Final Report

Grant/Contract Title:
PHYSICAL CRYPTOGRAPHY: A NEW APPROACH TO KEY GENERATION AND DIRECT ENCRYPTION

Principle Investigator:
Prof. Horace P. Yuen
Electrical Engineering & Computer Science
Northwestern University
2145 Sheridan Rd.
Evanston, IL 60208

Grant/Contract Number: FA9550-06-1-0452
AFOSR
**Physical Cryptography: A New Approach to Key Generation and Direct Encryption**

Northwestern University, EECS Dept., 2145 Sheridan Rd, Evanston, IL

Air Force Office of Scientific Research

Approved for public release; distribution unlimited

**Security Classification of:**
- a. Report: unclassified
- b. Abstract: unclassified
- c. This Page: unclassified

**Limitation of Abstract:**
- Same as Report (SAR)

**Number of Pages:** 17

**Name of Responsible Person:**

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18
Abstract:

The security of key generation and direct encryption in quantum and physical cryptography have been investigated. It is found that similar to the situation of conventional mathematics based cryptography, fundamental and meaningful security levels for either the data bits or the cipher seedkey bits have not been quantified for any concrete cipher except the one-time pad. Attempts were made in our study to rectify the situation, especially for the αη cryptosystem and the more powerful and complicated CPPM cryptosystem.

Some success can be obtained under certain assumptions such as the attack’s inability to entangle across many modes which are quite realistic, and the availability of unlimited bandwidth to the user which are not. It is concluded that much further effort is required for meaningful security quantification in concrete cryptosystem of any quantum or classical variety.

Table of contents

I. Introduction — page 3
II. Problem Description — page 3
III. Technical Accomplishments — page 9
IV. Other Results — page 13
V. References — page 15
VI. Executive Summary — page 16
   a. Personnel — page 16
   b. Doctoral Dissertation — page 17
   c. List of Publications — page 17
I. Introduction.

In this report we will summarize the main results that have been obtained from the work supported by this AFOST Grant. Most of them have been published in refereed journals, some accepted but yet to appear some results have been published in a conference proceeding volume and some are given the archive quant-ph yet to be submitted for journal publication. Some have appeared in a PhD dissertation which can be obtained through Northwestern University. A small number of our results appear in this report for the first time. The above publicity available papers would be listed in the Executive Summary.

In section II the general problems that were studied would be reviewed. The main results would be described in Section III. In section IV mention is made on the problems that were looked into with yet no significantly new results. An indication of the main open issue is also given.

II. Problem Description

The emerging development of classical-noise cryptography [1, 2] and quantum cryptography [3] suggests that a new way of building cryptosystems may be based on physical effects on electromagnetic signals that lend a qualitatively different layer of security from standard cryptosystems based on purely mathematical relations. Furthermore, physical cryptography may provide information-theoretically secure mechanisms for fresh key generation, which is impossible in standard cryptography where the user Bob and the attacker Eve have the same observation, i.e., \( Y_B = Y_E \) corresponding to any data \( X_A \) transmitted by Alice. For a detailed explanation, see refs. [4, 5] (Note that physical cryptography does not mean physical layer encryption, which is currently based on standard cryptography.). Similarly, information-theoretically secure direct encryption schemes against known-plaintext attacks may be possible [4]. If such cryptosystems can be operated realistically with high efficiency, they would provide new cryptographic capability and may replace or strengthen cryptosystems in current use.

There are two established approaches to physical cryptography. The first is based on classical noise that Eve has to suffer for whatever reason [1] - the only specific protocol that has been proposed is the so-called YK protocol [2] which has been \( \alpha \eta \) further studied theoretically and
experimentally to only a limited extent. The second is quantum cryptography [3] based on BB84/Ekert type protocols, which has received extensive development due to its promise of “unconditional security”. However, it is also necessarily inefficient from the weak signals, in addition to associated quantum sensitivity problems [5]. Furthermore, no unconditionally secure concrete quantum protocol has ever been even just proposed that takes into account all the side information Eve may obtain during execution [6], finite bit-sequence statistical fluctuation, as well as imperfections in any realistic implementation.

