An Introduction to Practical Quantum Key Distribution

Omar Amer, Vaibhav Garg, and Walter O. Krawec

Abstract

Quantum technology, and especially quantum computing, is advancing rapidly. Our every-day secure communication infrastructure relies heavily on public key cryptography; unfortunately, many public key schemes are in fact insecure against quantum algorithms. Furthermore, adversaries who, today, capture information encrypted using classical key distribution systems, can wait and, as soon as quantum computers of sufficient power become available, they may decipher all previously captured information. While some new “post-quantum” public key systems are assumed to be secure against quantum computers, this is only an assumption and, even if the assumption remains valid, still results in systems that are less efficient for end-users than the more commonly used public key systems employed today.

While the advent of quantum technology may seem detrimental to security, in fact, one may harness the power of quantum through Quantum Key Distribution (QKD) for stronger security guarantees. Such QKD protocols allow two (or more) parties to establish a shared secret key, secure against an all-powerful adversary. Perhaps surprising is that QKD technology is already here both experimentally and commercially.

This tutorial serves as an introduction to basic quantum key distribution along with QKD technology from a practical perspective. Quantum communication is a highly interdisciplinary field of research and one of the goals of this paper is to introduce a larger set of researchers and practitioners to its study in the hopes of furthering its progress and eventual wide-scale adoption.

I. INTRODUCTION

Currently the security of much of our communication infrastructure depends entirely on certain computational assumptions. As a single example, consider SSL/TLS [1], [2], the protocol used whenever

O. Amer is with the University of Connecticut
V. Garg is with Comcast Cable
W. O. Krawec is with the University of Connecticut
one visits an https site. This secure communication protocol utilizes several classical public key cryptographic primitives as building blocks to establish a private channel between two end-users. However, it is a mathematical fact that the security of any classical key distribution system (either those in use now, or those which have yet to be developed) must, by necessity, rely on assumed limits on an adversary’s power (e.g., one must assume certain problems are difficult to solve computationally). Should these assumptions prove false, our ability to communicate securely using these protocols and systems will come to a sudden end. Not only this, but secrets which were protected using these systems may no longer remain secure and would be open to attack. Finally, even if these computational assumptions remain true, the advent of quantum computers may lead to security breaks or inefficient constructions. It is difficult to say for certain when quantum computers, on a scale large enough to harm our current-day secure communication systems, will appear; however, if their appearance comes suddenly, organizations will have little time to adapt to their creation.

However, unlike classical key distribution systems, Quantum Key Distribution (QKD) protocols allow for the establishment of a secret classical key, secure against computationally unbounded adversaries; that is, security is proven in terms of adversaries limited only by the laws of physics, rather than the assumption that a mathematical problem is difficult to solve. This is in strict contrast to classical key distribution protocols, where it is proven that security always requires some computational limitation be placed on the power of the adversary. QKD attains this goal by the careful utilization of a quantum communication channel, in addition to several classical post-processing methods. Indeed, QKD protocols operate, necessarily, through careful mixture of new quantum communication methods, and already established classical cryptographic primitives.

Originally developed in 1984 by Bennett and Brassard [3] (though quantum cryptography was actually first proposed in the 1970’s by S. Wiesner [4]), and also independently in 1991 by Ekert [5], QKD protocol design and analyses has flourished as a field yielding numerous protocols, security analyses, and practical implementation methodologies. However, despite this, all QKD protocols follow the same general methodology involving a quantum communication stage followed by a classical post processing stage to yield the secret key.

At its core a QKD system, which allows a communication network to establish a secure secret key between two parties $A$ and $B$, consists of a quantum communication channel along with a classical communication channel. One of the fascinating, and from a cryptographic perspective, useful, properties
of quantum communication is that any attempt by a party (say an adversary) to gain information on the data sent through a quantum channel, necessarily causes a disturbance which may be detected by honest parties. The more information an adversary attempts to gain, the more noise in the channel she will create. Unlike in classical communication, with quantum communication, there is a direct correlation between observed quantum channel noise, and adversarial information gain. Furthermore, any attack on a quantum channel cannot be performed passively; instead it must be active as one cannot copy quantum bits with certainty due to the “no-cloning theorem.” [6]

Finally, unlike quantum computers which are difficult to implement in practice due, in large part, to the requirement that many hundreds or thousands of quantum bits must interact with one another, quantum key distribution is much easier to implement as it requires only single qubits sent. Due to this, there are many experimental and commercial applications of QKD. This also includes satellite communication [7] and mobile freespace QKD, even between a ground station and an airplane in flight [8]. These are just a few of the numerous instances of QKD put into practice recently which we will review in this work.

This survey will analyze QKD technology, both in theory and in practice. Unlike other surveys of QKD, we focus primarily on giving a high-level overview of quantum communication for readers not familiar with quantum physics or information theory. We also focus on practical applications of QKD technology. We will begin with an introduction to the basic resources necessary for QKD to operate. We will investigate, in detail, source and receiver methods along with various methods of encoding data in practice. We will also discuss free space and fiber communication. A discussion on classical-quantum network architectures and protocols is also covered. Finally, we cover current experimental and commercial efforts in QKD technology along with current telecom company investments. This paper aims to be a high-level overview of QKD technology; for technical details, we provide references throughout which provide more in-depth overview of individual aspects of the topic covered. For more detailed, technical, surveys of QKD, the reader is refered to [9], [10], [11], [12].

A. Communication Resources

At a high-level, QKD protocols require two communication channels: a quantum communication channel and an authenticated classical channel.

Quantum Communication Channel: allows quantum information to be sent from Alice (A) to Bob (B). Practical quantum channels are often photon channels, however even here there are different encoding methods (and may also lead to higher-dimensional quantum systems depending on the encoding used).
A qubit is simply a mathematical abstraction for a two-level quantum system; physically a qubit may be implemented using photons in a variety of encoding manners. Common methods including a photon’s polarization, phase encoding, or time-bin encoding. We will discuss these methods in detail in Section 2.

Security proofs assume the adversary has full access to the quantum channel. Furthermore, we assume the worst case in that the adversary (E) can also replace the natural, potentially noisy, quantum channel connecting A to B with a noise-less quantum channel; this assumption is to E’s benefit and we must assume that any noise observed in the quantum channel is not due to natural noise in the channel (inherit in any practical quantum channel, e.g., fiber lines), but is actually due to E’s attack. Thus, one generally does not “calibrate” a system to ignore observed natural noise.

There are three common attack models considered in QKD security proofs: Individual Attacks, Collective Attacks, and General Attacks. The last, general attacks, are the most powerful and allow the adversary to attack in any way allowed by quantum physics. Due to the difficulty in analyzing general attacks, often, security is proven against collective attacks (attacks where the adversary probes each signal sent independently and identically). For some protocols, such as BB84, security against collective attacks implies security against general attacks through de Finetti arguments [13] or postselection techniques [14]. Individual attacks are the weakest attack model - they assume E attacks independently and identically however, unlike collective attacks, she must also measure her quantum ancilla before A and B use their derived key. For more information on these different attack models, the reader is referred to [9].

Authenticated classical channel: allows A and B to send authenticated messages to one another (e.g., any message B receives on this channel can be verified by him that it actually came from A). This channel is only authenticated and not secret. Thus, it is a classical communication channel on which A and B, the honest participants of the QKD protocol, may read and write to; however E, the adversary can only read from this channel (i.e., any message sent on this channel is disclosed to E). This channel is often referred to as the public communication channel as anything sent on it becomes public knowledge (or at least to a public of three: A, B, and E).

Authentication can be done in an information theoretical secure manner (i.e., security may be guaranteed against all-powerful adversaries). See [15] for an example authentication method that is information theoretic secure. Note that there may be other methods of authentication used and so the exact details of this channel are generally outside of scope for QKD protocol design.

The authenticated channel is used, at a minimum, for the following necessary tasks in a QKD protocol:
1) Sifting and Error-rate Estimation: $A$ and $B$ must use the authenticated channel to agree on the current error rate of the quantum channel; this is generally done by choosing a random sample of outcomes and disclosing full information for this sample over the authenticated channel. $E$ cannot tamper with this data, however it is not secret and so these sampled iterations must be discarded.

2) Error Correction: $A$ and $B$ must run an error-correcting protocol allowing $A$ and $B$ to agree on a single final key.

3) Privacy Amplification: $A$ and $B$ must run a privacy amplification procedure in order to minimize $E$’s information.

The above tasks, which we will discuss again momentarily in more detail, are all performed with the aid of the authenticated channel. Note that this channel is *not secret* and so any message sent over it is potentially read, in full, by $E$; however $E$ cannot forge messages of her own.

**B. Basic Quantum Information Theory**

It will be useful in later sections for readers to have at least some understanding of the general framework and terminology of quantum information theory. Here we discuss briefly some of the postulates of quantum information theory [16], as well as a number of useful results in the field and the implications of said results on the work we will be discussing. We present these postulates and results in terms specific to our application. In the interest of accessibility we do not, any more than is necessary, delve too deeply into the mathematical framework underpinning these postulates and results. With that said, interested readers should consult [16] for a more in depth discussion of the concepts mentioned below.

1) Quantum States

In classical information theory, there are many different phenomena we can use to implement our fundamental unit of data, the bit. Regardless of the medium used, we can consider all bits equivalent, and thus abstract away the specifics of the implementation. Likewise, in quantum information theory, we often choose to abstract away the specifics of the implementation, and instead represent our system with qubits. A qubit, like a classical bit, can exist in a 0 or 1 state, however, unlike a classical bit, it is also possible for a qubit to exist in an infinite number of *super-position* states.

The first postulate of quantum information theory tells us that any closed quantum state may be represented mathematically as an element of a Hilbert space. If the dimension of the system, $n$, is finite (e.g., a quantum bit is a two-dimensional system), this is equivalent to representing the quantum state as a normalized vector in $\mathbb{C}^n$. 

5
When analyzing non-trivial systems, this vector notation quickly becomes unwieldy. To remedy that, we instead use *bra-ket* notation (introduced by Dirac and is a play on the word “bracket”). A *ket* is simply a variable representing a column vector. For example, we write \( |0\rangle \) (read "ket-zero") to mean the 0 state we discussed earlier, and \( |1\rangle \) to represent the 1 state. These states are chosen so that they are orthonormal and normalized, and span \( \mathbb{C}^2 \). Thus, they constitute an orthonormal basis. These states in particular form what’s usually called the computational basis:

\[
Z = \{ |0\rangle , |1\rangle \}.
\]

Additionally, we use *bras*, denoted \( \langle \cdot | \) to mean the conjugate transpose (in the finite dimensional case) of the corresponding ket \( |\cdot\rangle \). Note that this also leads to the following notation used to represent the inner-product of two vectors \( |\psi\rangle \) and \( |\phi\rangle \) as simply \( \langle \psi|\phi\rangle \).

