Randomness Characterization through Bayesian Model Selection

Rafael Díaz Hernández Rojas

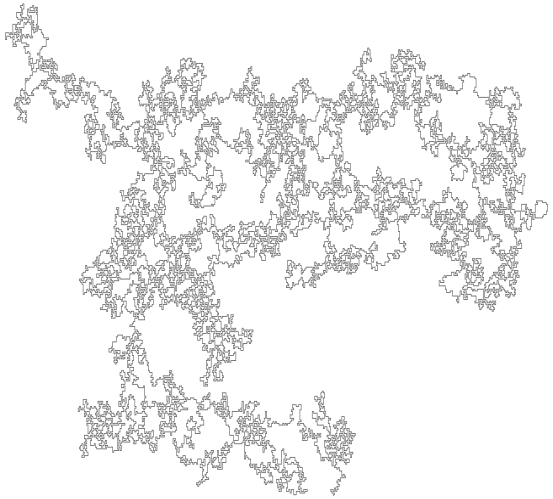
Sapienza University of Rome

Isaac Pérez Castillo (IF-UNAM)
Jorge Hirsch, Alfred U'Ren, Aldo Solís, Alí Angulo (ICN-UNAM)
Matteo Marsili (ICTP, Italy)

AISIS, UNAM. October 2019

How to tell if a number sequence is random?

$\hat{s} = HHTTTHTHT \ldots THTT$



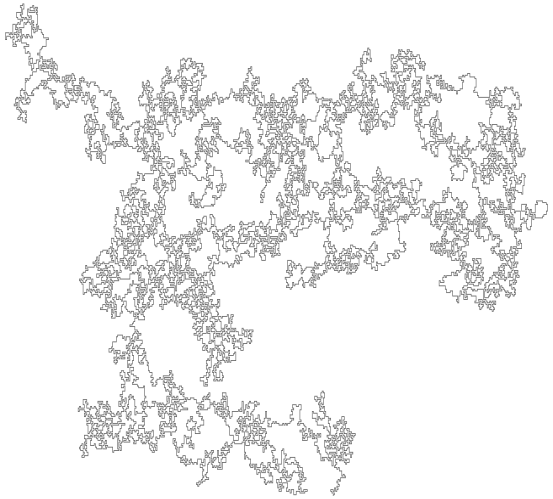
$\hat{s} = LLRRRLRLR \ldots RLRR$

- Dynamical systems mappings
- Spin systems
- Correlated photons
- Particles decay

$\hat{s} = 110001010 \ldots 0100$

How to tell if a number sequence is random?

$\hat{s} = HHTTTHTHT \ldots THTT$



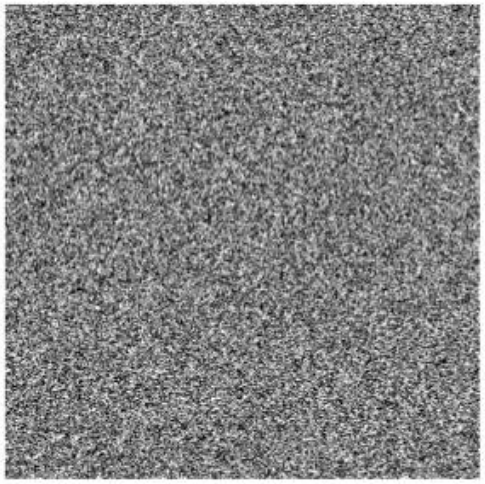
$\hat{s} = LLRRRLRLR \ldots RLRR$

- Dynamical systems mappings
- Spin systems
- Correlated photons
- Particles decay

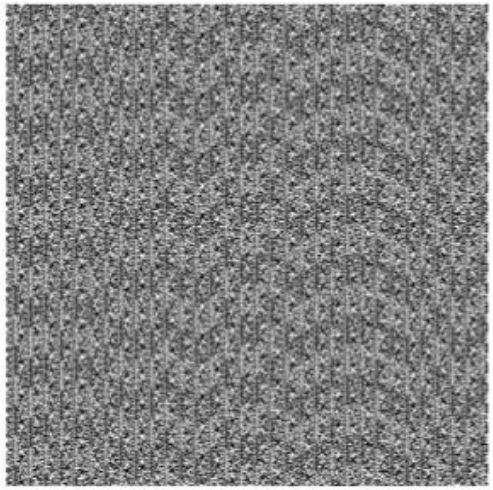
$\hat{s} = 110001010 \ldots 0100$

- Monte Carlo methods
- Cryptography
- Probabilistc algorithms

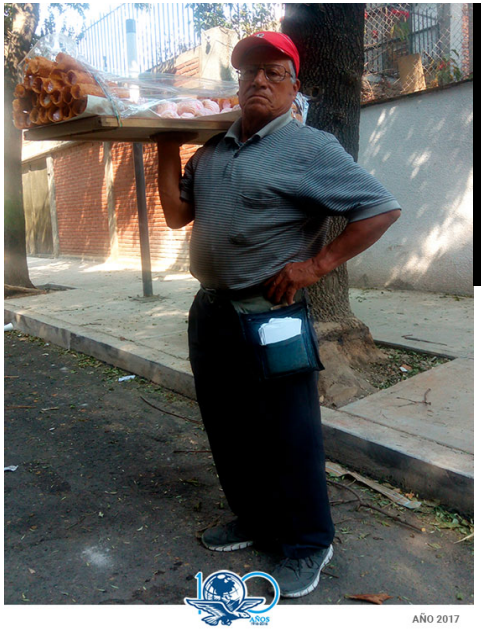
Pseudo-random bits



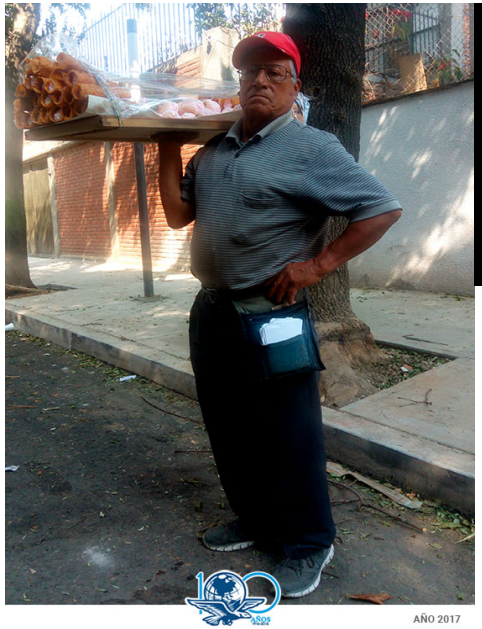
rand() in PHP on Windows



(Maximally) random sequences

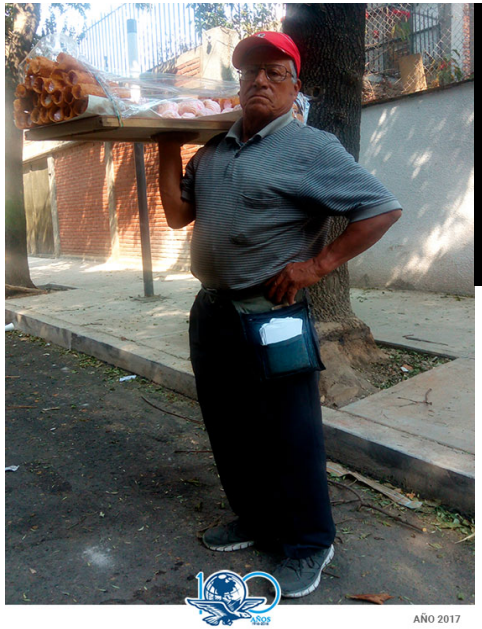


(Maximally) random sequences



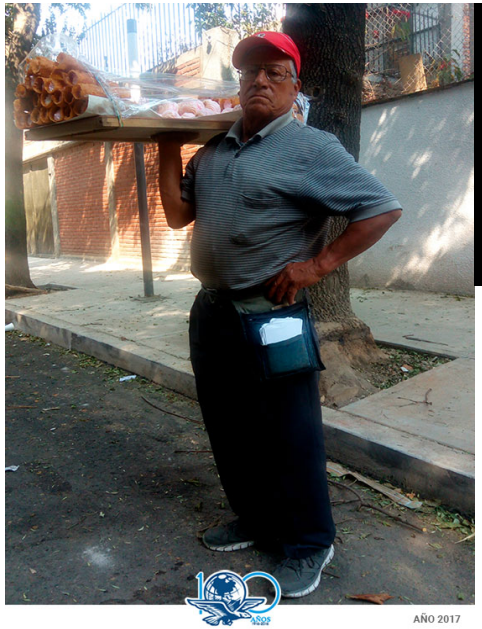
$$\hat{s} = \texttt{HHTTTHTHT...THTT}$$

(Maximally) random sequences



$$\hat{s} = \texttt{01001101...0010}$$

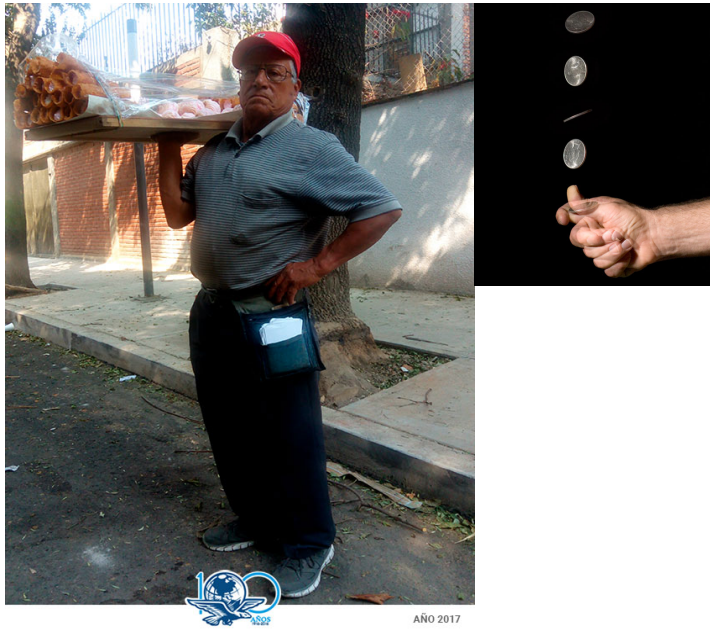
(Maximally) random sequences



$$\hat{s} = \texttt{01001101...0010}$$

What if the coin is biased?

