# Entropy and Simulation of No-Signaling Models 

Adam Brandenburger
Stern School of Business
Tandon School of Engineering
NYU Shanghai
New York University

Pierfrancesco La Mura
HHL-Leipzig Graduate School of Management
Leipzig

# Simulating Physics with Computers 

Richard P. Feynman<br>Department of Physics, California Institute of Technology, Pasadena, California 91107

You see, nature's
unpredictable; how do you expect to predict it with a computer? You can't, -it's unpredictable if it's probabilistic. But what you really do in a probabilistic system is repeat the experiment in nature a large number of times. If you repeat the same experiment in the computer a large number of times (and that doesn't take any more time than it does to do the same thing in nature of course), it will give the frequency of a given final state proportional to the number of times, with approximately the same rate (plus or minus the square root of $n$ and all that) as it happens in nature.

The only difference between a probabilistic classical world and the equations of the quantum world is that somehow or other it appears as if the probabilities would have to go negative, and that we do not know, as far as I know, how to simulate. Okay, that's the fundamental problem.

Empirical model:

|  | $(0,0)$ | $(1,0)$ | $(0,1)$ | $(1,1)$ |
| :--- | :---: | :---: | :---: | :---: |
| $(a, b)$ | $f_{1}$ | $f_{2}$ | $f_{3}$ | $f_{4}$ |
| $\left(a^{\prime}, b\right)$ | $f_{5}$ | $f_{6}$ | $f_{7}$ | $f_{8}$ |
| $\left(a, b^{\prime}\right)$ | $f_{9}$ | $f_{10}$ | $f_{11}$ | $f_{12}$ |
| $\left(a^{\prime}, b^{\prime}\right)$ | $f_{13}$ | $f_{14}$ | $f_{15}$ | $f_{16}$ |

Bell model:

|  | $(0,0)$ | $(1,0)$ | $(0,1)$ | $(1,1)$ |
| :--- | :---: | :---: | :---: | :---: |
| $(a, b)$ | $1 / 2$ | 0 | 0 | $1 / 2$ |
| $\left(a^{\prime}, b\right)$ | $3 / 8$ | $1 / 8$ | $1 / 8$ | $3 / 8$ |
| $\left(a, b^{\prime}\right)$ | $3 / 8$ | $1 / 8$ | $1 / 8$ | $3 / 8$ |
| $\left(a^{\prime}, b^{\prime}\right)$ | $1 / 8$ | $3 / 8$ | $3 / 8$ | $1 / 8$ |

PR Box:

|  | $(0,0)$ | $(1,0)$ | $(0,1)$ | $(1,1)$ |
| :--- | :---: | :---: | :---: | :---: |
| $(a, b)$ | $1 / 2$ | 0 | 0 | $1 / 2$ |
| $\left(a^{\prime}, b\right)$ | $1 / 2$ | 0 | 0 | $1 / 2$ |
| $\left(a, b^{\prime}\right)$ | $1 / 2$ | 0 | 0 | $1 / 2$ |
| $\left(a^{\prime}, b^{\prime}\right)$ | 0 | $1 / 2$ | $1 / 2$ | 0 |

Phase-space model:

|  | $a$ | $a^{\prime}$ | $b$ | $b^{\prime}$ |
| :--- | :--- | :--- | :--- | :--- |
| $p_{0}$ | 0 | 0 | 0 | 0 |
| $p_{1}$ | 0 | 0 | 0 | 1 |
| $p_{2}$ | 0 | 0 | 1 | 0 |
| $p_{3}$ | 0 | 0 | 1 | 1 |
| $p_{4}$ | 0 | 1 | 0 | 0 |
| $p_{5}$ | 0 | 1 | 0 | 1 |
| $p_{6}$ | 0 | 1 | 1 | 0 |
| $p_{7}$ | 0 | 1 | 1 | 1 |
| $p_{8}$ | 1 | 0 | 0 | 0 |
| $p_{9}$ | 1 | 0 | 0 | 1 |
| $p_{10}$ | 1 | 0 | 1 | 0 |
| $p_{11}$ | 1 | 0 | 1 | 1 |
| $p_{12}$ | 1 | 1 | 0 | 0 |
| $p_{13}$ | 1 | 1 | 0 | 1 |
| $p_{14}$ | 1 | 1 | 1 | 0 |
| $p_{15}$ | 1 | 1 | 1 | 1 |