Additionally, we will often talk about quantum systems that are composed of multiple, smaller systems. To model this mathematically, the tensor product is used, denoted

\[
|\psi\rangle_A \otimes |\phi\rangle_B = |\psi\rangle_A |\phi\rangle_B = |\psi,\phi\rangle_{AB},
\]

where subscripts represent the different subsystems of the joint system. Note that if the joint system is composed of subsystems A, with dimension \( n \), and B, with dimension \( m \), then the joint system \( AB \) is dimension \( nm \). Thus \( k \) qubits are of dimension \( 2^k \).

Qubits, commonly used in QKD protocols (at least in the ideal description), are two dimensional quantum states, but quantum states can be of higher dimension as well. Some protocols make use of these higher dimensional quantum states [17], [18], and it is common to assume that the adversary has access to systems of arbitrary dimension as well.

2) Measurement

To extract information from quantum systems, we need to conduct a *measurement* of the state. Unlike in classical communication, measuring a quantum state is potentially destructive and results in a probabilistic outcome. The simplest type of measurement is a projective measurement, taken with respect to some basis, for example, \( \{ |v_0\rangle , |v_1\rangle \} \). The measurement postulate tell us what the probability of seeing a particular outcome when measuring a state is, and what the state collapses to after seeing that outcome. In particular, given a state \( |\psi\rangle \), the probability of observing \( |v_i\rangle \) if making a measurement in the above basis, is \( |\langle v_i|\psi\rangle|^2 \). Following this measurement, the state collapses to the observed basis state.
As a more concrete example, consider the Hadamard (X) basis, consisting of states $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Now, if we measure the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state in the Hadamard basis. We will see $|+\rangle$ with probability 1, and the state remains $|+\rangle$. Now, if instead we measure in the computational, Z basis, we will observe either $|0\rangle$ or $|1\rangle$, with probability $\frac{1}{2}$ each, and the state, previously an equal superposition of the two, will collapse to whichever basis state was observed.

Crucially, the information about what the state was previously is destroyed in this process, and there is no way to reverse the measurement. Consider, for example, that we had observed $|0\rangle$, itself an equal superposition of the X basis states $|+\rangle$ and $|-\rangle$. Even though the state was originally $|+\rangle$, if we now measure in the X basis, we see $|+\rangle$ with $\frac{1}{2}$ and $|-\rangle$ with $\frac{1}{2}$ probability. Once a measurement is made in a certain basis, there is no way to reverse the measurement and find out what the result may have been in a different basis; the original state has been destroyed and you are left only with the collapsed state — a fact that is crucial to many QKD protocols.

3) No-Cloning Theorem

The No-Cloning Theorem [6] tells us that, for an unknown quantum state, there is no sequence of quantum operations one can perform on an unknown quantum state to make a perfect copy of it. That is, there is no sequence of operations such that, starting with an unknown state $|\psi\rangle$ and a known “empty” quantum state $|0\rangle$ (causing the joint state to be $|\psi\rangle \otimes |0\rangle$) to become the state $|\psi\rangle \otimes |\psi\rangle$, where the originally “empty” register now has a perfect, and independent, copy of the original unknown state.

The implication of this, for the field of QKD, is huge. As we discussed earlier in this section, measuring a quantum state results in an irreversible change in the state, and as we’ll discuss later, this phenomena is a large part of what makes QKD secure. Without the no cloning theorem, however, destructive measurement would be less impactful. Consider an adversary that was not bounded by the theorem, and could, after intercepting a quantum state en route from one party to another, make an exact copy of it without disturbing the original. Such an adversary could, among other things, conduct their attacks on these copied qubits so that Alice and Bob are kept oblivious of any interference into the protocol. Indeed, there would not be any interference as the adversary would not be interacting with the qubits that Alice and Bob see, simply their own copies. Thus, this forces all adversaries in quantum communication, to be active - they cannot copy quantum data to attack offline at a later time.
4) Entanglement

Quantum entanglement deals with correlations between disparate quantum systems that hold even when physically separated. Informally, when two quantum systems are entangled there is a correlation between their states such that you cannot describe the state of one system without making reference to the state of the other. Moreover, operations applied to one system can have an effect on the state of the other system, even when separated by distance. For example, consider an entangled pair of qubits, with one of the qubits at a station on earth and the other held at a satellite orbiting earth. If the joint system is in some superposition of states, and entangled, then researchers on earth can measure the qubit that they have access to and know what the qubit on the satellite’s state is. Entanglement is crucial to a class of QKD protocol, and one could also expect that an adversary might use entanglement to correlate their system in some way with the results of Alice and Bob’s measurements and choices. Even further, by using Quantum Teleportation protocols [19], two parties can use a shared entangled pair of quantum states along with a authenticated, classical communication channel to communicate arbitrary quantum states even when separated by great distance and not connected by a quantum communication channel.

C. A High-level Picture of QKD

QKD protocols operate in two stages, the first of which is the quantum communication stage; this is followed by the classical post-processing stage. Both stages are required to successfully establish a shared key between two end users.

First Stage: Quantum Communication Stage: The goal of the first stage, the quantum communication stage, is for A and B to establish a raw key $r k_A$ and $r k_B$ respectively. This key should be partially correlated (there may be errors in their raw keys due to natural noise or noise induced by an adversary). It is also only partially secret ($E$ may have some information on the raw key due to her attack). Thus, the raw-key cannot be used directly as a cryptographic key.

The raw key is created (or distilled) through use of both the quantum channel and the classical authenticated channel. This stage of a protocol is also the one that varies greatly across different QKD protocols (the classical post processing stage is usually standard and independent of the actual protocol itself).

This stage of a protocol typically operates over numerous, independent, iterations. On each iteration, A and B will attempt to distill one new raw key bit (thus, the size of the raw key after $N$ iterations is no greater than $N$ and typically less since some iterations will lead to inconclusive results and must
later be discarded as we will see). A QKD protocol specifies the steps for each party to take on each iteration. As an example, consider the BB84 protocol [3] shown below:

**Quantum communication stage of the BB84 protocol:** The encoding of classical to quantum data is performed using Table I. Note that step 3, the use of the authenticated channel, can actually be performed once “in bulk” at the very end of the quantum communication stage as opposed to every iteration as shown here.

Repeat for each iteration \( i = 1, 2, \cdots, N \):

1. \( A \) chooses a candidate raw-key bit \( rkb_A \) for this iteration uniformly at random. She also chooses a basis to use \( Z \) or \( X \), choosing the \( Z \) basis with probability \( p_Z \) and the \( X \) basis with probability \( p_X = 1 - p_Z \). She then sends to \( B \) a qubit encoding her raw key choice in the given basis. (Actual qubit encodings as photons are discussed in Section 2.)

2. \( B \) chooses, independently of \( A \), a basis to measure the incoming qubit in (either \( Z \) or \( X \)) - he chooses the basis using the same probability distribution as \( A \), namely \( p_Z \) and \( p_X \). He will then measure the qubit he received in the chosen basis and translate the outcome to his raw key bit \( rkb_B \) (again, using Table I).

<table>
<thead>
<tr>
<th></th>
<th>( Z ) Basis</th>
<th>( X ) Basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical 0</td>
<td>(</td>
<td>0\rangle )</td>
</tr>
<tr>
<td>Classical 1</td>
<td>(</td>
<td>1\rangle )</td>
</tr>
</tbody>
</table>

**TABLE I**

Showing the encoding of classical to quantum data used by the BB84 protocol. Note that the \( X \) basis is spanned by states \( |+\rangle \) and \( |-\rangle \). Measuring a \( Z \) basis state in the \( Z \) basis yields a deterministic result (similarly for the \( X \) basis); measuring in the incorrect basis (e.g., measuring a qubit prepared as a \( Z \) basis state such as \( |0\rangle \) in the \( X \) basis) yields a random result (in this case, \( |+\rangle \) or \( |-\rangle \) will be observed with probability \( 1/2 \)).

**Stage 2: Classical Post-Processing:** After the quantum communication stage has been executed for \( N \) iterations, the parties now in the second stage, classical post-processing, begins. First, they must establish a raw-key through **sifting** and conduct parameter estimation. To sift the data, for each round \( A \) and \( B \)
disclose their choice of basis using the authenticated channel. If they chose different bases, they will discard the iteration; otherwise their resulting raw key bits are appended to their respective raw keys. Note that, if $p_Z = 1/2$, then half the iterations are discarded; however one may optimize the choice of these settings to produce more efficient systems. Note also that any adversary now learns the correct basis information, however due to the before mentioned no cloning theorem, this does not immediately help an adversary learn the correct transmitted bit (as the adversary could not clone the original qubit to perform the correct measurement now - for instance, if the adversary measured in a randomly chosen basis, the adversary may learn that she chose the correct basis but this is after the fact and she will not always correctly measure; thus she gains partial information from this, but not full).

Next they begin error correcting as well as parameter estimation by first broadcasting a subset $\tau \subset \{1, 2, \cdots N\}$ of size $m$, chosen by one of the parties, over the authenticated channel. For every index $i \in \tau$, all measurement data (including potentially the raw key bit itself) for iteration $i$ of the quantum communication stage is revealed. This allows $A$ and $B$ to establish statistics on the noise properties of the quantum channel (such as, what is the probability of a $|0\rangle$ being measured as a $|1\rangle$ or a $|−\rangle$ being measured as a $|+\rangle$; both of these are considered noise). If the noise level is “too high” parties will signal to abort the protocol. The threshold for which the quantum bit error rate (QBER, or, the probability that a state prepared in a basis and measured in the same basis is not observed to be in the original, prepared state) is too high (called a QKD protocol’s noise tolerance) depends on the quantum communication stage of the protocol and is a key statistic required in any security proof of a QKD protocol. As an example, for BB84 without additional advanced classical post-processing, the theoretical noise tolerance is 11% [20] (with additional classical post-processing such as classical advantage distillation, this tolerance can be increased to over 20% [21], [22], [23]). Note that we cannot distinguish between natural noise and noise induced by an adversary. Thus, any talk of “noise” in a QKD protocol must, necessarily, assume the worst case that the noise is completely adversarial. Natural noise, therefore, is considered also noise induced by a potential adversary.

If the noise is not over the limit, then an error correction protocol, such as Cascade [24] or LDPC [25], is run, again using the authenticated classical channel. This allows $B$ to “fix” the errors in his raw key. After this, except with negligible probability, $A$ and $B$ now have a correlated raw key - a string of classical bits. This error correction protocol, however, leaks extra information to $E$ (this extra information is added on top of whatever information she already potentially has on the raw key based on
her attack performed during the actual quantum communication stage (i.e., her probing of the quantum communication channel while it was in use).

Finally, a privacy amplification protocol is run which takes as input the error corrected raw-key and outputs a (potentially much smaller) secret key which may then be used for other cryptographic purposes. The security property of privacy amplification (and, thus, a QKD protocol) guarantees that, except with negligible probability, the secret key that is output is indistinguishable from a uniformly generated random key independent of any adversary; furthermore, this adversary has no computational assumptions placed on it [26]. Privacy amplification is generally implemented using two-universal hash functions, taking the error-corrected raw key, and hashing it down to a smaller, secret key [27]. The size of the final secret key depends on how much information $E$ has on the raw-key. Unlike classical communication, with quantum communication, one may bound $E$’s maximal information based only on observed noise. Thus, the noisier the channel, the more information $E$ potentially has, the smaller the final secret key will be.

The authenticated channel generally uses an information theoretic secure (i.e., “perfectly secure”) Message Authentication Code (MAC) requiring a pre-shared secret key. How this initial key is installed in the system depends on the application - for instance, it can be hard-coded in the initial point-to-point link. Note that, to retain perfect security, classical key distribution cannot be used to establish this initial key. For any such MAC, whenever a message is authenticated, some amount of the key used for the creation of the MAC tag is depleted, and so with a fixed amount of key information there is a limited number of messages that can be sent. Note, however, that unlike information theoretically secure encryption (e.g., OTP), the size of the key for authentication can be much smaller than the message being authenticated (roughly the log of the message size)[15]. Thus, following the successful completion of a QKD protocol, a suitably sized portion of the resulting secret key (after privacy amplification) may be used to “refill” this authentication key material, leaving additional bits left over to be used for any other cryptographic task the user wishes. That is, QKD will produce a new shared secret key that is significantly longer than the key required to authenticate messages and so a portion of it may be used to replace the old, used, shared authenticated key (needed to repeat QKD) while having new key material left over for the user’s application. See Figure 1 for a general outline of the operation of a QKD protocol.

1) Efficiency and Noise Tolerance

Two important characteristics of any QKD protocol is its efficiency and its noise tolerance. QKD protocols operate in blocks of size $N$ (typically $N = 10^4$ up to $N = 10^9$). After a block has been sent
Fig. 1. An overview of a QKD protocol. On step (1), users begin with a small pre-shared secret key $k_0$. They then perform the quantum communication stage (QCS) of the protocol which is the only part that actually uses quantum resources. They then sift this data and conduct parameter estimation leading to the generation of a raw-key, as well as a bound on the amount of information an adversary could have on the key. They then run error correction followed by privacy amplification (which involves choosing a random two-universal hash function $f(\cdot)$ and sending it to $B$) resulting in a new secret key $k_1$ which consists of the unused portion of $k_0$ (some of $k_0$ will be used for the authenticated classical channel) denoted $k'_0$ combined with the newly generated secret key after privacy amplification, denoted $f(rk_A)$.

Through the quantum communication stage, it is processed in bulk through sampling, error correction, and privacy amplification leading to a new secret key which is added to the key pool. However, as the noise increases, the secret key produced shrinks in size, decreasing efficiency (as one must still produce a block of size $N$ before processing). Once the noise surpasses the threshold, no secret key can be distilled from the block.

Consider a QKD protocol operating in blocks of size $N$ and, so, the raw-key, after sampling, is of size $N' = N - m$, where $m$ is the size of the sample used for estimating the noise in the channel (as discussed in the previous subsection). From this, error correction and privacy amplification will yield a secret key of size $\ell(N')$. Note that, to produce a block of useful data of size $N$, would require $N_{\text{sig}} \geq N$ signals (qubits) passing from $A$ to $B$. Due to photon loss (to be discussed in Section 2), and incompatible basis choice in the protocol (e.g., if running BB84, $A$ and $B$ choose to use different bases), $N_{\text{sig}}$ is generally orders of magnitude larger than $N$, while $N$ may be $10^4$ to $10^9$ bits [28].

The size of the block, $N$, has impacts on many of the practical aspects of QKD, which are often referred to as finite-key effects. These effects include the probability of error correction failing, the amount of
uncertainty one has regarding the adversary’s information on the block and therefore the amount of key material lost to privacy amplification. These effects are lessened by having a larger block size, but there is a practical downside to having a large $N$. Larger $N$ generally means a larger input into the error correction and privacy amplification algorithms, which means these stages become computationally intensive and time-consuming, and depending on the computational ability of your hardware then having too large of an $N$ can actually reduce the number of key bits generated per second, even while increasing the total number of key bits generated from the data.

The security guarantee of a QKD protocol is as follows [27]: Assuming $A$ and $B$ do not abort (which must be taken into account [29]), then let $K_A$ and $K_B$ be the final keys output by $A$ and $B$ respectively (after error correction and privacy amplification - thus these are the keys intended for actual use in other cryptographic applications), then, with high probability, the protocol is correct (i.e., $Pr(K_A \neq K_B) \leq \epsilon$), and the final key is $\epsilon$ close to a truly uniform random string independent of any adversary system. Through the quantum communication stage, we are able to bound the amount of information Eve has on our key string so that we are actually able to run privacy amplification and achieve this guarantee. For more information on this theoretical notion of QKD security, the reader is referred to [27], [30], [31], [14], [32], [33], [9], [29]. In particular, [27], [29] discuss the information theoretic derivation of QKD security and the notion of composability while the Devetak-Winter key-rate equation, often used to derive asymptotic key-rate expressions for QKD protocols, is found in [32]. Finite key effects lead to increased complications [33] in security analyses, however an exact discussion on this is outside the scope of this survey.

Computing $\ell(N') = |K|$, where $K$ is an $\epsilon$ correct and $\epsilon$ secret key, leads to the efficiency of the system. In particular, one is often interested in the key-rate defined as:

$$r = \frac{\ell(N')}{{N_{sig}}}$$

where, again, $N_{sig}$ is the number of actual signals that passed between users to produce a raw-key block of size $N'$ (which, further, is processed to produce a secret key of size $\ell(N')$). We will consider this ratio later when discussing actual QKD implementations over fiber. For the time being, however, it is illustrative to consider a theoretical case of perfect qubit channels in the asymptotic limit (i.e., for $N_{sig}$ arbitrarily large). In the asymptotic limit for the BB84 protocol, it is known that [27], [20]:

$$r = \frac{\ell(N')}{{N_{sig}}} \approx 1 - 2h(Q),$$
Fig. 2. Theoretical maximal key-rate of the BB84 protocol (Equation 2) as the noise in the quantum channel increases. The noise tolerance is 11% after which BB84 cannot extract any key.

where $Q$ is the quantum bit error rate of the channel (e.g., the probability that a $|0\rangle$ flips to a $|1\rangle$ or a $|−\rangle$ flips to a $|+\rangle$) and $h(·)$ is the binary entropy function: $h(Q) = -Q \log_2 Q - (1 - Q) \log_2 (1 - Q)$. Note that we consider the biased, asymmetric version of BB84 first introduced in [34] whereby basis choices are biased leading to negligible loss due to incompatible basis choices (and in this case it is in fact necessary to consider the noise in each basis independently). This is possible in the asymptotic limit, though in finite key settings, the bias will greatly affect security as it will lead to decreased efficiency (due to incorrect choices) or worse parameter estimation (due to statistical fluctuations in parameter estimation from finite samples). However, for this purpose, note that when the error rate is 0, the key rate is 1 (as $E$ has no information on the raw key and error correction is not performed, thus every qubit sent leads to a secure key bit). Furthermore, the key-rate $r$ (Equation 2) is positive for all $Q \leq 11\%$. This 11% is referred to as BB84’s noise tolerance, at least in this ideal scenario. Other protocols have higher or lower noise tolerances (but, potentially other benefits/weaknesses also). As the noise increases in a quantum channel, the efficiency of the protocol decreases as more information is leaked due to error correction and, furthermore, privacy amplification will output an even smaller secret key in order to “erase” $E$’s information. A graph of Equation 2 is shown in Figure 2.

Of course, this is only considering ideal scenarios. In practice, one must not only worry about “noise” but also device imperfections, practical attack scenarios, and photon loss. These lead to security analyses with greater complexity, yet the underlying theme is the same: namely, estimate the channel characteristics, from this use information theoretic properties to bound an adversary’s information, and, finally, use this
bound to determine how much privacy amplification should shrink the raw key by, leading to a protocol’s key-rate. Notice that QKD technology relies not only on quantum communication and information science, but also classical communication technology and classical cryptography. Both together are necessary to achieve perfect security.

II. Q U A N T U M  K E Y  I N F R A S T R U C T U R E

Any network can be thought of as a set of interconnected nodes, and quantum key distribution networks are no different. At its most basic core, a node consists of a source and a receiver. The exact requirements of the source and receiver depend on the specific QKD paradigm being used, but in general it must be capable of emitting the desired quantum states and performing measurements on the received states. Connecting each node will be a quantum point-to-point link; two of the most commonly used channels are fiber and free-space.

Finally, there are two primary classes of QKD protocols: *Discrete Variable* (DV-QKD) and *Continuous Variable* (CV-QKD) (first introduced in [35], [36], [37]). DV protocols encode key information in discrete variables such as the polarization of a single photon; CV protocols, however, encode information in continuous variables such as the quadratures of quantized electromagnetic modes [38]. In this section we focus primarily on DV-QKD; subsequent sections are applicable to either model. The primary difference between the two lie in the detection methodologies, with discrete variable involving photon counting/detection and continuous variable utilizing homodyne or heterodyne detection. For more information on CV-QKD, the reader is referred to [35], [36], [37], [39], [40], [41], [42], [43].

DV-QKD was the first model of QKD established in theory with the BB84 [3] protocol. Ideally, in any DV-QKD implementation, single-photon sources would be used to encode a single quantum state. On the receiving end, a measurement is done using a photon counter, yielding a potential raw key bit. One benefit to these systems is their ability, in the absence of any noise or device imperfections, to lead to a fully correlated key between $A$ and $B$. Specific DV-QKD protocols include BB84, as already discussed in the introduction, B92 [44], SARG04 [45], E91 [5], COW [46], DPS [47], the three-state BB84 [48] protocol, the six-state BB84 protocol [49], and MDI QKD [50][51] (and there are numerous other protocols at this point with varying theoretical and/or experimental advantages and disadvantages).

