A CCA2 secure Code based encryption scheme in the Standard Model

Preetha Mathew K, C. Pandu Rangan Sachin Vasant, Sridhar Venkatesan

***Abstract*—This paper proposes an encryption scheme secure against chosen cipher text attack, built on the Niederreiter encryption scheme. The security of the scheme is based on the hardness of the *Syndrome Decoding* problem and the *Goppa Code Distinguishability* problem. The scheme uses the techniques provided by Peikert and Waters using the lossy trapdoor functions. Compared to the existing IND-CCA2 secure variants in standard model due to Dowsley et.al. and Freeman et. al. (using the κ repetition paradigm initiated by Rosen and Segev), this scheme is more efficient as it avoids κ repetitions.**

***Index Terms*—Standard Model, CCA-2 security, Neiderreiter Cryptosystem, Syndrome Decoding, Code Indistinguishability.**

I. Introduction

The strongest and commonly accepted notion of security for a public key encryption system is that of indistinguishability of messages under adaptive chosen ciphertext attacks. This is denoted as IND-CCA2. Security in this setting means that an adversary obtains no information about encrypted messages provided, the corresponding ciphertexts are not permitted to query the decryption oracle. Chosen-cipher text security, which guarantees confidentiality of encrypted messages against any adversary with polynomially bounded computing power, even though the decryption oracle is provided, has become the de-facto standard notion of security for public key encryption under active attacks.

There are two models based on which the security of the cryptosystem can be argued, namely (1) Random Oracle Model and (2) Standard Model. Random oracle model guarantees security in an idealized world where all parties get black box access to a truly random function. Therefore a proof in random oracle model can be a heuristic argument.The schemes that are provably secure under standard model (without random oracle) is highly prefered.

In the paper lossy trapdoor functions and its applications [12] the authors presented a black box construction of IND-CCA2 secure encryption scheme based on lossy TDFs and all-but-one trapdoor functions, with a witness recovering decryption algorithm. The decryption first recovers the randomness that was used to create the ciphertext, and then tests the validity of the ciphertext simply by re-encrypting the message under retrieved randomness. This paper investigates the usage of code-based assumption for the witness recovery, to obtain a IND-CCA2 secure encryption scheme in the standard model.

Code-based cryptography was initiated by the seminal paper due to McEliece [13], who presented a cryptosystem, based on the hardness of both the *Bounded Decoding* problem and the *Goppa Code Distinguishability* problem. Initially, the scheme did not fascinated the crypto community because of the large public and private key-sizes. Niederreiter proposed a cryptosystem, that is dual of the McEliece Cryptosystem [8]. The number-theoretic schemes that shown to be weak against an attack due to Shor [15], and McEliece and Niederreiter cryptosystems are found to have resistance against such attacks (when using Goppa codes). Therefore the above systems are considered as the candidates for future cryptographic systems that will resist quantum computer attack, which are termed as *Post-Quantum Cryptography*.

Related Work: Recently Dowsley et al. [3] showed that a randomized version of the McEliece cryptosystem is κ-repetition CPA secure and obtain a CCA2 secure scheme in the standard model. Rosen et al. [14] initiated the study of the one-wayness under correlated products and Freeman et al. [7] propose instantiation of lossy trapdoor functions and correlation-secure trapdoor functions. They proposed a correlation-secure trapdoor functions based on the hardness of syndrome decoding, thereby, obtaining a CCA-2 secure encryption scheme in the standard model.

Our Contributions: The currently existing variants of the Niederreiter and McEliece cryptosystem that are IND-CCA2 secure in the standard model [3], [7], are based on the κ repetition paradigm [14]. Such cryptosystems lead to extremely large keys, encrypting time and ciphertext size. The scheme by Freeman et al. [7], the parameters are chosen using the deterministic version of Niederreiter. The parameters that are generally used in deterministic Niederreiter-type cryptosystems are vulnerable to an attack proposed by Håstad, [9]. Hence, the parameters that are generally used for the above construction, requires a large (n, k) resulting in large key-sizes. If randomised version of Neiderreiter is used, the κ-repetition paradigm is required, which need 2κ pairs of keys. Therefore, the cryptosystem due to κ repetition of Niederreiter the key size is very very large and is similar to the system proposed in [3].

Our scheme uses several ideas from [12]. We use strongly unforgeable one-time signature (OTS) to handle malleability related issues as in [12], [14]. Also there are two injective functions on the verification key and the message. This novel approach leads to the elimination of the κ-repetition. Note that the instantiation of the protocol in [12] in a direct way leads to the scheme similar to [3] and it will involve κ-repetition. Thus, for practical code-based cryptosystems such a $\kappa-$ repetition

This work is partially funded by Indian Statistical Institute, Chennai.