(Maximally) random sequences



$\hat{s} = \texttt{01001101...0010}$

What if the coin is biased?

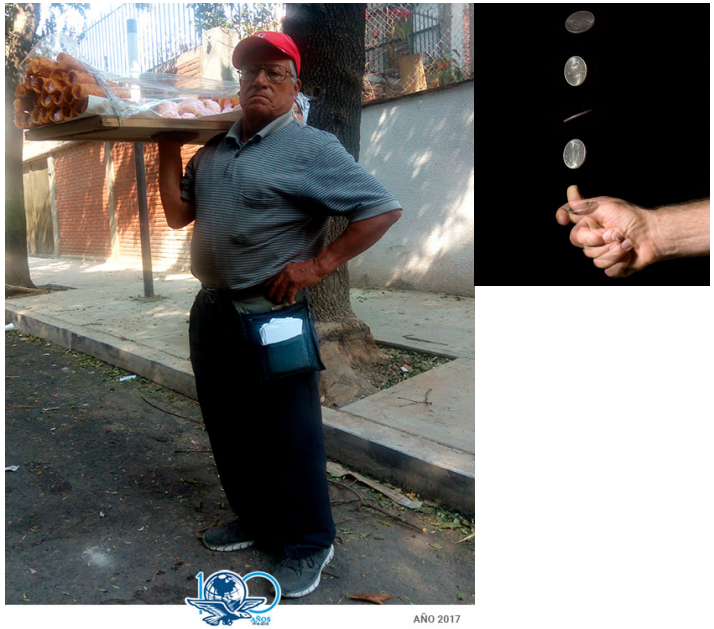
$H[X] \sim$ Measure of randomness

$$H[X] = -\sum_x p_x \log_2 p_x$$

$$0 \leq H[X] \leq \log_2 |\mathbf{X}|$$

$$H_{\max} = \log_2 |\mathbf{X}| \iff p_x = \frac{1}{|\mathbf{X}|}$$

(Maximally) random sequences



$$\hat{s} = \texttt{01001101...0010}$$

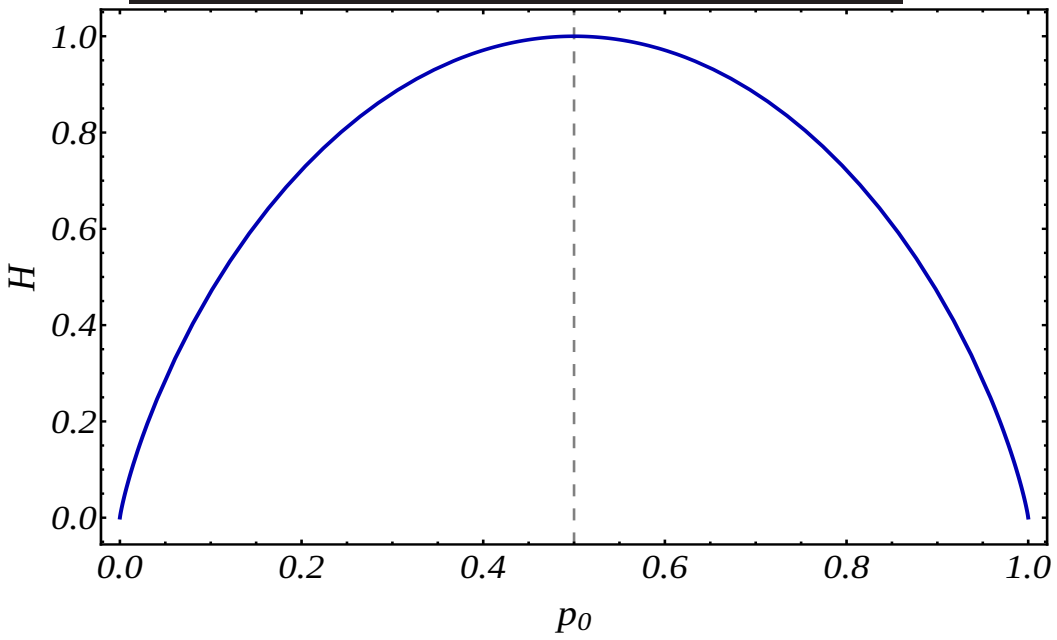
What if the coin is biased?

$H[X] \sim$ Measure of randomness

$$H[X] = -\sum_x p_x \log_2 p_x$$

$$0 \leq H[X] \leq \log_2 |\mathbf{X}|$$

$$H_{\max} = \log_2 |\mathbf{X}| \iff p_x = \frac{1}{|\mathbf{X}|}$$



$$H[X] = -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0)$$

Pragmatic approach: NIST battery of tests

$$\hat{s} = \texttt{01001101...0010} \quad \Longrightarrow \begin{cases} \text{Same frequency of `0' and `1' } (k_0 \approx k_1) \\ \text{Longest string of consecutive 0's} \\ \text{Fourier transform } \sim \text{ white noise} \\ \vdots \end{cases}$$

Each property is analysed as an *hypothesis test* $\Longrightarrow$ obtain a p-value

Pragmatic approach: NIST battery of tests

$$\hat{s} = \texttt{01001101...0010} \quad \Longrightarrow \begin{cases} \text{Same frequency of `0' and `1' } (k_0 \approx k_1) \\ \text{Longest string of consecutive 0's} \\ \text{Fourier transform } \sim \text{ white noise} \\ \vdots \end{cases}$$

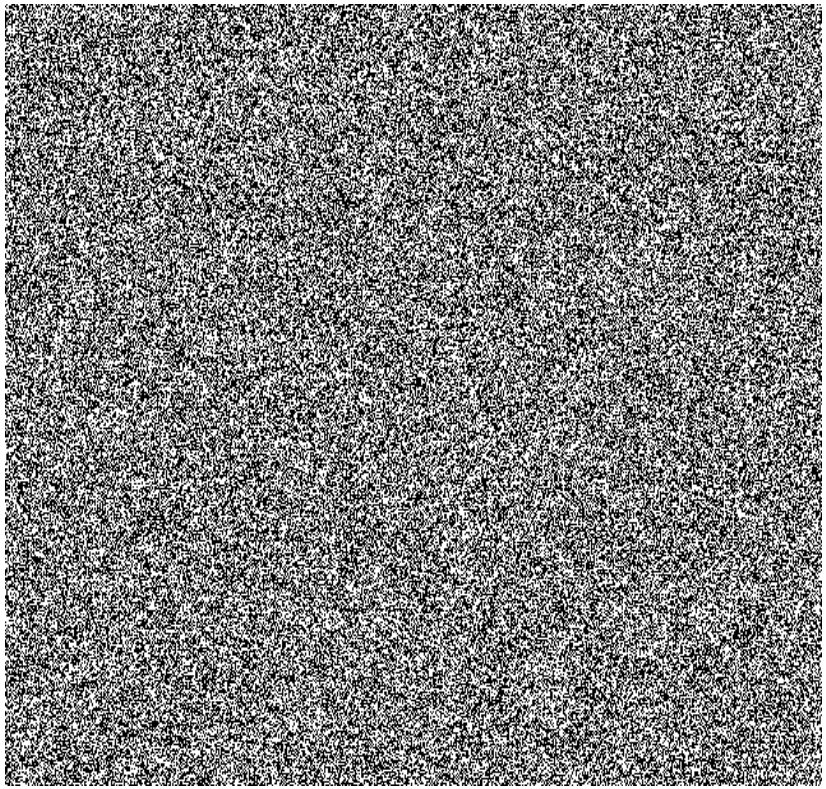
Each property is analysed as an *hypothesis test* $\Longrightarrow$ obtain a p-value

- **If** "random" $\Longrightarrow$ properties examined with the tests.
- Properties $\not\Longrightarrow$ randomness
- *Frequentist* approach based on p-values