|  | $(0,0)$ | $(1,0)$ | $(0,1)$ | $(1,1)$ |
| :--- | :---: | :---: | :---: | :---: |
| $(a, b)$ | $f_{1}$ | $f_{2}$ | $f_{3}$ | $f_{4}$ |
| $\left(a^{\prime}, b\right)$ | $f_{5}$ | $f_{6}$ | $f_{7}$ | $f_{8}$ |
| $\left(a, b^{\prime}\right)$ | $f_{9}$ | $f_{10}$ | $f_{11}$ | $f_{12}$ |
| $\left(a^{\prime}, b^{\prime}\right)$ | $f_{13}$ | $f_{14}$ | $f_{15}$ | $f_{16}$ |


|  | $a$ | $a^{\prime}$ | $b$ | $b^{\prime}$ |
| :--- | :--- | :--- | :--- | :--- |
| $p_{0}$ | 0 | 0 | 0 | 0 |
| $p_{1}$ | 0 | 0 | 0 | 1 |
| $p_{2}$ | 0 | 0 | 1 | 0 |
| $p_{3}$ | 0 | 0 | 1 | 1 |
| $p_{4}$ | 0 | 1 | 0 | 0 |
| $p_{5}$ | 0 | 1 | 0 | 1 |
| $p_{6}$ | 0 | 1 | 1 | 0 |
| $p_{7}$ | 0 | 1 | 1 | 1 |
| $p_{8}$ | 1 | 0 | 0 | 0 |
| $p_{9}$ | 1 | 0 | 0 | 1 |
| $p_{10}$ | 1 | 0 | 1 | 0 |
| $p_{11}$ | 1 | 0 | 1 | 1 |
| $p_{12}$ | 1 | 1 | 0 | 0 |
| $p_{13}$ | 1 | 1 | 0 | 1 |
| $p_{14}$ | 1 | 1 | 1 | 0 |
| $p_{15}$ | 1 | 1 | 1 | 1 |

Theorem (Abramsky-Brandenburger 2011): An empirical model (formulated much more generally) can be realized by a phase-space model with signed probabilities if and only if the empirical model satisfies no signaling.

But what do negative probabilities --- even if unobserved --- mean?
Can they be given an operational interpretation within classical physics?

[^0]Theorem (Abramsky-Brandenburger 2011): An empirical model (formulated much more generally) can be realized by a phase-space model with signed probabilities if and only if the empirical model satisfies no signaling.

But what do negative probabilities --- even if unobserved --- mean?
Can they be given an operational interpretation within classical physics?

The answer in Abramsky-Brandenburger (2014) is to push the minus sign in from probabilities to events

Think of an urn of colored balls with + or - painted on them!

A red-green and blue-green color blind person picks balls out of an urn and reports: "One-quarter of the time, the ball is red or green, and one-third of the time it is blue or green."

Is this possible?
If so, then we get
$1 / 4+1 / 3=\operatorname{Prob}($ Red $)+\operatorname{Prob}($ Green $)+\operatorname{Prob}($ Blue $)+\operatorname{Prob}($ Green $)=1+\operatorname{Prob}($ Green $)$
from which
$\operatorname{Prob}($ Green $)=-5 / 12!$
$\operatorname{Prob}($ Red $)=8 / 12$
$\operatorname{Prob}($ Blue $)=9 / 12$


An answer: The balls come with + or - signs


We calculate the empirical probability of e.g. the event \{Red, Green\} as:

(\# Red ${ }^{+}$balls - \# Red ${ }^{-}$balls) + (\# Green ${ }^{+}$balls - \# Green${ }^{-}$balls)



There is no information flow between Alice's location and Bob's location, which may be space-like separated

So, this is a bona fide physical (and non-quantum) simulation

How can we assess the 'efficiency' of this simulation?
Let's start by calculating entropies
The standard axioms for Renyí entropy extend to signed probability measures and, if we require sufficient smoothness, yield a parametric family of entropy functionals (Brandenburger and La Mura 2015):

$$
H_{\alpha}(p)=-\frac{1}{\alpha-1} \log \left(\sum_{i}\left|p_{i}\right|^{\alpha}\right) \text { where } \alpha>1
$$

Proposition: Fix a phase-space model with signed probability measure $p$ and the associated probability model with signed events and nonnegative probability measure $q$. Then $H_{\alpha}(q) \geq H_{\alpha}(p)$ for all $\alpha>1$, with strict inequality if at least one component of $p$ is strictly negative.

Sample $n$ times from a multinomial distribution $p=\left(p_{1}, \ldots, p_{k}\right)$

For each $i=1, \ldots, k$, let $f_{i}$ be the \# of times the $i$ th outcome is obtained and write $r_{i}=f_{i} / n$

Sanov's Theorem: The probability of obtaining $r=\left(r_{1}, \ldots, r_{k}\right)$ is given by

$$
\log \operatorname{Pr}(r ; p, n)=-n H_{1}(r, p)+o(n)
$$

where

$$
H_{1}(r, p)=\sum_{i=1}^{k} r_{i} \log \frac{r_{i}}{p_{i}}
$$

is the Kullback-Leibler divergence (relative entropy).

Consider two multinomial distributions

$$
\begin{aligned}
& p=\left(p_{1}, p_{2}, p_{3}, \ldots, p_{k}\right) \\
& p^{\prime}=\left(p_{1}-\varepsilon, p_{2}+\varepsilon, p_{3}, \ldots, p_{k}\right)
\end{aligned}
$$

where $p_{1}>p_{2}$ and $\varepsilon<p_{1}-p_{2}$
It is standard to show that $H_{1}\left(p^{\prime}\right)>H_{1}(p)$

## Proposition: Let

$$
p^{\prime \prime}=\left(p_{1}-1 / 2 \varepsilon, p_{2}+1 / 2 \varepsilon, p_{3}, \ldots, p_{k}\right)
$$

so that $p^{\prime \prime}$ is equidistant from $p$ and $p^{\prime}$. Then $H_{1}\left(p^{\prime \prime}, p^{\prime}\right)<H_{1}\left(p^{\prime \prime}, p\right)$ for small $\varepsilon$. Therefore, by Sanov's Theorem, $p^{\prime \prime}$ is more likely to be observed under $p^{\prime}$ than under $p$.

# Simulating Physics with Computers 

Richard P. Feynman<br>Department of Physics, California Institute of Technology, Pasadena, California 91107

unpredictable; how ds -it's unpredictable probabilistic system i times. If you repeat th times (and that doesn't for "deviations" an it acu. to the same thing in nature of course), it will give the frequency of a give.. final state proportional to the number of times, with approximately the same rate (plus or minus the square root of $n$ and all that) as it happens in nature.

The only difference between a probabil We saw that this world and the equations of the quantum world is that someho can be achieved appears as if the probabilities would have to go negative, and tl by pushing the know, as far as I know, how to simulate. Vimy, negative sign problem.

For "fluctuations," perhaps, but we saw some evidence that this may not be true with a computer? You can't, $t$ what you really do in a in nature a large number of nuter a large number of . inwards

You see, nature's


[^0]:    Abramsky, S., and A. Brandenburger, "The Sheaf-Theoretic Structure of Non-Locality and Contextuality," New Journal of Physics, 13, 2011, 113036; Abramsky, S., and A. Brandenburger, "An Operational Interpretation of Negative Probabilities and No-Signalling Models," in van Breugel, F., E. Kashefi, C.