DOI: 10.5176_2010-2283_2.1.144

paradigm needs to be avoided.
Our contributions in this paper are:
(i) Proposal of efficient variant of the Niederreiter scheme, that are not based on the κ repetition paradigm, and
(ii) formal argument of their security against IND-CCA2 adversary in the standard model.
An analogous idea for selective provision of trapdoor was also used by Agrawal et. al. [1] in the lattice-based setup, for simulation of the key-extraction phase in their proof of CPA security of a (H)IBE in the standard model. They use lattices built from two parts called right and left lattices. A trapdoor for the left lattices is used as the master secret in the real system and enables one to generate private keys for all identities. A trapdoor for the right lattice is only used to generate private keys for all identities. A trapdoor for the right lattice is only used in the proof of selective security and enables the simulator to generate private keys for all identities except for one. They used right lattices to achieve the targetted ID method of proving, where the key extraction simulation extracts private keys for all IDs except the targetted ID. In our case, $f_1,\ f_2$ are the two injective functions that achieve the same purpose, but the details and computations are entirely different from [1].
Organization of the paper: Section 2 provides the hardness assumptions used in the paper and the basic code-based cryptosystems (McEliece and Niederreiter). Section 3 gives the proposed scheme, the proof of security, the secure parameters for the cryptosystems, and the comparison with existing schemes. The paper is concluded in section 4.

II. Preliminaries

A. Notation

If x is a string, then $|x|$ denotes its length, while $|S|$ represents the cardinality of the set S. If $n \in \mathbb{N}$ then 1^n denotes the string of length n. $s \in_R S$ denotes the operation of choosing an element s from a set S uniformly at random. $w \leftarrow \mathcal{A}(x, y, ...)$ represents the running of algorithm $\mathcal{A}$ with inputs $x, y, ...$ and producing output w. We write $w \leftarrow \mathcal{A}^{\mathcal{O}}(x, y, ...)$ for representing an algorithm $\mathcal{A}$ having access to oracle $\mathcal{O}$. We denote by Pr[E] the probability that the event E occurs. For a matrix M, its transpose is represented by M^T, its inverse is represented by M^{-1}. If a and b are two strings of bits, we denote their bitwise XOR by $a \oplus b$. $\mathcal{U}_n$ is an oracle that return a random element of $\{0,1\}^n$.
Since, the proposed cryptosystems are code-based, a few notations regarding coding theory are introduced. A binary linear-error correcting code of length n and dimension k or a $[n, k]$- code is a *k-dimensional* subspace of $\mathbb{F}_2^n$. If the minimum hamming distance between any two codewords is d, then the code is a $[n, k, d]$ code. The hamming weight of a codeword x, $\mathsf{wt}(x)$, is the number of non-zero bits in the codeword. For $t \leq \lfloor \frac{d-1}{2} \rfloor$, the code is said to be t-error correcting if it detects and corrects errors of weight at most t. The generator matrix $G \in \mathbb{F}_2^{k \times n}$ of a $[n, k]$ linear code C is a matrix of rank k whose rows span the code C. The parity-check matrix $H \in \mathbb{F}_2^{n-k \times n}$ of a $[n, k]$ code C is a matrix satisfying $HG^T = 0$. Hence, code C can be defined as $\{mG : \forall m \in \mathbb{F}_2^k\}$ or $\{c : Hc^T = 0\}$.

B. Definition of the Security Notions

The IND-CCA2 security for any Public-Key Encryption Scheme (*PKE*) is defined as follows:

Definition 1: (IND-CCA2 security). For a two -stage adversary $\mathcal{A}$ = ($\mathcal{A}_1$,$\mathcal{A}_2$) against *PKE* we associate the following experiment $\mathsf{Exp}^{cca2}_{\mathsf{PKE},\mathcal{A}}(n)$:

(pk,sk) $\leftarrow$ *Gen(1^n)*
(m$_0$, m$_1$, state) $\leftarrow \mathcal{A}_1^{\mathsf{Dec}(\mathsf{sk},\cdot)}$*(pk)* s.t. $|m_0| = |m_1|$
b $\in_R$ *(0,1)*
c* $\leftarrow$ *Enc(pk,m$_b$)*
$b' \leftarrow \mathcal{A}_2^{\mathsf{Dec}(\mathsf{sk},\cdot)}$ (c^*,$state$)
if b=b$'$ return 1 else return 0

The adversary $\mathcal{A}_2$ is not allowed to query *Dec(sk,·)* with *c**. We define the advantage of $\mathcal{A}$ in the experiment as

$$\mathit{Adv}^{cca2}_{\mathit{PKE},\mathcal{A}}(n) = |Pr[\mathit{Exp}^{cca2}_{\mathit{PKE},\mathcal{A}}(n) = 1] - \tfrac{1}{2}|$$