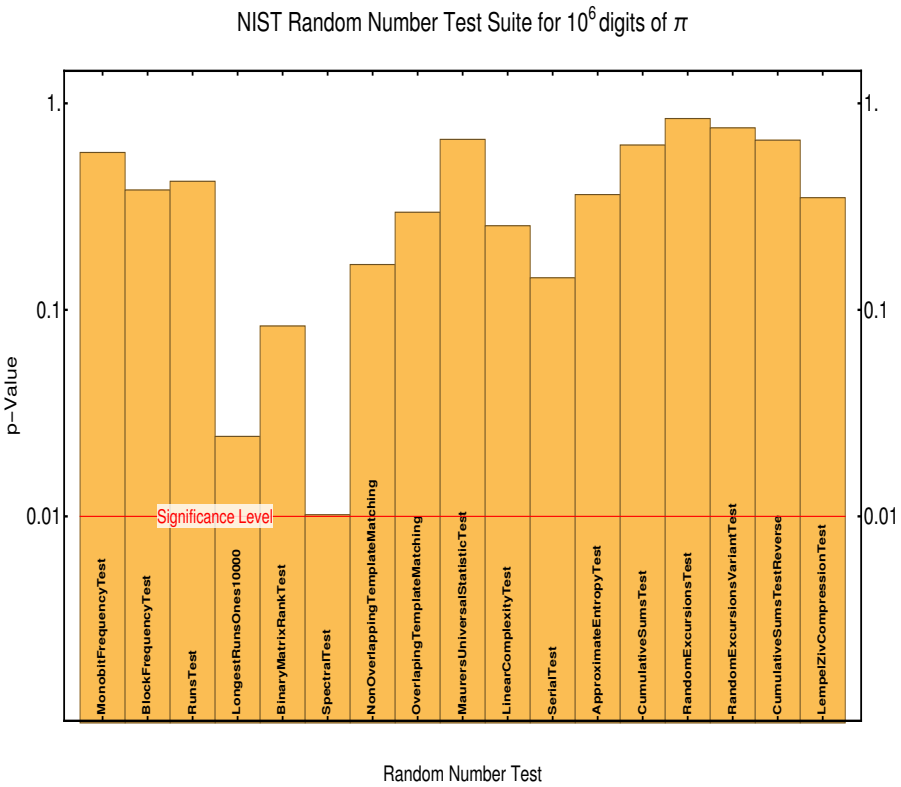
R. L. Wasserstein and N. A. Lazar, "The ASA's statement on p-values: context, process, and purpose", The American Statistician, 129–133 (2016)

M. Baker, "Statisticians issue warning on p-values", Nature **531**, 151 (2016)

Is π random?



(a) Binary representation of the first 302,500 digits of π.



(b) Results of the 15 NIST tests

Randomness as "incompressibility"

Algorithmic Information Theory

- $\hat{s}$ is random **iff** the "shortest" *algorithm* to generate it is `print`($\hat{s}$).
- AIT (Chaitin, Kolmogorov, Somolonoff): a *mathematically formal* theory that identifies (computationally) random $\sim$ incompressible.
- There is NO *general* algorithm capable of assessing whether *any* sequence is random.
- ... *but* Borel's Normality criterion

C. S. Calude, *Information and randomness: an algorithmic perspective*, 2nd Edition (Springer, 2010)

Randomness as "incompressibility"

Algorithmic Information Theory

- $\hat{s}$ is random **iff** the "shortest" *algorithm* to generate it is `print`$(\hat{s})$.
- AIT (Chaitin, Kolmogorov, Somolonoff): a *mathematically formal* theory that identifies (computationally) random $\sim$ incompressible.
- There is NO *general* algorithm capable of assessing whether *any* sequence is random.
- ... *but* Borel's Normality criterion

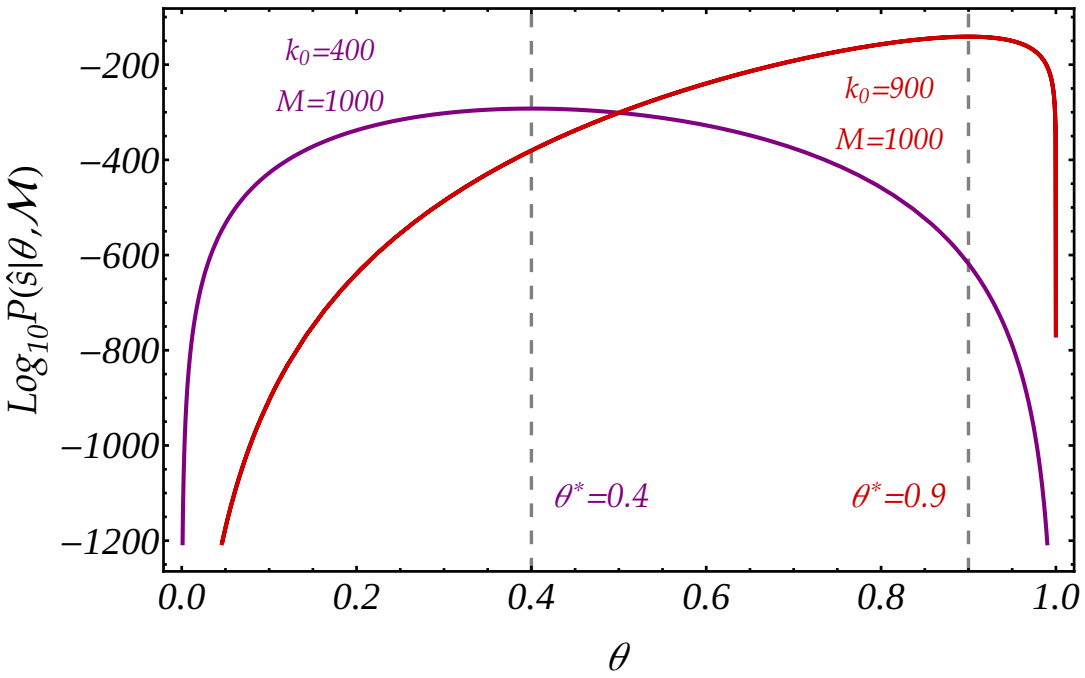
C. S. Calude, *Information and randomness: an algorithmic perspective*, 2nd Edition (Springer, 2010)

$$\pi = 4\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}, \qquad \pi = 2\cdot\frac{2}{\sqrt{2}}\cdot\frac{2}{\sqrt{2+\sqrt{2}}}\cdot\frac{2}{\sqrt{2+\sqrt{2+\sqrt{2}}}}\cdots$$

Primer on statistical inference for bit sequences

$|\hat{s}| = M$ bits, with k_0 and k_1 being the frequencies of `0`'s and '`1`'s; $k_0 + k_1 = M$.

$$\mathcal{M}: \quad p_0 = \theta; \quad p_1 = 1 - \theta \qquad \Longrightarrow \qquad P(\hat{s}|\theta, \mathcal{M}) = \theta^{k_0}(1-\theta)^{k_1} \ .$$



Maximization of $P(\hat{s}|\theta, \mathcal{M})$

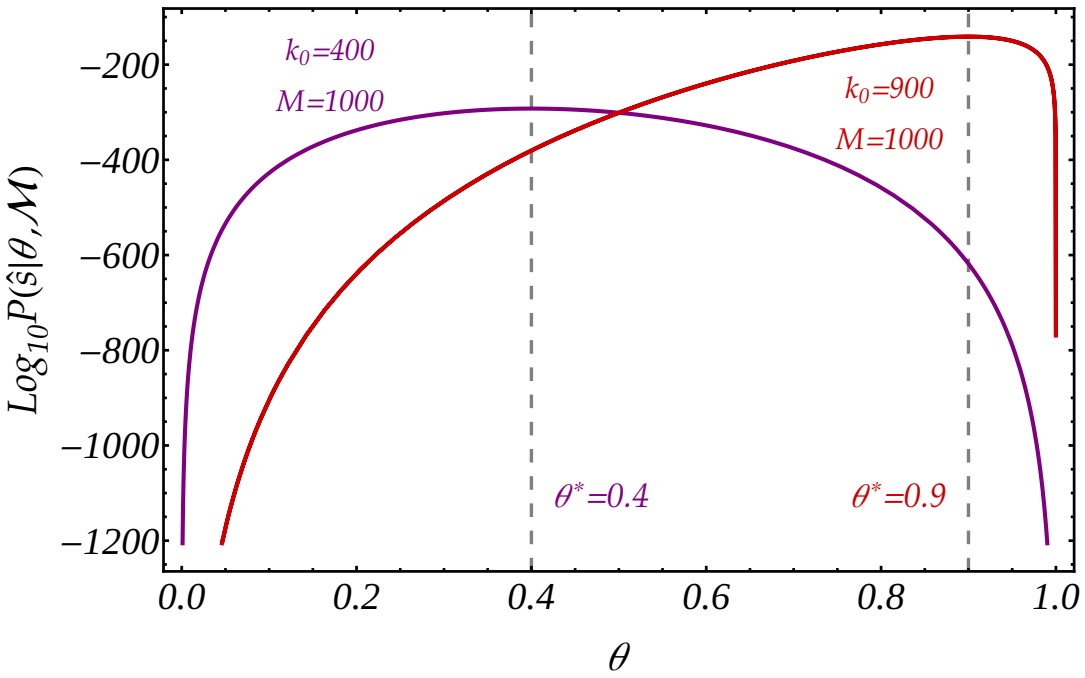
$$\theta^* = k_0/M$$

Fair "coin" (**RNG**)
$\Longrightarrow \theta^* = 0.5$

Primer on statistical inference for bit sequences

$|\hat{s}| = M$ bits, with k_0 and k_1 being the frequencies of `0`'s and '`1`'s; $k_0 + k_1 = M$.

$$\mathcal{M}: \quad p_0 = \theta; \quad p_1 = 1 - \theta \qquad \Longrightarrow \qquad P(\hat{s}|\theta, \mathcal{M}) = \theta^{k_0}(1-\theta)^{k_1} \, .$$



