Relations between matroids and mutual information with application to MAC polar codes

Emmanuel Abbe, EPFL

December 4, 2009
Information...
Information Measures: one terminal

\[ P \in M_1(\mathcal{X}) \]

\[ \rightarrow \quad H(P) = \mathbb{E}_P \log \frac{1}{P} \]
\[
P \in M_1(\mathcal{X}) \quad \leftarrow \quad \text{distribution of } X
\]

\[
\rightarrow \quad H(P) = \mathbb{E}_P \log \frac{1}{P} = H(X)
\]

min. avg. nb. of bits to describe
lower bound on the compression of

Emmanuel Abbe
Information / Matroids / Links / Polar Codes
Information Measures: one terminal

\[ P \in M_1(\mathcal{X}) \quad \leftrightarrow \quad \text{distribution of } X \]

\[ \rightarrow \quad H(P) = \mathbb{E}_P \log \frac{1}{P} = H(X) \]

\[ \rightarrow \quad \text{min. avg. nb. of bits to describe } X \]

\[ \rightarrow \quad \text{lower bound on the compression of } X \]
\( X \sim P \)

\( W \in M_1(\mathcal{Y} | \mathcal{X}) \)
Information Measures: one terminal

\[ X \sim P \]

\[ W \in M_1(Y|X) \quad \leftarrow \quad \text{channel (distribution)} \]
\[ X \sim P \quad \text{input} \quad X \xrightarrow{W} Y \quad \text{output} \]

\[ W \in M_1(Y|X) \quad \text{channel (distribution)} \]

\[ I(P, W) = \mathbb{E}_\mu \log \frac{\mu}{\mu_X \times \mu_Y}, \quad \mu = P \circ W \]

\[ I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \]
Information Measures: one terminal

\[ X \sim P \quad \text{input} \quad X \xrightarrow{W} Y \quad \text{output} \]

\[ W \in M_1(Y|X) \quad \leftarrow \quad \text{channel (distribution)} \]

\[ \rightarrow \quad I(P, W) = \mathbb{E}_\mu \log \frac{\mu_X}{\mu_X \times \mu_Y}, \quad \mu = P \circ W \]

\[ = \quad I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \]

\[ \rightarrow \quad \text{mutual information} \]
Define the uniform mutual information of a channel $W \in M(\mathcal{Y}|\mathcal{X})$ by $I_U(W) := I(U_\mathcal{X} \circ W)$, where $U_\mathcal{X}$ is the uniform distribution on $\mathcal{X}$.

$I_U(W) = 1 \iff W =$ perfect channel

$I_U(W) = 0 \iff W =$ pure noise channel
Define the **uniform** mutual information of a channel \( W \in M(\mathcal{Y}|\mathcal{X}) \) by \( I_U(W) := I(U_X \circ W) \), where \( U_X \) is the uniform distribution on \( \mathcal{X} \).

\[
I_U(W) = 1 \quad \iff \quad W = \text{perfect channel}
\]

\[
I_U(W) = 0 \quad \iff \quad W = \text{pure noise channel}
\]

→ We know how to transfer information with low complexity on these extremal channels, and polarization bring them
A multiple access channel (MAC) with $m$ users is an element of $M_1(Y|X^m)$,

$$(X_1, \ldots, X_m) \xrightarrow{W} Y$$

**Definition**

The mutual information collection of a MAC $W \in M(Y|X^m)$ is

$$\{ I(X[S]; Y X[S^c]) , \quad S \subseteq \{1, \ldots, m\} \}$$

where $(X[1 \ldots m], Y)$ has joint distribution $(P_1 \times \ldots \times P_m) \circ W$.
A multiple access channel (MAC) with $m$ users is an element of $M_1(Y|X^m)$,

$$(X_1, \ldots, X_m) \xrightarrow{W} Y$$

**Definition**

The mutual information collection of a MAC $W \in M(Y|X^m)$ is

$$\{ I(X[S]; Y X[S^c]) , \ S \subseteq \{1, \ldots, m\} \}$$

where $(X[1 \ldots m], Y)$ has joint distribution $(P_1 \times \ldots \times P_m) \circ W$.

Define $\rho : 2^m \rightarrow \mathbb{R}$

$$S \mapsto I(X[S]; Y X[S^c])$$
Operational meaning:

\[ 0 \leq \sum_{s \in S} R_s \leq \rho(S), \quad \forall S \in 2^m \]

leads to the capacity region of non cooperating users over a memoryless MAC.

