Inequalities for entropies and dimensions

alexander.shen@lirmm.fr,
www.lirmm.fr/~ashen

LIRMM CNRS & University of Montpellier

Computability in Europe 2023

Three approaches to measure information (Kolmogorov, 1965)

- combinatorial: n values of attribute A in a database table mean that this attribute carries $\log_2 n$ bits of information
- probabilistic (Shannon): random variable ξ with n values $(p_1 + \ldots + p_n = 1)$ carries

  $$H(p) = p_1 \log \frac{1}{p_1} + \ldots + p_n \log \frac{1}{p_n}$$

  bits of information
- algorithmic: the amount of information (*complexity*) $\mathrm{C}(x)$ in a bit string x is the minimal bit length of a program that produces x

Three approaches to measure information (Kolmogorov, 1965)

- combinatorial: n values of attribute A in a database table mean that this attribute carries $\log_2 n$ bits of information
- probabilistic (Shannon): random variable ξ with n values ($p_1 + \ldots + p_n = 1$) carries

$$H(p) = p_1 \log \frac{1}{p_1} + \ldots + p_n \log \frac{1}{p_n}$$

bits of information
- algorithmic: the amount of information (*complexity*) $\mathrm{C}(x)$ in a bit string x is the minimal bit length of a program that produces x

Three approaches to measure information (Kolmogorov, 1965)

- combinatorial: n values of attribute A in a database table mean that this attribute carries $\log_2 n$ bits of information
- probabilistic (Shannon): random variable ξ with n values $(p_1 + \ldots + p_n = 1)$ carries

  $$H(p) = p_1 \log \frac{1}{p_1} + \ldots + p_n \log \frac{1}{p_n}$$

  bits of information
- algorithmic: the amount of information (*complexity*) $\mathrm{C}(x)$ in a bit string x is the minimal bit length of a program that produces x

Three approaches to measure information (Kolmogorov, 1965)

- combinatorial: n values of attribute A in a database table mean that this attribute carries $\log_2 n$ bits of information
- probabilistic (Shannon): random variable ξ with n values ($p_1 + \ldots + p_n = 1$) carries
$$H(p) = p_1 \log \frac{1}{p_1} + \ldots + p_n \log \frac{1}{p_n}$$
bits of information
- algorithmic: the amount of information (*complexity*) $\mathrm{C}(x)$ in a bit string x is the minimal bit length of a program that produces x

Direct connections

- (CP) n values: $H(\xi) \leqslant \log_2 n$ (achieved when all outcomes are equiprobable)
- (PA) m independent trials of a random variable ξ: with high probability the complexity of the outcome is close to $mH(\xi)$
- (CA) there are at most 2^n strings of complexity less than n; every element of a simple set with N elements has complexity not exceeding $\log_2 N$ (almost)

Direct connections

- (CP) n values: $H(\xi) \leqslant \log_2 n$ (achieved when all outcomes are equiprobable)
- (PA) m independent trials of a random variable ξ: with high probability the complexity of the outcome is close to $mH(\xi)$
- (CA) there are at most 2^n strings of complexity less than n; every element of a simple set with N elements has complexity not exceeding $\log_2 N$ (almost)

Direct connections

- (CP) n values: $H(\xi) \leqslant \log_2 n$ (achieved when all outcomes are equiprobable)
- (PA) m independent trials of a random variable ξ: with high probability the complexity of the outcome is close to $mH(\xi)$
- (CA) there are at most 2^n strings of complexity less than n; every element of a simple set with N elements has complexity not exceeding $\log_2 N$ (almost)

Direct connections

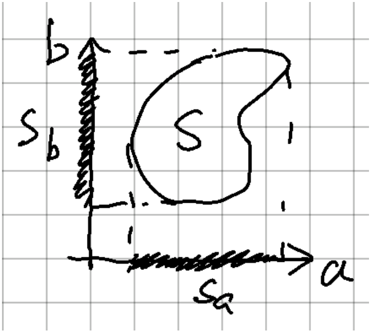
- (CP) n values: $H(\xi) \leqslant \log_2 n$ (achieved when all outcomes are equiprobable)
- (PA) m independent trials of a random variable ξ: with high probability the complexity of the outcome is close to $mH(\xi)$
- (CA) there are at most 2^n strings of complexity less than n; every element of a simple set with N elements has complexity not exceeding $\log_2 N$ (almost)

Direct connections

- (CP) n values: $H(\xi) \leqslant \log_2 n$ (achieved when all outcomes are equiprobable)
- (PA) m independent trials of a random variable ξ: with high probability the complexity of the outcome is close to $mH(\xi)$
- (CA) there are at most 2^n strings of complexity less than n; every element of a simple set with N elements has complexity not exceeding $\log_2 N$ (almost)

