Long Distance Continuous-Variable Quantum Key Distribution with a Gaussian Modulation

Paul Jouguet,[1,2] Sébastien Kunz-Jacques,[3] and Anthony Leverrier[3]

[1] *Institut Telecom / Telecom ParisTech, CNRS LTCI, 46, rue Barrault, 75634 Paris Cedex 13, France*

[2] *SeQureNet, 23 avenue d'Italie, 75013 Paris, France*

[3] *ICFO-Institut de Ciencies Fotoniques, 08860 Castelldefels (Barcelona), Spain*

(Dated: December 12, 2011)

We designed high-efficiency error correcting codes allowing to extract an errorless secret key in a Continuous-Variable Quantum Key Distribution (CVQKD) protocol using a Gaussian modulation of coherent states and a homodyne detection. These codes are available for a wide range of signal-to-noise ratios on an Additive White Gaussian Noise Channel (AWGNC) with a binary modulation and can be combined with a multidimensional reconciliation method proven secure against arbitrary collective attacks. This improved reconciliation procedure considerably extends the secure range of CVQKD with a Gaussian modulation, giving a secret key rate of about 10^{-3} bit per pulse at a distance of 120 km for reasonable physical parameters.

I. INTRODUCTION

Quantum Key Distribution (QKD) [1] is the first real-life application of quantum information. It allows two distant parties, Alice and Bob, to establish an *unconditional* [2] secret key through the exchange of quantum states even in the presence of an eavesdropper, with the help of a classical auxiliary authenticated communication channel [3]. The first QKD vendors, ID Quantique [4] and MagiQ Technologies [5], developed systems based on encoding the information on discrete variables such as the phase or the polarization of single photons. The limitation of these technologies is mainly due to the speed and the efficiency of the single photon detectors. Recently created companies, Quintessence Labs [6] and SeQureNet [7], are developing a new generation of systems that allow to get rid of these limitations by encoding the information on continuous variables (CV) such as the quadratures of coherent states.

In the standard protocol [8], one needs to prepare Gaussian modulated coherent states and to measure them with a homodyne or a heterodyne [9] detection which requires only standard telecommunication parts. With the current proof techniques, using a Gaussian modulation is optimal as regards the theoretical secret key rate. In particular, security against collective attacks is well understood [10, 11], even in the finite-size regime [12], and collective attacks are known to be asymptotically optimal [13]. However, since the efficiency of the current reconciliation protocols for Gaussian variables drops dramatically in the regime of low signal-to-noise ratios (SNRs), new protocols using specific non-Gaussian modulations, either discrete [14] or continuous [15], have been developed. The idea of these modulations is that they are compatible with high-performance error correction, making possible for the protocol parties to extract efficiently the information available in their raw data. This is in strong contrast with the Gaussian modulation for which no efficient reconciliation procedure was available until now. In theory, protocols with a non-Gaussian modulation therefore increase the achievable secure distance of CVQKD. They have, however, not yet been demonstrated experimentally. Indeed, for long distances, that is low transmission of the quantum channel, the optimal modulation variance is typically lower for non-Gaussian modulations (in particular for the four-state protocol [14]) than for a Gaussian modulation. This makes the design a stable continuous-variable system able to operate at large distances difficult. Even if this effect is mitigated for the eight-dimension protocol [15], the modulation allowing for the largest variance remains the Gaussian one [8].

In this paper, we exhibit high-efficiency error correcting codes which can be combined with a multidimensional reconciliation scheme [16]. This allows, for the first time, to distill a secret key from a CVQKD protocol with a Gaussian modulation in the regime of very low SNR, and paves the way for future experimental demonstrations of CVQKD over much larger distances than the current record of 30 km [17, 18].

In Section II, we explain how the problem of reconciling Gaussian variables can be translated into a channel coding problem on the Binary Input Additive White Gaussian Noise Channel (BIAWGNC), for which we describe very low rate error correcting codes in Section III. Combining these tools, we are able to efficiently reconcile data at low SNRs. Finally, we show in Section IV the consequences of these new developments on the performance of the Gaussian protocol over long distances.

II. THEORY OF RECONCILIATION OF GAUSSIAN VARIABLES

The data reconciliation step is critical in CVQKD: the distance of the chosen error-correction scheme to the

Shannon bound affects both the key rate and the range of the protocol. Of considerable importance is the problem of the reconciliation of correlated Gaussian variables. This is indeed the scenario considered in the GG02 protocol [8] where Alice's coherent states are modulated with a bivariate Gaussian distribution in phase space. Different approaches have been explored to increase the reconciliation efficiency for a Gaussian modulation, especially in the regime of low SNR.

