Scalable One-Time Pad—From Information Theoretic Security to Information Conservational Security

Wen-Ran Zhang

Department of Computer Science
Georgia Southern University, Statesboro, GA USA
wrzhang@georgiasouthern.edu

Abstract. Whereas it is widely deemed an impossible task to scale One-Time Pad (OTP) without sacrificing information theoretic security or network traffic, this paper presents a paradigm of Scalable One-Time Pad (S-OTP) ciphers based on information conservational computing/cryptography (ICC). Applicability of the new paradigm is analysed. It is shown that ICC enables data compression with quantum-fuzzy collective precision to reduce or scale a key length to a minimum that used to be deemed impossible. Based on ICC, it is shown that, with a local IEEE binary64 standard computer associated with quantum key distribution (QKD), S-OTP enables secure transmission of long messages or large data sets with significant traffic reduction for post-quantum cryptography. Quantum crypto machinery is proposed. Some open topics are identified for further investigation.

Keywords: Information Conservational Security; Data Compression or Scaling; Quantum-Fuzzy Collective Precision; Post-Quantum Cryptography; Scalable One-Time Pad

1 Introduction

Cryptography is essential for the security of digital communication. However, many commonly used cryptosystems will be completely broken by a *quantum* algorithm for integer *factorization* [1] once large quantum computers are commercially applicable. Post-quantum cryptography is to counter such quantum attacks and to keep digital communication secure [2]. A key for success is to identify mathematical operations for which quantum algorithms offer little advantage in speed, and then to build cryptographic systems around them. Although progress has been made most proposed methods incur serious costs, especially in network traffic. A major challenge is to reduce encryption key length without increasing data length.

One-Time Pad (OTP) [3] [4] is often regarded the only cipher with proven information theoretic security [5] [6]. It can now be used together with quantum key distribution (QKD)—a well-developed application of quantum cryptography. QKD uses quantum communication to establish a shared key between two parties— sender Alice and receiver Bob. The key is then shared. If a third party Eve tries to eavesdrop on the communication between Alice and Bob, the quantum communication will fail for security protection [8]. Once the key is established, it is typically used as a symmetric key for digital communication such as using OTP. Since OTP is quantum proof to quantum factorization, it is a good candidate for post-quantum cryptography [2]. Unfortunately, the key requirement of equal or greater length than the original message hinders the general application of OTP even though QKD is a well-developed partner technology. As a result, OTP is generally limited at present time to transmitting relatively short messages with high security requirement.

In his classical paper [5], Shannon identified three general approaches to cryptography: (1) concealment systems, (2) privacy systems, (3) "true" secrecy systems where the meaning of the message is concealed by cipher, code, etc. Shannon deemed concealment systems primarily a psychological problem, and privacy systems a technological one. OTP considers only the third type of "true" secrecy systems. History shows that, when Shannon invented OTP in 1946 [6], the first computer was not out yet. Since then, computing technology advanced beyond anyone's imagination. Although it is proven [7] that any cipher with the perfect secrecy property must use keys with effectively the same requirements as OTP keys, these proofs, however, did not take later data compression technological development into consideration that can conceal the meaning of a message. One such development is IEEE binary64 double-precision floating-point format for a wide dynamic range of numeric values. Another is information conservational computing/cryptography (ICC) [9,10] with quantum-fuzzy collective precision. With the new technological advances, information conservational data compression can be incorporated into OTP as an extension to information theoretic security. Thus, *scalable one-time pad* (S-OTP) does not attempt to falsify Shannon's theorem, it bypasses the assumptions of the theorem by reducing the message length with information conservational data compression. Now we have the question: *Can S-OTP achieve what is deemed impossible?*

This paper presents S-OTP for efficient use of QKD on a local computer. It makes OTP practically applicable with much shorter keys for transmitting compressed long messages or large data sets without increasing network traffic for quantum-proof digital communication. Different versions of S-OTP ciphers are analyzed and compared. Information conservational security conditions are established. Collective precision is proposed. Quantum machinery development of S-OTP is briefly discussed.

This paper is organized in five sections. Section II presents the theoretical basis with illustrations on ICC. Section III examines the applicability and optimization of the S-OTP paradigm. Section IV presents an architectural design of S-OTP quantum dream machinery. Section V draws a few conclusions.

2 Basic Concepts

2.1 Information Conservational Security

It is shown [3] [4] that an OTP cipher is information theoretically secure and unbreakable [5] [6] provided that the message to be ciphered is unknown to attackers, and a cipher key meets *the four conditions of OTP:* (a) truly random; (b) never reused;

(c) kept secret from all possible attackers; (d) of equal or greater length than the message.

While most ciphers can be broken, no efficient method can be found to break the OTP cipher. Thus, the OTP cipher remains the only theoretically unbreakable one. Based on the four security conditions, we consider the scalability of OTP.

Definition 1(a). *Information conservational transformation* (ICT) is referred to as a set of set-theoretic or information theoretic mathematical functions that forms an transformation T to transform the bit pattern of a long message in form F_1 to a shorter pattern in form F_2 systematically such that there exists a reverse transformation T' that recovers F_1 from F_2. Formally we have: $T(F_1)$: $F_1 \rightarrow F_2$ such that $\exists T'$ and $T'(F_2)$: $F_2 \rightarrow F_1$.

Definition 1(b). *Scalability* is referred to as using ICT once or multiple times systematically to transform a long message or large data set into one or a series of short forms such that cipher keys are reduced to practical lengths or to a minimum for enciphering the short forms as OTP pads for secure transmission. In this case, An OTP pad is called a scalable OTP (S-OTP) pad. In S-OTP, a key is assumed reusable if the reuse can be concealed in another unbreakable S-OTP pad.

Based on Definitions 1a and 1b, we extend *information theoretic security* (ITS) of OTP to *information conservational security* (ICS) of S-OTP.

Definition 2a. An S-OTP cipher is said having information conservational security (ICS) provided that: (a) The key length required is significantly shorter than the original message due to ICT; (b) The shorter key does not weaken the ITS of OTP; (c) It must reduce network traffic.

Definition 2b. Given $0<i<N$, a *minimum length form* is a message form $F_x = (X, \{x_i/X\})$ that cannot be further reduced in binary length through ICT in theory. An *absolute minimum length form* is the minimum form when N = 2.

It could be argued that S-OTP is just OTP plus data compression, and there is nothing new. The author wishes it is a valid argument. However, the reality is that: (1) Information conservation or preservation has been a long sought goal in physics and information theory [11,12], and it is essentially information theoretic in nature; (2) The key length problem has been a well-known long standing impasse; (3) ICS is a systematic extension to ITS to scale down the length of a message to be enciphered [9,10].

The inception of ICS accounts for the new development in computing technology. Double precision floating-point numbers of IEEE binary64 is used as a technological basis.

Theorem 1a (Possibility Theorem). ICT for ICS is possible based upon OTP and IEEE binary64. Formally, let $\{x\}$ be the data set of a long integer L representing a sufficiently long message divided into sections, some set $\{x_i\}$, $0<i<N$, exists such that $(X, \{x_i/X\})$ is significantly shorter than the long integer L, where $X = \sum_i x_i$ is a math summation (not XOR), and $\{x_i/X\}$ a percentage distribution.

Proof. *To show possibility, let L=16K bits divided into 32 of 512-bit sections $\{x_i\}$. That leads to one 64-bit double precision floating point summation $X=\sum_i x_i$ and 32 of 64-bit percentage distributions $\{x_i/X\}$ $0<i<32$, total 2K+64 bits vs. 16K, a nearly 8-fold reduction of key length and network traffic. This could be further hierarchically scaled (reduced) to a minimum or until the pair $(X, \{x_i/X\})$ is short enough to be randomized and enciphered with a significantly shorter key. Thus, ICS of S-OTP can be equivalent to ITS in terms of un-breakability, provided computational precision is satisfied.* (Note: As a possibility theorem, not an optimality theorem, a generic example is sufficient for its proof. However, the possibility opened the door to the scalability of OTP.) ■

Theorem 1b (Minimum Length Theorem). For any message in an intermediate ICT tranformation, given the number of data divisions N, *a minimum length* form $F_x = (X, \{x_i/X\})$ exists in theory where $length(X) = length(x_i/X), 0<i<N$. An *absolute minimum length* also exists in theory, where length(X) = length(three double precision floating point numbers), that would be 192 bits for IEEE binay64, assuming double precision floating point format.

Proof. *Given any ICT T, we must have a form in length $F_x = (X, \{x_i/X\}), 0<i<N$, such that at certain point we must have $T(F_x):F_x \to F_y$ and $length(F_y) = length(F_x)$ because (1) given $0<i<N$, $length(\{x_i/X\})$ is irreducible; (2) if $length(X) \leq$ length(one double precision floating point number) it becomes irreducible either. Then we must have an absolute minimum length for N, where one double precision floating point number is for the summation, two are for the distribution double precision floating point format, that is 64 bits each and 192 bits for three of them assuming IEEE binary64 standard.* ■

Theorem 1c (Reachability of Minimum). Given the number of data divisions N≥2, a minimum or absolute minimum length form can be reached through a recursive ICT transformation

Proof. *It follows from the proof of Theorem 1b.* ■

2.2 Method1: Add, Divide, and Conquer

The rationale of S-OTP is that, given an unsigned big integer *L* representing the long message or large data item *D* to be transmitted, *L* can be divided into a set of shorter long integers $\{x_i\} = x_1, x_2, .., x_i, .. x_n$ representing sectors or sections of *D* to be transmitted. The summation $X = x_1 + x_2 + ... + x_i + ... + x_n$ can be obtained which could be represented as a long integer or a floating-point decimal much shorter than *L* to transmit. The percentage distribution $\{x_i/X\}$ is a type of most primitive information conservational key that can be encrypted and transmitted together with *X* in ciphertext for recovering $\{x_i\}$ to *L* and then *D* in the receiver side. This leads to S-OTP-Method1—a one key cipher.

<u>S-OTP1-Method1</u>

Assume sender Alice and receiver Bob share a private key K distributed through QKD.

<u>Part I. Encryption</u>

Step 1. Let math summation $X = \sum_i x_i$ (not XOR).
Step 2. Calculate percentage distribution $\{x_i/X\}$;
Step 3. Encrypt the text $U = \{X, \{x_i/X\}\}$ with one key *K* to ciphertext $E=K \oplus U$ where $\oplus$ is *XOR* (not math summation).
Step 4. Alice Transmits *E* to Bob.

<u>Part II. Decryption</u>

Step 1. Use *K* to decipher *E* to obtain *X and* $\{x_i/X\}$;
Step 2. Use $\{x_i/X\}$ to decrypt the summation *X* and recover $\{x_i\}$;
Step 3. Recover transmitted message from $\{x_i\}$ with concatenation.

2.3 Illustration of S-OTP1-Mehrod1

Assuming the plaintext data *D* to be transmitted is represented by the big integer *L = 1048549998213983988*, we divide *L* into the three sections *1048549, 998213, and 983988.* Assume sender Alice and receiver Bob share a private key K distributed through QKD.

Part I - Encryption

(1) Let $x_1 = 1048549$, $x_2 = 998213$, $x_3 = 983988$, and
(2) $X = x_1 + x_2 + x_3 = 3030750$;
(3) Calculate percentage distribution $\{x_i/X\}=\{34.5970\%, 32.9362\%, 32.4668\%\}$;

(4) Encrypt the plaintext U={3030750, {34.5970%, 32.9362%, 32.4668%}} to result in ciphertext E=KεU;
(5) Transmit E to Bob;

Part II - Decryption:

(1) Use K to recover U={3030750, {34.5970%, 32.9362%, 32.4668%}};
(2) Use U to recover x_1=1048549, x_2=998213, and x_3=983988;
(3) L = concatenate(x_1, x_2, x_3) = 1048549998213983988;
(4) Recover D from L.

It can be argued that S-OTP$_1$-Mehrod1 does not reduce the key length. The counter argument is that, as a simple example the illustration is already in a minimum form. It does not really need S-OTP because a single OTP is sufficient. As proven in Theorem 1, S-OTP does reduce key length the same way with sufficiently long dada sections. The question is: could there be a precision problem?

2.4 Method2—ICC with Collective Precision

Percentage distribution has its own limitation due to sequential computation. When the math summation gets huge that is usually the case, the precision of a single percentage will be a problem. The computation of such a percentage can be avoided with the massive parallel collective precision property of ICC [9,10]. In ICC a big total can be divided into many subtotals or integers representing data sections. If each subtotal is further divided into bipolar import-export values, each value can be normalized by its corresponding column subtotal. A information conservational matrix can then be derived through column-major normalization for massive parallelism and collective precision without using the grand total.

ICC is made achievable with bipolar fuzzy sets [13-17]. Bipolar fuzzy set theory forms an equilibrium-based mathematical abstraction—a set theoretic or information theoretic extension to fuzzy set theory [18]. It is a generalization of truth-based computing which can still be used freely as long as equilibrium conditions are not violated. Bipolar fuzzy set theory was once rescued by Zadeh [19].

In this subsection, we show an ICC example. We then examine and explain the properties of the example in next two subsections. A key concept in ICC is an information conservational bipolar matrix *M*. With *M* an energy or information total or summation can be decrypted through equilibrium-based rebalancing to result in all the subtotals in parallel with percentage distribution in collective precision built into *M*. This makes it possible to develop digital or quantum machinery with massive parallelism that is not achievable with linearly normalized percentage distribution.

M consists of bipolar elements. The energy and/or information of a bipolar (import-export) element or variable $x = (a, b)$ is defined as the length of a bipolar interval where a is negative and b positive.

Energy of x: $\varepsilon|x| = \varepsilon|(a, b)| = b - a = |a| + |b|$. (1)

For instance, $\varepsilon|(-2.5, 3.5)|=3.5-2.5=2.5+3.5=6$.

A 3-partner US-China-EU trade example is used to illustrate the basic idea of ICC with collective precision. First, the 3-parners' bipolar import-export data for 2014 are shown in **Fig. 1a** as a cognitive map (CM) in million Euros. The total energy/information in the trade scenario is characterized by the total import/export

$\varepsilon|(-3030750, +0)| = \varepsilon|(-0, +3030750)| = 3030750$.

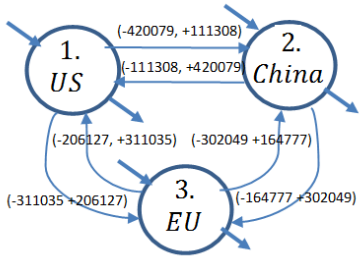
Using collective bipolar interaction in ICC, accurate calculation can be carried out with the bipolar quantum cellular automaton (BQCA) $E(t+1) = M \times E(t)$ based on a column-major normalized bipolar cognitive map matrix M that does not need the calculation of percentage distribution. (Note: The illustrations in this paper are in fix-point format for readability vs. floating-point format.)

In this ICC example $E(1)$ is the transpose of the initial bipolar column vector with certain total energy/information. A cognitive map (CM) C is referred to as a bipolar or unipolar conceptual graph or an import/export network. M is obtained based on column-major normalization of an i/o-consistent interactive CM in which all elements are directly or indirectly interrelated. In this example,

$$C(t) = \begin{bmatrix} (0,0) & (-420{,}079, +111{,}308) & (-311{,}035, +206{,}127) \\ (-111{,}308, +420{,}079) & (0,0) & (-164{,}777, +302{,}049) \\ (-206{,}127, +311{,}035) & (-302{,}049, +164{,}777) & (0,0) \end{bmatrix}.$$

$$M = normalize(C^T(t)) = \begin{bmatrix} (0.000\ 0.000) & (-0.112\ 0.421) & (-0.209\ 0.316) \\ (-0.401\ 0.106) & (0.000\ 0.000) & (-0.307\ 0.167) \\ (-0.297\ 0.197) & (-0.165\ 0.303) & (0.000\ 0.000) \end{bmatrix}.$$

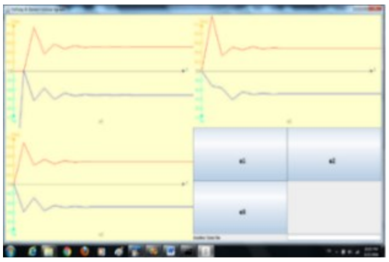
Equilibrium-based rebalancing is illustrated in **Fig. 1b** and curved in **Fig. 1c**. **Fig. 1d** verifies equilibrium-based rebalancing with sequential computing. **Fig. 1e** shows 200% is balanced to a perfect percentage distribution built in M. Thus, matrix M can be deemed the encryption of a percentage distribution. With M any total information such as $\varepsilon|(-3030750, 0)| = 3030750$ or $\varepsilon|(-100, +100)| = 200$ can be rebalanced.



(a)

t	E(t+1) = M(t) × E(t) e1	e2	e3
1	(-3030750.000, 0.000)	(-0.000, 0.000)	(-0.000, 0.000)
2	(0.000, 0.000)	(-321727.188, 1214205.875)	(-595794.188, 899022.625)
3	(-647442.250, 955835.875)	(-375738.375, 333436.250)	(-297782.688, 420514.375)
4	(-377520.813, 377520.813)	(-630613.250, 522677.375)	(-579544.938, 542872.625)
5	(-620577.625, 583280.625)	(-455013.250, 460129.250)	(-463296.094, 448453.000)
6	(-483180.125, 483180.125)	(-514797.813, 527851.500)	(-508652.563, 513087.688)
7	(-543767.625, 548278.250)	(-487545.188, 486926.438)	(-481218.563, 483013.688)
8	(-512763.688, 512763.688)	(-506231.313, 504652.594)	(-497437.406, 496901.063)
9	(-530639.875, 530094.375)	(-495690.813, 495765.688)	(-489388.063, 489170.938)
10	(-521049.438, 521049.469)	(-500810.844, 501001.781)	(-493386.656, 493451.531)
11	(-525948.625, 526014.563)	(-498152.719, 498143.656)	(-491231.906, 491258.156)
12	(-523370.719, 523370.719)	(-499627.813, 499604.750)	(-492391.719, 492383.844)
13	(-524756.813, 524748.875)	(-498834.719, 498835.813)	(-491788.219, 491785.063)
14	(-524021.125, 524021.125)	(-499249.156, 499251.906)	(-492102.594, 492103.531)
..	..	..	..
29	(-524274.313, 524274.313)	(-499106.281, 499106.281)	(-491993.781, 491993.781)
30	(-524274.250, 524274.250)	(-499106.281, 499106.281)	(-491993.781, 491993.781)
31	(-524274.250, 524274.250)	(-499106.250, 499106.250)	(-491993.750, 491993.750)
32	(-524274.250, 524274.250)	(-499106.250, 499106.250)	(-491993.750, 491993.750)

(b)



(c)

Partner	Import-Export of 2014	Total Volumes	Percentage (%)
US	(-731114 +317435)	1048549	1048549/3030750 = 34.5970%
China	(-276085 +722128)	998213	998213/3030750 = 32.9362%
EU	(-508176 +475812)	983988	983988/3030750 = 32.4668%
Total	(-1515375 +1515375)	3030750	100.0000%

(d)

t	E(t+1) = M(t) × E(t) Rebalancing 100 Percentage to Actual Distribution US	China	EU
1	(-100.0000, 100.0000)	(-0.0000, 0.0000)	(-0.0000, 0.0000)
2	(0.0000, 0.0000)	(-50.6783, 50.6783)	(-49.3217, 49.3217)
3	(-52.9004, 52.9004)	(-23.3993, 23.3993)	(-23.7003, 23.7003)
4	(-24.9127, 24.9127)	(-38.0530, 38.0530)	(-37.0343, 37.0343)
5	(-39.7215, 39.7215)	(-30.1952, 30.1952)	(-30.0833, 30.0833)
6	(-31.8852, 31.8852)	(-34.4024, 34.4024)	(-33.7125, 33.7125)
7	(-36.0322, 36.0322)	(-32.1528, 32.1528)	(-31.8150, 31.8150)
..	..	..	..
22	(-34.5969, 34.5969)	(-32.9362, 32.9362)	(-32.4669, 32.4669)
23	(-34.5971, 34.5971)	(-32.9361, 32.9361)	(-32.4668, 32.4668)
24	(-34.5970, 34.5970)	(-32.9362, 32.9362)	(-32.4668, 32.4668)
25	(-34.5970, 34.5970)	(-32.9362, 32.9362)	(-32.4668, 32.4668)

(e)

Figure 1. (a) Bipolar CM of 2014 US-China-EU trade (in Million Euros); (b) Rebalancing of total import/export to an equilibrium state; (c) Curves of the rebalancing; (d) Digital computing; (e) Quantum-fuzzy rebalancing of 200%

While sequential computing does not support parallel processing, equilibrium-based rebalancing can balance any total information to a perfectly equilibrium state with percentage distribution coded in M in a single iterative and massively parallel process without the need for the calculation of individual percentages. Although a perfect equilibrium-state is neither practical nor desirable, equilibrium-based rebalancing provides an information conservational approach to post-quantum cryptography described as S-OTP$_1$-Method2. Most importantly, it finds a way for collective precision.

S-OTP$_1$-Method2

Assume key K_1 is shared by sender Alice and receiver Bob through QKD.

Part I. Encryption

Step 1. Data Transformation. *Given binary data D to be transmitted, let the unsigned integer number set $\{d_i\} = \{d_1, d_2, .., d_i, .., d_n\}$, represent the data sections of D. Let the sum $X=d_1+d_2+...+d_i ...+ d_n$.*

Step 2. Bipolar Cognitive Mapping. *Construct an i/o-consistent BCM C based on $\{d_i\}$ such that $\{d_i\}$ is decomposed into an unbalanced relational data set $\{e_{ij}\} = \{(e_{ij}^-,e_{ij}^+)\}$ where each bipolar link weight $e_{ij} = (e_{ij}^-,e_{ij}^+)$ and $|d_i| \equiv \sum_j |\varepsilon| e_{ij}$ (energy/information of row i) with ratio $|e_{ij}^-|/|e_{ij}^+| > l$, a threshold for non-zero bipolar elements. Thus, $\{e_{ij}\}$ forms a BCM C with total information $X=\sum |d_i|$. (Note: C is not unique – an area of further research where bipolar linguistic fuzzy sets can be used for the optimization of l and C.)*

Step 3. Bipolar Energy/Information Normalization. *Normalize C^T (transpose of C) to an information conservational matrix M (a bipolar quantum-fuzzy logic gate (BQFLG) or a bipolar quantum-fuzzy cognitive map (BQFCM)) under the conditions of Eq. (3) such that the BQCA $E(t+1)=M \times E(t)$ is asymptotic to an equilibrium state [10,11].*

Step 4. Data Encryption. *Use K_1 to encipher $U=\{X,M\}$ to $E=U\oplus K_1 = \{X,M\}'$.*

Step 5. Transmit the pair E=$\{X,M\}'$.

Part II. Decryption

Step 1. *Use K_1 to decrypt E to $\{X,M\}$; use K_2.*

Step 2. *Use M to decipher and depolar X to recover $\{d_i\}$;*

Step 3. *Recover D from $\{d_i\}$ with concatenation.*

Applying S-OTP$_1$-Method2 we have the decryption example in **Fig. 1**. The total information of the last row of **Fig. 1b** approximate to exactly the same result as that of S-OTP$_1$-Mehrod1:

$d_1=|\varepsilon|(-731114, +317435)=1048549$;
$d_2=|\varepsilon|(-276085, +722128)=998213$;
$d_3=|\varepsilon|(-508176, +475812)=983988$
$D=Concatenate(d_1, d_2, d_3)=1048549998213983988$.

2.5 The Nature of Information Conservation

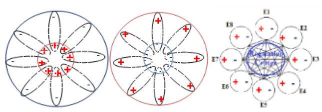
Given an $n\times n$ square bipolar interactive matrix M and an $n\times 1$ column bipolar vector $E(t)$ such that $E(t+1)=M \times E(t)$, if $\forall j$, the absolute energy/information subtotal $|\varepsilon_{col}|M_{*j}(t)$ of each column j of M (but not necessarily each row) equals 1.0, or $|\varepsilon_{col}|M_{*j}(t)\equiv 1.0$, M is defined as an information conservational bipolar quantum logic gate (BQLG) matrix or a bipolar quantum-fuzzy cognitive map (BQFCM) [9,16], and we must have the bipolar quantum cellular automata (BQCA):

$$|\varepsilon|E(t+1) = |\varepsilon|(M \times E(t)) \equiv |\varepsilon|E(t). \quad (2)$$

Eq. (2) leads to a general-purpose BQCA theory – an equilibrium-based unification of matter and antimatter for ICC. Computationally, a BQCA can be regulated to achieve information conservation, regeneration, degeneration and oscillation. BQCA is thus a type of quantum-cellular model (**Fig. 2**). This leads to Method2 and the theory of ICC.

The transpose $C^T(t)$ is used to obtain its column-major normalized BQLG matrix *M* for ICC. The normalization follows Eq. (3). Any man-made i/o-consistent CM can always be designed and normalized to *M* for a BQCA to be asymptotic to a bipolar equilibrium state even though some link weights are weaker and need more iterations (*t*) to be balanced. This property provides a basis for quantum and post-quantum cryptography.

$$\boldsymbol{M(i,j) = (C^T(i,j))/|\varepsilon_{col}|(C^T{}_{*j}).} \qquad \textbf{(3)}$$



(a) (b) (c)

Figure 2. A BQCA unification of matter and antimatter atoms (adapted from [11])

In Eq. (3), the denominator $|\varepsilon_{col}|(C^T{}_{*j})$ denotes the absolute energy/information subtotal of column *j* in C^T. But the notation $|\varepsilon_{col}|(M_{*j})$ denotes the normalized absolute energy/information subtotal of column *j* of matrix *M*.

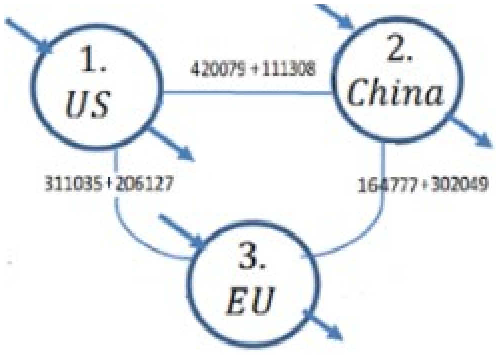
2.6 The Digital Nature of S-OTP-Method2

Notably, S-OTP-Method2 is based on bipolar equilibrium-based rebalancing**.** Bipolarity is a quantum feature that form the bipolar reality of negative-positive particles. The bipolar property, however, can be depolarized for digital cryptography.

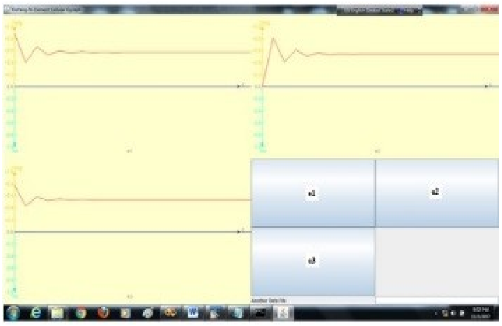
A unipolar CM can be revealed from a bipolar one with depolarization. Since a bipolar representation is a generalization of unipolar representation and subsumes unipolar cases, all the elements of a polarized map can simply have zero negative energy/information which leads to the simplified CM as in **Fig. 3** coded as a unipolar matrix *C(t)*—a positive relation that does not distinguish import and export with symmetrical subtotals.

Depolarization leads to a unipolar cipher named S-OTP$_{1R1}$-Method2 which is basically the same as S-OTP$_1$-Method2 except using a positive *CM* and a positive matrix *M*. **Fig. 3** shows a decryption example using S-OTP$_{1R1}$-Method2 where in the last row we have the same result as for the bipolar case.

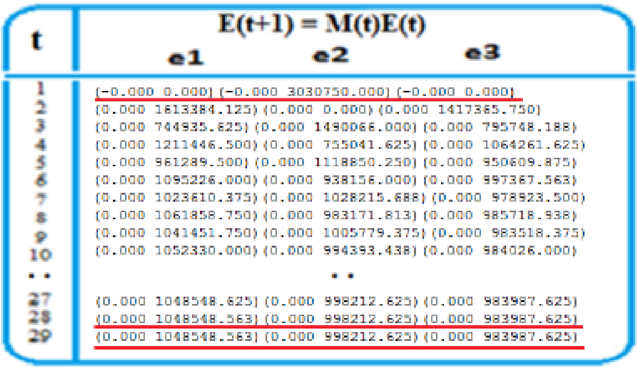
$d_1 = |\varepsilon|(-0, +1048549) = 1048549$;
$d_2 = |\varepsilon|(-0, +998213) = 998213$;
$d_3 = |\varepsilon|(-0, +983988) = 983988;$
$D = Concatenate(d_1, d_2, d_3) = 1048549998213983988.$



(a)



(c)



t	E(t+1) = M(t)E(t) e1	e2	e3
1	[-0.000 0.000]	[-0.000 3030750.000]	[-0.000 0.000]
2	(0.000 1613384.125)	(0.000 0.000)	(0.000 1417365.750)
3	(0.000 744935.625)	(0.000 1490066.000)	(0.000 795748.188)
4	(0.000 1211446.500)	(0.000 755041.625)	(0.000 1064261.625)
5	(0.000 961289.500)	(0.000 1118850.250)	(0.000 950609.875)
6	(0.000 1095226.000)	(0.000 938156.000)	(0.000 997367.563)
7	(0.000 1023610.375)	(0.000 1028215.688)	(0.000 978923.500)
8	(0.000 1061858.750)	(0.000 983171.813)	(0.000 985718.938)
9	(0.000 1041451.750)	(0.000 1005779.375)	(0.000 983518.375)
10	(0.000 1052330.000)	(0.000 994393.438)	(0.000 984026.000)
..		..	
27	(0.000 1048548.625)	(0.000 998212.625)	(0.000 983987.625)
28	(0.000 1048548.563)	(0.000 998212.625)	(0.000 983987.625)
29	(0.000 1048548.563)	(0.000 998212.625)	(0.000 983987.625)

(b)

Figure 3. Information-conservational unipolar rebalancing: (a) depolarized CM; (b) Positive distribution; (c) Positive curve (scaled)

2.7 Two Puzzles Explained

(A) How can matrix $C(t)$ in symmetry ($C(t)(i,j)= C(t)(j,i)$) be used in cryptography? The answer is that, although matrix $C(t)$ is symmetrical, a column-major normalized M can be non-linear and asymmetrical because the normalization is by dividing its column subtotal of $C^T(t)$ (data section subtotal), but not by the global total (corresponding to the overall summation). For instance,

$$C(t) = \begin{bmatrix} 0 & 531587 & 517162 \\ 531587 & 0 & 466826 \\ 517162 & 466826 & 0 \end{bmatrix}; \qquad M = \begin{array}{c} 0 \\ 1 \\ 2 \end{array}\begin{bmatrix} 0.000 & 0.532 & 0.526 \\ 0.507 & 0.000 & 0.474 \\ 0.493 & 0.468 & 0.000 \end{bmatrix};$$

where C is symmetrical but M is not. The non-linear asymmetrical property of M can be characterized with a set of linear equations. Let the three subtotals (or data sections) be x, y, and z, respectively, for the 3×3 matrix M we have $m_{10}\times x - m_{01}\times y = 0$; $m_{20}\times x - m_{02}\times z = 0$; *and* $m_{21}\times y - m_{12}\times z = 0$; *and* $m_{ij} \neq m_{ji}$. The set of equations have infinite number of solutions because all column coefficients of M correlate non-linearly with each other due to non-linear normalization based on different local column subtotals. This is fundamentally different from percentage distribution where all percentages are normalized with a global total and linearly independent.

(B) If a unipolar positive matrix is sufficient why do we need a bipolar equilibrium-based matrix in cryptography? There are two top answers to this question: (1) The universe consists of negative-positive particles. Without bipolarity there would be no bipolar information conservation and bipolar quantum computing. Thus, bipolarity leads to a quantum model compatible to digital computing (further discussed later). (2) A bipolar matrix avoids large denominators, doubles the number of elements in a unipolar matrix, doubles the parallel computing power, and doubles collective precision with equilibrium-based rebalancing (further discussed later).

2.8 Security of Method1 and Method2

Method1 is based on percentage distribution. It provides a basis for both theoretical analysis and practical development. The goal is to search for secure information conservational S-OTP ciphers by analyzing different approaches which may or may not be secure.

Theorem 2a. Under the conditions of Definition 2, S-OTP$_1$-Method1 is information conservationally not secure.

Proof. *It follows from that the transmitted message consists of numerical meta data with fixed format. Such knowledge could potentially weaken the ITS of OTP.* ■

The above problem can be solved by adding random bits to the metadata as paddings before being enciphered that can be removed when being decrypted by receiver.

Theorem 2b. Under the conditions of Definition 2, S-OTP$_1$-Method1 would be information conservationally secure provided that a sufficient number of random bits are added as paddings to the metadata to be enciphered that can be removed when being decrypted by receiver.

Proof. With the provision, the conditions of S-OTP as defined in Definition 2 remain intact. ■

The one key version (S-OTP$_1$) suggests that, two different keys (S-OTP$_2$) might be considered for an information conservational solution.

S-OTP$_2$-Method1 and Its Revised Versions

(1) S-OTP$_2$-Mehrod1: Use key K_1 to encipher the summation $X = \Sigma_i x_i$ to X'; Use key K_2 to encipher the text of $\{x_i/X\}$ to $\{x_i/X\}$'; Transmit the packaged pair {X', $\{x_i/X\}$'} without key reuse;

(2) S-OTP$_{2\text{-}2}$-Mehrod1: First, use a random number of bits specified in key K_1 as random paddings for altering the numerical format of {X, $\{x_i/X\}$} to {X, $\{x_i/X\}$}' (e.g. 080203 stands for "Insert 2 random bits for every 8 bits after bit position 3.); Then, use K_2 as a key to encipher {X, $\{x_i/X\}$}' to {X, $\{x_i/X\}$}''; Transmit {X, $\{x_i/X\}$}'';

Theorem 3a. Under the conditions of Definition 2, S-OTP$_2$-Mehrod1 is information conservationally not secure.

Proof. *It follows the proof of S-OTP$_1$-Mehrod1.* ■

Theorem 3b. Under the conditions of Definition 2, S-OTP$_{2\text{-}2}$-Mehrod1 is information conservationally secure.

Proof. *With sufficient random bits as paddings specified by the first key and a regular second key, the message and its format are randomized and concealed in an unbreakable pad that does not weaken the security of OTP.* ■

Evidently, if a percentage distribution $\{x_i/X\}$ is replaced with an information conservational matrix *M* we will have different 1-key or 2-key versions of S-OTP-Mehrod2 with similar security conditions as that of S-OTP-Mehrod1.

3 Optimization

3.1 Minimal BQCA Theorem

Theorem 4. Mehtod1 is the minimal case of Method2.

Proof. *Mehtod2 entails an N×N square matrix multiplied by a column vector in an information conservational BQCA. When N×N is reduced to N×1, the matrix becomes a column vector of percentage distributions $w_i = \{x_i/X\}$ summing up to 1.0, the single number must be the summation X of N sections, such that the column vector multiplied by a single element matrix results in a column vector energy/information distribution $\{x_i\}$. The Matrix multiplication can be deemed the minimal BQCA which requires a final equilibrium state be reached in a single step with high precision such as* $\begin{pmatrix} w_0 \\ w_1 \\ w_{i+1} \\ w_n \end{pmatrix}[X] = \begin{pmatrix} x_0 \\ x_1 \\ x_{i+1} \\ x_n \end{pmatrix}$. ■

Theorem 4 proves that Method1 is suitable for reducing network traffic, and Method2 can be used for computational precision. Theorems 1-4 provide a basis for applicability and efficiency analysis of a new crypto paradigm using either Method1 (percentage distribution) or Method2 (information conservational matrix) or a combination of the two. Based on IEEE binary64 standard, double precision floating-point format provides us an upper limit for long messages. Major considerations are on the key length for enciphering both the math summation and the percentage distribution. According to IEEE binary64 standard, exponents range from −1022 to +1023 that allows the representation of numbers between 10^{-308} and 10^{308}, with full 15–17 decimal digits precision. By compromising precision, it allows even smaller values up to about 5×10^{-324}.

3.2 Applicability and Efficiency of Method1

Using S-OTP$_{2\text{-}2}$-Method1, the percentage distribution $\{x_i/X\}$ needs to be enciphered, where each double precision floating-point number x_i/X requires $2^6 = 64$ bits. While a $1M = 2^{20}$ bits message needs an impractical same length OTP key, if the 1M-bit message is divided into 512-bit sections, the division leads to $N = 2^{20}/2^9 = 2^{11}$ data sections with a math summation less than 512+11 = 523 bits. $N = 2^{11}$ double precision floating-point numbers are needed for the percentage distribution $\{x_i/X\}$ that entails $2^{11}2^6 = 2^{17}$=128k-bit key length. A 128k+523-bit key is a nearly 8-fold reduction in key length compared with the message length. The upper limit of the exponent is +1023 for signed integers based on IEEE binary64. At the limit, the key length saving approaches 16-fold. It seems to be a solved problem. However, there are still unsolved problems. First, 128K+523 bits data plus K_1 for random paddings is still too long to be a practical key length. Second, when the grand total is huge, the percentage distribution will have a precision problem because a percentage is normalized by the grand total as the denominator using percentage distribution.

3.3 Applicability and Efficiency of Method2

While for 1M-bit long messages the key length requirement for OTP is not practical, it is much less a problem with percentage distribution using Method1, but still a problem with Method2. Evidently, a 1M-bit message divided into 2K 512-bit sections would need a $2^{11}\times2^{11}$ sparsely populated information conservational matrix *M*. Assuming each column has an average of no more than 8 non-zero elements in 64-bit double precision floating-point format plus one index that leads to $8\times64=2^3\times2^6$ bits per column. A total of $8\times64\times2^{11}=2^{20}$ bits plus a 513-bit summation need to be transmitted in ciphertext—more than the original 1M bits.

While Method2 is inefficient and impractical, its information conservational property is still quite attractive. In terms of digital computing, its column-major normalization does not use the grand total but a much smaller section subtotal as the denominator and a much smaller sender designated percentage of a subtotal as the numerator. Remarkably, it can divide-and-conquer the high precision requirement into lower precision requirement. On the quantum side, its equilibrium-based rebalancing property reflects the bipolar reality of particle-antiparticle coexistence [9-12].

3.4 Hierarchical Optimization

Without entirely enciphering both a summation and its percentage distribution or matrix *M*, Method1 and Method2 cannot achieve ICS. Enciphering matrix *M* does not reduce key length. Thus, S-OTP$_{2\text{-}2}$-Method1 is the best candidate for hierarchical scalability toward a final solution. This leads to S-OTP$_H$ as illustrated in **Fig. 4**. It is made practical with IEEE binary64.

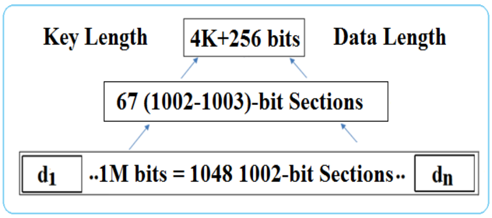


Figure 4. A 3-Level Hierarchy of S-OTP$_H$ for 1M bits

First, we assume that 1M = 1048576 bits data divided into 1048 1000-bit sections. We would have a maximum of 1012-bit summation X associated with a 1048×64 = $(2^{10} + 24) \times 2^6 = 2^{16} + (24 \times 2^6) = 64K + (16+8) \times 2^6 = 64K + 1K + 512$ bits percentage distribution D. The summation can be converted to a 64-bit double precision floating-point number. The pair {X, D} would consists of 64K + 1K + 576 bits = 67136 bits. Second, the 67136 bits can be divided into 67 of 1002-1003-bit sections. That leads to 67 of 64-bit double precision floating-point numbers for the percentage distribution plus a maximum of 1008-bit summation. Again, the summation can be converted to a 64-bit double precision floating-point number. The 67+1 double precision numbers need $68\times64 = 2^{12} + 256$ = 4K + 256 bits. Third, 4K + 256 bits can be further scaled to 1K-bits plus K_1 as a less than1K-bit randomizer to result in a 2K-bit key K_2. Evidently, due to the short length the two keys are no longer a drawback. Formally, we have S-OTP$_H$-Method1. Similarly, 1 Gaga bits = 1K Mega bits that entails a larger hierarchy.

S-OTP$_H$-Method1

For every 1M bits of data to be transferred, assume sender Alice and receiver Bob share two private keys K_1 and K_2 distributed through QKD.

Part I – Encryption

(1) L=1; if the data length is short enough for an OTP cipher key, print message "Please use OTP without hierarchy";
(2) L=L+1; determine summation X $=\{\sum_i x_i\}$ and derive the percentage distribution D=$\{x_i/X\}$;
(3) If the data is too long for an OTP cipher key and its length is reducible (>minimum), go to Step (2);
(4) If the data is too long for an OTP cipher key and its length is unreducible, stop and restart with different number of scalable pads;
(5) Apply S-OTP$_{2\text{-}2}$-Method1 to encipher {X,D,L} to {X,D,L}'' with key K_1 for adding random paddings and key K_2 as a cipher key;
(6) Transmit the ciphertext {X,D,L}'' to the receiver.

Part II - Decryption

(1) Decipher {X,D,L}'' to {X,D,L} with K_2 and K_1;
(2) Use X and D to find next layer $\{\{x_i\}$, D}, L = L-1,
if L>1, repeat step (2) until L = 1;
(3) Cast $\{x_i\}$ to string format $\{d_i\}$;
(4) Recover the original message or data set D by concatenating $\{d_i\}$.

Theorem 5. Under the conditions of Definition 2, S-OTP$_H$-Method1 is information conservationally secure.

Proof. *It follows from the proof of Theorem 1c and the information conservational security of S-OTP$_{2\text{-}2}$-Method1.* ■

3.5 Collective Precision

While Method1 uses percentage distribution, Method2 uses information conservational encryption. In Method1 each data section depends on a single percentage resulted from linear normalization by a grand total. When the data length is long, Method1 will have a precision problem. In Method2, each data section depends on all columns of matrix *M* resulted from column-major normalization by much smaller subtotals where percentage distribution is not directly calculated using the grant total. If each column has an average of $n > 2$ non-zero numbers, the precision requirement is n-times smaller. The larger the number n the more parallelism in high precision decryption. When n equals N, Method2 reaches maximum parallelism with N-fold precision enforcement for a positive matrix M and $2N$-fold for a bipolar matrix M. This observation leads to the inception of information conservational *collective precision*.

Observation 1: Asymptoticity. If M is information conservational, BQCA $E(t+1) = M \times E(t)$ is asymptotic to an equilibrium state determined by M [9,10].

Observation 2: Information Conservational Computing and Cryptography. If an original message D is converted to an energy/information total E through a BQCA transformation, the information conservational matrix M of the BQCA can serve as a key to decode the total information to the original message D in the receiver side [9]. However, to encrypt and transmit matrix M will cost more than to encrypt and transmit the original message. Thus, Method1 is more efficient than Method2 for encryption and transmission, but only Method2 can enable collective precision and efficient decryption.

Theorem 6. If M is information conservational, BQCA $E(t+1) = M \times E(t)$ can be used to derive the percentage distribution in an equilibrium state determined by the BQCA.

Proof. *Given 100 (percent), Theorem 6 follows the asymptoticity theorem [9] directly (see example in Fig. 1e).* ■

Theorem 7. A percentage distribution of N divisions can be converted to an $N \times N$ (unipolar or bipolar) information conservational matrix M for collective precision with maximum parallelism such that M is information conservational and BQCA $E(t+1) = M \times E(t)$ is asymptotic to an equilibrium state.

Proof. *Notice that M is normalized and information conservational but not unique. Theorem 7 follows from* $\begin{pmatrix} w_0 \\ w_1 \\ w_{i+1} \\ w_n \end{pmatrix}[E] = \begin{pmatrix} w_0E \\ w_1E \\ w_{(i+1)}E \\ w_nE \end{pmatrix}$ *because* $\begin{pmatrix} w_0 \\ w_1 \\ w_{i+1} \\ w_n \end{pmatrix}$ *is strictly proportional to* $\begin{pmatrix} w_0E \\ w_1E \\ w_{i+1}E \\ w_nE \end{pmatrix}$*. That is, M can be derived from either of them* ■

Based on the above findings we can conclude that, on the sender side, matrix M can be used for determining the percentage distribution with *N-2N fold* reduction of precision requirement due to column-major normalization (Re. Eq. (3)). On the receiver side, *M* can be used to decrypt a big total to subtotals (or data sections) with collective precision in a reverse way (**Fig. 1b** and **Fig. 3b**). Thus, Method1 and Method2 can be used in a combination. Method2 focuses on collective precision with ICC; Method1 focuses on secure and efficient data transmission, that lead to the block diagram design in Fig. 5 impact factor followed by an optimized algorithm that combines the advantages of Method1 and Method2 while eliminating their drawbacks.

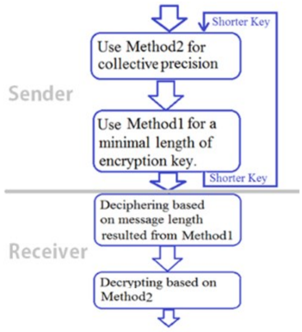


Figure 4. Sender and Receiver

S-OTP$_H$-Method1+2

For every 1M bits of data to be transferred, **a**ssume sender Alice and receiver Bob share two private keys K_1 and K_2 distributed through QKD.

Part I – Sender Side

(1) L=1; if the data length is short enough for an OTP cipher key, print message "Please use OTP" without hierarchy;

(2) L=L+1, compute summation X $=\{\sum_i x_i\}$, derive information conservational matrix M, and determine percentage distribution D=$\{x_i/X\}$ with BQCA E(t+1) = M×E(t) (see **Fig. 1e**);

(3) If the data is too long for an OTP cipher key and its length is reducible (>minimum), go to Step (2);

(4) If the data is too long for an OTP cipher key and its length is unreducible, go to Step (2) with a smaller N such that 0<i<N and N≥2;

(5) Apply S-OTP$_{2\text{-}2}$-Method1 to encipher {X,D,L} to {X,D,L}'' with key K_1 for adding random paddings and key K_2 as a cipher key;

(6) Transmit the ciphertext {X,D,L}'' to the receiver.

Part II – Receiver Side

(1) Decipher {X,D,L}'' to {X,D,L} with K_2 and K_1;

(2) Construct information conservational matrix M from D;

(3) Use X and M in a BQCA to find next layer $\{\{x_i\}$,M}, L=L-1;
if L>1, repeat step (3) until L=1;

(4) Cast $\{x_i\}$ to string format $\{d_i\}$;

(5) Recover the original message or data set D by concatenating $\{d_i\}$.

Theorem 8. Under the conditions of Definition 2, S-OTP$_H$-Method1+2 is information conservationally secure.

Proof. *Since Method2 is only used for collective precision on the sender side and parallel decryption on the receiver, the theorem follows from the proof of Theorem 1c and the security of S-OTP$_H$-Method1.* ■

3.6 Transmitting Large Dada Sets

With S-OTP$_H$-Method1+2 a large data set can be serialized as a number of Mega bits S-OTP pads, and each Mega bits can be securely transmitted with a 2K-4K bit short key that is practical with QKD while a 1M bit key with the same length as the message is obviously not practical.

3.7 A Consequence of Collective Precision

Collective precision adds a number of new features to hierarchical S-OTP. On the sender side, it can be used for testing everything efficiently and precisely to guarantee that the receiver side will get the correct message. On the receiver side it can be used to decrypt a summation efficiently with collective precision and in massive parallelism colluding with the sender side based on public protocols on data size and number of divisions N. These are necessary but auxiliary functions.

It can be observed that the percentage distribution $\{x_i/X\}$ is a major contributor to the key length requirement of S-OTP. If it does not have to be enciphered but transmitted in plaintext, we only need to cipher a short summation with a much shorter key. If 1Gaga bits divided into 1K mega divisions, each 1Mega division results in a 64-bit double precision summation, 1G bits with 1K such summations would only need 64K bits to be ciphered. Of course, the summations can be hierarchically scaled further.

Now, with collective precision, we have the challenging question: *Can the percentage distribution $\{x_i/X\}$ be securely transmitted in plaintext if $\{x_i\}$ are double precision floating-point numbers due to the reuse of a double precision floating-point key in multiplication or division operation instead of XOR?*

Whereas this paper has assumed that a transmitted message as a long binary integer *L* is divided into smaller integers $\{x_i\}$, and their summation is also an integer $X = \sum_i x_i$. Evidently, *X* can be guessed by attackers with a trial-error method to break S-OTP if an integer key is reused for $\{x_i\}$ without encrypting the percentage distribution $\{x_i/X\}$. Now with collective precision, floating point decimals can be used instead of integers.

Collective precision makes it possible to use double precision floating-point decimals as a reusable key for non-linear multiplication, division, addition, and/or subtraction. Such non-linear operations lead to decimal precision that cannot be guessed without knowing the reusable key—a double precision floating-point decimal. The summation of these decimals results in another double precision floating-point decimal that can be encrypted as an S-OTP pad. In this case, the percentage distribution $\{x_i/X\}$ could be misleading to attackers, and the final summation could be unguessable with a trail-error method due to floating point decimal precision and non-linear operation with a decimal key. This leads to the hypothesis for future research.

Hypothesis: *With double precision floating-point decimals for collective precision, the percentage distribution $\{x_i/X\}$ in S-OTP can be securely transmitted in plaintext provided that (a) $\{x_i/X\}$ is not the actual percentage distribution but a misleading to attackers; (b) the summation $X = \sum_i x_i$ is enciphered as an unbreakable summation of double precision floating point decimal numbers.*

(Remark: This hypothesis could close a loophole in the proof of Theorem 8 of ref. [9].)

4 Quantum-Dream Machinery

Collective precision suggests that Method2 is suitable for research/development of bipolar quantum-digital machinery. While unipolar values are preferred by digital machines, the bipolar nature of S-OTP-Mehtod2 makes it suitable for developing quantum machinery with equilibrium-based bipolar quantum rebalancing and information conservation (**Fig. 6**). Encryption would be unnecessary for quantum computing and communication [8]. The quantum machine in **Fig. 6(a)** can be used, theoretically, in encryption and decryption for digital communication. Each column of an N×N matrix *M* may have a maximum of N non-zero elements for maximum parallelism. If N=1K or 2K, a math distribution among N sectors can be determined in one procedure on the sender side; or an information total can be quantum rebalanced to N subtotals in parallel without using percentage distribution.

While the bipolar quantum dream seems to be "far-fetched" in terms of quantum-digital compatibility, a newly reported discovery of a class of subatomic particles (fermions) named Angel Particles [20-22] injected new life into this line of research. The new discovery is a family of particle-antiparticle pairs expected to make quantum computing more practical and powerful. It strengthens the ontological basis of equilibrium-based bipolar quantum rebalancing. **Fig. 6(a)** shows the draft of a bipolar quantum-digital crypto machine. **Fig. 6(b)** shows bipolar quantum teleportation. **Fig. 6(c)** shows a bipolar qubit register. The dream machinery forms a quantum intelligence paradigm or S-OTP$_Q$ for further research.

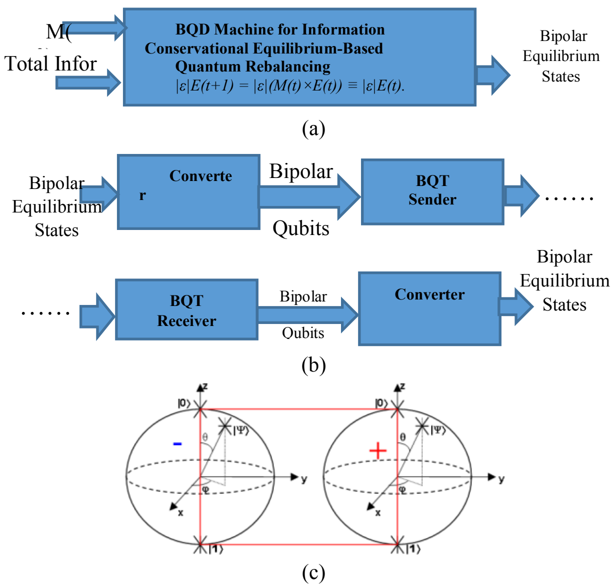


Figure 6. (a) Bipolar quantum-digital (BQD) computing; (b) Bipolar quantum teleportation (BQT); (c) Bipolar qubit register (adapted from [15,23])

5 Conclusions

S-OTP—scalable one-time pad has been presented based on ICC. Security conditions have been established. Collective precision has been proposed. It has been shown that

(1) S-OTP reduces key and data length significantly, and makes it possible for transmitting large data sets with ICS.
(2) Math summation without using big primes makes S-OTP quantum proof to quantum factorization (cf. [1, 9, 24-29]).
(3) ICC can be massively parallel, accurate, efficient, and suitable for developing quantum machinery with collective precision.

Whereas OTP is prevented from being widely used by its key length requirement, S-OTP gets around the problem through ICC without compromising OTP security and network traffic. Thus, the S-OTP paradigm qualifies itself as a unique extension from ITS to ICS for post-quantum cryptography (Fig. 7).

Floor-roof mysteries. According to the floor-roof theory of science [17], ITS of OTP is developed based on information theory rooted in probability and statistics—a floor of modern science; ICS of S-OTP is a set-theoretic development rooted in bipolar fuzzy sets and dynamic equilibrium—a roof of modern science. Thus, this work has opened some major challenges. Among them are the following floor-roof mysteries for future research:

(1) Is information conservation an information theoretic extension?
(2) If S-OTP is just OTP plus data compression, and there is nothing new, is S-OTP information theoretically secure? If not, how can OTP be secure?
(3) Shannon famously concluded on the impossibility for perfect secrecy beyond OTP with key length greater than or equal to the message to be ciphered [7]. Although this paper does not attempt to falsify Shannon's theorem, however, if S-OTP is secure, could Shannon's theorem be wrong?
(4) The data compression achieved in S-OTP is deemed by some researcher as an impossible task in violation of the most fundamental basics of Kolmogorov complexity [30,31]. But if it is really impossible, how can IEEE binary64 be a standard?
(5) Could modern science have been a well-founded building with a floor of observable truth but with a missing roof for equilibrium and information conservation [9,10,16,17]?

Floor-roof assertions. While the above mysteries are left for future research effort, we have the following floor-roof assertions:

(1) Can the floor perform some functions not performed by the roof? The answer is definitely YES.
(2) Can the roof perform some functions not performed by the floor? The answer is definitely YES.
(3) Can information conservational security on the roof solve some unsolved problems by information theoretic security on the floor? The answer should be logically YES.