We say that *PKE* is indistinguishable against adaptive chosen-cipher text attacks (IND-CCA2) if for all probabilistic polynomial time (PPT) adversaries $\mathcal{A}$ = ($\mathcal{A}_1$,$\mathcal{A}_2$) that makes a polynomial number of oracle queries the advantage of $\mathcal{A}$ in the experiment is a negligible function of n.
The security notion of One-time strongly unforgeable, or one-time existentially unforgeable under chosen message attack (EUF-1CMA) is as follows (based on [11]):

Definition 2: A signature scheme is said to secure under EUF-1CMA, if there exists no PPT algorithm $\mathcal{A}$, which has knowledge of only the verification key vk and the public parameters and access for just one query to the signature oracle to obtain a tuple(m', σ') , outputs a valid signature $(m, \sigma) \neq (m', \sigma')$ with a non-negligible probability.

C. Security assumptions

The following are some of the hard problems on which the security of the proposed cryptosystems is based.

Definition 3: ***Syndrome Decoding Problem.*** For some parameters $[n, k, 2t + 1]$ given an $a \in \mathbb{F}_2^{n-k}$ and a matrix $H \in \mathbb{F}_2^{n-k \times n}$, find a vector $e \in \mathbb{F}_2^n$ with weight $\mathsf{wt}(e) \leq t$ such that $He^T = a$.
The advantage of a PPT algortihm $\mathcal{D}$ of solving the problem is denoted by $\mathsf{Adv}^{\mathsf{SD}}_{\mathcal{D}}(n)$.

Assumption 1: For any probabilistic polynomial time algorithm $\mathcal{F}$, $\mathsf{Adv}^{\mathsf{SD}}_{\mathcal{F}}(C) < \epsilon_1(n, k)$ where $\epsilon_1(n, k)$ is a negligible value with respect to n and k.
For Goppa codes, there is a polynomial time bounded decoding/syndrome decoding algorithm. Thus, there is a preference for most code-based cryptosystems to use the Goppa code as a trapdoor.

Definition 4: ***Goppa code-distinguishability.*** For parameters $[n, k, 2t + 1]$ given a matrix $H \in \mathbb{F}_2^{n-k \times n}$, output 1 if H is a parity check matrix of a Goppa code, 0 if H is not a parity check matrix of any Goppa code.

The advantage of a PPT algorithm $\mathcal{D}$ of solving the problem is denoted by $\mathsf{Adv}^{\mathsf{CD}}_{\mathcal{D}}(n,k)$.

Assumption 2: For any probabilistic polynomial time distinguisher $\mathcal{D}$, $\mathsf{Adv}^{\mathsf{CD}}_{\mathcal{D}}(n,k) < \epsilon_2(n,k)$ where $\epsilon_2(n,k)$ is a negligible function if it is not a high rate goppa code, [4].

$$|Pr[\mathcal{D}(H)=1] - Pr[\mathcal{D}(M)=1]| < \epsilon_2(n,k)$$

where H is the parity check matrix of the Goppa code and $M \in_R \mathbb{F}_2^{n-k\times n}$.

D. Niederreiter's Cryptosystem

Niederreiter's cryptosystem [8] uses the hardness of *syndrome decoding* for its security. The scheme is given below,

- **Secret Key:**
  - C a binary t error correcting linear code.
  - a $(n-k)\times(n-k)$ non-singular matrix Q,
  - a $n\times n$ permutation matrix P
- **Public key:** $\widetilde{H} = QHP$, where H is a parity check matrix of C.
- **Encryption:** $c = \widetilde{H}m^T$, the message m is a word of length n and weight t.
- **Decryption:** $m = P^{-1}\mathsf{Decode}_H(Q^{-1}c)$.

It has been proved that Niederreiter and McEliece cryptosystem have equivalent security properties [10].

III. Variant of the Niederreiter Cryptosystem

A. Proposed Scheme

The proposed system is based on the construction presented by Peikert and Waters [12] for CCA-2 secure encryption scheme in the standard model, using lossy trapdoors and all-but-one trapdoor functions. This scheme uses two functions that are injective Trapdoor functions (TDFS) which require the decoding trapdoor for parity check matrices of the Goppa codes.

The use of a one-time signature for non-malleability of the ciphertext is a paradigm initiated by Dolev et al. [2]. Thus our scheme uses the following:

- A one-time strongly unforgeable signature scheme $\mathcal{OS}(\mathsf{KeyGen}_{\mathcal{OS}}, \mathsf{Sign}_{\mathcal{OS}}, \mathsf{Verfiy}_{\mathcal{OS}})$, for the security parameter κ such that
  - $\mathsf{KeyGen}_{\mathcal{OS}}(\kappa)$ outputs the key-pair $(\mathsf{vk},\mathsf{sk})$, where vk is the verifcation key, and sk is signing key. The size of the domain of the keys $|\mathsf{vk}|, |\mathsf{sk}| = \kappa$
  - $\mathsf{Sign}_{\mathcal{OS}}(\mathsf{sk}, M)$ outputs a signature σ on a message M
  - $\mathsf{Verfiy}_{\mathcal{OS}}(\mathsf{vk}, M, \sigma)$, verifies the signature σ on message M using the verification key vk. If the signature is valid then the verification algorithm outputs VALID, else it outputs INVALID.
- The injective functions $f_1,\ f_2$, map the verification key, to corresponding matrices. With the use of such functions, we are in possession of a tool that makes the ciphertext dependent on the verification key also, but without compromise in security. Such a setup, does away with most malleability issues.
- The private-keys with regard to $\widetilde{H}_2$ & $\widetilde{H}_3$ are used in the decryption oracle, just as in [12].

A more formal description of the scheme is as follows:

System Parameters. The system paramters are as follows: Let D_{vk} denotes the domain of the signature and verification keys.

- Parameters of the code n,k,t for any $[n,k,2t+1]$ linear code, with n,k determined by the security parameter κ, and $t = \frac{n-k}{\log_2 n}$.
- A hard-core predicate $h : \mathbb{F}_2^n \to \{0,1\}^l$, where l is the length of the message, and $l < n$,
- A one-time strongly unforgeable signature scheme $\mathcal{OS}(\mathsf{KeyGen}_{\mathcal{OS}}, \mathsf{Sign}_{\mathcal{OS}}, \mathsf{Verfiy}_{\mathcal{OS}})$, for the security parameter κ.
- An injective function $f_1 : D_{\mathsf{vk}} \to \mathbb{F}_2^{n-k\times n-k}$ which takes verification key as input, and gives the random matrix Q_2 that blinds the parity check matrix in the Niederreiter cryptosystem.
- An injective function $f_2 : D_{\mathsf{vk}} \to \mathcal{P}_{n\times n}$ which takes the verification key as input and output a $n\times n$ permutation matrix P_2, for the Niederreiter cryptosystem.

Key Generation. For the security parameter 1^κ, the KeyGen is as follows:

- $(\mathsf{vk}^*, \mathsf{sk}^*) \leftarrow \mathsf{KeyGen}_{\mathcal{OS}}(1^\kappa)$.
- Randomly select two distinct $[n,k,2t+1]$Goppa codes,and a a $[n,k,2t+1]$ linear code whose parity check matrices are H_1, H_3, R respectively.
- Randomly select $Q_1, Q_3 \in_R \mathbb{F}_2^{n-k\times n}$.$Q_3$ should be invertible,$P_1 \in_R \mathbb{F}^{n\times n}$.
- Define $\widetilde{H}_1 = Q_1H_1P_1$,$\widetilde{H}_2 = Q_2H_2P_2 \oplus R$, $\widetilde{H}_3 = Q_3R$.
- Compute $Q_2 = f_1(\mathsf{vk}^*)$ and $P_2 = f_2(\mathsf{vk}^*)$,.

Thus, we have :

- **Public Keys:** $\widetilde{H}_1,\ \widetilde{H}_2,\ \&\ \widetilde{H}_3$.
- **Secret Keys:** $H_1,\ Q_1,\ P_1,\ (\mathsf{vk}^*, \mathsf{sk}^*)$(hence $Q_2,\ P_2),\ H_2,\ R,\ \&\ Q_3$

Encryption: On an message $m \in \{0,1\}^l$, the following steps constitute the encryption algorithm:

- Generate $r \in_R \mathbb{F}_2^n$, with $wt(r) \le t$.
- $(\mathsf{vk},\mathsf{sk}) \leftarrow \mathsf{KeyGen}_{\mathcal{OS}}(1^\kappa)$, and compute $Q_{\mathsf{vk}} = f_1(\mathsf{vk})$ and $P_{\mathsf{vk}} = f_2(\mathsf{vk})$.
- Define $K_1 = Q_{\mathsf{vk}}\widetilde{H}_1P_{\mathsf{vk}}$, $K_2 = Q_{\mathsf{vk}}\widetilde{H}_2P_{\mathsf{vk}}$ and $K_3 = Q_{\mathsf{vk}}\widetilde{H}_3P_{\mathsf{vk}}$.
- Define $c_1 = K_1r^T, c_2 = K_2r^T, c_3 = K_3r^T$ and $c_4 = m \oplus h(r)$, where $h(r) \neq 0$.
- Compute $\sigma = \mathsf{Sign}_{\mathcal{OS}}(\mathsf{sk}, (c_1,c_2,c_3,c_4))$, i.e., the one-time signature on (c_1,c_2,c_3,c_4) (where (c_1,c_2,c_3,c_4) is denoted as M) using the signing key sk.

The ciphertext that is sent is $c = (\mathsf{vk}, c_1, c_2, c_3, c_4, \sigma)$.