A first approach called *Slice Reconciliation* was proposed in [19, 20] and implemented in [17, 18] but the efficiency of this method currently limits the protocol range to about 30 km. Another method is to encode the information on the sign of the Gaussian modulated value. However, since we deal with centered Gaussian variables, the uncertainty on the sign increases at low SNRs because most values have small amplitude. Another class of protocols use *post-selection* [21–24] by working only with high-amplitude data but the security is not proven against general collective attacks.

In [16], the idea of reducing the Gaussian variables reconciliation problem to the channel coding problem is introduced. One first uses a d-dimensional rotation to build a virtual channel close to the BIAWGNC from the physical Gaussian channel. This means that d consecutive instances of the physical channel are mapped to d approximate copies of a virtual BIAWGN channel, which are used to perform the error correction and eventually distill the actual secret. The final reconciliation efficiency one obtains with such a scheme depends on two things:

- The intrinsic efficiency of the error correcting code used on the virtual channel *on the BIAWGNC* (such an efficiency is given for example in Table I).
- The quality of the approximation between the virtual channel and the BIAWGNC (for the scheme given in [16], the quality of this approximation increases with the dimension d).

One can therefore improve the reconciliation efficiency of the global scheme by working on two things: designing codes with higher efficiencies on the BIAWGNC and increasing the dimension of the scheme.

Let us now explain in more details the setting defined in [16]: Alice, the sender, and Bob, the receiver, are given two n-dimensional real vectors $\mathbf{x}$ and $\mathbf{y}$ and can use a public authenticated channel to agree on a common bit string $\mathbf{u}$. For this, one of the parties (say Alice in the direct reconciliation scheme) sends to the other additional information describing a function f such that $f(\mathbf{x}) = \mathbf{u}$; the other party (Bob) applies this function to his data to get $\mathbf{v} := f(\mathbf{y})$; this way, a virtual communication channel with input $\mathbf{u}$ and output $\mathbf{v}$ is defined. The explicit construction of [16] aims at creating a virtual channel that is close to the BIAWGNC since very efficient codes are available for that channel.

Alice and Bob are given two d-uplets $\mathbf{x}$ and $\mathbf{y}$ corresponding to correlated Gaussian vectors (this is valid for CVQKD with a Gaussian modulation and a Gaussian optimal attack). This means that one can introduce constants t, t' and σ, σ' such that one has $\mathbf{y} = t\mathbf{x} + \mathbf{z}$ with $\mathbf{x} \sim \mathcal{N}(0,1)^d$, $\mathbf{z} \sim \mathcal{N}(0,\sigma^2)^d$ in the direct reconciliation case and $\mathbf{x} = t'\mathbf{y} + \mathbf{z}'$ with $\mathbf{y} \sim \mathcal{N}(0, 1+\sigma^2)^d$, $\mathbf{z}'^d \sim \mathcal{N}(0,\sigma'^2)^d$ in the reverse reconciliation case. Since the two scenarios are similar, we consider without loss of generality only the direct reconciliation one here. Furthermore, up to a simple renormalization, one can fix $t = 1$.

Alice chooses a random element $\mathbf{u} \in \{-1/\sqrt{d}, 1/\sqrt{d}\}^d$ with the uniform distribution on the d-dimensional hypercube and sends $\mathbf{r} = \mathbf{u}.\mathbf{x}^{-1}$ to Bob through the public channel (a "multiplication" and its inverse "division" operator are assumed to exist on d-dimensional vectors - more on this below). Then Bob computes $\mathbf{v} := \mathbf{r}.\mathbf{y}$. Let us analyse the noise $\mathbf{w}$ on this virtual channel:

$$\begin{aligned}\mathbf{w} &:= \mathbf{v} - \mathbf{u} \\ &= \mathbf{r}\mathbf{y} - \mathbf{u} \\ &= \mathbf{u}.\mathbf{x}^{-1}(\mathbf{x} + \mathbf{z}) - \mathbf{u} \\ &= \mathbf{u}\frac{\mathbf{z}}{\mathbf{x}} \sim \mathbf{u}\frac{\mathbf{z}}{||\mathbf{x}||}\end{aligned}$$

where the last equality holds in law and is due to the spherical symmetry of the distributions of $\mathbf{z}$ and $\mathbf{x}$ and their independence. Since the norm of $\mathbf{x}$ is transmitted, the channel considered is a Fading Channel with Known Side Information as defined in [25], the fading coefficient being the norm of $\mathbf{x}$, which follows a $\chi(d)$ distribution with d degrees of freedom. Since the distribution $\chi(d)$ gets closer to a Dirac distribution when d goes to infinity, one should use the highest dimension possible in order to obtain the degenerate version of the Fading Channel with Known Side Information where all the fading coefficients are equal to 1, that is, the BIAWGNC. Unfortunately, the required division operator only exists in dimensions 1, 2, 4 and 8 (where it can be built from the algebraic structure of $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ and $\mathbb{O}$ respectively), so that it is not possible to use the above algorithm in arbitrary dimension.

