Quantum Steganography

$$|\psi\rangle$$

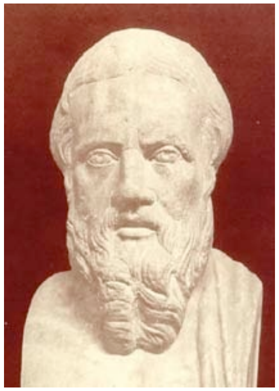
Todd A. Brun and Bilal A. Shaw

November 12, 2009

KITP, Santa Barbara

Outline

- A Brief History of Sneakiness
- Disguising Information as Errors
  - Noiseless Binary Symmetric Channel (BSC)
  - Noisy BSC
  - Key usage
  - Noiseless Depolarizing Channel
  - Noisy Depolarizing Channel
  - Effects of Eve’s Monitoring
  - Security
- Hiding Information in Error Syndromes
  - Classical Noiseless Case
  - Quantum Noiseless Case
  - Questions in Channels with Noise
- Conclusion and Open Questions



HERODTUS

485 – 420 BCE

... dreading therefore each of these things, he [Aristagoras] meditated a revolt: for it happened at the same time that a messenger with his head **punctured** (**steganographic message**) came from Susa from Histiaeus, urging Aristagoras to revolt from the king. For Histiaeus, being desirous to signify to Aristagoras his wish for him to revolt, had no other means of signifying it with safety, because the roads were guarded; therefore, having shaved the head of the most trustworthy of his slaves, he marked it, and waited till the hair was grown again: as soon as it was grown again (**encoding**), he sent him to Miletus without any other instructions than that when he arrived at Miletus he should desire Aristagoras to shave off his hair (**decoding**) and look upon his head: the punctures, as I said before, signified a wish for him to revolt. [*The Histories of Heredotus*]

Steganography = steganos+graphia

steganos = "covered"

graphia = "writing"

For when Xerxes had determined to invade Greece, Demaratus, who was then at Susa, and had heard of his intention, communicated it to the Lacedaemonians. But he was unable to make it known by any other means, for there was great danger of being detected; he therefore had recourse to the following contrivance: having taken a folding tablet, he scraped off the wax, and then wrote the king's intention (**steganographic message**) on the wood of the tablet; and having done this, he melted the wax again over the writing (**encoding**), in order that the tablet, being carried with nothing written on it, might occasion him no trouble from the guards upon the road. When it arrived at Sparta, the Lacedaemonians were unable to comprehend it; until, as I am informed, Gorgo, daugther of Cleomenes, and wife to Leonidas, made a suggestion, having considered the matter with herself, and bade them scrape off the wax (**decoding**), and they would find letters written on the wood. They, having obeyed, found and read the contents, and forwarded them to the rest of the Greeks. These things are reported to have happened in this manner.[*The Histories of Heredotus*

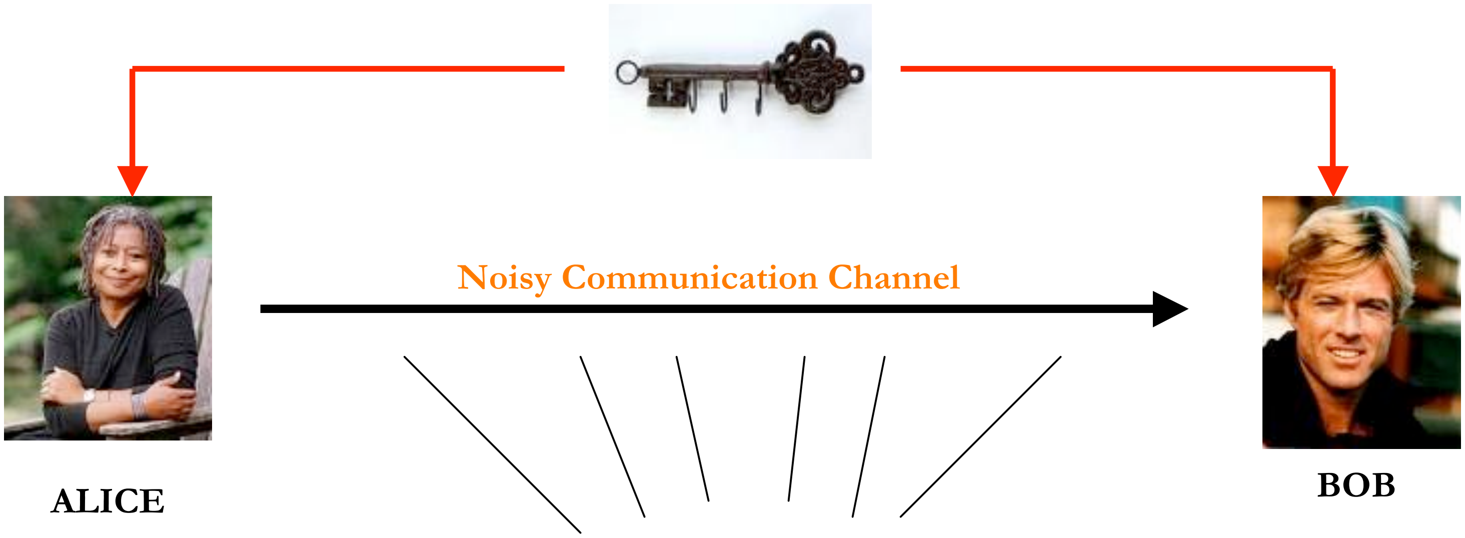
Johannes Trithemius

Profession: Abbot and occultist

Steganographia

Written 1500; published Frankfurt 1606

Gustavus J. Simmons' Model of Steganography



Story Line: Eve is the warden of a prison. She has incarcerated Alice and Bob in two separate cells. However, she does allow them to communicate.

Prior to their incarceration Alice and Bob shared a secret key which they now use to devise an escape plan, by sending hidden messages in their communiqués. These messages must look perfectly innocent to Eve, or she will prevent them from communicating.

CAN ALICE AND BOB DUPE EVE?

