Tutorial: Device-independent random number generation

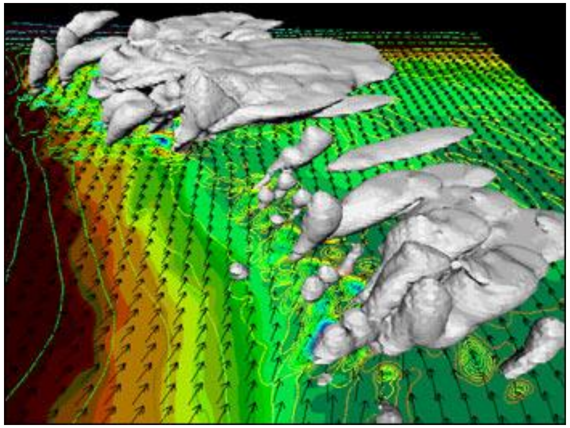
Roger Colbeck
University of York

Outline

- Brief motivation of random number generation
- Discuss what we mean by a random number
- Discuss some ways of generating them leading up to device-independent protocols
- Explain the main ideas behind a device-independent random number generator
- Discuss what it means for a protocol to be secure
- Briefly mention related tasks

Why are random numbers important?

gambling | simulations | cryptography



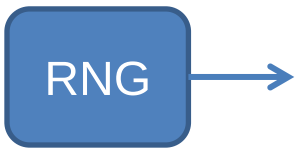
Random number generation

UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME

QUANTUM COMMUNICATIONS HUB

EPSRC
Engineering and Physical Sciences Research Council

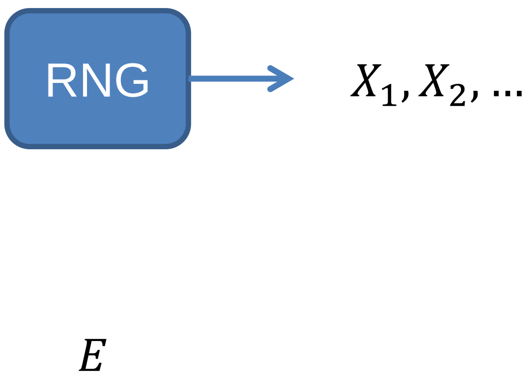
What is a random number?



$X_1, X_2, \ldots$

- Unpredictable by anyone (independent of everything else)
- Uniformly distributed

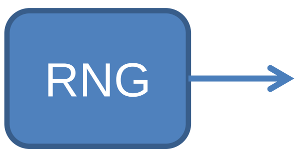
What is a random number?



- More formally, we can say

  X_j is a uniform random bit (with respect to E) if $P_{X_j|E} = P_{X_j} = \frac{1}{2}$ where E represents 'everything else' (includes $X_1, \ldots, X_{j-1}$)

What is a random number?



$X_1, X_2, \dots$

E

- Quantum case

$$\sum_x \frac{1}{|X|} |x\rangle\langle x|_A \otimes \rho_E$$

What do we want in a random number generation protocol?

- Secure
- Reliable
- Easy to implement
  - Technologically feasible
  - Requires few devices
- Have a fast rate

Security

- Protocol should come with a rigorous, precisely formulated security proof and statement of validity
  - E.g., if the protocol is used correctly, then no adversary can learn the random numbers even given unlimited time/resources (unless physics is wrong)

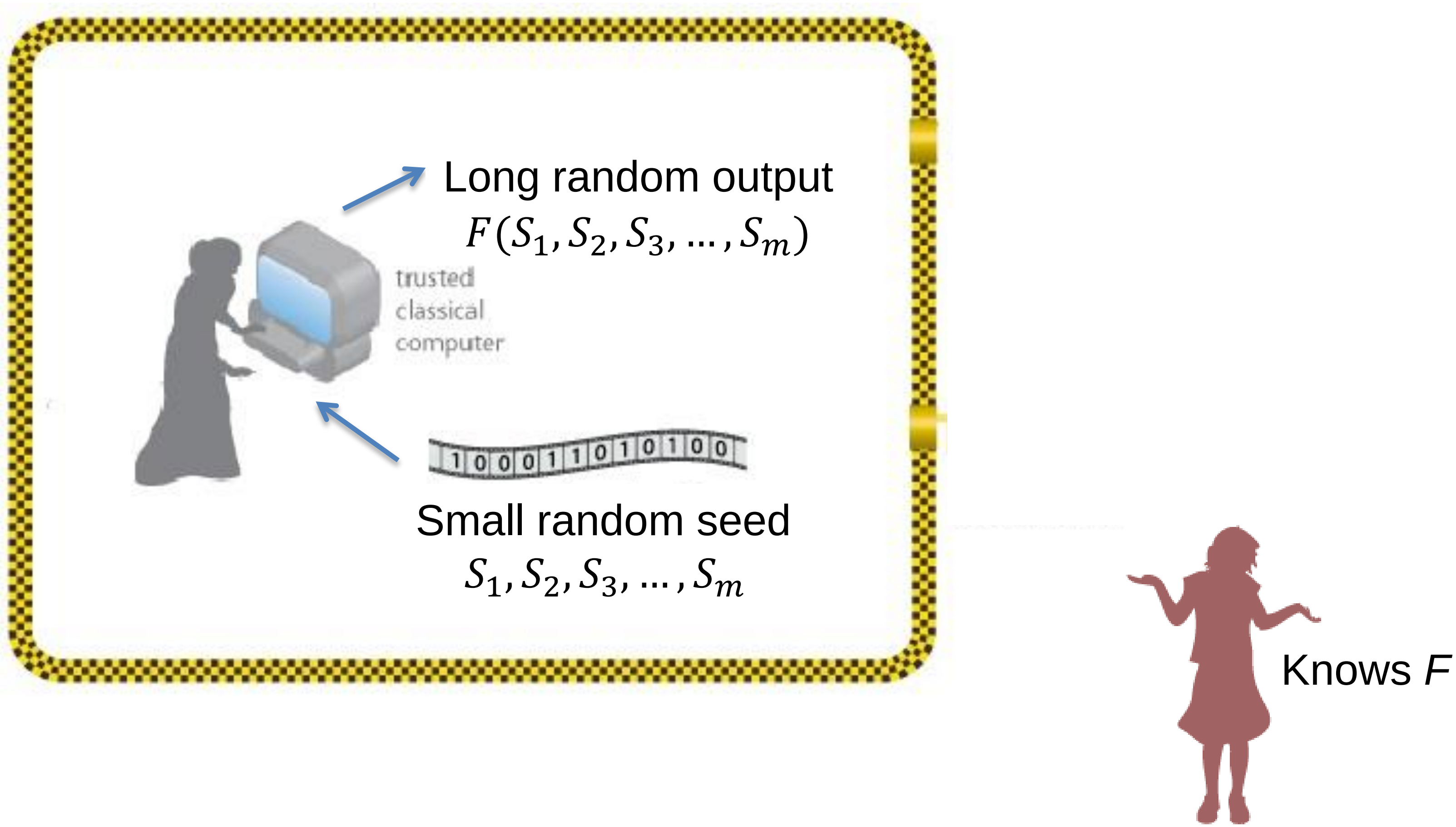
Security

- Protocol should come with a rigorous, precisely formulated security proof and statement of validity
  - E.g., if the adversary is limited to have particular computational resources, the random string can be treated as random for a certain amount of time.

How might we generate random numbers?

How might we generate random numbers? Classical case

How might we generate random numbers? Classical case

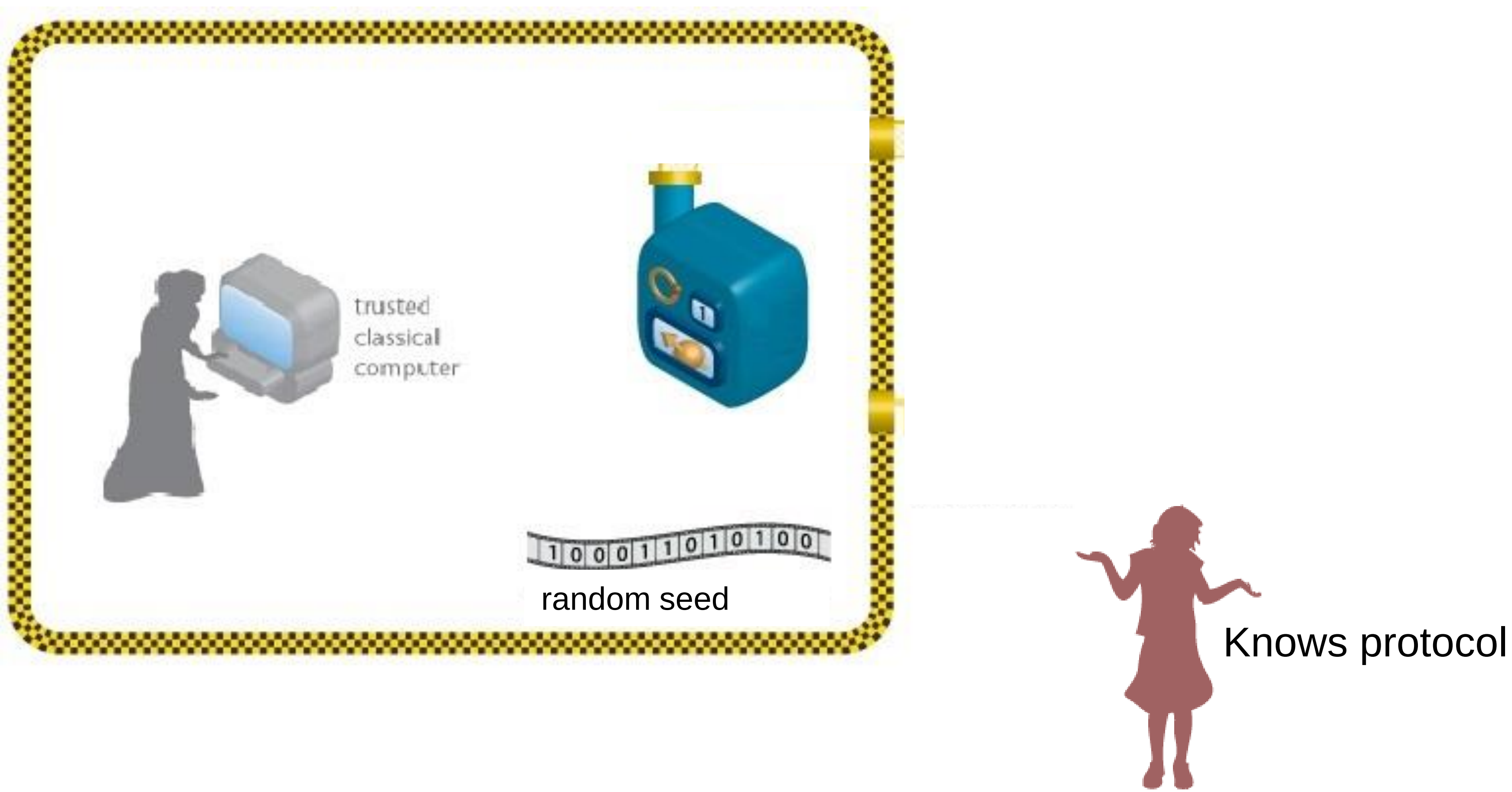


Classical case

Drawbacks:

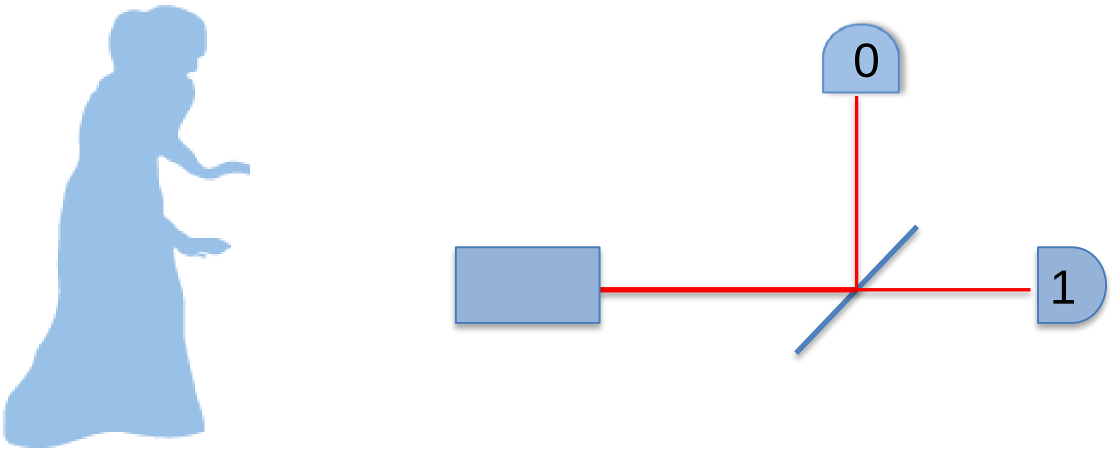
- Cannot have unconditional security
- In general, we cannot prove hardness of breaking the protocol

Trusted quantum case



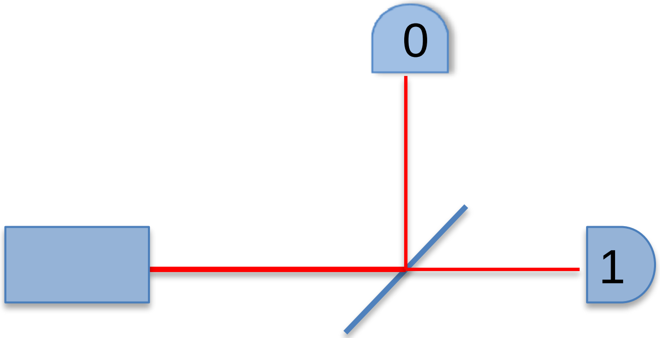
Trusted quantum case

For example: use a beamsplitter



Trusted quantum case

For example: use a beamsplitter

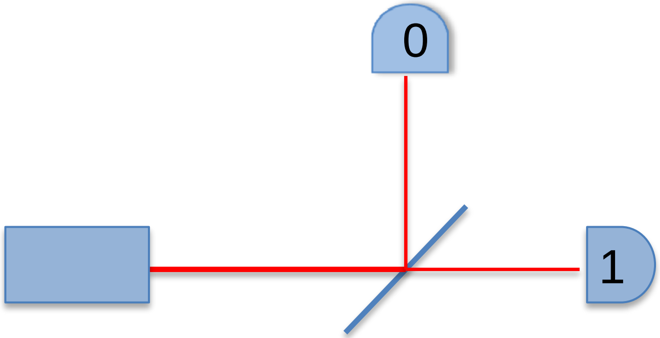


This might be ok if:

- Trust the equipment
- Ensure that it doesn't change over time

Trusted quantum case

For example: use a beamsplitter

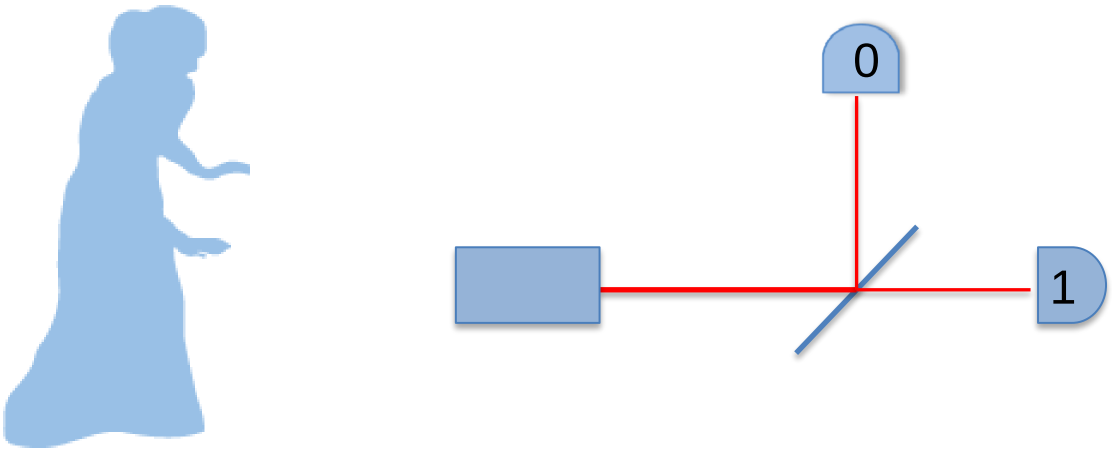


This might be ok if:

- Trust the equipment
- Ensure that it doesn't change over time
- (Trust the physics and that it is complete)

Trusted quantum case

For example: use a beamsplitter



Ideally we would like a certificate that outputs are random

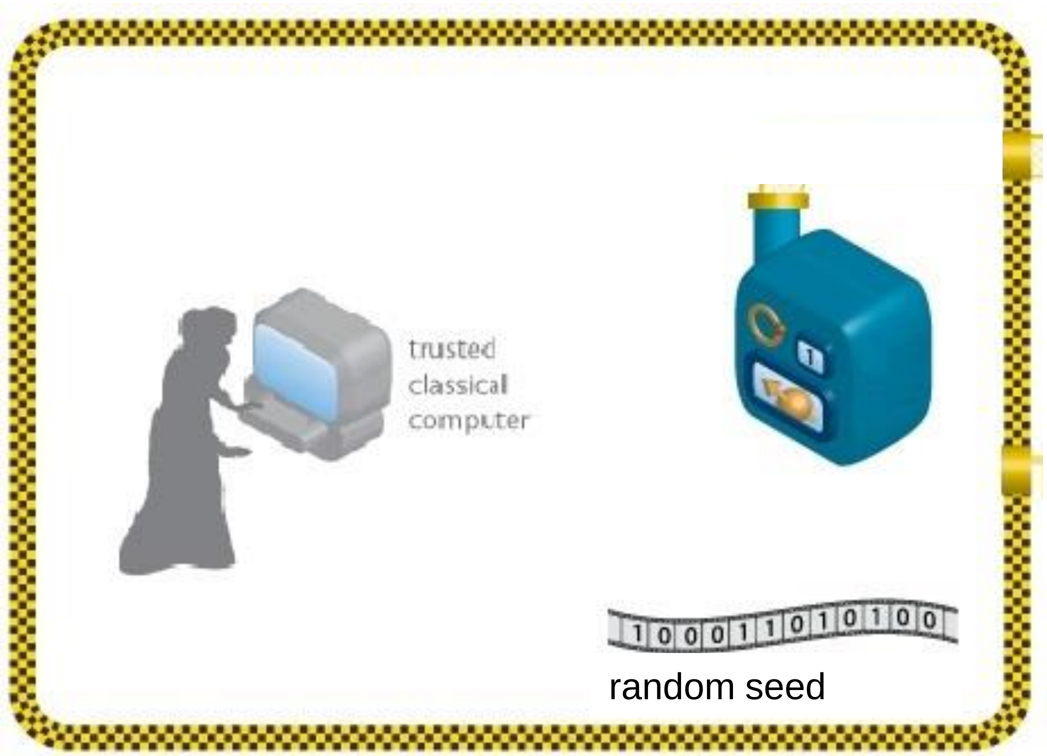
Trusted quantum case

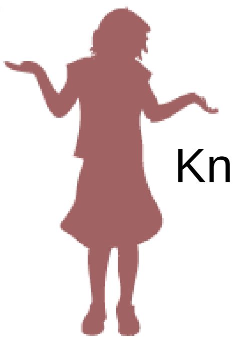
Removes classical drawbacks; in particular, can have security based on physics.

New drawbacks:

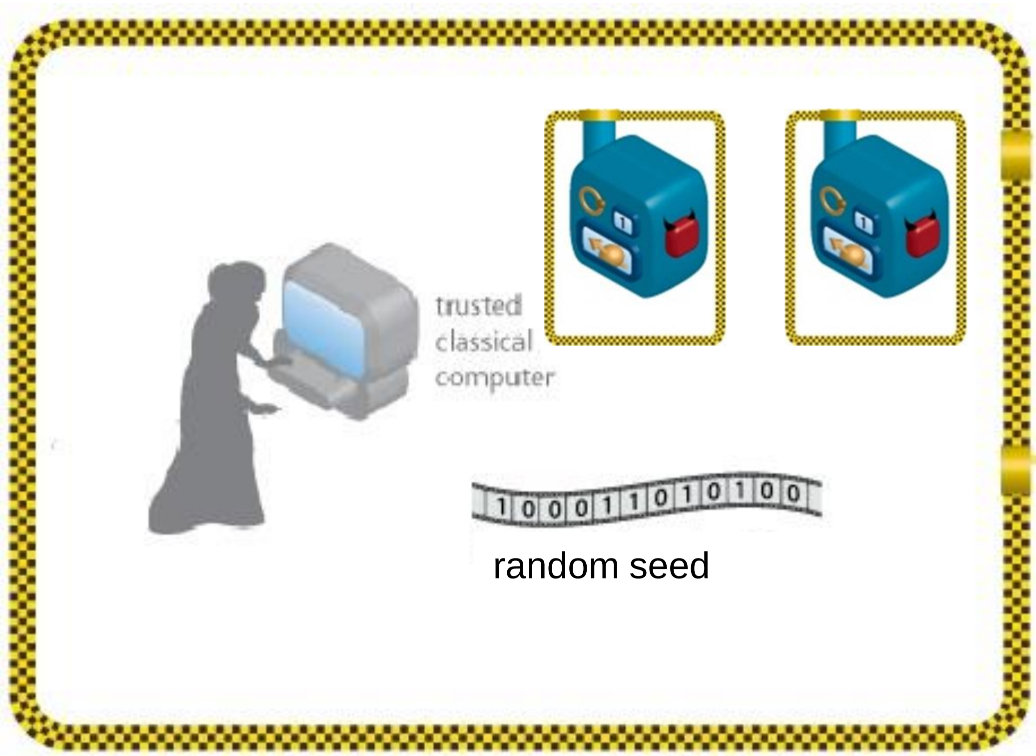
- Technologically harder to implement (but not too bad)
- Security relies on the devices behaving correctly

The setup (quantum)





The setup (device-independent)

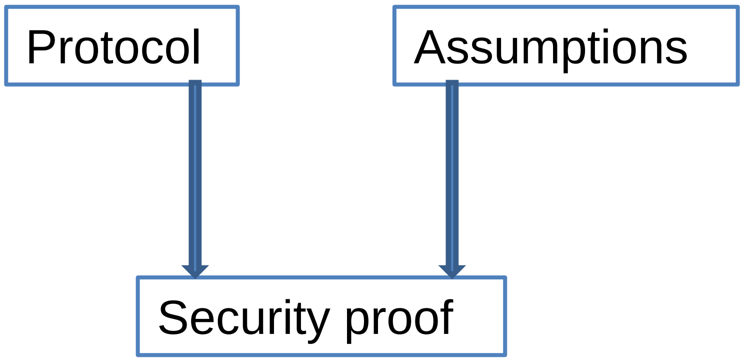


Want to generate longer random string

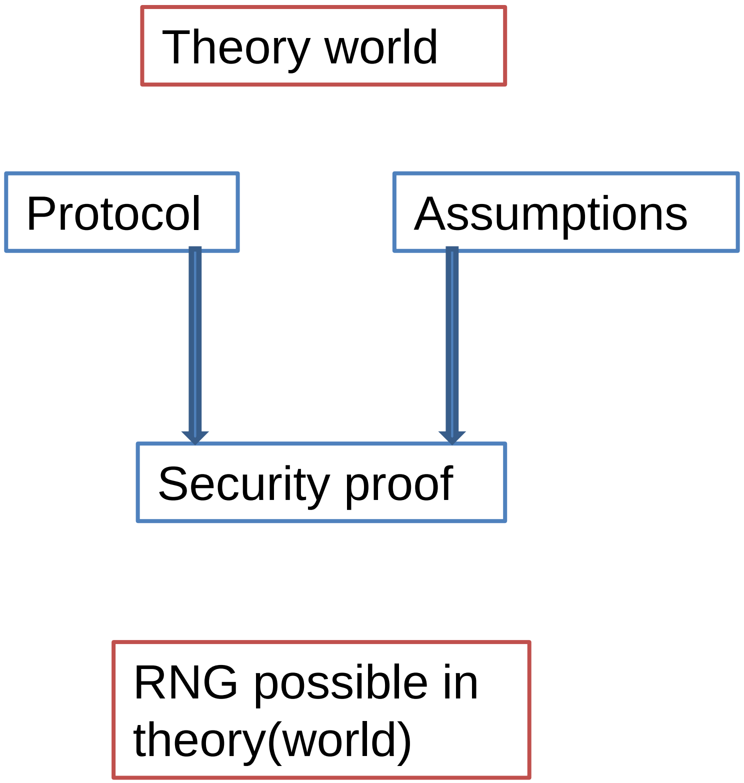
Device-independence

- No assumptions made about the workings of the devices used
- However, we do need some assumptions, in particular, both strong lab walls and initial randomness [necessary for cryptography]

Security proofs



Security proofs



QUANTUM COMMUNICATIONS HUB

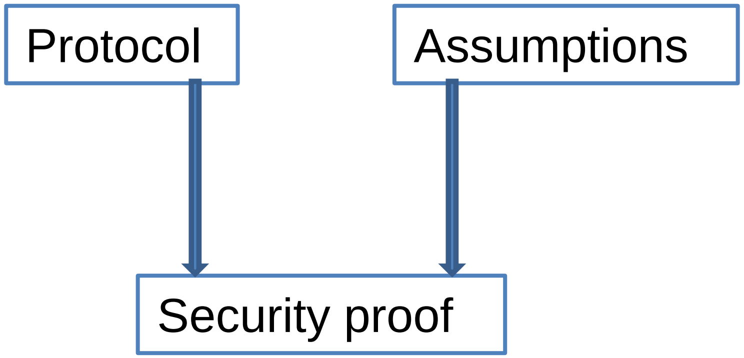
UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME

EPSRC Engineering and Physical Sciences Research Council

Security proofs

Theory world

Real world



Is our theory world proof relevant in the real world?

RNG possible in theory(world)

Security proofs

Weaker assumptions → More security

Security proofs

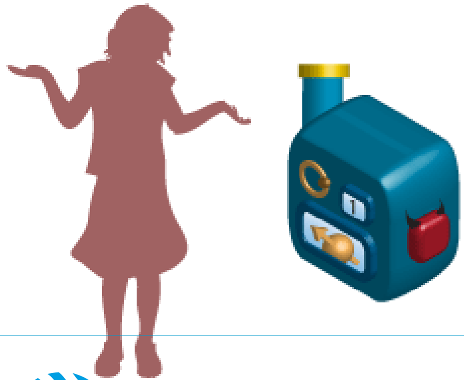
Weaker assumptions → More security

- Device-independence tries to remove all the assumptions on the devices
- Removes this mismatch problem between the real world and theory world

Security proofs

Weaker assumptions → More security

- No assumptions on devices means the security proof has to work even with maliciously constructed devices.



Security proofs

Weaker assumptions → More security

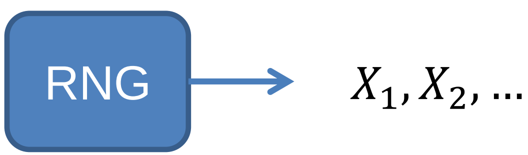
- Protocol remains secure if devices stop working properly or are tampered with

- Protocol checks the workings of the devices on-the-fly (hence, self-testing)

Device-independence: main ideas

- Don’t trust devices, so have to test them

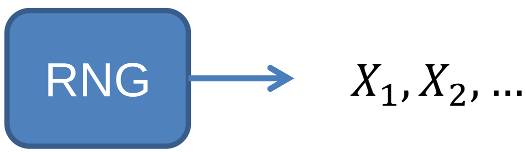
How can we test the devices?



How can we test for randomness?

- **Overlapping permutations:** Analyse sequences of five consecutive random numbers. The 120 possible orderings should occur with statistically equal probability.
- **Ranks of matrices:** Select some number of bits from some number of random numbers to form a matrix over {0,1}, then determine the rank of the matrix. Count the ranks.
- **Monkey tests:** Treat sequences of some number of bits as "words". Count the overlapping words in a stream. The number of "words" that don't appear should follow a known distribution.
- **The craps test:** Play 200,000 games of craps, counting the wins and the number of throws per game. Each count should follow a certain distribution.

How can we test for randomness?



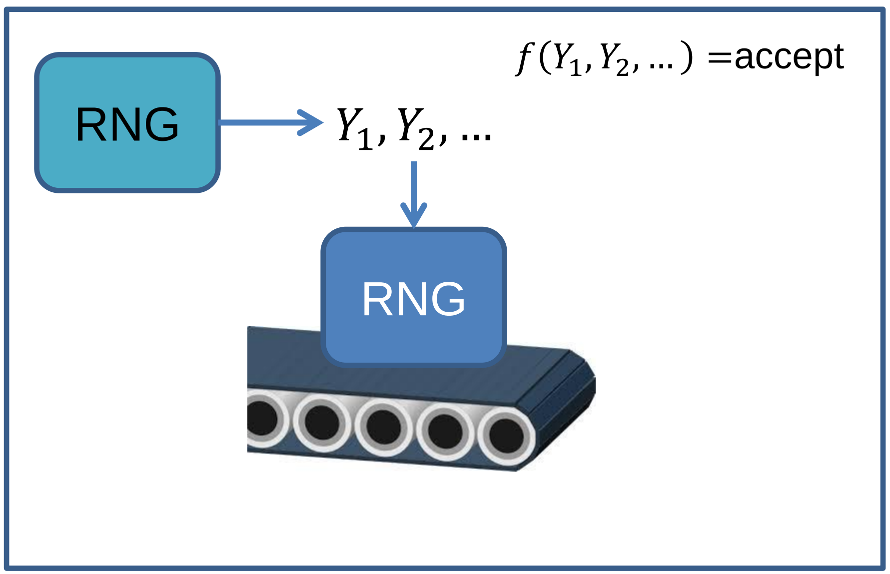
- There is no good test that acts only on the outputs.

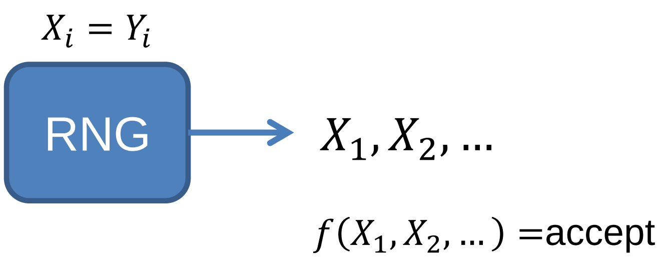
- No f such that

$$f(X_1, X_2, \ldots) = \begin{cases} \text{accept} \\ \text{reject} \end{cases}$$

with accept only if the sequence is random.

How can we test for randomness?





More advanced test

- There is no good test with only one device

$X_1, X_2, \ldots$

$A_1, A_2, \ldots$ (Random)

$$f(A_1, A_2, \ldots, X_1, X_2, \ldots) \in \{\text{pass, fail}\}$$

Adversary knows f
Adversary can supply pre-programmed classical device that will always pass

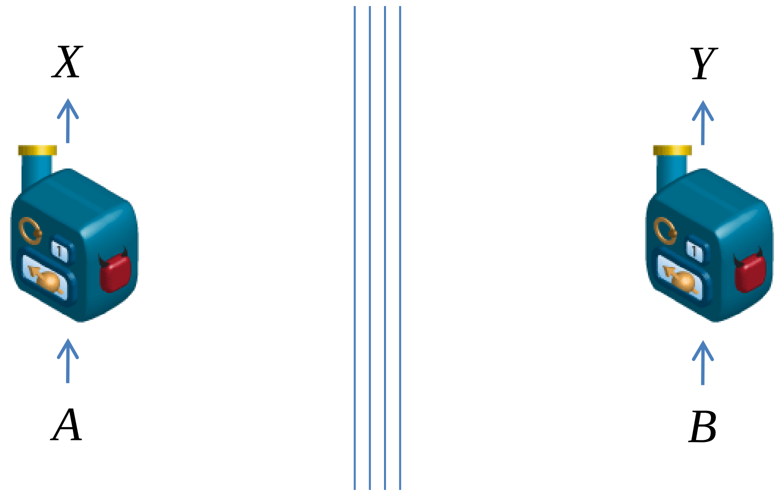
Device-independent randomness expansion: main ideas

Bell inequality violation → Non-classical behaviour

(loophole-free)

Device-independent randomness expansion: main ideas

- Bell-inequality violation



$P_{XY|AB}$ violates a Bell inequality
A and B random
Devices cannot communicate

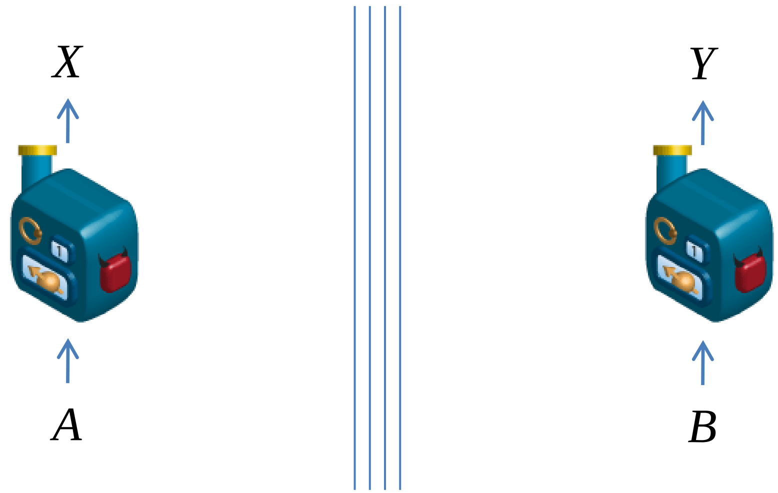
Bell's theorem

Eve cannot know X
X not function of A

Roughly the idea of Ekert 91, although note that we're not making key here

Device-independent randomness expansion: main ideas

- Bell-inequality violation



$P_{XY|AB}$ violates a Bell inequality
A and B random
Devices cannot communicate

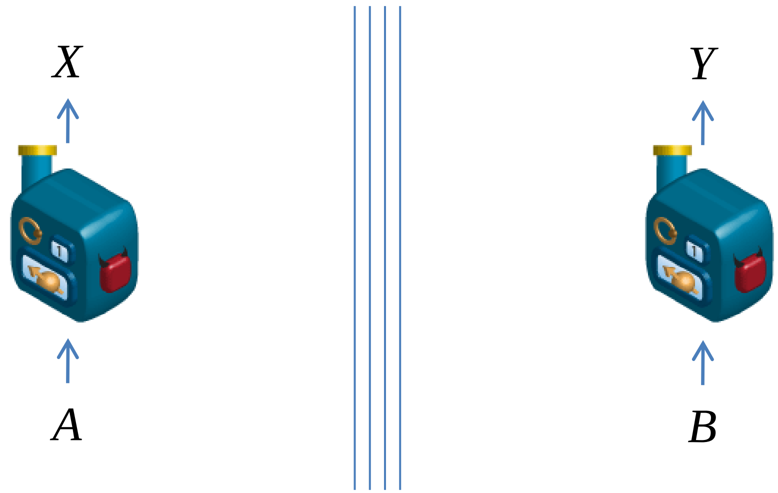
Bell's theorem

Eve cannot know X
X not function of A

- Doesn't mean that X is perfectly random

Device-independent randomness expansion: main ideas

- Bell-inequality violation



$P_{XY|AB}$ violates a Bell inequality
A and B random
Devices cannot communicate

↓ Bell's theorem

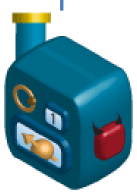
Eve cannot know X
X not function of A

- E.g. CHSH game winning probability

Device-independent randomness expansion: main ideas

- CHSH game

$X \in \{0,1\}$



$A \in \{0,2\}$

$Y \in \{0,1\}$

$B \in \{1,3\}$

Win if
$X = Y$ for $(\mathrm{A}, \mathrm{B}) = (0,1), (2,1)$ or $(2,3)$
$X \neq Y$ for $(A, B) = (0,3)$.

- $P_{cl} \leq \frac{3}{4}$ $\qquad$ $P_{qm} \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 0.85.$

(Bell value 2) $\qquad$ (Bell value $2\sqrt{2}$)

Device-independent randomness expansion: main ideas

$X \in \{0,1\}$

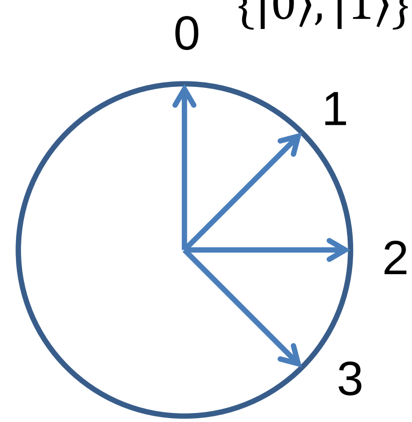
$A \in \{0,2\}$

$Y \in \{0,1\}$

$B \in \{1,3\}$

Win if
$X = Y$ for $(\mathrm{A}, \mathrm{B}) = (0,1), (2,1)$ or $(2,3)$
$X \neq Y$ for $(A, B) = (0,3)$.

$\{|0\rangle, |1\rangle\}$



$\{|+\rangle, |-\rangle\}$

- $P_{qm} \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 0.85$

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Device-independent randomness expansion: main ideas

Maximum quantum violation → Devices share max entangled (pure) state

↓

No entanglement with Eve

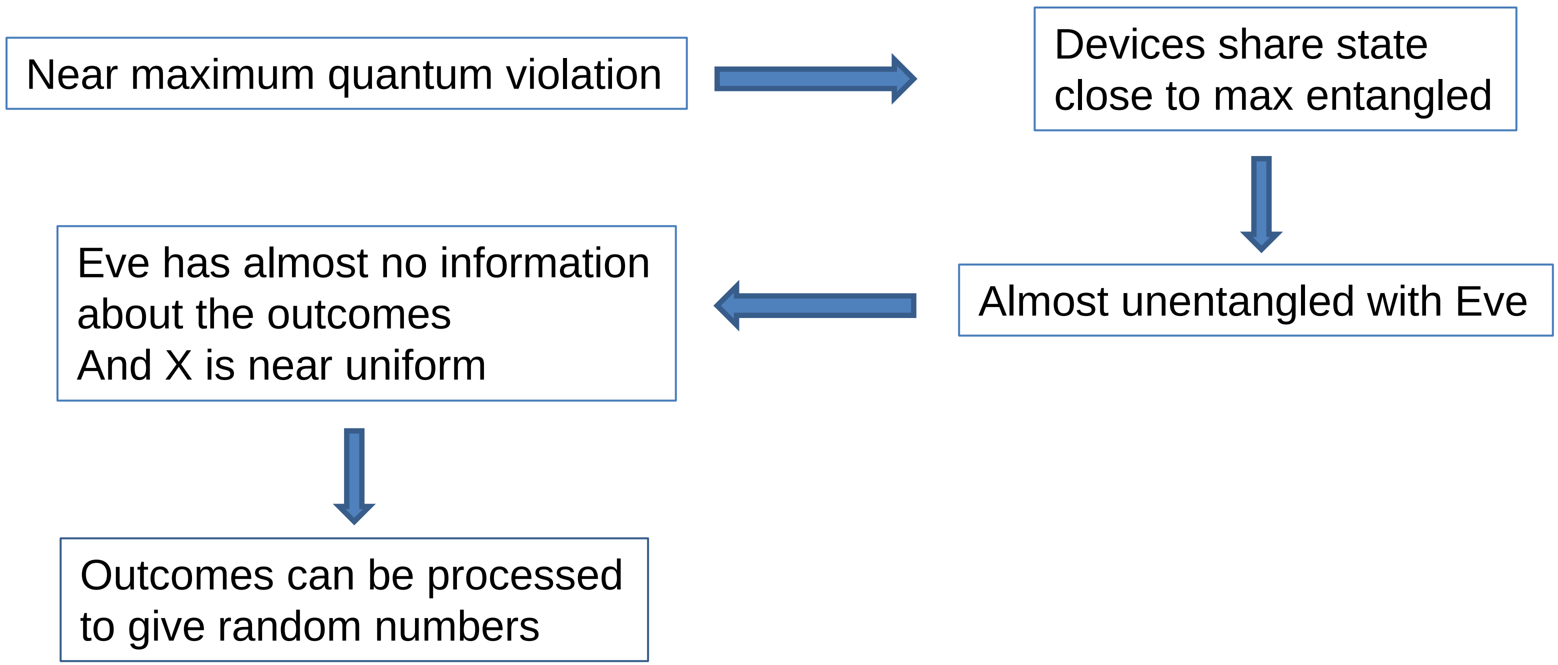
$$|\psi\rangle_{AB} \otimes |\phi\rangle_E$$

← Eve has no information about the outcomes
And X is uniform

↓

Outcomes can be used as random numbers

Device-independent randomness expansion: main ideas



Device-independent randomness expansion: main ideas

Near maximum quantum violation

↓

Eve has almost no information about the outcomes
And X is near uniform

↓

Outcomes can be processed to give random numbers

Connecting Bell violation with Eve's knowledge

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{2}-\varepsilon$	ε	ε	$\frac{1}{2}-\varepsilon$
	1	ε	$\frac{1}{2}-\varepsilon$	$\frac{1}{2}-\varepsilon$	ε
2	0	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$	ε
	1	ε	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$

$$P_{\text{win}} = 1 - 2\varepsilon$$

How much can Eve know about X?

Connecting Bell violation with Eve's knowledge

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{2}-\varepsilon$	ε	ε	$\frac{1}{2}-\varepsilon$
	1	ε	$\frac{1}{2}-\varepsilon$	$\frac{1}{2}-\varepsilon$	ε
2	0	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$	ε
	1	ε	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$

$$P_{\text{win}} = 1 - 2\varepsilon$$

How much can Eve know about X?

$$P_{XY|AB} = \sum_z p_z P_{XY|ABz}$$

Convex combination

Quantum-realizable distributions

Connecting Bell violation with Eve's knowledge

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{2}-\varepsilon$	ε	ε	$\frac{1}{2}-\varepsilon$
	1	ε	$\frac{1}{2}-\varepsilon$	$\frac{1}{2}-\varepsilon$	ε
2	0	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$	ε
	1	ε	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$

$$P_{\text{win}} = 1 - 2\varepsilon$$

How much can Eve know about X?

$$P_{XY|AB} = \sum_z p_z P_{XY|ABz}$$

Convex combination

Any non-signalling distribution

Connecting Bell violation with Eve's knowledge

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{2}-\varepsilon$	ε	ε	$\frac{1}{2}-\varepsilon$
	1	ε	$\frac{1}{2}-\varepsilon$	$\frac{1}{2}-\varepsilon$	ε
2	0	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$	ε
	1	ε	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$

$$P_{\text{win}} = 1 - 2\varepsilon$$

How much can Eve know about X?

$$P_{XY|AB} = \sum_z p_z P_{XY|ABz}$$

Convex combination

Any non-signalling distribution

$$P_{XY|AB} = (1-4\varepsilon)\begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + \varepsilon\left(\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \right)$$

Eve has no knowledge about X

Eve knows X perfectly

Connecting Bell violation with Eve's knowledge

$P_{XY\|AB}$		$B=1$, $Y=0$	$B=1$, $Y=1$	$B=3$, $Y=0$	$B=3$, $Y=1$
A	X				
0	0	$\frac{1}{2}-\varepsilon$	ε	ε	$\frac{1}{2}-\varepsilon$
	1	ε	$\frac{1}{2}-\varepsilon$	$\frac{1}{2}-\varepsilon$	ε
2	0	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$	ε
	1	ε	$\frac{1}{2}-\varepsilon$	ε	$\frac{1}{2}-\varepsilon$

$$P_{\text{win}} = 1 - 2\varepsilon$$

How much can Eve know about X?

$$P_{XY|AB} = \sum_z p_z P_{XY|ABz}$$

Convex combination

Any non-signalling distribution

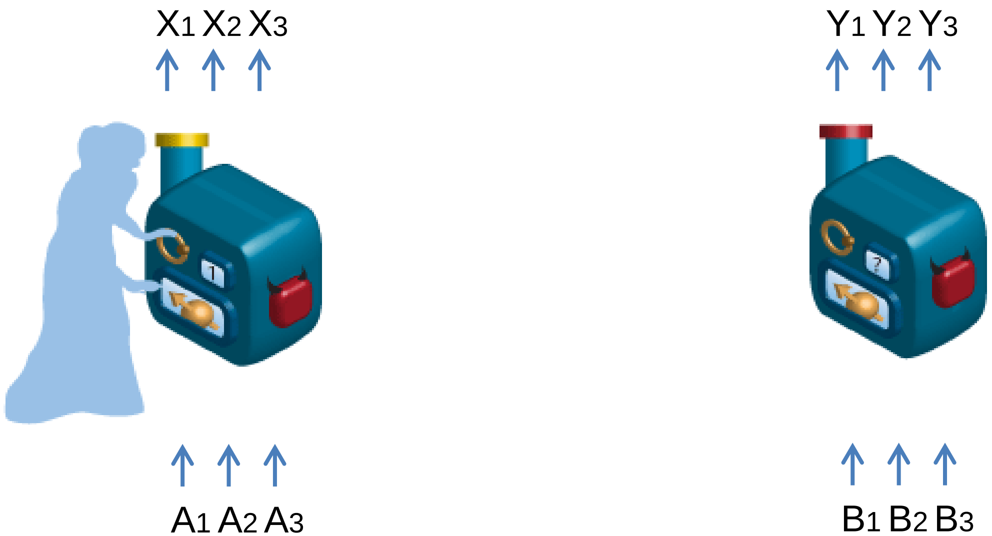
$$P_{XY|AB} = (1-4\varepsilon)\begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + \varepsilon\left(\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}\right)$$

Eve has no knowledge about X

Eve knows X perfectly

Non-signalling Eve can guess X with probability $4\varepsilon + \frac{1}{2}(1-4\varepsilon) = \frac{1}{2} + 2\varepsilon$

