June 9, 2013. First draft. (For distribution at the seminar at Uppsala University, Sweden, June 11, 2013).

Cracking the Bennett-Riedel secure scheme and a critical analysis of their claims about the Kirchhoff-law-Johnson-noise system

Laszlo B. Kish [1,*], Derek Abbott [2], Claes-Göran Granqvist [3]

[1] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128, USA [2] School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, South Australia, Australia [3] Department of Engineering Sciences, The Ångström Laboratory, Uppsala University, P. O. Box 534, SE-75121 Uppsala, Sweden

Content (page numbers are approximate)

Cracking the Bennett-Riedel secure scheme and a critical analysis of their claims about the Kirchhoff-law-Johnson-noise system

Laszlo B. Kish [1,*], Derek Abbott [2], Claes-Göran Granqvist [3]

[1] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128, USA

[2] School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, South Australia, Australia

[3] Department of Engineering Sciences, The Ångström Laboratory, Uppsala University, P. O. Box 534, SE-75121 Uppsala, Sweden

Abstract

Recently, Bennett and Riedel (BR) (http://arxiv.org/abs/1303.7435) has claimed that, in the Kirchhoff-law-Johnson-noise (KLJN) classical statistical physical key exchange method, thermodynamics (statistical physics) is not essential and that the KLJN scheme provides no security. They attempt to prove the no-thermodynamics view by proposing a dissipation-free deterministic key exchange method with two batteries and two switches. After showing that the BR scheme is unphysical and that some elements of the assumptions violate basic protocols of secure communications, we crack their system by passive attacks in eight different ways, with 100% success probability, and show that the same cracking methods do not work against the KLJN scheme due to Johnson noise and the Second Law of Thermodynamics. We critically analyze the other claims of BR; among others, we prove that their equations (1-3) describing zero security are incorrect for the KLJN scheme. We give mathematical security proofs for each BR attack type and conclude that the information theoretic (unconditional) security of the KLJN method has not successfully been challenged.

Introduction

Information theoretic (that is, unconditional) security [1] means that the stated security level, either perfect or imperfect (as in any physical system [2]) holds even for cases when Eve's abilities are limited only by the laws of physics. First, in 1984, quantum key distribution (QKD) [2] has been claimed to possess unconditional security and much later, in 2005, a classical physical alternative, the Kirchhoff-law-Johnson-noise (KLJN) scheme [2] appeared as a competing initiative.

Very recently, QKD co-founder Charles Bennett [3] coauthored a manuscript [4] with Jess Riedel where they present a serious criticism of the KLJN system and deny its security at the idealized conditions. Bennett and Riedel (BR) claim that, in the KLJN scheme, thermodynamics is not essential and that the method provides no security. They attempt to prove their "no-thermodynamics" claim by showing a dissipation-free deterministic key exchange method with just two batteries and two switches. Moreover, among other serious statements [4], BR claims that the quasi stationary (no-wave) limit of electrodynamics is not suitable for information transfer thus this (required) assumption [2] for the (perfect) security of the KLJN system is unphysical. Our present paper is the response to these claims. We have found some of the claims unphysical and most of the claims incorrect.

In this introductory chapter, we are making preparations for the next chapter where we will fully crack the BR system in different ways and respond to the BR arguments about the KLJN system.

First we mention the currently ongoing debates about the security of QKD because of the BR claim that the security of QKD is robust.

Then we briefly outline the Kirchhoff-law-Johnson-noise (KLJN) secure key distribution scheme and its main features.

Following that, we describe the "thermodynamics-free" key exchange system by Bennett-Riedel (BR) (a version of Davidee Antille's patented system) and the related argumentations in their papers.

1.1 Is the security of quantum encryption indeed robust?

Bennett and Riedel write [4]: "*we emphasize that quantum key distribution has been shown to be robust with imperfect components against very general attacks*". We see this situation very differently thus here we quickly mention currently ongoing debates in the QKD field.

Currently, there is an ongoing debate [5-8] about the fundamental security/non-security of existing QKD schemes. This debate was initiated by Yuen [5,8] and later joined by Hirota [6] in claiming that the security of existing quantum key distribution schemes is questionable or poor. Recently, Renner [7] also entered the discussion to defend the old security claims. (Note, Yuen [9], and Zubairy and his coworkers [10] have proposed new advanced schemes for non-QKD type secure quantum communications.)

However, the BR claim of QKD's *robust security with imperfect elements* [4] is proven to be incorrect. QKD has been cracked in much simpler ways by utilizing the imperfect nature of building elements, such as nonlinearity. Practical quantum communicators — including several commercial ones — have been fully cracked as shown in numerous recent papers [11-25]. Vadim Makarov, who is one of the leading quantum crypto crackers, says in Nature News that "*Our hack gave* 100% *knowledge of the key, with zero disturbance to the system*" [11]. This claim hits at the foundations of quantum encryption schemes because the often-claimed basis of the security of QKD protocols is the assumption that any eavesdropping activity will disturb the system enough to be detected by the communicating parties (Alice and Bob). An important aspect of these quantum-hacking attacks is the extraordinary (100%) success ratio of extracting the "secure" key bits by Eve, which indicates that the security is *not even imperfect* but simply *no-existent* against these types of attacks until proper defense units or protocol modifications are added to the scheme to restore the information theoretic security they supposedly had before these attacks were known.

In conclusion, on the contrary of the claim in [4], quantum key distribution has been shown to be *vulnerable* with imperfect components against proper attacks.

1.2 The KLJN secure key exchange system

The Kirchhoff-Law-Johnson-noise (KLJN) key distribution scheme [2, 26-39] is a statistical physical alternative of QKD and its security is based on Kirchhoff's Loop Law and the Fluctuation-Dissipation Theorem. More generally, it is founded on the Second Law of Thermodynamics, which indicates that the security of the ideal scheme is as strong as the impossibility to build a perpetual motion machine of the second kind. Potential and unique technical applications include non-counterfeitable hardware keys and credit cards via realizing Physical Uncloneable Functions (PUF) [35]; unconditionally secure computers, hardware and instruments [35,36]; and unconditionally secure smart grid [37-39]. The short summary below is based on the survey in [2].

1.1.1 The core (idealized) KLJN system and its security

The working principle of the ideal KLJN system [2,26] is as follows. The core KLJN system, without the defense circuitry (current-voltage measurement/comparison, filters, etc.) against invasive and non-ideality attacks, is shown in Fig. 1. At the beginning of each bit exchange period (BEP), Alice and Bob connect their randomly chosen resistor, R_A and R_B, respectively, to the wire line. These resistors are randomly selected by the switches from the set $\left\{R_L, R_H\right\}$, $\left(R_L \neq R_H\right)$, where the elements represent the low, *L*, and high, *H*, bit values, 0 and 1, respectively.

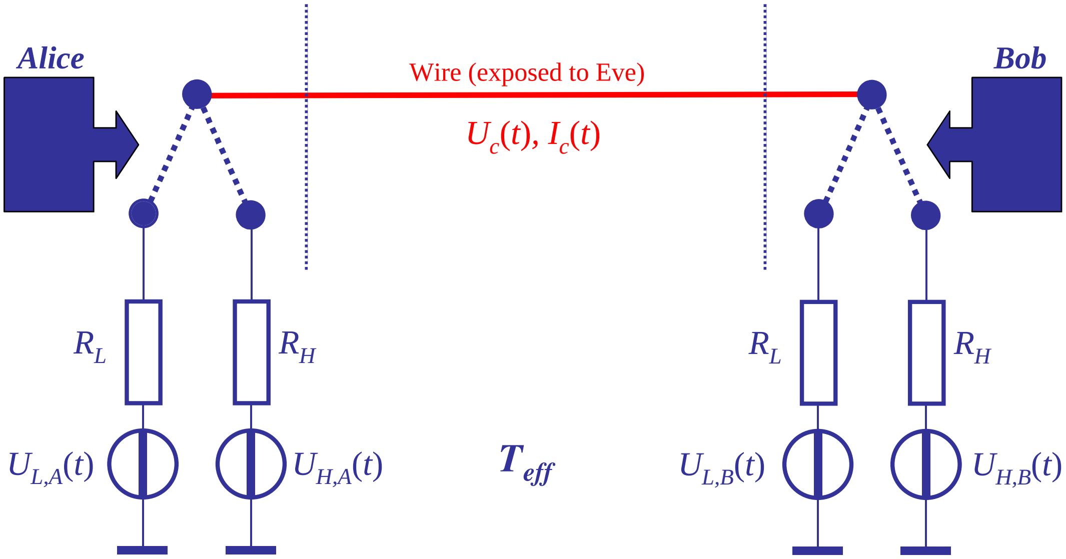

Fig. 1. Outline of the core KLJN system. Parasitic elements leading to non-ideal features and defense circuit block (current/voltage monitoring/comparison) against invasive attack are not shown/discussed here.

The Gaussian voltage noise generators—delivering white noise with publicly agreed bandwidth—represent an enhanced thermal (Johnson) noise at a publicly agreed high effective noise-temperature T_{eff} where their noises are statistically independent from each other, $\langle U_A(t)U_B(t)\rangle = 0$, and from the noise during the former BEP. During the first experimental demonstration by Mingesz, et al [29], the noise-temperature range of $8*10^8\mathrm{K} \le T_{eff} \le 8*10^{11}\mathrm{K}$ was used, which made the cable temperature insignificant even when the cable resistance was not zero.

