# Information Theoretic Security: From Classical to Quantum

*Hesham El Gamal*

# Is It Important?

- Critical information transmitted over (wireless) networks

  --bank account number, credit card number, SSN, etc..
- Concerns with security seems to be growing exponentially fast
- Is wireless less secure than wireline?
- What about the information theoretic approach?

# Cryptographic Applications

- Confidentiality: Alice's message to Bob should be kept confidential from Eve.
- Data Integrity: Bob must be sure that Alice's message has not been altered.
- Authentication: Bob must be sure that Alice was the one transmitting the message.
- Non-repudiation: Alice should not be able to claim that she did not send the message.

# Historical Background

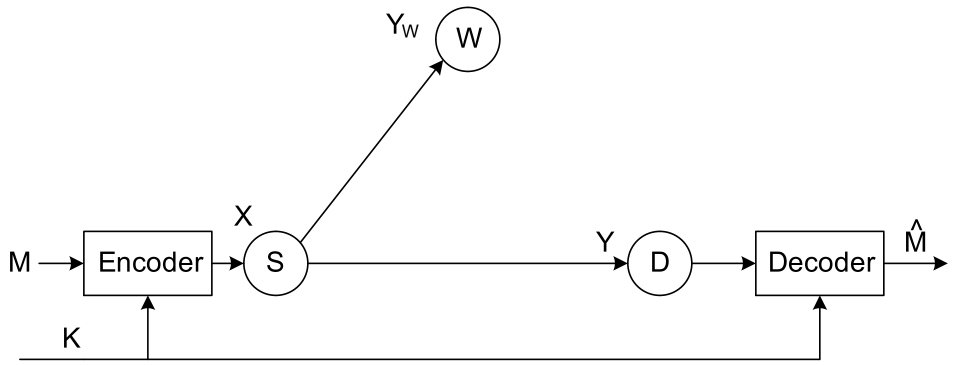
- One of the earliest crypto-systems is often attributed to Julius Caesar
  - The message is encrypted by shifting each letter by a fixed number of places
  - Decryption is performed by shifting back by the same number of places.
  - For example, with three places, a becomes D, b become E, c becomes F, etc…
  - Can be formalized as a (mod 26) shift scheme.

***Very easy to break if the protocol is known***

# Kerckhoffs’s Principle

- In assessing the security of a crypto-system, one should always assume that the enemy knows the method being used
  - More robust assumption (people can defect or be captured by the enemy)
  - Results in the key-based security paradigm.
  - The security of the system depends on the key not the protocol.

# Shannon's Model




$$Y = Y_W = X$$

- Use a private key K to encrypt and decrypt the message M.
- **Noiseless transmission**
- ***Somehow***, the private key is kept confidential from Eve.

# Perfect Secrecy

- It is natural to define perfect secrecy by the condition that, for all cryptograms the *a posteriori* probabilities are equal to the *a priori* probabilities independently of the values of these.
  - In this case, intercepting the message has given the crypto-analyst no information.
  - On the other hand, if the condition is not satisfied, certain key and message choices may occur for which the enemy's probabilities do change. This in turn may affect his actions and thus perfect secrecy has not been obtained.

***The fundamental notion of equivocation***

# Main Result

***Perfect secrecy is possible if and only if***

$$H(K) \geq H(M)$$

- *This result is positive since it establishes the possibility of perfect secrecy (one time pad).*
- *This result is negative since it seems impossible to share a secret key of that size.*
- *Optimal source coding is very critical for secrecy.*

# Public Key Cryptography

- The distribution of secret keys is challenging.
  - For example, if the one time pad was possible, one should send the information over the secure channel.
- Is secure communication possible without the exchange of a private key?
- In a landmark paper, Diffie and Hellman formalized the public key cryptography paradigm as a solution to this problem.
- This paradigm hinges on the concept of a one way function
  - y=f(x).
  - f(.) is one-to-one.
  - It is easy to compute y.
  - Inverting f(.) is essentially impossible unless you know the secret.
  - Only the legitimate receiver knows ***the combination to the lock.***

# The RSA Public Key Cryptosystem

- Invented by Rivest, Shamir, and Adleman in 1977.
- Widely used in electronic commerce protocols.

The receiver generates (e, d, n), publishes (e, n) as public key, leaves d confidential.

| Public key (e, n) | Private key d |
|---|---|
| **Encryption:** | **Decryption:** |
| To send message m | To recover m |
| Compute $C = m^e \mod n$ | Compute $m = C^d \mod n$ |

# State of the Art

- Public-Key approaches are used to distribute private keys.
- The bulk of the data is encrypted using a private (symmetric) key scheme.
- This approach is motivated by the relatively high complexity of public-key protocols.

***Assumes, implicitly, that the separation between error control coding and cryptography is (near) optimal***

# The Quantum Problem

## How to guarantee secrecy?

-- Assumes an eavesdropper with a limited computational power.

-- Assumes that some mathematical problems are difficult to solve
e.g., The RSA approach assumes that it is difficult to factorize large prime numbers