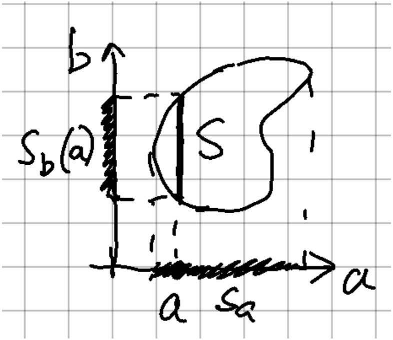
Indirect connections: similar properties (1)

- $H(\xi, \eta) \leqslant H(\xi) + H(\eta)$
- $\mathrm{C}(x, y) \leqslant \mathrm{C}(x) + \mathrm{C}(y) + O(\log n)$ for bit strings x, y of length at most n
- $\log_2 S \leqslant \log_2 S_a + \log_2 S_b$



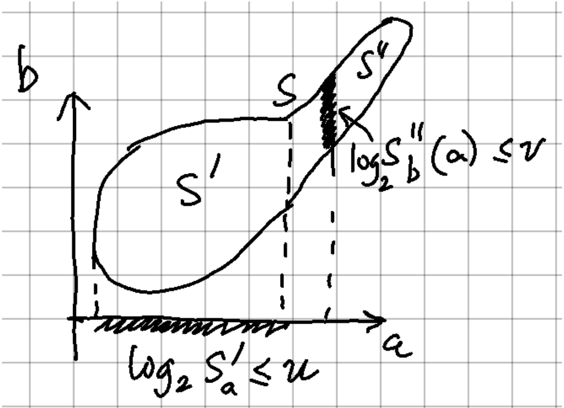
Indirect connections: similar properties (2)

- $H(\xi,\eta) \leqslant H(\xi) + H(\eta|\xi)$
- $\mathrm{C}(x,y) \leqslant \mathrm{C}(x) + \mathrm{C}(y|x) + O(\log n)$ for strings x, y of length at most n
- $\log_2 S \leqslant \log_2 S_a + \log_2 \max_a S_b(a)$



Indirect connections: similar properties (3)

- $H(\xi) + H(\eta|\xi) \leqslant H(\xi,\eta)$
- $\mathrm{C}(x) + \mathrm{C}(y|x) \leqslant \mathrm{C}(x,y) + O(\log n)$ for strings x, y of length at most n
- $\log S \leqslant u + v \Rightarrow S$ can be split into $S = S' \cup S''$ such that $\log_2 S'_a \leqslant u$ and $\max_a \log_2 S''_b(a) \leqslant v$



General results

- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- ...and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

General results

- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- ...and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

General results

- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- ...and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

General results

- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- …and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

General results

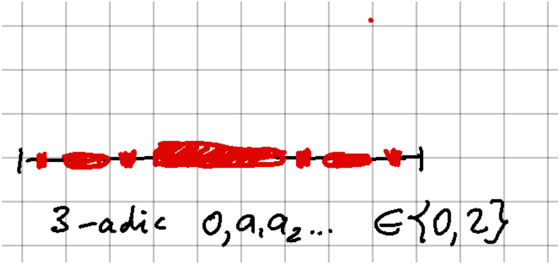
- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- ...and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

General results

- The same linear inequalities are true for Shannon entropy and Kolmogorov complexity (Romashchenko)
- This class has combinatorial interpretation (Vereshchagin et al., 2000s)
- …and in terms of subgroup sizes (Chan, Yeang, 2000s)
- The same class for space-bounded complexities (Gács et al., 2020)
- this paper: *one more result of this type using dimensions*

Algorithmic dimension theory

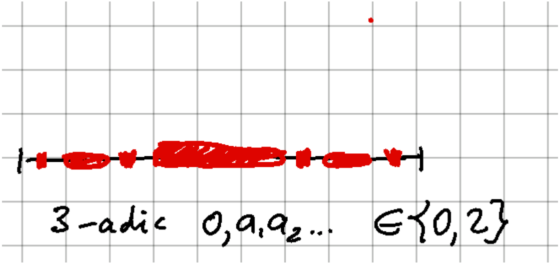
- Hausdorff and packing dimensions



- dimension = $\log 2/\log 3$, "use only 2 digits out of 3": to specify any point with a given precision we need less information (factor $\log 2/\log 3$) than in general case

Algorithmic dimension theory

- Hausdorff and packing dimensions



- dimension = $\log 2/\log 3$, “use only 2 digits out of 3”: to specify any point with a given precision we need less information (factor $\log 2/\log 3$) than in general case

