Physical-Layer Cryptography Through Massive MIMO

Thomas R. Dean, Member, IEEE, and Andrea J. Goldsmith, Fellow, IEEE

Abstract—We propose the new technique of physical-layer cryptography based on using a massive MIMO channel as a key between the sender and desired receiver, which need not be secret. The goal is for low-complexity encoding and decoding by the desired transmitter-receiver pair, whereas decoding by an eavesdropper is hard in terms of prohibitive complexity. The decoding complexity is analyzed by mapping the massive MIMO system to a lattice. We show that the eavesdropper’s decoder for the MIMO system with M-PAM modulation is equivalent to solving standard lattice problems that are conjectured to be of exponential complexity for both classical and quantum computers. Hence, under the widely-held conjecture that standard lattice problems are hard to solve, the proposed encryption scheme has a more robust notion of security than that of the most common encryption methods used today such as RSA and Diffie-Hellman. Additionally, we show that this scheme could be used to securely communicate without a pre-shared secret and little computational overhead. Thus, by exploiting the physical layer properties of the radio channel, the massive MIMO system provides for low-complexity encryption commensurate with the most sophisticated forms of application-layer encryption that are currently known.

Index Terms—Cryptography, Lattices, MIMO, Quantum Computing

I. INTRODUCTION

The decoding of massive MIMO systems forms a complex computational problem. In this paper, we exploit this complexity to form a notion of physical-layer cryptography. The premise of physical-layer cryptography is to allow the transmission of confidential messages over a wireless channel in the presence of an eavesdropper. We present a model where a given transmitter-receiver pair is able to efficiently encode and decode messages, but an eavesdropper who has a physically different channel must perform an exponential number of operations in order to decode. This allows for confidential messages to be exchanged without a shared key or key agreement scheme. Rather, the encryption exploits physical properties of the massive MIMO channel.

Our MIMO wiretap channel model for communication is shown in Figure 1. Here, a parallel channel decomposition allows for two users, Alice and Bob, to communicate with an overhead of only performing linear precoding and receiver shaping of their MIMO channel, assumed known to both of them. To an eavesdropper, Eve, who has a different channel, this decomposition does not aid in the ability to decode the channel with linear complexity. In particular, we prove that it is exponentially hard for the eavesdropper to decode Alice’s transmitted vector in our system model even if it knows the channel between Alice and Bob. We refer to the channel encryption key as a Channel State Information- or CSI-key. The model requires both the transmitter and receiver to have perfect knowledge of the channel, but this knowledge does not need to be kept secret. For decoding by Eve to be hard, our model requires a maximum on the SNR that Eve maintains and that Alice and Bob use a large constellation size, where the required constellation size is related to the number of transmit antennas.

To characterize the hardness of decoding MIMO systems, we use the method of reductions found in the computational study of cryptography. A brief description of this method will be provided below, while a more detailed overview of this approach to cryptography can be found in [1]. In the method of reductions, we suppose that we have access to an oracle, i.e., a black box which, given a channel and a received vector, returns the proper transmitted vector in a single operation. If the existence of such an oracle would imply efficient solutions (i.e., ones whose runtime is bounded to within a polynomial factor of the length of the input) to problems which are known (or conjectured) to be hard (i.e., ones whose runtime is bounded to within an exponential factor of the length of the input), then it follows that such an oracle cannot (or is conjectured to not) exist.

We show in this work that under proper conditions, the complexity of decoding MIMO systems by an eavesdropper can be related to solving standard lattice problems. The con-
connection between MIMO and lattices is not new: for example, see Damen et al. \[2\], where the maximum-likelihood decoder is related to solving the Closest Vector Problem. Problems on lattices have been widely studied in cryptography and other fields and many standard lattice problems are conjectured to be hard ([3]–[13]). Lattice problems are also conjectured to be hard even when solved using quantum computers when they exist ([13], [13]). Creating efficient cryptosystems that achieve security given the presence of quantum computers is currently an active area of research, since most cryptosystems today, such as RSA (a public-key algorithm named after its inventors, Rivest, Shamir, and Adleman, [14]) or the Diffie-Hellman key exchange (see [14]), could efficiently be broken by a quantum computer ([16]–[18]). Our physical-layer cryptosystem provides quantum-resistant cryptography by exploiting the hardness associated with the eavesdropper’s decoding in a massive MIMO system.

The idea of exploiting properties of the physical layer to achieve secrecy is not new and dates back to Shannon’s notion of information theoretic secrecy [19] and Wyner’s wiretap channel [20] (for a survey on the subject see [21]). In information theoretic secrecy, the goal is to communicate in a manner such that legitimate users may communicate at a positive rate, while the mutual information between the eavesdropper and the sender is negligibly small. Note that since in our model Bob and Eve have statistically-identical channels, information theoretic secrecy is not possible without a key rate, coding, or using properties of Alice’s and/or Bob’s channel in the transmission strategy. Along these lines, Shimizu et al. [22], have suggested using properties of the radio propagation channel to achieve information theoretic secrecy. These notions all differ from ours as we consider encryption based on computational complexity at the eavesdropper’s decoder rather than through equivocation at the eavesdropper related to entropy and mutual information. We believe this work makes significant progress in addressing some of the challenges that have been identified in applying physical layer security to existing and future systems [23].

The remainder of this paper is organized as follows. Section II outlines our system model and the underlying assumptions upon which the security of our system is based. In Section III we discuss lattices, lattice problems, and lattice-based cryptography in order to provide the background for our main result, which is stated in Section IV and proved in Appendix A. In Section V we discuss additional notions of security for our model, including how to achieve security under adversarial models commonly considered by cryptographers.

II. Problem Formulation

A. The Wiretap Model

Consider an \( n \times m \) real-valued MIMO system consisting of \( n \) transmit antennas and \( m \) receive antennas:

\[
y = Ax + e,
\]

(1)

where \( x \in \mathbb{R}^n \), and \( A \in \mathbb{R}^{n \times m} \) is the channel gain matrix. Each entry of the channel gain matrix is drawn i.i.d. from the Gaussian distribution with zero mean and standard deviation \( k/\sqrt{2}\pi \). This distribution is henceforth written \( \Psi_k \). The vector \( e \in \mathbb{R}^m \) is the channel noise with each entry i.i.d. \( \Psi_a \).

It is assumed that \( A \) is known to both the transmitter and receiver. If we constrain the vector \( x \) to use a discrete, periodic constellation, then the set of received points becomes analogous to points on a lattice, perturbed by a Gaussian random variable.

We assume that there are an arbitrary number of receive antennas, restricted to an amount within a polynomial factor of the number of transmit antennas, \( n \). By making this assumption, we are considering the advantage an eavesdropper would gain by having an arbitrarily large number of receive antennas, but we are assuming that building a receiver with an exponential number of antennas relative to those at the legitimate user’s transmitter would be prohibitively expensive. Assuming that certain requirements on SNR and constellation size are met, as described below, the security of the system can be quantified solely by the number of transmit antennas. This number plays the role of the security parameter commonly used by cryptographers in designing cryptosystems. Essentially, we are saying that decoding the system requires a number of operations that is exponential in \( n \) and we parameterize the remaining variables in our system based on \( n \).

We consider real systems with the transmitted signal constellation, \( \mathcal{X} \), defined as the set of integers \([0, M]\). Lattices can easily be scaled and shifted, so we use this constellation without loss of generality over all possible M-PAM constellations. Our analysis assumes channels with a zero-mean gain, but this is for simplicity of exposition as the proof given in Appendix A can easily be generalized to systems with non-zero mean. We consider uncoded systems, but our results can be extended to coded systems as we discuss in Section V.

Let the vector \( a_i \in \mathbb{R}^n \) denote the gains between the transmitter and the \( i \)th receive antenna, let \( e_i \) denote the noise sample at this antenna, and let \( x \) represent the transmitted vector which is drawn from \( \mathcal{X}^n \). The \( i \)th receive antenna gets a noisy, random inner-product of the form

\[
y_i = \langle a_i, x \rangle + e_i.
\]

(2)

Our work can easily be extended to complex channels, where the real and imaginary portions of the channel gain matrix and noise are drawn independently, and M-QAM symbols are sent, using an equivalent constellation as in the real case.

Our model considers static channels. However, our model can easily be extended to time-varying channels: consider a single message is transmitted \( t \) times, no more than once per channel coherence period. Each transmission is equivalent to Eve receiving the original message with \( t \) times as many receive antennas which does nothing to aid in the computational complexity of her decoding.

Let user \( A \) have \( n \) transmit antennas which are used to send a message to user \( B \) who has a number of receive antennas that is within a polynomial factor of \( n \). Let the channel between \( A \) and \( B \) be represented by \( A \), where each entry is i.i.d. Gaussian \( \Psi_k \). The noise at each receive antenna is drawn from \( \Psi_{Ma} \).
The results in this paper show that MIMO decoding can be related to solving standard lattice problems when a certain minimum noise level and constellation size are met. If the noise power is below the required level, efficient decoding methods such as the zero-forcing decoder could be applied to our system. In other words, if these conditions are not met, then our results provide no insight on the complexity of decoding, and hence on the security of the MIMO wiretap channel. Specifically, for some arbitrary \( m > 0 \), we require the following constraints on the transmission from user \( A \) to user \( B \):

\[
\text{Minimum Noise: } m\alpha/k^2 > \sqrt{n} \tag{3}
\]

\[
\text{Constellation Size: } M > m2^{n\log\log n/\log n} \tag{4}
\]

where the parameter \( m \) may be chosen by a user or system designer in order to take off the SNR requirement for the size of the constellation.

Now consider an eavesdropper, \( \text{EVE} \), which has \( \text{poly}(n) \) receive antennas, and receives message \( x \) with channel represented by \( B \), where each entry is again i.i.d. \( \Psi_k \). Let the channel have noise be drawn from \( \Psi_{M,2} \), where \( \beta \geq \alpha \). In other words, the eavesdropper must meet at least the minimum noise requirement stated above. Discussion on how to ensure this requirement is met is also provided in Section V.

In order to send message \( x \) to user \( B \), user \( A \) performs a linear precoding as described in [25]. Let the singular value decomposition of \( A \) be given as \( A = UV^H \). \( A \) now sends \( \tilde{x} = Vx \). Upon receiving a transmission from user \( A \), user \( B \) computes \( \tilde{y} = U^H\tilde{x} \). It is easy to show that this expands to \( \tilde{y} = \Sigma x + \tilde{e} \). Since \( \Sigma \), representing the singular values of \( A \), is a diagonal matrix, \( B \) can efficiently estimate \( x \) with linear complexity in \( n \). Notice that \( U \) is unitary so \( \|e\| = \|\tilde{e}\| \).

Now consider the message received by \( \text{EVE} \):

\[
\tilde{y} = BVx + \tilde{e}. \tag{5}
\]

Note that \( V \) consists of the right singular vectors of \( A \), which is independent of \( B \) and unitary. Gaussian random matrices are orthogonally invariant [26], so since \( V \) is unitary, multiplying by \( V \) returns the matrix to an identical, independent distribution. In other words, the entries of \( BV \) are i.i.d., following the same distribution as \( B \).

As the main result of this work, we will prove that the computational complexity for \( \text{EVE} \) to efficiently recover \( x \) can be mapped to the problem of solving standard lattice problems which are conjectured to be computationally hard.

\section{Definitions}

\subsection{Definitions}

The definitions will be used in the proofs of our results. We define a \textit{negligible} amount in \( n \), denoted \( \text{negl.}(n) \) as any amount being asymptotically smaller than \( n^{-c} \) for any \( c > 0 \). Similarly a \textit{non-negligible} amount is one that is at least \( n^{-c} \) for some \( c > 0 \). We define a \textit{polynomial} amount, denoted as \text{poly}(\( n \)), as an amount that is at most \( n^c \) for some \( c > 0 \). An expression is \textit{exponentially small} in \( n \) when it is at most \( 2^{-\Omega(n)} \), and \textit{exponentially close} to 1 when it is \( 1 - 2^{-\Omega(n)} \). A probability is \textit{overwhelming} if it holds with probability \( 1 - n^{-c} \) for some \( c > 0 \). An algorithm is \textit{efficient} or \textit{efficiently computable} if its run time is within some polynomial factor of the parameter \( n \). An algorithm is \textit{hard} if its run time is at least \( 2^n \). We assume that, as input, algorithms accept real numbers approximated to within \( 2^{-n^c} \) for some \( c > 0 \). \( \mathbb{Z} \) represents the set of all integers and \( \mathbb{Z}_q \) represents the set of integers modulo \( q \). Given two probability density functions \( \phi_1, \phi_2 \) on \( \mathbb{R}^n \), we define the \textit{total variational distance} as

\[
\Delta(\phi_1, \phi_2) = \frac{1}{2} \int_{\mathbb{R}^n} |\phi_1(x) - \phi_2(x)| \, dx. \tag{6}
\]

\section{C. MIMO Signal Distributions}

In this subsection, we define various distributions that are used in our problem. Specifically, we discuss distributions of lattice points and distributions which can be empirically related to received MIMO signals.

First, in the proof of our work, we will need to generate lattice points according to a Gaussian distribution. Since a lattice is a discrete set of points, we define a \textit{discrete Gaussian distribution}, \( D_{A,\alpha} \), for any countable set \( A \) and parameter \( \alpha \) as

\[
\forall x \in A, D_{A,\alpha}(x) = \frac{\Psi_\alpha(x)}{\Psi_\alpha(A)}. \tag{7}
\]

We now define the distribution, \( D_{M,\alpha,k} \) on \( \mathbb{R}^n \times \mathbb{R} \). The distribution \( D_{M,\alpha,k} \) is the distribution of channel gains and the received signal from a single antenna in a MIMO system for a transmitted vector. This distribution is defined for a single antenna so a MIMO system with \( m \) receive antennas would get \( m \) samples from this distribution. This distribution is the input to the MIMO decoding problem which we will define in this section. For some arbitrary \( x \in \mathcal{X}^n \), we define the distribution \( D_{M,\alpha,k} \) as the distribution on \( \mathbb{R}^n \times \mathbb{R} \) given by drawing \( a \) as defined above, choosing \( e \sim \Psi_{M,\alpha} \) and outputting \( (a, y = (a, x) + e) \). We could alternatively express our system as drawing \( e \sim \Psi_{\alpha} \) and outputting

\[
\text{Figure 2. A Gaussian distribution on a two-dimensional lattice}
\]
\( (\mathbf{a}_i, y_i = \langle \mathbf{a}_i, \mathbf{x} \rangle / M + e_i) \), which would result in the same received SNR.

**D. The MIMO Decoding Problem**

Given these distributions, we now precisely define MIMO decoding for the eavesdropper in the MIMO wiretap channel, which we denote as the MIMO-Search problem. The search problem asks us to recover the transmitted vector \( \mathbf{x} \) without error. In Section V, we discuss how to use the hardness of the problem to achieve cryptographically secure systems, and provide a comparison between cryptographer's notions of security with information theoretic ones. In Appendix A, we prove that the search problem is as hard as solving certain lattice problems. We also use the term “MIMO decoding problem” to refer to the search problem.

We wish to show that the MIMO decoding problem, defined below, is hard to solve, i.e., that this decoding is of exponential complexity in the number of transmit antennas. We say that an algorithm solves this problem if it returns the correct answer with a probability greater than \( 1 - n^{-c} \), for some \( c > 0 \).

**Problem Definition 1.** MIMO-Search\(_{M,\alpha,k}\). Let \( M \geq 2 \), \( \alpha \in (0, 1) \), \( k \in \mathbb{R} \), \( n > 0 \). Given a polynomial number of samples of \( A_{M,\alpha,k} \), output \( \mathbf{x} \).

The MIMO search problem above will be related to solving standard lattice problems (discussed in Section III) through standard reductions. In Section III and Appendix B we will also discuss the complexity of solving these lattice problems. The result of our reduction implies that the complexity of decoding a MIMO system grows exponentially in the number of transmit antennas, i.e., MIMO decoding is at least as hard as the lattice problems to which we relate them. More precisely, we state this in a contrapositive manner below. This contrapositive statement allows us to conjecture a lower bound on the complexity of solving the MIMO decoding problems, based on the conjectured hardness of solving lattice problems.

**Main Result.** Let \( M > m 2^n \log \log n / \log n \) for some \( m > 0 \), \( \alpha \in \mathbb{R} \), and \( k \in \mathbb{R} \) be such that \( m\alpha / k^2 > \sqrt{n} \). Given access to an efficient algorithm that can solve \( \text{MIMO-Search}_{M,\alpha,k} \), there exist efficient classical and quantum solutions to standard lattice problems, which are conjectured to be hard.

We briefly introduce a second problem, the MIMO-Decision problem. This problem asks whether or not received samples are transmitted from a MIMO system (with known channel gain matrix) or are generated from a Gaussian distribution. Similar decision problems are common in cryptography, for example, in [15]. Regev shows that the decision variant of the LWE problem is hard, and uses this fact to achieve semantic security – effectively hiding information in a random variable that is uniformly random. Due to the exponential requirement on the constellation size in our main result, a reduction from the MIMO-Decision problem to the MIMO-Search problem is not apparent. It is an open problem to show whether or not this problem is hard, and if so, how the result could be used to construct a secure cryptosystem.

**Problem Definition 2.** MIMO-Decision\(_{M,\alpha,k}\). Let \( M \geq 2 \), \( \alpha \in (0, 1) \), \( k \in \mathbb{R} \), \( n > 0 \). Given a polynomial number of samples, distinguish between samples of \( A_{M,\alpha,k} \) and samples of \( \Psi_{M(\alpha+k)} \).

**E. Assumptions**

In this work we offer a proof of security by showing that breaking our cryptosystem is at least as hard as solving well-known lattice problems in the worst case. These problems are conjectured to require exponential running time. This conjecture and the large body of research behind it is discussed in the following section. For a more through treatment of the hardness of these problems, see [3]–[12]. Our proof is also based upon the following assumptions:

- We assume that the Gaussian channel noise has sufficiently large power so that \( m\alpha / k^2 > \sqrt{n} \). If the noise is too small, it is possible that a subexponential algorithm exists to recover the message (for motivation of this belief consider the subexponential attack on the LWE problem with low noise described by [27]). One possible way to ensure this requirement is met is to add noise at the transmitter. This intentionally degrades the communication SNR for a trade-off of security. A further discussion and characterization of our system in terms of received SNR is found in Section V.

- We assume that the constellation size of the system is large relative to the number of transmit antennas. Specifically, in order for the proof of Theorem 1 to hold, we require that \( M > m 2^n \log \log n / \log n \). Unlike the noise requirement, it is unclear that sub-exponential decoding immediately follows if this bound is not strictly met; it is an open problem as to whether or not this requirement is needed.

- We assume that each entry in the channel gain matrix is independent and Gaussian. More importantly, we assume that the two receivers of Bob and Eve, have independent channels. The fact that channels between different antennas are independent is well justified in most scattering environments. For example, in a uniform scattering environment, it has been shown that channels are independent over a distance of \( 0.4 \lambda \), see e.g. [28] or [29]. This independence analysis is extended to MIMO channels in [30] and [31].

**III. LATTICES**

We now provide an overview of lattices and lattice-based cryptography. This section contains all of the concepts used in the cryptography literature that are required in the proof of our main result. We first define several problems on lattices that are all conjectured to be hard to solve and are all used in the proof of our main result to show that MIMO decoding is at least as hard as solving standard lattice problems. We next provide a discussion on the complexity of solving these lattice problems, followed by a discussion on the Learning With Errors (LWE) problem. The Learning With Errors problem has a striking similarity to the problem of MIMO decoding. We follow a very similar approach to show the hardness of
MIMO decoding as is used to show the hardness of LWE decoding.

Lattice-based cryptography has generated much research interest in recent years. Cryptographers’ interest in lattices largely began with the surprising result of Ajtai [5] which created a connection between the average-case and worst-case complexity of lattice problems. Cryptographers have used these results to create a wide variety of cryptographic constructions which enjoy strong proofs of security. Important to this paper is the work of Micciancio and Regev [13] which explores Gaussian measures on lattices. The Learning With Errors problem, first introduced by Regev [13], shows that recovering a point that is perturbed by a small Gaussian amount from a random integer lattice can be related to approximating solutions to standard lattice problems. This problem is discussed more thoroughly below. The hardness reduction in this work very closely follows the work in [13]. For a survey on lattice-based cryptography, see [3].

A lattice is a discrete periodic subgroup of \( \mathbb{R}^n \) that is closed under addition. Alternatively, a lattice can be defined as the set of all integer combinations of \( n \) linearly independent vectors, known as a basis,

\[
L(A) = \left\{ Ax : \sum_{i=1}^{n} x_i A_i : x \in \mathbb{Z}^n \right\}. \quad (8)
\]

A lattice is not defined by a unique set of basis vectors. Any basis \( A \) multiplied by a unimodal matrix results in an alternative basis representation of the same lattice. We can define a set of lengths, in the \( l_2 \)-sense, \( \lambda_i(A) \), for \( i \in \{1, \ldots, n\} \), as the radius of the smallest ball around the origin that contains \( i \) linearly independent vectors. These lengths are known as the successive minima. It is easy to show that there exists vector \( \|v_i\| = \lambda_i(A) \) for all \( 1 \leq i \leq n \) and that \( \lambda_1(A) \geq \min_i \| A_i \| \), where \( A \) is the Gram-Schmidt orthonormalization of the basis \( A \), where the norm is taken as the \( l_2 \) norm.

The dual of a lattice \( L(A) \in \mathbb{R}^n \), \( L^*(A) \), is the lattice given by all \( y \in \mathbb{R}^n \) such that for every \( x \in L(A) \), \( (x,y) \in \mathbb{Z} \). Since \( A \) is full rank, \( L((A^T)^{-1}) \) is the dual of the lattice \( L(A) \).

### A. Lattice Problems

This section considers a set of standard problems on lattices that are conjectured to be computationally hard. These problems will be used in the characterization of the computational complexity of MIMO decoding. These problems are defined with an approximation factor, \( \gamma(n) \geq 1 \), and input given in the form of an arbitrary basis. The precise definition of the approximation factor varies for each problem, but is precisely stated in each definition below. The approximation factor plays an important role in the computational complexity required to solve a given problem. In general, for very large approximation factors, the following lattice problems can be efficiently computed. For small approximation factors, the problems are conjectured to require exponential running time. The problem of MIMO decoding will be related to solving lattice problems with an approximation factor of \( \gamma = n/\alpha \), which is conjectured to be hard. The relation between the approximation factor and the complexity of solving the problem is discussed in detail in Appendix B. For a survey on the hardness of these problems, see [3] or [4].

The first problem we describe is the shortest vector problem. In general, these problems can be efficiently solved if \( \gamma \) is at least \( O(2^n) \), but are conjectured to require exponential run time for exact solutions or even for polynomial approximation factors. A close variant of this problem, the decision Shortest Vector Problem (GapSVP), asks whether or not the shortest vector generated by the lattice is shorter than some distance \( d \). In general, for an arbitrary basis, short vectors are hard to find.

**Definition 1.** GapSVP\(_\gamma\) is defined as follows: given an \( n \)-dimensional lattice \( L(A) \) and a number \( d > 0 \), output YES if \( \lambda_1(A) \leq d \) or NO if \( \lambda_1(A) > \gamma(n) \cdot d \).

The following problem, the shortest independent vectors problem (SIVP), is often referred to as a lattice basis reduction.

**Definition 2.** SIVP\(_\gamma\) is defined as follows: given an \( n \)-dimensional lattice \( L(A) \), output a set of \( n \) linearly independent vectors of length at most \( \gamma(n) \cdot \lambda_n(A) \).

One well-known algorithm for finding approximate solutions to the SIVP problem is the Lenta-Lenta-Lovász (LLL) lattice basis reduction algorithm, which creates an “LLL-reduced” basis in polynomial time. For large values of \( n \), the LLL algorithm returns a basis that is exponentially larger than the shortest possible basis of the lattice. It is generally conjectured that no polynomial time algorithm could approximate SIVP to within a polynomial factor of \( n \). For a more in-depth discussion of lattice basis reduction algorithms and their complexity, see [3].

The following two lattice problems are important in the reduction given in this paper. They are also used by Regev in [13] in his proof of the security of LWE. The first problem, the Bounded Distance Decoding (BDD) problem, is equivalent to decoding a linear code, where the received vector is at most some distance \( d \) from the nearest codeword. Exact solutions to this problem are conjectured to require exponential run time. By requiring \( d < \lambda_1/2 \), we ensure that the closest vector is unique. The Closest Vector Problem (CVP) is identical to this problem but without a bounding distance. The closest vector problem arises in many other contexts, for example it is equivalent to decoding a linear code, and also the maximum-likelihood decoding problem in MIMO systems.

**Definition 3.** The BDD\(_{L(A),d}\) problem is defined as follows: given an \( n \)-dimensional lattice \( L(A) \), a distance \( 0 < d < \lambda_1(A)/2 \), and a point \( x \in \mathbb{R}^n \) which is at most distance \( d \) from a point in \( L(A) \), output the unique closest vector in \( L(A) \).

Informally, we can relate MIMO decoding to BDD (or more generally to the Closest Vector Problem), as follows. Linear precoding is well known to simplify encoding and decoding of a MIMO system to be of polynomial complexity in \( n \). In terms of lattice problems, we say that this linear precoding transforms the lattice basis into one in which the Closest Vector Problem is easy to solve. This is very closely related to the cryptographic notion of the trapdoor function: a function
that is easy to compute in one direction, but computationally infeasible to invert without a key (which serves as the “trapdoor”) \cite{13}. The linear precoding transformation applied to the MIMO channel effectively creates a spatially-varying trapdoor in that it allows the MIMO channel to be efficiently inverted only at Bob’s location. In our model, both communicators must have knowledge of the channel in order to perform the parallel decomposition via a singular value decomposition (SVD) associated with linear precoding. The eavesdropper may also learn this channel and hence learn the “key”, but this does not allow it to decode the message with lower complexity.

We finally define the discrete Gaussian sampling problem (DGS), which, for small values of $r$, can be reduced to solving GapSVP and SIVP. See \cite{13} for this reduction. This problem asks us to sample a Gaussian distribution, with standard deviation $r/\sqrt{2\pi}$, with support over a lattice.

**Definition 4.** The DGS$_r$ problem is defined as follows: given an $n$-dimensional lattice $L(A)$ and a number $r$, output a sample from $D_{L,r}$, the discrete distribution defined in Section II.3.

We borrow the following claim from \cite{33}, which gives an efficient algorithm to solve DGS$_r$ for large $r$. Intuitively, solving DGS$_r$ for small values of $r$ becomes a computationally harder task because it reveals information about the short vectors in the lattice. We use the following claim in the proof of our main theorem, as our main theorem requires sampling Gaussian distributions on lattices.

**Claim 1.** \cite{33} Theorem 4.1. There exists a probabilistic polynomial-time algorithm that, given an $n$-dimensional lattice $L(A)$, and a number $r > \lambda_n(L(A)) \cdot \omega(\sqrt{\log n})$, outputs a sample from a distribution that is within negligible distance of $D_{L(A),r}$.

**B. Learning With Errors**

We now give an overview of the Learning With Errors problem. The reader may note the similarities between this problem and the problem of MIMO decoding. In particular, the Learning with Errors problem is very similar to the problem presented in this paper except that Learning with Errors is set entirely over integer fields, whereas MIMO decoding is set in the reals. We take advantage of the results related to the Learning with Errors problem to show the hardness of MIMO decoding.

The Learning with Errors (LWE) problem was first presented by Regev in \cite{13}. The problem allows one to have an arbitrary number of samples of “noisy random inner products” of the form

$$\langle a, y \rangle = \langle a, x \rangle + e,$$

where $a \in \mathbb{Z}_q^n$, $x \in \mathbb{Z}_q^n$, and $e \in \mathbb{Z}_q$. Each $a$ is random from the uniform distribution over $\mathbb{Z}_q^n$, and each $e$ is drawn from a discrete Gaussian distribution with a small second moment (relative to $q$). This problem is an extension of the classic learning-parity with noise problem of machine learning. As with many hard problems, the LWE problem has two variants: the search and decision variants. These two problems are related in complexity. In the search variant of the problem, the goal is to recover the vector $x$; whereas the decision variant asks us to distinguish between samples of $y$ and a random distribution. If we fix the number of samples to which one has access, the LWE problem becomes analogous to finding the closest vector in an integer lattice. In \cite{13}, Regev proves that the LWE problem is at least as hard as solving GapSVP$_{n/c}$ and SIVP$_{n/c}$ in the worst case, but requires the use of a quantum computer. Peikert in \cite{24} demonstrates a reduction that does not require a quantum computer for the case where $q = O(2^n)$.

If we fix the number of samples available to the receiver, then this problem becomes equivalent to decoding a random linear code. Consider the case where we take a random generator matrix $A \in \mathbb{Z}_q^{m \times n}$, a vector $t = Ax + e \in \mathbb{Z}_q^m$ and we wish to recover the vector $x$. If we restrict $m < \text{poly}(n)$, and provide the proper distribution on $e$, then decoding this code is as hard as solving worst-case lattice problems within a factor of $\gamma < n/c$.

We briefly summarize Regev’s reduction as follows. If an efficient algorithm exists to solve LWE, then this can be used to construct an efficient algorithm that solves the BDD problem for any $n$-dimensional lattice. This step is somewhat surprising as it allows us to use the LWE algorithm, which operates on an integer lattice reduced modulo $q$, to solve BDD for any real lattice. How this step is accomplished is not entirely intuitive, but requires one to convert a single instance of BDD into an arbitrary number of samples of the distribution defined in the LWE problem. This step also requires an input samples of a number of lattice points drawn from a discrete Gaussian distribution with a large second moment. With an efficient algorithm to solve BDD, it is possible to generate a distribution of lattice points which has a smaller second moment than input in the previous step. The above process can now be iterated with this new distribution of lattice points to solve BDD with a larger bounding distance than before. Repeating this step will eventually result in a distribution of lattice points with a very small second moment, which will reveal information about the shortest vectors of the lattice, allowing us to approximate solutions to GapSVP and SIVP.

This step, using an LWE algorithm to solve any instance of the BDD problem, turns out to be extremely useful to characterize the hardness of MIMO decoding. In particular, we will show that if we have an efficient algorithm for MIMO decoding, then this algorithm can be used to efficiently to solve instances of the BDD problem. Then, with an efficient BDD algorithm, Regev’s quantum reductions to worst-case lattice problems follow. Additionally for large $M$, Peikert’s classical reduction follows.

**C. Lattice-Based Cryptography**

In recent years, lattice-based cryptography has become a very attractive field for cryptographers for a number of reasons. The security guarantees provided by many lattice-based schemes far exceed that of many modern schemes such as RSA and Diffie-Hellman, since lattice problems enjoy an average-to-worst case connection, as discussed in Appendix B. Lattice problems also appear to be resistant to quantum
computers. The creation of such computers poses a significant challenge to the state of modern cryptography, since a quantum computer could break most modern number-theoretic schemes. In addition, many proposed lattice schemes are competitive with, or better than, many modern number-theoretic schemes in terms of both key size and the computational efficiency of encryption and decryption \[37\]. For these reasons, a number of lattice-based cryptography standards are being developed or have been developed recently by both the IEEE and the financial industry \[38\], \[39\].

Lattice-based cryptography provides cryptographers with a wide variety of tools to create many different cryptographic constructions. We here reference some of these constructions as it is possible that some of them could be applied to the MIMO decoding problem or even that the MIMO decoding construction could inspire entirely new cryptographic constructions. Importantly, the LWE problem can be extended to cyclotomic rings \[40\], which can improve the efficiency (in terms of key size and complexity of the encryption and decryption operations) of cryptosystems based on LWE. In \[13\], a system that is secure against Chosen-Plaintext Attack (CPA-secure) is presented and in \[24\] a Chosen-Ciphertext Attack secure (CCA-secure) system is presented. In addition, the LWE problem has been used to create a wide range of useful cryptographic primitives such as identity-based encryption \[41\], oblivious transfer \[42\], zero-knowledge systems \[43\], and pseudorandom functions \[44\]. The LWE problem was also used to construct a fully homomorphic cryptosystem \[45\], solving a long-standing open problem in cryptography.

**IV. MAIN THEOREM**

In this section we state our main theorem regarding the computational hardness of MIMO – Search\(_{M,\alpha,k}\). We show that given an efficient algorithm which can solve this problem, there exist efficient solutions to the problems of GapSVP\(_{n/\alpha}\) and SIVP\(_{n/\alpha}\). The reduction is based off the work of \[13\] and involves a step using a quantum computer. In \[24\] this step is removed, but the hardness is only related to the GapSVP problem. The relation between the reductions used to show the hardness of MIMO decoding and the reductions used to show the hardness of LWE is shown in Figure 3. Examples of parameters which meet the requirements stated in Theorem 1 are shown in Table 1. The proof of this theorem is found in Appendix A.

**Theorem 1.** MIMO – Search\(_{M,\alpha,k}\) to GapSVP\(_{n/\alpha}\) and SIVP\(_{n/\alpha}\). Let \(m > 0\), \(\alpha \in \mathbb{R}\), \(k \in \mathbb{R}\), \(M > m^{2n \log \log n / \log n}\), be such that \(m \alpha / k^2 > \sqrt{n}\). Assume we have an efficient algorithm that solves MIMO – Search\(_{M,\alpha,k}\), given a polynomial number of samples from \(A_{\alpha,k}\). Then there exists an efficient quantum algorithm that, given an \(n\)-dimensional lattice \(\mathcal{L}(A)\), solves the problems GapSVP\(_{n/\alpha}\) and SIVP\(_{n/\alpha}\). Hence, since GapSVP\(_{n/\alpha}\) and SIVP\(_{n/\alpha}\) are conjectured to be hard, it is also conjectured that MIMO – Search is hard.

An outline of the reductions is shown in Figure 3. Figure 3 also relates the work of Regev and Peikert to our work. The steps used to prove the theorem are outlined as follows:

- We first show that, given the MIMO oracle as described in Section II, we can solve problems where the coefficients of the channel gain matrix are instead drawn from a discrete Gaussian distribution as described in Lemma 1 in Appendix A.
- We begin by reducing a lattice basis \(A\) by using the Lenstra-Lenstra-Lovász (LLL) lattice-basis reduction algorithm. We then, using the procedure described in \[33\], create a discrete Gaussian distribution on this lattice, with a second moment around the length of the largest vector given in the reduced basis. We use this as the starting point for the iterative portion of the algorithm.
- The main step in the proof is given in Lemma 7, where we use this MIMO decoding oracle to solve the BDD problem given access to a DGS oracle. This allows us to directly use the results from Regev \[13\] and Peikert \[24\].
- In Lemma 8, from \[13\], the BDD oracle is used to (quantumly) solve DGS\(_{L,r,\eta(L)}(\mathcal{D})\), that is return samples of \(D_{L,r}\). Note that we can efficiently sample from \(D_{L,r}\) for \(r > \eta(L)\). In Lemma 7 we set parameters so that \(\sqrt{\Delta d} > \sqrt{n}\), then we can reduce the value of \(r\) to below the value for which we could previously efficiently sample, that is we can construct a distribution that is more narrow than previously possible.
- The BDD and DGS oracles can now be applied iteratively, shrinking the second moment of the discrete Gaussian distribution with each iteration. Eventually, the distribution becomes narrow enough to reveal information about the shortest vectors of the lattice, thereby solving the GapSVP\(_{n/\alpha}\) and SIVP\(_{n/\alpha}\) problems.

**V. SECURITY AND THE MIMO WIRETAP CHANNEL**

In the previous sections we have shown that the MIMO wiretap model is secure in the sense that the decoding complexity of any eavesdropper is exponential in the number of transmit antennas. In this section, we relate this notion of security for our system to other notions of security used by information theorists and cryptographers.
Example sets of parameters for various numbers of transmit antennas. In order for the security proofs contained in this paper, the system must meet these minimum constellation size and maximum SNR values. Here we set \( m = 1 \).

Until this section, we have referred to the third user on Alice and Bob’s public channel to be an eavesdropper. This implies that the eavesdropper is passive and has no ability but to receive and decode information. In practice, this model is limited and, in order to provide more practical and robust notions of security, a more powerful adversarial model should be considered. The need for using more robust security models in the study of physical-layer security was recently discussed in [23]. In this work the author suggests bridging the gap between notions of security in information theory and cryptography in order to make physical-layer schemes more widely accepted by the security community. Considering an adversary which, for example, has the ability to manipulate, inject, alter or duplicate information is considered essential when a cryptographer develops and designs a cryptosystem, but is rarely, if ever, considered in information-theoretic settings. In this section, we discuss stronger adversarial models which are commonly considered in cryptography and show how to use the hardness result derived in Section IV in order to construct secure schemes under these models. We note that there are many other adversarial models studied in the field of cryptography which we do not discuss here. A more complete discussion of these models is provided in [1].

In this section, we present two ways in which Alice and Bob can securely use our result. In the first scheme, we show how Alice may securely transmit a message to Bob. This scheme achieves both the notion of Chosen Plaintext Security (also known as Semantic Security) as well as Chosen Ciphertext Security. However, the downside of this scheme is that it relies on additional cryptographic primitives and assumptions: this scheme is secure in what cryptographers call the Random Oracle Model, which we describe in Section V.B. In Section V.C, we present a simple scheme which allows Alice and Bob to agree on a secret key which, in order to be practical, must take on additional assumptions about Eve’s ability to approximately decode. Constructing a cryptosystem which is provably secure using only the assumptions listed in Section II is not readily apparent. We believe that the schemes proposed are practical and that applying these notions of security in the context of physical-layer security is not only novel, but makes substantial progress in bridging ideas from information-theoretic security and cryptography, as discussed in [23].

### A. Information-Theoretic Security

We have shown that it is hard for Eve to decode Alice’s message. However, we defined decoding as exactly recovering the message without error. This is not the same as stating that, information-theoretically, Eve receives no information about the message. Consider that Eve applies some suboptimal decoder that runs in polynomial time (for example, successive interference cancellation or zero-forcing decoding). Eve would then get some estimate for Alice’s message with some bit error rate that is potentially less than half. In some sense, this is insecure as it implies that Eve in fact receives information about Alice’s message. In this section, we provide a means to use our system in a manner that is in fact secure under cryptographic notions of security rather than information-theoretic ones.

Under the widely accepted conjecture that lattice problems cannot efficiently be approximated to even subexponential factors, Eve can at best recover a constellation point that is exponentially far away from Alice’s transmitted vector. Effectively, this means that for a single channel use, Eve can reduce the number of possible transmitted vectors from \( M^n \) to \( 2^{O(n)} \).

We could use this result in an information-theoretic sense along with our hardness result to say that there is now a positive secrecy rate in the wiretap channel, which could be accomplished through either a key rate, coding, or using a transmission strategy that takes into account the channel characteristics of Bob and/or Eve (e.g. puts a null in the direction of Eve). While this approach would extend existing results on information-theoretic secrecy, it is different from our approach in two main aspects, as we now describe.

First, assume that the average received SNR at Bob’s location is identical to the average received SNR at Eve’s location and, at both locations, the sufficient conditions of our hardness result holds. Since the channel gain matrices between Alice and Bob and Alice and Eve have the same distribution, on average they have the same number of spatial streams that can be supported and at the same SNR. In other words, the mutual information between the channel input and the output at Bob and at Eve is the same, unless the input is encoded in a manner that allows Bob to decode whereas Eve cannot. This will reduce the mutual information between Alice and Eve, possibly to zero. On the other hand, our approach shows that there is (likely) no efficient algorithm that allows Eve to decode, and thus, practically, if Eve is subject to complexity constraints she will receive less information than Bob. To our knowledge, this notion of information exchange, based on the conjectured complexity of decoding, is novel.

Second, relying on the information theoretic secrecy argument does not provide a constructive result. By this we mean that it merely shows the existence of a scheme that allows for Alice and Bob to reliably and securely communicate, rather than actually providing such a scheme, as is the case for most information-theoretic results. In contrast, the cryptographic notion of security that we use in fact gives a practical scheme for secure communication.

### B. Secure Message Transmission in the Random Oracle Model

We have shown that it is computationally infeasible for Eve to exactly recover Alice’s transmitted message, but Bob can decode in polynomial time. This implies that (assuming lattice

### Table I

| \( n \) | \( \log_2 M \) | \( \text{SNR (dB)} \) |
|---|---|---|
| 80 | 3.57 | 87.1 |
| 128 | 5.13 | 139.2 |
| 196 | 6.81 | 210.7 |
| 256 | 7.68 | 272.2 |

For secure communication.
problems are hard) the MIMO channel naturally forms what cryptographers call a secure one-way trapdoor function\(^2\). In this subsection, we construct a secure cryptosystem by using a standard cryptographic construction, known as OAEP+ (Optimal Asymmetric Encryption Padding). OAEP+ was introduced in \([46]\) as an improvement to the original OAEP scheme given in \([47]\). The results of \([46]\) allow for the construction of a secure cryptographic system given any one-way permutation. The results imply that, in the random oracle model, the resulting system has the properties of Chosen Ciphertext (CCA) Security and Chosen Plaintext (CPA) Security. Precisely, the security of our system follows from the following theorem, which is proven in \([46]\) for any one way permutation. The fact that the MIMO-Search problem is one way, under the assumption that lattice problems are hard, is given by Theorem \([45]\). Theorem 2 gives strong evidence that at least one bit (and likely more) will remain computationally hidden from Eve, and thus Eve will learn no information about the transmitted message.

Theorem 2. \([57]\), Thm. 3 If the underlying trapdoor permutation is one way, then OAEP+ is secure against adaptive chosen ciphertext attack in the random oracle model.

We now give an informal treatment of CPA and CCA Security, as well as the random oracle model. For a formal treatment, see \([11]\) or \([48]\). A reader familiar with these notions may wish to skip ahead to Algorithm 1.

Succinctly, the condition necessary to achieve security under Chosen Plaintext Attack (CPA security – here we describe the notion referred to as \(\text{IND-CPA}\)) is that an adversary must be unable (without violating our computational hardness assumptions) to match up pairs of plaintext (unencrypted data) and ciphertext (encrypted data). This notion of CPA security was formally introduced in \([49]\). There are many reasons why such a notion of security is needed. For example, consider that an adversary knows that a user communicates with only a known small subset of possible messages. If the adversary can match plaintext and ciphertext, then the adversary can exhaustively search over the possible transmitted messages and learn the message.

We illustrate the value of this notion of security with an example from our model in Figure 1. Sending non-random messages, or messages from a source with a small entropy, could be problematic from a security perspective. Consider, for example, that Alice is casting a vote in an election which she sends to the election official, Bob. In the election, there are a small number, \(v\), of candidates. Alice votes by sending some message \(x_i\), for \(i \in [0, v]\), which is predefined, and publicly known, for each candidate. Eve could receive \(y = Bx_i + e\), and compute the \(l_2\) norm between \(y\) and each possible \(\{Bx_0, \ldots, Bx_{v-1}\}\) and tell with some certainty

\[^2\] A one-way function is a function which can be evaluated in polynomial time, but requires exponential complexity to invert. Such functions are conjectured to exist, but proving their existence is an open problem and would prove \(P \neq NP\).

\[^3\] The MIMO channel maps messages from bits to real values and thus is a function and not a permutation. However, the MIMO search problem asks Eve to decode a MIMO signal back to bits, and thus we can view the process of transmitting, detecting and decoding as a permutation of the message bits.
Algorithm 1 MIMO-OAEP+
Let $K = n \log M$ be the number of bits transmitted per MIMO channel use. Alice wishes to send Bob a message, $m$, that is $\eta = K - 2n$ bits. Assume Alice and Bob both have access to three random oracle functions: $G : \{0,1\}^n \to \{0,1\}^\eta$, $H' : \{0,1\}^{\eta+n} \to \{0,1\}^n$, $H : \{0,1\}^{n+\eta} \to \{0,1\}^n$. Alice computes $s \in \{0,1\}^{n+\eta}$, $t \in \{0,1\}^n$, $x \in \{0,1\}^K$ as shown under Encrypt below. Alice then multiplies $x$ by the right singular vectors of Bob’s channel and transmits the message to Bob. Bob recovers $x$ through his channel and then recovers $m$ using the procedure shown under Decrypt. Bob verifies that $c = H'(t||m)$. If these quantities are not equal then Bob has not properly received Alice’s message and he rejects.

| Encrypt | Decrypt |
|---------|---------|
| $s = (G(r) \oplus m) \parallel H'(r||m)$ | $s = x[0, \ldots, \eta + n - 1]$ |
| $t = H(s) \oplus r$ | $t = x[\eta + n, \ldots, k]$ |
| $x = s||t$ | $r = H'(s) \oplus t$ |
| $m = G(r) \oplus s[0, \ldots, \eta - 1]$ | $c = s[\eta, \ldots, \eta + n - 1]$ |
| $c = H'(t||m)$ | |

Bob receives this message through a channel with Gaussian noise and thus has some probability of error that he does not receive this message perfectly. In fact, given the large noise requirement in Theorem 2, it can be shown that the symbol error rate at Bob’s location will be close to one, even when Bob has a large number of antennas.

In order to allow Alice and Bob to communicate, we must reduce the rate of communication between them; however, in order to allow for Theorem 2 to hold, we must still maintain the large constellation size and noise requirement. Both of these objectives can be accomplished through the use of error-correcting codes. Note, however, that if we encode the message $x$ with a code that can correct up to $e$ errors, then this also aids Eve in recovering the plaintext. Because Eve cannot apply ML decoding, she will experience an increased error rate over Bob. We can use this fact to construct a code which Bob can decode but Eve cannot.

1) Eavesdropper Decoding: We now consider the limits to which Eve can estimate the message $x$ which is not precoded for her channel. The analysis in this section describes how much Eve’s estimate of $x$ differs from its maximum-likelihood, which we denote as $\hat{x}$. We obtain a lower bound on this distortion by considering state-of-the-art algorithms which are used to attack lattice-based cryptography schemes. We begin by stating a conjecture about the limits of approximating lattice algorithms. This conjecture is widely accepted to be true (see [43], for example) and discussed in more detail in Appendix B:

**Conjecture 1.** There is no polynomial time algorithm which can approximate SIVP to within polynomial factors.

Alice transmits a message $x \in [0, M]^n$, which is distributed uniformly, across her channel. Eve receives this message through her channel and can apply her favorite lattice-basis reduction algorithm to decoding the message, but is unable to obtain a lattice basis that is within a polynomial factor of the shortest basis. She can use this shorter basis along with Babai’s nearest plane algorithm [52] to recover a vector, $x'$, that is a super-polynomial distance away from $x$, that is $||x' - x|| > \sqrt{\omega(\log n)}$. The fact that Eve can recover this vector $x'$ implies that Eve can in fact learn a non-negligible amount of information about Alice’s message.

In order to more precisely determine the length of the basis which Eve can practically find, one can consider the current state-of-the-art for lattice basis reduction algorithms. For a study on limits of lattice basis reduction algorithms, see [53], for a more general overview see [3]. In [53], the authors suggest that the closest Eve can decode relative to the original message (given a realistic amount of computational resources) is bounded by $||x' - x|| > 2^{\sqrt{n} \log M \log \log \log n}$. According to [53], this is an extremely conservative estimate, and we claim that it would require either an infeasible amount of computational resources or a substantial advancement in lattice-basis reduction algorithms for Eve to learn more about the original message than this bound indicates.

In order to prevent Eve from being able to recover a coded version of the message $x$, we apply a code which has a minimum distance $d_{\text{min}} \leq 2^{2\sqrt{n} \log M \log \log n}$. With high probability, Eve will be unable to recover an error-free estimate of $x$, and will be unable learn any information about the message $m$ sent using Algorithm 1.

2) Correctness: We now show that, if Alice encodes the output of Algorithm 1 with an error-correcting code of $d_{\text{min}} = 2^{\sqrt{n} \log M \log \log n}$, then given a sufficient number of receive antennas, Bob will be able to receive and decode the vector $x$ without error, thus allowing him to recover the message $m$ from Algorithm 1. We proceed by choosing the parameters $M = 2^{n \log \log n / \log n}$, $m = 1$, $\alpha = 1$, and $k = 1$, but this analysis is easily generalized to other parameters which meet our security condition.

Assume that Bob has $n_r = t \cdot n$ receive antennas for some $t > 2$. With probability at most $2e^{-n/2}$, the smallest singular value in the channel between Alice and Bob will be at least $\sqrt{(t - 2)n}$ (see for examplecite{rudelson}, Eq. 2.3). This implies that the noise variance in each of the decomposed channels between Alice and Bob will decrease by at least a factor of $\sqrt{(t - 2)n}$ compared to the ambient channel noise. For Bob to be able to decode, we now require the following condition to hold with some reasonable probability:

$$\frac{M \alpha}{\sqrt{(t - 2)n}} < 2^{2\sqrt{n} \log M \log \log n},$$

where the left-hand side is an upper bound of the noise variance in each of Alice and Bob’s parallel channels, and the right-hand side is the minimum distance of the error correcting code employed by Alice. Substituting the values of our constants, this gives us:

$$t > \frac{2.86}{n}$$
and thus, for any \( t > 2 \), the probability of having a symbol error when Bob decodes is bounded below by:

\[
erfc\left(\frac{m_e}{2.86}\right) = \text{negl}(n),
\]

where \( \text{erfc}(\cdot) \) represents the complimentary error function. We note that the condition that \( t > 2 \) is only necessary so that, with overwhelming probability, there will be no ill conditioned channels in the parallel decomposition between Alice and Bob. For large \( n \), the probability of error will remain small for values of \( t \) close to 1.

### C. Secret Key Agreement and Computational Secrecy Capacity

We now proceed to describe a secret key transmission protocol. In the context of our model in Figure 1, such a protocol allows Alice to transmit a “secret” to Bob in the form of bits undecodable by Eve. Information-theoretic protocols which achieve this secret key transmission based on equivocation at the eavesdropper are known, for example [51]. Here we present a protocol in which the key is kept secret under computational rather than information-theoretic assumptions.

We call the number of bits Alice can securely transmit to Bob per channel use under computational assumptions the **Computational Secrecy Capacity**.

Alice transmits a message \( m \in [0, M]^n \) chosen uniformly at random through her channel. This message has \( n \log M \) bits of entropy. By similar reasoning to the previous subsection, Eve receives this message through her channel and recovers a vector, \( m' \), such that \( \|m - m'\| > 2\sqrt{n \log M \log 1.005} \). The fact that Eve can recover this vector \( m' \) implies that Eve can in fact learn a non-negligible amount of information about Alice’s message. Here we argue that, in the context of our proposed schemes, this information is in fact useless unless Eve can violate our computational assumptions.

Achieving this bound in terms of Eve’s ability to decode the transmitted message implies that the entropy associated with Eve’s estimate of \( m \) is at least \( 2\sqrt{n \log M \log 1.005} \). By applying the leftover hash lemma [55], it follows that Alice can transmit to Bob \( 2\sqrt{n \log M \log 1.005} \) bits of information per message without Eve learning any (non-negligible) amount of information about the message. Thus, the security of the bits transmitted via the scheme described in Algorithm 2 immediately holds. The quantity \( 2\sqrt{n \log M \log 1.005} \) is the Computational Secrecy Capacity of our model in Figure 1.

Table II gives examples of the computational secrecy capacity for various parameters.

**Algorithm 2 Key-Agreement Scheme**

Alice wishes to send Bob \( \eta \) secret bits. Alice generates some number \( c = c(n) \), such that \( 2c \sqrt{n \log M \log 1.005} > \eta \), of random messages \( m \in [0, M]^n \) and sends them to Bob over the MIMO channel with channel parameters meeting the constraints in Theorem 1. Alice and Bob ensure that the message is exchanged without error for example through channel coding. Alice and Bob then hash the message (after decoding if channel coding is used), using a universal hash which outputs \( \eta \) bits and use the result as their secret.

| \( n \) | \( \log_2 M \) | SNR (dB) | Computational Secrecy Capacity |
|---|---|---|---|
| 80 | 33.7 | 87.1 | 8.80 |
| 128 | 51.3 | 139.2 | 13.75 |
| 196 | 75.4 | 210.7 | 20.62 |
| 256 | 90 | 272.2 | 26.50 |

For unit channel gain variance, the minimum constellation size, and the minimum SNR that meets the noise requirements for the hardness condition to hold. Computational Secrecy Capacity gives the number of bits Alice can securely transmit to Bob per channel use using Algorithm 1.

We note that if Alice and Bob use channel coding to ensure for reliable communications, then the rate of the coding scheme employed would also effect the number of bits Alice and Bob could securely exchange per channel use. That is, if Alice and Bob used a scheme that could correct up to \( \epsilon \) bit errors, then the bound on \( \eta \) must be reduced by \( \epsilon \) bits. Similarly, this quantity, \( 2\sqrt{n \log M \log 1.005} \), also serves maximum number of bit errors that can be allowed to be corrected in any code chosen between Alice and Bob in the scheme given in Section V.B.

### D. Computational Attacks

We wish to briefly note the importance of maintaining the security parameters as stated in Theorem 1. When the minimum noise requirement is not met, it is possible that attacks follow on our system. By attack, we mean that an adversary could, by applying a sub-exponential algorithm in order to decode. As an example, in [56], the authors show an attack on a version of our system with security parameters not meeting those defined in Theorem 1. Finally, we note that it is not apparent that an attack follows immediately for smaller values of \( M \) than required in Theorem 1, but we leave it as an open question as to whether or not it can be shown that Eve can successfully decode with non-exponential complexity for smaller values of \( M \), or if a smaller bound on the requirement for \( M \) can be found that still entails decoding to be hard for Eve.

### VI. Conclusion

We have demonstrated that the complexity of an eavesdropper decoding a large-scale MIMO systems with M-PAM modulation can be related to solving certain lattice problems which are widely conjectured to be hard. This suggests that the complexity of solving these problems grows exponentially.
with the number of transmitter antennas. Unlike the computationally hard problems underlying many of the most common encryption methods used today, such as RSA and Diffie-Hellman, it is believed that the underlying lattice problems are hard to solve using a quantum computer, and thus this scheme presents a practical solution to post-quantum cryptography.

It is not new to exploit properties of a communication channel to achieve security; however, to our knowledge, this is the first scheme which uses physical properties of the channel to achieve security based on computational complexity arguments. Indeed, the notion of the channel is not typically considered by cryptographers. We thus describe our system as a way of achieving physical-layer cryptography.

Further novel to our scheme is the role that the channel gain matrix plays in decoding. A transmitted message can only be decoded by a user with the corresponding channel gain matrix. The channel gain matrix, or more specifically the precoding of the message using the right-singular vectors of the channel gain matrix, essentially plays the role of a secret key in that it allows for efficient decoding at the receiver. However, this value does not need to be kept secret, nor does it play the traditional role of a public key. We term this type of key as the Channel State Information- or CSI-key. In cryptography terminology, this system is a trapdoor function, for which the trapdoor varies both spatially and temporally. The fact that this is a new type of cryptographic primitive suggests the possibility of entirely new cryptographic constructions.

We have used the hardness result, in conjunction with a new notion of computational secrecy capacity, to construct a method in which two users can perform a key-agreement scheme, without a pre-shared secret. In addition, we give a scheme that allows Alice and Bob to securely communicate in the presence of an eavesdropper. We relate the parameters required to maintain security to SNR requirements and constellation size and show that they are practical to achieve assuming a system with enough transmitter antennas and the corresponding number of receivers, and relatively large constellation sizes.

APPENDIX A

PROOF OF MAIN THEOREM

Theorem 1. MIMO – Search \( M, α, k \) to GapSVP\(_{n/α} \) and SIVP\(_{n/α} \). Let \( α > 0 \), \( m > 0 \), \( k > 0 \) be such that \( mα/k^2 > \sqrt{n} \), and \( M > m2^{n \log \log n} / \log n \). Assume we have access to an oracle that solves MIMO – Search\(_{M, α, k} \). Given a polynomial number of samples from \( A_{M, α, k} \). Then there exists an efficient quantum algorithm that given an \( n \)-dimensional lattice \( L(A) \), solves the problems GapSVP\(_{n/α} \) and SIVP\(_{n/α} \). Additionally, there exists a classical solution to GapSVP\(_{n/α} \).

Proof: The lemmas required to prove this theorem are given below. We summarize the proof of the main theorem as follows.

1) We first show that, given the MIMO oracle as described in Section II, we can solve problems where the coefficients of the channel gain matrix are instead drawn from a discrete Gaussian distribution as described in Lemma 1.

2) We begin with an arbitrary lattice basis \( A \) and apply the LLL algorithm. We then create a discrete Gaussian distribution on this lattice, with a second moment around the length of the largest vector given in the reduced basis. We use this as the starting point for the iterative portion of the algorithm.

3) In Lemma 7, we use this MIMO decoding oracle to solve the BDD problem. The input to this problem is an (arbitrary) \( n \)-dimensional lattice \( L(A) \), a number \( r > \sqrt{2ηr} (L(A)) \), and a target point \( y \) within distance \( d < Mα/k^2 r \sqrt{2} \) (the bounding distance) of \( L(A) \). We take this instance of a BDD problem and, using a number of samples from the distribution \( D_{L,r} \), we are able to construct a number of samples in the form of \( (A, y = (a, x) + c) \), in the exact form of distribution expected by MIMO decoding oracle using Lemma 1. Here, returning the correct vector \( x \) solves the BDD problem. We now have an oracle which solves the BDD problem for arbitrary lattices.

4) In Lemma 8, from [13], the BDD oracle is used to (quantum) solve DGS\(_{L,r} \), that is return samples of \( D_{L,r} \). Note we can efficiently sample from \( D_{L,r} \) for \( r > ηr(L) \). If, in Lemma 7 we set parameters so that \( \sqrt{2d} > \sqrt{n} \), then we can reduce the value of \( r \) to below the value for which we could previously efficiently sample, that is we can construct a distribution that is more narrow than previously possible.

5) In [13], the steps of Lemma 7 and 8 are iteratively applied, resulting in a more narrow distribution of lattice points. Eventually, this distribution becomes narrow enough to reveal information about the shortest vectors of the lattice, solving the GapSVP\(_{n/α} \) and SIVP\(_{n/α} \). We refer the reader to [13] for the rigorous treatment of this process.

6) We refer the reader to [24] for the classical reduction, which requires an oracle to solve the BDD problem. Replacing Regev’s LWE-based BDD oracle with our MIMO-based BDD oracle, the classical reduction follows.

A. Smoothing Parameter.

Before we prove the main theorem, we review the smoothing parameter and state some of its properties that we will require in our proof. The smoothing parameter was introduced in [34] and is an important property of the behavior of a discrete Gaussian distribution on lattices. It is precisely defined as follows.

Definition 5. For an \( n \)-dimensional lattice \( L(A) \) and a real \( ϵ > 0 \), the smoothing parameter, \( η_ϵ(L(A)) \), is the smallest \( α \) such that \( Ψ_α(L(A) \setminus \{0\}) \leq ϵ \).

The smoothing parameter defines the smallest standard deviation such that, when the inverse is sampled over the dual \( L^∗(A) \), all but a negligible amount of weight is on the origin. More intuitively, it is the width at which a discrete Gaussian measure begins to behave as a continuous one. The motivation for the name ‘smoothing parameter’ is given in [34]. For \( α > \sqrt{2ηr(L)} \), if we sample lattice points from \( D_{L,α} \) then add Gaussian noise \( Ψ_α \), then the resulting distribution is
at most distance $4\epsilon$ from Gaussian. We borrow the following two technical claims from \cite{13}.

**Claim 2.** \cite{13} Lemma 3.2. For any $n$-dimensional lattice $\mathcal{L}(A)$, $\eta_\epsilon(\mathcal{L}) \leq \sqrt{n}/\lambda_1(\mathcal{L}^+(A))$ where $\epsilon = 2^{-n}$.

More generally, we can characterize the smoothing parameter for any $\epsilon$.

**Claim 3.** \cite{13} Lemma 3.3. For any $n$-dimensional lattice $\mathcal{L}(A)$ and $\epsilon > 0$,
\[
\eta_\epsilon(\mathcal{L}) \leq \sqrt{\frac{\ln(2n(1+1/\epsilon))}{\pi}} \cdot \lambda_n(A). \tag{13}
\]

Equivalently, for any superlogarithmic function $\omega(\log n)$, $\eta_\epsilon(\mathcal{L}) \leq \omega(\log n) \cdot \lambda_n(A)$.

We will also note the following property of the smoothing parameter, which follows from the linearity of lattices. If we scale the basis of a lattice, then all of the successive minima will scale by the same amount:

**Claim 4.** For any $n$-dimensional lattice $\mathcal{L}(A), \epsilon > 0$, and $c > 0, \eta_\epsilon(c \cdot \mathcal{L} \cdot A) = c \cdot \eta_\epsilon(\mathcal{L} \cdot A)$.

### B. Preliminary Lemmas

Before we can proceed with the main part of the proof showing the reduction from standard lattice problems to the MIMO decoding problem, we require several preliminary lemmas. We will first show, in Lemma 1, that it is sufficient to solve the MIMO decoding from a discrete Gaussian distribution rather than a continuous one. This distribution is used as input to the MIMO oracle in Lemma 7. We define the distribution $D_{M, \alpha, k}$, which is the discrete analog of the distribution $A_{M, \alpha, k}$. Given an arbitrary lattice $\mathcal{L}(A)$, and a number $r > \sqrt{2}\eta_\epsilon(\mathcal{L}(A))$, we first sample a vector $a$ from the distribution $D_{\mathcal{L}(A), r}$, and a point $e$ from the distribution $\psi_{a}$. We now output:
\[
\left(\frac{ka}{r}, y = \left(\frac{ka}{r}, x\right) / M + e\right) \tag{14}
\]

**Lemma 1.** Continuous-to-Discrete Samples. Given an oracle which can solve MIMO - Decision$_{M, \alpha, k}$, there exists an efficient algorithm to recover $x$ given samples from $D_{M, \alpha, k}$.

**Proof:** We first claim that every point in the distribution $D_{\mathcal{L}(A), r}$ is proportional to $\psi_{r}$, to within a negligible amount. This effectively follows from the fact that we choose $r$ to be larger than the square root of two times the smoothing parameter of the lattice (formally, see equation 11 of \cite{13} Claim 3.9)).

We now create a unitary transformation that can be applied to both $a$ and $y$, creating $\tilde{a}$ and $\tilde{y}$, so that the support of $\tilde{a}$ is effectively $\mathbb{R}^n$. We do not have to actually cover all of $\mathbb{R}^n$, nor do we in fact have to come close to doing so. Our algorithms, by assumption, can only approximate points in $\mathbb{R}^n$ to within a factor of $2^{-n^c}$, for some $c > 0$. Thus, we only need to have a support in which no point is more than a distance of $2^{-n^c}$ away from any point in $\mathbb{R}^n$. Then, by the fact that we choose a unitary transformation, all norms and inner products will be preserved, and we will achieve our desired result. We describe an appropriate transformation as follows.

Take two samples from the discrete distribution, call them $(a_i, y_i)$ and $(a_j, y_j)$. Now generate two numbers $c_i, c_j \in \mathbb{Z}_{2^n}$, uniformly, and output
\[
\left(\frac{c_i a_i + c_j a_j}{c_i + c_j}, y_i + y_j\right) = \left(\frac{c_i a_i + c_j a_j}{c_i + c_j}, \frac{c_i y_i + c_j y_j}{c_i + c_j}\right) = \left(\frac{c_i e_i + c_j e_j}{c_i + c_j}, \frac{c_i y_i + c_j y_j}{c_i + c_j}\right)
\]

since each $e$ is generated i.i.d., it is not hard to see that the noise has the correct distribution. Similarly, since each $a$ is i.i.d., moments of the quantity $\frac{c_i e_i + c_j e_j}{c_i + c_j}$ will be unchanged, and this quantity will be proportional to the desired Gaussian distribution. We have now increased the support of the distribution. Indeed, the support of the quantities $\frac{c_i e_i + c_j e_j}{c_i + c_j}$ covers every point in the interval $[0, 1]$ to within a factor of $2^{-n^c}$, and this new distribution will be indistinguishable from the distribution expected by the MIMO oracle.

We next state the following claim which is proven in \cite{13} and shows that a small change in $\alpha$ results in a small change in the distribution of $\Psi_\alpha$. This claim is required in the proofs of Lemmas 2 and 3.

**Claim 5.** \cite{13} Claim 2.2. For any $0 < \beta < \alpha \leq 2\beta$, \[
\Delta(\Psi_\alpha, \Psi_\beta) \leq 9 \left(\frac{\alpha}{\beta} - 1\right). \tag{15}
\]

We next show that given a vector $x$, it is easy to verify whether or not it is the correct solution to the MIMO-Search problem:

**Lemma 2.** Verifying solutions of MIMO - Search$_{M, \alpha, k}$. There exists an efficient algorithm that, given $x'$ and a polynomial number of samples from $A_{x, \alpha, k}$, for an unknown $x$, outputs whether $x = x'$ with overwhelming probability.

**Proof:** Let $\xi$ be the distribution on $y - \langle a, x'\rangle$. The same distribution can be obtained by sampling $e \sim \Psi_\alpha$ and outputting $e + \langle a, x - x'\rangle$. In the case $x = x'$, this reduces to $e$, and the distribution on $\xi$ is exactly $\Psi_\alpha$. In the case where $x \neq x'$, $\|x - x'\| > 1$ by the restriction on our choices of $x$. The inner product of $\langle a, x - x'\rangle$ is Gaussian with zero mean and a standard deviation of at least $k/\sqrt{2\pi}$, and the standard deviation on $e + \langle a, x - x'\rangle$ must be at least $\sqrt{\alpha^2 + k^2}/\sqrt{2\pi}$. We now must distinguish between the random variables of $\Psi_{\alpha}$ and $\Psi_{\sqrt{\alpha^2 + k^2}}$.

Assuming that $k^2$ is non-negligible in $n$, then given an arbitrary number of samples within a polynomial factor of $n$, we can distinguish between the two distributions with overwhelming probability by estimating the sample standard deviation.

The following lemma is used from \cite{13} Lem. 3.7. In \cite{13}, the lemma applies to the case of LWE over an integer field, this proof is repeated in this appendix in order to demonstrate that it follows for the case of MIMO channels, given Lemma 7. Specifically, this lemma shows that if we can solve the MIMO problems with noise parameter $\alpha$, then we can solve the problems given samples with noise drawn according to $\Psi_{\beta}$ for any $\beta \leq \alpha$. 


Lemma 3.  [13 Lem 3.7], Error Handling for $\beta \leq \alpha$. Assume we have access to an oracle which solves MIMO – Search$_{M,\alpha,k}$ by using a polynomial number of samples. Then there exists an efficient algorithm that given samples from $A_{x,\beta,k}$ for some (unknown) $\beta \leq \alpha$, outputs $x$ with overwhelming probability.

Proof: Assume we have at most $n^c$ samples for some $c > 0$. Let $Z$ be the set of all integer multiples of $n^{-2c}2^2$ between 0 and $\alpha^2$. For each $\gamma \in Z$, do the following $n$ times. For each sample, add a small amount of noise sampled from $\Psi_{\sqrt{\gamma}}$, which creates samples in the form $A_{x,\sqrt{\gamma},\alpha,k}$. Apply the oracle and recover a candidate $x'$. Use Lemma 2 and check whether $x' = x$. If yes, output $x'$, otherwise continue.

We now show the correctness of this algorithm. By Lemma 2, a result can be verified to be a correct solution of the algorithm, we output samples that are close to zero mean and standard deviation $\sigma$. Then there exists an efficient algorithm that given samples from $A_{x,\beta,k}$ for some (unknown) $\beta \leq \alpha$, outputs $x$ with overwhelming probability.

Claim 6. [32], [36]. For some $\mathcal{L}(A)$, we apply Schnorr’s variant of the LLL algorithm, and obtain a new, shorter basis for this lattice $A$. The norms of the new basis vectors in this lattice, given by $\sigma_1, \ldots, \sigma_n$, are bounded by:

$$\sigma_n < 2^{n \log \log n / \log n} \lambda_n$$

and

$$\sigma_1 < 2^{n \log \log n / \log n} \lambda_1$$

The rest of this proof proceeds with the following assumption:

$$1 \leq \frac{\lambda_n}{\lambda_1} < 2^{n \log \log n / \log n}$$

The lower bound is evident from Minkowski’s bound, and the upper bound comes from the fact that we have reduced the basis by applying the LLL algorithm. While no such upper bound would exist on an arbitrary lattice, were this ratio to be bigger than this bound, then the LLL algorithm would have returned exactly the shortest basis and we would have already exactly solved the GapSVP and SIVP problems.

C. Reducing MIMO Decoding to Standard Lattice Problems

We now begin the main procedure of the reduction. We begin by taking the basis and applying the LLL algorithm.

Lemma 6. We start with an arbitrary lattice $\mathcal{L}(A)$, and apply the LLL algorithm to the basis $A$. We now the use procedure given in [33] and Schnorr’s variant of LLL to get a distribution $D_{L,r}$ for some $r > 2^{n \log \log n / \log n} \lambda_n$.

The following lemma is the main mathematical contribution of this work, and allows the MIMO oracle to be used in place of the LWE oracle in the framework of Regew’s reduction for LWE. From Figure 3, this replacement implies that an efficient solution of the MIMO decoding problem would also provide an efficient solution for standard lattice problems. We briefly restate notation defined in Section II: $A_{M,\alpha,k}$ is the distribution of channel gains and the received signal from a single antenna in a MIMO system, and $D_{A,\alpha}$ is the discrete Gaussian distribution drawn over lattice $A$ with variance proportional to $\alpha$.

Lemma 7. MIMO – Search$_{M,\alpha,k}$ to BDD$_{L,r}$. Let $\alpha > 0$, $k > 0$, $m > 0$, and $M > m 2^{n \log \log n / \log n}$. Assume we have access to an oracle that, for all $\beta \leq \alpha$, finds $x$ given a polynomial number of samples from $A_{M,\beta,k}$ (without knowing $\beta$). Then there exists an efficient algorithm that given an $n$-dimensional lattice $\mathcal{L}(A)$, a number $r > \sqrt{2^n/|\mathcal{L}(A)|}$, and a target point $y$ within distance $d < M \sigma / (k^2 \sqrt{2})$ of $\mathcal{L}(A)$, where $\sigma$ is the smallest eigenvalue of $A^T A$, returns the unique $x \in \mathcal{L}(A)$ closest to $y$ with overwhelming probability.

Proof: We describe a procedure that, given $y$, outputs a polynomial number of samples from the distribution of $D_{x,\beta,k}$. Then, using the MIMO-Search oracle returns the closest point $x = A s$. 

We now state the following claim about a polynomial-time lattice basis reduction algorithm, the LLL algorithm, given in [32] and improved by Schnorr in [36].
First, using Claim 1, sample a vector \( v \in L^* (A) \) from \( D_{L^*,r} \). We now output
\[
\left( \frac{k}{r}, k \left< \frac{v-A^{-1}y}{r} \right>/M + ke/r \right) \tag{19}
\]
We note that the right-hand side is equal to
\[
k \left< \frac{v-A^{-1}y}{r} \right>/M + ke/r = \\
<kv/r,s>/M + k < v-A^{-1}\delta>/rM + ke/r
\]
We are guaranteed that \( |\delta| < r \). We first note that the quantity \( k < v-A^{-1}\delta>/r \) is distributed according to \( D_{kl^*/r,\xi} \), where \( \xi < kd/\sigma \). We note that \( \sigma \) relates to the maximum ‘skew’ that results from inverting the matrix \( \xi < kd/\sigma \)
We now add some noise \( \psi_\alpha \) so that the discrete nature of \( D_{kl^*/r,\xi} \) is effectively washed out and we are left with a distribution that is essentially Gaussian as expected by the MIMO oracle. That is the distribution of the noise is within negligible total variational distance of \( \psi_\beta \) for \( \beta = \sqrt{\xi^2 + \alpha^2}/2 < \alpha \). Condition will be true precisely when the condition given in Lemma 5 is met. We can see that this holds:
\[
\frac{1}{\sqrt{1/k^2 + (\sqrt{2\xi/M\alpha})^2}} \geq k/\sqrt{2} > \eta_r(L^*(kA/r)) \tag{20}
\]
Therefore, the distribution in equation 16 is the distribution expected by the MIMO oracle and we can recover the vector \( x \).

In order to relate the MIMO problem to standard lattice problems, we need the following lemma, given in [13], which uses a quantum computer.

**Lemma 8.** [13 Lem. 3.14], BDD\(_{L,\sigma,k}\) to \( D_{L^*,\sqrt{\pi}/(\sqrt{2}d)} \). There exists an efficient quantum algorithm that, given any \( n \)-dimensional lattice \( L \), a number \( d < \lambda_1 (A)/2 \), and an oracle that solves BDD\(_{L,d} \), outputs a sample from \( D_{L^*,\sqrt{\pi}/(\sqrt{2}d)} \).

In order for the reduction to hold, we must have \( d > \sqrt{n}/\sqrt{2} \), or the Gaussian distribution will actually grow in each iteration of the procedure. We show this in the following claim.

**Claim 7.** The inequality \( M\sigma/k^2 \sqrt{2} > \sqrt{n}/\sqrt{2} \) is true given the constraints stated in Theorem 1.

**Proof:** We first set the constraint that \( m\alpha/k^2 > \sqrt{n} \), and thus we only require that \( M\sigma/mr > 1 \). We can see this is true:
\[
\frac{M\sigma}{mr} > \frac{2n \log \log n/\log n}{\lambda_1} > \frac{2n \log \log n/\log n}{\lambda_1} > 1 \tag{21}
\]
Here, the first step simply applies the bound given on \( M/m \) in Theorem 1, and the second step follows from applying the bounds on \( r \) and \( \sigma \) from Lemma 7 and Claim 6. The final step follows from applying in equation 13 on the quantity \( \lambda_1/\lambda_2 \).

Finally, we require the following lemma from [13 Sec. 3.3], which uses both the BDD and DGS oracles iteratively to solve standard lattice problems. By setting \( d = M\sigma/k^2 \sqrt{2} \), we can iterate between the BDD and the DGS oracles, shrinking the Gaussian distribution with each step, to a limit, and the stated standard lattice problems.

**Lemma 9.** DGS\(_{L,\sqrt{\pi}/(\sqrt{2}d)} \) to standard lattice problems. For \( n \)-dimensional lattice \( L(A), \alpha > 0, m > 0, k \in \mathbb{R} \) such that \( m\alpha/k^2 > \sqrt{n} \), and \( M > m 2^n \log \log n/\log n \). Given an oracle which solves BDD\(_{L^*,d} \) and DGS\(_{L^*,\sqrt{\pi}/(\sqrt{2}d)} \), then there exists an efficient algorithm that given an \( n \)-dimensional lattice \( L(A) \), solves the problems GapSVP\(_{n/\alpha} \) and SIVP\(_{n/\alpha} \).

**Appendix B**

**Complexity of Lattice Problems**

The security of any lattice-based cryptosystem is based on the presumed hardness of lattice problems. In this subsection we limit our discussion to the GapSVP\(_{\gamma} \) and SIVP\(_{\gamma} \) problems.

One well-known algorithm for solving lattice problems is the LLL algorithm [32]. The algorithm solves SIVP\(_{\gamma} \) and can be adapted to solving many other lattice problems as well.

The algorithm runs in polynomial time but only achieves an approximation factor of \( O(2^n) \). There have been a number of improvements to this algorithm, such as Schnorr’s algorithm [36] but none that achieve small approximation factors to within even a polynomial factor of \( n \) that run in polynomial time. All known algorithms that return exact solutions to lattice problems in fact require a running time on the order of \( 2^n \), see for example [11] or [12]. The hardness of these problems leads to the following conjecture, stated in [3] Conj. 1.1:

**Conjecture 1.** There is no polynomial time algorithm that approximates lattice problems to within an approximation factor that is within a polynomial factor of \( n \).

Within certain approximation factors, the complexity class of solving lattice problems is known. It is NP-Hard to approximate GapSVP to within constant factors [5], [8]–[10]. For a factor of \( \sqrt{n} \), it belongs to the class NP\( \cap \)CoNP [6], [7]. It should be noted that our results are based on an approximation factor of \( n/\gamma \). While such a strong hardness result is not known for this regime, constructing algorithms to achieve such approximation factors within polynomial time seems to be out of reach. These results are summarized in Table III.

| \( \gamma \) | Hardness | Reference |
|---|---|---|
| \( O(1) \) | NP-Hard | [5] [8] [9] [10] |
| \( \sqrt{n} < \gamma < n \) | NP\( \cap \)CoNP | [6] [7] |
| \( \gamma \approx n \) | - | [13] [15] — This work |
| \( 2^{0.5 \log \log n/\log n} n \) | P | [23] [39] |

Another interesting result related to the hardness of lattice problems considers quantum computation. If a quantum computer were to be realized, this could have profound implications for the field of cryptography. A quantum computer could efficiently factor numbers and solve the discrete-logarithm problem, which would allow for virtually all key-exchange protocols to be broken in polynomial time [18]. In addition, algorithms such as Grover’s search algorithm could improve
exhaustive searches by a factor of a square root, weakening the security of systems like AES (the Advanced Encryption Standard, based on the Rijndael cipher) [17]. There are currently no known quantum algorithms that perform significantly better than the best known classical algorithms. However, it should be noted that quantum algorithms are far less understood and studied than classical algorithms, and thus we have less assurance that such an algorithm does not exist. We will still state the following conjecture, given in [3] Conj. 1.2, but note that this is a weaker conjecture than Conjecture 1:

**Conjecture 2.** There is no polynomial time quantum algorithm that approximates lattice problems to within polynomial factors.

Since we show in our main result that the hardness of decoding MIMO is based on lattice problems, this implies that MIMO decoding cannot be performed any faster on a quantum computer. In more general terms, this means that lattice-based cryptography currently provides promise for efficient constructions of classical cryptosystems that are secure against quantum computers.

Besides being conjectured to be hard, many lattice problems have an additional property that makes them attractive to cryptographers: for certain problems, there are connections between the average-case and worst-case complexities. This allows for the construction of systems which are based on robust proofs of security. This property and its significance is described next.

The worst-case complexity refers to the complexity of solving the problem for the worst possible input of a fixed size; whereas average-case complexity of a problem refers to the average complexity of solving a problem given some underlying distribution of inputs of a fixed size (typically uniformly random over all possible inputs). A worst-to-average case reduction gives a distribution of inputs for which the average complexity of solving a problem is as hard as the worst case complexity (potentially of a different problem). The connections between worst- and average-case complexity of certain lattice problems was first found by Ajtai [5]. Ajtai constructed a function that is one-way (that is, it can be computed in polynomial time, but is hard to invert) on average based on the worst-case hardness of lattice problems. This result was used by Ajtai and Dwork to construct a cryptosystem [35]. These worst-to-average case reductions were extended by many, but most important to this paper is the reduction found in [34].

Basing cryptographic systems on problems where a worst-to-average case reduction exists is an extremely strong guarantee of security. It means that it is at least as hard to break the cryptosystem as it is to solve any instance of the related problems. Such a strong guarantee is not provided by most cryptosystems today. For example, breaking a cryptosystem that is based on factoring (e.g. RSA) only implies a solution to factoring numbers of a specific form; namely, the specific form used to generate RSA keys and not a solution to worst-case factoring problems.

Our scheme does not have an average-to-worst case reduction, and it is not clear that such a reduction is possible. This means that there may certain structured inputs which may make MIMO decoding easy – for example consider the case that \( x \) is sparse. Finding a solution for decoding MIMO for sparse signals, would certainly not lead to a solution to approximating all lattice problems to within linear factors. A general polynomial-time algorithm that decodes all MIMO signals under the conditions given in this paper, would, however lead to a solution to approximating lattice problem, and this is unlikely.

**ACKNOWLEDGEMENT**

The authors would like to thank Dan Boneh for his discussions on lattice-based cryptography, Martin Hellman for his comments on a preliminary version of this work, Shlomo Shamai for discussions regarding information-theoretic secrecy in the context of our model, Mainak Chowdhury for discussions on algorithms for MIMO decoding and linear codes, and Jonathan Morin for his useful discussions and comments on a preliminary version of this work.

**REFERENCES**

[1] J.Katz, and Y. Lindell. *Introduction to Modern Cryptography*. New York: Chapman & Hall/CRC. 2007.

[2] M. O. Damen, H. El Gamal, and G. Caire, “On maximum likelihood detection and the search for the closest lattice point“, *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp.2389–2402, 2003.

[3] D. Micciancio and O. Regev, “Lattice-based Cryptography,” in Post-Quantum Cryptography, Berlin, German: Springer Berlin/Heidelberg, pp. 147–191, 2009.

[4] D. Micciancio and S. Goldwasser, *Complexity of Lattice Problems: A Cryptographic Perspective*. Boston, MA: Kluwer Academic Publishers, Mar. 2002.

[5] M. Ajtai, “Generating hard instances of lattice problems”, *Complexity of computations and proofs*, vol. 13 of Quad. Mat., pp. 1–32, 2004. Preliminary version in STOC 1996.

[6] O. Goldreich and S. Goldwasser, On the limits of nonapproximability of lattice problems. Journal of Computer and System Sciences, 60(3):540-563, 2000. Preliminary version in STOC 1998.

[7] D. Aharovon and O. Regev. Lattice problems in NP intersect coNP. *J. of the ACM*, 52(5):749–765, 2005. Preliminary version in FOCS 2004.

[8] S. Khot. Hardness of approximating the shortest vector problem in lattices. In *Proc. 45th Annual IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 126–135, 2004.

[9] I. Haviv and O. Regev. Tensor-based hardness of the shortest vector problem to within almost polynomial factors. In *Proc. 39th ACM Symp. on Theory of Computing (STOC)*, pp. 469–477, 2007.

[10] D. Micciancio. The shortest vector in a lattice is hard to approximate to within some constant. *SIAM J. on Computing*, 30(6): 2008–2035, 2001.

[11] D. Micciancio, and P. Voulgaris, Faster exponential time algorithms for the shortest vector problem. *Proc. of the Twenty-First Annual ACM-SIAM Symp. on Discrete Algorithms*, pages 1468–1480, 2010.

[12] M. Ajtai, R. Kumar, and D. Sivakumar. A sieve algorithm for the shortest lattice vector problem. *Proc. of the ACM Symp. on the Theory of Computing*, number 33, pages 601-610, 2001.

[13] O. Regev, “On lattices, learning with errors, random linear codes, and cryptography”, *Proc. 37th ACM Symp. on Theory of Computing (STOC)*, pp. 84–93, 2005.

[14] R. Rivest, A. Shamir, and L. Adleman, “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems”, *Communications of the ACM* 21 (2): 120–126, 1978.

[15] W. Diffie, and M. Hellman, "New directions in cryptography". *IEEE Transactions on Information Theory* 22 (6): 644–654, 1976.

[16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, MA: Cambridge Univ. Press, 2000.

[17] L. K. Grover, Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79(2):325, 1997.

[18] P. W. Shor, Polynomial-time algorithm for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comp.*, 26(5):1484–1509, 1997.