EVE IS A PASSIVE ADVERSARY

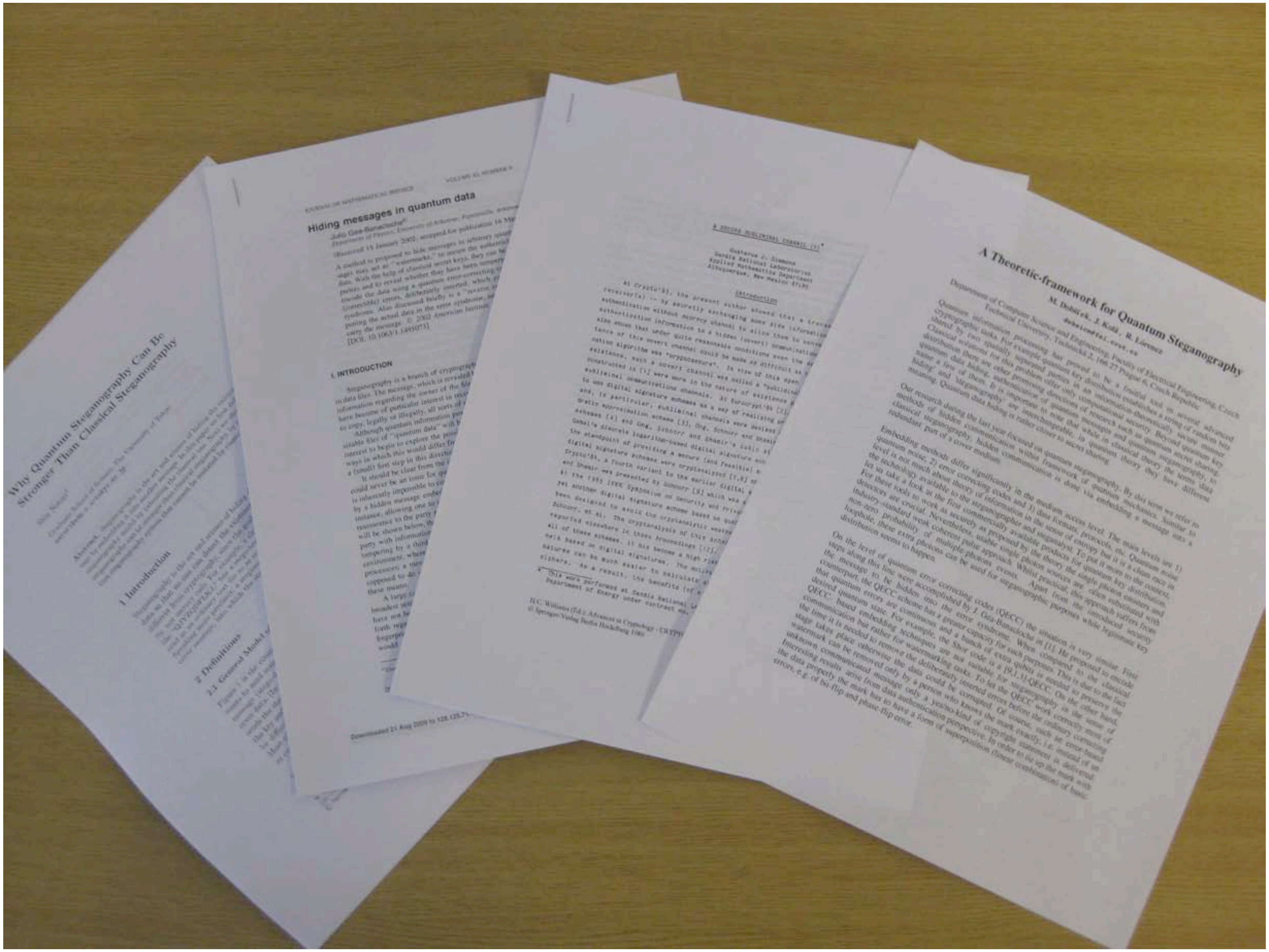
Gustavus J. Simmons. *The Prisoner's Problem And The Subliminal Channel.* Advances in Cryptology – CRYPTO 83, pp. 51-67

Requirements for Quantum Steganography

- What is required to make a quantum message look “innocent” to Eve?
- How can Alice and Bob hide information from Eve?
- How *much* information can they hide?
- Can they hide quantum information?

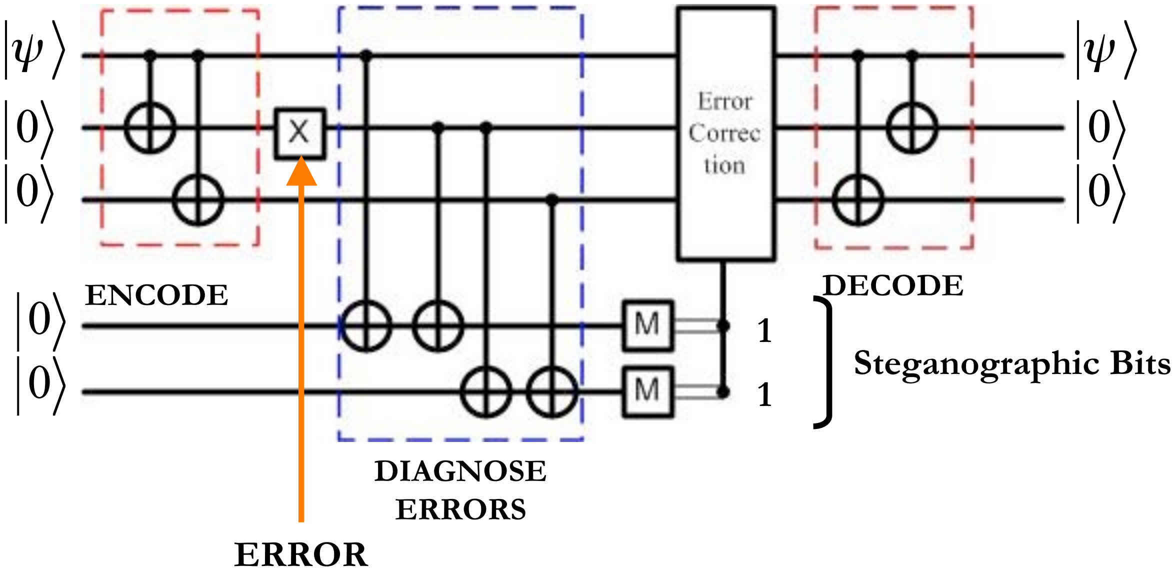
- We will start by looking at a model of *classical* steganography.

Past Work


Why Quantum Steganography Can Be Stronger Than Classical Steganography
Hiding messages in quantum data
Julio Gea-Banacloche
A SECURE SUBLIMINAL CHANNEL (?)
Gustavus J. Simmons
A Theoretic-framework for Quantum Steganography
M. Dobšíček, J. Kolá, R. Lórencz

Past Work

- Julio Gea-Banacloche
  - Hiding Messages in Quantum Data, *Journal of Mathematical Physics*, Volume 43, Number 9, September 2002.
- M. Curty and D. J. Santos
  - Protocols for Quantum Steganography, *2nd Bielefeld Workshop on Quantum Information and Complexity*, 12, 2004.
- Shin Natori
  - Why Quantum Steganography Can Be Stronger Than Classical Steganography, *Quantum Computation and Information*, **102**, 2006.
- M. Dobsiek
  - Simulation on Quantum Authentication. *Physics of Particles and Nuclei, Letters,* 2006.
- K. Martin
  - Steganographic Communication With Quantum Information. *Lecture Notes in Computer Science*, **4569**, 2008