Effective dimensions

- for $x \in (0,1)$ let $\mathrm{C}_n(x)$ be the complexity of 2^{-n}-approximation to x.
- *dim*$(x) = \liminf_n \mathrm{C}_n(x)/n$ [eff. Hausdorff]
- *Dim*$(x) = \limsup_n \mathrm{C}_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  dim$(A) = \max_{x \in A}$ *dim*(x)
  Dim$(A) = \max_{x \in A}$ *Dim*(x)
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $C_n(x)$ be the complexity of 2^{-n}-approximation to x.
- $dim(x) = \liminf_n C_n(x)/n$ [eff. Hausdorff]
- $Dim(x) = \limsup_n C_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  $dim(A) = \max_{x \in A} dim(x)$
  $Dim(A) = \max_{x \in A} Dim(x)$
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $C_n(x)$ be the complexity of 2^{-n}-approximation to x.
- *dim*$(x) = \liminf_n C_n(x)/n$ [eff. Hausdorff]
- *Dim*$(x) = \limsup_n C_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  dim$(A) = \max_{x \in A}$ *dim*(x)
  Dim$(A) = \max_{x \in A}$ *Dim*(x)
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $\mathrm{C}_n(x)$ be the complexity of 2^{-n}-approximation to x.
- *dim*$(x) = \liminf_n \mathrm{C}_n(x)/n$ [eff. Hausdorff]
- *Dim*$(x) = \limsup_n \mathrm{C}_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  dim$(A) = \max_{x\in A}$ *dim*(x)
  Dim$(A) = \max_{x\in A}$ *Dim*(x)
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $\mathrm{C}_n(x)$ be the complexity of 2^{-n}-approximation to x.
- $\mathit{dim}(x) = \liminf_n \mathrm{C}_n(x)/n$ [eff. Hausdorff]
- $\mathit{Dim}(x) = \limsup_n \mathrm{C}_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  $\mathit{dim}(A) = \max_{x \in A} \mathit{dim}(x)$
  $\mathit{Dim}(A) = \max_{x \in A} \mathit{Dim}(x)$
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $C_n(x)$ be the complexity of 2^{-n}-approximation to x.
- $\mathit{dim}(x) = \liminf_n C_n(x)/n$ [eff. Hausdorff]
- $\mathit{Dim}(x) = \limsup_n C_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  $\mathit{dim}(A) = \max_{x \in A} \mathit{dim}(x)$
  $\mathit{Dim}(A) = \max_{x \in A} \mathit{Dim}(x)$
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Effective dimensions

- for $x \in (0,1)$ let $\mathrm{C}_n(x)$ be the complexity of 2^{-n}-approximation to x.
- *dim*$(x) = \liminf_n \mathrm{C}_n(x)/n$ [eff. Hausdorff]
- *Dim*$(x) = \limsup_n \mathrm{C}_n(x)/n$ [eff. packing]
- For $A \subset (0,1)$ let
  dim$(A) = \max_{x\in A}$ *dim*(x)
  Dim$(A) = \max_{x\in A}$ *Dim*(x)
- locality paradox: how many / which elements are in A
- a point can have high effective dimension; a set of high (classical) dimension always contains a point of high effective dimension

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $\mathit{dim}^S(x), \mathit{Dim}^S(x), \mathit{dim}^S(A), \mathit{Dim}^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S \mathit{dim}^S(A)$,
  $\mathrm{Dim}(A) = \min_S \mathit{Dim}^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $\mathit{dim}^S(x), \mathit{Dim}^S(x), \mathit{dim}^S(A), \mathit{Dim}^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S \mathit{dim}^S(A)$,
  $\mathrm{Dim}(A) = \min_S \mathit{Dim}^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $dim^S(x), Dim^S(x), dim^S(A), Dim^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S dim^S(A)$,
  $\mathrm{Dim}(A) = \min_S Dim^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $\mathit{dim}^S(x), \mathit{Dim}^S(x), \mathit{dim}^S(A), \mathit{Dim}^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S \mathit{dim}^S(A)$,
  $\mathrm{Dim}(A) = \min_S \mathit{Dim}^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $\mathit{dim}^S(x), \mathit{Dim}^S(x), \mathit{dim}^S(A), \mathit{Dim}^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S \mathit{dim}^S(A)$,
  $\mathrm{Dim}(A) = \min_S \mathit{Dim}^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

Point-to-set principle (J.Lutz, N.Lutz)

- relativizing: $\mathit{dim}^S(x), \mathit{Dim}^S(x), \mathit{dim}^S(A), \mathit{Dim}^S(A)$
- point-to-set principle:
  $\dim(A) = \min_S \mathit{dim}^S(A)$,
  $\mathrm{Dim}(A) = \min_S \mathit{Dim}^S(A)$
- classical dimension =
  $= \min_{\text{oracle}} \max_{\text{point}}$ effective dimension
- opens a way to translate information inequalities into statements about dimensions
- for $A \subset [0,1] \times [0,1]$; projections A_1 and A_2

$$\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_1) + \mathrm{Dim}(A_2)$$

More examples

- $2\,\mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + \mathrm{C}(x_1, x_3) + \mathrm{C}(x_2, x_3)$
- $A \subset [0,1]^3$, 2-dimensional projections A_{12}, A_{23}, A_{13}:

$$2\,\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) + \mathrm{Dim}(A_{23})$$

- $C(x_1) + \mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + C(x_1, x_3)$
- for $A \subset [0,1]^3$: if (for some u and v)

$$\mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) \leqslant u + v$$

then A can be splitted into $A = A^{'} \cup A^{''}$ such that

$$\dim(A^{'}_1) \leqslant u, \quad \dim(A^{''}) \leqslant v$$

More examples

- $2\,\mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + \mathrm{C}(x_1, x_3) + \mathrm{C}(x_2, x_3)$
- $A \subset [0,1]^3$, 2-dimensional projections A_{12}, A_{23}, A_{13}:

$$2\,\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) + \mathrm{Dim}(A_{23})$$

- $C(x_1) + \mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + C(x_1, x_3)$
- for $A \subset [0,1]^3$: if (for some u and v)

$$\mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) \leqslant u + v$$

then A can be splitted into $A = A' \cup A''$ such that

$$\dim(A'_1) \leqslant u, \quad \dim(A'') \leqslant v$$

More examples

- $2\,\mathrm{C}(x_1,x_2,x_3) \leqslant \mathrm{C}(x_1,x_2) + \mathrm{C}(x_1,x_3) + \mathrm{C}(x_2,x_3)$
- $A \subset [0,1]^3$, 2-dimensional projections A_{12}, A_{23}, A_{13}:

$$2\,\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) + \mathrm{Dim}(A_{23})$$

- $C(x_1) + \mathrm{C}(x_1,x_2,x_3) \leqslant \mathrm{C}(x_1,x_2) + C(x_1,x_3)$
- for $A \subset [0,1]^3$: if (for some u and v)

$$\mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) \leqslant u + v$$

then A can be splitted into $A = A' \cup A''$ such that

$$\dim(A_1') \leqslant u, \quad \dim(A'') \leqslant v$$

More examples

- $2\,\mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + \mathrm{C}(x_1, x_3) + \mathrm{C}(x_2, x_3)$
- $A \subset [0,1]^3$, 2-dimensional projections A_{12}, A_{23}, A_{13}:

$$2\,\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) + \mathrm{Dim}(A_{23})$$

- $C(x_1) + \mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + C(x_1, x_3)$
- for $A \subset [0,1]^3$: if (for some u and v)

$$\mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) \leqslant u + v$$

then A can be splitted into $A = A' \cup A''$ such that

$$\dim(A_1') \leqslant u, \quad \dim(A'') \leqslant v$$

More examples

- $2\,\mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + \mathrm{C}(x_1, x_3) + \mathrm{C}(x_2, x_3)$
- $A \subset [0,1]^3$, 2-dimensional projections A_{12}, A_{23}, A_{13}:

$$2\,\mathrm{Dim}(A) \leqslant \mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) + \mathrm{Dim}(A_{23})$$

- $C(x_1) + \mathrm{C}(x_1, x_2, x_3) \leqslant \mathrm{C}(x_1, x_2) + C(x_1, x_3)$
- for $A \subset [0,1]^3$: if (for some u and v)

$$\mathrm{Dim}(A_{12}) + \mathrm{Dim}(A_{13}) \leqslant u + v$$

then A can be splitted into $A = A' \cup A''$ such that

$$\dim(A'_1) \leqslant u, \quad \dim(A'') \leqslant v$$

New description of information inequalities

- Every true linear inequality for entropies/complexites can be translated into a (true) statement about classical Hausdorff and packing dimensions as described
- If a linear inequality is *not* true for entropies/complexities, the corresponding statement about dimensions is also false

New description of information inequalities

- Every true linear inequality for entropies/complexites can be translated into a (true) statement about classical Hausdorff and packing dimensions as described
- If a linear inequality is *not* true for entropies/complexities, the corresponding statement about dimensions is also false

New description of information inequalities

- Every true linear inequality for entropies/complexites can be translated into a (true) statement about classical Hausdorff and packing dimensions as described
- If a linear inequality is *not* true for entropies/complexities, the corresponding statement about dimensions is also false

THANKS!