## Yet Another Statistical Analysis of the data of the (2015)

# "Loophole-Free" Bell-CHSH Experiments


Richard D. Gill

**Leiden University**
**Mathematical Institute**

*Combray*
*Causality Consultancy*

http://www.math.leidenuniv.nl/~gill
http://richardgill.nl


Shanghai / Beijing, September / October 2019

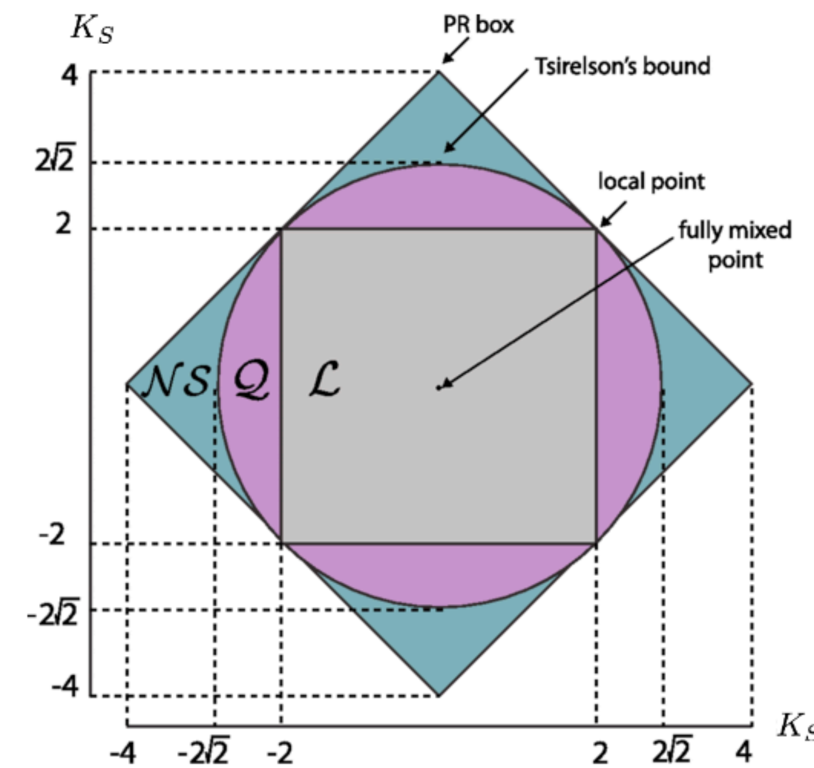
**I present novel statistical analyses of the data of the famous Bell-inequality experiments of 2015 and 2016: Delft, NIST, Vienna and Munich. Every statistical analysis relies on statistical assumptions. I'll make the traditional, but questionable, i.i.d. assumptions. They justify a novel (?) analysis which is both simple and (close to) optimal.**

**It enables us to fairly compare the results of the two main types of experiments: NIST and Vienna CH-Eberhard "one-channel" experiment with settings and state chosen to optimise the handling of the detection loophole (detector efficiency > 66.7%); Delft and Munich CHSH "two channel" experiments based on entanglement swapping, with the state and settings which achieve the Tsirelson bound (detector efficiency ≈ 100%).**

**One cannot say which type of experiment is better without agreeing on how to compromise between the desires to obtain high statistical significance and high physical significance. Moreover, robustness to deviations from traditional assumptions is also an issue**

# The local polytope

- The local polytope of a 2x2x2 experiment has exactly 8 facets, A. Fine (1982).
- They are the 8 one-sided CHSH inequalities
- They are necessary and sufficient for LR. There are no other 2x2x2 inequalities!
- CH, Eberhard, J are therefore *just* different ways to write CHSH !
- Yet with experimental data they give different results !?




**The diagram should be imagined as drawn on a plane in a higher dimensional space**
**The experimental data is a point close to, but not on, the plane**

# Raw counts

"One channel" experiment

| | | Settings | | | |
|---|---|---|---|---|---|
| | | 11 | 12 | 21 | 22 |
| Outcomes | dd | 141.439 | 146.831 | 158.338 | 8.392 |
| | dn | 73.391 | 67.941 | 425.067 | 576.445 |
| | nd | 76.224 | 326.768 | 58.742 | 463.985 |
| | nn | 875.392.736 | 874.976.534 | 875.239.860 | 874.651.457 |
| | Totals | 875.683.790 | 875.518.074 | 875.882.007 | 875.700.279 |