A new approach to physical cryptography, called KCQ (Keyed Communication in Quantum Noise) in the quantum domain but which also has a classical analog applicable to rf systems, has been developed both theoretically [4-5, 7-9] and experimentally [10-13]. It promises, with the help of a shared secret key between Alice and Bob, efficient and secure key generation and direct encryption not obtainable from other quantum schemes or classical noise schemes. In the course of its theoretical development, it was found that the foundations of symmetric-key cryptography and key generation have not been sufficiently developed for many purposes. It is the aim of our work to address some of these fundamental problems in general, and in conjunction, to develop further the KCQ security/efficiency study for the following two concrete schemes.

Consider the original experimental scheme \( \alpha \eta \) (called Y-00 in Japan) as described in [10] and depicted in Fig. 1. Alice encodes each data bit into a coherent state in a qumode, i.e., an infinite-dimensional Hilbert space (the terminology is analogous to the use of qubit for a two-dimensional Hilbert space), of the form (we use a single qumode representation rather than a two-qumode one for illustration)

\[
|\alpha_i\rangle = |\alpha_0 (\cos \theta_i + i \sin \theta_i)\rangle
\] (1)

where \( \alpha_0 \) is real, \( \theta_i = \pi \ell/M \), and \( \ell \in \{0, \ldots, 2M-1\} \). The \( 2M \) states are divided into \( M \) basis pairs of antipodal signals \( \{ |\pm \alpha_i\rangle \} \) with \( -\alpha_i = \alpha_{i+M} \). A seed key \( K \) of bit length \( |K| \) is used to drive a conventional encryption mechanism whose output is a much longer running key \( K' \) that
is used to determine, for each qumode carrying the bit $b \in \{0,1\}$, which pair $\{|\pm \alpha_i\rangle\}$ is to be used. The bit $b$ could either be part of the plaintext in a direct encryption system (as is the case in [10]) or it could be a raw key bit from a random number generator. Bob utilizes a quantum receiver to decide on $b$ knowing which particular pair $\{|\pm \alpha_i\rangle\}$ is to be discriminated. On the other hand, Eve needs to pick a quantum measurement for her attack in the absence of the basis knowledge provided by the seed or running key. The difference in their resulting receiver performances is a quantum effect that constitutes the ground both for making $\alpha\eta$ a random cipher for direct encryption and for possible advantage creation vis-à-vis key generation. To avoid confusion, we shall use the term ‘$\alpha\eta$’ to refer only to the direct encryption system following our practice in [11]. When we want to use the same system as part of a key generation protocol, we shall refer to it as ‘$\alpha\eta$-Key’ Generation’ or ‘$\alpha\eta$-KG’. KCQ key generation is further elucidated in [9].

Note that since the quantum-measurement noise is irreducible, such advantage creation may result in an unconditionally secure key-generation protocol. In contrast, in a classical situation including noise, the simultaneous measurement of the amplitude and phase of the signal, as realized by heterodyning, provides the general optimal measurement for both Bob and Eve; thus preventing any advantage creation under our approach that grants Eve a copy of the state for the purpose of bounding her information.
We have investigated \( \alpha \eta \) scheme [4-12] described in Fig. 1, and also the CPPM (Coherent Pulse Position Modulation) scheme [4-5] of Fig. 2 that is under current experimental development at Northwestern University and NUCrypt. In this approach, a large-energy short optical pulse is coherently divided and re-combined by beamsplitters whose transmittance coefficients are controlled by a shared secret key. In the absence of a bandwidth limitation, this \( M \)-ary modulation scheme allows much greater energy advantage to be created as compared to \( \alpha \eta \) which is a binary modulation scheme from the ‘users’ point of view.

![Figure 2: Overall schematic of the CPPM scheme. An \( L \)-ary pulse is spread out according to a secret key and recombined before direct detection by Bob.](image)

There are three logical steps for fresh key generation, quantum or classical, viz. advantage creation, error correction, and privacy amplification. In advantage creation, Alice and Bob make sure they have a “channel” \( (X_A, Y_B) \) between them that is better than Eve's \( (X_A, Y_E) \). In BB84, this is obtained via intrusion level estimation from the quantum information/disturbance tradeoff on Eve's intrusion. In classical noise protocols such as that in ref. [1], it is obtained from post-detection selection of \( X_A \). In KCQ protocols, it is obtained from the optimal quantum receiver performance difference that results from having versus not having knowledge of a key \( K \) [5].
Following the creation of advantage, the users employ a perhaps interactive error correction procedure to get an error-free bit string $X_C$ between themselves. This string is compressed in the privacy amplification procedure to generate a final fresh generated key $K_g$ on which Eve would have vanishingly small information, by eliminating Eve's possible knowledge on $X_C$ from $Y_E$ and any side information she gained during protocol execution.

For quantitative security analysis, Eve's information-theoretic (Shannon) entropy on $K_g$, $H(K_g)$ conditioned on all her knowledge, is usually taken as the security measure. Such use dates back to Shannon [14] and is used in classical-noise cryptography [2] as well as quantum cryptography [3]. However, it has been pointed out [4] that it is not a good measure to use in concrete realistic cryptosystems, for the following reason. Let $p_1 \geq p_2 \geq \ldots \geq p_M$ be Eve's “error profile” on $K_g$, i.e., her probability distribution $p(K_g)$ on the $m = \log_2 M$-bit string $K_g$ that has entropy $H_E(K_g)$. If $H_E(K_g) \leq 2^{-l}|K_g|$, it is possible that her maximum probability $p_1$ of correctly identifying the whole $m$-bit string $K_g$ is $p_1 \leq 2^{-l}$ [4-5]. Thus, if Eve knows 1 bit of Shannon information on a 100-bit string, it is possible that she can guess the whole string correctly with a probability $\sim 0.01$ due to the bit correlations, a disastrous breach of security. Other measures, such as Kolmogorov distance between $p(K_g)$ and the uniform distribution $p_i = 2^{-|K_g|}$ for all $i$, have similar problems. Since it is experimentally impossible to guarantee the above $l$ to be large in a concrete system due to imperfections, a different criterion has to be adopted for realistic applications.

We suggest that Eve’s $p_1$, the largest probability in her error profile, be used as the criterion, from which other measures such as $H_E$ and “trial complexity” may be bounded [4-5]. It is clear that $p_1$ itself has to be sufficiently small for meaningful security. Furthermore, it is the quantity of interest in both detection-theoretic and information-theoretic analysis of communication systems, and in the present case both in connection with error correction by the users and
incorporation of side information by Eve as described later in this whitepaper. In addition, $p_1$ is a more appropriate benchmark for privacy amplification than the Renyi entropy $R$ used in ref [15]. It would be useful to find algorithms that would generate an $m$-bit $K_g$ with uniform distribution, i.e. true random numbers, from a longer $n$-bit string $X_c$ characterized by some measure such as $H(X_c), R(X_c)$, or $p_1(X_c)$, if not by the whole distribution $p(X_c)$. This is the generalized privacy amplification problem. It is clear that $p_1$ controls the number of uniform bits that can be obtained from $X_c$, which is $l$ for $p_1 \leq 2^{-l}$, from an openly known compression function as in privacy amplification. This is because $p_1$ cannot be decreased by a deterministic transformation of $X_c$ [5]. Indeed, the privacy amplification theorem in ref. [15] that characterizes $X_c$ and $K_g$ by $R$ instead of $p_1$ can never lead to truly random $K_g$.

Security analysis of key generation systems is actually very much of a communication and information-theoretic nature, with the usual performance concerns for the users but the opposite concerns for the attacker, from the designer's point of view. This is true from the perspectives of detection, information and coding theory. Thus, upper bounds on the error rate are desired for the users, while lower bounds are wanted for the attacker to guarantee security. While there are some lower bound results in detection, information and coding theory, they are far less developed than upper bounds. For direct encryption, the security criterion of $H(X|Y_g)$ for information-theoretic security is not adequate for various reasons, while search complexity is the usual one used in practice for which there is little rigorous theoretical result.

In this work we tried to remedy the situation but the results are not yet sufficiently strong to establish rigorous security. Note, however, that rigorous security has never been established either in conventional cryptography or in quantum cryptography [6], despite claims to the contrary.
III. Technical Accomplishments

In this section we will describe the main technical results that were developed from the work supported by this Grant. The chronological order of the publications at the end of this report will be more or less followed in the following presentation of results.

III A. General Theory of Quantum-Noise Randomized Cyphers

This is mainly provided in ref [9], but some previous results in [7-8] are also relevant. In [9] the general description of a quantum-noise randomized cipher is provided after a review of the relatively unfamiliar subject of classical symmetric randomized cipher. Previously in [8], especially its appendix A, we have given a new quantitative relation between the data security and key security in a classical randomized cipher via Shannon entropies. In [7] we have described some basic quantitative features of $\alpha \eta$ for both direct encryption and key generation. Here in ref [9] we put $\alpha \eta$ within the framework of a general quantum noise randomized cipher and relate its basic parameters to its quantitative complexity-based security under an “intelligent” search attack.

Of equal importance is the definite refutation in Section V of [9] of the claim by the Japanese group [16] that $\alpha \eta (Y00)$ is a nonrandom cipher. It appears that the consensus has been reached in Japan to our favor despite the many papers of the Japanese group which has since become silent.

III B. Upper Bound on Eve’s Error Probability

An upper bound on Eve’s optional error probability $\bar{P}_e$ on the $\alpha \eta$ seedkey $K$ under known-plaintext attacks is the main result of the Ph.D. thesis of Ranjith Nair [17], which also contains very weak lower bounds on $\bar{P}_e$ for both known-plaintext and ciphertext-only attacks on $K$. The main conclusion is given in section 4.1 of [17]. It shows the key would be pinned down with high probability and $\bar{P}_e$ goes to zero as the data length gets long. However, for the numerical values of experimental $\alpha \eta$ the bound does not become valid until the data length $n \geq 10^7$. Thus,
it is an open question whether $\alpha \eta$ is much more secure for smaller and thus more practically reasonable $n$ in a known-plaintext attack.

It is important to compare with the corresponding case of a conventional cipher such as AES. When $n = |K|$, the seedkey length in such cases, the seedkey $K$ can be determined with certainty $P_e = 0$. That is why the security of conventional ciphers depends exclusively on complexity, that it is hard to find $K$ even though a unique solution exists. In contrast to conventional ciphers, on the other hand, $\alpha \eta$ is not fully secure against ciphertext-only attack on the seedkey $K$ even when the data is completely random to Eve. This problem is addressed in the next subsection.

**III C. Fast Correlation Attack on $\alpha \eta$**

In [18] a fast correlation attack (FCC) similar to the ones extensively studied for conventional stream ciphers was described for ciphertext-only attack on the $\alpha \eta$ seedkey, which can be adapted to known-plaintext attacks also. In response we have described several possible approaches in [19] for defending $\alpha \eta$ seedkey. It should be mentioned that the FCC in [18] and later improvements by the Japanese group still has an exponential complexity $2^{K/2}$ in general, and thus poses no real threat to an $\alpha \eta$ that operates easily with much longer key than $|K| = 100$.

The major theoretical solution against ciphertext-only attack is the use of Deliberate Signal Randomization (DSR) first described in [4]. In [19] we provide a quantitative description and show that ideally it would imply full information theoretic security on $K$ against such attacks. However, DSR requires true random numbers generated at a very high speed, ten times the $\geq 1\text{Gbps}$ data rate for the current experimental $\alpha \eta$ parameters. For the future data rate $R$ with $M \alpha \eta$ bases, the random number generation speed required in $R \log_2 M$. In addition, there is the quantitative problem of dealing with boundary effects on the PSK signal circle in $\alpha \eta$ at the receiver in a concrete realistic implementation.

An alternative approach is suggested in [9] where AES is employed for the ENC box of Fig. 1, in a configuration (Fig. 2 in [9]) that seems to be still a fair comparison to conventional AES. The
security is evidently much enhanced in comparison but the improvement is difficult to quantify as the performance of conventional cipher is not quantified.

**III D. Lower Bound on Eve’s Error Probability**

Some not yet written results are presented in this subsection, which are similar to that of ref [20] but in a more useful form in terms of Eve’s optimal error probability instead of the number of spurious keys. Note that a good lower bound on $P_e$ would establish the security of $\alpha\eta$ if the numerical values turn out favorable, in contrast to upper bounds on $P_e$ which could only rigorously establish insecurity.

