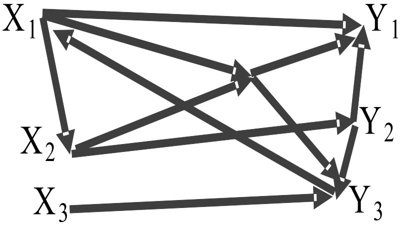
On Matroidal Solutions for Network Coding

Asaf Cohen, Michelle Effros, Salim ElRouayheb, Ralf Koetter

FLoWS PI Meeting, Austin TX, December 2007



Status Quo:

Network coding capacities are well-understood for some demands. Little theory is developed for general demands:

- Negative results.
- Suboptimal bounds.
- Network scaling laws.

New Insights:

Given an arbitrary solution for an arbitrary network, we try to show that there exists an equivalent (in terms of rates) matroidal solution. This may result in:

- A better understanding on which tools to apply. The set over which a solution is sought can be much smaller and more structured.
- Conclusions regarding very general networks.
- Usage of known results from matroid theory.

Important: Only an existence proof is required. Once such a result is established, we may limit our attention (and the toolbox we use) to matroidal solutions.

Polymatroids, Matroids and Entropy Vectors

- For a random vector X^n over a finite alphabet, the entropy function h maps each subset $\alpha \subseteq N = \{1, 2, \ldots, n\}$ to $\mathbb{R}^+$.
- Thus, each random vector is mapped to a point in $\mathbb{R}^{\mathcal{P}(N)}$, called an entropy vector, where $\mathcal{P}(N)$ is the power set of N.
- All these points satisfy:
  1. $h(\phi) = 0$,
  2. $h(I) \leq h(J)$ for $I \subseteq J \subseteq N$,
  3. $h(I) + h(J) \geq h(I \cup J) + h(I \cap J)$ for $I, J \subseteq N$.

  Hence (N, h) is a polymatroid with ground set N and rank h.
- Integer-valued polymatroids with a rank function bounded by the cardinality $h(I) \leq |I|$ are called matroids.

The Main Achievement Sought:

Conjecture

Any arbitrary network solution can be approximated with arbitrary precision (in terms of rates) by a matroidal solution over a possibly larger alphabet size.

That is, the entropy of any set of bits on the network is an integer.

Consequence

When seeking a solution for a network, we can limit the search to matroidal solutions.

Notes:

- The conjecture above aims at arbitrary random vectors, independent of any network setting or topology. In a sense, at the basis of the result stands a density result for distributions.
- Previous results regarding countable sets of entropy vectors which are dense in the set of all entropy vectors can point out both actual methods and general plausibility.

How it Works

- Approximate the original distribution by a dyadic distribution, with an excess entropy of at most 2 bits regardless of the alphabet size.
- Extend this result to show that all marginal distributions can also be made dyadic, with a negligible excess entropy as the alphabet size increases.
- Dyadic distributions translate to rational entropies, hence integer entropies for sufficiently large block size.
- The Huffman code of a dyadic distribution generates purely symmetric and independent bits.
- Since the conditional entropies are also approximated with an arbitrary precision, the resulting solution is also implementable.

Assumptions and Limitations

- We assume the links are noiseless (the common network coding setting).
- Large alphabet size means large block length, hence a large delay.
- We assume each source is i.i.d. and that the operations in the nodes are memoryless.

Initial Results

The following two theorems may lead to the required dyadic approximation:

Theorem

Any multivariate distribution over finite alphabets can be approximated by a dyadic distribution, with an excess entropy of at most 2 bits regardless of the alphabet size.

Theorem

Any point in the convex hull of K dyadic multivariate distributions can be approximated with an arbitrary precision by a dyadic distribution.

Using the theorem below, though, it is possible to show that each network solution has an integer-valued polymatroid approximation:

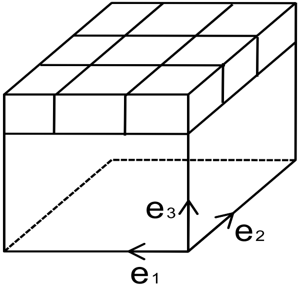
Theorem

The set of rational entropy vectors is dense in the cone of all entropy vectors. Namely, any random vector X^n can be approximated by $\tilde{X}^n$ such that $H(\tilde{X}_\alpha) \in \mathbb{Q}$ for any $\alpha \subseteq \{1, 2, \ldots, n\}$.

Tools

- Present each multi-variable distribution over a finite alphabet as a multi-dimsional cube.
- Approximate (in terms of entropy) the cube by several cubes, each having all non-empty entries equal to some negative power of two.
- Split each plane in the cube such that the row sums, column sums and total sum are powers of two.

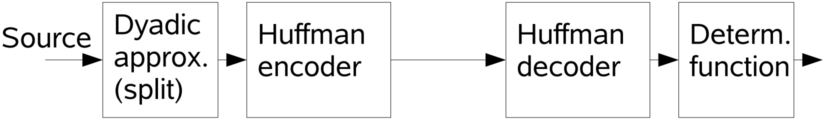
If these steps are done carefully, the excess entropy may be negligible as the block size increases.



				2^{-m}		
	2^{-m}		2^{-m}			
2^{-m}						2^{-m}
		2^{-m}		2^{-m}		
	2^{-m}				2^{-m}	
	2^{-m}				2^{-m}	
			2^{-m}			

Case Study: just one link...

Assume the simplest noiseless coding problem between a source and a destination. Can we construct an asymptotically optimal encoder which sends only purely symmetric and i.i.d. bits?
Note: the asymptotics are in the rate, not the bits' distribution!
Answer: Yes.



- Use a block of n source symbols (entropy $= nH$).
- Code the source according to its approximating dyadic distribution (entropy $\leq nH + 2$).
- At the decoder, apply a deterministic function to recover the original source.

Future Work

End of Phase:

- Continue the proof:
  - Show whether our proposed approximation indeed has negligible redundancy, and whether dyadic marginals result not only in i.i.d. bits per source, but also across the network.
  - Show whether negligibly small conditional entropies indeed result in an implementable solution.
- Extend the result to non-i.i.d. sources and non-memoryless nodes.
- Discuss concrete examples where the existence of a matroidal solution solves a problem or changes our perspective on it.
- Extend our understanding of subclasses of matroids and their relation to network solutions.

Community Challenge:

- Generalize known random coding techniques to matroidal solutions.
- Apply the results to different networking problems, such as noisy networks.