Maximization of $P(\hat{s}|\theta, \mathcal{M})$

$$\theta^* = k_0/M$$

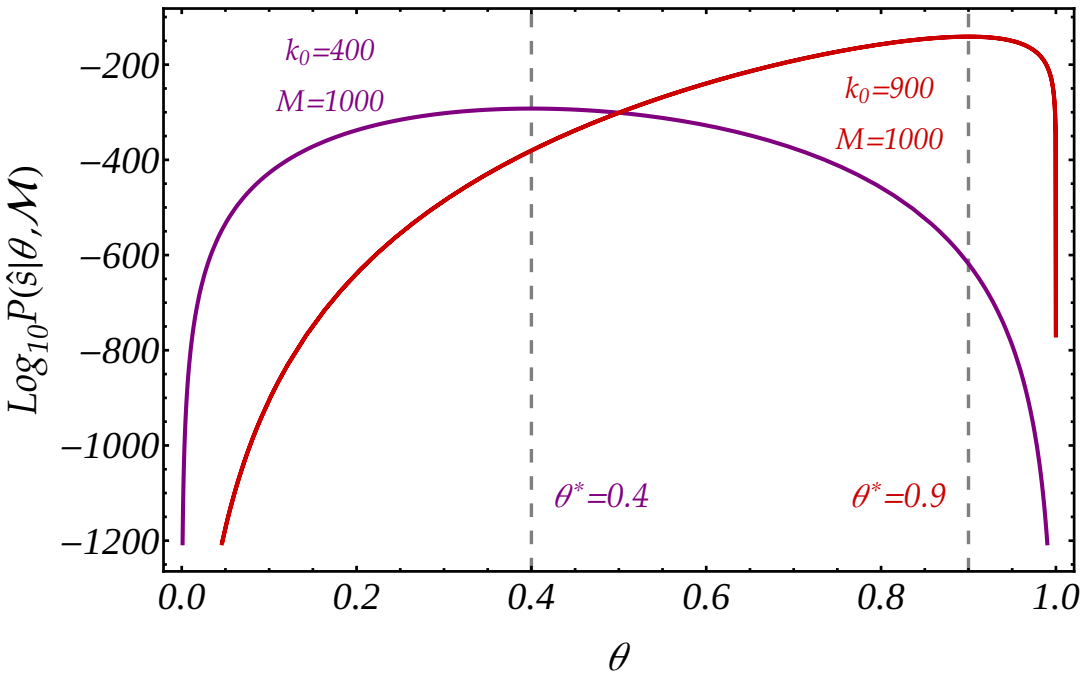
Fair "coin" (**RNG**)
$\Longrightarrow \theta^* = 0.5$

What happens if $\theta^* \approx 0.5$?

Primer on statistical inference for bit sequences

$|\hat{s}| = M$ bits, with k_0 and k_1 being the frequencies of `0`'s and '`1`'s; $k_0 + k_1 = M$.

$$\mathcal{M}: \quad p_0 = \theta; \quad p_1 = 1 - \theta \qquad \Longrightarrow \qquad P(\hat{s}|\theta, \mathcal{M}) = \theta^{k_0}(1-\theta)^{k_1} \, .$$



Maximization of $P(\hat{s}|\theta, \mathcal{M})$

$$\theta^* = k_0/M$$

Fair "coin" (**RNG**)
$\Longrightarrow \theta^* = 0.5$

What happens if $\theta^* \approx 0.5$?

...p-values

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Once a model $\mathcal{M}$ is chosen,
the usual question is:
How well does it describe the
set of observations $\hat{s}$?

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Once a model $\mathcal{M}$ is chosen, the usual question is:
How well does it describe the set of observations $\hat{s}$?
$P(\hat{s}|\mathcal{M}, \boldsymbol{\theta})$ or $P(\hat{s}|\mathcal{M})$

The **right** question is:
Given the observations $\hat{s}$ how likely it is that $\mathcal{M}$ is the *true* model?

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Once a model $\mathcal{M}$ is chosen, the usual question is:
How well does it describe the set of observations $\hat{s}$?
$P(\hat{s}|\mathcal{M}, \boldsymbol{\theta})$ or $P(\hat{s}|\mathcal{M})$

The **right** question is:
Given the observations $\hat{s}$ how likely it is that $\mathcal{M}$ is the *true* model?
$P(\mathcal{M}, \boldsymbol{\theta}|\hat{s})$ or $P(\mathcal{M}|\hat{s})$

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Once a model $\mathcal{M}$ is chosen, the usual question is:
How well does it describe the set of observations $\hat{s}$?
$P(\hat{s}|\mathcal{M}, \boldsymbol{\theta})$ or $P(\hat{s}|\mathcal{M})$

The **right** question is:
Given the observations $\hat{s}$ how likely it is that $\mathcal{M}$ is the *true* model?
$P(\mathcal{M}, \boldsymbol{\theta}|\hat{s})$ or $P(\mathcal{M}|\hat{s})$

Bayes Theorem

$$P(\mathcal{M}|\hat{s}) = \frac{P(\hat{s}|\mathcal{M})P(\mathcal{M})}{P(\hat{s})} = \frac{P(\hat{s}|\mathcal{M})P(\mathcal{M})}{\sum_i P(\hat{s}|\mathcal{M}_i)P(\mathcal{M}_i)}$$

Recipe for how to update our (un)certainty about a model – *hypothesis* – given some data.

Model selection as hypothesis test

A model, $\mathcal{M}$, defines a family of probability distributions, including its dependence on parameters, $P(\hat{s}|\boldsymbol{\theta}, \mathcal{M})$, and their distribution $P(\boldsymbol{\theta}|\mathcal{M})$.

Once a model $\mathcal{M}$ is chosen, the usual question is:
How well does it describe the set of observations $\hat{s}$?
$P(\hat{s}|\mathcal{M}, \boldsymbol{\theta})$ or $P(\hat{s}|\mathcal{M})$

The **right** question is:
Given the observations $\hat{s}$ how likely it is that $\mathcal{M}$ is the *true* model?
$P(\mathcal{M}, \boldsymbol{\theta}|\hat{s})$ or $P(\mathcal{M}|\hat{s})$

Bayes Theorem

$$P(\mathcal{M}|\hat{s}) = \frac{P(\hat{s}|\mathcal{M})P(\mathcal{M})}{P(\hat{s})} = \frac{P(\hat{s}|\mathcal{M})P(\mathcal{M})}{\sum_i P(\hat{s}|\mathcal{M}_i)P(\mathcal{M}_i)}$$

Recipe for how to update our (un)certainty about a model – *hypothesis* – given some data.

$$\{\mathcal{M}_\alpha\}_{\alpha=1}^N \xrightarrow[\hat{s}]{\text{Bayes}} \left\{P(\mathcal{M}_\alpha|\hat{s})\right\}_{\alpha=1}^N \Longrightarrow \mathcal{M}^* = \arg\max_{\mathcal{M}_\alpha} P(\mathcal{M}_\alpha|\hat{s})$$

Binary models: Likelihoods and inference

$\hat{s} = \texttt{0100110101...1110100101}$, $|\hat{s}| = M$ bits, with k_0 and k_1 the frequencies of '0's and '1's; $k_0 + k_1 = M$.

$$\mathcal{M}_0: \quad p_0 = p_1 = \frac{1}{2} \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_0) = \frac{1}{2^M};$$

$$\mathcal{M}_1: \quad p_0 = \theta; \quad p_1 = 1 - \theta \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_1) = \theta^{k_0}(1-\theta)^{k_1}.$$

Binary models: Likelihoods and inference

$\hat{s} = \texttt{0100110101...1110100101}$, $|\hat{s}| = M$ bits, with k_0 and k_1 the frequencies of '0's and '1's; $k_0 + k_1 = M$.

$$\mathcal{M}_0: \quad p_0 = p_1 = \frac{1}{2} \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_0) = \frac{1}{2^M};$$

$$\mathcal{M}_1: \quad p_0 = \theta; \quad p_1 = 1 - \theta \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_1) = \theta^{k_0}(1-\theta)^{k_1}.$$

$P(\theta|\mathcal{M}_0) = \delta(\theta - \frac{1}{2})$

$P(\theta|\mathcal{M}_1) = P_{\text{Jeff}}(\theta) = \frac{\Gamma(1)}{\Gamma^2(\frac{1}{2})\sqrt{\theta(1-\theta)}}$

Binary models: Likelihoods and inference

$\hat{s} = \texttt{0100110101...1110100101}$, $|\hat{s}| = M$ bits, with k_0 and k_1 the frequencies of '0's and '1's; $k_0 + k_1 = M$.

$$\mathcal{M}_0: \quad p_0 = p_1 = \frac{1}{2} \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_0) = \frac{1}{2^M};$$

$$\mathcal{M}_1: \quad p_0 = \theta; \quad p_1 = 1-\theta \quad \Longrightarrow P(\hat{s}|\theta, \mathcal{M}_1) = \theta^{k_0}(1-\theta)^{k_1}.$$

$P(\theta|\mathcal{M}_0) = \delta(\theta - \frac{1}{2})$

$P(\theta|\mathcal{M}_1) = P_{\text{Jeff}}(\theta) = \frac{\Gamma(1)}{\Gamma^2(\frac{1}{2})\sqrt{\theta(1-\theta)}}$

Likelihoods of the models $P(\hat{s}|\mathcal{M}) = \int \mathrm{d}\theta\, P(\theta|\mathcal{M}) P(\hat{s}|\theta, \mathcal{M})$

$$P(\hat{s}|\mathcal{M}_0) = \int \mathrm{d}\theta\, \delta\left(\theta - \frac{1}{2}\right) P(\hat{s}|\theta, \mathcal{M}_0) = \frac{1}{2^M}$$

$$P(\hat{s}|\mathcal{M}_1) = \frac{\Gamma(1)}{\Gamma^2(\frac{1}{2})} \int \mathrm{d}\theta\, \theta^{k_0 - \frac{1}{2}} (1-\theta)^{k_1 - \frac{1}{2}} = \frac{\Gamma(1)\Gamma(k_0 + \frac{1}{2})\Gamma(k_1 + \frac{1}{2})}{\Gamma^2(\frac{1}{2})\Gamma(M+1)}$$