Decryption. The decryption on the ciphertext $c = (\mathsf{vk}, c_1, c_2, c_3, c_4, \sigma)$ is done as follows:

if $(\mathsf{Verify}_{\mathcal{OS}}(\mathsf{vk}, (c_1,c_2,c_3,c_4), \sigma) \to \text{INVALID}))$
  return $\perp$
else
  Compute, $Q_{\mathsf{vk}} \leftarrow f_1(\mathsf{vk})$, $P_{\mathsf{vk}} \leftarrow f_2(\mathsf{vk})$.

if ($\mathsf{Decode}_{H_1}((Q_{\mathsf{vk}}Q_1)^{-1}c_1) \rightarrow \perp$)
return $\perp$.
else
$r \leftarrow P_{\mathsf{vk}}^T P_1^T \mathsf{Decode}_{H_1}((Q_{\mathsf{vk}}Q_1)^{-1}c_1)$
if($c_2 \neq Q_{\mathsf{vk}}\widetilde{H}_2 P_{\mathsf{vk}} r^T$ OR $c_3 \neq Q_{\mathsf{vk}}\widetilde{H}_3 P_{\mathsf{vk}} r^T$)
return $\perp$.
else
return $m = c_4 \oplus h(r)$
end

Correctness The reciever on getting the cipher text can verify the signature as the the verification key is attached along with the cipher text components, then decode the randomness with which the message is encrypted. The receiver can verify the consistency of the retrieved randomness using the components c_2, c_3. The message can be retrieved by $m = c_4 \oplus h(r)$

B. Proof for the Security of the system

The proof of security follows a game-based approach.It is claimed that every adversary has only a negligible advantage in the CCA-2 games under the standard model, provided the *Computational Syndrome Decoding* problem and *Goppa Code Distinguishability* are hard to solve, and the signature is one-time strongly unforgeable.

Theorem 1: The probability for any PPT adversary $\mathcal{A}$ of winning the $IND-CCA2$ under the standard model for the Niederreiter variant is within the range of,

$$\frac{1}{2} \pm \{\mathsf{Adv}^{cca2}_{\mathsf{PKE},\mathcal{A}}(n,k) + \frac{1}{2^\kappa} + Adv^{\mathsf{CD}}_{\mathcal{A}}(n,k) + Adv^{\mathsf{SD}}_{\mathcal{A}}(n,k) + Adv^{\mathsf{SD}}_{\mathcal{A}}(n,3k-2n)\}$$

provided the signature is one-time strongly unforgeable and the functions f_1 and f_2 are injective.

Proof: We build the proof as a sequence of games **Game0, Game1,**$\cdots$, where **Game0** is the IND-CCA2 game, as directly applied to the given scheme. Succesive games are obtained by small modifications of the preceding games, in such a way that the difference of the adversarial advantage in consecutive games is easily quantifiable. For the proof we assume two matrices Y and R, for which the syndrome decoding trapdoor is not known to the challenger.The Setup always first chooses a one-time signature key pair $(\mathsf{vk}^*, \mathsf{sk}^*)$ and in the Challenge using the encryption oracle it uses the one-time signature pair $(\mathsf{vk}^*, \mathsf{sk}^*)$ instead of the pair $(\mathsf{vk}, \mathsf{sk})$.The games are mentioned as follows:

Game 0. The Setup, Decryption and Challenge are generated as per the defined IND-CCA2 experiment. Setup uses, the setup algorithm as mentioned in the proposed scheme, to generate the public keys $\widetilde{H}_1, \widetilde{H}_2, \widetilde{H}_3$, and the secret keys $H_1, Q_1, P_1, \mathsf{vk}^*, \mathsf{sk}^*, R, Q_3$. The decryption oracle uses the decryption algorithm, making use of the decoding trapdoor corresponding to $\widetilde{H}_1$. The challenge ciphertext is generated by the challenger $\mathcal{C}$ on m_b, with $b \in_R \{0,1\}$, using the encryption algorithm of the proposed scheme, where m_0 & m_1 are sent by the adversary $\mathcal{A}$. Hence, the probability, of success in game 0, is probability of success of breaking the proposed scheme. Let, X_0 be the event that $\mathcal{A}$ wins the game. Then

$$|Pr[X_0] - \frac{1}{2}| = \mathsf{Adv}^{cca2}_{\mathsf{PKE},\mathcal{A}}(n,k) \quad (1)$$