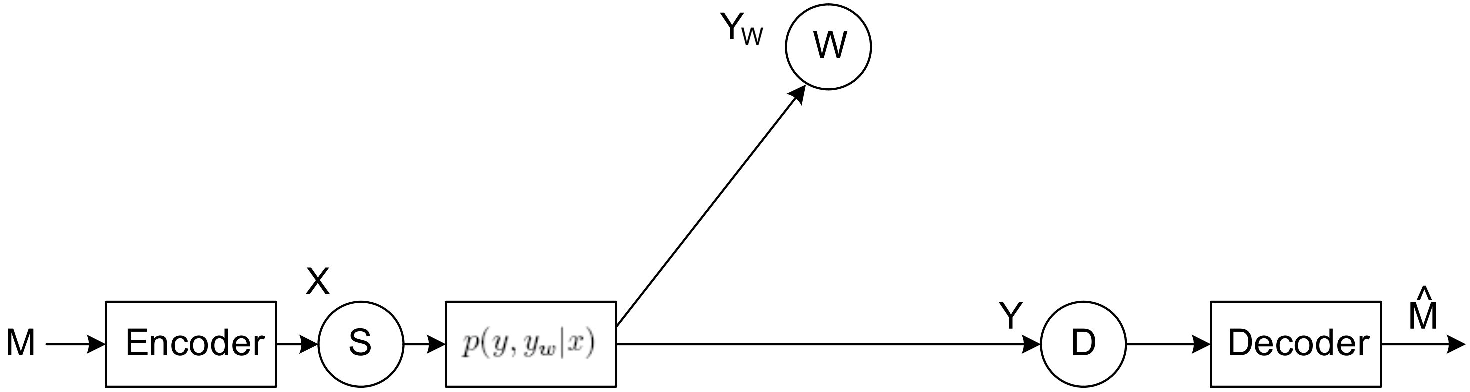
## Limitations

-- The difficulty assumption has not been proved in most cases.
-

***Computationally secure systems will eventually become obsolete!***

# ***Back to Information Theoretic Security***

# The Classical Wiretap Channel




- Takes the transmission uncertainty into consideration

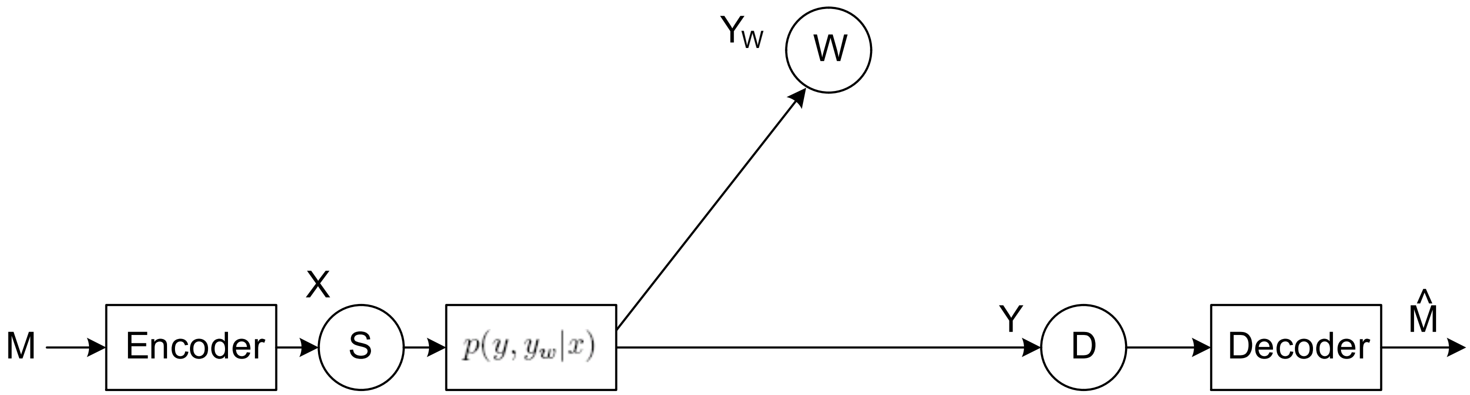
$$p(y, y_W|x)$$

- Achieves perfect secrecy if the main channel is less noisy

eg: $y = x + z, y_W = y + z'$

$$C_s = \max_{V \to X \to YY_W} [I(V;Y) - I(V;Y_W)]$$

# The Limitation




What if the wiretapper is less noisy?

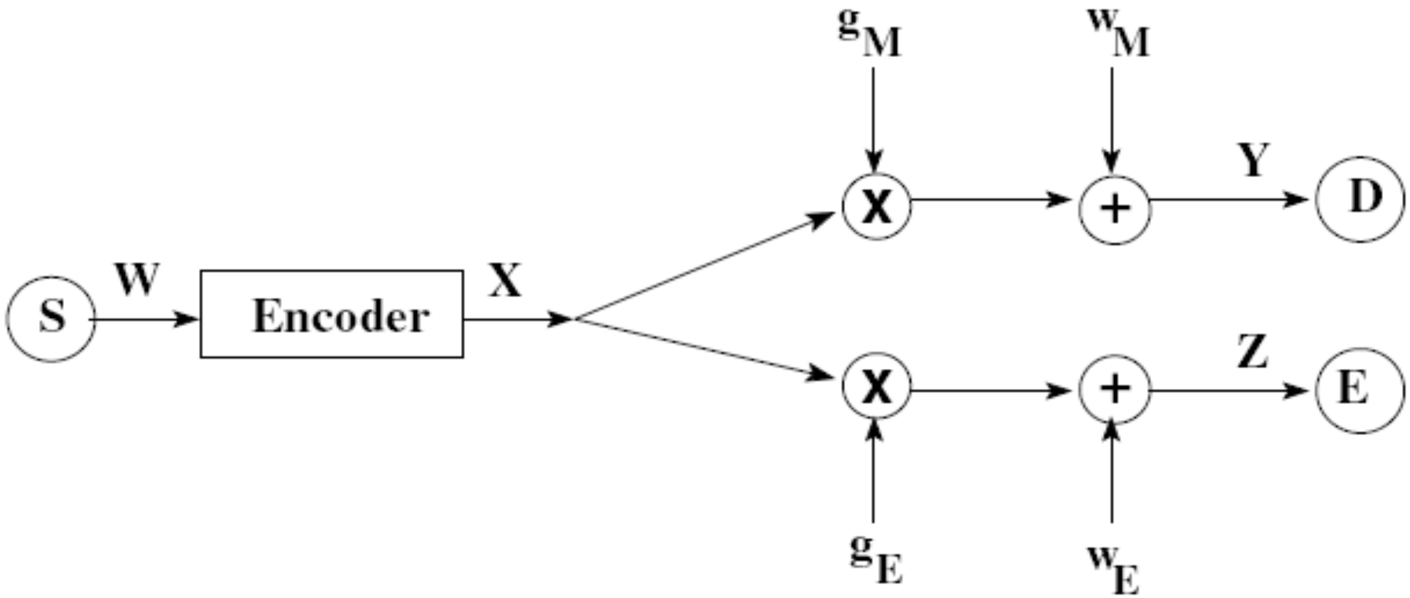
$$C_s = 0$$

# The Wireless Solution

***Can we leverage the wireless medium to avoid the limitation of the Classical Wiretap channel***

# ***Opportunism***

# Secrecy Capacity of Fading Channels




- $h_M(i) = |g_M(i)|^2 \qquad h_E(i) = |g_E(i)|^2$
- $w_M$ , $w_E$ – AWGN noise with unit variance
- Independent channel fading

# Full Transmitter CSI

- Both $h_M$ and $h_E$ known at transmitter
- Transmit only when $h_M > h_E$ (opportunistic secrecy)
- Use instantaneous power and rate adaptation