Vienna

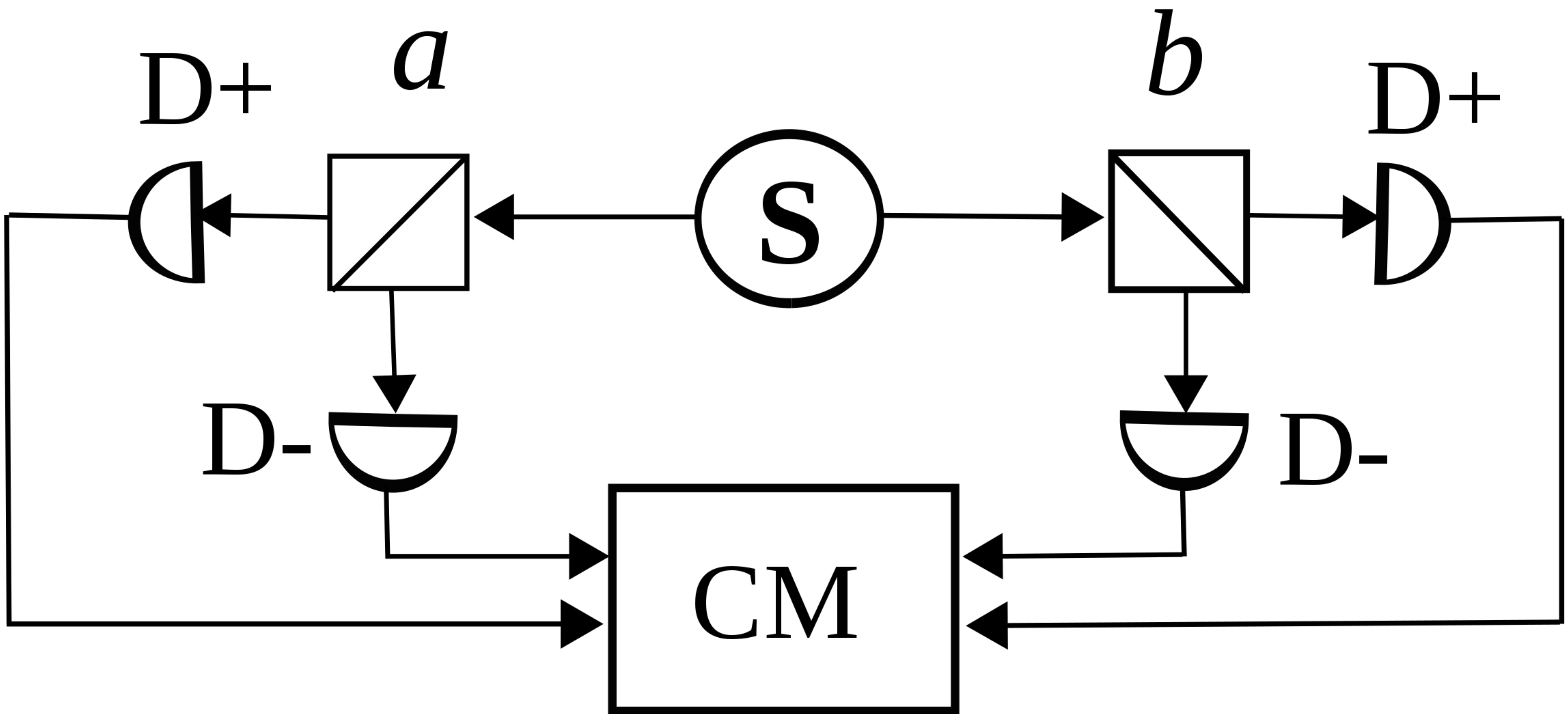
# Normalised counts

| | | Settings | | | |
|---|---|---|---|---|---|
| | | 11 | 12 | 21 | 22 |
| Outcomes | dd | 162 | 168 | 181 | 10 |
| | dn | 84 | 78 | 485 | 658 |
| | nd | 87 | 373 | 67 | 530 |
| | nn | 999.668 | 999.381 | 999.267 | 998.802 |
| | Totals | 1.000.000 | 1.000.000 | 1.000.000 | 1.000.000 |

| Normaliser | | Normalised | | |
|---|---|---|---|---|
| 1.000.000 | 1.000.000 | 1.000.000 | 1.000.000 | 1.000.000 |

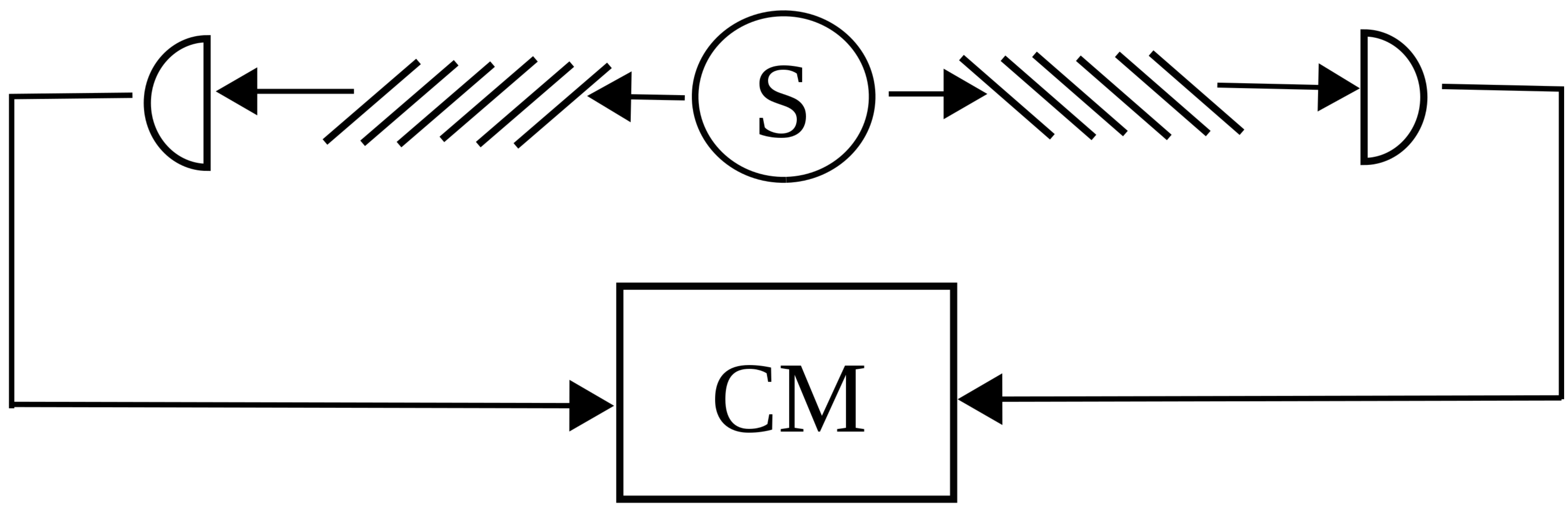
"d" = detection, "n" = no detection

**“Two channel” experiment (CHSH - Aspect, Weihs, …, Delft, Munich)**




***Clocked* experiment: outcomes on each side are “+”,”–“, or “0”**

**“One channel” experiment (Clauser-Horne, Eberhard, Vienna, NIST)**




**Outcomes on each side are “d” corresponding to “+” and “n” corresponding to “–” or “0”**

**Peter Bierhorst**

**Philippe Eberhard**

Lawrence Berkeley National Laboratory

Location
Berkeley, United States

Department
Physics Division

Position
Retired Senior staff

$$S = 2 + 4J$$

$$J = (S - 2)/4$$

- The experiments in Vienna and at NIST (Boulder, Colorado) do *not* use the singlet state
- They exploit the fact that QM *can* violate CHSH from 66% detector efficiency upwards

**Jason Semitecolos**

- Clauser-Horne (1974)
- Philippe H. Eberhard (1993)

**Experimental mathematics**

- Jan-Åke Larsson and Jason Semitecolos (2001)

**Proof of 67% bound**

- Peter Bierhorst (2016), “Geometric decompositions of Bell polytopes with practical applications”, *Journal of Physics A: Mathematical and Theoretical*

