constraints, meaning we are restricted to a subset of the full entropy vector space.
3. An important open question is whether or not $\mathrm{conv}(\Phi_N)$ is a polytope. If it is, then it admits a finite characterization in both a listing of its generating vertices and the inequalities characterizing its exposed faces, and it is of interest to obtain these listings. This is the subject of our current work.