Binary models: phase diagrams

Posterior:

$$P(\mathcal{M}_\alpha|\hat{s}) = \frac{P(\hat{s}|\mathcal{M}_\alpha)P_0(\mathcal{M}_\alpha)}{P_0(\hat{s})}$$
$$\propto P(\hat{s}|\mathcal{M}_\alpha)$$

$$\Longrightarrow \frac{P(\hat{s}|\mathcal{M}_0)}{P(\hat{s}|\mathcal{M}_1)}$$

Binary models: phase diagrams

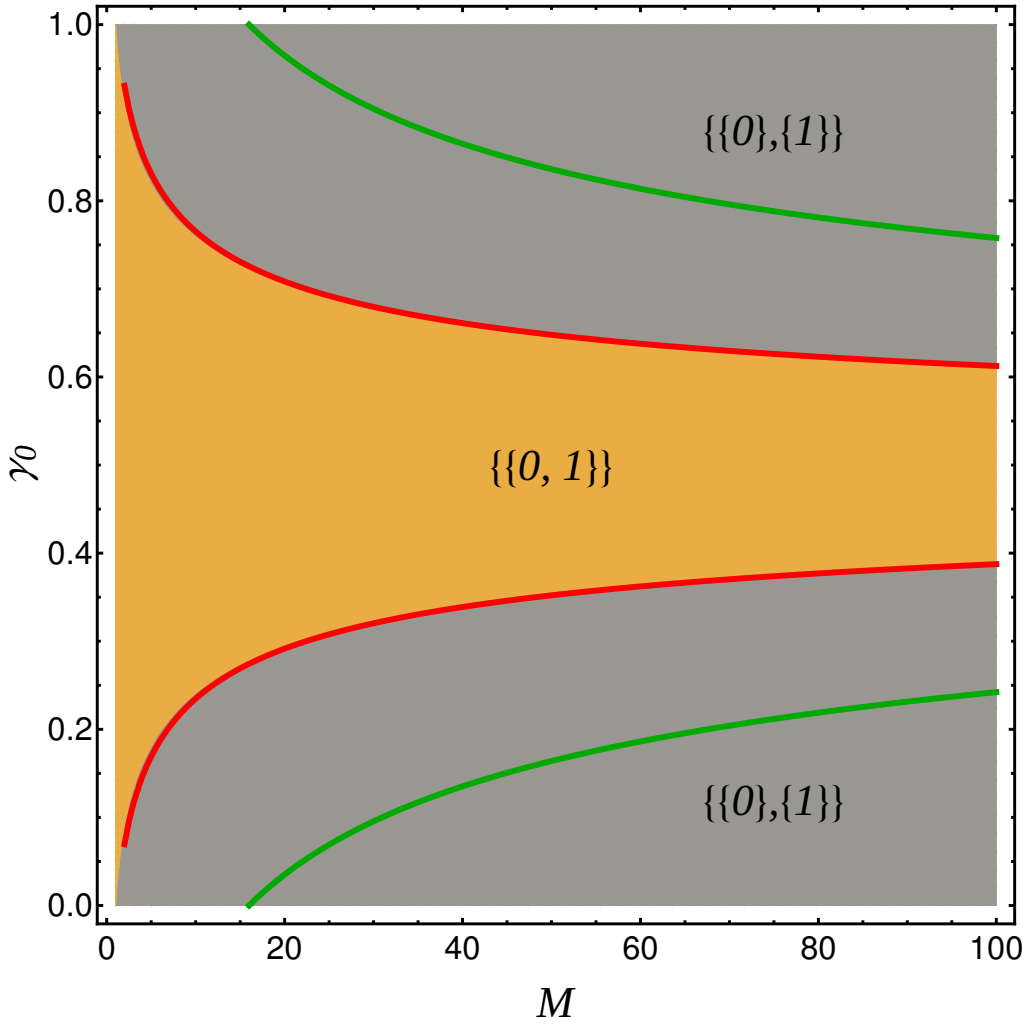
Posterior:

$$P(\mathcal{M}_\alpha|\hat{s}) = \frac{P(\hat{s}|\mathcal{M}_\alpha)P_0(\mathcal{M}_\alpha)}{P_0(\hat{s})}$$
$$\propto P(\hat{s}|\mathcal{M}_\alpha)$$

$$\Longrightarrow \frac{P(\hat{s}|\mathcal{M}_0)}{P(\hat{s}|\mathcal{M}_1)}$$

Region where $\frac{P(\mathcal{M}_0|\hat{s})}{P(\mathcal{M}_1|\hat{s})} > 1$

Region allowed by Borel Normality



Binary models: phase diagrams

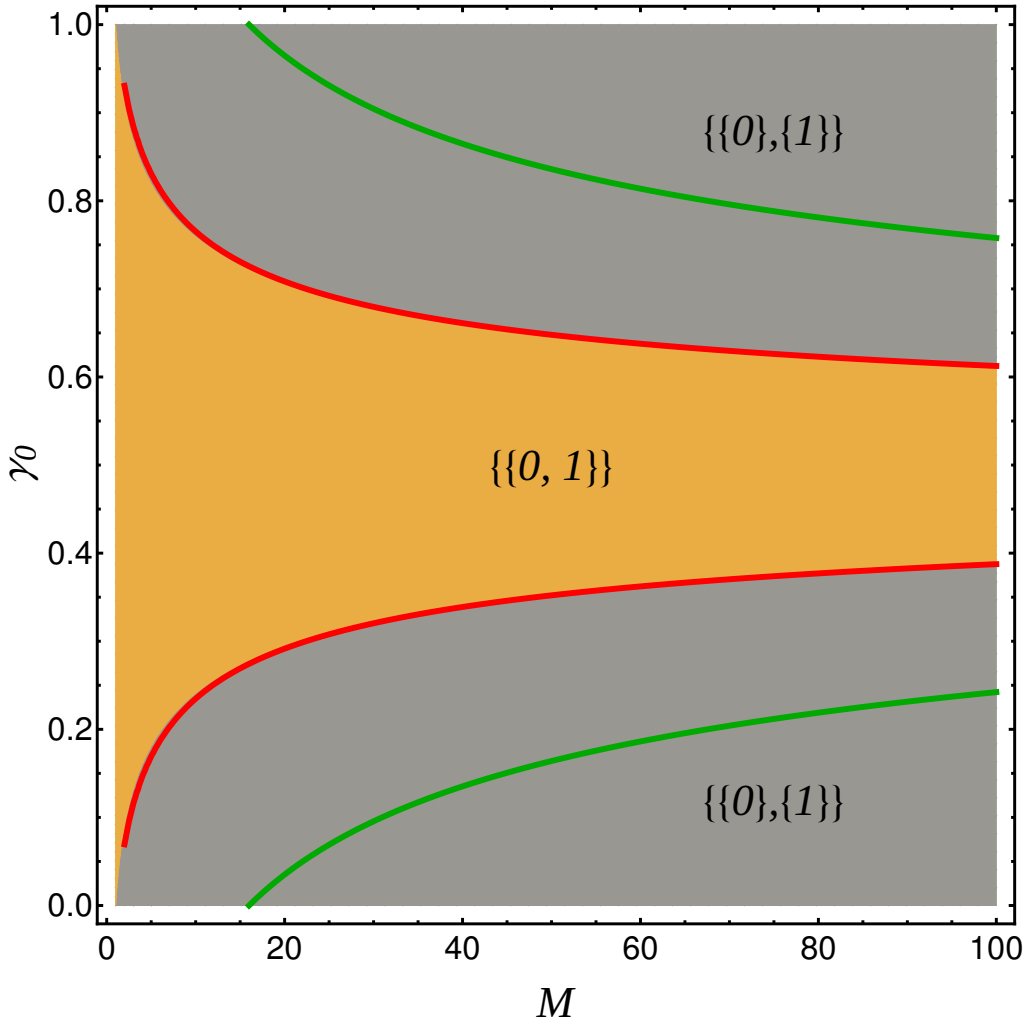
Posterior:

$$P(\mathcal{M}_\alpha|\hat{s}) = \frac{P(\hat{s}|\mathcal{M}_\alpha)P_0(\mathcal{M}_\alpha)}{P_0(\hat{s})}$$
$$\propto P(\hat{s}|\mathcal{M}_\alpha)$$

$$\Longrightarrow \frac{P(\hat{s}|\mathcal{M}_0)}{P(\hat{s}|\mathcal{M}_1)}$$

Region where $\frac{P(\mathcal{M}_0|\hat{s})}{P(\mathcal{M}_1|\hat{s})} > 1$

Region allowed by Borel Normality



$$\Xi = \{0,1\} \longrightarrow 2 \text{ partitions} \Longleftrightarrow 2 \text{ models}$$

$\mathcal{M}_0 \longleftrightarrow \{\{0,1\}\}$ $\qquad$ $\mathcal{M}_1 \longleftrightarrow \{\{0\},\{1\}\}$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

Partitions of a set $\Xi_\beta = \{0, \ldots, 2^\beta - 1\}$

A partition, α_K, is a grouping of the elements of Ξ_β in K disjoint, non-empty subsets, $\{\omega_K(r)\}_{r=1}^K \implies \alpha_K = \bigcup_{r=1}^K \omega_K(r)$.