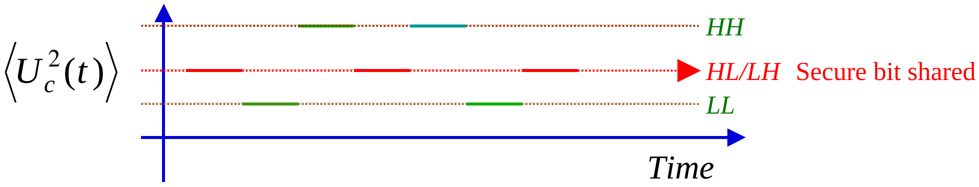

Fig. 2. The mean-square voltage (and current) has three different levels depending on the bit values and the intermediate value indicates a secure bit exchange.

For the evaluation of the bit status of the system, Alice, Bob (and Eve) can use the measurement of the mean-square voltage and/or that of the current, see Fig. 2 for the case of voltage. The situations *LH* and *HL* represent secure bit exchange [2,26], because Eve cannot distinguish between them through measurements and whenever Alice and Bob see the HH/LH situation they know that the other party has the complementary bit value, which means they learn about the full bit arrangement. Eve cannot extract this information because she does not know any of the bit values. In other words, a secure bit has been generated and shared. The bit situations *LL* and *HH* are insecure, which means these bits (50% of the executed BEPs) are discarded by Alice and Bob.

According to the Fluctuation-Dissipation Theorem, the power density spectra $S_{u,L}(f)$ and $S_{u,H}(f)$ of the voltages $U_{L,A}(t)$ and $U_{L,B}(t)$ supplied by the voltage generators in R_L and R_H are given by

$$S_{u,L}(f) = 4kT_{eff}R_L \text{ and } S_{u,H}(f) = 4kT_{eff}R_H \ , \qquad (1)$$

respectively.

In the case of secure bit exchange (*i.e.*, the *LH* or *HL* situation), the power density spectrum $S(f)$ and the mean-square amplitude $\left\langle U_{ch}^2 \right\rangle$ of the channel voltage $U_{ch}(t)$ and the same measures of the channel current $I_{ch}(t)$ are given as

$$\left\langle U_{ch,HL/LH}^2 \right\rangle = \Delta f \, S_{u,ch,HL/LH}(f) = 4kT_{eff} \frac{R_L R_H}{R_L + R_H} \Delta f \ , \qquad (2)$$

and

$$\left\langle I_{ch,HL/LH}^2 \right\rangle = \Delta f \ \ S_{i,ch,HL/LH}(t) = \frac{4kT_{eff}}{R_L + R_H} \Delta f \ , \qquad (3)$$

respectively, where Δf is the noise-bandwidth.

1.1.2 The security of KLJN is based on the Second Law of Thermodynamics

It should be observed that during the *LH* and *HL* cases, due to linear superposition, the spectrum given by Eq. (2) represents the sum of the spectra at two particular situations, *i.e.*, when only the noise generator of R_L is running one gets

$$S_{L,u,ch}(f) = 4kT_{eff}R_L \left(\frac{R_H}{R_L + R_H} \right)^2 \ , \qquad (4)$$

and when the noise generator of R_H is running one has

$$S_{H,u,ch}(f) = 4kT_{eff} R_H \left(\frac{R_L}{R_L + R_H} \right)^2 \ . \tag{5}$$

For Eve, to identify that which end has the R_L and R_H the measurement and evaluation of a physical quantity offering *directional information* is necessary. In the idealistic case, the only directional information is the *direction of the power flow* from Alice to Bob (or from Bob to Alice depending on the choice of positive current direction). However in thermal equilibrium, this power $P_{A \to B} = \langle U_{ch}(t) \ I_{ch,}(t) \rangle = 0$, which is required by the *Second Law of Thermodynamics*. In other words, the ultimate security of the KLJN system against passive attacks is provided by the fact that the power $P_{H \to L}$, by which the noise generator of resistor R_H is heating resistor R_L, is equal to the power $P_{L \to H}$ by which the noise generator of resistor R_L is heating resistor R_H [2,26,32]. Thus the fact that the net power flow $P_{A \to B} = P_{L \to H} - P_{H \to L} = 0$ can easily be shown from Eqs. (4,5) for the noise-bandwidth of Δf :

$$P_{L \to H} = \frac{S_{L,u,ch}(f)\Delta f}{R_H} = 4kT_{eff} \frac{R_L R_H}{(R_L + R_H)^2} \Delta f \ , \tag{6a}$$

and

$$P_{H \to L} = \frac{S_{H,u,ch}(f)\Delta f}{R_L} = 4kT_{eff} \frac{R_L R_H}{(R_L + R_H)^2} \Delta f \ . \tag{6b}$$

The equality $P_{H \to L} = P_{L \to H}$ is in accordance with the Second Law of Thermodynamics: the impossibility to build a perpetual motion machine of the second time. In other words, it is as difficult to crack the ideal KLJN system as to build a perpetual motion machine [4].

This security proof against passive (listening) attacks holds only for Gaussian noise (the same statistics thermal noise has), which has the well-known property that its power density spectrum or autocorrelation function already provides the maximum achievable information about the noise, and no higher order distribution functions or other tools (such as higher-order statistics) are able to serve with additional information.

The required duration of the bit exchange period (BEP) and the error probability of the bit exchange between Alice and Bob is determined by the following issue. In the case of the *LL* bit status of Alice and Bob, which is not secure situation, the channel voltage and current satisfy:

$$\langle U_{ch,LL}^2 \rangle = \Delta f \, S_{u,ch,LL}(f) = 4kT_{eff} \frac{R_L}{2} \Delta f \quad \text{and} \quad \langle I_{ch,LL}^2 \rangle = \Delta f \ S_{i,ch,LL}(t) = \frac{2kT_{eff}}{R_L} \Delta f \quad , \tag{7}$$

while, in the case of the other non-secure situation, the *HH* bit status, the channel voltage and current satisfy:

$$\left\langle U_{ch,HH}^2 \right\rangle = \Delta f \, S_{u,ch,HH}(f) = 4kT_{eff} \frac{R_H}{2} \Delta f \quad \text{and} \quad \left\langle I_{ch,HH}^2 \right\rangle = \Delta f \;\; S_{i,ch,HH}(t) = \frac{2kT_{eff}}{R_H} \Delta f \qquad (8)$$

During key exchange in this classical way, Alice and Bob must compare the predictions of Eqs. (2,26,34) with the actually measured mean-square channel voltage and current to decide if the situation is secure (LH or HL) while utilizing the fact that these mean-square values are different in each of these three situations (LL, LH/HL and HH). If the situation is secure, Alice and Bob will know that the other party has the inverse of his/her bit, which means, a secure key exchange takes place. To make an error-free key exchange, Alice and Bob must use a sufficiently large statistics, which means long-enough bit-exchange-period (BEP). Fortunately, the errors decay exponentially versus the duration of the BEP [34]. A new ("intelligent") KLJN protocol [31] that requires also circuit calculations by Alice and Bob is able to reduce the BEP without increasing the error probability.

1.1.3 On active (invasive) attacks and attacks utilizing non-idealities

It should be observed [2,26,28,29,32] that deviations from the shown circuitry—including invasive attacks by Eve, parasitic elements, delay effects, inaccuracies, non-Gaussianity of the noise, *etc.*—will cause a potential information leak toward Eve. The circuit symbol “line” in the circuitry represents an ideal wire with uniform instantaneous voltage and current along it. Fortunately the KLJN system is very simple, implying that the number of such attacks is strongly limited. The defense method against attacks utilizing these aspects is straightforward and it is generally based on the comparison of instantaneous voltage and current data at the two ends via an authenticated communication between Alice and Bob.

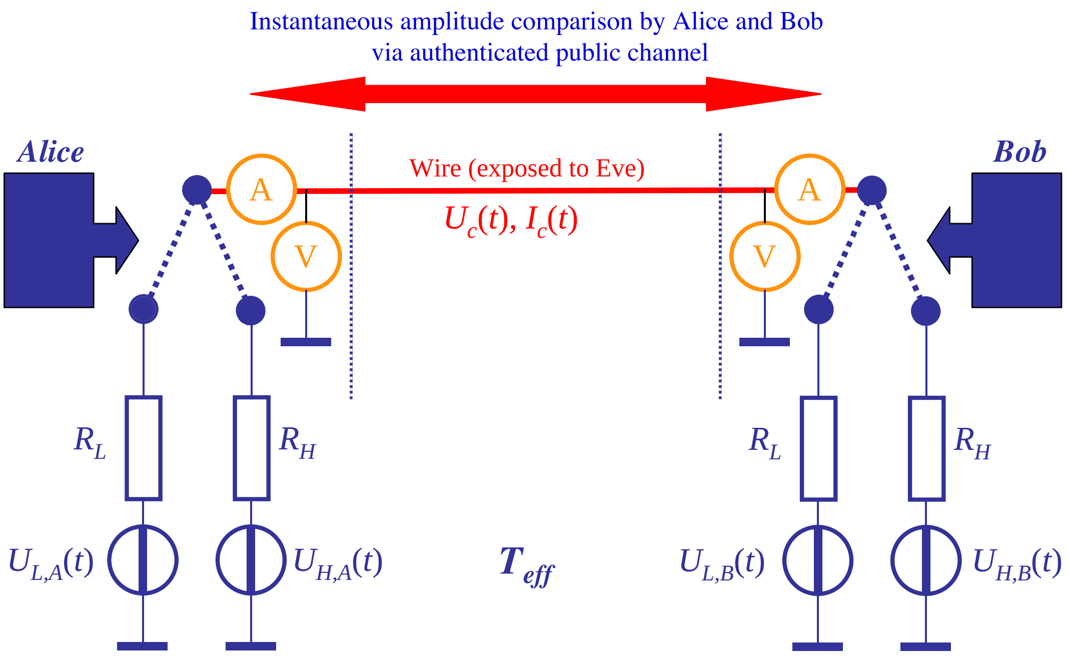

Figure 3. KLJN system minimally armed against invasive (active) attacks including the man-in-the-middle-attack. Note, important practical additions against hacking, such as line filters, blinding detectors, etc., are not shown.

These attacks [2,40,43,45] are not subject of the present paper and we refer to our relevant responses where they have been analyzed [2,32,41,42,44] and the misconceptions (and errors) are corrected. In survey [2], existing invasive attacks by other authors and us have been reviewed.

Finally, it is important to emphasize that, Alice and Bob knows Eve's best measurement information because that is represented by comparison of voltage and currents at the two ends. If Eve and them use the best available protocol and the security of a certain bit is compromised, this is *known also by Alice and Bob* therefore they can decide to discard the bit to have a clean secure key. This is a new and unique situation in cryptography, which raises a number of new research questions (see, for example [32]).

1.1.4 Near-to-perfect information theoretic security in practical KLJN systems

Of course, perfect security of any physical key exchanger exists only at the ideal (mathematical) conditions. For example, quantum encryption can theoretically offer perfect security only in the zero photon emission rate (that is zero bit exchange rate) and zero detector and channel noise limits, which are unphysical and can never happen in a real system. The KLJN system is not exception from this rule [2,26,32]: it offers perfect security only at zero bandwidth or distance due to transients, cable resistance and capacitance, etc. However, similarly to pro-QKD claims, the parameters of the KLJN building elements and protocol can be chosen so that the perfect security limit can arbitrarily be approached (though never reached). The general situation in the non-idealistic case is that a miniscule DC signal component buried in a large Gaussian noise must be detected by Eve from a small statistics limited by the BEP period. When the parameters approach the idealistic situation, the ratio of the DC signal amplitude and the RMS amplitude of the noise converges to zero in a *power law decay* (typically with exponents -1 or -2) [2,29,42,44] fashion versus the invested resources (such as wire volume, current/voltage resolution, BEP duration, etc).

When we want to compare the security of the shared key we must compare the distribution of the probability of successfully guessing each possible key sequence of the *N*-bit long key (2^N different sequences) with that of the perfect key with uniform distribution. For this, the *statistical distance* Δ [46] ,

$$\Delta\left(E,I\right)=\max_{j=1,\dots,2^N}\left[P(E_j)-P\left(I_j\right)\right] \ , \tag{9}$$

can be used, where *E* and *I* represent Eve's extracted key and the perfect key, respectively; and $P(E_j)$ and $P\left(I_j\right)$ are the probabilities of correct guessing with the *j*th version of Eve's key and that of the perfect key. For $\varepsilon \geq 0$ the key exchange has ε - *security* (see Hirota [6]) if

$$\Delta\left(E,I\right)\leq\varepsilon \quad . \tag{10}$$

KLJN provides identically distributed (IID) sequence of random variables as key bit values thus:

$$\Delta\left(E,I\right)=\max_{j=1,\dots,2^N}\left[P(E_j)-P\left(I_j\right)\right]=p^N-0.5^N \tag{11}$$

where *p* is Eve's probability of successfully guessing bits. In the *non-ideal cases* with information leak, *p* can be given as

$$p=0.5+q \ , \tag{12}$$

where $0<q<<0.5$ and $q=0$ would mean a perfectly secure key. The above-mentioned power scaling behavior of the ratio of Eve's extracted small DC signal amplitude (such as a correlation coefficient) and the large RMS amplitude of the noise burying it (due to the small statistics) versus the invested resources yields also a *power law decay* of *q* versus the invested resources, thus the deviation of the statistical distance Δ from the ideal zero value is reduced accordingly.

Here, as an example, we analyze Δ for the wire resistance attack ([40], corrected in [41]) while, in Chapter 2, we will show more examples.

In the case of *non-zero wire resistance,* if the capacitive effects are compensated or can be neglected due to the actual bandwidth, in the case of fixed distance and bandwidth, *q* is proportional to the inverse of the square of wire diameter *D*, that is, with the inverse of the wire's volume *V*, which is one of the *resources* to be used:

$$q = \vartheta_w V^{-1} \quad , \tag{13}$$

where ϑ_w is a constant valid for the wire resistance attack. Then

$$\Delta = \left(0.5+q\right)^N - 0.5^N = 0.5^N\left[\left(1+2q\right)^N - 1\right] \cong 2Nq0.5^N = 2N\vartheta_w V^{-1} 0.5^N \quad , \tag{14}$$

where the last approximation is valid for $q \to 0$. Eq. 14 indicates that Δ decays exponentially with increasing *N* and inversely with the wire volume *V*. The value shown by Eq. 14 is reached without privacy amplification.

For ε - *security* , $\Delta \le \varepsilon$, the required *q* is given as:

$$q(\varepsilon, N) = \frac{\vartheta}{V(\varepsilon, N)} \le \frac{\varepsilon}{2N} 2^N \quad . \tag{15}$$

At the experimental demonstrations [29], during wire resistance attacks with wire resistance of 2% of the loop resistance during secure bit exchange, $q = 0.025$ was measured. With a 1000 bit long shared key, the resulting security measure is $\varepsilon = 4.7 * 10^{-300}$ (10^{-299} - security). It is important to note that this experimentally demonstrated security level is reached *without privacy amplification*.

Finally, we note that there are advanced protocols that can enhance the security or limit the required resources in efficient ways while the scaling shown above does not change. Below is a short list of basic security features and advanced protocols proposed up to now:

a) Ideal KLJN system with passive attacks: The Second Law of Thermodynamics [2,26].

b) Non-ideal KLJN systems with passive or active (invasive) attacks:

- *Transient protocols*: Random-walk from equal resistances [31]; voltage ramping/timing [2,29].
- *Selecting the noise bandwidth* versus the value of wire resistance and wire capacitance [2,29].
- *General defense that works in any situation including hacking*: comparison of instantaneous voltage and current amplitudes and discarding any bits where they differ, or where they give out information to Eve, even if they give out wrong information. Note: specific protocols apply for different hacking attacks.
- *Privacy amplification* (XOR-ing key bit pairs [33]).

- *Enhanced KLJN protocols*, for example, the intelligent (iKLJN) and keyed (KKLJN) methods [31].

1.2 Summary of the Bennett-Riedel arguments regarding the KLJN system

Bennett and Riedel [4] (BR) have presented an extensive analysis, which is invalid but very useful to elucidate the differences between simplistic or irrelevant model approaches and the real physics of the KLJN system.

The outline of the BR claims is as follows. They state that the no-wave limit (quasi static electrodynamics) is unphysical for signal propagation. Based on this statement, they declare that Eve can separate and measure the "orthogonal" wave components propagating from Alice to Bob and from Bob to Alice, respectively. They also state that the KLJN system is deterministic thus, when Eve's measurements of the two wave components are limited only by the laws of physics, Eve has a full description of the whole system, including Alice's and Bob's history. To support the above claims BR state that thermodynamics and noise are not essential in the KLJN system (thermodynamics would kill determinism due to fluctuations) and to support this view, they construct a deterministic, thermodynamics-free key exchanger (originally of Davide Antilli), which looks similar to the KLJN system without resistor and where two of the four noise voltage generators are removed and the remaining ones are replaced by batteries with known identical voltage. Moreover, BR proposes and passive correlation-measurement based attack and an active current-extraction attack against the KLJN system.

After shortly describing the BR claims mentioned above, in chapter 2, we refuse all these claims and show the proper physics of the KLJN scheme.

1.2.1 BR claim: There is no information transfer in a wire in the no-wave (quasi static) limit

BR writes: "*We believe this no-wave limit is inappropriate and nonphysical for analyzing communication protocols (even as a mathematical idealization) because if propagating waves are excluded there is no way for information to get from Alice's side of the circuit to influence Bob's side, or vice versa.*"

Based on this belief, they declare that Eve can separate and measure the "orthogonal" wave components propagating from Alice to Bob and from Bob to Alice, respectively.

After surveying the relevant physics facts about waves, directional couplers for signal separation and the no-wave limit in section 2.1, we first refuse the above belief in section 2.2. Furthermore, we show what does physics say about signal propagation in the no-wave (quasi static) limit.

1.2.2 Bennett-Riedel theoretical claim: the KLJN system offers no security

BR set up three equations about the KLJN scheme by vaguely arguing about the KLJN system while invoking the deterministic nature of the Maxwell equations and *neglecting the stochastic nature of Johnson noise and the secret/random choice of the resistors*. Thus it is not surprising that they arrived at the conclusion that KLJN offers no security. Here we discuss only the first and the last equations (the second one is redundant) and BR's main conclusion.

The conditional information $H\left(F|G\right)$ represents the remaining uncertainty about the set of data *F* when the set of data *G* is known. If *G* completely determines *F* then $H\left(F|G\right)=0$ while, if *G* provides no information about *F* then $H\left(F|G\right)=H\left(F\right)$. BR's first equation is:

$$H\left(X|Z_A\right)=H\left(X|Z_A,Z_B\right)=H\left(X|Z,Y\right)\ , \tag{16}$$

where *X* a variable gives a full description of the physical quantities on Alice's side of Eve's location during the BEP, including waves traveling toward her, away from her, and all of her equipment, noise and memory; and *Y* is the same for Bob. Z_A and Z_B are the wave components propagating from Alice and Bob (as observed by Eve), respectively, and Z=(Z_A, Z_B) represents both. Note in the BR paper [4] either *Z* is incorrectly indexed or the *X* and *Y* must swap.

Eq. 16 (if valid, which assumes Z_A and Z_B can separately measured) means that the uncertainty about Alice's "full description" *X* does not change if Eve expands her knowledge of wave Z_A coming for Alice by the knowledge of wave Z_B coming from Bob, or even if this is expanded by the knowledge of the total description of Bob's data *Y*.

Note, the $H\left(X|Z_A\right)=H\left(X|Z_A,Z_B\right)$ part of Eq. 16 contradicts with BR's proposed "passive correlation attack" [4], which requires the knowledge of both Z_A and Z_B and thus implies $H\left(X|Z_A\right)>H\left(X|Z_A,Z_B\right)$.

The mutual information $I\left(X;Y\right)$ of *X* and *Y* measures how much the knowledge of *X* or *Y* tells us about the other.

As a consequence of Eq. 16, with further argumentations, BR deduces the following equation for the *conditional* mutual information between *X* and *Y* , conditional on *Z*:

$$I\left(X;Y|Z\right)=H\left(X|Z\right)-H\left(X|Z,Y\right)=0 \tag{17}$$

which, if valid, would mean that after measuring the two waves Z=(Z_A, Z_B), Eve's information about *X* (Alice's full description) is not increased by learning *Y* (Bob's full description). That would mean that after measuring the two waves Z=(Z_A, Z_B), Bob's information about Alice would not be more than Eve's information about here. The same argumentation would also work in the opposite direction thus the KLJN system would not offer any security.

We will see that BR's equations are invalid even in the wave limit (which is illegal operational conditions) due to multiple reflections, the secret reflection coefficients at the two ends, and the secret noise amplitudes, which always guarantee that Alice and Bob know more than Eve.

Most importantly, in the no-wave limit, these equations (16,17) *cannot even be written* because the propagating relaxations (which are not waves) Z_A and Z_B cannot separately be measured, just their sum.

1.2.3 The Bennett-Riedel "thermodynamics-free" key exchanger scheme

One of the major claims of Bennett and Riedel [4] is that, in the KLJN method, thermodynamics and noise are not essential for the security. They attempted to prove this claim by constructing a deterministic key exchange method with two voltage generators and two switches, see Fig 4. Note this scheme is known, it is called *Orlando System* and it is created and patented [47] by Davide Antilli in 2005. Because of the BR arguments, we call this scheme BR system in the present paper even if it is Antilli's invention.

In the idle mode between bit exchange periods (BEP) the switches are in the *I* position, see Fig. 4, thus the wire channel is grounded. At the beginning of the BEP, Alice and Bob randomly choose between the switch positions *L* or *H* representing the corresponding bit values. At the middle of the BEP, they change their bit value. If the chosen sequence of bit values happens to be the same then the voltage on the wire will be zero for half of the BEP and these BEP events will be disregarded. If the choices are different, then the voltage is U_0 for the whole BEP.

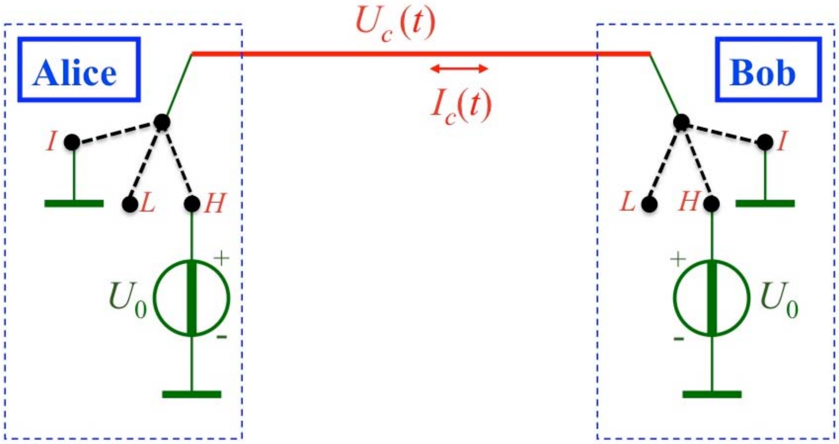

Figure 4. The (Antilli-) Bennett-Riedel system. While this is Antilli's invention, we call it BR system in the present paper.

BR has three statements about the system, which will be important later:

i) "*The wires and voltage sources are taken to be ideal, with zero thermal noise*", and as a consequence:

ii) "*Thermodynamics and noise do not play a role*."

iii) The BR system is secure in the "no-wave" limit provided in a special way: Eve waits with her measurements until the transients decay.

We will see that statements *i-iii* lead to an unphysical situation, namely, *Eve must wait for infinite time before she may start listening.*

Furthermore, note, *iii* is an illegal assumption in *unconditionally secure* communications because then Eve can only be limited by the laws of physics. Thus itself statement *iii* would imply only a conditional security.

In Sections 2.1 and 2.2 we show why the BR scheme is unphysical and also we fully crack it in various ways, while we show that the KLJN scheme stays unbroken due to the Second Law of Thermodynamics and noise.

1.2.4 The Bennett-Riedel wave-transient based attack before the steady state reached

BR writes [4]:

Thus, while the steady state mean square noise voltage in the original KLJN protocol does not allow Eve to distinguish between the LH and HL settings of Alice's and Bob's resistors she can distinguish them using (a) transient waves created by the switching action before the steady state is established.

For example Bob's resistor affects the phase and amplitude correlations between a right-traveling wave at time t and its left-traveling echo at time $t+\Delta$, *where* Δ *is the transit time from Eve to Bob and back, with the echo vanishing only if the resistor is perfectly impedance matched to his end of the line.*

Note, no concrete protocol with quantitative, testable evaluation scheme has been shown by BR. This is unfortunate because, by working out such a protocol one can see that in the no-wave limit, such transients would represent minuscule information for Eve about Alice's and Bob's status, even if such propagating signal components (not waves) could be measured, because of the limited information about the noise within a small fraction of its correlation time and the unknown additive noise and reflection at the other end, moreover, that even this minuscule information would converge to zero for decreasing noise bandwidth and/or reduced wire length. The difference of the actual statistical distance of the key from that of a perfectly secure key will vanish in an exponential fashion versus the required resources, similarly to QKD, see section 2.6.

In section 2.6, when we analyze this problem and provide a security proof, we will also discuss some of the extra transient handling protocols that eliminate the remaining information for Eve.

1.2.5 The Bennett-Riedel passive time-correlation attack in the no-wave limit

BR writes [4]:

Thus, while the steady state mean square noise voltage in the original KLJN protocol does not allow Eve to distinguish between the LH and HL settings of Alice's and Bob's resistors, she can distinguish them using (b) time correlations in the steady-state distribution of traveling waves resulting from the fluctuations that give rise to Johnson-Nyquist noise. For example Bob's resistor affects the phase and amplitude correlations between a right-traveling wave at time t and its left-traveling echo at time t+Δ, where Δ is the transit time from Eve to Bob and back, with the echo vanishing only if the resistor is perfectly impedance matched to his end of the line.

We will analyze this problem in section 2.7 and give a security proof showing that the difference of the actual statistical distance of the key from that of a perfectly secure key will vanish in an exponential fashion versus the required resources.

1.2.6 Current extraction/injection based active (invasive) attack

BR writes [4]:

...she (Eve) *could still learn the key by an active steady-state attack in which she would place a very high-resistance shunt between her node and ground, and monitor the direction of current flow into it. Of course Alice and Bob could try to detect this weak leakage current also, and abort the protocol if they found it. The result would be an unstable arms race, won by whichever side had the more sensitive ammeter, not the sort of robustness reasonably expected of a practical cryptosystem.*

Note, this attack of BR is valid only against the BR system because in the KLJN system the *direction* of the current flowing into the shunt resistor does not provide any information, as this is current is a Gaussian noise process with zero mean, which implies a perfect symmetry around zero. What BR would want to say for the KLJN system is to determine if the change of the RMS channel current will be dominant in the direction of Alice or Bob because the shunt resistor will cause a small imbalance between the mean-square currents in the two directions.

Because such miniscule current difference is very difficult to measure one of us (LK) has proposed a more efficient attack of this type in the foundation paper of KLJN [26]: a separate noise current generator was proposed instead of a shunt resistance and the evaluation of the cross-correlations between the injected current and the channel currents at the two sides of the injection. These cross-correlations determine which end has the low and which one the high resistance. This attack was already disregarded as inefficient already in the foundation paper because Eve would need ages to make a sufficient statistics for a reasonable decision however she only have the short duration τ of the BEP before the process ends. In section 2.8, we mathematically analyze this attack and give a security proof against it.

Discussions and Results

The flow of analysis and argumentation in this chapter are as follows. First, in section 2.1, we survey the well-know facts about the physics related to the no-wave (quasi static) limit of electrodynamics and facts about information transfer in that limit. Then, in section 2.2, we refuse the BR claim that there is no information transfer in the quasi-static (no-wave) limit. In 2.3, we analyze BR's equations [16,17] indicating zero security and we show that they are invalid for the KLJN system not only in the no-wave limit but also in the wave limit, though they are valid for the BR thermodynamics-free system. In 2.4, we show that the BR thermodynamic-free key exchanger is unphysical because transients will oscillate for infinite time in the wire. In 2.5, we analyze the real (physical) BR system and show eight different ways to fully crack it. We also show there that none of these ways of cracking work against the KLJN system due to the Johnson noise and the Second Law of Thermodynamics, which proves that thermodynamics is essential for the security of KLJN. In 2.6 we show that BR is incorrect by writing that the wave-transient attack would crack the KLJN system and we show that, here too, the statistical distance converges to the 2^{-N} limit (representing perfect security) versus the invested resources. In 2.7 and 2.8, we show why BR's passive correlation attack does not work in the KLJN system and why BR's current extraction attack fails to change the exponential convergence of the trace distance to the 2^{-N} limit. Finally, in section 2.9, we have some general remarks about hacking protection.

2.1 Physics facts: information, propagation, and wave couplers in the quasi static limit

In sections 2.1.1-2.1.4, we clarify what is a wave in physics; what are the conditions for the existence of a wave; what is a *quasi-static electrodynamics* [48] represented by circuit symbols; and discuss if electronic circuits are able to transfer signals and information in the quasi-static (no-wave) limit. We also discuss the type of delayed signal propagation at the no-wave limit and the inefficiency to separate directions with directional couplers [49] there.

2.1.1 The mathematical definition of waves in physics

The physical definition of a wave is a propagating amplitude disturbance $U(x,t)$ that is the solution of the wave equation:

$$c^2 \frac{\partial^2 U(x,t)}{\partial x^2} = \frac{\partial^2 U(x,t)}{\partial t^2} \quad , \tag{18}$$

where *c* is the phase velocity (propagation velocity if no dispersion is present). The essential physical dynamics of waves based on the oscillating wave energy between two energy types, such as the electrical and magnetic field energies during propagation. If only one of these energy types takes part in the propagation or if the propagation is not based on the bouncing of the energy between these two fields then the propagating field disturbance is not a wave but merely a *near-field* oscillation with retardation effects.

Let us discuss a wire with finite size *L*. The wave equation (18) has solutions only for frequencies

$$f \geq f_m = \frac{c}{2L} \quad . \tag{19}$$

In other words, propagating field disturbances with frequency components below the minimal wave frequency f_m are not waves. We agree with BR that there is a propagation and corresponding time delay (retardation) however the propagation entities are not waves but field relaxations with the consequences outlined below. Thus BR's wording about propagating "orthogonal" wave components that can be separated in the two directions is simply unphysical and leads to incorrect equations and conclusions. Furthermore, when KLJN operates in the "no-wave limit" of the KLJN system, that means the

$$f << f_m = \frac{c}{2L} \tag{20}$$

condition [2,26], thus Bennett-Riedel are correctly using the name *quasi-static* term to describe this situation. However, in the *quasi-static electrodynamics* [48] limit it is incorrect to classify the propagating disturbances as waves, because they are neither the solution of the wave equation nor the electrical and magnetic field has the wave energy bouncing back and forward between them during propagation.

2.1.2 The quasi-static limit of electrodynamics and electrical circuitry symbols with lumped elements

The *quasi-static electrodynamics* [48] and Eq. 20, is the base of the operation and circuit drawings of any electrical circuit with lumped elements. The physical implication is that, along a wire line in a circuit drawing and the corresponding wire line in the realized circuit, at a given moment, the instantaneous current and the voltage amplitudes are virtually homogeneous and retardation effects (including waves) can be neglected. Without this, everyday electrical engineering design of circuits with lumped elements would be invalid and impossible.

2.1.3 Signal propagation in the no-wave (quasi static) limit

After this preparation it is obvious that the BR claim that, without waves in the wire there is no information transfer, is not only unphysical but also contradicts to everyday experience. No landline phones, no computers and other electrical circuits with lumped elements would be able to function and process information if the BR claim would be true. In conclusion, the quasi-static (no wave) limit [48] is a physically *valid* working condition for the KLJN system and not unphysical as BR claims.

2.1.4 Further implications of the quasi-static (no-wave) limit: directional couplers, etc.

The implications are tremendous for a passive attack with *wave-based* directional couplers to extract and separate the signal components in the two directions. Wave-based directional couplers simply don't work in the quasi-static limit. Even *in the wave limit*, the cancellation of the irrelevant signal component would be strongly frequency dependent because it is determined by the successful destructive interference of wave components in the coupler [49]. The couplers with good directivity have $\lambda_0 / 4$ size, where $\lambda_0 = c / f_0$ and f_0 is the frequency of optimal working. For longer wavelengths (smaller frequencies) the system executes a Rayleigh scattering and, accordingly, the separation of intensities will decay in a power function scaling $\left(f^4 \right)$ fashion.

Finally, it is important to note that there are also *non-wave-based* directional couplers, which are able to separate signals coming from two different directions in the wire. These are working with lumped elements (transformers or active devices) and efficient in a wide range of frequency. Their working principle is to cancel the signal of the irrelevant direction by subtracting from the channel voltage another voltage that is induced by the channel current. However, they all fail with the KLJN key exchanger because, for a proper operation to learn Alice's voltage spectrum, the designer must know the exact value of the driving resistor of Alice. If instead, Bob's resistor value is used, then the resulting signal voltage will be different and its spectrum will match Bob's noise spectrum. This fact is *again* the consequence of the Second Law of Thermodynamics, which guarantees that the crosscorrelation of the channel voltage and channel current is zero, and this leads to statistically independent channel voltage and current due to their Gaussian nature. In conclusion, non-wave-based directional couplers provide no useful information for Eve.

2.2 Denial of the BR claim about no information transfer in the no-wave limit

As it has been already shown in section 2.1.3, there *is* information transfer in the no-wave limits and this fact is supported by common experience, see Eqs. 18-20. Therefore, the quasi-static limit is physical in an information processing system.

2.3 Invalidity of BR's equations and the correct equations

Below, we show that the BR equations in both the wave limit and the no-wave (quasi-static limit) are invalid for the KLJN system.

2.3.1 The wave limit and the Pao-Lo Liu key exchange system

It is important to note that the BR system's (Fig. 4) default operation is within the wave limit due to the abrupt switching of the voltage (see section 2.4) and the generated high-frequency products and the BR equations (16,17) are valid for the BR system, which indeed does not offer any security, see in section 2.5.

While the wave limit is illegal operational condition for the KLJN system, therefore unimportant, we mention that there is a software-based protocol working in the wave limit, the Pao-Lo Liu key exchange method [50-52], which is inspired by the KLJN system but it does not utilize the Second Law. Random number samples of infinitesimally slow noises (in the ideal situation) at Alice's and Bob's site are sent and reflected with random sign of the reflection coefficient. There, Alice's reflection coefficient and the noise intensity added by her are chosen so that, in the steady-state mode of ideal conditions, BR's proposed correlation attack [4] between the incoming and outgoing waves does not yield information for Eve. The relevant relation for the Liu protocol in the ideal situation is

$$H\left(X\middle|Z_A\right)=H\left(X\middle|Z\right)=H(X)>0 \qquad (21)$$

instead of the zero-security situation $H\left(X\middle|Z_A\right)=H\left(X\middle|Z\right)=0$ implied by the BR considerations and Eq, 17. Furthermore, surprisingly, the Liu system seems to satisfy

$$I\left(X;Y\middle|Z\right)=H\left(X\middle|Z\right)-H\left(X\middle|Z,Y\right)>0 \qquad (22)$$

in the steady state at the idealistic limit. The Liu system has other weaknesses stemming from the wave limit, that is, the distinct observability of Z_A and Z_B.

Finally, returning to the KLJN system but still staying in the wave limit, we have the following comments. If only the waves from Alice (Z_A) are known, that particular situation provides less information about Alice's total description than the situation when the waves (Z_B) from Bob are also known because Z_A has limited information about the reflection coefficient (and the resistance determining it) at Alice's side, while, in accordance with BR's correlation attack, the cross-correlation of Z_A and Z_B, if both are known, can potentially provide more information about that $H\left(X|Z_A\right) > H\left(X|Z\right)$. (Note, the BR attack and its justification contradicts to this part of their own equation (1) (Eq. 16), which claims that adding Z_B to the knowledge of Z_A does not help Eve). The duration of the BEP is limited by the KLJN protocol thus the $H\left(X|Z_A\right) > H\left(X|Z\right) > 0$ relation applies in Eq. 21. However, the example of the Liu system implies that the correct relation is: $H\left(X|Z_A\right) \geq H\left(X|Z\right) > 0$.

2.3.2 The BR equations in the no-wave (quasi static) limit

The BR equations do not exist in the quasi-static limit because Z_A and Z_B are not observable separately [49]. Directional couplers that are able to separate such waves would produce

$$Z_A^{'} = Z_A + (1-\kappa)Z_B \qquad (23)$$

$$Z_B^{'} = Z_B + (1-\kappa)Z_A \qquad (24)$$

outputs with $\kappa \propto 1/f^2$. The largest separation would be at the high cut-off frequency B_{kljn} of the noise bandwidth. As we already pointed out in section 2.1.4, this will lead to an unconditional ε - security $\left(\varepsilon \propto B_{kljn}^4\right)$, results of the same nature as Eqs.-14,15. The resources used are the duration τ of the BEP and the length of the key $\left(\varepsilon \propto 2^{-N}\tau^{-4}\right)$.

Finally, we set up the correct relations. The conditional information for the KLJN system is:

$$H(X) > H\left(X|U_c,I_c\right) > H\left(X|U_c,I_c,Z_A^*\right) >> H\left(X|U_c,I_c,Z_A^*,Y\right) > 0 \qquad (25)$$

where $U_c(x,t)$, $I_c(x,t)$ are the current and voltage amplitudes along the line in the *steady state* where the dependence on *x* is miniscule and converges to zero for $B_{kljn} \to 0$; $Z_A^*(x,t)$ is the initial transient disturbance running from Alice toward Bob until the Bob's end is reached and Bob's unknown noise is mixed to it. The last entropy term expresses the fact that Bob, by knowing his own total description is able to do an almost perfect guess [31] of Alice's description *X*, however this entropy is still larger than zero because there are still errors even in this case [31,34] thus a small uncertainty remains.

Correspondingly, the conditional mutual information satisfies:

$$I\left(X;Y\middle|U_c,I_c,Z_A^*\right)=H\left(X\middle|U_c,I_c,Z_A^*\right)-H\left(X\middle|U_c,I_c,Z_A^*,Y\right)\gg 0\ . \qquad (26)$$

2.4 Proving that the Bennett-Riedel key exchanger is unphysical

It is easy to see that BR's key exchanger is unphysical in its present form (Fig. 4). Let us consider, how long does Eve have to "graciously wait" for the end of the switching transients before she can measure? The answer is infinite time because the transient will bounce back from the two ends of the line: with the same sign from the open end and alternated sign from the end terminated by the battery.

This is a clear proof that without thermodynamics and the loss/energy dissipation it implies, even the BR key exchanger cannot function, even if we allow to violate the basic rules of security that Eve is allowed to measure whenever she can and wants, and instead we force Eve to wait until the transients decay, which then takes an infinitely long time.

In conclusion, the BR scheme is unphysical and, to make it physical, we must realize that there are losses in a "physical" physical systems and the related energy dissipation is controlled by thermodynamics.

2.5 Eight ways to crack the Bennett-Riedel key exchanger by passive attacks

Below, we show eight ways to crack the BR thermodynamics-free system with 100% success rate and point out that the same cracking methods do not work with the KLJN system due to thermodynamics and noise.

2.5.1 Five universal, energy/current-flow-analysis-attacks

To avoid the problem of waves, Alice and Bob uses proper voltage envelopes (they may also use filters) to avoid frequency components in the wave limit, or to keep them exponentially small, like at a Gaussian voltage envelope. Note, some loss is still needed for convergence and any real physical system has that.

These attacks are based on the fact that any wire has a geometrical capacitance and, to charge that, a current flow, energy flow and power flow are needed. That makes five different ways of efficient attack, see below. The measurement of the voltage and current and by their product gives the power flow and its direction (the quasi static analogy of the Poynting vector in electromagnetics), see Figure 5.

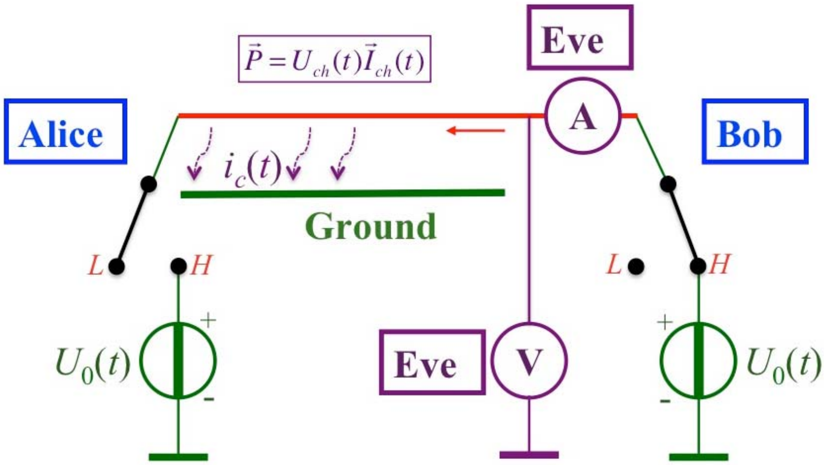

Figure 5. Universal, energy-flow-analysis-attack against the BR scheme (even in the no-wave transient limit provided by proper voltage time functions by eliminating high-frequency components). The direction of power flow, energy flow or that of the current fully characterizes the situation for Eve. Similarly the size of power flow, energy flow and current flow versus location also fully compromise the security.

The power vector is given as

$$\vec{P}(t) = U_c \vec{I}_c(t) \qquad , \tag{27}$$

and the energy vector is its integral over the BEP:

$$\vec{E} = \int_0^{\tau} \vec{P}(t) dt \quad . \tag{28}$$

The direction of $\langle \vec{P}(t) \rangle$ or that of $\vec{E}$ fully characterizes the situation for Eve. Similarly the size of the $\langle \vec{P}(x,t) \rangle$ $\vec{E}(x)$ and $\langle \vec{I}(x,t) \rangle$ vectors versus the location x also fully compromises the security. The further away from the connected voltage source these location-dependent quantities are evaluated the less are their values, which are zero at the other end.

2.5.2 Two transient-damping resistor attacks

To make the system physical and stop the transient after one return, Alice and Bob may use damping resistors to match the wave resistance of the cable, see Fig 6. This will cause a continuous noise current drive into the geometrical capacitance of the wire. We show two ways to utilize thermodynamics to crack this system in the steady state.

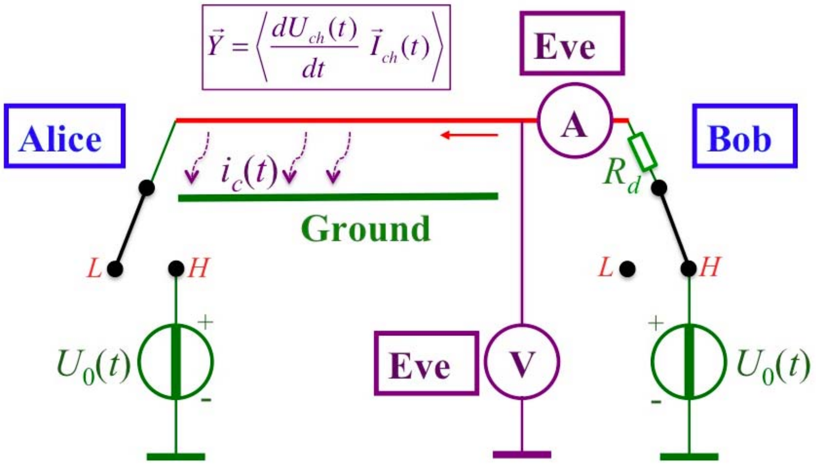
$\vec{Y}=\left\langle \frac{dU_{ch}(t)}{dt}\vec{I}_{ch}(t)\right\rangle$
Eve
Alice
A
Bob
R_d
$i_c(t)$
Ground
L
H
L
H
+
$U_0(t)$
-
Eve
V
+
$U_0(t)$
-

Figure 6. The transient-damping resistor version of the BR scheme and the capacitive noise current attack

This current is correlated with the time derivative of the channel voltage.

$$\vec{Y}=\left\langle \frac{dU_c(t)}{dt}\vec{I}_c(t)\right\rangle \qquad (29)$$

Both the sign of the cross-correlation $\vec{Y}$ and its value versus location fully informs Eve about the situation; its value is zero at the free end.

2.5.3 The wire resistance Johnson noise attack

Any wire will have non-zero resistance thus produce Johnson noise. Eve can simply measure the voltage noise between the wire and the ground at the two ends of the wire, see Figure 7.

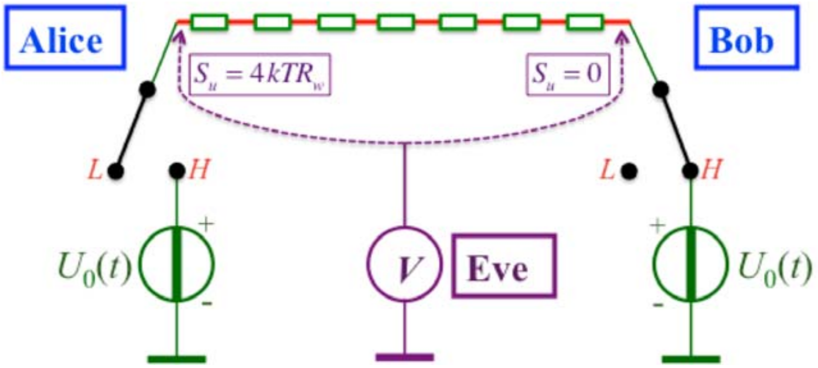

Figure 7. The BR scheme with finite wire resistance and the Johnson noise attack.

The free end will have

$$S_u = 4kTR_w \qquad (30)$$

voltage noise spectrum, while the connected end shows zero. Eve can fully crack the system. Note that this attack can be avoided if the connected end is using a large additive noise to bury the noise however then he former attacks will still crack the system with less effort.

2.5.4 The above attacks are inefficient against the KLJN system due to thermodynamics

It is easy to understand how does thermodynamics protects the KLJN system against the above attacks. Due to the resistors of Alice and Bob, the system is thermodynamic and produces Johnson noise. These noise voltages are much larger than the parasitic Johnson noise of the wire because the wire resistance must be small (max. 1-2% of $R_L + R_H$). Similarly the noise-bandwidth is chosen so that the capacitive currents are negligible compared to the channel current. The implication is that, when the above described attacks are used against the KLJN system, the measurement result of Eve will again be a small DC signal buried in a large noise, leading to similar relations as it was shown for the wire resistance voltage-drop, see Eqs. 12-15 with near-to-perfect information theoretic security.

2.6 On transient attacks against the KLJN scheme

This attack is different from other attacks in the literature and in this paper in the sense that, in the no-wave (quasi static) limit, where KLJN operates, no concrete realization has ever been proposed with a measurement and evaluation protocol. Therefore, at the moment, this attack exists only at the "philosophical level". However, KLJN researchers have realized from the beginning that transients pose vulnerabilities and, to reduce the potential information leak, various schemes have been proposed including ramping up/down of the noise, starting from zero noise amplitude (and velocity) and the adiabatic random walking of Alice's and Bob's resistance [31], which we are showing here for the sake of this debate. Before that, we note that, in our opinion, due to the *quasi-static* condition, which is manifested by the noise voltage envelope and filters, and due to the unknown resistances and noise at the two ends, the efficiency of any transient attacks is strongly limited because, as soon as the front of the noise propagation (not wave) reaches the other end, the flow of information about the particular noise is strongly reduced. Thus, even in the no transient protection limit, Eve has effectively only a very small sample of a noise with duration much shorter than its correlation time.

First, we describe the so far best-known transient protocol based on random walking resistances [31]. Alice and Bob randomly choose R_L or R_H for their R_A and R_B, respectively. Then, to execute the key exchange, they use continuously variable resistors (such as potentiometers, MOSFETs, etc.). If noise generators also used to enhance the noise-temperature than the band-limited white noise spectra of the noise-generators need to also be variable in a synchronized fashion so that the noise-temperature stays constant, at the publicly agreed value T_{eff}. Suppose also that the KLJN noise bandwidth is secured also by line filters at Alice's and Bob's ends. At the beginning of the KLJN-clock period, both Alice and Bob start with:

$$R_A(0) = R_B(0) = \frac{R_L + R_H}{2} \tag{31}$$

and they stay at this value until the noises equilibrates in the wire. Thus no informative transients can be observed right after connecting the resistors to the line because the bit values have not been realized yet. Then Alice and Bob execute independent, adiabatically slow continuum-time random walks with their resistor values (in a synchronized fashion with the spectral parameter of their noise generators). The random walks are executed so slowly that, from a thermodynamic point of view, the system is changing in the adiabatic limit: there is virtually thermal equilibrium in the line during the whole random-walk process.

There is a publicly pre-agreed time period t_r to execute these independent random walks. If within this time period Alice and Bob reach their randomly preselected R_A and R_B value, they stop the random walk and stay at this value. Then, after the t_r time period, they start the measurements, in the regular fashion. In this way, the transient effects and the information leak they may cause, are virtually kept away.

If, by the end of the t_r time period, either the random walk of Alice or Bob (or both) does not reach the randomly preselected resistance value, he/she (or both) submits a cancellation signal via an authenticated channel and the bit exchange process is immediately terminated; and a new independent KLJN-clock period starts in the way described above.

Concerning the security, the production of spurious frequency products will be proportional to the RMS speed v_{rms} of the random walk. Thus, it is reasonable to assume that, if a concrete attack type will be proposed, it will satisfy (see Eqs. 12,13):

$$q = \vartheta_{tr} v_{rms} \ , \tag{32}$$

where ϑ_{tr} is a constant relevant for the transient attack against this scheme. This is leading to unconditional ε - security $\left(\varepsilon \propto v_{rms}\right)$, with results of the same nature as in Eqs.-12-15 with statistical distance:

$$\Delta = \left(0.5 + q\right)^N - 0.5^N \cong 2Nq0.5^N = 2N\vartheta_{tr} v_{rms} 0.5^N \tag{33}$$

This value is reached without privacy amplification.

Here the resource used to approach the perfect security is the duration τ of the bit exchange period (BEP) because, when the random walk time dominates it, it is inversely proportional to v_{rms}. In other words, ay fixed key length, the "price" of increasing the security is reducing the speed of key exchange and $\varepsilon \propto 0.5^N \tau^{-1} N$ can again be arbitrarily small.

2.7 Why the BR passive correlation attack does not work against the KLJN scheme

The directional couplers have limited bandwidth, work in the wave limit and, with good directivity, they have $\lambda_0 / 4$ size, where $\lambda_0 = c / f_0$ and f_0 is the frequency of optimal working, see section 2.1.4. For much longer wavelengths (smaller frequencies, as in KLJN) the system executes a Rayleigh scattering and, accordingly, see Eqs. 23, 24, the passive correlation attack results in a correlation coefficient with f^4 power function scaling. This is leading to unconditional ε - security $\left(\varepsilon \propto f^4\right)$, with results of the same nature as in Eqs.-12-15 with statistical distance:

$$\Delta = \left(0.5 + q\right)^N - 0.5^N \cong 2Nq0.5^N = 2N\vartheta_{cr} B_{kljn}^4 0.5^N \tag{34}$$

where ϑ_{cr} is a constant defined similarly to Eq. 13. This value is reached without privacy amplification.

Here the resource used to approach the perfect security is the duration τ, $\left(\tau \propto 1 / B_{kljn}\right)$ of the bit exchange period (BEP) because that is inversely proportional to the highest frequency in the noise-bandwidth. In other words, at fixed key length, the "price" of increasing the security is reducing the speed of key exchange and $\varepsilon \propto 0.5^N \tau^{-4} N$ can again be arbitrarily small.

2.8 Why the current extraction/injection active attack does not work against KLJN

BR [4] proposes and active (invasive) attack against the BR system where Eve at the middle connects a grounded resistor to the line to extract some current and monitors the current direction in the wire.

...she (Eve) *could still learn the key by an active steady-state attack in which she would place a very high-resistance shunt be- tween her node and ground, and monitor the direction of current flow into it. Of course Alice and Bob could try to detect this weak leakage current also, and abort the protocol if they found it. The result would be an unstable arms race, won by whichever*

side had the more sensitive ammeter, not the sort of robustness reasonably expected of a practical cryptosystem.

We fully agree with the above assessment when it is discussing the BR system. However, this attack is inefficient against the KLJN system and this fact was pointed out already in the foundation paper [26] of the KLJN scheme. There [26], a technically more efficient attack of the same nature was proposed: Eve injects a stochastic current at the middle and monitors the cross-correlation of this current with the channel currents in the two directions. The correlation coefficient will be greater in the direction of the smaller resistance. This problem was later posed also by Reiner Plaga and Horace Yuen in private communications. Alice and Bob monitors the channel currents at the two ends and compare their instantaneous amplitudes via an authenticated public channel. If they differ, the bit exchange event is terminated and that bit is discarded.

The usual statement to justify this attack is that Eve may use miniscule current amplitudes, which are below the detection limit of the comparison by Alice and Bob. This argument does not work because Alice and Bob design their current resolution so that Eve, by using this attack, can't extract enough information. The channel current at Alice's side of Eve will be:

$$I_{cA}(t) = I_c(t) - \gamma I_E(t) \qquad (35)$$

and at Bob's side of Eve it is:

$$I_{cB}(t) = I_c(t) + (1-\gamma) I_E(t) \quad , \qquad (36)$$

where $I_E(t)$ is Eve's injected current and $(1-\gamma)/\gamma = R_A / R_B$. The cross-correlations with Eve's current during the BEP are:

$$\rho_A = \left\langle \left[I_c(t) - \gamma I_E(t) \right] I_E(t) \right\rangle_\tau = \left\langle I_c(t) I_E(t) \right\rangle_\tau - \gamma \left\langle I_E^2(t) \right\rangle_\tau = U_{cE\tau}(t) - \gamma \left\langle I_E^2(t) \right\rangle - \gamma U_{EE\tau}(t) \ , \qquad (37)$$

$$\begin{aligned} \rho_B &= \left\langle \left[I_c(t) + \left(1-\gamma\right) I_E(t) \right] I_E(t) \right\rangle_\tau = \left\langle I_c(t) I_E(t) \right\rangle_\tau + \left(1-\gamma\right) \left\langle I_E^2(t) \right\rangle_\tau = \\ &= U_{cE\tau}(t) + \left(1-\gamma\right) \left\langle I_E^2(t) \right\rangle + \left(1-\gamma\right) U_{EE\tau}(t) \ , \end{aligned} \qquad (38)$$

where $\langle \ \rangle_\tau$ stands for the finite time (τ) average, U for their noise components, and $\langle \ \rangle$ for the exact average (requiring infinite time). The dominant terms at the right-hand side of Eqs. 37,38 are the noise terms of the cross-correlations between Eve's current and the channel current, with mean-square amplitudes scaling with τ^{-1}. Because the RMS amplitude $I_{E,rms}$ of Eve's current is negligible compared to that of the channel current,

$$I_{E,rms} = \sigma I_{c,rms} \quad , \qquad (39)$$

where $\sigma << 1$, the last noise terms at the right-hand side of Eqs. 37,38 are negligible compared to the first noise terms. The detection problem is again the same as we have already seen at the wire resistance attack: a small DC component (the second term) in a large noise (the first term). Thus, q will again satisfy (see Eqs. 12,13):

$$q = \vartheta_{ci}\sigma \ , \tag{32}$$

where ϑ_{ci} is a constant relevant for this current injection/extraction attack at a fixed τ (note, ϑ_{ci} is inversely proportional to τ). This is leading to unconditional ε - security $\left(\varepsilon \propto 0.5^N \sigma\right)$, with results of the same nature as in Eqs.-12-15 with statistical distance:

$$\Delta = \left(0.5+q\right)^N - 0.5^N \cong 2Nq0.5^N = 2N\vartheta_{tr}\sigma 0.5^N \tag{33}$$

This value is reached without privacy amplification. At fixed key length, the resource utilized for approaching the perfect security is the resolution of Alice's and Bob's current comparison because σ must be chosen smaller than the relative current resolution to stay hidden during the current injection.

2.9 Remarks about potential hacking attacks

Mathematical models of physical systems and building elements are always approximate and security proofs can only be given for these model systems. Particularly dangerous are the elements that are directly exposed to Eve. Thus a commercial secure key exchanger must be carefully designed with considering all the foreseeable hacking attacks.

For example, a real KLJN system must be armed with extra circuitry and protocol steps against Makarov-style blinding attacks, out-of frequency range attacks, circulator-based attacks, etc.

Conclusions

We showed that thermodynamics, noise, and the Second Law of Thermodynamics (the impossibility to construct a perpetual motion machine of the second kind) are essential for the security of the KLJN classical physical key exchanger.

We gave mathematical security proofs for each BR attack and the results indicate that the security of the KLJN method has not successfully been challenged by BR.

Acknowledgements

LK is grateful for helpful discussions/consultations/critical remarks about the relevant security measures with Horace Yuen, Vadim Makarov, Renato Renner and Vincent Poor.

We are grateful to also to Charles Bennett and Jess Riedel for motivating our analysis by their paper.

References

1. Liang Y, Poor HV, Shamai S (2008) Information theoretic security. Foundations Trends Commun. Inform. Theory 5:355-580. DOI: 10.1561/0100000036.
2. Mingesz R, Kish LB, Gingl Z, Granqvist CG, Wen H, Peper F, Eubanks T, Schmera G (2013) Unconditional Security by the laws of classical physics. Metrology & Measurement Systems XX:3-16. DOI: 10.2478/mms-2013-0001. http://www.degruyter.com/view/j/mms.2013.20.issue-1/mms-2013-0001/mms-2013-0001.xml
3. Bennett CH, Brassard G (1984) Proc. International Conference on Computers, Signals, and Signal Processing, Bangalore, India. pp. 175-179.
4. Bennett CH, Riedel CJ (2013) On the security of key distribution based on Johnson-Nyquist noise. http://arxiv.org/abs/1303.7435
5. Yuen HP (2012) On the foundations of quantum key distribution — Reply to Renner and beyond. arXiv:1210.2804.
6. Hirota O (2012) Incompleteness and limit of quantum key distribution theory. arXiv:1208.2106v2.
7. Renner R (2012) Reply to recent scepticism about the foundations of quantum cryptography. arXiv:1209.2423v.1.
8. Yuen HP (2012) Security significance of the trace distance criterion in quantum key distribution. arXiv:1109.2675v3.
9. Yuen HP (2009) Key generation: Foundation and a new quantum approach. IEEE J. Selected Topics in Quantum Electronics 15, 1630.
10. Salih H, Li ZH, Al-Amri M, Zubairy MS (2013) Protocol for Direct Counterfactual Quantum Communication. Phys. Rev. Lett. 110:170502.
11. Merali Z (29 August 2009) Hackers blind quantum cryptographers. Nature News, DOI:10.1038/news.2010.436.
12. Gerhardt I, Liu Q, Lamas-Linares A, Skaar J, Kurtsiefer C, Makarov V (2011) Full-field implementation of a perfect eavesdropper on a quantum cryptography system. Nature Commun. 2; article number 349. DOI: 10.1038/ncomms1348.
13. Lydersen L, Wiechers C, Wittmann C, Elser D, Skaar J, Makarov V (2010) Hacking commercial quantum cryptography systems by tailored bright illumination. Nature Photonics 4:686-689. DOI: 10.1038/NPHOTON.2010.214.
14. Gerhardt I, Liu Q, Lamas-Linares A, Skaar J, Scarani V, Makarov V, Kurtsiefer C (2011) Experimentally faking the violation of Bell's inequalities. Phys. Rev. Lett. 107:170404. DOI: 10.1103/PhysRevLett.107.170404.
15. Makarov V, Skaar J (2008) Faked states attack using detector efficiency mismatch on SARG04, phase-time, DPSK, and Ekert protocols. Quantum Inf. Comp. 8:622-635.
16. Wiechers C, Lydersen L, Wittmann C, Elser D, Skaar J, Marquardt C, Makarov V, Leuchs G (2011) After-gate attack on a quantum cryptosystem. New J. Phys. 13:013043. DOI: 10.1088/1367-2630/13/1/013043.
17. Lydersen L, Wiechers C, Wittmann C, Elser D, Skaar J, Makarov V (2010) Thermal blinding of gated detectors in quantum cryptography. Opt. Express 18:27938-27954. DOI: 10.1364/OE.18.027938.
18. Jain N, Wittmann C, Lydersen L, Wiechers C, Elser D, Marquardt C, Makarov V, Leuchs G (2011) Device calibration impacts security of quantum key distribution. Phys. Rev. Lett. 107:110501. DOI: 10.1103/PhysRevLett.107.110501.
19. Lydersen L, Skaar J, Makarov V (2011) Tailored bright illumination attack on distributed-phase-reference protocols. J. Mod. Opt. 58:680-685. DOI: 10.1080/09500340.2011.565889.
20. Lydersen L, Akhlaghi MK, Majedi AH, Skaar J, Makarov V (2011) Controlling a superconducting nanowire single-photon detector using tailored bright illumination. New J. Phys. 13:113042. DOI: 10.1088/1367-

2630/13/11/113042.

21. Lydersen L, Makarov V, Skaar J (2011) Comment on “Resilience of gated avalanche photodiodes against bright illumination attacks in quantum cryptography”. Appl. Phys. Lett. 99:196101. DOI: 10.1063/1.3658806.
22. Sauge S, Lydersen L, Anisimov A, Skaar J, Makarov V (2011) Controlling an actively-quenched single photon detector with bright light. Opt. Express 19:23590-23600.
23. Lydersen L, Jain N, Wittmann C, Maroy O, Skaar J, Marquardt C, Makarov V, Leuchs G (2011) Superlinear threshold detectors in quantum cryptography. Phys. Rev. Lett. 84:032320. DOI: 10.1103/PhysRevA.84.032320.
24. Lydersen L, Wiechers C, Wittmann C, Elser D, Skaar J, Makarov V (2010) Avoiding the blinding attack in QKD; Reply (Comment). Nature Photonics 4:801-801. DOI: 10.1038/nphoton.2010.278.
25. Makarov V (2009) Controlling passively quenched single photon detectors by bright light. New J. Phys. 11:065003. DOI: 10.1088/1367-2630/11/6/065003.
26. Kish LB (2006) Totally secure classical communication utilizing Johnson(-like) noise and Kirchhoff’s law. Phys. Lett. A 352:178-182.
27. Cho A (2005) Simple noise may stymie spies without quantum weirdness. Science 309:2148; http://www.ece.tamu.edu/~noise/news_files/science_secure.pdf.
28. Kish LB (2006) Protection against the man-in-the-middle-attack for the Kirchhoff-loop-Johnson(-like)-noise cipher and expansion by voltage-based security. Fluct. Noise Lett. 6:L57-L63. http://arxiv.org/abs/physics/0512177.
29. Mingesz R, Gingl Z, Kish LB (2008) Johnson(-like)-noise-Kirchhoff-loop based secure classical communicator characteristics, for ranges of two to two thousand kilometers, via model-line, Phys. Lett. A 372:978-984.
30. Palmer DJ (2007) Noise encryption keeps spooks out of the loop. New Scientist, issue 2605 p.32; http://www.newscientist.com/article/mg19426055.300-noise-keeps-spooks-out-of-the-loop.html.
31. Kish LB (2013) Enhanced secure key exchange systems based on the Johnson-noise scheme. Metrology & Measurement Systems XX:191–204. open access: http://www.degruyter.com/view/j/mms.2013.20.issue-2/mms-2013-0017/mms-2013-0017.xml?format=INT
32. Kish LB, Horvath T (2009) Notes on recent approaches concerning the Kirchhoff-law-Johnson-noise-based secure key exchange. Phys. Lett. A 373:901-904.
33. Horvath T, Kish LB, Scheuer J (2011) Effective privacy amplification for secure classical communications. Europhys. Lett. 94:28002. http://arxiv.org/abs/1101.4264.
34. Saez Y, Kish LB (2013) Errors and Their Mitigation at the Kirchhoff-Law-Johnson-Noise Secure Key Exchange. http://vixra.org/abs/1305.0126. http://arxiv.org/abs/1305.4787.
35. Laszlo B. Kish, Chiman Kwan (2013) Physical uncloneable function hardware keys utilizing Kirchhoff-law-Johnson-noise secure key exchange and noise-based logic. http://vixra.org/abs/1305.0068; http://arxiv.org/abs/1305.3248
36. Kish LB, Saidi O (2008) Unconditionally secure computers, algorithms and hardware. Fluct. Noise Lett. 8:L95-L98.
37. Elias Gonzalez, Laszlo B. Kish, Robert Balog, Prasad Enjeti (2013) Information theoretically secure, enhanced Johnson noise based key distribution over the smart grid with switched filters. http://vixra.org/abs/1303.0094; http://arxiv.org/abs/1303.3262
38. Kish LB, Mingesz R (2006) Totally secure classical networks with multipoint telecloning (teleportation) of classical bits through loops with Johnson-like noise. Fluct. Noise Lett. 6:C9-C21.
39. Kish LB, Peper F (2012) Information networks secured by the laws of physics. IEICE Trans. Commun. E95-B:1501-1507.
40. Scheuer J, Yariv A (2006) A classical key-distribution system based on Johnson (like) noise – How secure? Phys. Lett. A 359:737-740.
41. Kish LB, Scheuer J (2010) Noise in the wire: The real impact of wire resistance for the Johnson(-like) noise based secure communicator. Phys. Lett. A 374:2140-2142.
42. Kish LB (2006) Response to Scheuer-Yariv: “A classical key-distribution system based on Johnson (like) noise – How secure?”. Phys. Lett. A 359:741-744.
43. Hao F (2006) Kish’s key exchange scheme is insecure. IEE Proc. Inform. Soc. 153:141-142.
44. Kish LB (2006) Response to Feng Hao’s paper “Kish’s key exchange scheme is insecure”. Fluct. Noise Lett. 6:C37-C41.
45. Liu PL (2009) A new look at the classical key exchange system based on amplified Johnson noise. Phys. Lett. A 373:901-904.

46. Arora S, Barak B (2009) Computational Complexity. Cambridge University Press. Cambridge.
47. Antilli D (2005) System and method for the propagation of deterministic and non-deterministic values by means of electrical conductors. Patent publication number: EP1952573 (A2). http://www.google.com/patents/EP1952573A2?cl=en
48. Pauli W (2000) Electrodynamics. Dover Publications, New York.
49. Matthaei GL, Young L, Jones EMT (1964) Microwave Filters, Impedance-Matching Networks, and Coupling Structures. McGraw-Hill, New York.
50. Liu PL (2009) A Key Agreement Protocol Using Band-Limited Random Signals and Feedback. IEEE J. Lightwave Tech. 27:5230-5234.
51. Liu PL (2009) Security risk during the transient in a key exchange protocol using random signals and feedback. Phys. Lett. A 373:3207–3211.
52. Kish LL, Zhang B, Kish LL (2010) Cracking the Liu key exchange protocol in its most secure state with Lorentzian spectra. Fluct. Noise Lett. 9:37-45