Example: \( m = 2 \):

\[ R_1 \leq I(X[1]; Y X[2]) \]
\[ R_2 \leq I(X[2]; Y X[1]) \]
\[ R_1 + R_2 \leq I(X[1] X[2]; Y) \]
Information Measures: multiple terminals

Operational meaning:

\[ 0 \leq \sum_{s \in S} R_s \leq \rho(S), \quad \forall S \in 2^m \]

leads to the capacity region of non cooperating users over a memoryless MAC.

Example: \( m = 2 \):

\[
\begin{align*}
R_1 & \leq I(X[1]; Y X[2]) \\
R_2 & \leq I(X[2]; Y X[1]) \\
R_1 + R_2 & \leq I(X[1] X[2]; Y)
\end{align*}
\]

Question: take uniform input distributions, what would be an extremal MAC?
A matroid $M$ is an ordered pair $(E, \mathcal{I})$, where $E$ is a finite set called the ground set and $\mathcal{I}$ is a collection of subsets of $E$ called the independent sets, which satisfies:

(I1) $\emptyset \in \mathcal{I}$.

(I2) If $I \in \mathcal{I}$ and $I' \subseteq I$, then $I' \in \mathcal{I}$.

(I3) If $I_1, I_2 \in \mathcal{I}$ and $|I_1| < |I_2|$, then there exists an element $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$. 
A matroid $M$ is an ordered pair $(E, \mathcal{I})$, where $E$ is a finite set called the ground set and $\mathcal{I}$ is a collection of a subsets of $E$ called the independent sets, which satisfies:

1. $(I1)$ $\emptyset \in \mathcal{I}$.
2. $(I2)$ If $I \in \mathcal{I}$ and $I' \subseteq I$, then $I' \in \mathcal{I}$.
3. $(I3)$ If $I_1, I_2 \in \mathcal{I}$ and $|I_1| < |I_2|$, then there exists an element $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$.

Examples:

1. **Vector matroids**: $E$ is the column set of a matrix (over a field), and independent sets defined by linearly independent columns.
2. **Graphic matroids**: $E$ is the set of edges of an undirected graph, and independent sets are collections of edges containing no cycle.
Definition

Let $M = (E, \mathcal{I})$. Define

- $\mathcal{D} = \mathcal{I}^c$, the collection of dependent sets
- $\mathcal{B}$, the collection of bases, i.e., maximal subsets of $E$ which are independent
- $\mathcal{C}$, the collection of circuits, i.e., minimal subsets of $E$ which are dependent.

Definition

We define a rank function $r : 2^m \to \mathbb{Z}_+$ such that for any $S \subseteq E$, $r(S)$ is given by the cardinality of a maximal independent set contained in $S$. 
The rank function satisfies the following properties.

\begin{itemize}
  \item [(R1)] If $X \subseteq E$, then $r(X) \leq |X|$.  
  \item [(R2)] If $X_1 \subseteq X_2 \subseteq E$, then $r(X_1) \leq r(X_2)$.  
  \item [(R3)] If $X_1, X_2 \subseteq E$, then 
  \[ r(X_1 \cup X_2) + r(X_1 \cap X_2) \leq r(X_1) + r(X_2). \]
\end{itemize}

Claim: this can also be used to define a matroid: an independent set is then a set with $r(X) = |X|$. 

Emmanuel Abbe
Information / Matroids / Links / Polar Codes
Theorem

Let $M$ be a matroid on $E$ with a set of bases $B$. Let $B^* = \{E - B : B \in B\}$. Then $B^*$ is the set of bases of a matroid on $E$. We denote this matroid by $M^*$ and call it the dual of $M$.

Lemma

If $r$ is the rank function of $M$, then the rank function of $M^*$ is given by

$$r^*(S) = r(S^c) + |S| - |E|.$$
Matroid representation

**Definition**
A matroid $M$ is representable over a field $F$ if it is isomorphic to a vector matroid over the field $F$.

If $A$ is a matrix representing $M$, we denote $M \cong M[A]$. A $\mathbb{F}_2$ representable matroid is called a **binary matroid**.

- The restriction of $M$ to $S$, is denoted by $M|S$ and means...
- The contraction of $M$ by $S$, is given by $M^*|S^c$
- A matroid $N$ that is obtained from $M$ by a sequence of restrictions and contractions is called a **minor** of $M$.