III. RECONCILIATION OF GAUSSIAN VARIABLES: IMPLEMENTATION WITH LDPC CODES

Low Density Parity Check (LDPC) codes (or Gallager codes) are linear error-correcting codes with a sparse parity check matrix. A good reference about general coding theory and LDPC codes is [25]. LDPC codes can be represented as bipartite graphs, one set of the nodes being the check nodes representing the set of parity-check equations which define the code; the other, the variable nodes which represent the elements of the codewords. Variables nodes and check nodes are connected through

edges. LDPC codes are commonly used in telecommunications since they perform very close to Shannon limit and can be decoded with a fast iterative message-passing decoder called Belief Propagation (BP) (in such a decoding scheme, information is propagated between variable and check nodes that are connected by edges). These codes are designed for a given channel and a given SNR. The rate of a code is defined as the ratio between the information bits and the total number of transmitted bits on the channel. A low rate code is therefore a code with a lot of redundancy bits. Correcting errors at very low SNRs implies to design codes with low rates since adding redundancy allows to correct more errors.

A standard way to characterize LDPC codes is the probabilistic method: an ensemble of LDPC codes $\mathcal{C}$ is characterized by the node degrees and one proves that good codes occur with high probability within this ensemble. A specific code is simply drawn randomly from this set. Then one can modify the node degrees and their probabilities of occurence to improve the performance of the codes of the ensemble. A well known method to optimize LDPC codes for a given rate and a given channel is to use a genetic algorithm called *Differential Evolution*. This method has been successfully applied for a wide range of channels: the Binary Erasure Channel (BEC) [26], the BIAWGNC [27] and the Binary Symmetric Channel (BSC) [28]. The cost function that is maximized using this algorithm is defined as the threshold value for the channel (*i.e.* the maximal value of the noise that can be corrected with a given code, e.g. the standard deviation σ of the noise for the BIAWGNC or the probability of error ϵ for the BSC) and *Discretized Density Evolution* is used to compute the threshold.

In CVQKD, we need low-rate and high-efficiency codes for the BIAWGNC since errors must be efficiently corrected at very low SNRs to increase the secure distance. *Multi-edge-type LDPC codes* [29] give simple structures allowing to operate very close to Shannon limit at very low SNRs (for another construction of low rate LDPC codes refer to [30]). In the multi-edge setting, several edge classes are defined on the bipartite graph; then every node is defined by its number of sockets in each class. Whereas for standard LDPC ensembles the graph connectivity is constrained only by the node degrees, the multi-edge-type setting allows a greater control over the graph because only sockets of the same class can be connected together. Unlike standard LDPC ensembles, this framework provides for example the possibility to use degree-1 edges which improves significantly the threshold.

Every known reconciliation technique for CVQKD with a Gaussian modulation achieves an efficiency less than or equal to 90% [17, 20, 31]. This efficiency parameter β (defined by $\beta(s) = R/C(s)$ for a SNR s where R is the code rate used for the reconciliation and C is the capacity of the Additive White Gaussian Noise Channel (AWGNC)) is critical since the asymptotic secure key rate in the reverse reconciliation scheme is given by $K = \beta I(x;y) - \chi(y;E)$, where both $I(x;y)$ (the mutual information between the two protagonists bit strings x and y) and $\chi(y;E)$ (the Holevo information between the eavesdropper and the receiver's data) are large compared to K. One should especially pay attention to the dependency of β on the SNR. In [17, 20, 31], the good efficiency values are obtained only for SNRs higher than 1 which is incompatible with long distances. In [16], a 90% efficiency is obtained for a 0.5 SNR which allows to extend the secure distance from 30 km to 50 km. In this paper, we obtain higher efficiencies for even lower SNRs which allows secure key distribution over longer distances.

R	s_{DE}	C_{th}	β_{DE}
0.1	0.156	0.10429	95.9%
0.05	0.074	0.05144	97.2%
0.02	0.029	0.02038	98.1%

TABLE I. SNR asymptotic thresholds (s_{DE}) on the BIAWGNC, corresponding channel capacities (C_{th}) and efficiencies (β_{DE}) given by Density Evolution for low rate multi-edge LDPC codes of rate R.

Let us now review low rate LDPC codes with a good efficiency available in literature. In [29], table IX, a 95.9% efficiency, rate 1/10 code for the BIAWGNC is described. This efficiency can be further improved through an optimization of the distribution coefficients as mentioned in [29]. Starting from the structure of this code we designed codes with lower rates and with higher asymptotic thresholds. Table I sums up the performances of this original code together with our set of new multi-edge LDPC codes (the actual structure of the rate 0.02 code is described as an example in Appendix A). In this table, R is the rate of the considered code, s_{DE} is the SNR threshold given by Discretized Density Evolution, C_{th} is the theoretical channel capacity for this level of noise and β_{DE} is the efficiency of the code. These results are valid in the asymptotic regime, *i.e.* for codes of infinite length. However, the efficiency that is obtained with codewords of length 2^{20} is within 1% of the asymptotic efficiency.

A. Simulation Results with Rotations on S^1, S^3 and S^7

Let us discuss the simulation results we obtained applying the multidimensional reconciliation scheme with the previous codes for a dimension $d = 2$, $d = 4$ and $d = 8$, for the sign coding technique ($d = 1$) and without using any additional information, *i.e.* when we try to use a code designed for the BIAWGNC with a Gaussian modulation.

Tables II and III summarize the efficiencies we obtained with respect to the Gaussian channel capacity with our multi-edge LDPC codes for a block size of 2^{20}. We obtained a quite high Frame Error Rate (FER) (about 1/3) but a null Bit Error Rate (BER) on the

R	s	$s_{d=1}$	$s_{d=2}$	$s_{d=4}$	$s_{d=8}$
0.1	0.271	0.187	0.169	0.163	0.161
0.05	0.123	0.082	0.077	0.076	0.075
0.02	0.047	0.030	0.029	0.029	0.029

TABLE II. SNR thresholds on the BIAWGNC for low rate multi-edge LDPC codes (size 2^{20}) using the multidimensional reconciliation scheme ($d = 1, 2, 4, 8$).

R	β	$\beta_{d=1}$	$\beta_{d=2}$	$\beta_{d=4}$	$\beta_{d=8}$
0.1	57.9%	80.8%	88.7%	92.1%	93.1%
0.05	59.7%	88.3%	93.5%	94.8%	95.8%
0.02	60.0%	93.1%	96.3%	96.6%	96.9%

TABLE III. Efficiencies (w.r.t. the BIAWGNC capacity) for low rate multi-edge LDPC codes (size 2^{20}) using the multidimensional reconciliation scheme ($d = 1, 2, 4, 8$)

blocks where the decoding succeeded. This means that concatenating our codes with very high rate codes like BCH codes to remove the residual errors (as was done in [17, 18]) is not necessary here.

Since the channel obtained with rotations is not exactly a BIAWGNC, the efficiencies β are always lower than the efficiencies predicted by density evolution on the BIAWGNC. However, increasing the dimension d of the rotations allows to get closer to the efficiency of the code on the BIAWGNC. This is expected since the norm of the input vector $u^d||x^d||$ of the virtual channel follows a distribution $\chi(d)$ (where d is the number of degrees of freedom), which gets closer to a Dirac when d tends to infinity.

Figure 1 compares the capacities of the BIAWGNC and the multidimensional virtual channels for $d = 1, 2, 4, 8$ as a function of the SNR.

B. Use of rotations in higher dimension spaces

As was explained in the previous section, the multidimensional reconciliation scheme is limited to dimensions $1, 2, 4$ and 8 because these are the only ones compatible with a division structure [16].

In [16], the following construction applicable to arbitrary dimension d is proposed. In the direct case, with the same notations as in paragraph II (where Alice has a vector $\mathbf{x}$, Bob a vector $\mathbf{y}$, and Alice uses $(\mathbf{x}, \mathbf{r})$ to 'virtually' send $\mathbf{u}$ to Bob), a random orthogonal transformation Q on $\mathbb{R}^d$ is drawn according to the Haar mesure, then Q is composed with the reflection S across the mediator hyperplane of $\mathbf{x}' = Q(\mathbf{x})$ and $\mathbf{u}$. The resulting matrix $R = S \circ Q$ sends $\mathbf{x}$ to $\mathbf{u}$ and $\mathbf{y}$ to a point close to $\mathbf{u}$, because R preserves the euclidean distance; R is revealed by Alice and plays the same role as the vector $\mathbf{r}$ in section II. The randomization provided by Q ensures that R does not reveal more information on $(\mathbf{x}, \mathbf{u})$ than the relation $R(\mathbf{x}) = \mathbf{u}$; in particular, all $\mathbf{u}$ are equally likely given R.

FIG. 1. (Color online) Ratios between the capacities of the multidimensional channels ($d = 1, 2, 4, 8$) and the BIAWGNC and between the BIAWGNC and the AWGNC with respect to the SNR

Q is built, for instance, as the orthogonal ('Q') part of the QR decomposition of a $d \times d$ matrix G of Gaussian normalized random values. This method has complexity $\mathcal{O}(d^3)$. All other known methods to draw random orthogonal matrices have the same complexity.

We propose a method that allows to reduce the complexity to $\mathcal{O}(d^2)$. Let us observe first that we have the choice of the encoding of R: we do not need to reveal it in matrix form. However, the encoding must not reveal anything about $\mathbf{u}$ except that R satisfies $R(\mathbf{x}) = \mathbf{u}$. For instance, with the first method, revealing separately Q and S instead of $R = S \circ Q$ is not a good idea since S leaks information about u: indeed, in high dimension d, two random independent vectors are approximately orthogonal and therefore their mediator hyperplane forms and angle of about $\pi/4$ with either vector.

Let us examine first how an orthogonal transform Q can be drawn according to the Haar measure with complexity $\mathcal{O}(d^2)$, using an adequate representation, the Householder decomposition. An orthogonal basis $\mathbf{e}_1, \ldots, \mathbf{e}_d$ is fixed. Let E (resp. F) be the span of $\mathbf{e}_1, \ldots, \mathbf{e}_d$ (resp. $\mathbf{e}_2, \ldots, \mathbf{e}_d$).

If $d = 1$, choose $+1$ or -1. If $d > 1$, choose a random vector $\mathbf{g}$ uniformly on $\mathcal{S}^{d-1}$, the unit sphere in $\mathbb{R}^d$ (it can be constructed as $\mathbf{g} = \mathbf{h}/||\mathbf{h}||$ where $\mathbf{h}$ has independent normalized Gaussian coordinates), and draw recursively a random orthogonal matrix Q' of dimension $d-1$, viewed as a transform of F. Q' is extended to E by setting $Q'(\mathbf{e}_1) = \mathbf{e}_1$. Let S be the reflection that sends $\mathbf{e}_1$ on $\mathbf{g}$, and define $Q = S \circ Q'$. Q' is itself a composition of $d-1$ reflections in spaces of dimensions $d-1, \ldots, 1$. Describing each reflection by its corresponding eigenvector for the eigenvalue -1, Q is described by d vectors of dimensions $d, d-1, \ldots, 1$, for a total of $\frac{d(d+1)}{2}$ coefficients. The decomposition is unique. Note that $Q(\mathbf{e}_1) = \mathbf{g}$.

This process can be adapted when a constraint $Q(\mathbf{x}) = \mathbf{u}$ is added, with $||\mathbf{x}|| = ||\mathbf{u}||$. If $d = 1$, choose $+1$ or -1 depending on $\mathbf{x} = \mathbf{u}$ or $\mathbf{x} = -\mathbf{u}$. Assuming $d > 1$, $\mathbf{g}$ is chosen uniformly at random among unit vectors s.t.

$$\mathbf{u} \cdot \mathbf{g} = \mathbf{x} \cdot \mathbf{e}_1 \tag{1}$$

where $\cdot$ is the dot product. This relation is required for Q to satisfy both $Q(\mathbf{x}) = \mathbf{u}$ and $Q(\mathbf{e}_1) = \mathbf{g}$. Starting from a Gaussian normalized vector $\mathbf{h}$, α is chosen uniformly so that $(\mathbf{h} + \alpha\mathbf{u}) \cdot \mathbf{u} = (\mathbf{x} \cdot \mathbf{e}_1) \times ||\mathbf{h} + \alpha\mathbf{u}||$ (this is a quadratic equation that has at least one solution except if

d	s	β
2	1.644	76.3%
4	1.336	85.7%
8	1.194	91.7%
16	1.144	94.3%
32	1.108	96.2%
64	1.097	96.9%

TABLE IV. SNR thresholds and channel efficiencies on the BIAWGNC for the rate 1/2 multi-edge LDPC code in Table VI of ref. [29] with respect to the dimension of the multidimensional reconciliation scheme

$\mathbf{h}, \mathbf{u}$ span the same line, and $\mathbf{e}_1, \mathbf{x}$ do not, which happens with probability 0). $\mathbf{g} = \frac{\mathbf{h}+\alpha\mathbf{u}}{||\mathbf{h}+\alpha\mathbf{u}||}$ is computed in linear time and satisfies (1).

For an arbitrary vector $\mathbf{v}$, write its decomposition on F, e_1 as $\mathbf{v} = \mathbf{v}_F + \mathbf{v}_{F^\perp}$. Q' is drawn recursively, satisfying $Q'(\mathbf{x}_F) = S(\mathbf{u})_F$. This is possible because $\mathbf{x} \cdot \mathbf{e}_1 = \mathbf{u} \cdot \mathbf{g} = \mathbf{u} \cdot S(\mathbf{e}_1) = S(\mathbf{u}) \cdot \mathbf{e}_1$ implies $\mathbf{x}_{F^\perp} = S(\mathbf{u})_{F^\perp}$ and $||\mathbf{x}_F|| = ||S(\mathbf{u})_F||$. Then as $Q'(e_1) = e_1$, $Q'(\mathbf{x}) = S(\mathbf{u})$.

Define $Q = S \circ Q'$ as before: $Q(\mathbf{x}) = S(Q'(\mathbf{x})) = \mathbf{u}$.

The algorithm still runs in $\mathcal{O}(d^2)$, and the decomposition does not reveal any side information because it is unique. Since the added constraint (1) is required for the relation $Q(\mathbf{x}) = \mathbf{u}$ to hold, one sees recursively that the process yields the correct distribution on O_d. Finally, given the d reflection vectors, computing $Q(\mathbf{z})$ for any $\mathbf{z}$ is also done in time $\mathcal{O}(d^2)$. Hence by revealing these vectors instead of Q in matrix form, one gets the desired $\mathcal{O}(d^2)$ algorithm.

Let us now consider the rate 1/2 multi-edge LDPC code given in Table VI of reference [29]. The SNR threshold given by Discretized Density Evolution is $s^* = 1.074$. The corresponding efficiency on the BIAWGNC is 98.2%. When using a Gaussian modulation, table IV shows the effect of the dimension d on the efficiency β of the reconciliation scheme. We can see that increasing the dimension above 8 when operating at a high SNR enables to increase significantly the efficiency, and therefore the key rate in QKD applications.

C. Dealing with a continuous range of SNR with puncturing, shortening and repetition

We designed good efficiency codes for a finite set of rates so far; we are going to show how to deal with a continuous range of SNRs with this finite set. Since we designed low rate codes with good efficiencies, we can apply the simple technique of repetition codes mentioned in [32]. It is shown that starting from a code of rate R achieving an efficiency $\beta(s)$ for a SNR s on the BIAWGNC, one can use a repetition scheme of length k to build a new code of rate $R' = R/k$ achieving an efficiency β' for a SNR $s' = s/k$ given by

$$\beta'(s/k) = \beta(s) \, \frac{\log_2(1+s)}{k \log_2(1+s/k)}$$

For example, using a repetition scheme of length 3 with our code of rate 0.02 and efficiency 98% for a SNR of 0.03, we can build a code of efficiency $\beta(0.01) = 0.98\frac{\log_2(1.03)}{3\log_2(1.01)} = 97\%$. We applied this technique with repetition factors of 2 and 4 with our code of rate 0.02 to obtain the codes of rates 0.01 and 0.005 given in Table V.

However, this technique allows a low efficiency loss only for very small SNRs. For higher SNRs, other techniques must be applied if we want to keep very good efficiencies. Puncturing and shortening for LDPC codes are a good way to adapt the rate of a code [33]. Let us start with a (n, k) code, *i.e.* a code of length n with $n - k$ bits of redundancy; the rate is $R = k/n$. Puncturing consists in deleting a predefined set of p symbols from each word, converting a (n, k) code into a $(n-p, k)$ code. Shortening means deleting a set of s symbols from the encoding process (or revealing s message bits in addition to the syndrome in each codeword), converting a (n, k) code into a $(n - s, k - s)$ code. With a combination of these techniques the rate obtained is

$$R = \frac{k-s}{n-p-s} \quad .$$

The loss of efficiency incurred is small for small relative variations of the code rate. Typically, one can achieve a decrease of 5% (though shortening) and an increase of 10% (through puncturing) of the code rate with an efficiency loss smaller than 1%.

IV. PRACTICAL USE FOR A CONTINUOUS-VARIABLE QUANTUM KEY DISTRIBUTION SYSTEM

In this section, we apply the techniques developed in the previous sections to CVQKD in order to increase the secure distance achievable. We have to take into account that our quantum channel is Gaussian so that code efficiencies must be computed w.r.t. this channel capacity:

$$\beta = \frac{R}{C_{AWGNC}}$$

where R is the rate of the code and C_{AWGNC} is the capacity of the AWGNC. As we can see on Figure 1, the capacity of the BIAWGNC is very close to the capacity of the AWGNC for small values of the SNR. We give the efficiencies we can achieve on the AWGNC for different SNRs in Table V.

Our set of codes allows us to correct errors with an efficiency of about 95% for some fixed low SNRs. Let us plot the secret key rate as a function of the SNR on Bob

R	β	s
0.5	93.6%	1.097
0.1	93.1%	0.161
0.05	95.8%	0.075
0.02	96.9%	0.029
0.01	96.6%	0.0145
0.005	95.9%	0.00725

TABLE V. SNR thresholds and channel efficiencies on the AWGNC of the multi-edge LDPC codes mentionned in this paper.

FIG. 2. (Color online) Optimal modulation variance with respect to the distance: $\eta = 0.6$, $V_{elec} = 0.01$, $\xi = 0.01$, $\alpha = 0.2dB/km$, $\beta = 95\%$ and $\beta = 90\%$ from top to bottom.

side for a given distance and assuming a fixed error correction efficiency β. This enables to determine for which particular SNR it is relevant to design error-correcting code in order to maximize the secret key rate. We do not consider in this paper finite-size effects [12], meaning that our figures represent the key rate in the regime of infinite block length. In order to take finite-size effects into account two approaches are possible: a theoretical one consists in improving the proofs and the bounds on the secret key rate [34], a more practical one consists in designing systems with sufficient hardware stability in order to compute keys on large blocks.

The modulation variance is restricted within the interval $[1, 100]$ (in shot noise units) since lower values make the experimental setup much more complex. Indeed, a very low modulation variance is not compatible with brighter synchronization and phase tracking signals, because of the limited extinction ratios of the optical modulators (30dB for the most common models). An attenuation of 0.2dB/km is assumed. The homodyne detection efficiency is set to 0.6, and a value of 1% of the shot noise is taken for the electronic noise of the homodyne detection [17, 18]. A conservative value of 4% of the shot noise as in the European project SECOQC (Secure Communication based on Quantum Cryptography) [18] is used for the excess noise in Figure 4 while a more optimistic figure of 1% is used in Figures 2, 3 and 5. This second value is also typical of a realistic CVQKD system [17].

Figure 2 shows the optimal variance modulation on Alice side with respect to the key rate as a function of the distance. Achieving a good reconciliation efficiency at any SNR allows to work with a high modulation variance. This compares favorably to previous schemes with a discrete modulation which require modulation variances 10 times lower than the ones shown here.

FIG. 3. (Color online) Secret key rate for collective attacks with respect to the SNR: $\eta = 0.6$, $V_{elec} = 0.01$, $\xi = 0.01$, $\alpha = 0.2dB/km$, $V_A \in \{1, 100\}$, $\beta = 95\%$ and $\beta = 90\%$, $D = 10, 20, 50, 100$ km.

FIG. 4. (Color online) Secret key rate for collective attacks with respect to the SNR: $\eta = 0.6$, $V_{elec} = 0.01$, $\xi = 0.04$, $\alpha = 0.2dB/km$, $V_A \in \{1, 100\}$, $\beta = 95\%$ and $\beta = 90\%$, $D = 10, 20, 50, 100$ km.

FIG. 5. (Color online) Secret key rate for collective attacks with respect to the distance: $\eta = 0.6$, $V_{elec} = 0.01$, $\xi = 0.01$, $\alpha = 0.2dB/km$, $V_A \in \{1, 100\}$, $SNR = 1.097, 0.161, 0.075, 0.029, 0.0145, 0.00725$, $\beta = 93.6\%, 93.1\%, 95.8\%, 96.9\%, 96.6\%, 95.9\%$ from left to right.

Figure 3 and 4, plotted respectively for an excess noise of 1% and 4% of the shot noise, show that an improvement on the reconciliation efficiency yields at any distance a wider range of SNR with a close-to-optimum secret key rate. Conversely, the range of distances where a given error-correcting code working close to its threshold SNR can be used to get an almost optimal key rate is increased.

Given these large distance ranges where an error-correcting code is usable, it becomes feasible to use a small family of error-correcting codes to perform the reconciliation step at *any* distance and *without* using rate-tuning techniques such as puncturing or shortening. Figure 5 shows the key rate and the maximum secure distance obtained with this simple approach and the codes of Table V. With an excess noise of 1% of the shot noise, a secure distance above 150 km is obtained (with an excess of noise of 4% and the same codes, the secure distance is above 140 km). This is a significant improvement over previous reconciliation techniques since a reconciliation efficiency of 90% for a SNR of 0.5 only allows a secure distance of about 50 km with a Gaussian modulation [16].

V. CONCLUSION

We designed high-efficiency error-correcting codes allowing to distribute secret keys with a continuous-variable quantum key distribution system using a Gaussian modulation over long distances. Our results give a secure distance above 150 km against collective attacks (in the asymptotic regime) and can be implemented with only software modifications in the experimental setups of [17] and [18].

ACKNOWLEDGMENTS

This research was supported by the ANR, through the FREQUENCY and HIPERCOM projects, and by the

European Union, through the FP7 project Q-CERT and ERC Starting Grant PERCENT.

Appendix A: A rate 1/50 multi-edge LDPC code ($\sigma^* = 5.91$ on the BIAWGNC)

Below is the description of a multi-edge LDPC ensemble of codes of rate $R = 0.02$. The left half of the array describes the multidegree distributions of variable nodes, and the right half the distribution of check node multidegrees. m stands for a multidegree distribution of probability ν_m at the variable nodes and μ_m at the check nodes. For instance, with probability 0.0225, a variable node has multidegree [2, 57, 0], *i.e.* it has 2 sockets for edges of type 0, 57 sockets for edges of type 1, and no socket of type 2. Check node probabilities sum to $1 - R = 0.98$ since there is 0.98 check node for 1 variable node.

ν_m	m			μ_m	m		
0.0225	2	57	0	0.010625	3	0	0
0.0175	3	57	0	0.009375	7	0	0
0.96	0	0	1	0.6	0	2	1
				0.36	0	3	1

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Reviews of Modern Physics **81**, 1301 (2009).

[2] That is one does not have to make any assumption on the capacities of the eavesdropper (calculation power, knowledge of efficient algorithms, amount of memory...) to prove the security of the established key.

[3] For a discussion about how such a channel can be established in practice, refer to [35].

[4] "ID Quantique," `http://www.idquantique.com`.

[5] "MagiQ Technologies," `http://www.magiqtech.com`.

[6] "Quintessence Labs," `http://www.quintessencelabs.com`.

[7] "SeQureNet," `http://www.sequrenet.com`.

[8] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).

[9] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Phys. Rev. Lett. **93**, 170504 (2004).

[10] R. Garcia-Patron and N. Cerf, Phys. Rev. Lett. **97** (2006).

[11] M. Navascués, F. Grosshans, and A. Acín, Phys. Rev. Lett. **97**, 190502 (2006).

[12] A. Leverrier, F. Grosshans, and P. Grangier, "doibase-10.1103/PhysRevA.81.062343 Phys. Rev. A **81**, 062343 (2010).

[13] R. Renner and J. I. Cirac, Phys. Rev. Lett. **102**, 110504 (2009).

[14] A. Leverrier and P. Grangier, Phys. Rev. Lett. **102**, 180504 (2009).

[15] A. Leverrier and P. Grangier, Phys. Rev. A **83**, 042312 (2011).

[16] A. Leverrier, R. Alléaume, J. Boutros, G. Zémor, and P. Grangier, Phys. Rev. A **77**, 42325 (2008).

[17] J. Lodewyck, M. Bloch, R. Garcia-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. Cerf, R. Tualle-Brouri, S. McLaughlin, *et al.*, Phys. Rev. A **76**, 42305 (2007).

[18] S. Fossier, E. Diamanti, T. Debuisschert, A. Villing, R. Tualle-Brouri, and P. Grangier, New Journal of Physics **11**, 045023 (2009).

[19] G. Van Assche, J. Cardinal, and N. J. Cerf, IEEE TRANS.INFORM.THEORY **50**, 394 (2004).

[20] M. Bloch, A. Thangaraj, and S. W. McLaughlin, CoRR **abs/cs/0509041** (2005).

[21] C. Silberhorn, T. C. Ralph, N. Lütkenhaus, and G. Leuchs, Phys. Rev. Lett. **89**, 167901 (2002).

[22] R. Namiki and T. Hirano, Phys. Rev. Lett. **92**, 117901 (2004).

[23] A. M. Lance, T. Symul, V. Sharma, C. Weedbrook, T. C. Ralph, and P. K. Lam, Phys. Rev. Lett. **95**, 180503 (2005).

[24] M. Heid and N. Lütkenhaus, Phys. Rev. A **76**, 022313 (2007).

[25] T. Richardson and R. Urbanke, *Modern Coding Theory* (Cambridge University Press, New York, NY, USA, 2008).

[26] M. A. Shokrollahi and R. Storn, in *Proceedings of the 2000 IEEE international conference on Symposium on Information Theory* (2000).

[27] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, IEEE Trans. Inform. Theory **47**, 619 (2001).

[28] D. Elkouss, A. Leverrier, R. Alléaume, and J. Boutros, in *Proceedings of the 2009 IEEE international conference on Symposium on Information Theory - Volume 3*, ISIT'09 (IEEE Press, Piscataway, NJ, USA, 2009) pp. 1879–1883.

[29] T. Richardson and R. Urbanke, "Multi-edge type LDPC codes," presented at the Workshop honoring Prof. Bob McEliece on his 60th birthday, California Institute of Technology, Pasadena, California (2002).

[30] I. Andriyanova and J. Tillich, CoRR **abs/1010.1911** (2010).

[31] G. Van-Assche, *Quantum Cryptography and Secret-Key Distillation* (Cambridge University Press, New York, NY, USA, 2006).

[32] A. Leverrier and P. Grangier, arxiv preprint:1002.4083 (2010).

[33] D. Elkouss, J. Martínez-Mateo, D. Lancho, and V. Martin, CoRR **abs/1006.2660** (2010).

[34] M. Berta, F. Furrer, and V. B. Scholz, CoRR **abs/1107.5460** (2011).

[35] S. Kunz-Jacques and P. Jouguet, CoRR **abs/1109.2844** (2011).