Game 1. **Restriction on the Decryption Oracle.** The decryption oracle is restricted by the Challenger, in the sense that, for the query ciphertext $c = (\mathsf{vk}, c_1, c_2, c_3, c_4, \sigma)$, where σ is a valid signature on c_1, c_2, c_3, c_4, if $\mathsf{vk} == \mathsf{vk}^*$ the Challenger aborts the game. We can assume that, this occurs with a probability of at most $\frac{1}{2^\kappa}$, due to the one-time unforgeability assumption of the signature used. Thus, for the event that $\mathcal{A}$ win the game Game 1,X_1,

$$|Pr[X_1] - Pr[X_0]| \leq \frac{1}{2^\kappa} \quad (2)$$

Game 2. **Simulation of the Decryption Oracle.** In the current game, the decryption oracle is simulated, without using the trapdoor related to $\widetilde{H}_1$. The decryption oracle uses the trapdoor related to the keys $\widetilde{H}_2$ and $\widetilde{H}_3$. The simulation is as follows:

Input: The ciphertext $c = (\mathsf{vk}, c_1, c_2, c_3, c_4, \sigma)$
Output: The message m.
if ($\mathsf{Verify}_{\mathcal{OS}}(\mathsf{vk}, c_1, c_2, c_3, c_4) == FALSE$)
Return $\perp$ /** signature not valid **/
else
if($\mathsf{vk} == \mathsf{vk}^*$)
ABORT /** for challenge cipher creation key is reserved **/
else
$Q_{\mathsf{vk}} \leftarrow f_1(\mathsf{vk})$, $P_{\mathsf{vk}} \leftarrow f_2(\mathsf{vk})$
Compute $y = Q_{\mathsf{vk}}^{-1}c_2 \oplus Q_3^{-1}Q_{\mathsf{vk}}^{-1}c_3$
if($\mathsf{Decode}_{H_2}(Q_2^{-1}y) \rightarrow \perp$)
Return $\perp$
else /** invalid cipher cannot be decrypted **/
Compute $r = P_{\mathsf{vk}}^T P_2^T \mathsf{Decode}_{H_2}((\overline{Q_2}^T)^{-1} y)$
if ($c_1 \neq \widetilde{H}_1 r^T$ OR $c_2 \neq Q_{\mathsf{vk}}\widetilde{H}_2 P_{\mathsf{vk}} r^T$ OR $c_3 \neq Q_{\mathsf{vk}}\widetilde{H}_3 P_{\mathsf{vk}} r^T$)
Return $\perp$
else
Return $m = c_4 \oplus h(r)$.
end

The above decryption oracle, makes use of the decoding trapdoor corresponding to $\widetilde{H}_2$ & $\widetilde{H}_3$. To elucidate on the computations of y and hence r, the elements c_2 is of the form $K_2 r^T$, i.e. $c_2 = (Q_{\mathsf{vk}}Q_2H_2(P_2P_{\mathsf{vk}})r^T \oplus Q_{\mathsf{vk}}RP_{\mathsf{vk}}r^T$, and $c_3 = (Q_{\mathsf{vk}}Q_3)RP_{\mathsf{vk}}r^T$, thus, $y = Q_{\mathsf{vk}}^{-1}c_2 \oplus Q_3^{-1}Q_{\mathsf{vk}}^{-1}c_3 \implies y = Q_2H_2(P_2P_{\mathsf{vk}})r^T$, from which r can be computed as shown in the algorithm.

It can be shown that the decryption oracle in this game exactly simulates the decryption oracle in Game 1. The aborting cases of the game is same

as that of Game 1. Therefore

$$Pr[X_2] = Pr[X_1] \tag{3}$$

Game 3. **Altering the** Setup **algortihm.** For the key generation, $\widetilde{H}_1 = Q_1 Y P_1$, where $Y \in_R \mathbb{F}_2^{n-k\times n}$. Hence, the Goppa code with partiy check matrix H_1 is replaced by a randomly selected matrix Y.The decryption oracle is same as that of Game 2. The, adversary is able to identify the inconsistency, by, the solution of the *Goppa Code distinguishability* problem for H and Y. Let X_3 be the event that the adversary wins Game 3. Then,

$$|Pr[X_3] - Pr[X_2]| \leq Adv_{\mathcal{A}}^{\mathsf{CD}}(n,k) \tag{4}$$

Game 4. **Challenge ciphertext and Challenge** In this case, the selection of the keys of the one-time signature $(\mathsf{vk}, \mathsf{sk})$, is restricted. Instead, of generating $(\mathsf{vk}, \mathsf{sk}) \leftarrow \mathsf{KeyGen}_{\mathcal{OS}}(1^\kappa)$, the challenger $\mathcal{C}$, selects $(\mathsf{vk}, \mathsf{sk}) = (\mathsf{vk}^*, \mathsf{sk}^*)$. The ciphertext is generated randomly as, $c_1, c_2, c_3 \in_R \mathbb{F}_2^{n-k}$ and $c_4 \in_R \{0,1\}^l$. The signature σ is generated on c_1, c_2, c_3, c_4 using the signing key sk^*.

