Multi-Mediated Semi-Quantum Key Distribution

Walter O. Krawec
Computer Science & Engineering Department
University of Connecticut
Storrs, CT 06269
Email: walter.krawec@uconn.edu

***Abstract*—A semi-quantum key distribution (SQKD) protocol allows two users A and B to establish a shared secret key that is secure against an all-powerful adversary E even when one of the users (e.g., B) is semi-quantum or classical in nature while the other is fully-quantum. A mediated SQKD protocol allows two semi-quantum users to establish a key with the help of an adversarial quantum server. We introduce the concept of a multi-mediated SQKD protocol where two (or more) adversarial quantum servers are used. We construct a new protocol in this model and show how it can withstand high levels of quantum noise, though at a cost to efficiency. We perform an information theoretic security analysis and, along the way, prove a general security result applicable to arbitrary MM-SQKD protocols. Finally, a comparison is made to previous (S)QKD protocols.**

I. Introduction

Semi-quantum key distribution (SQKD), originally introduced in 2007 by Boyer et al., [1] allows for the establishment of a secret key between two parties A and B such that security is guaranteed against an all-powerful adversary E (something impossible to achieve through classical communication alone) and where one party, typically B, is severely restricted in his quantum capabilities and is, in a way, "classical" in nature. Studying SQKD protocols allows us to investigate the question "how quantum does a protocol need to be to gain an advantage over its classical counterpart?" Beyond theoretical interests, there may also be practical benefits to such systems. Indeed, one may envision a future quantum communication network utilizing fully-quantum devices; however if some device malfunctions, one may be able to switch to a "semi-quantum" mode of operation to continue secure communication. We stress, however, that in this paper we are only interested in the theoretical advantages to SQKD, namely, to study the "gap" between quantum and classical protocols.

Normally, SQKD protocols operate with A being a "fully-quantum user" and B being a "classical user." A two-way quantum channel allows A to send qubits, prepared in any state to the classical user B. This user, then, can either `Measure and Resend` - that is, take the qubit, measure it in the Z basis (spanned by $\{|0\rangle, |1\rangle\}$) and resend a Z basis qubit to A; or he can `Reflect` - that is, disconnect from the quantum channel and "bounce" the qubit back to A undisturbed. Note that the classical user can only directly operate in the Z basis or he can simply disconnect from the channel. Clearly if both A and B were restricted to these two operations, the entire protocol would be mathematically equivalent to classical communication (and hence not perfectly secure). When a qubit returns to the fully-quantum user A, she may measure in any basis of her choice.

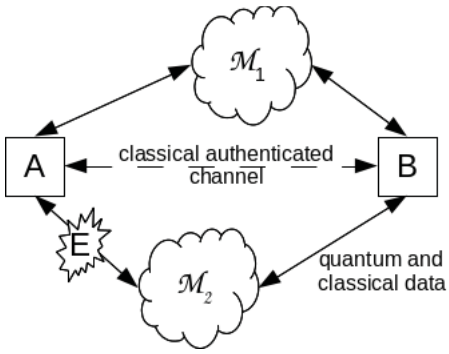

Fig. 1. Diagram showing parties involved. Two "classical" users A and B are connected to two fully-quantum servers $\mathcal{M}_1$ and $\mathcal{M}_2$. These servers may be adversarial, however honest mediators need not collaborate and can be, e.g., competing companies. E is a third-party adversary.

In [2], the idea of a *mediated SQKD* (M-SQKD) protocol was introduced where both A and B are "classical" in this sense, but, by utilizing the services of a fully-quantum server to prepare and measure in alternative bases, allowed A and B to establish a shared secret key. Security was proven in an information theoretic sense, even when this fully-quantum server was an all-powerful adversary. Since then, other mediated SQKD protocols have been proposed [3], [4] though without noise tolerance analyses.

In this paper, we revisit the idea of M-SQKD and introduce the idea of *multi-mediated SQKD* (MM-SQKD) where, as before, both A and B will be classical according to semi-quantum cryptographic definitions in [1], but they will now utilize the services of two (or more) adversarial quantum servers. Our goal will be to create a protocol that has a high tolerance to noise. Unlike M-SQKD, MM-SQKD protocols should take advantage of both servers to attain some quantum-level advantage (i.e., they should not simply use two servers in parallel). We achieve this goal in this paper. We make the assumption in this work that these adversarial servers may collaborate after the protocol is complete, but during the protocol's operation, they act independently of one another, except with regards to any classical messages sent (we formally define our attack model later in this work). This seems to be a safe assumption as attacking collaboratively during the quantum

978-1-7281-0960-2/19/$31.00 ©2019 IEEE

communication stage of our protocol would require them to somehow combine their quantum memories rapidly. If we assume that the two servers are spatially dislocated, this would be difficult or impossible to do without being detected (e.g., by monitoring the time between qubits sent by the server). Alternatively, the two servers could be competing companies and so would not be willing to collaborate to attack anyway. See Figure 1 for a diagram of the scenario we envision.

We make several contributions in this paper. First we introduce the idea of multi-mediated SQKD and design a new protocol for this scenario. Note that, as we wish to use both mediators to provide an advantage to noise tolerance, the trivial protocol of running two copies of a standard M-SQKD protocol will only result in an increase in overall efficiency but will not result in an increase in noise tolerance! Thus our protocol must take advantage of these two servers in a non-trivial manner. After presenting our protocol, we prove a very general security result, applicable to arbitrary MM-SQKD protocols using two *or more* servers thus providing an important theoretical result applicable to future work in this area (not just our specific protocol). Using this security result, we preform an information theoretic security analysis of our protocol using our attack model described earlier and discussed in more depth later. Finally, we evaluate our protocol's performance on a realistic channel scenario, comparing with prior work and commenting on future improvements that could be made. In particular we show our protocol's noise tolerance is as high as 18.7% (higher than the 13.04% tolerance from the M-SQKD protocol in [2], [5] and comparing very favorably with other (S)QKD protocols).

II. The Protocol

We present our protocol assuming both mediators are honest. Later, when we prove security, we will assume the mediators are adversarial. Our protocol operates as follows:

Quantum Communication Stage: The following process repeats until a sufficiently large raw-key has been established:

1. Mediator $\mathcal{M}_1$ and $\mathcal{M}_2$ each prepare, independently, Bell states of the form: $|\Phi^+\rangle = \frac{1}{\sqrt{2}}|0,0\rangle_{AB} + \frac{1}{\sqrt{2}}|1,1\rangle_{AB}$, sending the A particle to Alice and the B particle to Bob. Note that each mediator prepares their own two-qubit state. Thus, the joint state, assuming the mediators are honest, should be: $|\Phi^+\rangle \otimes |\Phi^+\rangle$. Both A and B receive two qubits each, one from each $\mathcal{M}_i$.
2. A (resp. B), on receiving a qubit from each mediator, will choose with probability p to `Measure and Resend` both qubits (resending back to the respective mediator from which it was received) or with probability $1-p$ to `Reflect` both qubits back to their respective mediators. Both parties choose independently. If A chooses `Measure and Resend`, she saves her measurement results in a bit-string $m_A \in \{0,1\}^2$; similarly, B will save his measurement results in m_B.
3. On receiving their qubits back from Alice and Bob, each mediator will, independently, perform a Bell measurement. If the outcome is $|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|0,0\rangle - |1,1\rangle)$, that mediator will send the classical message "$-$" while, for any other measurement result (the three other options), that mediator will send the message "$+$" Note that this communication need not be done in an authenticated manner (thus an adversary may tamper with this message).
4. A and B, on receiving a classical message from both mediators, will disclose to one-another, *using the authenticated classical channel connecting them*, the following information: (1) The messages they received from both mediators. This is done to ensure that a mediator (or an adversary) does not send different classical messages to each user (*if this is detected, users immediately abort*); (2) Their choice of `Measure and Resend` or `Reflect`; and (3) the parity of their measurement bits if applicable.
5. If both parties chose `Measure and Resend`, if both mediators send "$-$", and if both party's parity bit is 0, they will keep their measurement result to contribute a single raw-key bit. (Note that two qubits are needed for a single raw-key bit, thus we lose efficiency, though as we will see in our security analysis, we gain noise tolerance.) All other iterations, along with a randomly chosen sample of these raw-key bit iterations, will be used for parameter estimation to determine the channel noise.

Once a sufficient raw-key has been established, standard error correction and privacy amplification are run resulting in a secret key (see [6] for a description of these standard processes). The size of the secret key depends on the observed noise in the channel. Note that, as we are considering the asymptotic scenario, we may set p to be arbitrarily close to 1 to improve performance as was done in [2].

III. Security Analysis

We now turn to the security analysis of our protocol. We will assume the following security model: (1) All noise is the cause of the adversary (who we assume, in the worst case, are the servers). This is a standard assumption in QKD research [6]. (2) Both servers prepare states individually and later attack their returned qubits individually, resulting in a private quantum system for each server, however they can send secret classical information to each other during the protocol. We also assume they use collective attacks [6], that is each iteration is attacked identically and individually (possibly in a probabilistic manner), however they may keep their quantum memories and measure them at any future point in time coherently. (3) After the protocol has completed, however, the two servers may take their (unmeasured) quantum memory systems, resulting from their attacks, and collaboratively attempt to extract maximal information from both. (4) There may exist third-party adversaries (as depicted in Figure 1), however, in this preliminary work, a single adversary cannot attack the $\mathcal{M}_1$ and $\mathcal{M}_2$ regions simultaneously. This allows us to "absorb" any third-party attacks into one of the respective mediators. The more general case we leave as future work.

Collective attacks are a common assumption in QKD security proofs [6]; usually security against general attacks then follows [7], though we leave a complete proof of that as future work. In this case, we may use the Devetak-Winter

keyrate equation [8]. Let N be the size of the raw-key and $\ell(N)$ the size of the resulting secret key after privacy amplification. Then, the *key-rate* $\ell(N)/N$ as N approaches infinity is $\inf(S(A|E) - H(A|B))$ where the infimum is over all collective attacks which induce the observed channel statistics and where E is the quantum memory of the adversary (in our case $E = M_1M_2$) [8]. $S(A|E)$ is the conditional von Neumann entropy while $H(A|B)$ is the Shannon entropy. Computing $S(A|E)$ is the challenge in any (S)QKD security proof since $H(A|B)$ is a simple function of observed statistics.

A. State Preparation

We first prove a general result, applicable to any MM-SQKD protocol, concerning the form of the state that mediators may prepare. This result generalizes a result in [2] to the n-server case (prior work only considered $n = 1$).

Theorem 1. Consider an MM-SQKD protocol with classical A and B and with $n \geq 1$ (adversarial) quantum servers performing a (potentially collaborative) collective attack. If the protocol is such that the servers must prepare a quantum system, sending one qubit per server to A and B each, and if A and B either `Reflect` or `Measure and Resend` each qubit back to the originating server, then it is sufficient to analyze the case where the servers simply prepare the state:

$$|\psi_0\rangle = \sum_{i\in\{0,1\}^{2n}} \alpha_i |i\rangle, \tag{1}$$

where $\alpha_i \in \mathbb{R}_{\geq 0}$ and $\sum_i |\alpha_i|^2 = 1$. That is, there is no advantage to the servers in creating a state which is entangled with their private quantum memory at the initial stage of the protocol. Furthermore, if the n servers act independently, then it suffices to consider a state of the form: $|\psi_0\rangle = \bigotimes_{i=1}^{n} \left(\sum_{j,k\in\{0,1\}} \alpha_{j,k}^{(i)} |j,k\rangle\right)$, where $\alpha_{j,k}^{(i)} \in \mathbb{R}_{\geq 0}$ with the obvious normalization constraint.

Proof. Consider an arbitrary collective attack that the servers may perform. The state they prepare, then, may be written, without loss of generality, as: $|\phi_0\rangle = \sum_{i\in\{0,1\}^{2n}} \alpha_i |i\rangle \otimes |c_i\rangle$, where $|c_i\rangle$ are arbitrary normalized, but not necessarily orthogonal, states in some ancilla $\mathcal{H}_C$ owned jointly by the n adversarial servers. By absorbing any alternative phase into the vectors $|c_i\rangle$ we may, furthermore, assume each α_i is real and non-negative. Half the qubits are sent to A while the other half are sent to B.

Let $\Theta \in \{0,1\}^{2n}$ be A and B's choice of operation; namely if $\Theta_i = 1$ then whoever owns the i'th qubit performs the `Measure and Resend` operation on it; otherwise the `Reflect` operation is preformed. Using standard arguments (see [1] for the semi-quantum case), it is equivalent to consider the `Measure and Resend` operation as applying a CNOT gate to some local, private, ancilla register (that party may measure the ancilla later in the Z basis and this is equivalent to measuring the qubit immediately). Thus, in this case, following A and B's operation, the joint state is found to be: $|\phi_1\rangle = \sum_{i\in\{0,1\}^{2n}} \alpha_i |i\rangle_T \otimes |i \wedge \Theta\rangle_{AB} \otimes |c_i\rangle_C$, where $i \wedge \Theta$ is bit-wise "and" and where the AB register is held privately by A and B (separately depending on an irrelevant ordering of the qubits sent). At this point the T portion returns to the n servers who will then perform some further operations on it.

Consider, now, the restricted case shown in Equation 1. If the servers had sent this instead, the state, following A and B's operation would have been: $|\psi_1\rangle = \sum_{i\in\{0,1\}^{2n}} \alpha_i |i\rangle \otimes |i \wedge \Theta\rangle_{AB}$. Let V be the operator from $\mathcal{H}_T \to \mathcal{H}_T \otimes \mathcal{H}_C$ defined on basis states as $V|i\rangle = |i\rangle \otimes |c_i\rangle$. Clearly V is an isometry (and can thus be extended to a unitary operator) and is therefore something the servers could apply. Notice that $V|\psi_1\rangle = |\phi_1\rangle$. Thus, there is no advantage to sending the entangled initial state, instead they could prepare the simpler state shown in Equation 1, then when the qubits return, apply V before further attacking. The resulting states will be identical.

Of course, the above assumed the servers were collaborating. However if each server acts individually then the general state is actually of the form $|\phi_0\rangle = \bigotimes_{i=1}^{n}(\sum_{j,k} \alpha_{j,k}^{(i)} |j,k,c_{j,k}^{(i)}\rangle)$. In this case it is not difficult to see, by repeating the above steps, that the operator V may be written as $V = V_1 \otimes \cdots \otimes V_n$ where $V_i|j,k\rangle = |j,k,c_{j,k}^{(i)}\rangle$.

The same arguments can be used if some servers act alone while others collaborate thus completing the proof. □

Note that the above implies also that even a strategy where servers prepare unentangled states, it is sufficient to consider, instead, they send a state of the form in Equation 1. This is seen simply by first purifying the initial state and then applying the above theorem.

We add one further comment in the two-server case, namely that if A and B enforce a symmetry in their received states, (an assumption commonly made in most proofs of (S)QKD security), then the state the two servers send may be written:

$$|\psi_0\rangle = (\sqrt{c_0}|\Phi^+\rangle + \sqrt{c_1}|\Psi^+\rangle) \otimes (\sqrt{d_0}|\Phi^+\rangle + \sqrt{d_1}|\Psi^+\rangle). \tag{2}$$

where $|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$.

B. Modeling the Second Attack

When qubits return to each server, the server must perform some operation on it. Normally it should be a Bell measurement followed by a classical message reporting the outcome. However, an adversarial server should be allowed to perform any quantum operation on the returning qubits. Nonetheless, a single classical message should be sent to both parties. Furthermore, since we do not assume a broadcast channel, one server must act first - for simplicity we always assume this is $\mathcal{M}_1$, however our analysis is identical if it is $\mathcal{M}_2$; note that there would be no advantage to server's alternating turns randomly since they may collaborate at the end.

For single servers, it was shown in [2] that the return attack is modeled as a *quantum instrument* which, through standard techniques, may be dilated to an isometry U_1 mapping $\mathcal{H}_T$ to $\mathcal{H}_{cl} \otimes \mathcal{H}_{secret} \otimes \mathcal{H}_{M_1}$. Here $\mathcal{H}_T$ is the four-dimensional qubit space (modeling the two qubits returning to $\mathcal{M}_1$ from A and B); $\mathcal{H}_{cl}$ is a two dimensional Hilbert space spanned by

orthonormal basis $\{|\text{“}+''\rangle, |\text{“}-''\rangle\}$ which we use to denote the classical message sent to A and B (since the server cannot send different messages, as discussed, this two-dimensional space is sufficient); $\mathcal{H}_{secret}$ is a Hilbert space modeling the secret classical message sent to $\mathcal{M}_2$; finally $\mathcal{H}_{M_1}$ is $\mathcal{M}_1$'s quantum memory (note that the $\mathcal{H}_T$ portion may be absorbed into this space after U_1 is applied. The exact attack consists of $\mathcal{M}_1$ applying U_1, then measuring the cl space in the given basis. This determines the message actually sent and the post measurement state determines the state of the quantum memory in this event and the secret message to be sent. For greater details on how one may model a single server's attack using a quantum instrument and dilating to an isometry, see [2] which proved this for the single server case. Since our servers act individually (the secret message sent will not affect the proof), that result also applies here. The case for $\mathcal{M}_2$ is similar, though there U_2 is an isometry mapping $\mathcal{H}_T \otimes \mathcal{H}_{secret}$ to $\mathcal{H}_{cl_2} \otimes \mathcal{H}_{M_2}$. Note that the "secret" message sent becomes absorbed into $\mathcal{H}_{M_2}$. The subspace $\mathcal{H}_{cl_2}$ is spanned by the same basis as $\mathcal{H}_{cl}$ and is used to model $\mathcal{M}_2$'s message.

C. Key-Rate Computation

In light of our security assumption that both mediators are allowed to collaborate after the protocol's conclusion, our main goal in this section is to bound $S(A|M_1M_2)$, the entropy of A's raw-key register conditioned on both mediators' quantum memory systems; of course the entropy computation is only on iterations where a raw-key is distilled (not on iterations where parameter estimation is performed since these do not contribute to the final secret key). To do this, we must first construct a density operator modeling one iteration of the protocol, conditioning on that iteration being used to distill a raw-key.

In light of Theorem 1, and enforcing a symmetric attack, the joint state prepared by the mediators is shown in Equation 2. Note that, in our analysis, we will assume both systems are received by A and B simultaneously and, later, $\mathcal{M}_1$ and $\mathcal{M}_2$ receive their qubits back simultaneously. However this assumption is only made to ensure clarity in the presentation of our proof - if the systems are received at alternative times, since both mediators act independently and since A and B may postpone their classical "accept/reject" decisions until after both systems arrive and are processed, our security analysis holds in the more practical setting where this simultaneous arrival assumption is not made.

When A and B receive their qubits from both mediators, they will perform the `Measure and Resend` operation (again, we are conditioning on this iteration being used for raw-key distillation). Furthermore, we must condition on A (resp. B) receiving the same measurement outcome on their two qubits. Let's first consider A's measurement. If she measures and observes a $|0\rangle$ on both qubits, the joint state (before B's measurement) collapses to: $c_0d_0 |0000\rangle_{ABAB} + c_0d_1 |0001\rangle + c_1d_0 |0100\rangle + c_1d_1 |0101\rangle$. Alternatively, if she observes a $|1\rangle$ on both qubits: $c_0d_0 |1111\rangle_{ABAB} + c_0d_1 |1110\rangle + c_1d_0 |1011\rangle + c_1d_1 |1010\rangle$. Note that both of these are the normalized post-measurement states.

Next, B measures. Conditioned on A's raw key bit being 0, B will observe $|0\rangle$ in both his qubits with probability c_0d_0 and the post-measurement state will simply be $|0000\rangle_{ABAB}$. He will observe $|1\rangle$ in both qubits with probability c_1d_1 (again, conditioned on A's raw key-bit being 0) and the post-measurement state is simply $|0101\rangle$. Other cases are similar. Note that, since we are assuming all noise comes from the adversary's attack, if the state leaving A and B is, $|i,j,i,j\rangle$ then the state arriving at the two mediators will remain $|i,j,i,j\rangle$; noise will be modeled, to the adversary's advantage, by the attack operator.

Now, as discussed earlier, the attack operator of server $\mathcal{M}_1$ may be modeled as an isometry U_1. Without loss of generality, we may write the action of this operator as follows: $U_1 |a,b\rangle = |\text{“}+''\rangle |m_+\rangle |e^1_{a,b}\rangle + |\text{“}-''\rangle |m_-\rangle |f^1_{a,b}\rangle$. where $|m_\pm\rangle$ are states in $\mathcal{H}_{secret}$ which is the secret message sent by $\mathcal{M}_1$ to $\mathcal{M}_2$. The action of U_2, we write as follows: $U_2 |m_{cl}, m_{secret}, a, b\rangle = |\text{“}+''\rangle |g_{m_{cl},m_{secret},a,b}\rangle + |\text{“}-''\rangle |h_{m_{cl},m_{secret},a,b}\rangle$.

Each server applies their U_i operator, performs a measurement of the "cl" register, and sends the result to A and B (with $\mathcal{M}_1$ acting first). Conditioning on both servers sending "$-''$ and both A and B accepting, the final density operator describing the joint $A, B, \mathcal{M}_1$, and $\mathcal{M}_2$ system, denoted $\rho_{ABM_1M_2}$ is (note the "cl" space is no longer needed as, in this case, it is always "$-''$):

$$\frac{c_0d_0}{N}[\mathbf{00}]_{AB} \otimes [\mathbf{f^1_{0,0}}, \mathbf{f^2_{0,0}}] + \frac{c_0d_0}{N}[\mathbf{11}]_{AB} \otimes [\mathbf{f^1_{1,1}}, \mathbf{f^2_{1,1}}] \quad (3)$$
$$+\frac{c_1d_1}{N}[\mathbf{01}]_{AB} \otimes [\mathbf{f^1_{0,1}}, \mathbf{f^2_{0,1}}] + \frac{c_1d_1}{N}[\mathbf{10}]_{AB} \otimes [\mathbf{f^1_{1,0}}, \mathbf{f^2_{1,0}}],$$

where $[\mathbf{x}, \mathbf{y}] = |x\rangle\langle x| \otimes |y\rangle\langle y|$; $|f^2_{a,b}\rangle = |h_{\text{“}-'',m_-,a,b}\rangle$; and N is the normalization term:

$$N = c_0d_0(\langle f^1_{0,0}|f^1_{0,0}\rangle \langle f^2_{0,0}|f^2_{0,0}\rangle + \langle f^1_{1,1}|f^1_{1,1}\rangle \langle f^2_{1,1}|f^2_{1,1}\rangle)$$
$$+ c_1d_1(\langle f^1_{0,1}|f^1_{0,1}\rangle \langle f^2_{0,1}|f^2_{0,1}\rangle + \langle f^1_{1,0}|f^1_{1,0}\rangle \langle f^2_{1,0}|f^2_{1,0}\rangle)$$

At this point, we may use a theorem from [9] to compute $S(A|M_1M_2)$. This theorem is for general classical-quantum states and gives a worst-case bound. To simplify notation, let $F_{a,b} = \langle f^1_{a,b}|f^1_{a,b}\rangle \langle f^2_{a,b}|f^2_{a,b}\rangle$. Then we have the following lower-bound on $S(A|M_1M_2)$:

$$\frac{c_0d_0(F_{0,0}+F_{1,1})}{N}\left(H\left[\frac{F_{0,0}}{F_{0,0}+F_{1,1}}\right] - H[\lambda_1]\right) \quad (4)$$
$$+\frac{c_1d_1(F_{0,1}+F_{1,0})}{N}\left(H\left[\frac{F_{0,1}}{F_{0,1}+F_{1,0}}\right] - H[\lambda_2]\right),$$

where:

$$\lambda_1 = \frac{1}{2} + \frac{\sqrt{(F_{0,0}-F_{1,1})^2 + 4|\langle f^1_{0,0}|f^1_{1,1}\rangle|^2 |\langle f^2_{0,0}|f^2_{1,1}\rangle|^2}}{2(F_{0,0}+F_{1,1})}$$

$$\lambda_2 = \frac{1}{2} + \frac{\sqrt{(F_{0,1}-F_{1,0})^2 + 4|\langle f^1_{0,1}|f^1_{1,0}\rangle|^2 |\langle f^2_{0,1}|f^2_{1,0}\rangle|^2}}{2(F_{0,1}+F_{1,0})}$$

Note that we are using a stronger version of the Theorem in [9] derived in their proof. Also note that we only need to bound the

real part of the above inner-products since: $|\langle x,y|z,w\rangle|^2 = |\langle x|z\rangle|^2|\langle y|w\rangle|^2 \geq Re^2|\langle x|z\rangle| \cdot Re^2|\langle y|w\rangle|$ and the closer to $1/2$, λ_i is, the smaller $S(A|E)$ (giving a worst-case bound).

Parameter Estimation: To evaluate our key-rate expression, we must determine which of the many parameters appearing in the above equations, can be observed, or at least bounded based on observations. Clearly c_i and d_i are observable parameters. Indeed, c_0 is the probability that A and B's measurement results agree on the state sent by $\mathcal{M}_1$. Similarly for d_0 for the state sent by $\mathcal{M}_2$. Finally, it is not difficult to show that $\langle f^1_{a,b}|f^1_{a,b}\rangle$ is the probability that $\mathcal{M}_1$ announces "$-$" given that both A and B observed $|a\rangle$ and $|b\rangle$ respectively on the system received by $\mathcal{M}_1$. Similarly for $\langle f^2_{a,b}|f^2_{a,b}\rangle$ (*however, here, A and B must condition on $\mathcal{M}_1$ sending the message "$-$" to determine this statistic*) thus allowing A and B to observe $F_{a,b}$ needed above. The only remaining quantities are $Re\langle f^i_{0,0}|f^i_{1,1}\rangle$ and $Re\langle f^i_{0,1}|f^i_{1,0}\rangle$ which we will bound by considering the case when A and B both reflect.

Let us consider $\mathcal{M}_1$, the case for $\mathcal{M}_2$ will be similar. If A and B choose `Reflect`, then the state arriving at $\mathcal{M}_1$ is simply $\sqrt{c_0}|\Phi^+\rangle + \sqrt{c_1}|\Psi^-\rangle$ (any noise in the channel is modeled, to the adversary's advantage, using $\mathcal{M}_1$'s attack operator). In this case, after applying U_1, the state evolves to $|\text{"}-\text{"}\rangle \otimes |g_0\rangle + |\text{"}+\text{"}\rangle \otimes |h_0\rangle$, where $|h_0\rangle$ is irrelevant and: $|g_0\rangle = \left(\frac{\sqrt{c_0}}{\sqrt{2}}\left(|f^1_{0,0}\rangle + |f^1_{1,1}\rangle\right) + \frac{\sqrt{c_1}}{\sqrt{2}}\left(|f^1_{0,1}\rangle + |f^1_{1,0}\rangle\right)\right)$. Thus the probability that $\mathcal{M}_1$ sends the message $|\text{"}-\text{"}\rangle$, in the event A and B both `Reflect`, which we denote p_1^{err}, is simply $\langle g_0|g_0\rangle$. Expanding this and solving for the two inner-products appearing in λ_1 and λ_2, we find:

$$\begin{aligned} & c_0 Re\langle f^1_{0,0}|f^1_{1,1}\rangle + c_1 Re\langle f^1_{0,1}|f^1_{1,0}\rangle \qquad (5)\\ & = p_1^{err} - \frac{c_0}{2}(\langle f^1_{0,0}|f^1_{0,0}\rangle + \langle f^1_{1,1}|f^1_{1,1}\rangle)\\ & - \frac{c_1}{2}(\langle f^1_{0,1}|f^1_{0,1}\rangle + \langle f^1_{1,0}|f^1_{1,0}\rangle)\\ & - \sqrt{c_0c_1}Re(\langle f^1_{0,0}|f^1_{0,1}\rangle + \langle f^1_{0,0}|f^1_{1,0}\rangle)\\ & - \sqrt{c_0c_1}Re(\langle f^1_{1,1}|f^1_{0,1}\rangle + \langle f^1_{1,1}|f^1_{1,0}\rangle). \end{aligned}$$

We show how the technique of mismatched measurements [10], [11] can be applied here to directly observe those inner products appearing on the right-hand-side of the above equation.

Denote by $p^1_{0,R}$ to be the probability that $\mathcal{M}_1$ sends the message "$-$", conditioned on the event that B chooses `Reflect` (the "R") and that A chooses `Measure and Resend` *and* actually observes outcome $|0\rangle$ on the state arriving from $\mathcal{M}_1$. Tracing the evolution of the system after $\mathcal{M}_1$'s attack operator, conditioning on these events, we easily find the state, as it arrives to $\mathcal{M}_1$, to be: $\sqrt{c_0}|00\rangle + \sqrt{c_1}|01\rangle$, and so, after applying operator U_1, the probability $p^1_{0,R}$ is found to be: $p^1_{0,R} = c_0\langle f^1_{0,0}|f^1_{0,0}\rangle + c_1\langle f^1_{0,1}|f^1_{0,1}\rangle + 2\sqrt{c_0c_1}Re\langle f^1_{0,0}|f^1_{0,1}\rangle$ Generalizing to the other cases, we find for $x = 0, 1$ and $y = 1 - x$:

$$Re\langle f^1_{x,x}|f^1_{x,y}\rangle = \frac{p^1_{x,R} - c_0\langle f^1_{x,x}|f^1_{x,x}\rangle - c_1\langle f^1_{x,y}|f^1_{x,y}\rangle}{2\sqrt{c_0c_1}} \qquad (6)$$

$$Re\langle f^1_{x,x}|f^1_{y,x}\rangle = \frac{p^1_{R,x} - c_0\langle f^1_{x,x}|f^1_{x,x}\rangle - c_1\langle f^1_{y,x}|f^1_{y,x}\rangle}{2\sqrt{c_0c_1}} \qquad (7)$$

(Note that if $c_0 = 0$ or $c_1 = 0$, then these terms never appear in $\langle g_0|g_0\rangle$ and so we do not need to resort to mismatched measurements here.) The various $p^1_{i,j}$ values are defined analogously to $p^1_{0,R}$. Thus, Equation 5 has only two unknowns, only one of which is free. Finally, the above process may be repeated for $\mathcal{M}_2$. Thus, to compute our lower-bound on $S(A|M_1M_2)$, we minimize Equation 4 over two free parameters: $Re|\langle f^i_{0,1}|f^i_{1,0}\rangle|$ upper-bounded by $\sqrt{\langle f^i_{0,1}|f^i_{0,1}\rangle\langle f^i_{1,0}|f^i_{1,0}\rangle}$ (due to the Cauchy-Schwarz inequality). We take the minimum over all such parameters as we must assume the worst case in that each mediator chose an optimal attack. Computing $H(A|B)$ is trivial as we explain in the next section.

D. Evaluation

Our above derivation is valid for any attack in the considered security model. Based on observable parameters, A and B may evaluate a lower-bound on $S(A|M_1M_2)$ and then compute $H(A|B)$, thus giving a value for the key-rate r. To evaluate our bound, however, and to compare with other protocols, we must put numbers to these observable parameters. While in practice this would be done through actual measurements, we will, for this work, assume a standard depolarization channel and use this to determine these values. We will also assume that the adversarial servers act in a way so as to simulate an honest server in that all statistics should conform to a Bell measurement as prescribed by the protocol, even if the servers are doing something else malicious. *These are all enforceable conditions.*

First, we need c_i and d_i. However, these are simply the Z-basis noise in the channel connecting A and B to each server. We assume a depolarization channel: $\mathcal{E}_q(\rho) = (1-2q)\rho + \frac{q}{2}I$. We use $2q$ to remain consistent with work in [5] so as to immediately compare; it also makes sense to use $2q$ for a reason which will be clear momentarily. Note that I, above, is the 4×4 identity operator as we are working with two qubits. We will set $q = Q_1$ for the $\mathcal{M}_1$ channel and $q = Q_2$ for the $\mathcal{M}_2$ channel. Under this channel assumption, we have $c_0 = 1 - Q_1$, $c_1 = Q_1$, $d_0 = 1 - Q_2$, and $d_1 = Q_2$.

Next, we determine $\langle f^1_{a,b}|f^1_{a,b}\rangle$ which, as described earlier, is the probability that $\mathcal{M}_1$ sends the message "$-$" conditioned on A and B observing $|a\rangle$ and $|b\rangle$ respectively, from the system received from the first server. If we assume a depolarization channel, then the state actually arriving to $\mathcal{M}_1$, conditioned on this event, is: $\mathcal{E}_{Q_1}([\mathbf{a},\mathbf{b}]) = (1-2Q_1)[\mathbf{a},\mathbf{b}] + \frac{Q_1}{2}I$. An honest server would then perform a Bell measurement and send "$-$" only on receiving outcome $|\Phi^-\rangle$. Thus $\langle f^1_{a,a}|f^1_{a,a}\rangle = (1 - Q_1)/2$ and $\langle f^1_{a,1-a}|f^1_{a,1-a}\rangle = Q_1/2$. Of course an adversarial server could replace the noisy quantum channel with a perfect one, and perform any operation, however in a reasonable setting, A and B could expect, *and enforce*, that whatever attack is done, the observable statistics conform to this derivation. For $\mathcal{M}_2$, the process is identical and, in our

noise scenario, the value is the same but parameterized with Q_2.

Next, we need $p^1_{0,R}$ which is the probability that $\mathcal{M}_1$ sends "$-$" conditioned on B choosing `Reflect` and A observing $|0\rangle$. To determine this in our noise model, we trace the evolution of the system. On arriving at A and B, the two qubits are in a state $(1-2Q_1)[\mathbf{\Phi}^+]+\frac{Q_1}{2}I$. Now, conditioning on B choosing `Reflect` and A choosing `Measure and Resend` and actually observing $|0\rangle$, the state collapses to $(1-Q_1)[\mathbf{00}]+Q_1[\mathbf{01}]$. This system then returns to $\mathcal{M}_1$ passing through the depolarization channel again; thus, the state arriving at the server is: $(1-2Q_1)((1-Q_1)[\mathbf{00}]+Q_1[\mathbf{01}])+\frac{Q_1}{2}I$. If the server were honest, it would perform a Bell measurement on this; the probability of observing $|\Phi^-\rangle$ is then: $p^1_{0,R} = (1-2Q_1)(1-Q_1)/2+Q_1/2 = 1/2-Q_1(1-Q_1)$. Note that we do not need to assume this behavior on the servers, but users can enforce these observations. We find similar values for $p^1_{1,R}, p^1_{R,0}$ and $p^1_{R,1}$; similar values are found for $\mathcal{M}_2$.

Substituting this into Equations 6 and 7, we find those eight (four for $\mathcal{M}_1$ and four for $\mathcal{M}_2$) are all zero in this case. This simplifies Equation 5 to: $(1-Q_i)Re\langle f^i_{0,0}|f^i_{1,1}\rangle + Q_iRe\langle f^i_{0,1}|f^i_{1,0}\rangle = p^{err}_i - \frac{(1-Q_i)^2}{2} - \frac{Q_i^2}{2}$. One may bound $|\langle f^i_{0,1}|f^i_{1,0}\rangle| \leq Q_1/2$ using the Cauchy-Schwarz inequality. Thus, to evaluate $S(A|M_1M_2)$, we minimize Equation 4 over two free parameters: $Re\langle f^i_{0,1}|f^i_{1,0}\rangle$ subject to the constraint that they are within the range $[-Q_i/2, Q_i/2]$. The values of $Re\langle f^i_{0,0}|f^i_{1,1}\rangle$ are then determined by the above equation.