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \quad \xrightarrow[\text{simultaneously}]{\text{"read" } \beta \text{ bits}} \quad \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

Partitions of a set $\Xi_\beta = \{0, \ldots, 2^\beta - 1\}$

A partition, α_K, is a grouping of the elements of Ξ_β in K disjoint, non-empty subsets, $\{\omega_K(r)\}_{r=1}^K \implies \alpha_K = \bigcup_{r=1}^K \omega_K(r)$.

e.g., with $\beta = 2$ there are *six* partitions of $\Xi_2 = \{\texttt{0,1,2,3}\}$ in $K = 3$ subsets.

$$\alpha_3^{(1)} = \{\{\texttt{0}\}, \{\texttt{1}\}, \{\texttt{2,3}\}\}; \quad \alpha_3^{(2)} = \{\{\texttt{0}\}, \{\texttt{2}\}, \{\texttt{1,3}\}\}; \quad \alpha_3^{(3)} = \{\{\texttt{0}\}, \{\texttt{3}\}, \{\texttt{1,2}\}\}$$

$$\alpha_3^{(4)} = \{\{\texttt{1}\}, \{\texttt{2}\}, \{\texttt{0,3}\}\}; \quad \alpha_3^{(5)} = \{\{\texttt{1}\}, \{\texttt{3}\}, \{\texttt{0,2}\}\}; \quad \alpha_3^{(6)} = \{\{\texttt{2}\}, \{\texttt{3}\}, \{\texttt{0,1}\}\}$$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \quad \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \quad \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

One partition $\iff$ one model

Partitions represent different ways to assign *biases* to strings.

$$\mathcal{M}_{\alpha_K^{(l)}} : \quad p_j = \frac{\theta_r}{\left|\omega_K^{(l)}(r)\right|}; \quad \forall j \in \omega_K^{(l)}(r); \quad r = 1, \ldots, K\,.$$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \quad \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \quad \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

One partition $\iff$ one model

Partitions represent different ways to assign *biases* to strings.

$$\mathcal{M}_{\alpha_K^{(l)}} : \quad p_j = \frac{\theta_r}{\left|\omega_K^{(l)}(r)\right|}; \quad \forall j \in \omega_K^{(l)}(r); \quad r = 1, \ldots, K\,.$$

$$\alpha_3^{(3)} = \{\{\texttt{0}\}, \{\texttt{3}\}, \{\texttt{1,2}\}\} \quad \implies \quad \mathcal{M}_{\alpha_3^{(3)}} : \quad p_0 = \theta_0 \neq p_3 = \theta_1 \neq p_1 = p_2 = \frac{\theta_2}{2}\,;$$

Partitions as a tool to identify regularities

What about the sequence $\hat{s}_{\text{reg}} = \texttt{0101010101...010101}$? ¿¿¿It's random???

$$\hat{s} = \underbrace{\texttt{110}}_{\beta}\texttt{100101011...10110} \quad \xrightarrow[\text{simultaneously}]{\text{“read” } \beta \text{ bits}} \quad \begin{cases} \beta = 1, & \hat{s} = \texttt{110100101011...010110} \\ \beta = 2, & \hat{s} = \texttt{310223...112} \\ \beta = 3 & \hat{s} = \texttt{6453...15} \end{cases}$$

$$\beta = 2 \implies \hat{s}_{\text{reg}} = \texttt{11111...111}.$$

One partition $\iff$ one model

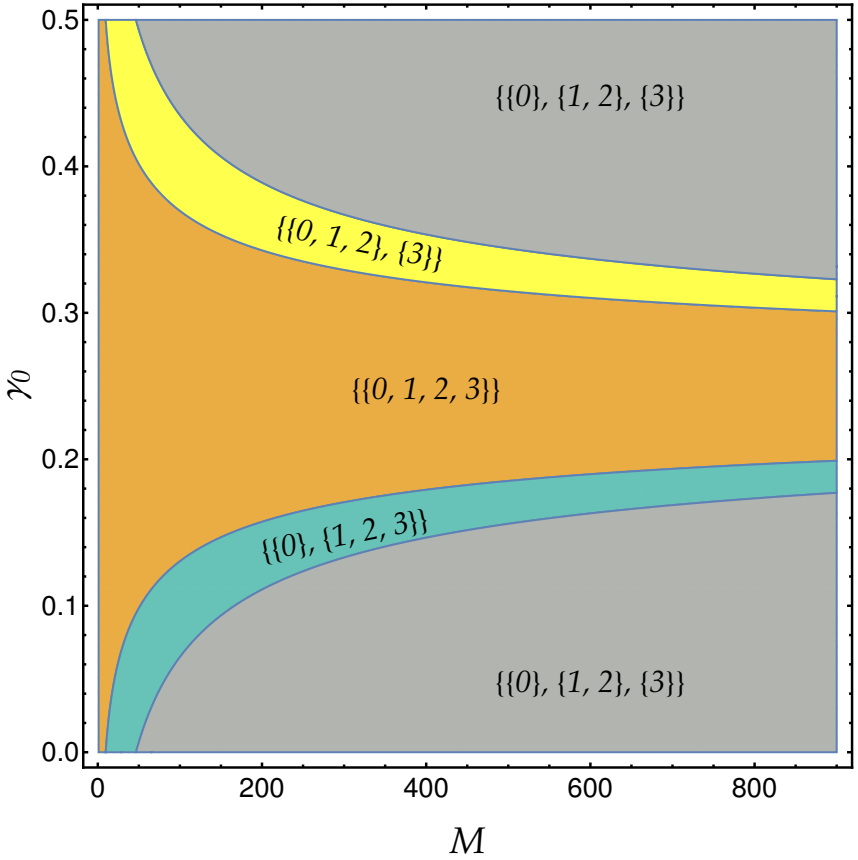
Partitions represent different ways to assign *biases* to strings.

$$\mathcal{M}_{\alpha_K^{(l)}} : \quad p_j = \frac{\theta_r}{\left|\omega_K^{(l)}(r)\right|}; \quad \forall j \in \omega_K^{(l)}(r); \quad r = 1, \ldots, K\,.$$

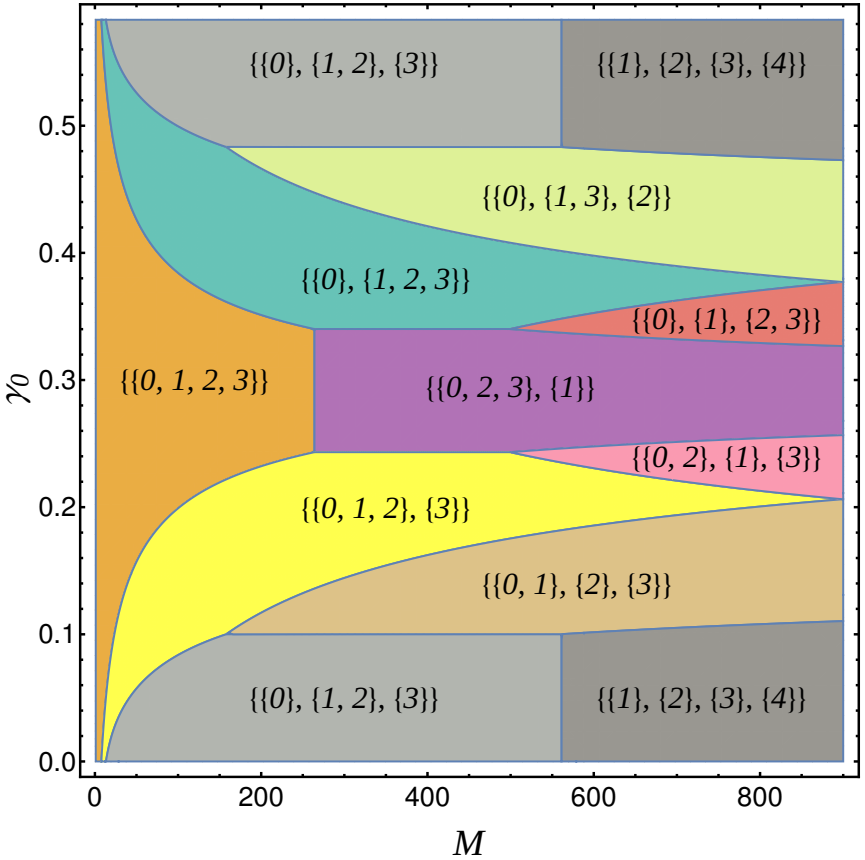
α_1 is **unique** $\implies \mathcal{M}_{\alpha_1} \equiv \mathcal{M}_{\text{sym}} : \; p_j = \frac{1}{2^\beta}$ is the **only** model describing a *maximally random* process .

Phase diagrams: $\beta = 2$

We can perform the model selection using only information about the frequencies $\{\gamma_i = \frac{\beta k_i}{M}\}_{i=0}^{2^\beta - 1}$ and the sequence length, M.



(a) $\gamma_1 = \gamma_2 = 1/4$.