The adversary is not allowed to query the decryption oracle for the challenge cipher text, and he is asked to distinguish between the correctly generated cipher text with the random cipher text. The decryption oracle is same that of Game 3. If adversary succeeds, implies that adversary is able to distinguish betweem a syndrome and a random vector, as the message is masked with the hard-core predicate of the code from the syndromes c_1, c_2, c_3. Therefore the distribution in this game varies only if the adversary is able to distinguish between a randomly generated challenge ciphertext and the ciphertext generated using the encryption oracle (as in Game 3). Fischer and Stern [6] have proven that, as long as syndrome decoding is hard, a syndrome is computationally indistinguishable from a randomly generated vector from the same space. Hence, to distinguish between the syndrome and randomly generated vector, the adversary has to solve the instance of syndrome decoding problem (i.e., the syndrome decoding assumption should not hold for the adversary)or by concatenation of c_1, c_2, c_3 using the key obtained by concatenanating the columns of K_1, K_2, K_3 (a $3(n-k) \times n$ matrix), thus requiring the solution of an instance of $\mathsf{SDP}(n, 3k-2n)$. Hence, for the event of winning the Game4 X_4

$$|Pr[X_4] - Pr[X_3]| \leq Adv_{\mathcal{A}}^{\mathsf{SD}}(n,k) + Adv_{\mathcal{A}}^{\mathsf{SD}}(n, 3k-2n) \tag{5}$$

Also, it can be seen that the challenge ciphertext is completely independent of the target plaintext m_b. Hence,

$$Pr[X_4] = \frac{1}{2} \tag{6}$$

Adding and substituting (1) to (6) we obtain

$$\begin{aligned} \mathsf{Adv}_{\mathsf{PKE},\mathcal{A}}^{cca2}(n,k) &\leq \frac{1}{2^\kappa} + Adv_{\mathcal{A}}^{\mathsf{CD}}(n,k) + Adv_{\mathcal{A}}^{\mathsf{SD}}(n,k) \\ &\quad + Adv_{\mathcal{A}}^{\mathsf{SD}}(n, 3k-2n) \\ \implies Pr[X_0] &\leq \frac{1}{2} + \{\frac{1}{2^\kappa} + Adv_{\mathcal{A}}^{\mathsf{CD}}(n,k) + \\ &\quad Adv_{\mathcal{A}}^{\mathsf{SD}}(n,k) + Adv_{\mathcal{A}}^{\mathsf{SD}}(n, 3k-2n)\} \\ \& \; Pr[X_0] &\geq \frac{1}{2} - \{\frac{1}{2^\kappa} + Adv_{\mathcal{A}}^{\mathsf{CD}}(n,k) + \\ &\quad Adv_{\mathcal{A}}^{\mathsf{SD}}(n,k) + Adv_{\mathcal{A}}^{\mathsf{SD}}(n, 3k-2n)\} \end{aligned}$$

Therefore we get the probability that an adversary wins the game to be in the range ■

From the result, we obtain that the advantage of the adversary, depends on the advantage of solving the *Goppa Code distinguishability* problem and *Syndrome Decoding* problem. For, parameters (n,k) for which $\mathsf{CD}(n,k)$, $\mathsf{SD}(n,k)$, & $\mathsf{SD}(n, 2k-3n)$ are hard, the advantage for the adversary is negligible.

C. Parameters

From, the previous section, we have seen that the selection of parameters is important in defining the negligible advantage an adversary has in solving the syndrome decoding problem. Clearly, for $\mathsf{SD}(n, 3k-2n)$ to be hard, we have to select a $k > \frac{2n}{3}$. Since, the required codes need not have a very high rate, the distinguisher attack [4] does not hold. The table I, presents the (n,k) parameters (where the error-correcting capacity is $t = \frac{n-k}{log_2 n}$), and the binary work factor for syndrome decoding for (n,k) and $(n, 3k-2n)$. Here binary work factor, is $\log_2(time\ taken)$. The work factors are taken according to the lower bound complexity given in [5]. For, the given parameters, Goppa codes are *indistinguishable* [4].

TABLE I
PARAMETERS FOR THE GIVEN SCHEME, AND CORRESPONDING WORK FACTORS FOR SOLUTION OF SYNDROME DECODING PROBLEM

(n,k)	**Security factor for** (n,k)	**Security factor for** $(n, 3k-2n)$
(2048,1696)	86.8	100.79
(4096,3604)	128.5	292.53

D. Comparison with other schemes

The parameters given in the table I, are in fact the parameters that are generally used for any version of the Niederreiter cryptosystem. Hence, it can be seen that the proposed scheme is IND-CCA2 secure in the standard model, without much change in the parameters. The comparison of the proposed schemes with existing schemers are presented in table II.

IV. CONCLUSION

In the paper, we propose an efficient IND-CCA2 secure code-based encryption scheme in the standard model. The scheme is the first such scheme, that does not use the κ repetition paradigm [14]. Thus, the scheme has avoided the

TABLE II
COMPARISON WITH OTHER CODE-BASED CCA-2 CRYPTOSYSTEMS.