On the source side, Alice must encode quantum states in single photons to send to Bob (or, at least, limit the preparation of multiple photons). If Alice sends two or more photons, encoded in the same basis, Eve could capture the second photon and later, when basis information is revealed, learn complete
information on the encoded information — this is the so-called photon-number-splitting (PNS) attack [52], [53]. Thus, at least with BB84, whenever two or more photons are sent in a single iteration of the quantum communication stage, one must assume that Eve has full information on that round (recall that basis information is later revealed during the BB84 protocol, so if $E$ can capture and store the second photon, she can later make the correct measurement to learn $A$’s raw-key bit without creating any observable noise). Therefore, the probability that multiple photons are sent must be small.

In practical implementations, one common source is a weak attenuated laser. Such a source actually emits $n$ photons according to a Poisson distribution:

$$p(n|\mu) = \frac{e^{-\mu} \mu^n}{n!}$$

where $\mu$ is the intensity (mean number of photons per pulse) of the laser (which may be set by the user). For standard QKD protocols, one typically sets the intensity, $\mu$, to be small so that, on average, the laser source actually emits less than one photon per pulse and the probability of sending two or more photons on any round (thus leaking vital information to Eve), is small. However, the lower the intensity, the more likely a vacuum event will occur (i.e., with high probability, no photon will be emitted) which causes the round to become useless. Therefore, used with plain BB84, the compromise is too low of an intensity and the key-rate will drop due to the fact that little information is carried between the users (i.e., on most rounds $A$ sends no photons so nothing is distilled); conversely, too high of an intensity and the key rate will also be low as Eve will have more information due to the high probability of multi-photon events. As an example, if we select an average intensity of $.1$, then the probability of sending multiple photons is $.004$, a single photon is $.0904$, and a vacuum is $.904$. If we select a higher intensity, such as $.4$, then the probability of sending vacuum states is lower, at $.67$, while the probability that we send single photons is $.26$, but this comes at the cost of sending multi-photon signals and leaking information to Eve with probability $.07$. However, these issues can be mitigated through decoy-state protocols [54]. We discuss the decoy state protocol later, however the overall idea is to randomize the intensity $\mu$, thus allowing one to more accurately detect PNS attacks and derive tighter bounds on $E$’s information. With this, and due to their wide-spread use and ease of implementation, attenuated lasers are the most commonly used source in DV-QKD implementations.

The above method is used for prepare-and-measure type QKD systems. These systems have the source preparing a qubit, transmitting it over a channel, while the receiver immediately measures on receipt. An alternative approach is to use entanglement-based QKD, such as E91 [5] and BBM92 [55], where
a source (which may be Alice, or even an adversary), prepares an entangled qubit pair. One qubit will travel to Alice while the other travels to Bob both of whom measure the qubit on receipt. Such pairs are typically produced through a process called spontaneous parametric down-conversion (SPDC) [56], the exact process, however, is outside the scope of this review. Some important considerations, however, are that entanglement-based protocols were shown to tolerate higher levels of loss if the source is placed midway between A and B [57], [58]. Additionally, the use of entanglement allows for device independent or measurement device independent protocols, which can be very powerful in mitigating certain classes of attacks, as discussed later.

On the receiver side (regardless of the choice of prepare-and-measure or entanglement-based), Bob must be able to detect the arrival of photons. In DV-QKD, this is accomplished through photon counters, typically single photon avalanche photodiodes (SPAPDs). Several important parameters characterize a photon counter, namely:

- **Quantum Efficiency**: The probability that the photon counter actually clicks (i.e., signals) when a photon hits it.
- **Dark Count**: The probability that a photon counter clicks when no photon has hit it (a false signal).
- **Dead Time**: The time required for the detector to reset after a click.
- **Afterpulsing**: The probability that a single photon causes multiple clicks
- **Timing Jitter**: The uncertainty regarding the time at which a detector detected a photon

Different APD technology exists with different characteristics. For instance InGaAs APDs, often used as they operate at telecom wavelengths, have an efficiency of 20% (i.e., they only register a photon 20% of the time) and a dark count on the order of $10^{-7}$ (that is, they falsely detect a photon with probability $10^{-7}$).[59]. Although there has been research done into other types of SPDs, avalanche photo-diodes (APDs) are the largely used class of SPDs for DV QKD protocols. Perhaps their largest advantage over other types of SPDs is that both silicon APDs (Si APDs) and indium gallium arsenide APDs (InGaAs APDs) can be operated at temperatures of 250 K, making them far more practical for commercial appliances than other detectors, and InGaAs. Si APDs have a higher quantum efficiency than InGaAs APDs, but Si APDs are only capable of detecting photons with wavelengths between in the range of 400 to 1000 nm. InGaAs, on the other hand, have the advantage of being functional for wavelengths between 950 and 1600 nm (importantly, this includes standard telecom wavelengths). By using parametric upconversion to convert higher wavelength signals to a wavelength in bands appropriate for Si APDs, it
is possible to use the more efficient Si APDs to detect photons at telecom wavelengths. There are losses in efficiency and increases in noise associated with this conversion process, but the method has been refined to be 89% efficient in some cases.

For applications in which the cryogenic temperatures necessary are not inhibitory, superconducting nanowire SPDs (SNSPDs) are an area of much interest. Recent developments in the technology have produced SNSPDs that can operate at greater than 90% efficiency at a large range of wavelengths, including telecom wavelengths with low noise and dark count rates. These devices have also been commercialized by many companies specializing in QKD technologies, and are readily available as fully enclosed systems with closed cryogenic cooling systems [59].

Regardless of the source used to prepare photons, there are, then, several options to actually encode qubit information onto them (which also dictates how to utilize the photon counter to perform the correct measurement in order to decode the information). Three primary methods are polarization encoding, phase encoding, and time-bin encoding.

**Polarization Encoding:** In polarization encoding, a photon is emitted from a source, the polarization of which is altered to encode the desired qubit. Typically, a qubit in the state $|0\rangle$ will be *horizontally polarized* while a qubit in the state $|1\rangle$ will be *vertically polarized*. Other superposition states may be represented with a suitable choice of polarization. For instance, the superposition state $|+\rangle$ can be represented as a 45 degree polarization. The generation of a particular polarization state can be achieved using multiple laser sources (e.g., if one requires a system capable of producing a state $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$ then four laser sources are used) - see [60], [61], [62], [63] for some experimental implementations with this configuration. However such an apparatus leads to potential side-channel attacks due to the different optical properties of each individual laser which an adversary could utilize to determine the choice of state sent [64]. An alternative approach is to use non-passive optical components, such as electro-optic modulators, to vary the polarization of a photon emitted by a single laser (see [65], [66], [67] for some experimental implementations using this form of state preparation).

Measurements in a polarization encoding scheme may also be done through passive or active means. Normally, both systems consist of passing the photon through a polarized beamsplitter (PBS) causing any photon horizontally polarized to pass through one output mode while any photon polarized vertically will output through the other. At the end of each output, a single photon counter (SPC) is placed which will “click” if it is hit by a photon (as there are two outputs to the polarizing beamsplitter, two SPC’s are
Fig. 3. Showing how measurements may be performed in a polarization based encoding scheme. A polarized beam splitter (PBS) allows a qubit to pass one arm only if it is polarized horizontally (denoted here as a state $|0\rangle$) or the other arm if it is polarized vertically (denoted $|1\rangle$). Single photon counters (SPC’s) can detect the presence of a photon. If $|p\rangle$, the input, is in state $|0\rangle$, it will always be detected by SPC1. On the other hand, if $|p\rangle = |+\rangle$, then either SPC1 or SPC2 will detect the photon (with 50/50 probability), but not both (abstractly, this is implementing a $Z$ basis measurement of an $X$ basis state). This same circuit can be used to implement an $X$ basis measurement through the use of a half wave-plate (HWP) which rotates the polarization of the incoming qubit by 45 degrees (thus converting a $|+\rangle$ to a $|0\rangle$ and a $|-\rangle$ to a $|1\rangle$).

Fig. 4. A potential implementation of a phase-encoded QKD system with Alice on the left and Bob on the right. Each party uses an unbalanced interferometer where one arm of the input beamsplitter (BS) is longer than the other. If both parties choose their phases ($\phi$ for $A$ and $\phi'$ for $B$) so that the difference is 0 or 180 degrees, the output of the second interferometer is deterministic; otherwise it is random.

needed). To measure in the $X$ basis, one may simply place a half-wave-plate to rotate the polarization 45 degrees and then repeat the above (or use electro-optic modulators). See Figure 3. Passive optics may be used to make basis measurement choices via the use of a beamsplitter whereas active systems may employ electro-optic modulators before passing through the PBS circuit. For more information on recent experiments in polarization encoding, and the actual apparatus used, the reader is referred to [64], [68], [69], [70], [62], [61], [65], [66], [67].
**Phase Encoding:** Phase-coding [71], [72] is based on the use of Mach-Zehnder interferometers[73]. A prepares a single qubit and passes it through the interferometer. The photon passes the first beamsplitter causing it to travel in a superposition down both arm \( a \) and arm \( b \); the interferometer must be unbalanced so that one arm, say \( a \) is longer than the other to avoid interference at the second beamsplitter. Along arm \( a \), the phase is modulated. The signal is recombined at the second beamsplitter. The receiver is similar, though the phase chosen is independent of \( A \)’s choice. When both \( A \) and \( B \) set their phase modulators so that the difference in shift is either 0 or 180 degrees, the output of the beamsplitter on the receiving end will be deterministic. Otherwise, the photon, on leaving the second beamsplitter in the receiver's apparatus will randomly hit either photon counter 0 or photon counter 1. See figure 4. An interesting use of unbalanced interferometers in the use of QKD design is the so-called “plug-and-play” system, also used commercially [74].

**Time-Bin Encoding:** Another popular method of physically realizing a qubit on a photon is through time-bin encoding [75]; this also has the added benefit of being able to reliably create high-dimensional quantum states [17], [76], [77], [18] which can be used in protocols to achieve higher noise tolerance than similar qubit-based protocols achieve. Here, typically, a single photon is passed through an unbalanced interferometer. That is, a single photon is passed through a beamsplitter, causing it to travel in a superposition down both a short arm \( (a) \) and a long arm \( (b) \). Since one arm is longer, traveling down this path requires additional time. The signals are recombined at a second beamsplitter. However, since the photon has not yet been observed, it is now in a superposition of time slots - traveling in a superposition of the form \( \frac{1}{\sqrt{2}} |t_0\rangle + \frac{1}{\sqrt{2}} |t_1\rangle \). Furthermore, the relative phase along an arm may be altered, causing the creation of any qubit state of the form:

\[
\frac{1}{\sqrt{2}} |t_0\rangle + \frac{e^{i\phi}}{\sqrt{2}} |t_1\rangle,
\]

for user-controlled \( \phi \). Measuring in the “time basis” \( |t_0\rangle, |t_1\rangle \) may be done simply by recording the time of detection. Measuring in alternative bases (needed for QKD security) may be done by passing the photon through a second interferometer [17].