Theorem (Tutte)
A matroid is binary if and only if it has no minor that is $U_{4,2}$. $U_{4,2} = 4$ el. ground set and bases are the 2 el. sets.
A matroid $M$ is representable over a field $F$ if it is isomorphic to a vector matroid over the field $F$. If $A$ is a matrix representing $M$, we denote $M \cong M[A]$. A $\mathbb{F}_2$ representable matroid is called a binary matroid.

- The restriction of $M$ to $S$, is denoted by $M|S$ and means...
- The contraction of $M$ by $S$, is given by $M^*|S^c$
- A matroid $N$ that is obtained from $M$ by a sequence of restrictions and contractions is called a minor of $M$.

Theorem (Tutte) A matroid is binary if and only if it has no minor that is $U_{4,2}$.

$U_{4,2} = 4$ el. ground set and bases are the 2 el. sets
A polymatroid is a finite set $E$ equipped with a function $f : 2^m \rightarrow \mathbb{R}$, such that

\begin{align*}
(F1) & \quad f(\emptyset) = 0. \\
(F2) & \quad \text{If } X_1 \subseteq X_2 \subseteq E, \text{ then } f(X_1) \leq f(X_2). \\
(F3) & \quad \text{If } X_1, X_2 \subseteq E, \text{ then } f(X_1 \cup X_2) + f(X_1 \cap X_2) \leq f(X_1) + f(X_2).
\end{align*}

Such a $f$ is called a $\beta$-rank function.

A matroid is a polymatroid for which $f$ is integer valued and bounded.
links...
Let $E$ a finite set and $X[E] = \{X_i\}_{i \in E}$ be a random vector with distribution $P_E$.
Let $h(I) := h(X[I])$.

**Theorem (Lovász ’82, ...)**

$h(\cdot)$ is a $\beta$-rank function.
Hence, $(E, h)$ is a polymatroid.

**Definition**

A (poly)matroid $M$ is entropic if $M \cong M[h]$.

Some ref.: Han, Fujishige, Zhang, Matús and Yeung
If $|E| \leq 3$, all matroids are entropic, but otherwise...
Let $E$ be a finite set, and $X[E] \xrightarrow{W} Y$.

**Theorem (Hanly et al. ’94, ...)**

\[ \rho(S) = I(X[S]; Y | X[S^c]) \] is a $\beta$-rank function on $E$. Hence, $(E, \rho)$ is a polymatroid.

**Definition**

A (poly)matroid $M$ is MAC if $M \cong M[\rho]$

A (poly)matroid $M$ is BUMAC if it is MAC and if $P_1, \ldots, P_m$ are the uniform distributions on $X = \mathbb{F}^2$.

If $|E| \leq 3$, all matroids are BUMAC, but otherwise...
Polar codes application

Single-user setting:

turn $n$ independent channel uses into $n$ successive extremal channels,

→ either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$. 
Single-user setting:

turn $n$ independent channel uses into $n$ successive extremal channels,

$\rightarrow$ either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$.

Multi-user MAC setting:

turn $n$ independent channel uses into $n$ successive extremal MACs.
Single-user setting:

- Turn $n$ independent channel uses into $n$ successive extremal channels,
- Either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$.

Multi-user MAC setting:

- Turn $n$ independent channel uses into $n$ successive **extremal MACs**
- $\rho(S) \in \mathbb{Z}_+ \iff$ BUMAC matroids [EA and Telatar ’09]
Results

Theorem

A matroid is BUMAC if and only if it is binary.
Theorem

A matroid is BUMAC if and only if it is *binary*.

Theorem

A BUMAC matroid is “equivalent” to a *linear deterministic channel*:

if $M = M[W]$, and $A$ represents $M$, we have

$$I(AX[E]; Y) = \text{rank} A,$$

where $Y$ is the output through $W$. 
Example

Let a BUMAC with 2 users: \((X[1], X[2]) \xrightarrow{W} Y\) s.t.
\[I(X[1]; Y | X[2]) = I(X[2]; Y | X[1]) = I(X[1]X[2]; Y) = 1.\]

This defines a BUMAC matroid \(M\) on \(E = \{1, 2\}\) given by the ranks
\[(\emptyset, 1, 2, 12) \xrightarrow{r} (0, 1, 1, 1).\]

Then \(M\) is binary (thm 1), and in this case represented by
\[A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\]
Moreover, we have (thm 2)
\[I(A \begin{bmatrix} X[1] \\ X[2] \end{bmatrix}; Y) = 1, \text{ i.e., } I(X[1] + X[2]; Y) = 1.\]
proofs