<u>Theorem</u>

:

$$C_s^{(F)} = \max_{P(h_M,h_E)} \int_0^\infty \int_{h_E}^\infty \Big[\log(1 + h_M P(h_M, h_E)) - \log(1 + h_E P(h_M, h_E))\Big] f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E,$$

$$\textit{such that} \qquad \mathbb{E}\{P(h_M, h_E)\} \leq \bar{P}.$$

# Achievability Scheme

- Codeword rate = $\log\left(1 + h_M P(h_M, h_E)\right)$
- Achievable perfect secrecy rate (at any instant) =

$$[\log\left(1 + h_M P(h_M, h_E)\right) - \log\left(1 + h_E P(h_M, h_E)\right)]^+$$

$$\text{where } [x]^+ = \max\{0, x\}$$

- Averaging over all channel realizations ($h_M$,$h_E$)

$$R_s^{(F)} = \iint [\log\left(1 + h_M P(h_M, h_E)\right) - \log\left(1 + h_E P(h_M, h_E)\right)]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E$$

# Finally

Choose the optimal power allocation policy to maximize the perfect secrecy rate!

$$P(h_M, h_E) = \frac{1}{2}\left[\sqrt{\left(\frac{1}{h_E} - \frac{1}{h_M}\right)^2 + \frac{4}{\lambda}\left(\frac{1}{h_E} - \frac{1}{h_M}\right)} - \left(\frac{1}{h_M} + \frac{1}{h_E}\right)\right]^+$$

$$\text{where } \lambda \ \text{ satisfies } \mathbb{E}\{P(h_M, h_E)\} = \bar{P}.$$

***Different from the celebrated water-filling solution***

# Only Main Channel CSI

- Transmitter only knows $h_M$
- Use an ergodic scheme
- Idea: Hide secure message across different fading states
- Instantaneous power and rate adaptation

**Theorem**

$$C_s^{(M)} = \max_{P(h_M)} \iint \left[\log\left(1 + h_M P(h_M)\right) - \log\left(1 + h_E P(h_M)\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \ ,$$

$$\textit{such that} \qquad \mathbb{E}\{P(h_M)\} \leq \bar{P}.$$

# Achievability Scheme

- Codeword rate = $\log(1 + h_M P(h_M))$
- Average throughput of main channel

$$\iint \log\left(1 + h_M P(h_M)\right) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E$$

***One can achieve this throughput without rate adaptation***

# Why Rate Adaptation is Critical?

- When $h_M < h_E$, mutual info between source & eavesdropper is upper-bounded by $\log(1 + h_M P(h_M))$
- Hence, Information accumulated by the eavesdropper

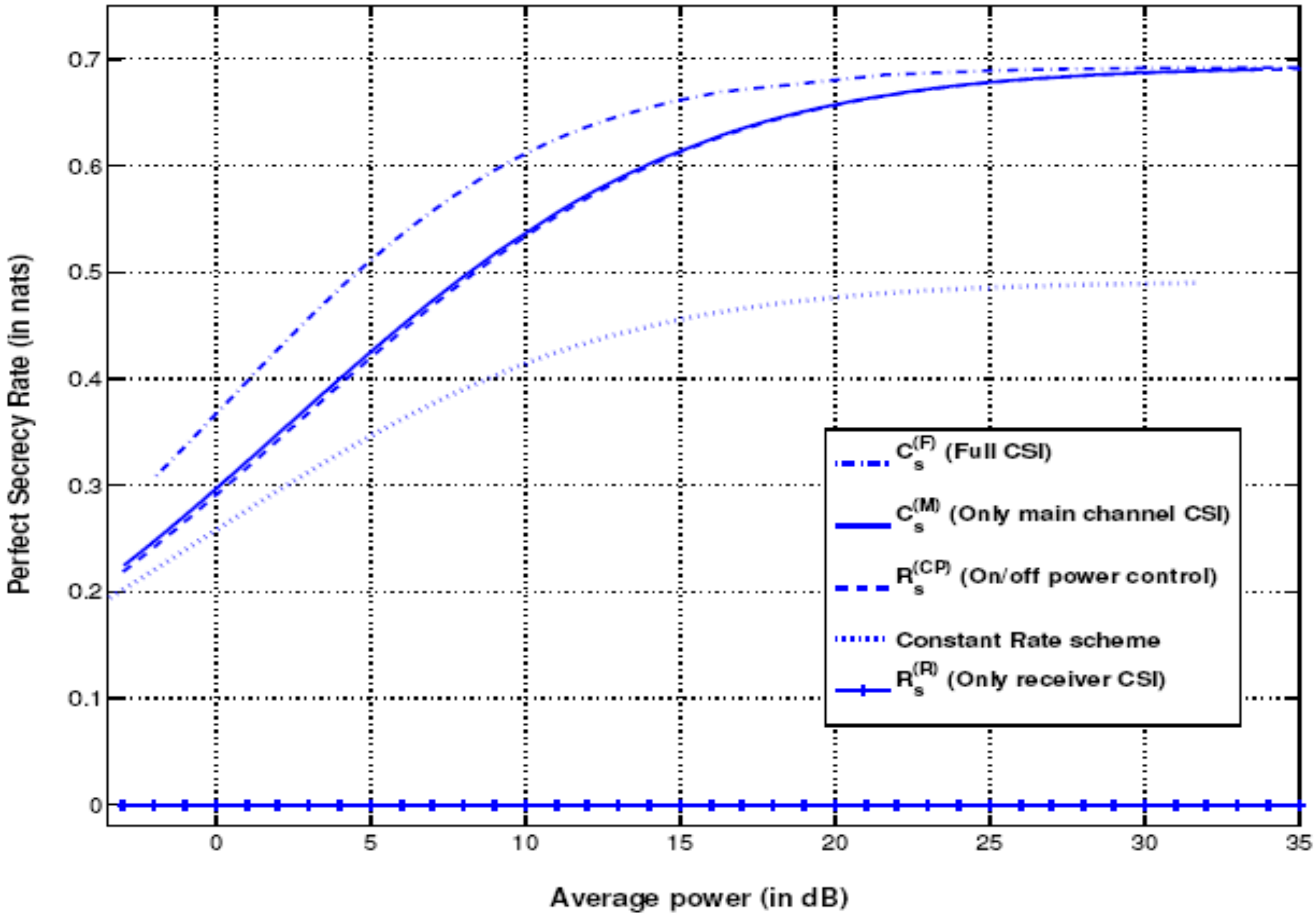
$$\iint \log\left(1 + \min\{h_M, h_E\} P(h_M)\right) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E$$

- Hence the achievable perfect secrecy rate is

$$R_s^{(M)} = \iint \left[\log\left(1 + h_M P(h_M)\right) - \log\left(1 + h_E P(h_M)\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E$$

# Performance Comparison

<u>Symmetric scenario</u>: ( $\mathbb{E}\{h_M\} = \mathbb{E}\{h_E\} = 1$ )



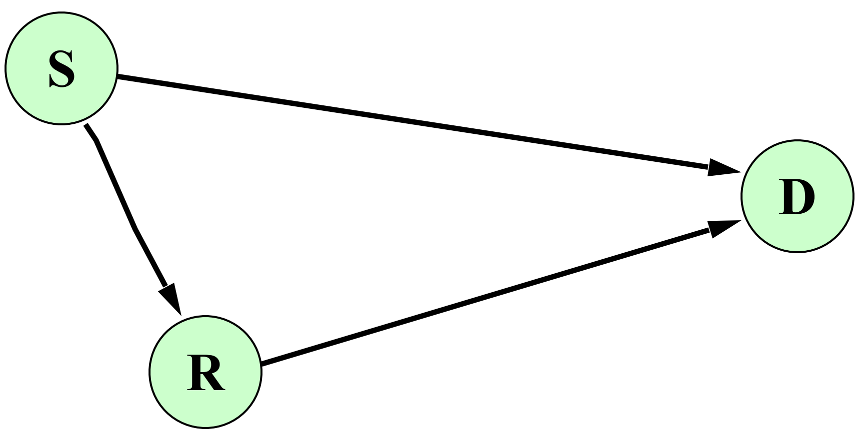
# Summary

- Positive secrecy capacity even when $\mathbb{E}\{h_E\} \geq \mathbb{E}\{h_M\}$
- Fading has a positive impact on secrecy capacity!
- Presence of eavesdropper CSI at transmitter does not increase secrecy capacity at high SNR!
- Knowledge of main channel CSI at transmitter is crucial!
- Rate adaptation is crucial for facilitating secure communication over ***slow*** fading channels!
- Noise insertion enhances the achievable secrecy rate in fast fading channels

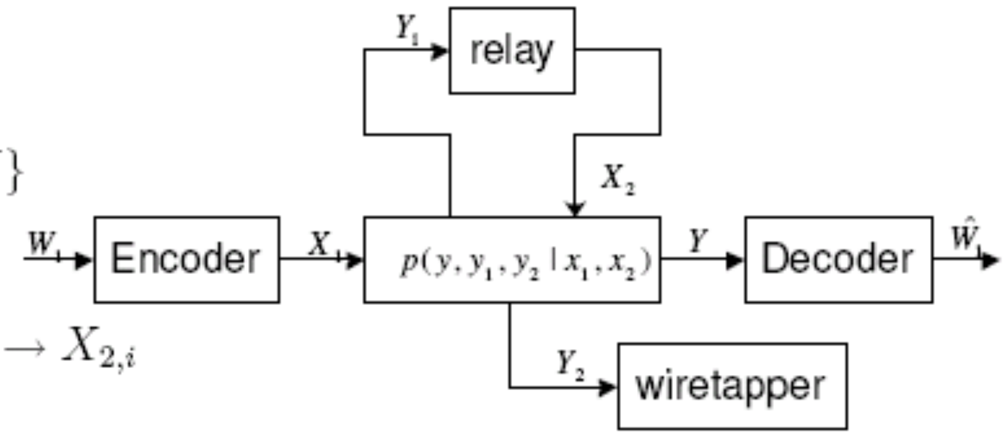
# ***Cooperation***

# Long History!

- Almost all works on cooperative communications start from the relay channel.
- The capacity of the Gaussian relay channel remains unknown and we will not seek to characterize it here
- **What if we add a secrecy constraint to the problem?**



# Cooperation for Secrecy




The relay-wiretap channel

Message set: $W_1 \in \mathcal{W}_1 = \{1, 2, \cdots, M\}$

Encoding function: $f_n : W_1 \to X_1^n$

Relay function: $\varphi_i$: $(Y_{1,1}, Y_{1,2}, \cdots, Y_{1,i-1}) \to X_{2,i}$

Decoding function: $\phi$: $\mathcal{Y}^n \to \mathcal{W}_1$

Error probability: $P_e^n = \sum_{w_1 \in \mathcal{W}_1} \frac{1}{M} \Pr\{\phi(\mathbf{y}) \neq w_1 | w_1 \text{ was sent}\}$

Message rate: $R_1 = \frac{1}{n} \log_2 M$

Equivocation rate: $R_e = \frac{1}{n} H(W_1 | \mathbf{Y}_2)$

What is the tradeoff of $(R_1, R_e)$? How can the relay help the source-destination pair?

# The First Step: Decode and Forward

Block Markov coding and backward decoding

| | Block 1 | Block 2 | Block N-1 | Block N |
|---|---|---|---|---|
| $\underline{x}_1$ | $\underline{x}_1(b(1),1)$ | $\underline{x}_1(b(2),b(1))$ | $\underline{x}_1(b(N-1),b(N-2))$ | $\underline{x}_1(1,b(N-1))$ |
| $\underline{x}_2$ | $\underline{x}_2(1)$ | $\underline{x}_2(b(1))$ | $\underline{x}_2(b(N-2))$ | $\underline{x}_2(b(N-1))$ |

Decoding at the relay (→)

$$R_1 \leq I(X_1; Y_1|X_2)$$

Decoding at the destination (←)

$$R_1 \leq I(X_1, X_2; Y)$$

$$nR_e = H(W_1|\mathbf{Y}_2) \geq H(\mathbf{X}_1) + H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) - H(\mathbf{Y}_2) - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2)$$

$H(\mathbf{X}_1)$: $nR_1$

$H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) - H(\mathbf{Y}_2)$: $\geq -nI(X_1, X_2; Y_2) + n\delta_n$

$H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2)$: To drive this term down, rate pair given W1 needs to be inside of the capacity region of wiretapper

# The Rate-equivocation Region

***Theorem:***

The rate pairs in the closure of the convex hull of all $(R_1, R_e)$ satisfying

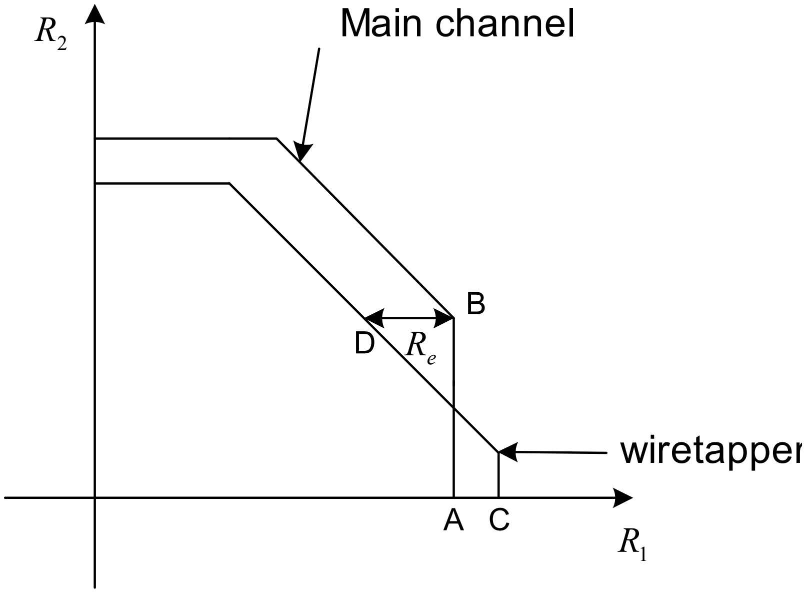
$$\begin{aligned} R_1 &< \min\{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\}, \\ R_e &< R_1, \\ R_e &< \left[\min\{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\} - I(X_1, X_2; Y_2)\right]^+, \end{aligned}$$

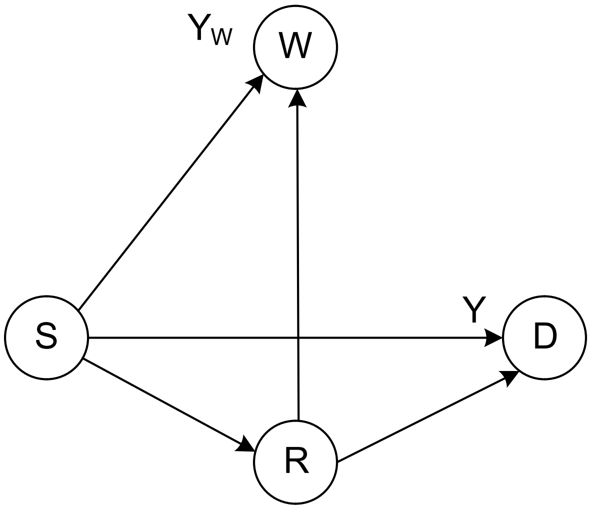
for some distribution $p(x_1, x_2, y_1, y_2, y) = p(x_1, x_2)p(y_1, y_2, y|x_1, x_2)$ are achievable.

Advantage: maximization over $p(x_1, x_2)$, beamforming

Disadvantage: bottleneck at the relay

# What If the Relay Can Not Decode?

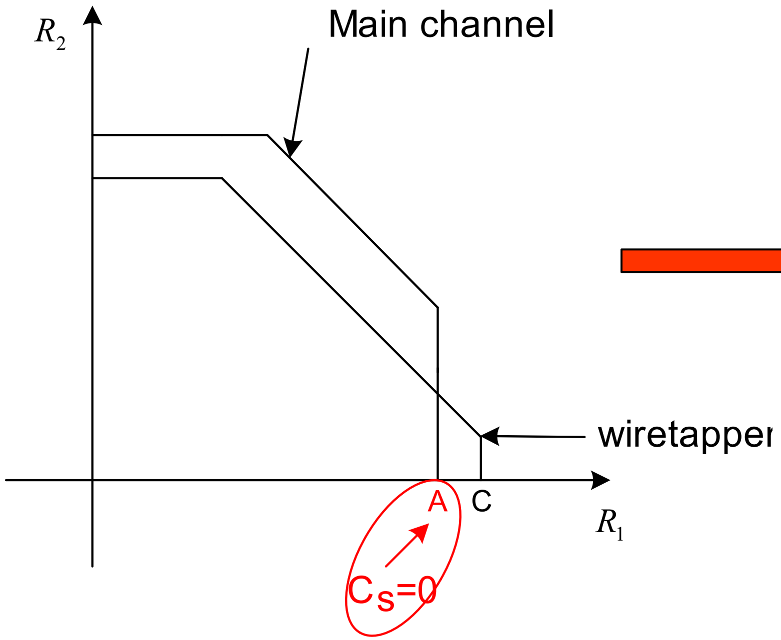


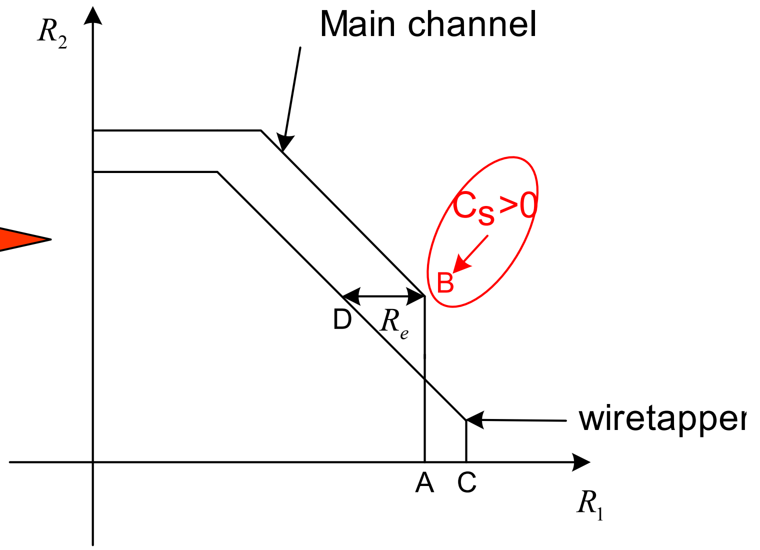




The relay just sends noise!
Doesn't even listen.

- Simple, works for relay node with practical constraints
- Special for the relay channel with secrecy constraints

# Noise Forwarding: Case 1




If no relay, work on A

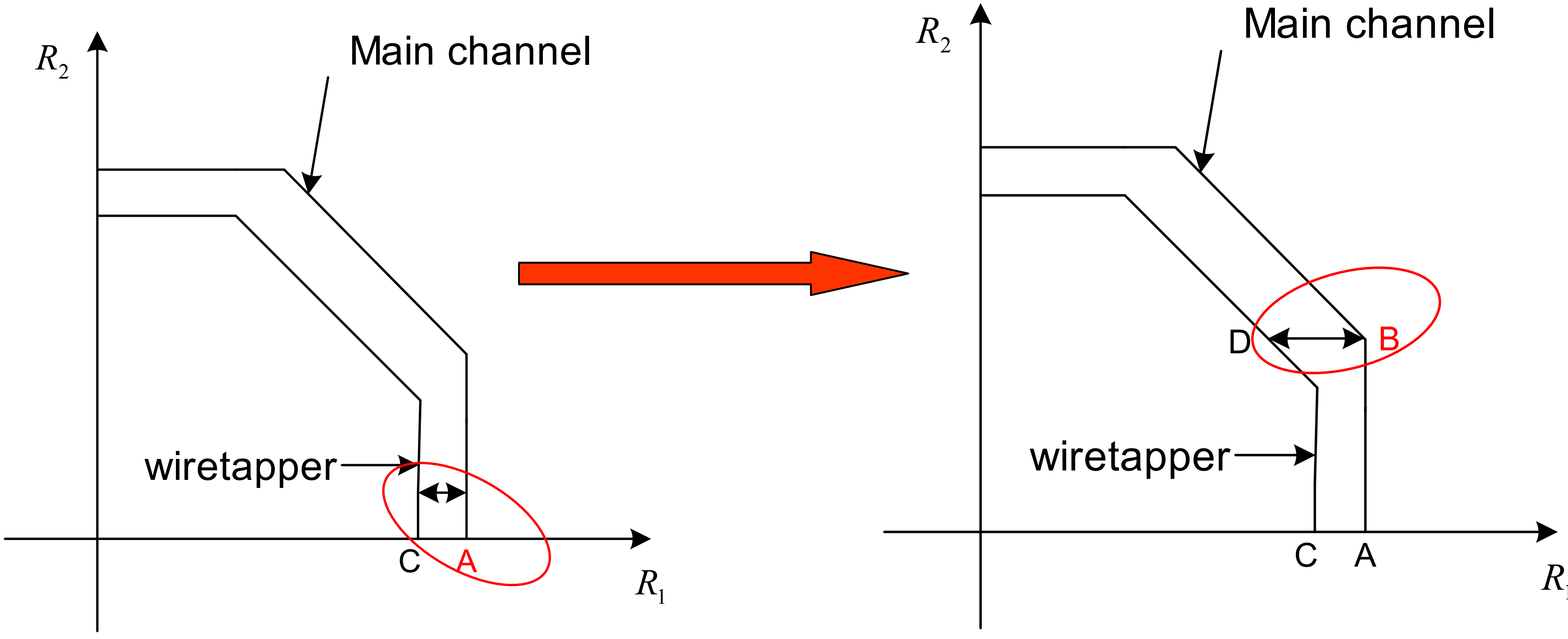
$C_s = 0$

With relay, coordinate to work on B

$C_s > 0$

From zero to positive

# Noise Forwarding: Case 2




If no relay, work on A

With relay, coordinate to work on B

From small to large

# The Rate-equivocation Region

***Theorem:***

The rate pairs in the closure of the convex hull of all $(R_1, R_e)$ satisfying

$$\begin{aligned} R_1 &< I(X_1;Y|X_2), \\ R_e &< R_1, \\ R_e &< \big[\min\{I(X_2;Y), I(X_2;Y_2|X_1)\} + I(X_1;Y|X_2) \\ &\quad - \min\{I(X_2;Y), I(X_2;Y_2)\} - I(X_1;Y_2|X_2)\big]^+, \end{aligned}$$

for some distribution $p(x_1, x_2, y_1, y_2, y) = p(y_1, y_2, y|x_1, x_2)\ p(x_1)p(x_2)$

are achievable.

# The Deaf Helper Phenomenon

We further require perfect secrecy at the relay node $I(W_1; Y_1) = 0$

Corollary:

The achievable perfect secrecy rate of the NF scheme with an additional security constraint at the relay node is

$$R_s = \max_{p(x_1)p(x_2)} \min\{R_{1,e}, R_{2,e}\}$$

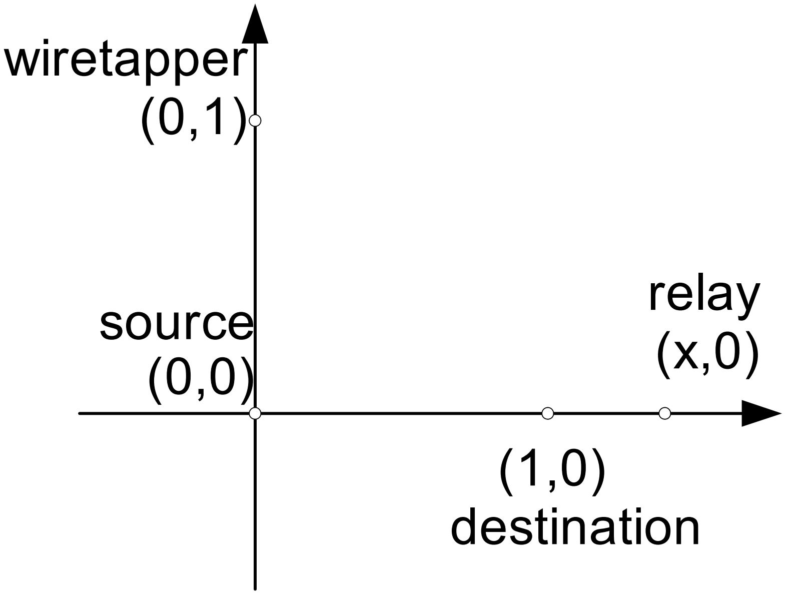
where:

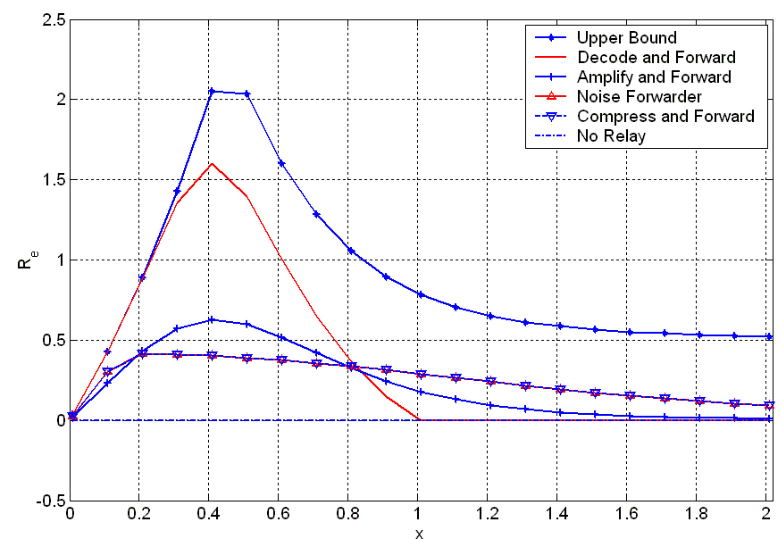
$$R_{1,e} = \left[\min\{I(X_2;Y), I(X_2;Y_2|X_1)\} + I(X_1;Y|X_2) - \min\{I(X_2;Y), I(X_2;Y_2)\} - I(X_1;Y_2|X_2)\right]^+$$

$$R_{2,e} = [I(X_1;Y|X_2) - I(X_1;Y_1|X_2)]^+$$

***The relay is still able to help even when it is totally ignorant of the source's messages!***

# Performance



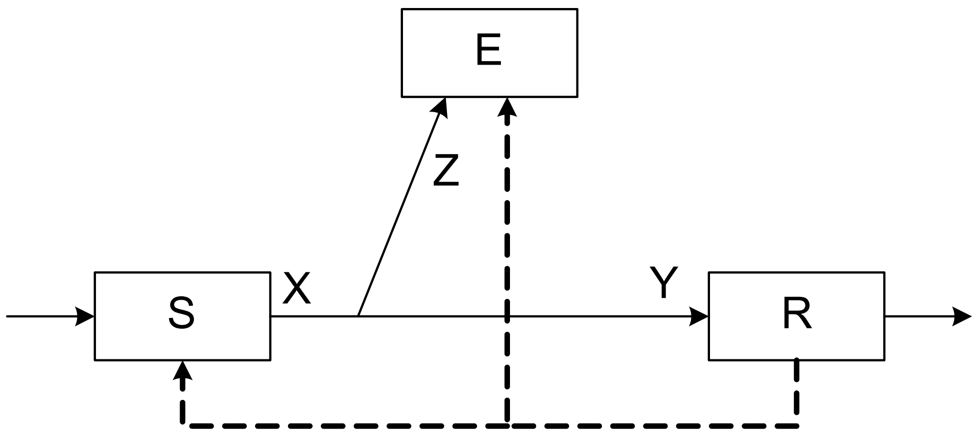




AWGN channel $h_{ij} = d_{ij}^{-\gamma}$

1. When the relay is close to the source, DF achieves the upper-bound
2. When x>1, DF doesn't work, NF has the best performance

# *Feedback*

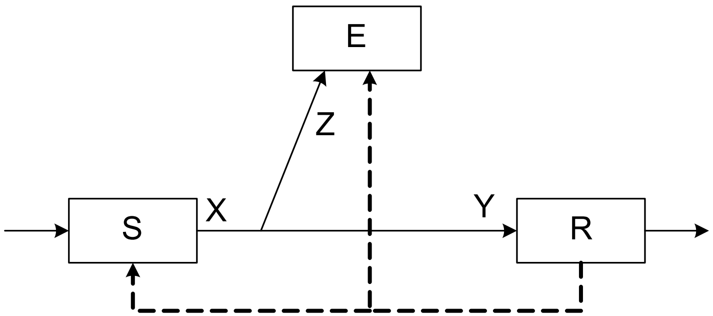
# Feedback for Secrecy: Public Discussion




The receiver and transmitter can discuss over a separate public channel

The eavesdropper has ***ONLY*** full read access to this public channel

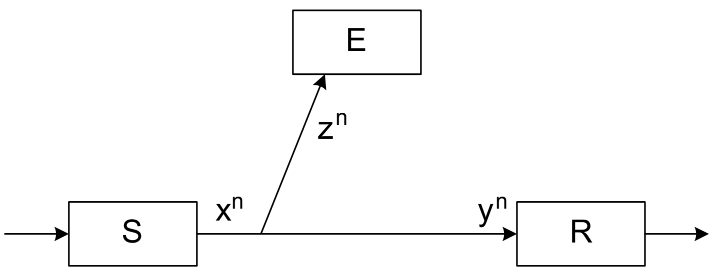
# Example




$X \longrightarrow Y$ BSC with $\epsilon$

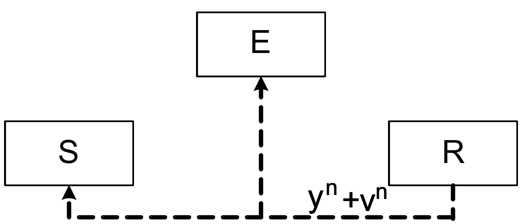
$X \longrightarrow Z$ BSC with $\delta$

$$\delta \leq \epsilon \qquad C_s = 0$$




S randomly generates $\mathbf{x}^n$ with $P(x(i) = 1) = 1/2$

transmits it over the BSC channel




R generates information containing $\mathbf{v}^n$

sends $\mathbf{y}^n + \mathbf{v}^n$ over the public channel

Equivalent channels

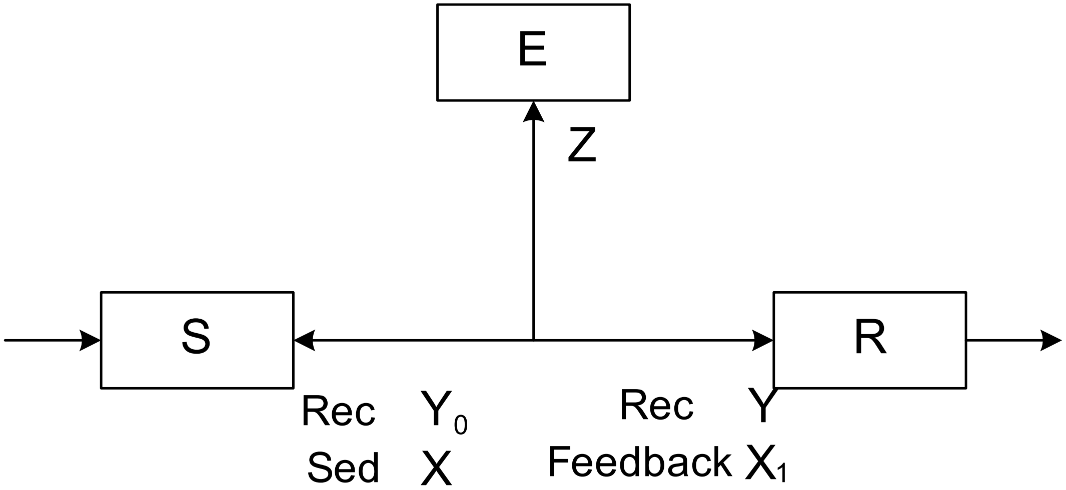
S $\quad \mathbf{v}^n + \mathbf{y}^n + \mathbf{x}^n$

E $\quad \mathbf{v}^n + \mathbf{y}^n + \mathbf{z}^n$

$Y \longrightarrow X$ BSC with $\epsilon$

$Y \longrightarrow Z$ BSC with $\epsilon + \delta - 2\epsilon\delta$

# A More Realistic Model




Feedback at time $i$ : $$X_1(i) = \Psi(Y^{i-1})$$

Received noisy feedback: $$y_0(i) = x(i) + x_1(i) + n_0(i)$$

Transmission at time $i$ : $$x(i) = f(i, w, y_0^{i-1})$$

Received noisy signals: $$y(i) = x(i) + x_1(i) + n_1(i)$$
$$z(i) = x(i) + x_1(i) + n_2(i)$$

We consider modulo-additive channel

# Secrecy Capacity

Theorem:

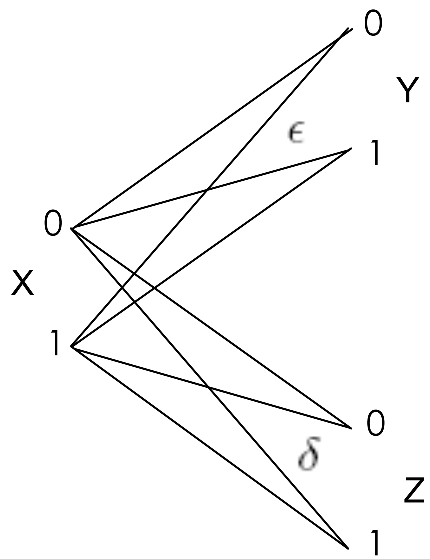
The secrecy capacity of the discrete memoryless modulo-additive wiretap channel with noisy feedback is

$$C_s^f = C$$

where C is capacity of the main channel in the absence of the wiretapper.

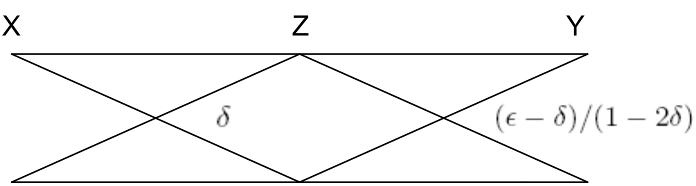
Noisy feedback increases the secrecy capacity to the capacity of the main channel

# Example



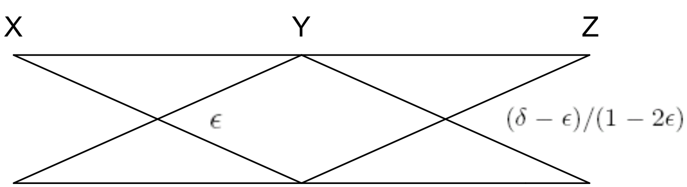

No feedback $C_s = [H(\delta) - H(\epsilon)]^+$

| | Noiseless feedback | Noisy feedback |
| --- | --- | --- |
| $\epsilon = \delta = 0$ | $C_s^p = 0$ | $C_s^f = 1$ |
| $0 < \delta < \epsilon < 1/2$ n1,n2 independent | $C_s^p = H(\epsilon + \delta - 2\epsilon\delta) - H(\epsilon)$ | $C_s^f = 1 - H(\epsilon)$ |




$0 < \delta < \epsilon < 1/2$
n1=n2+n'

$C_s^p = 0$ $\quad$ $C_s^f = 1 - H(\epsilon)$




$0 < \epsilon < \delta < 1/2$
n2=n1+n'

$C_s^p = H(\delta) - H(\epsilon)$ $\quad$ $C_s^f = 1 - H(\epsilon)$

# In a nutshell….

***The wireless medium can be leveraged to facilitate secure communications.***

***But, it relies on statistical channel assumptions***

# QKD: The Quantum Solution

Two new resources

Entanglement

Measurement Postulate

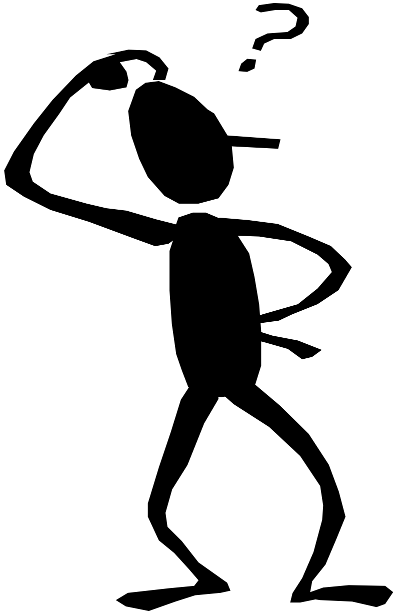
1. Alice chooses random data and basis bits.
2. Alice encodes the data bits in quantum states.
3. Alice sends the encoded states (qubits) to Bob.
4. Bob announces and the reception and measures in random basis.
5. Alice announces the correct bases.
6. Alice and Bob discard the bits corresponding to mismatched bases.
7. Alice and Bob announce and compare a subset of bits (checks)
8. Abort if the number of mismatched checks is larger than a threshold

# Quantum Advantage/Limitation

- Provable Security with **no assumptions**.
- **Only line of sight or fibre links** with relatively short.

- **What about the QIoE?**

THE OHIO STATE UNIVERSITY



The ``key'' maybe in combining the wireless and quantum worlds