Device-independent randomness expansion protocol: Main ideas



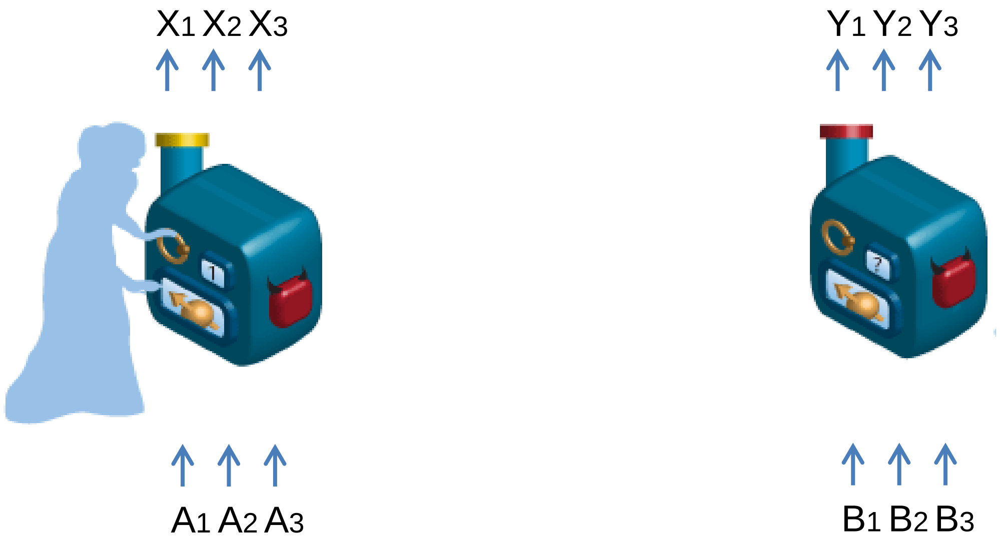
- Doing CHSH test costs randomness
- We want expansion

UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME

QUANTUM COMMUNICATIONS HUB

EPSRC Engineering and Physical Sciences Research Council

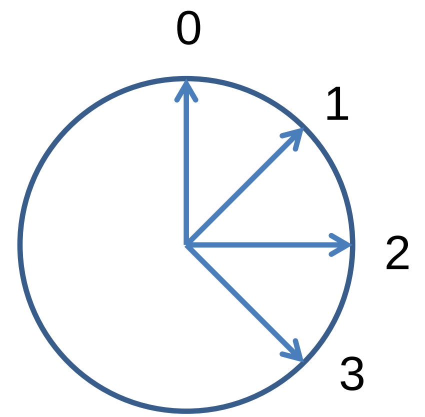
Device-independent randomness expansion protocol: Main ideas



- Divide rounds into "test rounds" (T) and "generation rounds" (G)
- Test rounds are a small subset that cost randomness
- On the generation rounds, fixed inputs are used (no cost), e.g., (try to) measure in $\{|0\rangle, |1\rangle\}$ basis on both

Protocol structure

	A	X		B	Y
G	0	1		0	1
T	2	0		1	1
G	0	1		0	1
T	0	0		1	0
T	2	0		3	0
G	0	1		0	1
G	0	0		0	1
G	0	1		0	0
G	0	1		0	1
G	0	0		0	0
T	0	1		3	0



Use T rounds to check CHSH wins and error rate. For these

If $(\mathrm{A}, \mathrm{B}) = (0,1), (2,1)$ or $(2,3)$, want $X = Y$

If $(A, B) = (0,3)$ want $X \neq Y$

Error rate too high → abort

Protocol structure

	A	X		B	Y
G	0	1		0	1
T	2	0		1	1
G	0	1		0	1
T	0	0		1	0
T	2	0		3	0
G	0	1		0	1
G	0	0		0	1
G	0	1		0	0
G	0	1		0	1
G	0	0		0	0
T	0	1		3	0

Raw string is processed to give final random string

$S_A = 1110110\ldots$

Randomness extraction

$01101\ldots$

NB: randomness extraction needs a short random seed.

UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME
QUANTUM COMMUNICATIONS HUB
EPSRC Engineering and Physical Sciences Research Council

Proof ingredients

- Protocol acts like a filter: for a significant probability of not aborting, the devices must have a large Bell inequality violation almost every time.

- Large Bell inequality violations implies difficulty for Eve to guess.

- If Eve cannot guess the output well, then we can compress the string to one she cannot guess at all. [via randomness extractor]

Randomness accounting

- Randomness input:
  - To choose the test rounds
  - To choose the tests (2 bits per test)
  - To seed the randomness extractor
- Randomness output:
  - If all goes well about 1 bit per round
- Few test rounds, short seed extractors → expansion

Security definition

- What does it mean for a protocol to be secure?
- Define ideal
- Imagine Alice will randomly decide either to perform the real protocol or the ideal.
- The real protocol is secure if it is virtually impossible to distinguish the two.

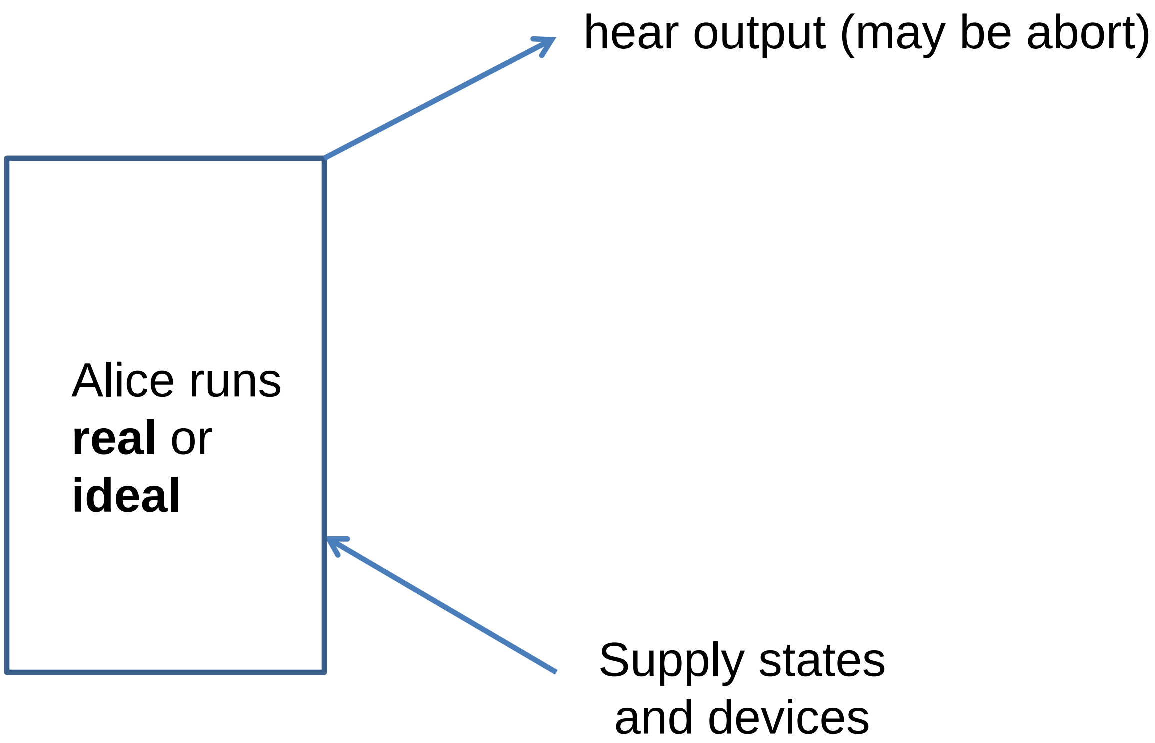
Composable security

- Larger protocol
  - 1.
  - 2.
  - …
  - n. Call randomness expansion sub-protocol
  - n+1.
  - …

Either use **Real** expansion sub-protocol, or **Ideal**

How well can we tell the difference?

Security definition



The ideal

- We want the final state to have the form

$$\tilde{\rho}_{AE} = \sum_{x} \frac{1}{|X|} |x\rangle\langle x|_A \otimes \rho_E$$

The ideal

- We want the final state to have the form

$$\tilde{\rho}_{AE} = \sum_{x} \frac{1}{|X|} |x\rangle\langle x|_A \otimes \rho_E$$

- However, we **don’t** simply define the ideal to output a state of this form.
- (It would be easy to distinguish this from the real protocol, e.g. by forcing real to abort)

The ideal

- Instead, take the ideal protocol to be the real protocol modified such that if it does not abort, right at the end Alice replaces her output by a perfect random string.

$$\sum_x \frac{1}{|X|} |x\rangle\langle x|_A \otimes \rho_E$$

The ideal

- With the ideal defined in this way, it is impossible to distinguish the real and ideal based on abort.

- Only way to distinguish is if both:
  - The protocol does not abort; and
  - The output can be distinguished from a perfect random string

$$D\left(\rho_{AE}, \sum_{x} \frac{1}{|X|} |x\rangle\langle x|_A \otimes \rho_E\right) > 0$$

real

The ideal

- Thus, the security statement is a bound on the *a priori* probability that the protocol does not abort and the output can be distinguished from perfect randomness over all possible devices.

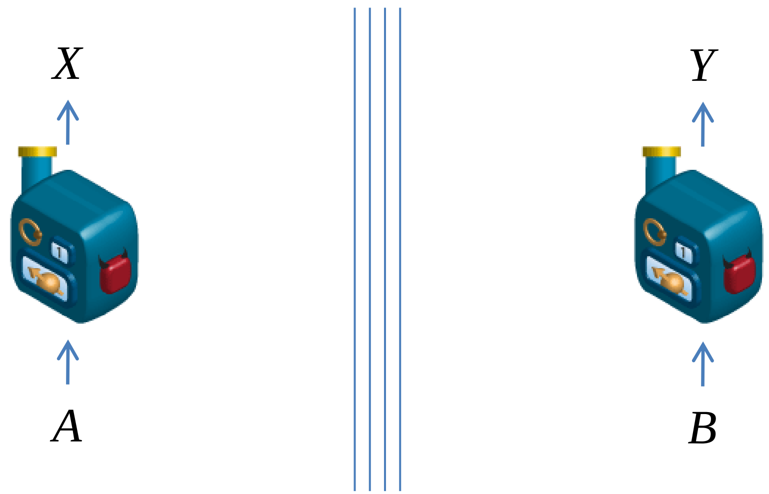
- NB: we don't make statements of the form "Given the protocol did not abort, the output is secure (except with very small probability)"

Technological aspects

- We have theoretical proofs: what about in practice?

Technological aspects

- What about in practice?
- Key ingredient is a Bell inequality violation
  - Need to close detection loophole

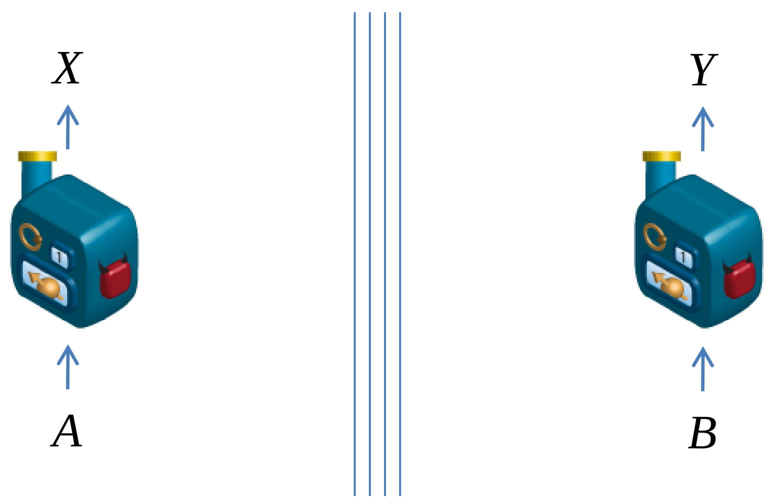


$P_{XY|AB}$ must violate a Bell inequality
In order to verify this, have to include failure to detect events

Technological aspects

- What about in practice?
- Key ingredient is a Bell inequality violation
  - Need to close detection loophole

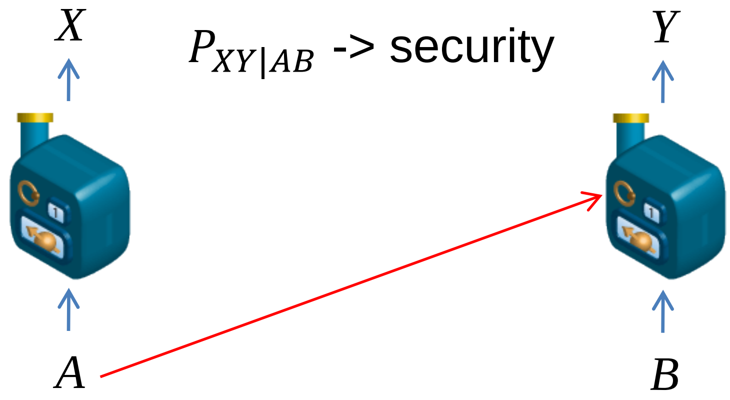
  NB: easier to do this than for QKD



$P_{XY|AB}$ must violate a Bell inequality
In order to verify this, have to include failure to detect events

Technological aspects

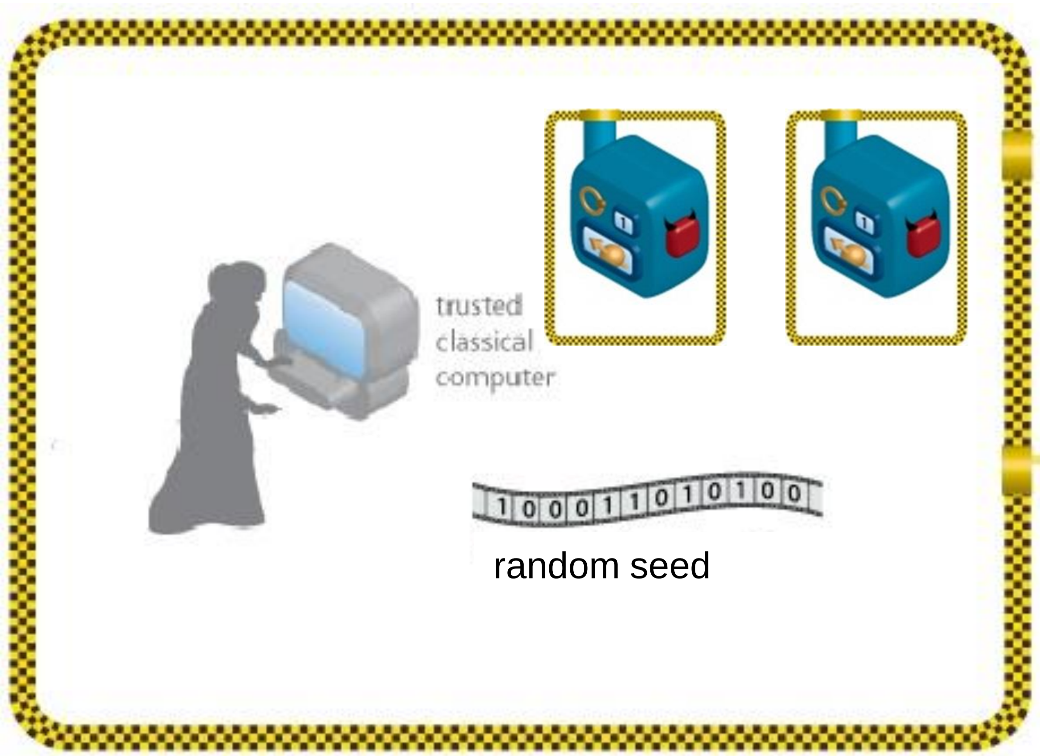
- What about in practice?
  - Need to close detection loophole
  - (Note: no need to close locality loophole; although it doesn't hurt)



Technological aspects

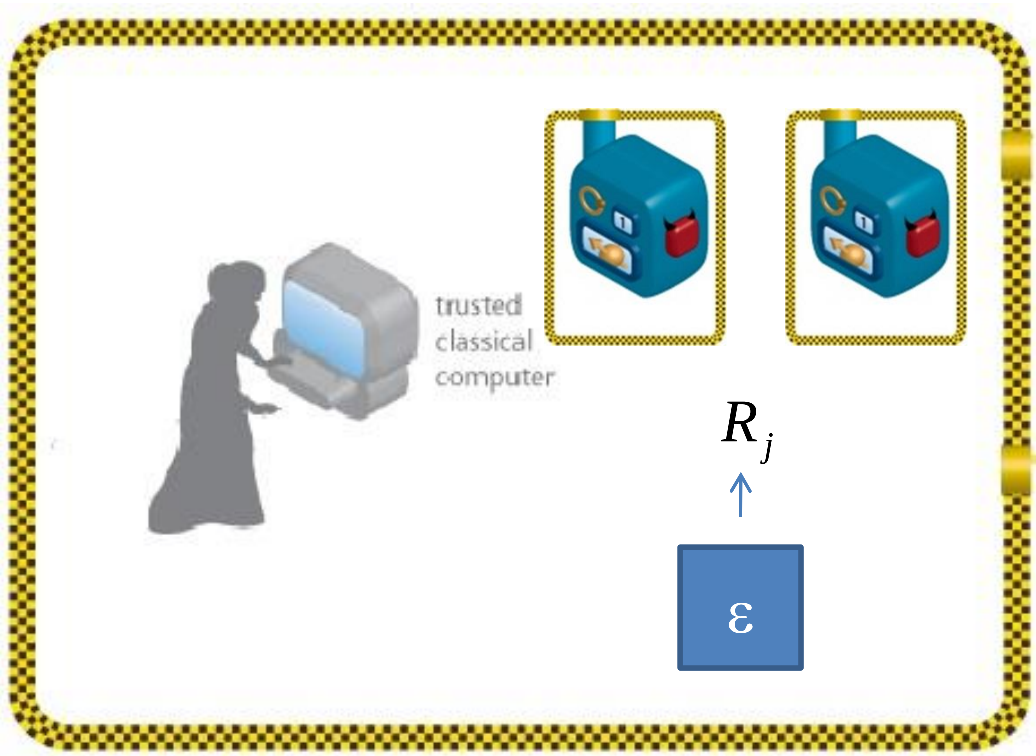
- What about in practice?
  - Need to close detection loophole
  - (Note: no need to close locality loophole; although it doesn't hurt)
  - Need them to be faster to compete with current approaches

Some references



C, Thesis U.Camb. 2007,
CK, JPhysA **44**, 095305 2011
Pironio+, Nature **464**, 1021 2010
PM, PRA **87**, 012336, 2013
FGS, PRA **87**, 012335, 2013
VV, Phil Trans **370**, 3432, 2012
CY, STOC 14
MS, STOC 14, arXiv:1411.6608
ARV, later today

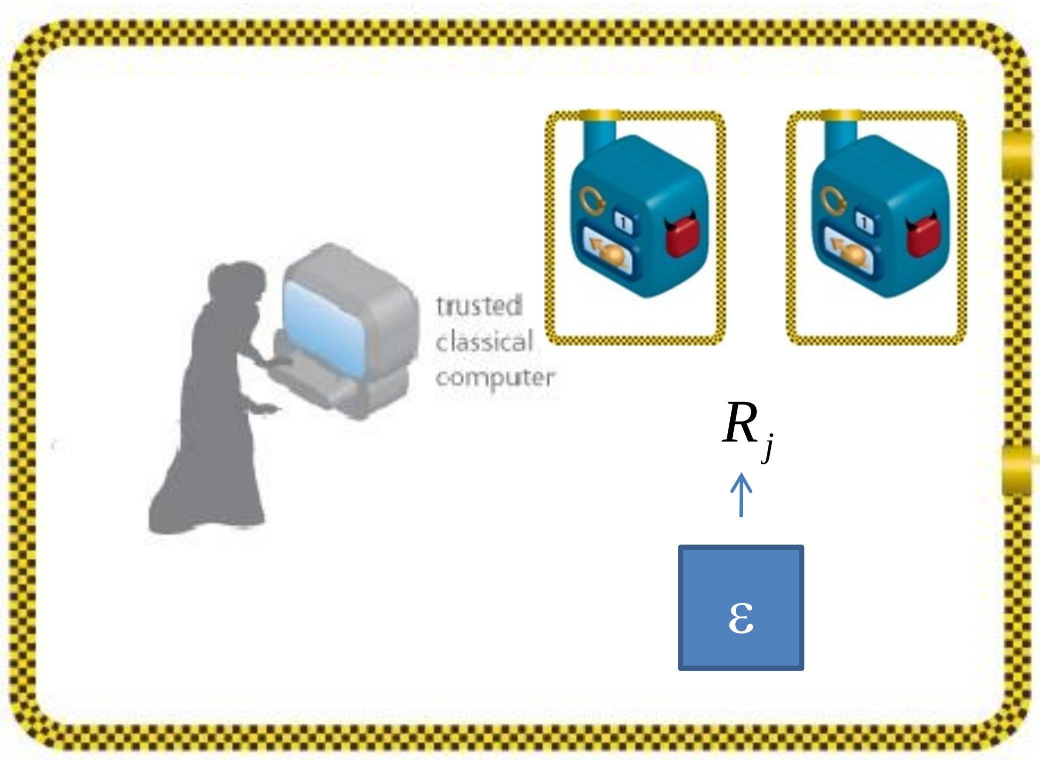
Related task: Randomness Amplification



Imperfect randomness:
- Looks random to Alice
- Partly correlated with other information (that may be held by Eve)

Want to generate perfect randomness

Related task: Randomness Amplification



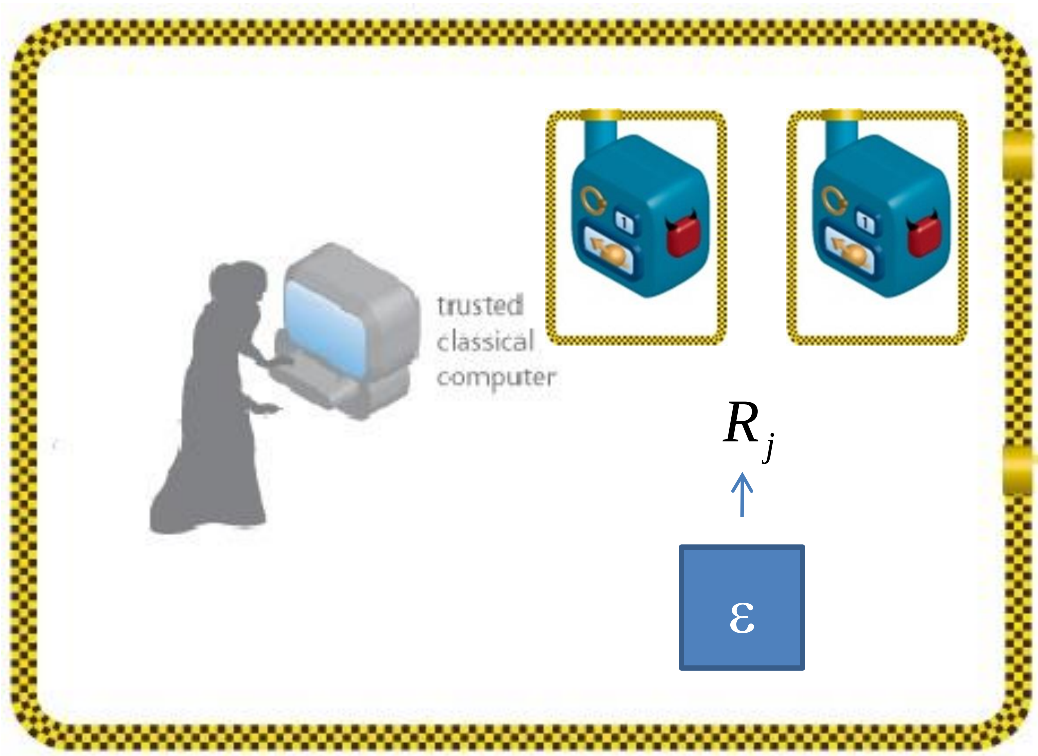
Imperfect randomness:
- Looks random to Alice
- Partly correlated with other information (that may be held by Eve)

E.g., Santha-Vazirani source [FOCS 84]
Limitation to the bias of each bit conditioned on previous ones and adversary.

$$P_{R_j|W} \in \left[\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon\right]$$

Want to generate perfect randomness

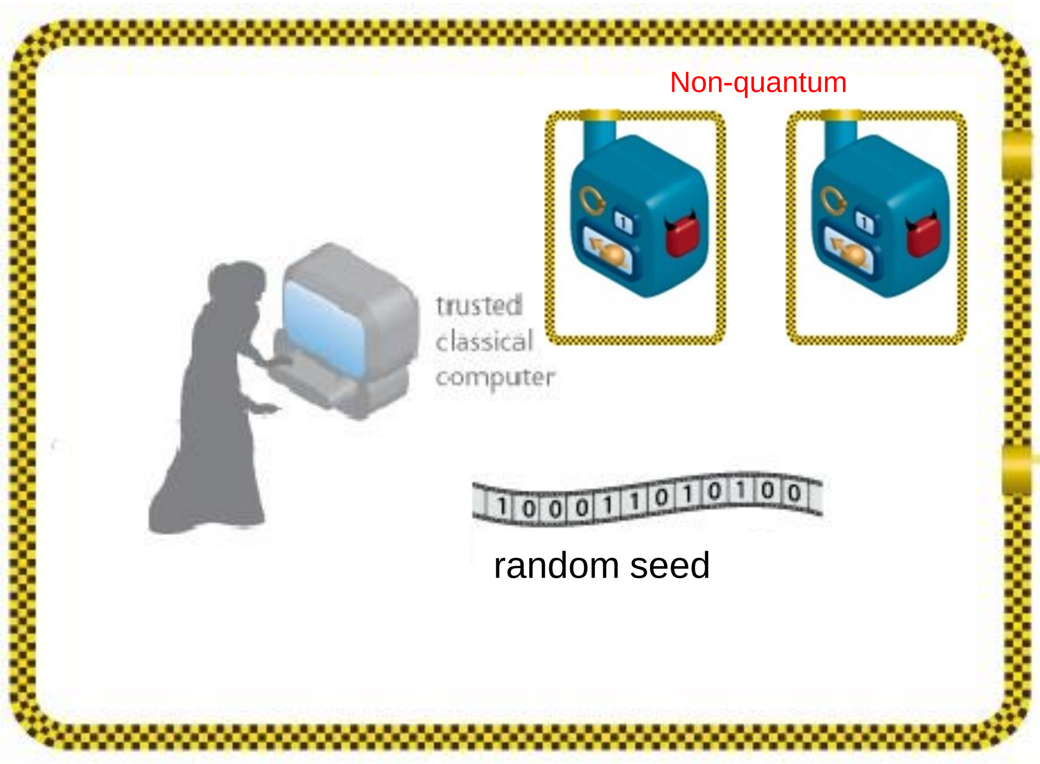
Related task: Randomness Amplification



CR, N.Phys **8** 450 (2012)
Gallego+, N. Commun **4**, 2654 (2013)
Brandao+, N.Commun **7**, 11345 (2016)
CY, STOC 14
CSW, arXiv:1402.4797

Want to generate perfect randomness

Another interesting scenario



Randomness expansion against non-signalling eavesdroppers

Summary

- Classical protocols aim to provide time-limited security

- Standard quantum protocols allow this to be upgraded to unconditional security

- Device-independent protocols allow security against device failure or tampering

more security

fewer assumptions

UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME

QUANTUM COMMUNICATIONS HUB

EPSRC Engineering and Physical Sciences Research Council

Summary

- Advantages:
  - **weaker assumptions -> more security**
  - certify security on-the-fly (calibration errors automatically caught).
- Open challenges
  - Increased speed
  - Sensible ways to reuse untrusted devices
  - Can we get security against no-signalling adversaries?