Scheme	Public key (bits)	Secret key (bits)	Cipher text	Encrypt Com-plexity	Decrypt com-plexity
Dowsley et al.	$2\kappa\times$NP	$2\kappa\times$NS	$\kappa\times$ NC	$\kappa\times$NE	1 ND
Freeman et al.	$2\kappa\times$MP	$2\kappa\times$MS	$\kappa\times$MC	$\kappa\times$ME	1 MD
Proposed Scheme	$3\times$NP	$3\times$NS	$3\times$NC	$3\times$NE	1 ND

M - McEliece, N - Niederreiter, E - Encryption, D - Decryption
P - public-key size S - secret key size

inherent costs incurred by the existing schemes [3], [7] and is more efficient, because it requires at most three repetitions of the underlying Niederreiter encryption scheme and any one-time strongly unforgeable signature.

REFERENCES

[1] Shweta Agrawal, Dan Boneh, and Xavier Boyen. Efficient lattice (H)IBE in the standard model. In Henri Gilbert, editor, *EUROCRYPT*, volume 6110 of *Lecture Notes in Computer Science*, pages 553–572. Springer, 2010.

[2] Danny Dolev, Cynthia Dwork, and Moni Naor. Nonmalleable cryptography. *SIAM J. Comput.*, 30(2):391–437, 2000.

[3] Rafael Dowsley, Jörn Müller-Quade, and Anderson C. A. Nascimento. A CCA2 secure public key encryption scheme based on the McEliece assumptions in the standard model. In Marc Fischlin, editor, *CT-RSA*, volume 5473 of *Lecture Notes in Computer Science*, pages 240–251. Springer, 2009.

[4] J.-C. Faugére, A Otmani, L. Perret, and J.-P. Tillich. Algebraic Cryptanalysis of McEliece variants with compact keys – toward a complexity analysis. In *SCC '10: Proceedings of the 2nd International Conference on Symbolic Computation and Cryptography*, pages 45–55, RHUL, June 2010.

[5] Matthieu Finiasz and Nicolas Sendrier. Security bounds for the design of code-based cryptosystems. In Mitsuru Matsui, editor, *ASIACRYPT*, volume 5912 of *Lecture Notes in Computer Science*, pages 88–105. Springer, 2009.

[6] Jean-Bernard Fischer and Jacques Stern. An efficient pseudo-random generator provably as secure as syndrome decoding. In *EUROCRYPT*, pages 245–255, 1996.

[7] David Mandell Freeman, Oded Goldreich, Eike Kiltz, Alon Rosen, and Gil Segev. More constructions of lossy and correlation-secure trapdoor functions. In Phong Q. Nguyen and David Pointcheval, editors, *Public Key Cryptography*, volume 6056 of *Lecture Notes in Computer Science*, pages 279–295. Springer, 2010.

[8] Niederreiter H. Knapsack-type cryptosystems and algebraic coding theory. *Prob Contr Inform Theor 15*, pages 159 – 166, 1986.

[9] Johan Håstad. Solving simultaneous modular equations of low degree. *SIAM J. Comput.*, 17(2):336–341, 1988.

[10] Yuan Xing Li, Robert H. Deng, and Xin mei Wang. On the equivalence of McEliece's and Niederreiter's public-key cryptosystems. *IEEE Transactions on Information Theory*, 40(1):271–, 1994.

[11] Rafael Misoczki and Paulo S. L. M. Barreto. Compact McEliece keys from Goppa Codes. In Michael J. Jacobson Jr., Vincent Rijmen, and Reihaneh Safavi-Naini, editors, *Selected Areas in Cryptography*, volume 5867 of *Lecture Notes in Computer Science*, pages 376–392. Springer, 2009.

[12] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In Cynthia Dwork, editor, *STOC*, pages 187–196. ACM, 2008.

[13] McEliece R.J. A public-key cryptosystem based on algebraic coding theory. *JPL DSN Progress Report*, pages 114–116, 1978.

[14] Alon Rosen and Gil Segev. Chosen-ciphertext security via correlated products. In Omer Reingold, editor, *TCC*, volume 5444 of *Lecture Notes in Computer Science*, pages 419–436. Springer, 2009.

[15] Peter W. Shor. Polynominal time algorithms for discrete logarithms and factoring on a quantum computer. In Leonard M. Adleman and Ming-Deh A. Huang, editors, *ANTS*, volume 877 of *Lecture Notes in Computer Science*, page 289. Springer, 1994.

Authors

Preetha Mathew K

Research Scholar in code based cryptosystems in Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu India

Sachin Vasant, Sridhar Venkatesan
Department of Mathematics and Computer Applications, PSG College of Technology, Coimbatore, Tamil Nadu, India

C. Pandu Rangan
Professor in Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu India