The almost-entropic regions are not semialgebraic

Carolina Mejia, J. Andres Montoya
Universidad Nacional de Colombia, Bogota, Colombia.

October 26, 2015

Abstract

We prove that the almost-entropic region of order four is not semialgebraic, we get as a corollary Matus’ Theorem, which asserts that the almost-entropic regions of order larger than four are not polyhedral. We discuss the algorithmic consequences of our result.

Key words: Entropy, Information theory, entropic regions, information inequalities

Subject code classification: 68Qxx, 94Axx

We prove that the almost-entropic region of order four is not semialgebraic. It means that the set of entropic polymatroids of order 4 cannot be defined by a finite list of polynomial inequalities. Semialgebraic sets are closed under projections, and it implies that the almost entropic regions of order larger than four are non-semialgebraic. Notice that linear inequalities are a special type of polynomial inequalities. Thus, our result implies that the almost-entropic regions of order larger than four are not polyhedral, and it also suggests that those regions are undecidable and cannot be defined in a suitable way.

Consider the problem

Problem 1 (Min-Semialgebraic)

• Input: \((A, f, n)\), where \(A \subseteq \mathbb{R}^n\) is a convex, compact and semialgebraic set, and \(f : \mathbb{R}^n \to \mathbb{R}\) is a semialgebraic function

• Problem: Compute \(\min_{x \in A} \{f(x)\}\).

It follows from the work of Laserre \cite{9} that the above problem can be efficiently solved, and it follows from our previous work \cite{4} and \cite{5} that if the almost-entropic were semialgebraic, then many optimization problems coming from network coding \cite{14}, secret sharing \cite{2}, database theory \cite{6} (and more) could be reduced to the problem Min-Semialgebraic. Thus, it would be great news if we could prove that the almost-entropic regions are semialgebraic. Unfortunately, the results consigned in this paper rule out such a possibility.
1 Introduction

Proviso. In this work we only consider random variables whose support is a finite set.

Let $\overline{X} = (X_1, ..., X_n)$ be a $n$-tuple of random variables, it determines a
$n$-order entropic function $h_{\overline{X}} : (\mathcal{P}([n]) - \emptyset) \to \mathbb{R}$, which is defined by:

\[ h_{\overline{X}}(I) = H(X_I) \]

where $X_I$ denotes the tuple $(X_i)_{i \in I}$ and $H(X_I)$ its Shannon entropy. That is:
the entropic function (vector) $h_{\overline{X}}$ is constructed by collecting together the
entropies of the $2^n - 1$ nonempty subtuples of the tuple $\overline{X}$.

The set of all the entropic vectors of order $n$ spans a solid convex cone of
dimension $2^n - 1$ [14]. We use the symbol $\Gamma_n^*$ to denote this cone and the symbol
$\Gamma_n^{\infty}$ to denote its topological closure. The former set is called the entropic region
of order $n$, while the later is called the almost-entropic region of order $n$. Notice
that any element of $\Gamma_n^*$ can be expressed as a linear combination $\sum_{I \subseteq [n]} w_I e_I$,
where given $I \subseteq [n]$, the symbol $e_I$ denotes the vector of $\mathbb{R}^{2^n - 1}$ that is defined
by the condition

For all $h_{\overline{X}} \in \Gamma_n^*$, we have that $\langle h_{\overline{X}}, e_I \rangle = H(X_I)$

From now on, we fix $n = 4$. A typical element of $\overline{\Gamma}_4$ is a 15-dimensional
vector, say

\[ (h_1, h_2, h_3, h_4, h_12, h_13, h_14, h_{23}, h_{24}, h_{34}, h_{123}, h_{124}, h_{134}, h_{234}, h_{1234}) \]

where given $\{i_1, ..., i_k\} \subseteq \{1, 2, 3, 4\}$, the entry $h_{i_1, ..., i_k}$ is equal to $\langle h, e_{\{i_1, ..., i_k\}} \rangle$.
Given $i, j, k, l \in \{1, 2, 3, 4\}$, and given $h \in \overline{\Gamma}_4$ we set

- $I_h(i : j) = h_i + h_j - h_{ij}$.
- $I_h(i : j | k) = h_{ik} + h_{jk} - h_{ijk} - h_k$.
- $H_h(i | j, k, l) = h_{ijkl} - h_{jkl}$

Let $\overline{X} = (X_1, ..., X_4)$, and suppose that $h_{\overline{X}}$ is the entropic vector that is
associated to the tuple $\overline{X}$, notice that:

- $I_h(i : j)$ is equal to the mutual information of $X_i$ and $X_j$, and it is also
equal to $\langle h_{\overline{X}}, e_{\{i\}} + e_{\{j\}} - e_{\{i,j\}} \rangle$. We use the symbol $I(i : j)$ to denote
the vector $e_{\{i\}} + e_{\{j\}} - e_{\{i,j\}}$.
- $I_h(i : j | k)$ is the mutual information of $X_i$ and $X_j$ given $X_k$, and it is
also equal to $\langle h_{\overline{X}}, e_{\{i,k\}} + e_{\{j,k\}} - e_{\{i,j,k\}} - e_{\{k\}} \rangle$. We use the symbol
$I(i : j | k)$ to denote the vector $e_{\{i,k\}} + e_{\{j,k\}} - e_{\{i,j,k\}} - e_{\{k\}}$, and we use
the symbol $H(i | j, k, l)$ to denote the vector $e_{\{i,j,k,l\}} - e_{\{j,k,l\}}$. 

2
If one wants to describe the set $\Gamma_4^*$, he can choose to work either with the canonical basis $\{e_I : \emptyset \neq I \subseteq \{1, 2, 3, 4\}\}$, or with some other (more appropriate) basis. Matus and Studenyi (see [11]) introduced the so called natural basis, which is given by:

\[
\begin{align*}
v_1 &= -I(1 : 2) + I(1 : 2 | 3) + I(1 : 2 | 4) + I(3 : 4). \\
v_2 &= I(1 : 2 | 3). \\
v_3 &= I(1 : 2 | 4). \\
v_4 &= I(1 : 3 | 2). \\
v_5 &= I(1 : 4 | 2). \\
v_6 &= I(2 : 3 | 1). \\
v_7 &= I(2 : 4 | 1). \\
v_8 &= I(3 : 4 | 1). \\
v_9 &= I(3 : 4 | 2). \\
v_{10} &= I(3 : 4).
\end{align*}
\]

\[
\begin{align*}
v_{11} &= I(1 : 2 | 3, 4) = e_{\{1,3,4\}} + e_{\{2,3,4\}} - e_{\{1,2,3,4\}} - e_{\{3,4\}}. \\
v_{12} &= H(1 | 2, 3, 4). \\
v_{13} &= H(2 | 1, 3, 4). \\
v_{14} &= H(3 | 1, 2, 4). \\
v_{15} &= H(4 | 1, 2, 3).
\end{align*}
\]

Notice that $v_1$ is the vector encoding the famous Ingleton inequality [7]. Ingleton inequality is a linear rank inequality, but it is not a linear information inequality, it means that given a linear rank function $h$, the inequality $\langle h, v_1 \rangle \geq 0$ holds, while there exists an entropic function $h$ such that $\langle h, v_1 \rangle < 0$. We will use the symbol $\mathcal{I}$ to denote the vector $v_1$.

Notice also that for all $i \geq 2$, the vector $v_i$ encodes a Shannon inequality asserting that a certain Shannon information measure is positive, that is: for all $i \geq 2$, and for all almost entropic vector $h$ it happens that $\langle h, v_i \rangle \geq 0$.

Given $A \subseteq \mathbb{R}^n$, the symbol $A^\circ$ denotes its polar set. Recall that the polar of $A$ is the set

$$\{v \in \mathbb{R}^n : \forall w \in A \langle (v, w) \rangle \geq 0\}$$

We are interested in the set $\left(\Gamma_4^*\right)^\circ$, which is the polar of the almost-entropic region $\Gamma_4^*$, and which encodes the linear inequalities that are satisfied by the joint entropies of all the 4-tuples of random variables.

## 2 Semialgebraicity

We know that the set $\Gamma_4^*$ cannot be defined by a finite list of linear inequalities [11]. It does not rule out the possibility of defining the set $\Gamma_4^*$ by a finite list of polynomial inequalities. Notice that the set $\{(x, y) : x^2 + y^2 \leq 1\}$, which is defined by a single polynomial inequality, cannot be defined by a finite list of linear inequalities. Chan and Grant observed that there exist polynomial information inequalities, which are not entailed by any finite set of linear information inequalities [11], they asked if the entropic regions can be defined by a finite list of polynomial inequalities, that is, they asked if the entropic regions are basic semialgebraic sets.
Definition 2 Given $A \subset \mathbb{R}^n$, we say that $A$ is a basic semialgebraic set, if and only if, there exist polynomials $p_1(X_1, \ldots, X_n), \ldots, p_k(X_1, \ldots, X_n)$ such that

$$A = \left\{(a_1, \ldots, a_n) \in \mathbb{R}^n : \bigwedge_{i \leq k} (p_i(a_1, \ldots, a_n)) \geq 0 \right\}$$

and we say that $A$ is a semialgebraic set, if and only if, the set $A$ is a finite union of basic semialgebraic sets.

Semialgebraic sets have the following pleasant features (the interested reader can consult the reference [13]):

- Given a closed convex set $A$, the set $A$ is semialgebraic, if and only if, the set $A^\circ$ is semialgebraic.
- Semialgebraic sets are closed under projections.
- If $A \subset \mathbb{R}^n$ is semialgebraic, and $L$ is two-dimensional plane contained in $\mathbb{R}^n$, then $A \cap L$ is also semialgebraic.

Then, we have that if we could prove that there exists a two-dimensional plane $L$, contained in $\mathbb{R}^{15}$, and such that $(\Gamma_4^\ast)^\circ \cap L$ is not semialgebraic, we would get as a corollary that the almost-entropic regions of order larger than four are not semialgebraic. Notice that any polyhedral set is semialgebraic, it means that we would also get as a corollary the famous Matus’ Theorem.

3 The theorem

We begin with the following observation:

Suppose that $L$ is a two-dimensional plane contained in $\mathbb{R}^{15}$, and that we have fixed a rectangular system of coordinates for $L$. Suppose that we have a sequence $\{v_i\}_{i \geq 1}$, which is contained in $(\Gamma_4^\ast)^\circ \cap L$, and let $f : \mathbb{N} \rightarrow \mathbb{R}$ be a function such that $v_i$, as an element of $L$, is equal to $(i, f(i))$. Then, if the function $f$ is not semialgebraic, and the sequence $\{v_i\}$ is close enough to the border of $(\Gamma_4^\ast)^\circ \cap L$, the set $(\Gamma_4^\ast)^\circ$ cannot be a semialgebraic set.

To make the above observation become a functional one, we have to define the meaning of being close enough to the border of $(\Gamma_4^\ast)^\circ \cap L$, as well as the meaning of function $f$ being non-semialgebraic. So, let us proceed.

Given $p \in \mathbb{R}^2$ and given $B \subset \mathbb{R}^2$, we use the symbol $d(p, B)$ to denote the euclidean distance between the point $p$ and the set $B$.

Lemma 3 Let $A \subset \mathbb{R}^2$ be a closed convex set, and let $\{(i, f(i))\}_{i \geq 1}$ be a sequence contained in $A$. Suppose that $R$ is a ray (or a line) that is contained in the complement of $A$, and let $k : \mathbb{N} \rightarrow \mathbb{R}$ be the sequence defined by

$$k(i) = d((i, f(i)), R)$$

If $k$ is an exponential decaying sequence, then the set $A$ is not semialgebraic.
Proof. Let us suppose that, for all \( x \geq 1 \), the inequality \( f(x) \leq 1 \) holds. Given \( i \geq 1 \), we use the symbol \( a_i \) to denote the point in \( R \) that is closest to \((i, f(i))\).

Let \( g : \mathbb{R}^+ \to \mathbb{R} \) be the function defined by

\[
g(t) = \frac{1}{d(a_1 + t \cdot (a_2 - a_1), A)}
\]

We have that \( g \) is a well defined function of exponential growth.

Suppose that \( A \) is a semialgebraic, then, it follows easily from The Tarski-Seidenberg Theorem (see [12]) that function \( g \) has a first order definition over the real-closed field. Then, if \( A \) is semialgebraic, the function \( f \) is definable in first order logic, and then semialgebraic. We arrived to a contradiction given that exponential growing functions cannot be semialgebraic. Then, the lemma is proved.

We say that a ray \( R \) is a forbidden ray, if and only if, \( R \cap (\Gamma^* \cdot \Gamma) = \emptyset \). We need to find a sequence of information inequalities, which is contained in some two-dimensional plane, and which approaches a forbidden ray at exponential speed.

Set:

- \( a = v_3 \).
- \( J = v_4 + v_5 + v_6 + v_7 \).
- \( d = v_5 + v_7 \).

The sequence

\[
\{ I - d + \frac{1}{2^s - 1} a + \frac{s2^{s-1}}{2^s - 1} J \}_{s \geq 1}
\]

is an infinite sequence of information inequalities discovered by Dougherty et al [3]. We use the symbol \( dfz_s \) to denote the inequality

\[
I - d + \frac{1}{2^s - 1} a + \frac{s2^{s-1}}{2^s - 1} J
\]

and we use the term DFZ-sequence to denote the sequence \( \{ dfz_s \}_{s \geq 1} \).

Given \( Q \), a two-dimensional plane contained in \( \mathbb{R}^5 \), and given \( v \in \mathbb{R}^5 \), we use the symbol \( Tr_r(Q) \) to denote the translation of \( Q \) by the vector \( v \). Let \( Q \) be the plane spanned by the vectors \( a \) and \( J \), and let \( L \) be equal to \( Tr_{I-d}(Q) \). Notice that the DFZ-sequence is included in \( (\Gamma^* \cdot \Gamma) \cap L \). If we use \( I - d \) as the origin of \( L \), and the lines \( \{ I - d + sa : s \geq 0 \} \) and \( \{ I - d + sJ : s \geq 0 \} \) as its coordinate axis, we get a parametrization of \( \{ dfz_s \}_{s \geq 1} \), over the plane \( L \), that looks like \( \left( \frac{1}{2^s - 1}, \frac{s2^{s-1}}{2^s - 1} \right)_{s \geq 1} \) and that asymptotically behaves like \( \left( \frac{1}{s}, \frac{s}{2} \right) \). Thus, the DFZ-sequence is a plane sequence that approaches the ray

\[
R = \{ I - d + \lambda J : \lambda \geq 1 \}
\]

at exponential speed. It remains to be proved that \( R \) is contained in the complement of \( (\Gamma^* \cdot \Gamma) \cap L \). Thus, we focus our attention on the following problem.
**Problem 4** Prove that for all $\lambda > 0$, the vector $\mathcal{I} + \lambda \mathcal{J}$ is not an information inequality.

Given an entropic function $h$, if the inequality $\langle \mathcal{I}, h \rangle < 0$ holds, we say that $h$ is an Ingleton violating entropic function. Suppose that for all $\lambda > 0$, the vector $\mathcal{I} + \lambda \mathcal{J}$ is not an information inequality, we have:

- Either, there exists an Ingleton-violating entropic function $h$, such that the equality $\langle \mathcal{J}, h \rangle = 0$ holds, or.
- There exists a sequence of entropic polymatroids $\{h_n\}_{n\geq 1}$, such that for all $n \geq 1$ it happens that:
  1. $\langle h_0, \mathcal{I} \rangle = -1$.
  2. $\langle h_n, \mathcal{J} \rangle \leq \frac{1}{n}$.

It would be better for our purposes, if it were possible to find an Ingleton violating entropic function that is orthogonal to $\mathcal{J}$.

**Lemma 5** There does not exist an Ingleton violating entropic function $h$, such that $h$ is orthogonal to $\mathcal{J}$.

**Proof.** Suppose that $h$ is an Ingleton violating polymatroid, and suppose that $h$ is orthogonal to $\mathcal{J}$. Then, $h$ must be orthogonal to $I(1;3 \mid 2)$, $I(2;3 \mid 1)$, $I(1;4 \mid 2)$ and $I(2;4 \mid 1)$. It implies that $h$ is orthogonal to $d$, given that $d = I(1;4 \mid 2) + I(2;4 \mid 1)$. Then, we have that

$$
\langle dfz_s, h_0 \rangle = \left\langle \mathcal{I} - d + \frac{1}{2^s - 1}a + \frac{s2^{s-1}}{2^s - 1}\mathcal{J}, h_0 \right\rangle
= \langle \mathcal{I}, h_0 \rangle + \frac{1}{2^s - 1} (a, h_0)
$$

and then, it is clear that there must exist $s$ such that $\langle dfz_s, h_0 \rangle < 0$. ■

The above lemma asserts that the implication

$$
\langle \mathcal{J}, h \rangle = 0, \text{ implies that } \langle \mathcal{I}, h \rangle \geq 0
$$

holds true for any entropic vector $h$. It means that the pair $(\mathcal{J}, \mathcal{I})$ is a conditional information inequality in the sense of Kaced and Romashenko [8]. Recall that

**Definition 6** Given $v, w \in \mathbb{R}^{15}$, the pair $(v, w)$ is a conditional information inequality, if and only if, the implication

$$
\langle v, h \rangle = 0, \text{ implies that } \langle w, h \rangle \geq 0
$$

holds true for any entropic vector $h$. 


We have to prove that \((J, I)\) is essentially conditional, which means that for all \(N \geq 0\), the expression \(I + N J\) is not an information inequality.

**Lemma 7** The pair \((J, I)\) is an essentially conditional information inequality.

**Proof.** Given \(n \geq 0\), we have to exhibit a 4-tuple of random variables \(\vec{X} = (X_1, X_2, X_3, X_4)\) such that \(\langle h_{\vec{X}}, I + n J \rangle < 0\). We begin with a 4-tuple \(\vec{X}\), that is given by the following probability distribution:

1. The support of the distribution is constituted by the set 
   \(\{(a, b, c, d) : a, b \in \{0, 1\}, \ c = 0 \& d \in \{0, 1, 2, 3, 4\}\}\)

2. Given \(a, b \in \{0, 1\}\), and given \(d \neq 4\), we have that 
   \(\Pr (X_1 = a, X_2 = b, X_3 = 0, X_4 = d) = \frac{1}{32}\)

3. Given \(a, b \in \{0, 1\}\), we have that 
   \(\Pr (X_1 = a, X_2 = b, X_3 = 0, X_4 = 4) = \frac{1}{8}\)

Notice that those four random variables are pairwise independent, it implies that \(\langle h_{\vec{X}}, I \rangle = \langle h_{\vec{X}}, J \rangle = 0\). Given \(\delta \in [0, \frac{1}{32}]\), we define a perturbation of \(\vec{X}\), which is denoted with the symbol \(\vec{X}_\delta\), and which is defined by adding to the support of \(\vec{X}\) the four 4-tuples 

\[(1,0,1,0), (1,1,1,1), (0,0,1,2), (0,1,1,3)\]

each with probability \(\delta\). The sum of the probabilities must be equal to 1, therefore we set

\(\Pr (X_1 = 0, X_2 = 0, X_3 = 0, X_4 = 4) = \frac{1}{8} - 4\delta.\)

Some easy computations show that, when we make \(\delta\) tends to zero, we get

- \(\langle h_{\vec{X}_\delta}, I \rangle = -4\delta + O(\delta^2)\).
- \(\langle h_{\vec{X}_\delta}, J \rangle = O(\delta^2)\).

Then, it is clear that for all \(n \geq 0\), it is possible to find a \(\delta > 0\) such that \(\langle h_{\vec{X}_\delta}, I + n J \rangle < 0\), and then the lemma is proved.

We get as a corollary the main result of this paper

**Theorem 8** Given \(n \geq 4\), the set \(\Gamma_n^\ast\) is not semialgebraic.
Now, we can get as an easy corollary The theorem of Matus.

**Corollary 9** Given $n \geq 4$, the set $\Gamma_n^\ast$ is not polyhedral.

Those two results have unfortunate consequences, the later implies that many important optimization problems cannot be reduced to linear programming, while the former implies that those problems cannot be reduced to semi-algebraic optimization.

### 4 Concluding remarks

We knew that the almost-entropic regions cannot be defined by finite lists of linear inequalities. Now, we know that those regions cannot be defined by finite lists of polynomial inequalities. This later fact implies that those regions cannot be defined in first order logic over the real-closed field. Which is the next level of complexity to be tried? We think that the next question to be investigated is the decidability of the almost entropic regions, we conjecture that the set of almost-entropic vectors with integer components is not decidable.

**Acknowledgement.** The first author would like to thank the support provided by Colciencias. The second author would like to thank the support provided by Universidad Nacional de Colombia through the project Hermes 8943.

### References

[1] T. Chan, A. Grant. Non-linear Information Inequalities. Entropy 10(4): 765-775, 2008.

[2] L. Csirmaz. The Size of a Share Must Be Large. J. Cryptology 10(4): 223-231, 1997.

[3] R. Dougherty, C. Freiling, K. Zeger. Non-shannon Information Inequalities in Four Random Variables. [arXiv:1104.3602](https://arxiv.org/abs/1104.3602)

[4] A. Gomez, C. Mejia, J. A. Montoya. Linear Network Coding and the Model Theory of Linear Rank Inequalities. Proceedings of NETCOD 2014, Aalborg, Denmark.

[5] A. Gomez, C. Mejia, J. A. Montoya. Network Coding and the Model Theory of Linear information Inequalities. Proceedings of NETCOD 2014, Aalborg, Denmark.

[6] G. Gottlob, S. Lee, G. Valiant, P. Valiant. Size and Treewidth Bounds for Conjunctive Queries. J. ACM 59(3): 16-45, 2012.

[7] A. Ingleton. Representation of matroids. In Welsh, D.J.A. *Combinatorial mathematics and its applications*. Proceedings, Oxford, 1969. Academic Press. pp. 149–167, 1971.
[8] T. Kaced, A. Romashenko. Conditional Information Inequalities for Entropic and Almost Entropic Points. IEEE Transactions on Information Theory 59(11): 7149-7167, 2013.

[9] J. Lasserre. Moments and sums of squares for polynomial optimization and related problems. J. Global Optimization 45(1): 39-61, 2009.

[10] F. Matus. Infinitely many information inequalities. Proceedings of ISIT 2007, Nice, France.

[11] F. Matus and M. Studeny. Conditional independencies among four random variables I. Combinatorics, Probability and Computing, 4: 269-278, 1995.

[12] A. Tarski. A Decision Method for Elementary Algebra and Geometry. Univ. of California Press, 1951.

[13] L. Van der Dries. Tame Topology and o-minimal Structures. London Mathematical Society Lecture Note Series 248, 1997.

[14] R. Yeung. A First Course in Information Theory. Springer, Berlin, 2002.