**Proof of 67% bound**

**Jan-Åke Larsson**

# P.H. Eberhard (1993)

The vector $\psi$ turned out to be of the form

$$\psi = \frac{1}{2\sqrt{1+r^2}} \left| \begin{array}{c} (1+r)e^{-i\omega} \\ -(1-r) \\ -(1-r) \\ (1+r)e^{i\omega} \end{array} \right| , \tag{31}$$

which can be reached in the two-photon experiment considered in this paper by first superposing states $| \leftrightarrow \updownarrow >$ and $| \updownarrow \leftrightarrow >$ in unequal amounts,

$$\psi_0 = (1/\sqrt{1+r^2}) \left( | \leftrightarrow \updownarrow > + r \, | \updownarrow \leftrightarrow > \right) , \tag{32}$$

then rotating the planes of polarization of $a$ and of $b$ in setup $(\alpha_1, \beta_1)$ by the angles

$$\alpha_1 = (\omega/2) - 90^\circ , \tag{33}$$

$$\beta_1 = \omega/2 , \tag{34}$$

respectively, and using the values of $r$, $\omega$, and $\alpha_1 - \alpha_2$ ($\equiv \beta_1 - \beta_2$) given in Table II. Note that, for $\eta = 1$, the vector $\psi_0$ reduces to the value given by Eq. (1), and the angles $\alpha_1$, $\alpha_2$, $\beta_1$, and $\beta_2$ reduce to the values given by Eqs. (2)–(5).

TABLE II. Extreme conditions for a loophole-free experiment.

| $\eta$ (%) | $\zeta$ (%) | $r$ | $\omega$ (deg) | $\alpha_1 - \alpha_2$ (deg) |
|---|---|---|---|---|
| 66.7 | 0.00 | 0.001 | 0.0 | 2.2 |
| 70 | 0.02 | 0.136 | 3.4 | 21.4 |
| 75 | 0.31 | 0.311 | 9.7 | 32.0 |
| 80 | 1.10 | 0.465 | 14.9 | 37.9 |
| 85 | 2.48 | 0.608 | 18.6 | 41.5 |
| 90 | 4.50 | 0.741 | 20.9 | 43.6 |
| 95 | 7.12 | 0.871 | 22.1 | 44.7 |
| 100 | 10.36 | 1.000 | 22.5 | 45.0 |

| | | Settings | | | |
|---|---|---|---|---|---|
| | | 11 | 12 | 21 | 22 |
| Outcomes | dd | 141.439 | 146.831 | 158.338 | 8.392 |
| | dn | 73.391 | 67.941 | 425.067 | 576.445 |
| | nd | 76.224 | 326.768 | 58.742 | 463.985 |
| | nn | 875.392.736 | 874.976.534 | 875.239.860 | 874.651.457 |
| | Totals | 875.683.790 | 875.518.074 | 875.882.007 | 875.700.279 |

**Normaliser**

**1000000**

| | | Settings | | | |
|---|---|---|---|---|---|
| | | 11 | 12 | 21 | 22 |
| Outcomes | dd | 162 | 168 | 181 | 10 |
| | dn | 84 | 78 | 485 | 658 |
| | nd | 87 | 373 | 67 | 530 |
| | nn | 999.668 | 999.381 | 999.267 | 998.802 |
| | Totals | 1.000.000 | 1.000.000 | 1.000.000 | 1.000.000 |

| Normaliser | Normalised | | | |
|---|---|---|---|---|
| **1.000.000** | 1.000.000 | 1.000.000 | 1.000.000 | 1.000.000 |

| | Outcomes | Bob Setting 1 “ d ” | Bob Setting 1 “ n ” |
|---|---|---|---|
| Alice Setting 1 | “ d ” | 162 | 84 |
| | “ n ” | 87 | 999668 |
| | | rho11 = 0,999658 | |
| Alice Setting 2 | “ d ” | 181 | 485 |
| | “ n ” | 67 | 999267 |
| | | rho21 = 0,998895 | |

| | Bob Setting 2 “ d ” | Bob Setting 2 “ n ” |
|---|---|---|
| Alice Setting 1 | 168 | 78 |
| | 373 | 999381 |
| | rho12 = 0,999098 | |
| Alice Setting 2 | 10 | 658 |
| | 530 | 998802 |
| | rho22 = 0,997624 | |

**J =** 0,000007

**S = CHSH =** 2,000028

# Vienna

$$4\,J = (1 + a_1 + b_1 + z_{11})$$
$$- (1 - a_2 + b1 - z_{21})$$
$$- (1 + a_1 - b_2 - z_{12})$$
$$- (1 + a_2 + b_2 + z_{22})$$
$$= -2 + (z_{11} + z_{21} + z_{12} - z_{22})$$

$$4\ \rho_{11}$$
$$= (2 + 2\,z_{11}) - (2 - 2\,z_{11})$$
$$= 4\,z_{11}$$

$$4\,S = 4\ \text{CHSH}$$
$$= 4\,(\rho_{11} + \rho_{12} + \rho_{21} - \rho_{22})$$

$$S = z_{11} + z_{12} + z_{21} - z_{22}$$
$$= 2 + 4\,J$$

$$J = (S - 2) / 4$$

**4 x probs**

| | Outcomes | Bob Setting 1 “ d ” | Bob Setting 1 “ n ” | |
|---|---|---|---|---|
| Alice Setting 1 | “ d ” | 1 + a1 + b1 + z11 | 1 + a1 – b1 – z11 | 2 + 2 a1 |
| | “ n ” | 1 – a1 + b1 – z11 | 1 – a1 – b1 + z11 | 2 – 2 a1 |
| | | 2 + 2 b1 | 2 – 2 b1 | **4** |
| Alice Setting 2 | “ d ” | 1 + a2 + b1 + z21 | 1 + a2 – b1 – z21 | 2 + 2 a2 |
| | “ n ” | 1 – a2 + b1 – z21 | 1 – a2 – b1 + z21 | 2 – 2 a2 |
| | | 2 + 2 b1 | 2 – 2 b1 | **4** |

| Bob Setting 2 “ d ” | Bob Setting 2 “ n ” | |
|---|---|---|
| 1 + a1 + b2 + z12 | 1 + a1 – b2 – z12 | 2 + 2 a1 |
| 1 – a1 + b2 – z12 | 1 – a1 – b2 + z12 | 2 – 2 a1 |
| 2 + 2 b2 | 2 – 2 b2 | **4** |
| 1 + a2 + b2 + z22 | 1 + a2 – b2 – z22 | 2 + 2 a2 |
| 1 – a2 + b2 – z22 | 1 – a2 – b2 + z22 | 2 – 2 a2 |
| 2 + 2 b2 | 2 – 2 b2 | **4** |

## Estimation, standard errors, p-values

Routine M.L.E. (Sir R.A. Fisher 1921…)

Log Lik = $N$(“dd”|11)log(1 + $a_1$ + $b_1$ + $z_{11}$) + [15 more terms]

Parameters: $a_1$ $a_2$ $b_1$ $b_2$ $z_{11}$ $z_{12}$ $z_{21}$ $z_{22}$

Get m.l.e. of $z_{11} + z_{21} + z_{12} - z_{22}$

Get estimated standard error of estimated $z_{11} + z_{21} + z_{12} - z_{22}$ from inverse negative Fisher information matrix

Asymptotically optimal

[Linear constraints?]

$4\,\rho_{11} = (2 + 2\,z_{11}) - (2 - 2\,z_{11}) = 4\,z_{11}$

$4\,S = 4$ CHSH $= 4\,(z_{11} + z_{12} + z_{21} - z_{22})$

$$
\begin{aligned}
4\,J &= (1 + a_1 + b_1 + z_{11}) \\
&\quad - (1 - a_2 + b_1 - z_{21}) \\
&\quad - (1 + a_1 - b_2 - z_{12}) \\
&\quad - (1 + a_2 + b_2 + z_{22}) \\
&= -2 + (z_{11} + z_{21} + z_{12} - z_{22})
\end{aligned}
$$

$S = z_{11} + z_{12} + z_{21} - z_{22} = 2 + 4\,J$

$J = (S - 2) / 4$

**Modern approach:**
**algebraic geometry, computer algebra**

**Poor man’s solution:**
**two stage, generalised, least squares**
**Asymptotically just as good as MLE!**

**Also possible: amusing hybrid solutions**
***Also* asymptotically optimal**

# Theory, next 6 slides:

A standard Bell-type experiment with

- two parties,
- two measurement settings per party,
- two possible outcomes per measurement setting per party,

generates a vector of $16 = 4 \times 4$ numbers of outcome combinations per setting combination.

This can be applied to the two-channel experiments with no "no-shows", and to the one-channel experiments, and to the two-channel experiments with "–" and "no-show" combined

The four sets of four counts can be thought of as four observations each of a multinomially distributed vector over four categories.

Write $X_{ij}$ for the number of times outcome combination $j$ was observed, when setting combination $i$ was in force.

Let $n_i$ be the total number of trials with the $i$th setting combination.

The four random vectors $\vec{X}_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4})$, $i = 1, 2, 3, 4$, are independent each with a Multinomial$(n_i; \vec{p}_i)$ distribution, where $\vec{p}_i = (p_{i1}, p_{i2}, p_{i3}, p_{i4})$.

The 16 probabilities $p_{ij}$ can be estimated by relative frequencies $\widehat{p}_{ij} = X_{ij}/n_i$ which have the following variances and covariances:

$$\mathrm{var}(\widehat{p}_{ij}) = p_{ij}(1 - p_{ij})/n_i,$$

$$\mathrm{cov}(\widehat{p}_{ij}, \widehat{p}_{ij'}) = -p_{ij}p_{ij'}/n_i \quad \text{for } j \neq j',$$

$$\mathrm{cov}(\widehat{p}_{ij}, \widehat{p}_{i'j'}) = 0 \quad \text{for } i \neq i'.$$

The variances and covariances can be arranged in a $16 \times 16$ block diagonal matrix $\Sigma$ of four $4 \times 4$ diagonal blocks of non-zero elements.

Arrange the 16 estimated probabilities and their true values correspondingly in (column) vectors of length 16.

I will denote these simply by $\widehat{p}$ and $p$ respectively.

We have $\mathrm{E}(\widehat{p}) = p \in \mathbb{R}^{16}$ and $\mathrm{cov}(\widehat{p}) = \Sigma \in \mathbb{R}^{16\times 16}$.

We are interested in the value of one particular linear combination of the $p_{ij}$, let us denote it by $\theta = a^\top p$.

We know that four other particular linear combinations are identically equal to zero: the so-called no-signalling conditions.

This can be expressed as $B^\top p = 0$ where the $16 \times 4$ matrix $B$ contains, as its four columns, the coefficients of the four linear combinations.

We can sensibly estimate $\theta$ by $\widehat{\theta} = a^\top \widehat{p} - c^\top B^\top \widehat{p}$ where $c$ is any vector of dimension 4. For whatever choice we make, $\mathrm{E}\widehat{\theta} = \theta$.

We propose to choose $c$ so as to minimise the variance of the estimator. This minimization problem is a well-known problem from statistics and linear algebra ("least squares").

Define

$$\mathrm{var}(a^\top \widehat{p}) = a^\top \Sigma a = \Sigma_{aa},$$

$$\mathrm{cov}(a^\top \widehat{p}, B^\top \widehat{p}) = a^\top \Sigma B = \Sigma_{aB},$$

$$\mathrm{var}(B^\top \widehat{p}) = B^\top \Sigma B = \Sigma_{BB};$$

then the optimal choice for $c$ is

$$c_{\mathrm{opt}} = \Sigma_{aB} \Sigma_{BB}^{-1}$$

leading to the optimal variance

$$\Sigma_{aa} - \Sigma_{aB} \Sigma_{BB}^{-1} \Sigma_{Ba}.$$

In the experimental situation we do not know $p$ in advance, hence also do not know $\Sigma$ in advance. However we can estimate it in the obvious way ("plug-in") and for $n_i \to \infty$ we will have, just as in the previous section, an asymptotic normal distribution for our "approximately best" Bell inequality estimate, with an asymptotic variance which can be estimated by natural "plug-in" procedure, leading again to asymptotic confidence intervals, estimated standard errors, and so on.

The asymptotic width of this confidence interval is the smallest possible and correspondingly the number of standard errors deviation from "local realism" the largest possible.

The fact that $c$ is not known in advance does not harm these results.

The methodology is called "generalized least squares". It would be nice to use these estimates as the first step of a one-step Newton-Raphson iteration, and subsequently compute the Wilk's generalised log likelihood test, evaluated through its asymptotic chi-square distribution. This typically gives better approximation but is of course asymptotically equivalent

# Results

**https://rpubs.com/gill1109/OptimizedVienna**

**https://rpubs.com/gill1109/OptimizedNIST**

**https://rpubs.com/gill1109/OptimizedMunich**

**https://rpubs.com/gill1109/OptimizedDelft**

# OptimizedVienna.R

richard

2019-10-16

```r
## Comparison of CHSH and J for recent Bell experiments,
## together with optimally noise-reduced versions of both.
## Theory: https://pub.math.leidenuniv.nl/~gillrd/Peking/
Peking_4.pdf

## In short: assume four multinomial samples,
## estimate covariance matrix of estimated relative
frequencies,
## use sample deviations from no-signalling to optimally reduce
## the noise in the estimate of Bell's S or Eberhard's J

## AKA: generalized least squares

############# VIENNA #############

## The basic data, four 2x2 tables

table11 <- matrix(c(141439,   73391,   76224,   875392736),
    2, 2, byrow = TRUE,
    dimnames = list(Alice = c("d", "n"), Bob = c("d", "n")))
table12 <- matrix(c(146831,   67941,   326768,   874976534),
    2, 2, byrow = TRUE,
    dimnames = list(Alice = c("d", "n"), Bob = c("d", "n")))
table21 <- matrix(c(158338,   425067,   58742,   875239860),
    2, 2, byrow = TRUE,
    dimnames = list(Alice = c("d", "n"), Bob = c("d", "n")))
table22 <- matrix(c(8392,   576445,   463985,   874651457),
    2, 2, byrow = TRUE,
    dimnames = list(Alice = c("d", "n"), Bob = c("d", "n")))
```

```r
table11
##        Bob
## Alice       d         n
##      d 141439     73391
##      n  76224 875392736
table12
##        Bob
## Alice       d         n
##      d 146831     67941
##      n 326768 874976534
table21
##        Bob
## Alice       d         n
##      d 158338    425067
##      n  58742 875239860
table22
##        Bob
## Alice       d         n
##      d   8392    576445
##      n 463985 874651457
## Check of the total number of trials

# "The number of valid trials is N = 3 502 784 150"
sum(table11) + sum(table12) + sum(table21) + sum(table22)
## [1] 3502784150
```

```r
## The same data now in one 4x4 table

tables <- cbind(as.vector(t(table11)), as.vector(t(table12)),
as.vector(t(table21)), as.vector(t(table22)))
dimnames(tables) = list(outcomes = c("++", "+-", "-+", "--"),
                      settings = c(11, 12, 21, 22))
tables
##          settings
## outcomes        11        12        21        22
##       ++    141439    146831    158338      8392
##       +-     73391     67941    425067    576445
##       -+     76224    326768     58742    463985
##       -- 875392736 874976534 875239860 874651457
## The total number of trials for each setting pair

Ns <- apply(tables, 2, sum)
Ns
##        11        12        21        22
## 875683790 875518074 875882007 875700279
## observed relative frequencies, one 4x4 matrix

rawProbsMat <- tables / outer(rep(1,4), Ns)
rawProbsMat
##          settings
## outcomes           11           12           21           22
##       ++ 1.615183e-04 1.677076e-04 1.807755e-04 9.583188e-06
##       +- 8.380993e-05 7.760091e-05 4.853017e-04 6.582675e-04
##       -+ 8.704512e-05 3.732282e-04 6.706611e-05 5.298445e-04
##       -- 9.996676e-01 9.993815e-01 9.992669e-01 9.988023e-01
```

```r
## Convert the relative frequencies to one vector of length 16

VecNames <- as.vector(t(outer(colnames(rawProbsMat),
rownames(rawProbsMat), paste, sep = "")))
rawProbsVec <- as.vector(rawProbsMat)
names(rawProbsVec) <- VecNames

VecNames
##  [1] "11++" "11+-" "11-+" "11--" "12++" "12+-" "12-+"
"12--" "21++" "21+-"
## [11] "21-+" "21--" "22++" "22+-" "22-+" "22--"
rawProbsVec
##         11++         11+-         11-+         11--
12++
## 1.615183e-04 8.380993e-05 8.704512e-05 9.996676e-01
1.677076e-04
##         12+-         12-+         12--         21++
21+-
## 7.760091e-05 3.732282e-04 9.993815e-01 1.807755e-04
4.853017e-04
##         21-+         21--         22++         22+-
22-+
## 6.706611e-05 9.992669e-01 9.583188e-06 6.582675e-04
5.298445e-04
##         22--
## 9.988023e-01
```

```r
## Building up the 4 no-signalling constraints, combined in
one 16 x 4 matrix "NS"

Aplus <- c(1, 1, 0, 0)
Aminus <- - Aplus
Bplus <- c(1, 0, 1, 0)
Bminus <- - Bplus
zero <- c(0, 0, 0, 0)
NSa1 <- c(Aplus, Aminus, zero, zero)
NSa2 <- c(zero, zero, Aplus, Aminus)
NSb1 <- c(Bplus, zero, Bminus, zero)
NSb2 <- c(zero, Bplus, zero, Bminus)
NS <- cbind(NSa1 = NSa1, NSa2 = NSa2, NSb1 = NSb1, NSb2 =
NSb2)
rownames(NS) <- VecNames
NS
##       NSa1 NSa2 NSb1 NSb2
## 11++     1    0    1    0
## 11+-     1    0    0    0
## 11-+     0    0    1    0
## 11--     0    0    0    0
## 12++    -1    0    0    1
## 12+-    -1    0    0    0
## 12-+     0    0    0    1
## 12--     0    0    0    0
## 21++     0    1   -1    0
## 21+-     0    1    0    0
## 21-+     0    0   -1    0
## 21--     0    0    0    0
## 22++     0   -1    0   -1
## 22+-     0   -1    0    0
## 22-+     0    0    0   -1
## 22--     0    0    0    0
```

```r
## Build the 16x16 estimated covariance matrix of the 16 observed
relative frequencies

cov11 <- diag(rawProbsMat[ , "11"]) - outer(rawProbsMat[ , "11"],
rawProbsMat[ , "11"])
cov12 <- diag(rawProbsMat[ , "12"]) - outer(rawProbsMat[ , "12"],
rawProbsMat[ , "12"])
cov21 <- diag(rawProbsMat[ , "21"]) - outer(rawProbsMat[ , "21"],
rawProbsMat[ , "21"])
cov22 <- diag(rawProbsMat[ , "22"]) - outer(rawProbsMat[ , "22"],
rawProbsMat[ , "22"])

Cov <- matrix(0, 16, 16)
rownames(Cov) <- VecNames
colnames(Cov) <- VecNames
Cov[1:4, 1:4] <- cov11/Ns["11"]
Cov[5:8, 5:8] <- cov12/Ns["12"]
Cov[9:12, 9:12] <- cov21/Ns["21"]
Cov[13:16, 13:16] <- cov22/Ns["22"]

## The vector "S" is used to compute the CHSH statistic "CHSH"
## The sum of the first three sample correlations minus the fourth

## Note: the experiment is designed to favour use of Eberhard's J !

S <- c(c(1, -1, -1 ,1), c(1, -1, -1 , 1), c(1, -1, -1, 1), - c(1,
-1, -1, 1))
names(S) <- VecNames
CHSH <- sum(S * rawProbsVec)
CHSH
## [1] 2.000028
```

```r
## Compute the estimated variance of the CHSH statistic,
## its estimated covariances with the observed deviations
from no-signalling,
## and the 4x4 estimated covariance matrix of those
deviations.
## We'll later also need the inverse of the latter.

varS <- t(S) %*% Cov %*% S
covNN <- t(NS) %*% Cov %*% NS
covSN <- t(S) %*% Cov %*% NS
covNS <- t(covSN)

InvCovNN <- solve(covNN)

## Estimated variance of the CHSH statistic,
## and estimated variance of the optimally "noise reduced"
CHSH statistic.

varCHSH <- varS
varCHSHopt <- varS - covSN %*% InvCovNN %*% covNS

## The variance, and the improvement as ratio of standard
deviations

varS
##              [,1]
## [1,] 1.078084e-11
sqrt(varCHSH / varCHSHopt)
##          [,1]
## [1,] 2.055586
```

```r
## The coefficients of the noise reduced CHSH statistic and the
resulting improved estimate

Sopt <- S - covSN %*% InvCovNN %*% t(NS)
Sopt
##          11++      11+-      11-+ 11--     12++      12+-
12-+ 12--
## [1,] 2.011642 -0.5819319 -0.406426    1 2.302625 -1.418068
0.720693    1
##          21++      21+-      21-+ 21--      22++       22+-
22-+ 22--
## [1,] 2.188951 0.7825251 -1.593574    1 -4.503218 -0.7825251
-0.720693   -1
CHSHopt <- sum(Sopt * rawProbsVec)
CHSHopt
## [1] 2.000028
## p-values assuming approximate normality for testing CHSH
inequality

pnorm((CHSH - 2)/ sqrt(varCHSH), lower.tail = FALSE)
##              [,1]
## [1,] 5.43703e-18
pnorm((CHSHopt - 2)/ sqrt(varCHSHopt), lower.tail = FALSE)
##               [,1]
## [1,] 4.745794e-69
```

```
## Now we repeat for the Eberhard J statistic
## First, the coefficients in the vector "J"
## and the observed value of the statistic

J <- c(c(1, 0, 0 ,0), c(0, -1, 0 ,0), c(0, 0, -1, 0), c(-1,
0, 0, 0))
names(J) <- VecNames
sum(J * rawProbsVec)
## [1] 7.26814e-06
## Next, its estimated variance and resulting p-value

varJ <- t(J) %*% Cov %*% J
sum(J * rawProbsVec) / sqrt(varJ)
##          [,1]
## [1,] 12.10426
pnorm(sum(J * rawProbsVec) / sqrt(varJ), lower.tail = FALSE)
##              [,1]
## [1,] 5.013606e-34
## The covariances between J and the observed deviations
from no-signaling
## The variance of the usual estimate of J and of the
improved estimate of J
## The improvement as a ration of standard deviations

covJN <- t(J) %*% Cov %*% NS
covNJ <- t(covJN)
varJopt <- varJ - covJN %*% InvCovNN %*% covNJ

varJ
##              [,1]
## [1,] 3.605539e-13
sqrt(varJ / varJopt)
##          [,1]
## [1,] 1.503676

## The coefficients of an improved estimataor of Eberhard's J

Jopt <- J - covJN %*% InvCovNN %*% t(NS)
Jopt
##           11++      11+-        11-+ 11--       12++      12+-
12-+
## [1,] 0.2529105 -0.395483 -0.3516065    0 0.3256562 -0.604517
-0.06982674
##      12--      21++         21+-        21-+ 21--        22++
22+-
## [1,]    0 0.2972378 -0.05436871 -0.6483935    0 -0.8758045
0.05436871
##             22-+ 22--
## [1,] 0.06982674    0
## Observed estimate of J, and improved estimate of J

sum(J * rawProbsVec)
## [1] 7.26814e-06
sum(Jopt * rawProbsVec)
## [1] 6.997615e-06
## p-values based on J and on improved J
## Note that the p-value based on improved J is the same as that
of improved CHSH

pnorm(sum(J * rawProbsVec) / sqrt(varJ), lower.tail = FALSE)
##              [,1]
## [1,] 5.013606e-34
pnorm(sum(Jopt * rawProbsVec) / sqrt(varJopt), lower.tail =
FALSE)
##              [,1]
## [1,] 4.745794e-69
## The p-value of the optimized J got a lot better, even though
the estimate got a bit smaller
## The optimization procedure for CHSH made an enormous
difference
## The deviation from no-signalling is small; it is responsible
for these small changes
```

**https://rpubs.com/gill1109/OptimizedVienna**

**https://rpubs.com/gill1109/OptimizedNIST**

**https://rpubs.com/gill1109/OptimizedMunich**

**https://rpubs.com/gill1109/OptimizedDelft**

# Maybe you prefer a Bayesian approach…

PHYSICAL REVIEW A **99**, 022112 (2019)

**Very strong evidence in favor of quantum mechanics and against local hidden variables from a Bayesian analysis**

Yanwu Gu,[1,2,*] Weijun Li,[2,†] Michael Evans,[3,‡] and Berthold-Georg Englert[2,1,4,§]
[1]*Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542, Singapore*
[2]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore*
[3]*Department of Statistical Sciences, University of Toronto, Toronto, Ontario M5S 3G3, Canada*
[4]*MajuLab, CNRS-UCA-SU-NUS-NTU International Joint Unit, Singapore*

(Received 31 August 2018; published 13 February 2019)

The data of four recent experiments—conducted in Delft, Vienna, Boulder, and Munich with the aim of refuting nonquantum hidden-variables alternatives to the quantum-mechanical description—are evaluated from a Bayesian perspective of what constitutes evidence in statistical data. We find that each of the experiments provides strong, or very strong, evidence in favor of quantum mechanics and against the nonquantum alternatives. This Bayesian analysis supplements the previous non-Bayesian ones, which refuted the alternatives on the basis of small *p* values but could not support quantum mechanics.

DOI: 10.1103/PhysRevA.99.022112

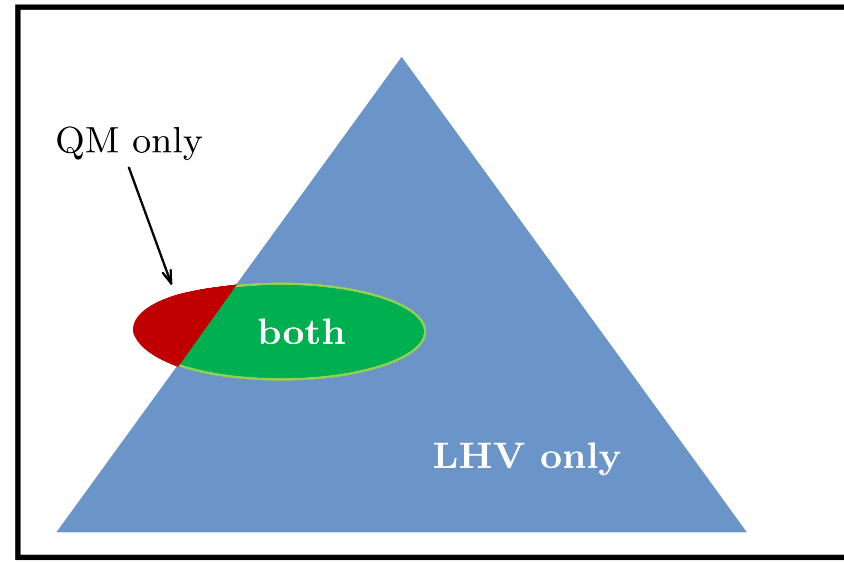



FIG. 1. Symbolic sketch of the eight-dimensional set of permissible probabilities. The points inside the ellipse symbolize probabilities accessible by quantum mechanics (QM); the triangle encloses the probabilities permitted by local hidden variables (LHV). There are no QM probabilities in the blue portion of the LHV set and no LHV probabilities in the red part of the QM set. The green overlap region contains the probabilities that are possible both for QM and LHV.

[22] We trust that runs tests have confirmed this statistical independence. We cannot perform run tests ourselves on all data sets as we do not know the sequence of detection events for some of them. For all experiments evaluated, we only explore the total counts of recorded detection events (Tables II, VII, X, and XIV), not the order in which they were recorded. The list of counts is a sufficient statistic if the events are statistically independent, and only then.

[43] A remark similar to that in Ref. [22] applies: We cannot test this assumption on the basis of the data in Table X as one would need to compare the data recorded at different intervals of the data-collection period. We trust that the data have been tested for drifts in the experimental parameters and none were found within each of the three datasets.

**Gu et al. assumption:**
**“QM” = “QM of two qubits”, *arbitrary* density matrix,**
***perfect pre-specified* spin measurements**

# Postscript

- What should we believe now?
- Should we trust the data?
- According to Bednorz, Adenier and Khrennikov, Graft, Santos, Hnilo, Fodje, … no.
- I think that the observed anomalies are not important and are anyway irrelevant if we use martingale tests, possibly adding room for imperfect random number generation
- Let's trust the experiments … what then?

# Erwin Schrödinger

- **I don't like it, and I'm sorry I ever had anything to do with it.**

  [About the probability interpretation of quantum mechanics.] Epigraph, without citation, in John Gribbin, In Search of Schrödinger's Cat: Quantum Physics and Reality (1984), v, frontispiece.

- **If all this damned quantum jumping were really here to stay, I should be sorry, I should be sorry I ever got involved with quantum theory.**

  As reported by Heisenberg describing Schrödinger's time spent debating with Bohr in Copenhagen (Sep 1926). In Werner Heisenberg, Physics and Beyond: Encounters and Conversations (1971), 75. As cited in John Gribbin, Erwin Schrodinger and the Quantum Revolution.

- **God knows I am no friend of probability theory, I have hated it from the first moment when our dear friend Max Born gave it birth. For it could be seen how easy and simple it made everything, in principle, everything ironed and the true problems concealed. Everybody must jump on the bandwagon [Ausweg]. And actually not a year passed before it became an official credo, and it still is.**

  Letter to Albert Einstein (13 June 1946), as quoted by Walter Moore in Schrödinger: Life and Thought (1989) ISBN 0521437679

# The experiments of 2015 convinced me … <u>*rebrand*</u> "spooky action at a distance" …

- Entanglement is an asset, not a horror
- We call it "spooky" because our mammal brains, trained by evolution, can't "understand" it any way except as the work of a *potentially* malevolent *agent*
- "Spooky" is an inadequate translation of "<u>spukhaft</u>". We have to say it in German.
- "<u>Passion</u> at a distance" is better
- More precise: "(Martingale like) <u>disciplined</u> passion at a distance"? No, it won't catch on …

- **Auserlesene / engelhafte 'spukhafte Fernwirkung'** (exquisite / angelic "action at a distance")

# … and … Belavkin's "eventum mechanics" is the way to go.

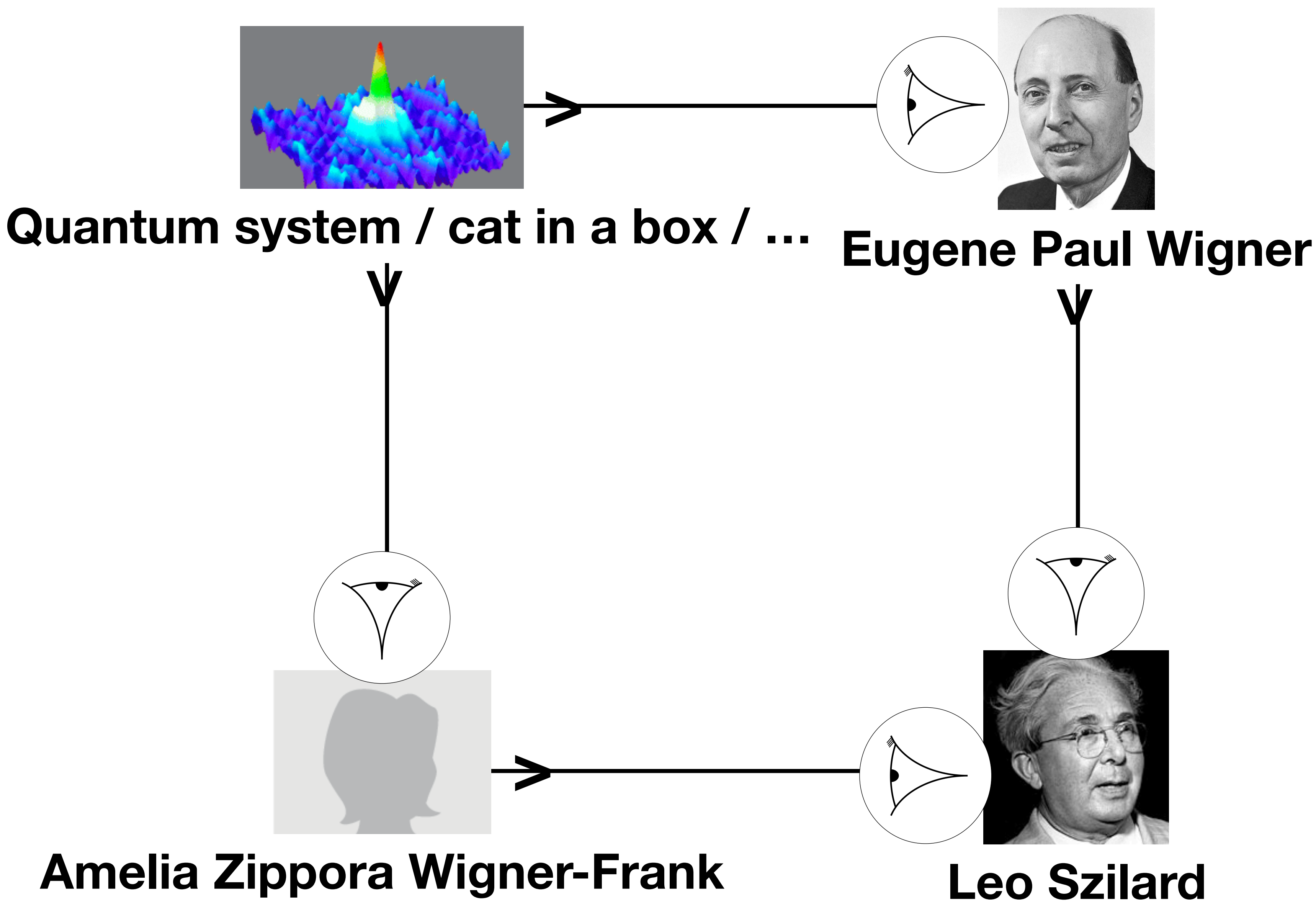
- It's a "collapse theory"
- It is therefore "non-local"
- It can be made Lorentz invariant!
- Some famous recent works confirm me in my opinions:
  - Daniela **Frauchiger** & Renato **Renner**

    [My title] *Schrödinger's cat, the Wigners, and the Wigners' friend*

  - Gilles **Brassard** & Paul **Raymond-Robichaud**

    "*The equivalence of local-realistic and no-signalling theories".* Abstract: We provide a framework to describe all local-realistic theories and all no-signalling operational theories. We show that when the dynamics is reversible, these two concepts are equivalent. In particular, this implies that unitary quantum theory can be given a local-realistic model.

# The Wigners' friend



# My **<u>prejudice:</u>** The clicks are “real”, the rest … a construction of our minds

- It is allowed to imagine that more stuff is real
- Such a “dilation” need not be unique
- “QM without collapse”, or Unitary QM - several theories, best known being MW and Quantum Qubism
- MW is many words
- QB is subjective Bayes … but I’m a frequentist … usually Bayes and frequentist inference agree … it’s really interesting when they disagree !!!
- Quantum Buddhism gives yet further insights

# F&R: The Wigners' friend

- QM *without collapse* + MW  implies only the wave function is real

- QM *without collapse* + Qbism  implies nothing is real

- **<u>My conclusion: QM without collapse is non-sense!</u>**

# B&RR

- They insist on irreversibility!
- Change definitions of everything
- It’s brilliant but … it’s very technical and very long

- **<u>My conclusion: we must trash ‘irreversibility’!</u>**

- “Spukhafte fernwerkung” is for real and … Exquisite? Angelic?
- Collapse is real
- Recommendation: take a look again at Belavkin’s “Eventum Mechanics”
- We must keep questioning the very words which we use (Eastern thought / Western post-modernism) … and remember what we are … *nothing* is real - QBism!
- I think that *both* QBism and Eventum Mechanics are self-consistent; their empirical predictions are (so far) identical. So it’s a matter of taste. Non-local collapse masked by irreducible randomness … or quantum Buddhism (we only have our sensory impressions and our inter-subjective confidence in predictions of future sensory impressions)

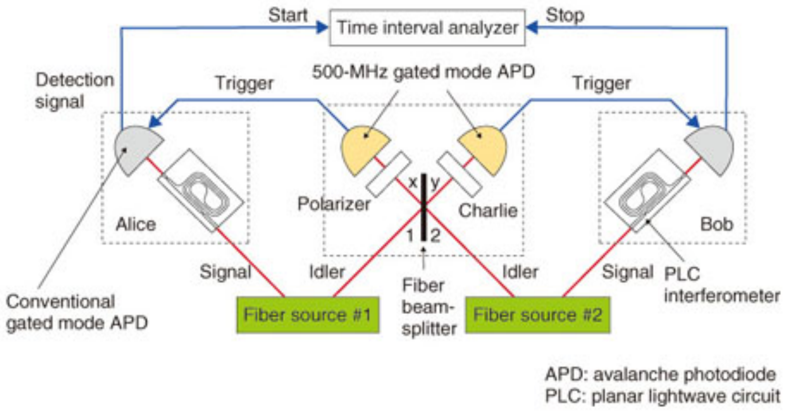
Everything is a construction of our minds - there is nothing else

Beware: every word is a “model”

All models are wrong, some are useful

Universiteit Leiden

COMBRAY




# Postscript

I cannot say that action at a distance is required in physics. But I cannot say that you can get away with no action at a distance. You cannot separate off what happens in one place with what happens at another – John Bell

https://www.youtube.com/watch?v=V8CCfOD1iu8

Nature produces chance events (irreducibly chance-like!) which can occur at widely removed spatial locations without anything propagating from point to point along any path joining those locations. … The chance-like character of these effects prevents any possibility of using this form of non locality to communicate, thereby saving from contradiction one of the fundamental principles of relativity theory according to which no communication can travel faster than the speed of light – Nicolas Gisin

Quantum Chance: Nonlocality, Teleportation and Other Quantum Marvels. Springer, 2014