**A. Quantum Communication over Fiber: Decoy State Efficiency**

Regardless of the source and encoding method, two general mediums for quantum communication are available: fiber and freespace. Fiber involves the transmission of quantum information, encoded using photons in one of the various methods described above, over a fiber channel. While past experimental
implementations used dark-fiber, today it is possible to perform quantum communication over fiber also carrying classical information [78], [79], [80], [81], [82]. One prime limiting factor to the use of fiber is the high probability of photon loss as distance increases (indeed, the probability of photon loss is modeled as $1 - 10^{-\alpha \ell/10}$ where $\ell$ is the distance between users and $\alpha$ is typically .1 to .25 dB/km [28]. Note that one cannot repeat quantum signals (as this would involve a measurement - thus decohering the state); while quantum repeaters are a theoretical possibility, they require advanced technology including perfect, or near-perfect, quantum memories which are far from practical today [83], [84]. This, therefore, limits the distance allowed between two QKD nodes connected by fiber. Note that quantum repeaters operate through entanglement swapping via Bell state measurements and do not violate the no-cloning theorem [85], [10].

Not only does loss limit raw key distillation due to $B$ not receiving any information, one must also assume that any loss is due to an adversary’s attack, thus greatly shrinking the size of the final secret key after privacy amplification. The reason for this can be understood from an examination of some common attacks, such as a photon number splitting attack (PNS)[86] and the unambiguous state discrimination attack[87]. In the former, Eve blocks laser pulses that contain only one photon, and siphons off one of the photons from each other the remaining photons allowing her to have more information on each of the successful key rounds than she would have otherwise. Even worse, the B92 protocol, without mediation, is susceptible to the unambiguous state discrimination attack, in which Eve is able to only allow photons through during rounds in which she has full information on the qubit.

Besides photon loss, the use of practical sources, which as discussed produce multiple photons with non-zero probability, is also a bottleneck to efficient QKD over long distances. To mitigate this concern, new protocols were developed which do not leak full information when multiple photons are emitted [45] (as, with these, key information is encoded in the basis choice, not the actual state). An alternative approach is the decoy state protocol which, using BB84-style encoding, allows users $A$ and $B$ to actually estimate the probability of single-photon events and, in particular, channel statistics within those events and to detect photon number splitting (PNS) attacks [54], [88], [89], [28]. Due to its relatively simple hardware requirements, this decoy state protocol is commonly used in experimental and commercial implementations.

The idea behind the decoy-state protocol is, in hind-sight, relatively simple, yet extremely powerful. In essence, instead of $A$ using a fixed intensity setting for her laser source (the $\mu$ in Equation 3), she alters
it choosing, randomly, from three possible settings: signal (the highest), decoy (generally on the order of $10^{-1}$), and decoy-vacuum (some implementations cannot truly be set to vacuum, and so instead their intensity is set to something on the order of $10^{-4}$ - essentially, with this setting $A$ always sends a vacuum state). Since the intensity setting is unknown to $E$ (who cannot distinguish what intensity $A$ actually used on any particular iteration), this allows $A$ and $B$ to get a better idea as to what $E$’s attack strategy is when multiple, or no, photons are emitted. That is, $A$ and $B$ will later sample and collect statistics, separately for all three intensities. If $E$ is performing a PNS attack, there will be a detectable change in behavior between the signal and decoy events (as, normally, $E$ will block single-photon events). Beyond detecting this attack, the decoy state protocol allows one to determine accurate bounds on single-photon errors, needed to derive a more optimistic upper-bound on $E$’s information gain. By determining a more accurate bound on $E$’s information, one need not necessarily shrink the raw key by as much during the privacy amplification process.

For a recent analysis of the decoy-state protocol for practical experimental sources, the reader is referred to [28]. Here we provide a high-level overview of the protocol. In addition to choosing random bases each iteration, as is standard in BB84 (see Section 1), $A$ also chooses one of the three intensity settings $\mu$. The process repeats until a sufficiently large set of raw-key material has been produced (typically $10^s$ for $s = 4, 5, \ldots, 9$ depending on a user specified parameter $s$). Once this block of data has been produced, $A$ and $B$ will sample on a randomly chosen subset of their data, estimating noise and loss levels for each of the three intensity settings. Since $E$ cannot determine what intensity was used, her attack cannot depend on the intensity setting and this allows users to determine good bounds on the noise in single-photon events (even though they cannot measure this exactly as they can never be certain of when a single photon leaves $A$’s lab) [90], [28]. From this, one can calculate exactly the key-rate (Equation 1):

$$r = q(Q_1(1 - h(e_1)) - Q_\mu h(E_\mu)),$$

where $q$ is the probability that $A$ and $B$ choose the same basis; $e_1$ is the estimated single-photon error; $Q_1$ is the estimated probability that a single-photon event leads to a detection at $B$’s end; $Q_\mu$ is the observed probability that an iteration leads to a detection at $B$’s end during rounds in which the mean photon number was $\mu$ (so it is the average over all photon numbers sent by $A$ during rounds where the intensity was $\mu$); and $E_\mu$ is the observed error (again, it is the average of the error over all photon numbers sent by $A$). Recall that $A$ sends multiple photons with certain probabilities and it is never known by users
when a single-photon event occurred. Thus, while $E_\mu$ and $Q_\mu$ can be measured, $E_1$ and $Q_1$ cannot be directly measured. Instead, the seminal result of the decoy state protocol is that they can be bounded leading to more optimistic key-rate computations (and, thus, greater allowed distances between users). Indeed, experimentally, the record distance for a QKD protocol over fiber is 421km [91] using a decoy state protocol introduced recently in [92] (which is a more advanced variant of the one we discussed here).

1) Device Independence

So far we have briefly mentioned some attacks on QKD systems and in a later section we will look at these attacks in more detail. Attacks on QKD systems side-step the theoretical unconditional security guarantees by exploiting imperfections in the devices used to implement the protocols. In an effort to mitigate against the existence of these imperfections, Device Independent (DI)-QKD has been proposed and is an active area of research [93], [94], [95], [96], [97]. These protocols are devised in such a way as to be secure even in the case where the adversary has full information over the devices in Alice and Bob’s nodes. Such protocols generally involve the creation of entangled photons and verifying they are maximally entangled through Bell inequality violations [98]. Using the fact that maximally entangled systems cannot be entangled with any other systems (e.g., cannot be correlated with any adversary), Alice and Bob are then able to be certain that Eve’s system is independent of the key.

An alternative to fully DI-QKD is Measurement Device Independent (MDI) QKD[50], [51], [99]. Unlike DI-QKD, in which it is assumed that all of Alice and Bob’s apparatus are untrustworthy, in MDI-QKD we must assume that at least their state preparation devices can be trusted but not the measurement devices. With this assumption, MDI-QKD can guarantee security against any and all known and yet unknown side-channels in measurement devices (as these are assumed under complete control of the adversary). This guarantee is achieved by having Alice and Bob communicate with an untrusted third party, Charlie, rather than each other. This third party (who may be the adversary) is responsible for performing measurements on the qubits received from $A$ and $B$, reporting the outcomes. Even if this third party is adversarial, security can be guaranteed. Since $A$ and $B$ do not need any measurement devices (only qubit sources), side-channel attacks against measurement devices are eliminated. The current record distance for MDI-QKD is 404km [100]. More recently, a new protocol called Twin-Field-QKD [101] [102], which in some cases can be considered a variant of the MDI-QKD protocol, has been shown to be capable of breaking the 421km record held by the decoy state protocol in [91]. A TF-QKD protocol
generally operates with both users $A$ and $B$ having their own sources and preparing a phase randomized pulse to a centralized measurement device (which may even be adversarial; which is why some TF-QKD protocols may be considered a form of MDI protocol as proven in [103]).

**B. Freespace Quantum Communication**

Complimentary to fiber optic channels is the use of free space QKD. While fiber optic channels exhibit an exponential increase in loss over distance generally due to Rayleigh scattering loss, freespace communication’s loss rate, generally due to diffraction, is only quadratic giving it a significant advantage; see [10] for information on the loss rates of freespace quantum communication versus fiber. Free space QKD may allow us to conduct QKD over vast distances. As would be expected, however, communicating quantum states over free space is a more complex task than simply using fiber optic channels, and so the hardware needed to do so is more complex and has its own associated disadvantages. Here we discuss the implications of the various possible configurations of free-space communication.

A free space communication link is generally composed of at least two entities: namely ground station A and a satellite SAT (though it may also be some other remote entity, such as a plane in flight [8]). We can add a third entity, ground station B, and use SAT as a kind of trusted node (described later in this section) to establish communication between ground stations A and B; this kind of network was used in the ground-breaking Micius satellite demonstration where QKD secured communication was conducted over 7,500 km between Beijing and Vienna [104]. The two main decisions that must be made when configuring a free space network are the location of the satellite and the division of labor [105], [106], [107], [108].

First we will discuss the location of the satellite, for which we consider two main options. The satellites can be placed in low earth orbit (LEO), or geostationary orbit (GEO) [107]. In LEO, the satellite is located at an altitude of 160km to 3000km, orbiting the earth at very high speeds. The low altitude of LEO satellites comes with the advantage of causing less loss than is found with GEO satellites, but LEO satellites must move at relatively high velocities to maintain their orbits, thus they have a much tighter temporal window in which a link can be established with a given ground station. The regularity and frequency of these opportunities can also vary greatly depending on the exact orbit of the satellite. For example, the Micius satellite, which is in a 500km LEO sun-synchronous orbit, is able to link with a ground station in Xinglong for five minutes every night. A satellite in the International Space Station orbit would be limited to only 150 links over the course of a year. A GEO satellite, on the other hand, is
relatively static, and so can continuously service the same stations that wish to use it to establish QKD generated keys with one another. Of course, GEO satellites must be located at an altitude of 35,786 km, and so experience significantly more loss than LEO satellites. In practice, LEO satellites have thus far been the more commonly used satellites for QKD, though there have been some preliminary tests of GEO systems[109]. Additionally, there has been recent work in showing the feasibility of medium earth orbit (MEO) QKD, where it has been shown that QKD can be conducted with satellites situated in orbits of 7000km [110], [111]. MEO is less lossy than GEO, as one might except, while also being more reliably available than LEO.

The simplest forms of satellite based QKD have ground stations each independently conducting QKD protocols with the satellite to establish keys with it [108]. If a pair of ground stations wish to establish a key between themselves, they notify the satellite which can securely communicate a key with each of them individually. For example, if stations A and B each have communicated keys $K_a$ and $K_b$ with the satellite respectively, then the satellite broadcasts to each station (publicly over an authenticated but insecure channel) $K_a \oplus K_b$, allowing A to recover $K_b$ and B to recover $K_a$ (anyone listening to $K_a \oplus K_b$ learns neither). In this case, the satellite is acting as a trusted node.

Trusted node networks [112], [113] are a way to bridge two quantum end-points allowing for increased distances (especially in fiber links), or for key-establishment between two parties who do not have current line-of-sight (for instance, in the satellite case). A trusted node $T$ operates the same protocol as an end-user, establishing two separate keys with $A$ and $B$. Call these keys $k_{AT}$ (between $A$ and $T$) and $k_{TB}$ (between trusted node $T$ and user $B$). The trusted node then typically broadcasts the bit-string $k_{AT} \oplus k_{TB}$ (that is, the bit-wise XOR of the strings). This allows $A$ to compute $k_{TB}$ (which $B$ already knows) thus giving both parties a shared key. The broadcast of $k_{AT} \oplus k_{TB}$ does not give an adversary additional information on $k_{TB}$ (since $k_{AT}$, being a secret key, is uniformly random and independent of any adversary’s information). The disadvantage, of course, is that $T$ must be trusted (and secured from outside hacking) as this entity holds the secret key used by the users.

To agree on these keys, one may choose between three main link types: uplink, downlink, and retro-reflector simulated downlink. In uplink configurations, a source at the ground station transmits the quantum states to a receiver located aboard the satellite, while in downlink configurations the roles are reversed, with the satellite transmitting the states to the ground station. Uplinks, at the cost of having their detectors be at risk of irradiation from cosmic rays, have the advantage of alleviating the need to locate a moving
quantum signal at, a potentially complicated task, while downlinks suffer fewer losses than uplinks [114]. Downlinks are the commonly recommended configuration for satellite QKD, and currently the only demonstrated configuration [115]. A third possibility is the use of retro-reflectors along with an uplink to simulate a downlink configuration [116]. In this scenario, both the source and the receiver are located at the ground station. The ground station transmits a signal to the retro-reflector aboard the satellite where the signal is modulated and reflected back to the receiver, thus simulating a downlink without needing to place a transmitter on the satellite.

A common issue with satellite QKD has been that their operation has been limited to the night, when the apparatus are not prohibited by daylight. Recent work [117], [118], aided by the use of telecommunications wavelengths and more advanced detectors using integrated photonics [118], [119], [120], has shown that it is feasible to achieve long distance quantum communication in both urban and non-urban environments while experiencing loss that is comparable to the loss experienced in LEO systems.

Finally, as a futuristic goal, these kinds of configurations can be combined and chained as parts of larger networks [121]. Satellites can conduct QKD protocols between themselves to transmit keys from more distant regions of the network. A large network may consist of fiber linked metropolitan networks communicating with both LEO and GEO satellites, those satellites themselves also communicating with other metropolitan linked networks and other LEO and GEO satellites ultimately resulting in much larger, QKD secured networks than the ones currently in place. Of course, this is a long-term goal.

C. Loss and Key-Rate

The loss of a quantum channel, as it turns out, is a fundamental limitation on key-generation rates. Regardless of the protocol, and even augmented with unlimited two-way classical communication, it was shown in [122], [123] that the secure key generation rate of any QKD protocol is upper-bounded by a function only the loss in the channel. If $\eta$ is the transmittance of the channel, then the so-called PLOB bound (after the authors of [123] shows the key-rate is proportional to $-\log_2(1 - \eta)$). This rate may be overcome using quantum repeater networks, or, more feasible today, through the use of the newly developed Twin-Field QKD (TF-QKD) protocols [101][102]. The key-rate for TF-QKD protocols generally scale according to to the square-root of $\eta$ in theory. In practice, recent experimental results have already broken the PLOB bound [124] leading to exciting possibilities in scaling QKD infrastructure without necessarily requiring advanced quantum repeaters.
III. INTEGRATING QKD TECHNOLOGY

While much research has been done in developing the theory and practice of QKD communication between two points, there has also been much work done in actually integrating this technology into current, or future, networks allowing the technology to be used by a wide range of users and services. This section will focus on the current state-of-the-art in this area, discussing QKD applications in current communication protocols and also proposed future QKD network architectures. Also mentioned will be a discussion of existing, or past, city-wide QKD networks and their architectures. Finally, we explain various practical security concerns that appear when integrating QKD into current systems along with potential counter-measures.

A. Basic QKD Applications

Much work has been done in integrating QKD into common secure communication protocols, such as IPSec and TLS. The quantum key material generated by QKD can be used both for message authentication and message confidentiality in both protocols. With small changes, such as a method for checking whether there is a path for QKD to occur between two hosts, or a protocol for configuring the parameters of QKD for two hosts, the replacement of public key agreement with quantum key agreement in these protocols is easily realizable.

Some early work in this area includes Ref. [125], where Mink et al., discuss that QKD could be integrated into TLS or IPSec [2], or used for One-Time Pad as an encryption algorithm within these protocols by allowing, as part of the handshake in TLS and IPSec, hosts to agree on using One-Time Pad and Quantum Key information. In [126], a protocol for configuring QKD between two hosts as part of the TLS handshake is suggested in which two QKD capable parties agree on the method of agreeing on a quantum key, the length of the resultant key, and other pertinent parameters. In general, QKD generated key material can be used in place of symmetric and public keys in most applications and protocols, assuming there is a method for clients to agree on and run QKD protocols with each other.

Further, and more generically, as stated in [127], any protocol that requires a pre-shared key, such as TLS and IPSec above, or Kerberos [2], can instead use a key that was distilled from QKD. In the case of TLS, this results in provably quantum resistant security at the transport layer, similarly in the case of IPSec we obtain security at the IP packet layer. For Kerberos, we achieve single sign on, authenticated, encrypted communication between local and remote hosts using session keys that are generated from the
key data produced from QKD between the local and remote hosts that are therefore once again quantum resistant.

Another area of interesting work in this space is the design and implementation of QKD networks that allow for the use of other cryptographic primitives. For example, in [128], the hub and spoke architecture proposed, allows for quantum authenticated key exchange (QAKE), quantum secret sharing (QSS) and quantum digital signing (QDS) protocols between \( n \) parties, with each linked by a quantum channel to a single trusted node. Designing the network architectures and protocols so that they can function in tandem using the same resources helps to amortize the cost of implementing these specialized networks and therefore make their adoption more practical.

B. Network Architectures

How best to design and implement QKD networks is of course an important question. To ensure security in a large network, if two nodes in the network that are not directly connected would like to agree on a key, then each node in the path must also agree on a key with its neighbor using QKD to securely transmit the key along the path; furthermore, these nodes must be trusted, leading to a trusted node network (note that the links connecting nodes, of course, need not be trusted). How the underlying QKD nodes operate in such a scenario, is of course a question of interest. Similarly, if in practice we want quantum key information to be available to a wide number of parties at a reasonable rate, it is important we design the network with these capabilities in mind.

To that end, Tysowski et al., [127] proposed a multi-layered network architecture as a scalable solution for connecting many sites with QKD (see Figure 5). In their model, QKD is used to connect some number of physically secure sites, each with their own LAN containing potentially thousands of hosts. Briefly, the layers of this network, from top to bottom, are the host layer, the key management system (KMS), the quantum network layer (QNL), and the quantum link layer (QLL).

The top layer of the network, the host system, represents all the top level hosts, clients, and applications making use of this network. The users on this layer, when wanting to communicate with other hosts at different sites, make a request to the key management system one layer below to issue a QKD generated key and information about how to generate the same key on a different site. Our requesting party sends this information to the host it is attempting to communicate with, and that host gives that information to its own KMS and retrieves the same QKD generated key. With this, any process requiring a symmetric key can be conducted between these two hosts.
The KMS system is responsible for coordinating the generation and management of different QKD data pools with the QNL, as well as generating session keys from those pools and providing them to the host layer. Depending on the necessities of the specific protocols, the KMS can apply various key-extraction algorithms to the data in the QKD pools to create symmetric keys of various length and with different lifetimes. It is the KMS’ duty to ensure that QKD pools remain synchronized between different sites, and when more key data is needed, the KMS requests that the QNL coordinate QKD between the sites in need of additional data.

When the QNL receives these requests, it coordinates the QLL into conducting QKD from one node to another. Importantly, the QNL is flexible with regards to different QKD Node network layouts. In a
fully connected, meshed network, it is simple for the QNL to instruct Node A and Node B to simply run a QKD protocol between them. With other network topologies, such as a ring or a star, the QNL needs to route the key over various trusted nodes. Of note is that the sites can also be used as trusted nodes themselves, if necessary. The routing protocols proposed for the QNL will take into account the capacities between the links of the quantum network, as well as the demand for keys between each of the sites to react to the needs of the network and optimize accordingly.

Finally, on the bottom of this stack is the QLL. It is the QLL which actually utilizes the quantum channels and conducts QKD between two linked nodes (as discussed in Section 2) to generate key bits for use on the higher levels of the protocol. This can be done with any standard QKD protocol which the QLL is equipped to handle, and in fact it may be the case that based on the status of the network (noise level, demand, etc.), the QNL may instruct the QLL to run different QKD protocols at different times and locations. For a description of what hardware is actually used for this, see Section 2.

As an alternative, [129] proposes a three layer architecture consisting of a quantum, key management, and key supply layer that uses a centralized management server for the entire network. Like in [127], the quantum layer can be arranged in an arbitrary topology.

The key bits are then passed to the corresponding key management agents on the key management layer, where they are associated with the proper headers. The header information includes timing information, source and destination information, key meta-data such as the size and type, as well as other information pertaining to the higher layers of the network. On this layer, key synchronization and identification is realized using authenticated channels. The header and authentication information is sent to the central key management server. When a client attempts to make a secure connection with another client, the key supply agents on the key supply layer make note of the request, updating the header files of the corresponding key files and identify the keys, and transmit the updated headers to the key management server before finally supplying the keys to the clients. The main benefit of the centralized key management server here is the ability to have a central authority tracking and validating the key life cycles, which may be more practical than having each client be responsible for that in more mobile or dynamic networks, such as the secure mobile phone network discussed in [129].

1) City-wide Networks

As QKD research has matured, there have been many large scale tests of QKD networks to compliment the experimental results. The earliest of these was put into place by DARPA in Boston in 2003 [130]. The
Boston network implemented both fiber link and free space links by using a mixture of the classic BB84 protocol and entanglement based protocols. The DARPA network used a total of 10 trusted nodes to connect Harvard and Boston University, and there has recently an initiative in Europe, called OpenQKD, aiming to support and coordinate a number of QKD projects across Europe to encourage practical applications in the field [131]. Further, in China there have been a number of large scale tests of practical, city-wide and mobile network [132].

A more advanced QKD network was put into place in Vienna through the European SECOQC project (SEcure COmmunication based on Quantum Cryptography) through the efforts of 41 research and industrial organizations [70]. The Vienna network incorporated 6 nodes constructed from hardware from a multitude of different sources (including the Cerberis QKD Blade and a prototype from Toshiba, both of which we discuss in Section 4) with the total network spanning 165km in total. In this network, neighboring nodes conduct standard QKD protocols to establish keys between them. For more distant nodes, the nodes between them are used as trusted nodes.

Another of the SECOQC’s innovations is their node specification and the complementary quantum point to point protocol (Q3P) [70]. The SECOQC network is comprised of 6 nodes following this specification, each constructed from different types of QKD hardware, but all of them working to achieve the three main capabilities of the SECOQC node modules. Namely, each node must allow for point-to-point information theoretic secure communication between nodes that are directly linked to that node, each node must be able to compute an efficient path from itself to any other reachable node in the network, and finally, each node must be able to transport secret key material securely over this path using the trusted nodes along the path.

Further, these nodes together take part in the QKD-network layer protocol (QKD-NL). This protocol is responsible for defining and broadcasting the necessary information to properly represent the state of the network and route packets across the network. As noted in [70], it would be possible to use IPv4 or IPv6 to accomplish this goal, but as the routing information packets are authenticated using QKD generated key bits, QKD-NL is designed to be more economic in when and how link announcements are broadcast to be more economic with the key-bits, therefore resulting in an overall more efficient network.

In China, there have been large-scale tests of MDI-QKD reliant networks, in a star shaped network in which a single MDI server is located in the center of a ring of parties. When two parties would engage, they instead transmit information to the MDI server which facilitates the key distribution. Of note is
that in this protocol, only the central MDI server needs to be equipped with single photon detectors, which are often the most expensive component in a QKD network. Such a network was tested in the QKD network Hefei, China. Also in China, a series of metropolitan QKD networks were connected over 2000km with the use of a QKD backbone. Connecting networks in Beijing, Hefei, Shanghai, and Jinan, this backbone operators using the same principles as the Vienna network by utilizing 32 trusted nodes [132]. Importantly, these networks have begun being tested for use in practical metropolitan applications, paving the way for future advances as the practical limitations of the networks are investigated.

C. QKD Security: Theoretical and Practical

As is the case with many other systems, when implementing designs that are theoretically secure, there can be vulnerabilities discovered due to the use of imperfect devices. QKD is no exception to this. There have been discovered several attacks against the actual devices which we discuss here. We also discuss potential counter-measures.

1) Side-channels in the Source

As discussed in Section 2, a weak coherent photon source works by attenuating the laser beam so that with high probability it emits one or no photons, and with lower probability emits multiple photons. In the case that such a device is used, Eve is able to compromise the security of the system by blocking pulses sent from $A$ to $B$ that contain only one photon, and siphoning off one of the photons contained in the higher intensity pulses. She can store these extra photons until the protocol is completed, at which point she is able use the information transmitted during the key distillation phase (which, importantly, discloses the correct basis information) to measure the photons in the correct bases and thus distill her own copy of the secret key, completely undetected.

To counter this attack, the decoy state protocol was introduced which we discussed in Section 2. Its security depends on $E$ not being able to determine what $A$’s intensity setting is (which randomly changes each iteration). However, as shown in [133], even in cases when decoy states are being employed, flaws in the implementation can lead to Eve discovering and abusing side channels to distinguish between rounds in which Alice prepares a decoy state and rounds in which she does not. By doing this, Eve is able to run a PNS attack that bypasses the supposed security afforded by the decoy states.

Further, in discrete variable protocols utilizing some kind of weakly coherent source (a common implementation choice in QKD), Alice must attempt to choose an optimal mean photon number $\mu$ that should be low enough to prevent PNS attacks but high enough to counteract the loss in the channel. In
it was shown that by manipulating the channel into temporarily exhibiting a higher level of loss, an attacker could trick Alice into choosing a non-optimal photon number. In many QKD schemes, a basic assumption is that the mean photon number is low enough that bursts of multiple photons happen infrequently enough that it is safe to disregard them in the security analysis. By making Alice increase the mean photon number, Eve makes the assumption invalid and can, as a result, gain more information on the system than should be possible otherwise, in some cases gaining full information on the key.

2) Side-channels in the Detector

A clever attacker may also be capable of extracting auxiliary information about the system by exploiting insecurities in the practical implementation of the detector. Trojan-horse attacks, as described in [135], are one such attack. In standard QKD protocol’s, components of Alice and Bob’s QKD apparatus must be precisely configured in accordance with the bases they will be using. In a Trojan-horse attack, Eve gains information on these configurations by transmitting a high intensity pulse of light into the apparatus of one of the parties and measuring the reflected light after it has passed back through the apparatus to gain information on the system. In [135], it was shown that an attacker could use commonly available materials to carry out such an attack against commercially available QKD devices (specifically the Clavis2 from IDQuantique - we discuss commercial QKD devices in Section 4) and successfully recover Bob’s bases choices while remaining undetected. In the same paper, it was shown that a viable method for protecting against these attacks could be realized by introducing a filter to the QKD apparatus to block pulses of light with high intensities, a key component in remaining undetected in this attack.

In some cases, even if precautions have been taken to protect against possible exploits, an attacker may still be able to disable these precautions by damaging components by use of a laser, as shown in [136]. One key example of the damage that can be dealt and the resulting possible exploits is the desensitization of detectors that can cause not only a denial of service attack, but can potentially open the door for a Trojan-horse [135] attack, allowing an attacker to recover the entire QKD key.

In one such attack, by thermally blinding avalanche photon detectors (APDs) used in commercial QKD systems [137], Eve was able to recover the entire secret key. In these commercial systems, precautions were already taken to prevent Eve from blinding detectors (making them unable to detect single photons, instead only detecting bright pulses), however in [137], a more complicated blinding attack was proposed in which the photon-detectors were subjected to heating, which bypassed the precautions put in place. After blinding Bob’s detectors, Eve measures Alice’s signals in whatever basis she chooses and sends
the result to Bob as a bright pulse. As a result of the blinding, Bob can only detect these pulses if his basis choice agrees with Eve’s, and so his raw key will be fully known to Eve. This attack does not at all contribute to the error rate used to detect an attack in common QKD attacks, and it is not yet clear how best to protect against this and the broader range of blinding attacks. One possible solution, briefly mentioned in [137], is to design protocols and implementations that can operate in both single-photon mode and bright pulse mode.

Finally, this class of attack was in fact realized in realistic conditions against the 290 meter QKD network in Singapore in 2012 [138]. Eve’s attacking apparatus, small enough to fit in a briefcase, was able to reliably control Bob’s measurement results while introducing a negligible amount of extra loss in the channel and only a small amount (well below the tolerance level) of noise. As a result, Eve was able to gain full information on the final secret key against two parties engaging in a QKD session without either party detecting her interference.

It is possible to close the side channel used in [138], and it is usually possible to close other side channels that could be utilized in similar attacks through various filters and countermeasures. However, alternatively, one may move to various device independent models, discussed earlier, which closes these side-channels by moving certain “weak points” (such as measurement devices) to the adversary itself.

IV. CURRENT COMMERCIAL SYSTEMS

At this point QKD technology has advanced to a point where several commercial companies exist producing a wide-range of quantum cryptographic devices. In this section, we survey several of the major efforts in this commercial area. This is done to provide the reader with a view as to the upcoming progress of actual QKD technology beyond academic experimental implementations and to show that this technology has the potential to seriously impact every-day communication. We make no claims as to the benefits of one company’s products over another nor do we survey their advantages and disadvantages. We merely survey what is currently possible from a commercial perspective. We also survey recent telecom investments in QKD technology and applications. Note that this is a rapidly changing area and this survey was conducted in early 2019.

A. ID Quantique

One of the largest commercial supplier of QKD solutions is the Swiss company ID Quantique. They provide a range of devices for single photon detection, quantum random number generation (QRNG), and
QKD designed to be easily integrated with current data center technology. We summarize the product lines offered using the information provided on their website [59] below.

1) Random Number Generation

True random number generation is vital to achieve security in a variety of systems, including QKD. For this reason random number generators based on properties of quantum mechanics, are currently being commercialized.

ID Quantique first introduced random number generators, based on quantum effects, in 2001. Currently they offer their “Quantis” line of products as their quantum random number generator solutions. These products make use of the inherently random nature of photons incident on a semi-transparent mirror being transmitted or reflected to generate a truly random stream of data. These devices are capable of generating between 4 and 16 MB/s of truly random bits.

In more detail, their devices consist of three primary “modules” [139]. The first is the “Optical Subsystem” consisting of a single photon source, a beam splitter, and two single-photon detectors. The photon source emits a single photon which passes through the beam splitter causing it to be detected at one of the two detectors. The timing and acquisition of these events is managed by the second module, the “Synchronization and Acquisition Subsystem” which simply times the photon emissions to ensure, among other things, that the detector dead-time is taken into account, and also collects data on which of the two detectors “clicked.”

Due to imperfections in sources and detectors as described in the previous section, multiple, or no, photons may be emitted; furthermore, the detector may click regardless of whether a photon hits it or not (i.e., a detector’s dark count rate and efficiency - see Section 2). Furthermore, no two detectors are manufactured the same and so they will each have their own particular properties. Because of this, a third module, the “Processing and Interfacing Subsystem,” is responsible for any post-processing necessary to rule out double-clicks and to handle biases in the outcomes. This module is also responsible for creating a pool of post-processed bits which may be interfaced to from outside applications to stream the generated, unbiased, random numbers.

Quantis is available both as a single unit, mountable on any PC board, or as a stand-alone PCI card. Furthermore, it has received certifications and validations from numerous agencies including NIST SP800-22 Test Suite Compliance. For more information, the reader is referred to the company’s white paper [139].
2) Quantum Key Distribution

ID Quantique distributes the “Cerberis QKD Blade” as their QKD solution [140]. This product is capable of distributing 20,000 keys (256-bit AES keys) per hour over a distance of 50km, and 2,000 keys per hour over a distance of 100km. Cerberis, currently in its third generation, is designed in an ATCA form factor and is thus can be used with any ATCA-compatible shelf. Different form-factors are also under development. The case is also designed to be tamper resistant and will zero-out any internal key material on detection of case tampering.

The company also manufactures the “Clavis” QKD system which is an open development platform for R&D applications along with educational and training purposes [141]. This system consists of two stations: a transmitter (Clavis-A) and a receiver unit (Clavis-B) each of which contains all necessary optical and electronic circuitry needed for QKD. The internal optical devices, however, must be controlled by an external computer interfaced through an Ethernet connection allowing for its reconfiguration for alternative QKD protocols.

B. Toshiba

The Japanese company, Toshiba, has also been investigating QKD technology. Rather than simply commercializing existing technology, Toshiba has been developing their own proprietary single photon detectors to be used with their own T12 [142] QKD protocol.

Toshiba recently announced a class of proprietary self-differencing (SD) [143] single photon detectors (SPDs) [144] which has been shown to achieve secret key rates equal to or greater than standard avalanche photodiodes when used for QKD over distances of 50 kilometers or less. These SD SPDs work by combining standard APDs with a SD circuit which processes the signal output by the APD to discern when signals are hidden in the noise and when the noise may result in a false positive by comparing it with the output of an identical APD shifted by some number of clock cycles, hence the term self-differencing.

Toshiba has also developed the T12 [142] protocol. This protocol is an extension of BB84 with biased basis choice and decoy states. In experimental results, using their SD SPDs in conjunction with the T12 protocol, Toshiba was able to achieve key rates of 2.20, 1.09, 0.40 and 0.12 Mbps at fiber lengths of 35, 50, 65, and 80 kilometers respectively with a QKD node operating at room temperature.

Toshiba’s QKD prototype is based on this technology, and advertises 1 Mbps key bit rates at 50km with functionality up to 100km. The system is contained in a standard 19” server rack mount. Importantly, as Toshiba notes, when used at distances less than 50km (and perhaps longer in conjunction with range
lengthening technology such as trusted node architectures), the key bit rates of 1 Mbps make it practical to use this prototype for applications such as video conferencing, or as part of a multiparty key distribution network.

C. Qubitekk

Qubitekk, is a company based in California, USA offering commercial QKD equipment [145]. Their primary system, designed primarily for the Industrial Control Systems community, is meant to be “set and forget” and can be used either in parallel to, or instead of, a traditional public key infrastructure. One unique property of their systems is it is based on quantum entanglement as opposed to standard prepare-and-measure systems (see Section 2). Their architecture requires three units: one “transmitter” and two “receivers.” The transmitter, which sits between the two receivers, is responsible for creating two entangled particle pairs, with one particle each being sent to a receiver. The receivers, controlled by the end-users $A$ and $B$, perform measurements on the received states yielding a key.

Communication on their system can be done using traditional optical fibers with key rates of 100kbps over 20km. Longer distances can be utilized, however with a drop in key-rate performance. Generally, keys should be used for AES256 encryption, however the application is outside the scope of the product. Keys are transferred through a serial interface using the Modbus protocol.

Beyond their primary commercial system, Qubitekk also produces R&D equipment. One product, the QKD Demonstrator, allows for experimental implementation of entanglement-based QKD protocols. As with their primary product, this system utilizes three boxes. The receivers support an HDMI output allowing users to visualize the key and error rate in real time. The system can be configured using a Raspberry Pi 3 but is limited to 1km at 1kbps.

Finally, Qubitekk also produces a “Hacker-box” which is meant as an add-on device for their QKD Demonstrator. This device simulates an adversary performing an intercept-resend attack in a single basis. To support this, the device consists of an optical switch; a polarizing beamsplitter; two single-photon counters; and one laser source capable of producing, with high probability, single photons.

Qubitekk, last year, has also created the “Quantum Grid” initiative in an effort to commercialize QKD technology developed for the US Department of Energy. This initiative asked for interested industry partners to participate in a cost-shared field trial program. However, while preliminary registration for interested participants has already expired (as of March 2018), it will be interesting to see what reports may come out of this.
D. Quantum XChange

Another US company entering the commercial QKD field is Quantum XChange. Based in Bethesda, MD, Quantum XChange (QX) offers to connect clients using QKD devices supplied by ID Quantique [146] and their own proprietary trusted node technology developed with Battelle [147]. This service was launched in June, 2018, but as of yet is not publicly available, instead being used by only a select group of early adopters.

More recently, QX, in partnership with ID Quantique has announced plans to use dark fiber purchased for the Zayo Group [148], along with their Trusted Node technology to launch a QKD network spanning from Boston to Washington D.C. This network, dubbed Phio [149], would offer QKD-as-a-service to clients across the eastern seaboard. At the time of writing, there is no information as to the specific capabilities (key-rates, error-rates, etc.) but this network would be the largest QKD network in North America and would allow, for example, QKD secured communication to be exchanged between traders on Wall Street and their home offices in New Jersey.

E. Joint Work: Fujitsu, NEC Corp

As mentioned previously in this document, one of the limiting factors of QKD is the efficiency of the single photon sources used to implement QKD protocols. In 2015 a joint effort with the University of Tokyo, Fujitsu, and NEC Corp created new single photon emitters which operate by exciting a quantum dot in a controlled manner causing it to emit photons for an infinitesimally small period of time. As the time period for which photons can be emitted is so small, this system reduces the probability of the source emitting multiple photons to that of one in a million, greatly increasing the efficacy of this source (recall, multiple photon emissions lead to leaked information to an adversary).

F. Telecom Investment in QKD

Many telecommunication companies, globally, have been showing interest in QKD, and other quantum related research. Here we summarize some of the major initiatives over the past decade.

United Kingdom: Beginning in 2016, British Telecom (BT) and Toshiba have collaborated to open a quantum communication showcase at a research and development lab owned by BT in Ipswich [150]. Though BT had been researching quantum communication for several years before this; Toshiba has also
been highly involved in quantum research as noted above. Both companies are currently collaborating to build a complete quantum communication network to connect Cambridge, Bristol, London, and Adastral Park.

Related to the above endeavor is the UK National Quantum Technologies Programme [151]. Begun in 2013, this program seeks to develop cross-collaboration among many different research and industry groups to develop marketable quantum technologies. While primarily consisting of academic groups, they are focused on deriving commercial systems and working with industry partners. They were also a lead partner with BT in developing a QKD link over a distance of 120km (running through several BT exchanges) in the Cambridge Metropolitan area [152].

**South Korea:** SK Telecom has been active in quantum related research and technology. As recently as February 2018, they invested US $65 million in ID Quantique for the development of quantum technologies relevant to telecom and also IoT markets. Before this, SK Telecom has been actively partnering with IDQ having invested $2 million in 2016 and in 2017 they developed the worlds smallest quantum random number generator using technology licensed from IDQ [153].

SK Telecom is also planning to use technology from ID Quantique in their future 5G network. Specifically, they plan to apply this technology, first, to the Seoul-Ansan section of its upcoming 5G network; they plan to then expand and apply this technology to their customer authentication server in early 2019 [154].

Having established their Quantum Technology Lab in 2011 specifically devoted to quantum technologies, SK Telecom has been actively involved in the research and development of quantum repeaters (needed to increase the range of QKD networks without the use of trusted nodes). They also applied quantum cryptography technology to their commercial LTE network in Sejong City [155], [156].

**Nippon Telegraph and Telephone Corporation (NTT):** NTT has been actively investing, since 2003 [157] in quantum related technology and research. Most recent, in 2015, their researchers proposed an “all photonic intercity QKD system.” This system requires $A$ and $B$ to send photons to a central server which will perform Bell measurements. The outcome of this measurement, announced by the server, allows for $A$ and $B$ to share a correlated bit, without the server actually knowing what that bit is. They have been active in other areas of protocol development also, including the development of a “Round-Robin
Differential Phase Shift” protocol which can operate without the need to constantly monitor error rates [158], [159].

In 2017, it was announced, in partnership with the National Institute of Informatics and the University of Tokyo, to construct a Quantum Neural Network (QNN) based computer. Their computer, using optical quantum computation, will also be made available to the public through a cloud-based interface. [160], [161] While not directly related to QKD, it shows their interest in general quantum technology.

**Deutsche Telekom:** Deutsche Telekom, partnering with South Korea Telecom, created the Quantum Alliance in 2017 to counter the threat posed by eventual development of quantum computers. Their initiative is currently recruiting partners, including network operators and equipment makers, to join their effort with a target to develop commercial products [162].

As of October 2018, they are also investing in ID Quantique (part of a joint agreement with SK Telecom) in order to strengthen the competitiveness and security in 5G services, the goal being to apply quantum cryptographic services, with hardware developed by ID Quantique, to future 5G operations [163].

Again, in partnership with SK Telecom, Deutsche Telekom is planning to incorporate quantum communication technology, consisting of QKD nodes and quantum random number generators, to a trial network. Their plan is to incorporate this technology for commercial communication within their network. [164], [165].

**V. CLOSING REMARKS**

Quantum key distribution is a highly promising field of research and development. From humble beginnings in the 1980’s, it is currently seeing massive interest in the research community with the development of more efficient protocols; the experimental community, with the demonstration of QKD over various communication modes, including fiber, over-the-air with mobile targets, and even satellites; and the commercial industry through the creation of several commercial QKD products in use today. It is also a technology that is highly cross-disciplinary with numerous advances made by fields not directly related to quantum physics. While still, perhaps, in the relative early stages, this technology continues to grow in prominence and, we suspect, will be highly visible in the very near future. There are still several fascinating problems that remain open, some of which were highlighted throughout this paper; of particular interest and importance are improving efficiency and distance along with combining alternative
technologies to adapt QKD to new operating conditions and applications. With this survey paper, by focusing on a high-level overview of this field, along with current practical implementations, we hope to expand interest of QKD to researchers and practitioners in many other fields so as to further speed up its growth and eventual wide-scale adoption.

ACKNOWLEDGMENTS

WK and OA would like to thank the Comcast Innovation Fund for their support. The authors would also like to thank Peter Quesada and Bahman Rashidi for reading early versions of this paper and providing very valuable feedback which greatly improved its quality.

REFERENCES