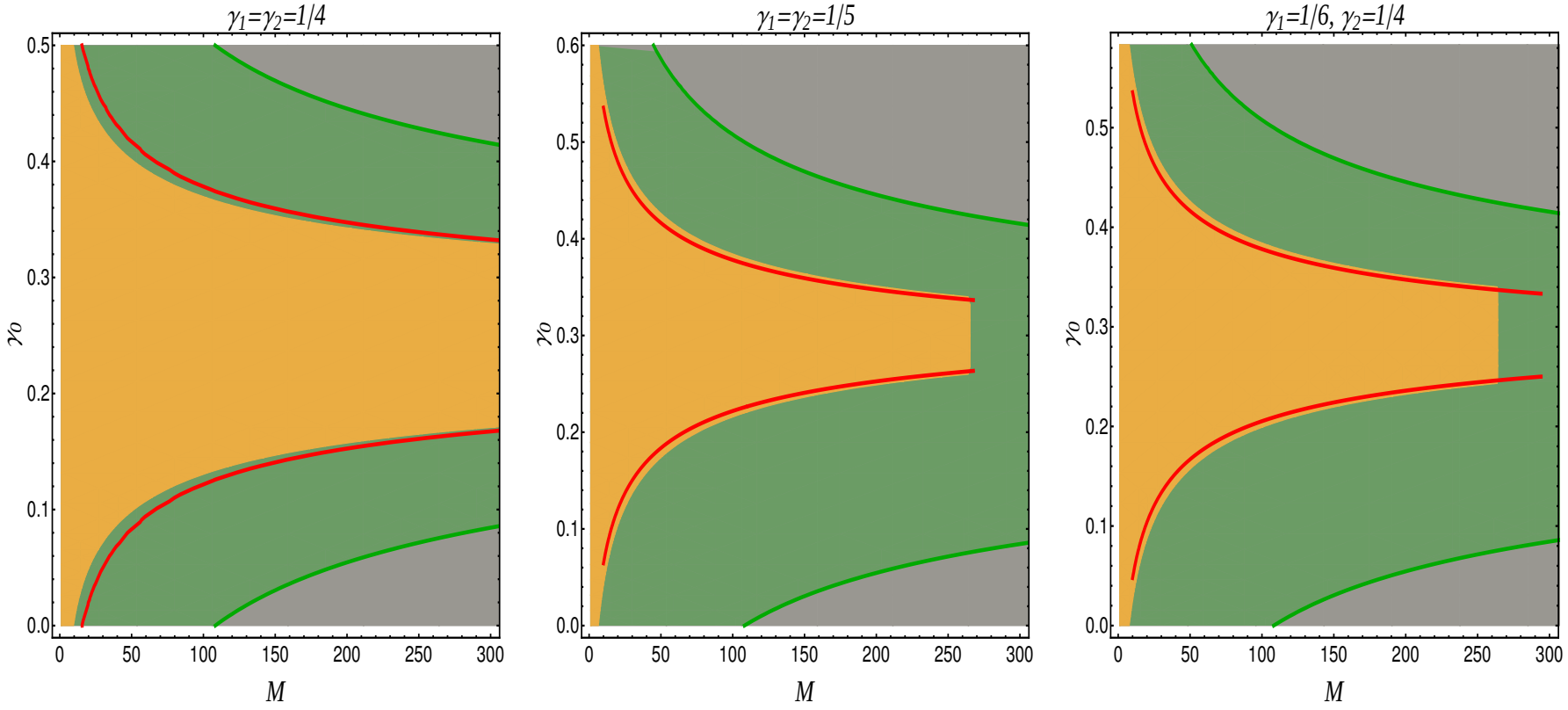


(b) $\gamma_1 = 1/6;\ \gamma_2 = 1/4$.

Comparing with Borel's Normality bounds

A sequence $\hat{s}$ is Borel-Normal if:

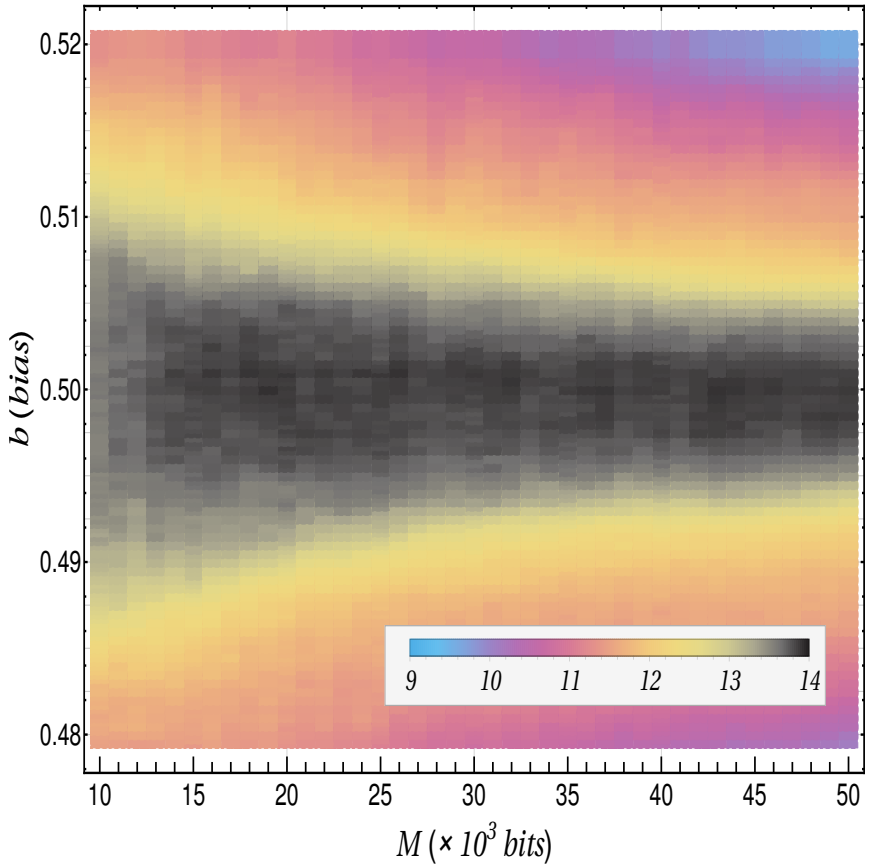
$$\left|\gamma_j^{(\beta)} - \frac{1}{2^\beta}\right| < \sqrt{\frac{\log_2 M}{M}}, \qquad \beta \leq \log_2 \log_2 M$$



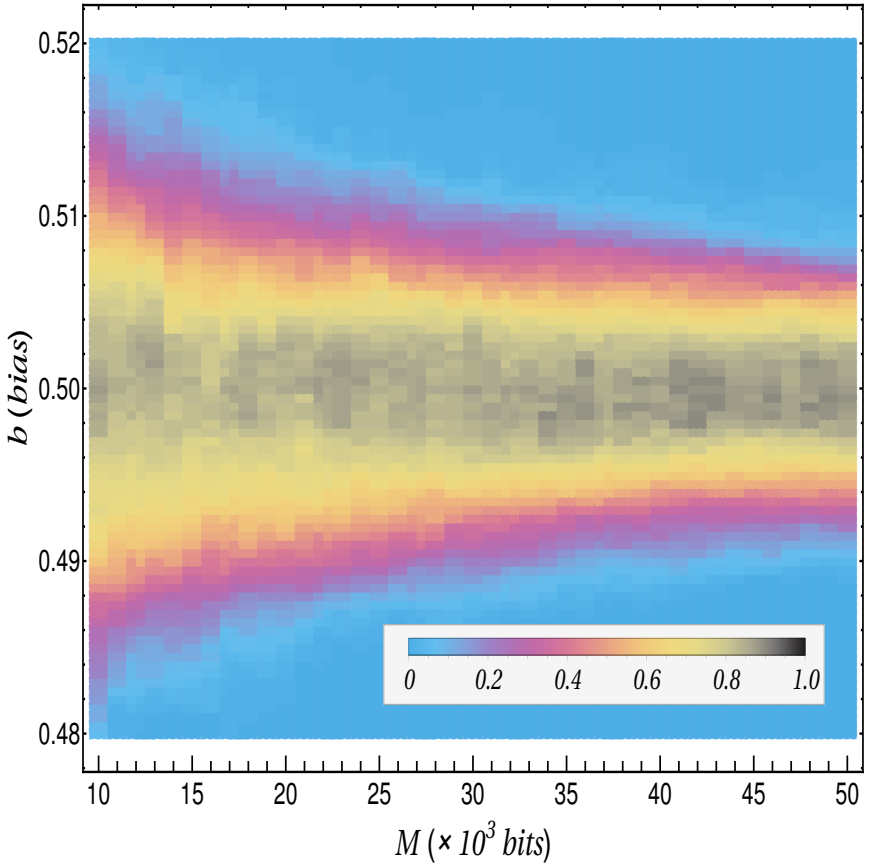
Region where $\mathcal{M}_{\text{sym}}$ is the likeliest , region allowed by Borel Normality test, and Approximated bounds obtained by an expansion of $\log \mathcal{M}_{\text{sym}}/\mathcal{M}_{\alpha_2}$.

Comparing with NIST tests

Used a RNG (Mathematica) to generate 100 bit sequences of different length, with $p_0 = b \in (0.48, 0.52)$.

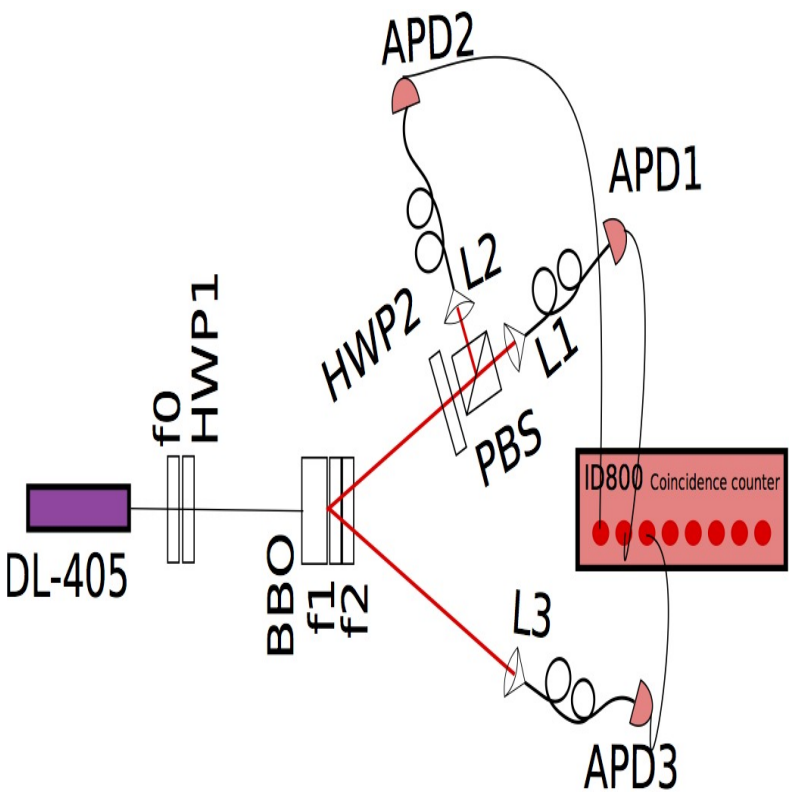


(a) Average of tests passed



(b) Fraction of times $\mathcal{M}_{\text{sym}}$ was selected using $\beta = \{1, 2, 3\}$.

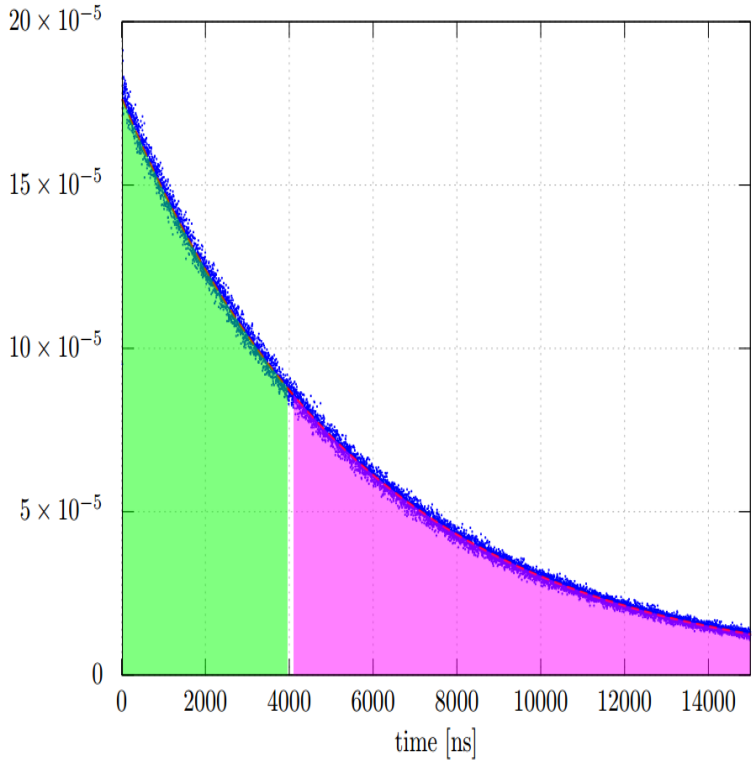
Analysing an experimentally generated sequence



Experimental setup to produce correlated photons.

A. Solis et al., "How random are random numbers generated using photons?", Physica Scripta **90**, 074034 (2015)

Analysing an experimentally generated sequence



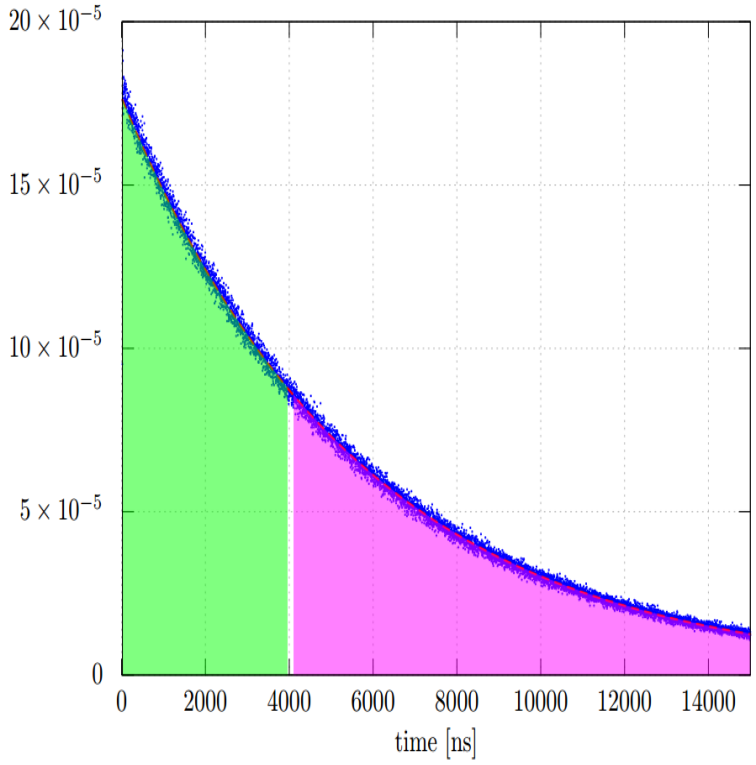
Distribution of Δt between the simultaneous detections. It is used to generate strings of 0 and 1.

$$M = 4 \times 10^9 \text{ bits!}$$

$$\beta_{max} \lesssim 5$$

A. Solis et al., "How random are random numbers generated using photons?", Physica Scripta **90**, 074034 (2015)

Analysing an experimentally generated sequence



Distribution of Δt between the simultaneous detections. It is used to generate strings of 0 and 1.

$$M = 4 \times 10^9 \text{ bits!}$$

$$\beta_{max} \lesssim 5$$

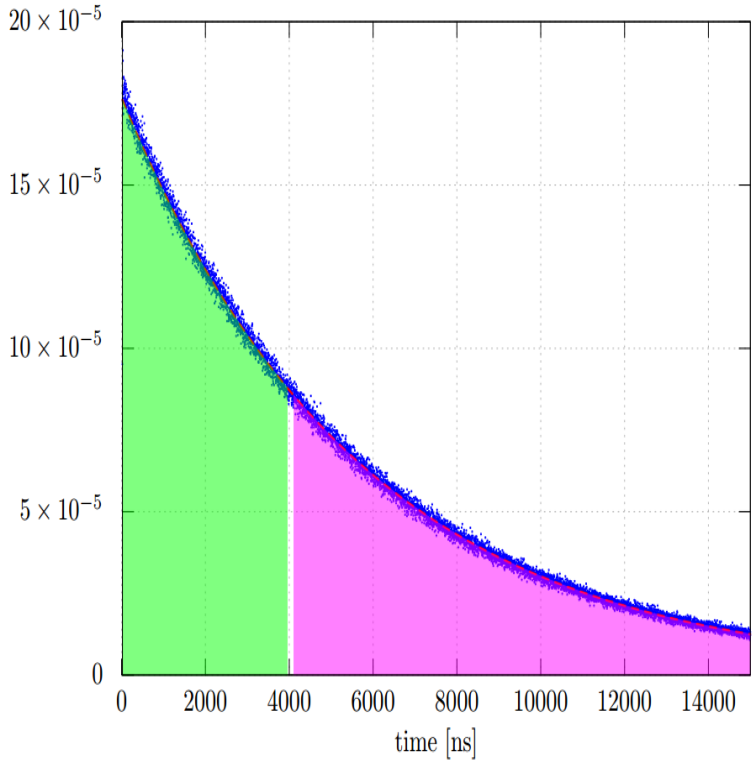
β	$P(\mathcal{M}_{\text{sym}}\|\hat{s})$	$\log_{10} \text{BF}_{\text{sym},\alpha}$
1	0.999965	4.45
2	0.999562	≥ 3.72
3	0.968353	≥ 2.01
4	0.46718	≥ 3.46

For $\beta = 4$, only models associated to partitions into $K = 2$ subsets were considered.

$$\left\{ \begin{matrix} 2^4 \\ 2 \end{matrix} \right\} = 32,767 \quad \Longrightarrow \quad P(\mathcal{M}) = 3 \times 10^{-5} \, .$$

A. Solis et al., "How random are random numbers generated using photons?", Physica Scripta **90**, 074034 (2015)

Analysing an experimentally generated sequence



Distribution of Δt between the simultaneous detections. It is used to generate strings of 0 and 1.

A. Solis et al., "How random are random numbers generated using photons?", Physica Scripta **90**, 074034 (2015)

$$M = 4 \times 10^9 \text{ bits!}$$

$$\beta_{max} \lesssim 5$$

β	$P(\mathcal{M}_{\text{sym}}\|\hat{s})$	$\log_{10} \text{BF}_{\text{sym},\alpha}$
1	0.999965	4.45
2	0.999562	≥ 3.72
3	0.968353	≥ 2.01
4	0.46718	≥ 3.46

For $\beta = 4$, only models associated to partitions into $K = 2$ subsets were considered.

$$\left\{ \begin{matrix} 2^4 \\ 2 \end{matrix} \right\} = 32,767 \quad \Longrightarrow \quad P(\mathcal{M}) = 3 \times 10^{-5}\,.$$

The device functions as a random *source*.

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Improving randomness characterization through Bayesian model selection

Received: 4 November 2016
Accepted: 7 April 2017
Published online: 08 June 2017

Rafael Díaz Hernández Rojas[1], Aldo Solís[2], Alí M. Angulo Martínez[2], Alfred B. U'Ren[2], Jorge G. Hirsch[2], Matteo Marsili[3] & Isaac Pérez Castillo[1,4]

Thank you!