The significance of this work is manefest in the above mysteries and assertions. The findings prove that S-OTP can bypass the assumptions of Shannon's theorem on OTP key length limitation [7] without falsifying it. The key is information conservation with double-precision floating-point format and quantum-fuzzy collective precision.

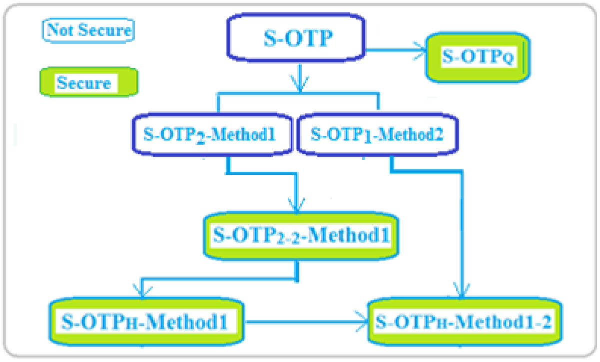


Figure 7. The road to information conservational security

REFERENCES

1. P. Shor (1994). Algorithms for quantum computation: Discrete logarithms and factoring. *Proc. 35th Annual Symposium on Foundations of Computer Science*, 1994, 124-134.
2. Daniel J. Bernstein & Tanja Lange (2017). Post-quantum cryptography. *Nature*, volume549, pages188–194 (14 September 2017)
3. *F. Miller (1882). Telegraphic code to insure privacy and secrecy in the transmission of telegrams.* C.M. Cornwell.
4. S. *Bellovin, M. Steven (2011). "Frank Miller: Inventor of the One-Time Pad". Cryptologia* ***35*** *(3): 203–222.*
5. C. Shannon (1949). "Communication Theory of Secrecy Systems". *Bell System Technical Journal* **28** (4): 656–715.
6. C. Shannon (1948). "A Mathematical Theory of Communication." *Bell System Technical Journal. 27 (3):379–423.*
7. C. Shannon (1949). "Communication Theory of Secrecy Systems". *Bell System Technical Journal* **28** (4): 656–715.
8. J. Park (1970). "The concept of transition in quantum mechanics". *Foundations of Physics*. 1: 23-33.
9. W.-R Zhang (2017). From Equilibrium-Based Business Intelligence to Information Conservational Quantum-Fuzzy Cryptography — A Cellular Transformation of Bipolar Fuzzy Sets to Quantum Intelligence Machinery. (*IEEE Explore 3/24/2017*) *IEEE Trans. on Fuzzy Systems***,** Volume: 26, Issue: 2, April 2018, 656 – 669.
10. W.-R. Zhang (2017). Programming the Mind and Decrypting the Universe—A Bipolar Quantum-Neuro-Fuzzy Associative Memory Model for Quantum Cognition and Quantum Intelligence. *Proc. of Int'l J. Conf. on Neural Networks (IJCNN 2017),* Anchorage, Alaska, USA, May 14–19, 2017, 1180 - 1187.
11. W.-R. Zhang (2012). YinYang Bipolar Atom – An Eastern Road toward Quantum Gravity. *J. of Modern Physics*, 2012, 3, 1261-1271. (open access) DOI: 10.4236/jmp.2012.329163.
12. W.-R. Zhang (2016). G-CPT Symmetry of Quantum Emergence and Submergence – An Information Conservational Multiagent Cellular Automata Unification of CPT Symmetry and CP Violation for Equilibrium-Based Many World Causal Analysis of Quantum Coherence and Decoherence. *J. of Quantum Infor. Sci.*, Vol. 6, No. 2, 2016, pp. 62-97. (open access) DOI: 10.4236/jqis.2016.62008.
13. W.-R. Zhang (1998). (Yin)(Yang) Bipolar Fuzzy Sets. *Proc. of IEEE World Congress on Computational Intelligence*, Anchorage, Alaska, May, 1998, Fuzz-IEEE pp835-840.
14. W.-R. Zhang and L. Zhang (2004). YinYang Bipolar Logic and Bipolar Fuzzy Logic. *Inform. Sciences*, Vol. 165, 3-4, 2004, pp265-287.
15. W.-R. Zhang (2011), *YinYang Bipolar Relativity: A Unifying Theory of Nature, Agents and Causality with Applications in Quantum Computing, Cognitive Informatics and Life Sciences*. IGI Global, Hershey and New York, 2011.

16. W.-R. Zhang (2016). Information Conservational YinYang Bipolar Quantum-Fuzzy Cognitive Maps – Mapping Business Data to Business Intelligence. *Proc. of IEEE World Congress on Computational Intelligence – Fuzz-IEEE*, Vancouver, July 2016, 2279-2286.
17. W.-R. Zhang (2018), The Road from Fuzzy Sets to Definable Causality and Bipolar Quantum Intelligence — To The Memory of Lotfi A. Zadeh. To appear, *Journal of Intelligent & Fuzzy Systems* xx (20xx) x–xx. DOI:10.3233/JIFS-172159, IOS Press
18. L. A. Zadeh (1965). Fuzzy sets. *Information and Control,* 8(3): 338–353.
19. L. A. Zadeh (2006). Fuzzy logic. *Scholarpedia,* 3(3):1766, Created: 10 July 2006, reviewed: 27 March 2007, accepted: 31 March 2008.
20. Stanford News (2017). *An experiment proposed by Stanford theorists finds evidence for the Majorana fermion, a particle that's its own antiparticle*. http://news.stanford.edu/2017/07/20/evidence-particle-antiparticle/, 7/20/2017.
21. S.-C. Zhang (2017). *Search and Discovery of the Majorana Fermion. Lecture, Institute for Theoretical Physics,* Stanford University. 2017. https://sitp.stanford.edu/video/shoucheng-zhang-search-and-discovery-majorana-fermion-2-2
22. Q. L. He; Lei Pan; Alexander L. Stern; Edward C. Burks; Xiaoyu Che; Gen Yin; Jing Wang; Biao Lian; Quan Zhou; Eun Sang Choi; Koichi Murata; Xufeng Kou; Zhijie Chen; Tianxiao Nie; Qiming Shao; Yabin Fan; Shou-Cheng Zhang; Kai Liu; Jing Xia; Kang L. Wang (2017). Chiral Majorana fermion modes in a quantum anomalous Hall insulator–superconductor structure. *Science* 21 Jul 2017: Vol. 357, Issue 6348, pp. 294-299
23. W.-R. Zhang (2013). Bipolar Quantum Logic Gates and Quantum Cellular Combinatorics — A Logical Extension to Quantum Entanglement, *J. of Quantum Information Science*, Vol. 3 No. 2, 2013, pp. 93-105. (open access) DOI: 10.4236/jqis.2013.32014.
24. National Cryptographic Solutions Management Office (NCSMO) (2015). *Cryptography Today*. (Date Posted: Jan 15, 2009; Last Modified: Aug 19, 2015; Last Reviewed: Aug 19, 2015)
25. D. Hayford (2014). The Future of Security: Zeroing in on Un-Hackable Data with Quantum Key Distribution. *WIRED*.
26. C. Dillow (2013), Unbreakable Encryption Comes to the U.S. *Forture.* OCTOBER 14, 2013, 9:00 AM EDT.
27. W. Diffie and Martin E Hellman (1976). Multiuser cryptographic techniques. In *Proceedings of national computer conference and exposition*, pages 109–112. ACM, 1976.
28. J. Chu (2016). The beginning of the end for encryption schemes? New quantum computer, based on five atoms, factors numbers in a scalable way. *MIT News.* March 3, 2016.
29. T. Monz, D. Nigg, E. A. Martinez, M. F. Brandl, P. Schindler, R. Rines, S. X. Wang, I. L. Chuang, R. Blatt (2016). Realization of a scalable Shor algorithm. *Science* 04 Mar 2016: Vol. 351, Issue 6277, pp. 1068-1070.
30. A. Kolmogorov (1963). "On Tables of Random Numbers". *Sankhyā Ser.* A. **25**: 369–375. MR 0178484.
31. A. Kolmogorov, (1998). "On Tables of Random Numbers". *Theoretical Computer Science*. **207**(2): 387–395.