ENCODE

ERROR

DIAGNOSE ERRORS

DECODE

Steganographic Bits

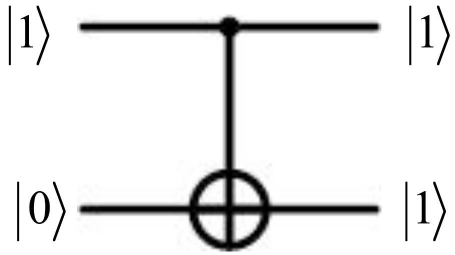
$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

ENCODE

$$|\psi_L\rangle = \alpha|000\rangle + \beta|111\rangle$$

$$|\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$$

Error	PC Anc 1	PC Anc 2	Bit String	Error Operator
0 0 0	0	0	00	I
0 0 1	0	1	01	X_1
0 1 0	1	1	11	X_2
1 0 0	1	0	10	X_3

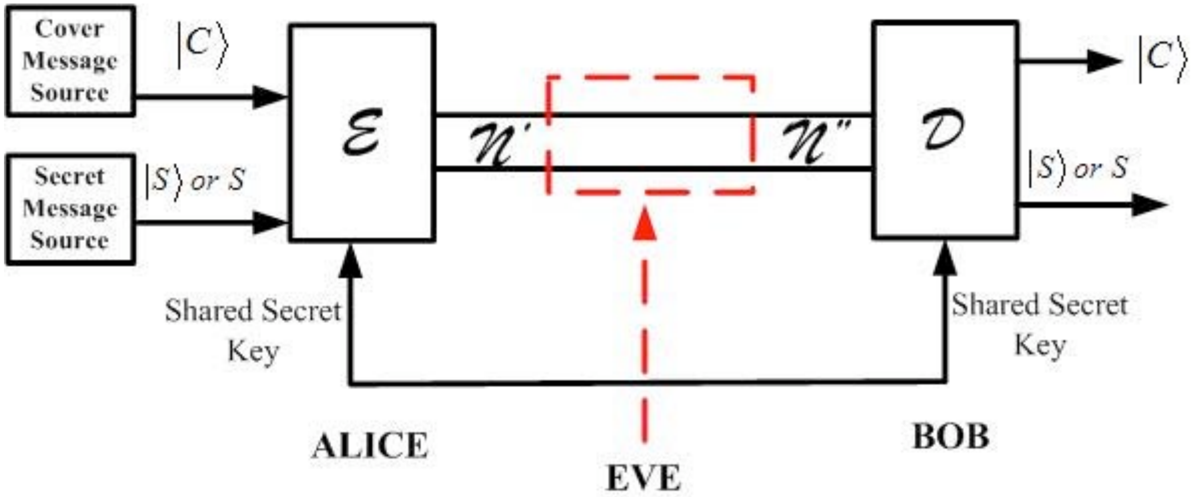


Control	Target	
$\|0\rangle$	$\|0\rangle$	$\|00\rangle$
$\|0\rangle$	$\|1\rangle$	$\|01\rangle$
$\|1\rangle$	$\|0\rangle$	$\|11\rangle$
$\|1\rangle$	$\|1\rangle$	$\|10\rangle$

- Here the requirement of an “innocent” message is fulfilled by having the stego bits appear like errors in an error-correcting code.
- However, in this naïve formulation, the errors don’t appear very natural. The error “rate” is far too high for the choice of code, and the probabilities of the three syndromes don’t match a natural choice of channel. If Eve is paying attention, she will be suspicious.
- If Eve guesses that there is information hidden, she can read it without difficulty.
- Is it possible to send *quantum* information in this way?

Let’s try a more sophisticated take on the same idea.

General Protocol for Quantum Steganography



Hiding Classical Information in the Binary Symmetric Channel

$$\begin{tikzcd} \rho \arrow[r, "\mathcal{N}"'] & \mathcal{N}_p\rho \end{tikzcd}$$

The binary symmetric channel has a probability *p* to flip each bit that passes through it. (We will write this using quantum notation to make it easy to go to the quantum case later.)

$$\mathcal{N}_p\rho \equiv (1-p)\rho + pX\rho X$$

$$\mathcal{R}\rho \equiv \frac{1}{2}\left(\rho + X\rho X\right)$$

$$\mathcal{N}_p\rho \equiv (1-2p)\rho + 2p\mathcal{R}\rho$$

Instead of flipping a bit with probability *p*, we can think of replacing bits by completely random bits with probability 2*p*.

$$\mathcal{N}_p\rho \equiv (1-r)\rho + r\mathcal{R}\rho$$

$$I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Pauli Matrices

Classical Protocol

1. Alice and Bob choose an [*N,k*] ECC to encode *k* bits of covertext. This is the *outer code*. (It is important to choose *N* sufficiently large.)
2. Using the secret key, Alice and Bob determine the number *M* of randomized bits using the binomial expansion with probability 2*p*. They want *M* to be big enough to hold the entire message with probability close to 1, so the should have *N* > *M*/2*p* by a sufficient margin. If *M* is bigger than the message, Alice pads with zero bits.
3. Again using the secret key, Alice chooses a random subset of *M* bits out of the *N*-bit string. She substitutes her *M* stego bits for these bits.
4. Alice applies *M* bits of a one-time pad to the stego bits, again using bits from the shared key. To Eve, who lacks the key, these now appear to be random bits.
5. Alice transmits the string to Bob. He pulls out the subset, applies the one-time pad, and reads off the message.

Noisy Case

- If there is noise in the channel, essentially the same protocol is used. However, Alice first encodes her stego bits using an ECC. This is the *inner code*.
- The concatenation of two binary symmetric channels with bit flip probabilities *p* and *q* is still a binary symmetric channel. The resulting bit flip probability is

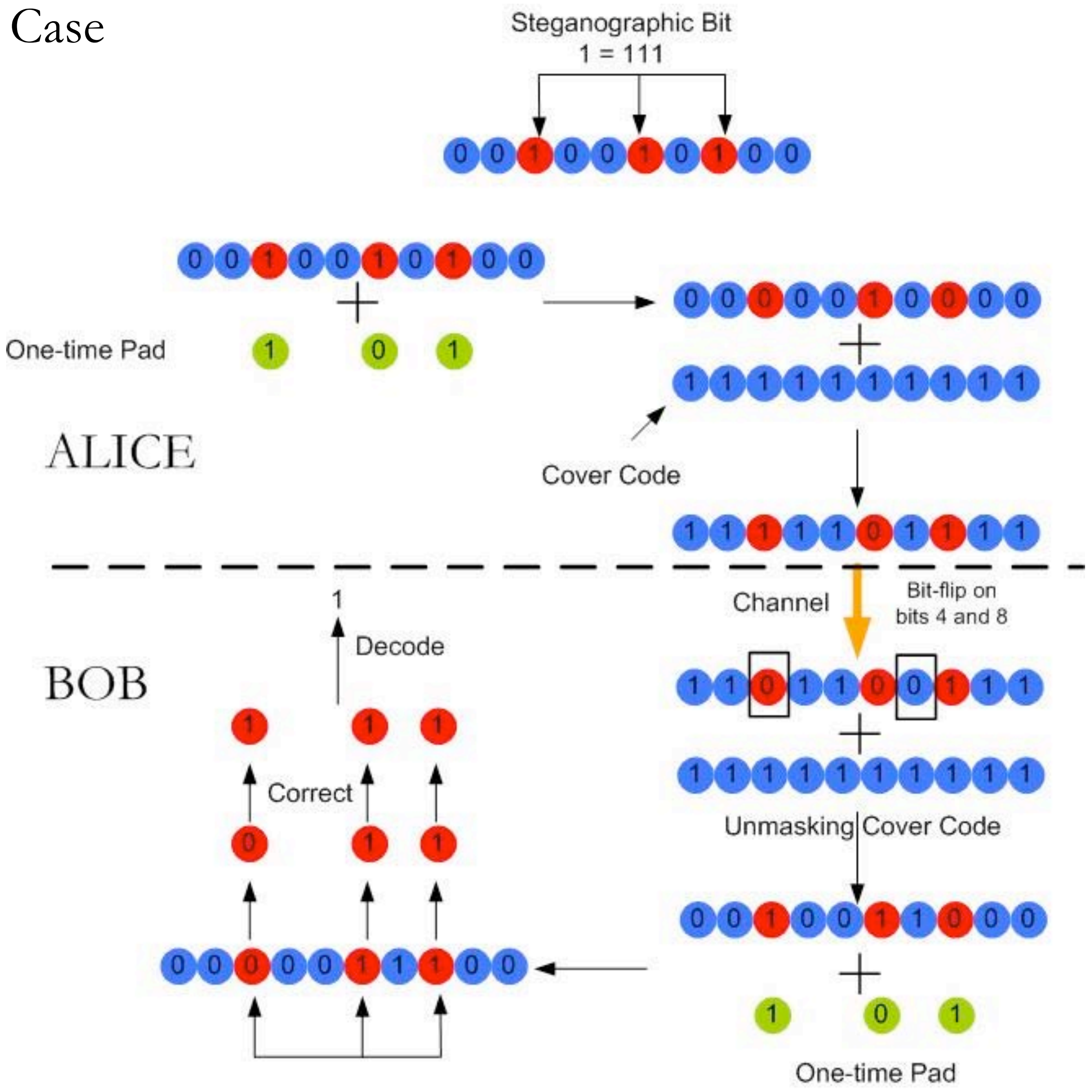
$$p' = p + q - 2pq = p + q(1-2p).$$

- We define the increase in the noise rate over the physical noise rate *p* to be

$$p' = p + \delta p,$$

$$\delta p = q(1-2p).$$

Noisy Case



Steganographic Bit
1 = 111

One-time Pad

ALICE

Cover Code

Channel

Bit-flip on bits 4 and 8

BOB

Decode

Correct

Unmasking Cover Code

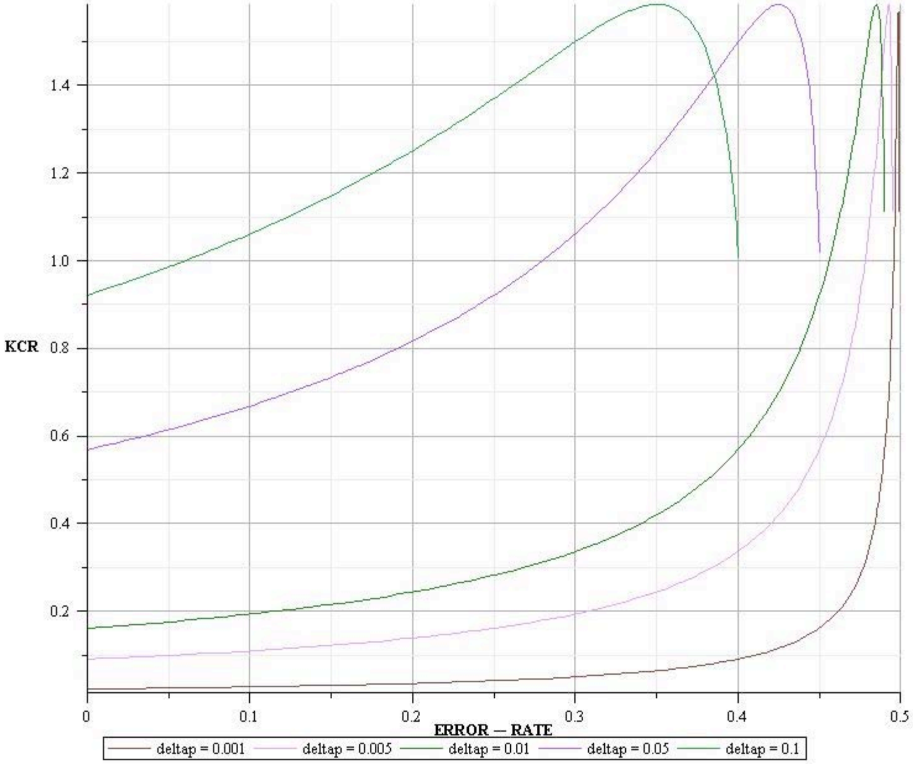
One-time Pad

Key Consumption Rate

$$|\mathcal{K}| = \log \binom{N}{M} + M$$

Number of bits for shared one-time pad

Number of bits to specify the subset that comprises the inner-code



$$|\mathcal{K}| \approx N \log N - M \log M - (N-M)\log(N-M) + M$$

Stirling's Approximation

$$\boxed{M \approx 2Nq \qquad q = \frac{\delta p}{1-2p}}$$

$$\boxed{\mathcal{K}_p^{\delta p} = \left(\frac{2\delta p}{1-2p} \log \frac{\beta}{\left(\frac{\delta p}{1-2p}\right) \beta^{\left(\frac{1-2p}{2\delta p}\right)}} \right)}$$

$$\beta = \frac{1-2(p+\delta p)}{1-2p}$$

Quantum Case

This classical protocol for the binary symmetric channel naturally goes over to the quantum case for the depolarizing channel.

Let’s see how this works.

An operator ρ is the density operator associated to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the conditions:

(1) (Trace Condition) ρ has trace equal to one.

(2) (Positivity Condition) ρ is a positive operator.

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^+$$

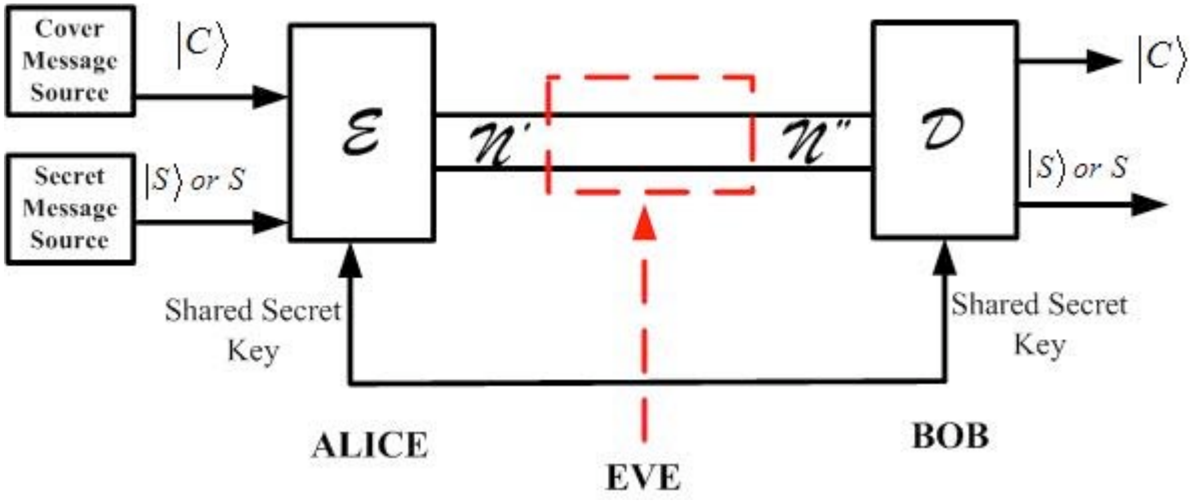
$$\sum_k E_k^+ E_k = I$$

A quantum channel $\mathcal{E}$ is a completely-positive, trace-preserving map.

Complete-Positivity: If $\mathcal{E}$ maps density operators of system Q to density operators of S, then $\mathcal{E}(A)$ must be positive for any positive operator A. Furthermore, if we introduce an extra system R of arbitrary dimensionality, it must be true that $(\mathcal{I} \otimes \mathcal{E})(A)$ is positive for any positive operator A on the combined system RQ, where $\mathcal{I}$ denotes the identity map on system R.

Trace-Preserving:

$$Tr(\mathcal{E}(\rho)) = 1 = Tr(\rho), \forall \rho$$



Bloch Sphere

$$|+\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

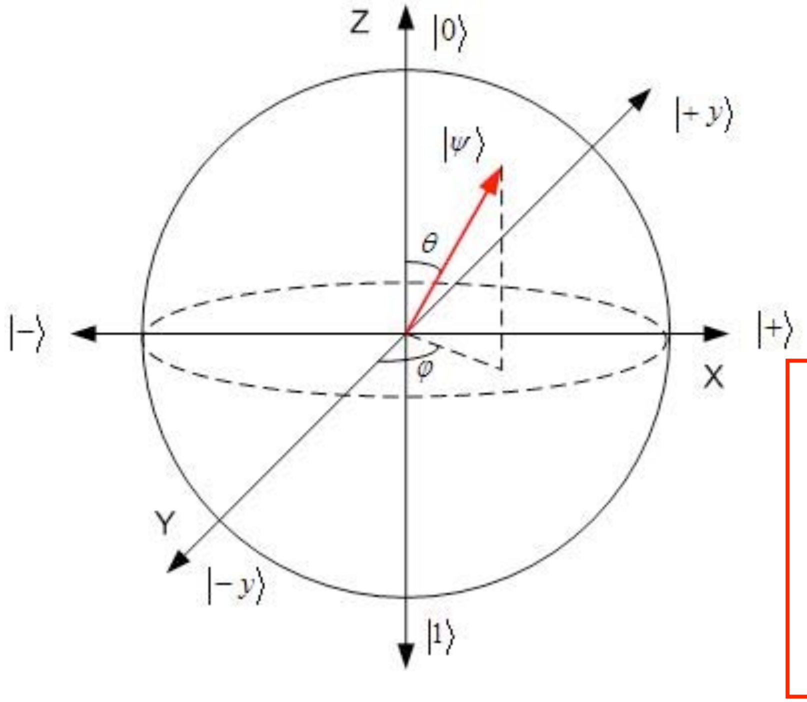
$$|-\rangle \equiv \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$|+y\rangle \equiv -i|0\rangle + |1\rangle$$

$$|-y\rangle \equiv i|0\rangle + |1\rangle$$

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$|1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$$I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

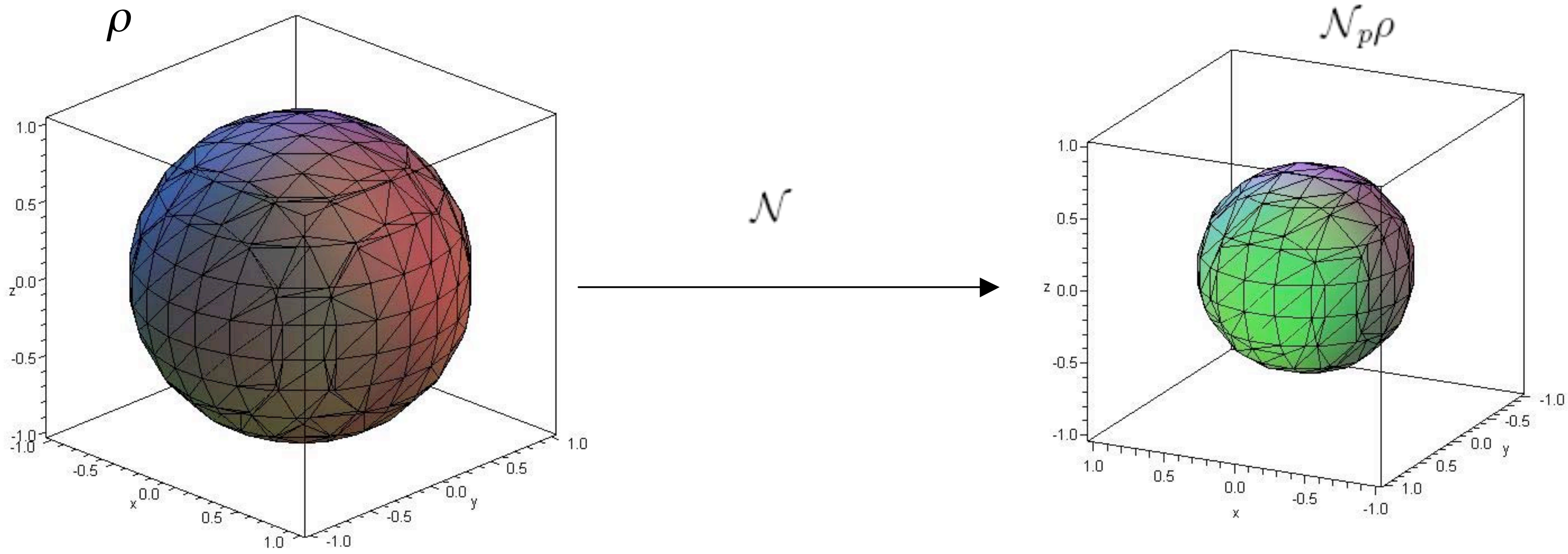
Pauli Matrices

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$

$$0 \le \theta \le \pi$$

$$0 \le \varphi < 2\pi$$

Depolarizing Channel



$$\mathcal{N}_p\rho \equiv (1-p)\rho + \frac{p}{3}\Big(X\rho X + Y\rho Y + Z\rho Z\Big)$$

$$\mathcal{T}\rho \equiv \frac{1}{4}\Big(\rho + X\rho X + Y\rho Y + Z\rho Z\Big)$$

$$I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Pauli Matrices

$$\mathcal{N}_p\rho = \Big(1 - \frac{4}{3}p\Big)\rho + \frac{4}{3}p\mathcal{T}\rho$$

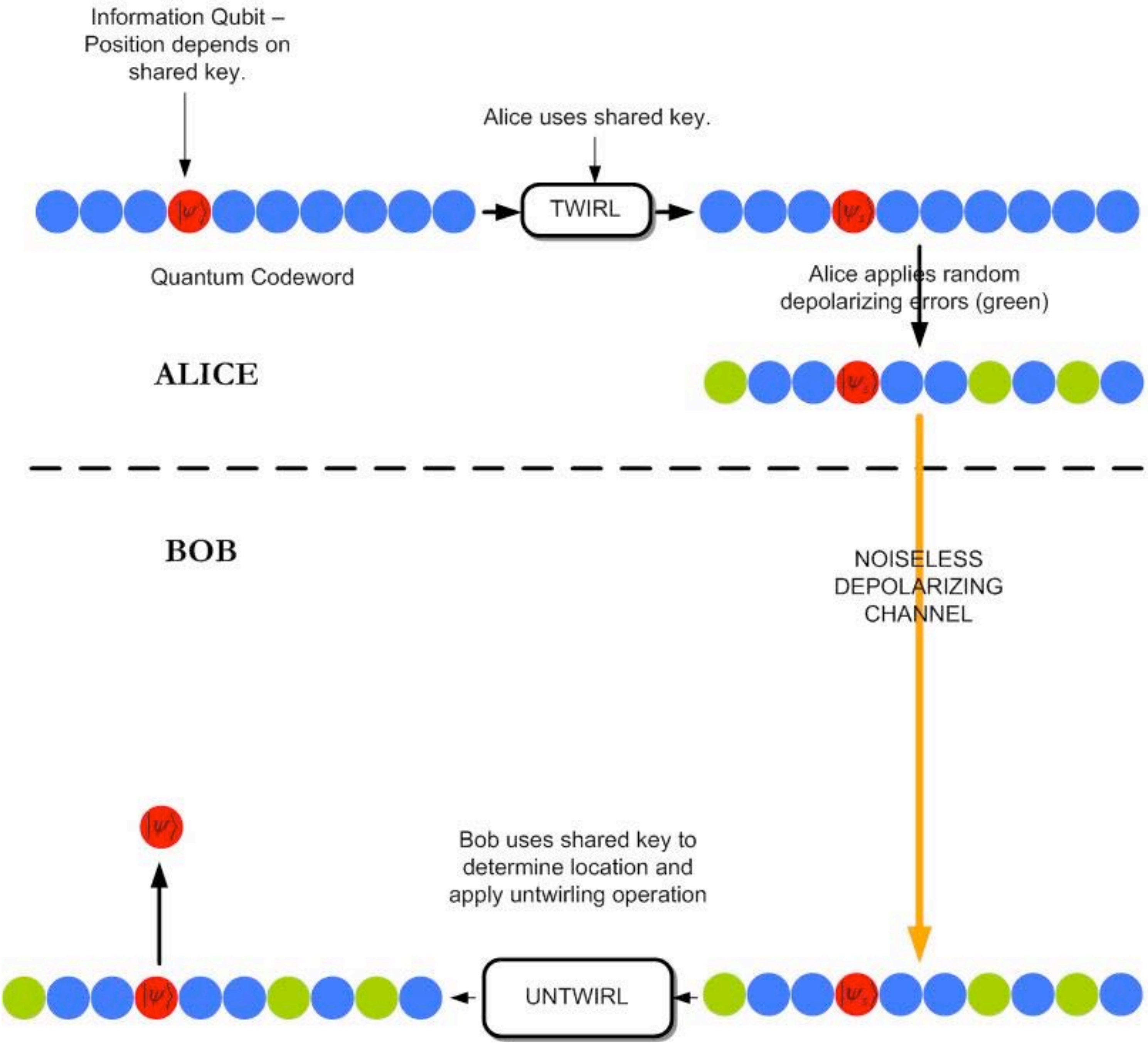
Twirled qubit
masquerading as a
depolarizing error

$$\mathcal{N}_p\rho \equiv (1-r)\rho + r\mathcal{T}\rho$$

Quantum Protocol

1. Alice and Bob choose an [[*N,k*]] QECC to encode *k* qubits of covertext.
2. Using the secret key, Alice and Bob determine the number *M* of maximally mixed bits using the binomial expansion with probability 4*p*/3.
3. Again using the secret key, Alice chooses a random subset of *M* qubits out of the *N*-qubit string. She substitutes her *M* stego qubits for these.
4. Alice applies 2*M* bits of a one-time pad to "twirl" the stego qubits, again using bits from the shared key. To Eve, who lacks the key, these now appear to be maximally mixed qubits.
5. Alice transmits the string to Bob. He pulls out the subset, applies the one-time pad, and recovers the message.

If the channel contains intrinsic noise, Alice first protects her stego qubits in a QECC (the *inner code*).


Information Qubit –
Position depends on
shared key.
Alice uses shared key.
$|\psi\rangle$
TWIRL
$|\psi_s\rangle$
Quantum Codeword
Alice applies random
depolarizing errors (green)
ALICE
$|\psi_s\rangle$
BOB
NOISELESS
DEPOLARIZING
CHANNEL
$|\psi\rangle$
Bob uses shared key to
determine location and
apply untwirling operation
$|\psi\rangle$
UNTWIRL
$|\psi_s\rangle$

Effects of Monitoring

- In the classical case, it was assumed that Eve was a passive observer--she monitored the channel, but did not actively alter what flowed through it. In the quantum case, however, measuring the channel collapses the state. Monitoring is active!
- Alice and Bob can therefore only share quantum information if Eve doesn't always measure the syndromes. For instance, if she measures only 1 codeword in *n* at random, Alice and Bob can spread their quantum information out over multiple codewords and use a QECC to correct any corruption due to Eve's measurements.
- In another scheme, Alice and Bob share EPR pairs rather than a classical random key. In this case, they can send quantum information by teleportation, using stego qubits only to transmit classical information (which is robust against Eve's measurements).

How innocent is “innocent enough?”

- Suppose that Eve knows the physical channel error rate to be *p*. Can Alice and Bob still send information? If δp is “sufficiently small,” Eve will not be able to detect the difference.
- We can bound her ability using the diamond norm:

$$\left\| \mathcal{N}_{p+\delta p}^{\otimes N} - \mathcal{N}_{p}^{\otimes N} \right\|_{\Diamond} \leq \delta p \sqrt{\frac{N}{p(1-p)}}, \qquad \delta p << p.$$

- So Alice and Bob can send an arbitrarily large amount of information if they spread it over a sufficiently large covertext.
- If Eve’s knowledge of the channel is imperfect (as in practice it must be), Alice and Bob can communicate at a finite rate by choosing a small enough δp.

Another approach: hiding information in error syndromes

- While the protocols described before are very simple and elegant, they do not approach the amount of information Alice and Bob *could* send (at least in the noise-free case).
- Instead of hiding information in a random subset of *M* bits, Alice could encode information in *all possible* strings of weight *M*/2. The difference is *M* versus

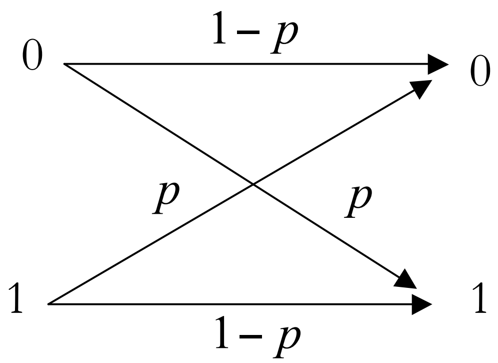
$$\log_2\binom{N}{M/2} \sim (M/2)\log_2(N).$$

- We can think of this as hiding information in the error syndromes rather than the errors themselves.
- This approach will probably also generalize to a broader class of channels.

CLASSICAL STEGANOGRAPHY

Three-Bit Repetition Code – Noiseless Channel

Error	Corrupted Codeword	Syndrome	Probability of Error	Probability
No Error	000	s_0	$(1-p)^3+p^3$	p_0
Error on bit 1	001	s_1	$p(1-p)^2+p^2(1-p)$	p_1
Error on bit 2	010	s_2	$p(1-p)^2+p^2(1-p)$	p_2
Error on bit 3	100	s_3	$p(1-p)^2+p^2(1-p)$	p_3



Binary Symmetric Channel with bit-flip rate $0 < p < 1/2$.

A key-set S that contains bit-strings distributed uniformly.

A generating function f takes input key-bits and generates an ordered pair (j, k), with a non-uniform probability q_{jk} where j, k take values from the set {0, 1, 2, 3} that correspond to syndromes.

Eve is a passive observer.

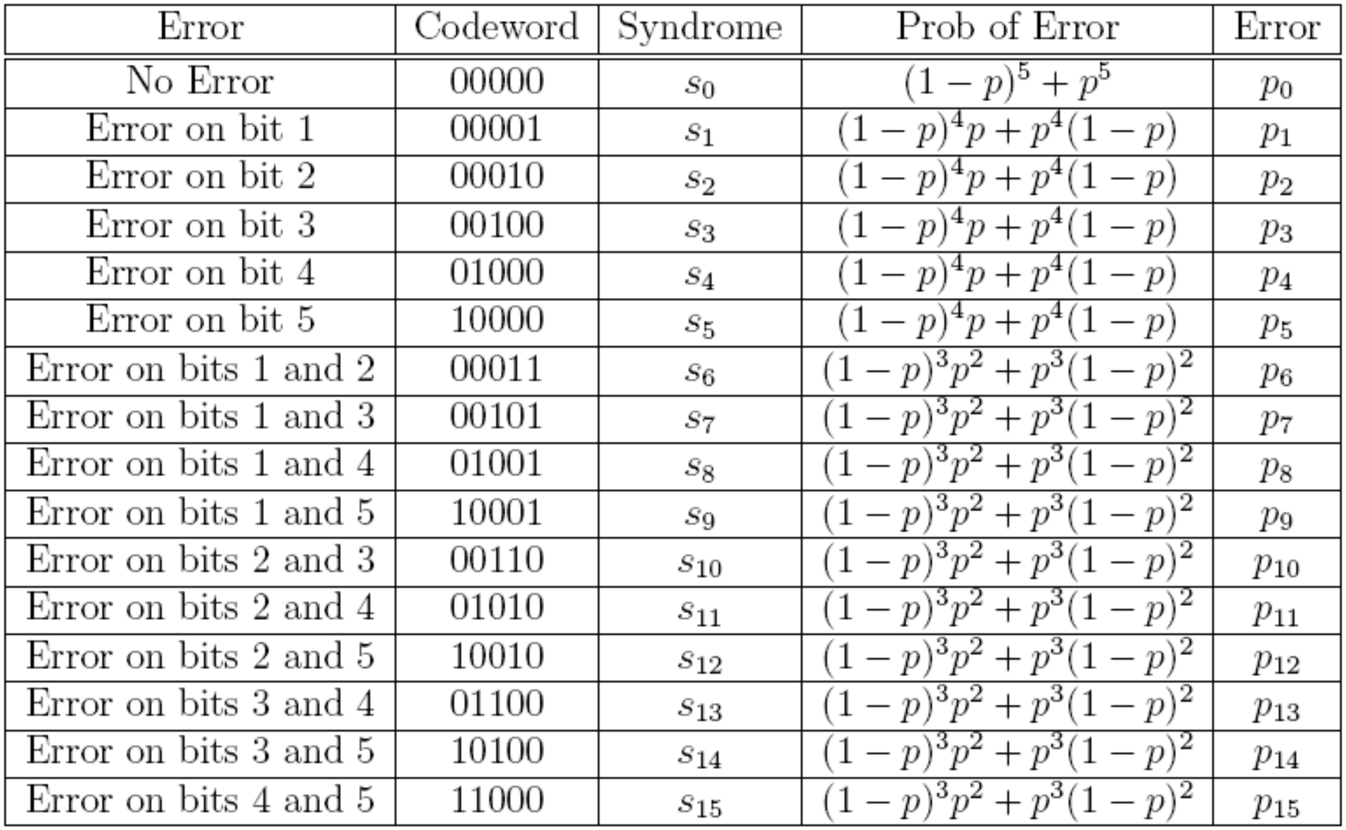
Alice wants to transmit a single steganographic bit to Bob.

$$\left.\begin{array}{l} 0 \rightarrow s_j \\ 1 \rightarrow s_k \end{array}\right\} \quad q_{jk}$$

$$\left.\begin{array}{l} 0 \rightarrow s_k \\ 1 \rightarrow s_j \end{array}\right\} \quad q_{kj}$$

Five-Bit Repetition Code

[5, 1, 5]



Error	Codeword	Syndrome	Prob of Error	Error
No Error	00000	s_0	$(1-p)^5 + p^5$	p_0
Error on bit 1	00001	s_1	$(1-p)^4 p + p^4(1-p)$	p_1
Error on bit 2	00010	s_2	$(1-p)^4 p + p^4(1-p)$	p_2
Error on bit 3	00100	s_3	$(1-p)^4 p + p^4(1-p)$	p_3
Error on bit 4	01000	s_4	$(1-p)^4 p + p^4(1-p)$	p_4
Error on bit 5	10000	s_5	$(1-p)^4 p + p^4(1-p)$	p_5
Error on bits 1 and 2	00011	s_6	$(1-p)^3 p^2 + p^3(1-p)^2$	p_6
Error on bits 1 and 3	00101	s_7	$(1-p)^3 p^2 + p^3(1-p)^2$	p_7
Error on bits 1 and 4	01001	s_8	$(1-p)^3 p^2 + p^3(1-p)^2$	p_8
Error on bits 1 and 5	10001	s_9	$(1-p)^3 p^2 + p^3(1-p)^2$	p_9
Error on bits 2 and 3	00110	s_{10}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{10}
Error on bits 2 and 4	01010	s_{11}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{11}
Error on bits 2 and 5	10010	s_{12}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{12}
Error on bits 3 and 4	01100	s_{13}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{13}
Error on bits 3 and 5	10100	s_{14}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{14}
Error on bits 4 and 5	11000	s_{15}	$(1-p)^3 p^2 + p^3(1-p)^2$	p_{15}

Noise Model for Binary Symmetric Channel with Error-Rate p.

Classical (Noiseless) Protocol

Alice and Bob are simulating a channel with entropy *s*. For a codeword of length *N*>>1, the channel will be dominated by 2^{Ns} “typical errors.” These errors can be enumerated $\{e_j\}$ for j=0,…, 2^{Ns}-1. (These can be approximated as equally likely.)

1. Alice encodes the covertext in an [*N,k*] ECC. This code has distinct syndromes s_j for each of the typical errors e_j.
2. Alice’s stego text is ~*Ns* bits. Alice applies a one-time pad to these bits to get a random string of bits. These are interpreted as representing a random number *j* in binary form.
3. Alice applies error e_j to the codeword.
4. Alice transmits the codeword to Bob. If Eve examines it, she will find a typical error.
5. Bob decodes and finds the syndrome s_j. He maps s_j to *j*, applies the one-time pad to *j*, and reads the message.

Quantum (Noiseless) Protocol

- Alice and Bob are simulating a quantum channel with entropy *s*. The channel will be dominated by 2^{Ns} “typical errors.” These errors are represented by operators $\{E_j\}$ for j=0,…, 2^{Ns}.
- Alice will encode the covertext in an $[[N,k]]$ QECC. Assume (for now) this code is nondegenerate, and hence has distinct syndromes s_j for each of the typical errors E_j.
- It is possible to choose the encoding and decoding unitaries for this code such that

$$U_D = U_E^{-1}, \qquad |\Psi\rangle = U_E\left(|\psi\rangle \otimes |0\rangle^{\otimes N-k}\right),$$

$$U_D\left(E_j|\Psi\rangle\right) = |\psi\rangle \otimes |s_j\rangle.$$

Quantum Protocol (Continued)

1. Alice prepares *k* qubits of covertext in a state $|\psi_c>$.
2. Alice's stego text is ~*Ns* qubits, in a state

$$\left|\psi_s\right\rangle = \sum_k \alpha_k \left|k\right\rangle.$$

3. Alice twirls these qubits to get a maximally mixed string, and appends *N*-*Ns*-*k* ancilla qubits in the state |0>, to get a total register of *N*-*k* qubits.
4. Alice applies a unitary U_s to this register that maps

$$U_s\left(\left|j\right\rangle \otimes \left|0\right\rangle^{\otimes N-Ns-k}\right) = \left|s_j\right\rangle.$$

5. Alice now applies the encoding unitary U_E to the covertext plus the register. The state will look to Eve like

$$\rho = 2^{-Ns} \sum_k E_j \left|\Psi_c\right\rangle\left\langle\Psi_c\right| E_j^t.$$

6. Alice transmits the codeword to Bob. Bob applies the decoding unitary U_D and the inverse of U_S, and discards the covertext and ancilla qubits. Using the shared key, he undoes the twirling operation and recovers the stego qubits.

The Noisy Case

- It is clear that this encoding will fail if the channel contains intrinsic noise. In the plausible case where $p \gg \delta p$, the typical errors used by Alice would be swamped by the noise in the channel.
- This can be countered by some form of error correction. But since we are encoding in syndromes, rather than directly in bits, it is not clear how to design codes that give good performance.
- One approach borrows from our earlier protocol--Alice chooses a random subset of the bits (or qubits) and encodes only in this subset. This allows Alice and Bob to ignore physical errors outside this subset. However, some error correction is still necessary.
- Without knowing the best form of encoding, we can only roughly estimate the best possible rate of stego transmission in the noisy case.

Conclusion and Future Work

- Developed a classical and quantum model of steganography for the binary symmetric and depolarizing channels. The stego bits or qubits appear to Eve to be random errors because she lacks the shared secret key.
- Also developed an alternative formulation encoding information in syndromes. This is certainly more efficient in the noiseless case.
- Calculated the key-consumption rate
- Showed that these protocols have steganographic security if a long enough covertext is used.

- For a general channel, what encoding wil stay closest to what Eve expects?
- What is the capacity of quantum steganographic channels with noise?
- If Bob knows the covertext ahead of time, can additional information be sent?
- What optimal measurements can Eve perform to gain knowledge of a quantum steganographic channel? Can we prove that if Eve knows the channel exactly, Alice and Bob cannot communicate at a finite rate?
- For which quantum error-correcting codes do we get the best rates?

Capacity of Steganographic Channels

Jeremiah J. Harmsen, *Member, IEEE*, and William A. Pearlman, *Fellow, IEEE*

***Abstract*—This work investigates a central problem in steganography, that is: How much data can safely be hidden without being detected? To answer this question, a formal definition of steganographic capacity is presented. Once this has been defined, a general formula for the capacity is developed. The formula is applicable to a very broad spectrum of channels due to the use of an information-spectrum approach. This approach allows for the analysis of arbitrary steganalyzers as well as nonstationary, nonergodic encoder and attack channels. After the general formula is presented, various simplifications are applied to gain insight into example hiding and detection methodologies. Finally, the context and applications of the work are summarized in a general discussion.**

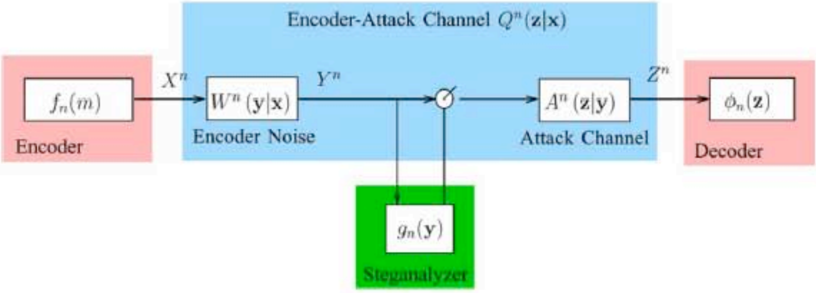


Fig. 1. Steganographic channel.