The only thing remaining to compute the actual key-rate is $H(A|B)$. Let $p^{\text{key}}_{i,j}$ be the probability that A's raw-key bit is i and B's raw-key bit is j. From Equation 3, we have: $p^{\text{key}}_{0,0} = p^{\text{key}}_{1,1} = \frac{(1-Q_1)^2(1-Q_2)^2}{4N}$, and $p^{\text{key}}_{0,1} = p^{\text{key}}_{1,0} = \frac{Q_1^2Q_2^2}{4N}$, where: $N = \frac{(1-Q_1)^2(1-Q_2)^2+Q_1^2Q_2^2}{2}$. And so: $H(A|B) = H(p^{\text{key}}_{0,0}, \cdots, p^{\text{key}}_{1,1}) - H(p^{\text{key}}_{0,0}+p^{\text{key}}_{1,0})$.

Evaluating our key-rate under this noise scenario reveals a maximal noise tolerance of 18.7%. This is higher than the 13.04% allowed by the original M-SQKD protocol [5] (under the same channel assumptions). It is also higher than the noise tolerance of the original SQKD protocol which can tolerate up to 11% noise [9]. Finally it is higher than many fully-quantum protocols (though certain higher-dimensional QKD protocols can tolerate up to 50% noise [12]).

What is lost, however, is efficiency. Consider *effective key-rate* which multiplies the computed key-rate by the probability of any particular iteration actually yielding a raw-key bit. We compute our effective key-rate and compare with the effective key-rate of the M-SQKD protocol from [5] in Figure 2. We note that, for low noise levels, the M-SQKD protocol is more efficient, however it cannot tolerate high levels of noise. One may consider more complex strategies in practice, switching from M-SQKD to MM-SQKD depending on the noise level.

IV. Closing Remarks

We presented a new model of semi-quantum cryptography, namely the *multi mediator* model and designed a new protocol in this scenario. Along the way we proved a general security result applicable to arbitrary MM-SQKD protocols which will help future researchers. Many interesting open problems remain - of particular interest would be designing a more efficient protocol in this model. Also, developing strategies to swap between M and MM modes of operation based on (perhaps changing) noise levels would also be very interesting.

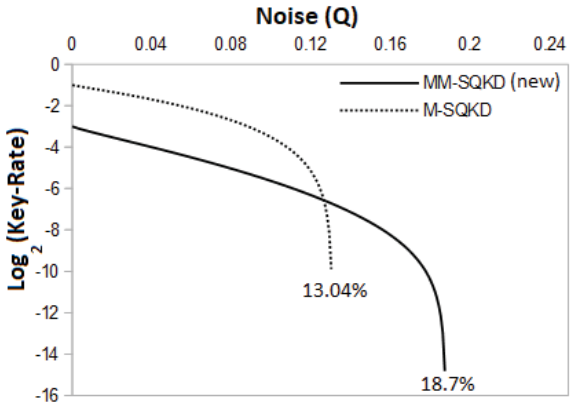

Fig. 2. Comparing the *effective* key rates of our new MM-SQKD protocol with the M-SQKD protocol in [5]. The M-SQKD protocol is more efficient but it cannot tolerate high noise levels.

Acknowledgments: WK is partially supported by the NSF under grant number 1812070.

References

[1] M. Boyer, D. Kenigsberg, and T. Mor, "Quantum key distribution with classical bob," *Phys. Rev. Lett.*, vol. 99, p. 140501, Oct 2007.

[2] W. O. Krawec, "Mediated semiquantum key distribution," *Physical Review A*, vol. 91, no. 3, p. 032323, 2015.

[3] Z.-R. Liu and T. Hwang, "Mediated semi-quantum key distribution without invoking quantum measurement," *Annalen der Physik*, vol. 530, no. 4, p. 1700206, 2018.

[4] P.-H. Lin, C.-W. Tsai, and T. Hwang, "Mediated semi-quantum key distribution using single photons," *Annalen der Physik*, p. 1800347, 2019.

[5] W. O. Krawec, "An improved asymptotic key rate bound for a mediated semi-quantum key distribution protocol," *Quantum Information and Computation*, vol. 16, no. 9 & 10, pp. 0813–0834, 2016.

[6] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, pp. 1301–1350, Sep 2009.

[7] M. Christandl, R. Konig, and R. Renner, "Postselection technique for quantum channels with applications to quantum cryptography," *Phys. Rev. Lett.*, vol. 102, p. 020504, Jan 2009.

[8] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, vol. 461, no. 2053, pp. 207–235, 2005.

[9] W. O. Krawec, "Quantum key distribution with mismatched measurements over arbitrary channels," *Quantum Information and Computation*, vol. 17, no. 3 and 4, pp. 209–241, 2017.

[10] S. M. Barnett, B. Huttner, and S. J. Phoenix, "Eavesdropping strategies and rejected-data protocols in quantum cryptography," *Journal of Modern Optics*, vol. 40, no. 12, pp. 2501–2513, 1993.

[11] S. Watanabe, R. Matsumoto, and T. Uyematsu, "Tomography increases key rates of quantum-key-distribution protocols," *Physical Review A*, vol. 78, no. 4, p. 042316, 2008.

[12] H. Chau, "Quantum key distribution using qudits that each encode one bit of raw key," *Phy. Rev. A*, vol. 92, no. 6, p. 062324, 2015.