The story unfolds as follows. In ref [2], it was claimed that $\alpha\eta$ is insecure under heterodyne attack from their estimates of Eve’s mutual information and some analogy with Shannon’s unicity distance [14] for conventional ciphers. In response [20] we showed that their estimates of mutual information are overly optimistic for Eve and their analogy with Shannon ‘random cipher’ does not go through. As it turns out, the Shannon unicity distance $d$ [14], which he defined to be the data length at which the key of the cipher can be found, is not a useful concept because it can be rigorously shown to be infinite in almost all practical cases. It has to be extended to be a function $d(p)$ which is the data length at which the key $K$ can be found with probability $p$. The original $d$ is thus $d(0)$. With such a more meaningful definition there are no available rigorous results in the literature and it is not clear what significance Shannon’s estimate has, i.e., at what $p$ his estimate is valid.

It is exactly for this reason that Hellman, the co-inventor of public key cryptography, introduces the average number of superior key $\bar{N}_k$, the number of possible keys given the data, and lower bound it as a function of the system parameters. [22]. We have generalized his result for conventional ciphers to randomized ones applicable to $\alpha\eta$ via Theorem 2 of [20]. It is then easy to show the results in [21] fit in the discussion of $\alpha\eta$ exactly as Shannon’s [14] fit in Hellman’s [22]. As a lower bound on $\bar{N}_k$, such result could not in principle imply insecurity of the cryptosystem.
Eve’s error probability is still missing in the description of $\overline{N}_k$, a lower bound on which is still not enough to establish security in a meaningful operational sense. This is both because the security may be compromised if $\overline{N}_k$ is not very large, and especially because it is Eve’s success or error probability for a given data length $n$ that determines security. It turns there is a lower bound to $\overline{P}_e$ [23] corresponding to the Hellman result on $\overline{N}_k$. We have generalized it to cover $\alpha\eta$ as in [20], which is given as follows. Let $H(X^n)$ be the data $n$-bit entropy, $Y^n$ the $n$-sequence of continuous-variable heterodyne measurement result of Eve, and $S^n$ the $\alpha\eta$ PSK signal random variable. Then one has in general

**Theorem 1:**

$$\overline{P}_e \geq \frac{H(X^n) + H(K) + I(S^n;Y^n) - 1}{\log |K| - 1}$$  \hspace{1cm} (2)

Where $I(S^n;Y^n)$ is the mutual information. For $\alpha\eta$ it follows from (2),

**Corollary 1:**

$$\overline{P}_e \geq \frac{n(1 - U) + |K| - 1}{\log |K| - 1}$$  \hspace{1cm} (3)

Where $U = I(S_i;Y_i)$ is the single measurement mutual information which is independent of $i$. The result (3) on $P_e$ is analogous to equation (24) in [20], both being too weak to imply meaningful security for $\alpha\eta$.

**III E. Security in Key Generation**

This has been extensively analyzed and reported in ref. [5]-[6]. The main conclusion is that the quantitative security of a generated key in BB84 type protocols is completely inadequate in practice, as shown in Appendices I and II of ref. [5] and in ref [6]. Indeed, the case as we now
understand is even more damning. It can be shown that for all subsets of the generated key $K$ which are shorter than the seedkey length $|K|$ and for the whole key, Eve’s probability for success in estimating $K$ from her probe is much larger than that of a mere pseudo-random number generator such as a linear feedback shift register, for all the numerical values studied till now on concrete realistic BB84 protocols. This result also leads to the conclusion that the key security cannot be separated from the so-called composition problem, in which the generated key $K$ is used in a given context and the security in such generation/application combined context is what counts. However, the composition security of BB84 was incorrectly asserted as detailed in ref. [6].

The positive new observation in [5] that is very encouraging for the KCQ approach is that one may assume the KCQ seedkey is never available to Eve and not just during her quantum measurement stage. It is a useful Gedanken device to grant $K$ to Eve after her quantum measurement to demonstrate the possibility of key generation, but realistically there is no reason why Eve would ever know $K$ in any uncontrived scenario, not to mention all situations.

There are various new results in [5] not contained in [4], which we would not review here and would just leave [5] for reading. However, we would like to mention section II of [5] that describes the use of a pseudo-random number generator for bases determination in a qubit protocol similar to BB84. It is important because the KCQ idea can be implemented which gives the possibility of not employing intrusion level estimation in BB84 at all.

IV. Other Results

In addition to subsection IID, the list of publications includes all the specific readily usable results we have obtained from the work supported by this grant. We have also looked extensively into two areas where no major result has been obtained but which are very important areas to explore.

The first one concerns true random number generation at high speed via optical heterodyne detection of the vacuum. We believe heterodyne detection is better than homodyne in this regard because there are two degrees of freedom in heterodyne versus one in homodyne, and more
practically easy in heterodyne. We have looked into possible algorithms for generating true random numbers when realistic devices are used in the heterodyning. It appears a major development is required as all the conventional results are primarily related to complexity obtained from known “computationally hard problems” and are thus inapplicable to true random number generation.

We also investigated the security of CPPM for key generation as well as direct application. As reported in [5], it is found that a further parameter needs to be adjusted to get the great performance in the infinite bandwidth limit. For realistic bandwidths coding must be employed as the number of signals grow exponentially. Our estimate of a transmitter photon number ~100 that leads to a 20dB advantage over Eve when a Reed-Solomon code is employed is predicated on the assumption that Eve has the seedkey after her quantum measurement. As discussed in subsection IIIE, we do not actually think that is a reasonable assumption in real applications and security should be possible with larger signal energy without it. As a new Grant is being started on the CPPM scheme, we would develop the security analysis without such an assumption and also with smaller number of possible CPPM signals corresponding to our in-principle demonstration experiment that is being planned.

In conclusion, the major open theoretical questions remain for establishing meaningful but quantifiable security criterion for both KCQ key generation and direct encryption, and applying them to $\alpha\eta$ and CPPM.
V. References

VI. Executive Summary

We have obtained upper and lower bounds on Eve’s optimum error probability in finding the seedkey of a general randomized cipher applicable to the \(\alpha\eta\) cryptosystem, under both ciphertext-only attacks and known-plaintext attacks. The upper bound is a major part of a Ph.D. dissertation partially supported by this grant; the other major part of the thesis involves various results on quantum-noise randomized ciphers supported by DARPA which ended at about the time this grant started. The lower bound is entirely the result of this Grant and described in this report for the first time.

We have also developed various security results on key generation, which show in particular the lack of adequate security in concrete practical BB84 key generation. The security situation of KCQ generation, which is much more efficient and practical than BB84, has also been analyzed. There are a variety of other minor results described in the Technical Accomplishment Section of this Final Report.

The main overall conclusion is that there is actually yet no meaningful quantifiable security level in quantum and physical cryptography, for any cryptosystem that has been studied thus far. The situation is exactly the same in conventional cryptography. A lot more fundamental theoretical investigation is needed for the security quantification in quantum and physical cryptography.

Personnel

1. Horace P. Yuen, PI, Professor of Electrical Engineering and Computer Science, Professor of Physics and Astronomy, Northwestern University.
2. Ranjith Nair, graduate student, postdoc, and Research Assistant Professor, Northwestern University.
4. Koichi Yamazaki, visitor, Professor of Engineering, Tamagawa University, Japan.
5. Osamu Hirota, visitor, Professor of Engineering, Tamagawa University, Japan.
6. Masanao Ozawa, visitor, Professor of Mathematics, Nagoya University, Japan.
7. Mauro D’Ariano, visitor, Professor of Physics, University of Pavia, Italy.
Doctoral Dissertation

Ranjith Nair, Quantum Detection and Coding with Applications to Quantum Cryptography, December, 2006, Department of Electrical Engineering and Computer Science, Northwestern University.

List of Publications

Journal Papers:

3. “Comment on exposed-key weakness of $\alpha \eta$”; Ranjith Nair and Horace P. Yuen; Phys. Lett. A 372 (2008) 7091-7096

Papers in Book:
