Cloud-assisted Asynchronous Key Transport with Post-Quantum Security

Gareth T. Davies[1], Herman Galteland[2], Kristian Gjøsteen[2], and Yao Jiang[2]
[1]Bergische Universität Wuppertal, Germany.
davies@uni-wuppertal.de
[2]Norwegian University of Science and Technology, NTNU, Norway.
{herman.galteland,kristian.gjosteen,yao.jiang}@ntnu.no

December 5, 2019

Abstract

In cloud-based outsourced storage systems, many users wish to securely store their files for later retrieval, and additionally to share them with other users. These retrieving users may not be online at the point of the file upload, and in fact they may never come online at all. In this asynchoronous environment, key transport appears to be at odds with any demands for forward secrecy. Recently, Boyd et al. (ISC 2018) presented a protocol that allows an initiator to use a modified key encapsulation primitive, denoted a blinded KEM (BKEM), to transport a file encryption key to potentially many recipients via the (untrusted) storage server, in a way that gives some guarantees of forward secrecy. Until now all known constructions of BKEMs are built using RSA and DDH, and thus are only secure in the classical setting.

We further the understanding of secure key transport protocols in two aspects. First, we show how to generically build blinded KEMs from homomorphic encryption schemes with certain properties. Second, we construct the first post-quantum secure blinded KEMs, and the security of our constructions are based on hard lattice problems.

Keywords: Lattice-based cryptography, NTRU, Group Key Exchange, Blinded Key Encapsulation, Forward Secrecy, Cloud Storage, Post-quantum cryptography

1 Introduction

Consider the following scenario: a user of a cloud storage service wishes to encypt and share a file with a number of recipients, who may come online to retrieve the file at some future time. In modern cloud storage environments, access control for files is normally done via the storage provider's interface, and the user is usually tasked with performing any encryption and managing the resulting keys. However the users do not trust the server, and in particular may be concerned that key compromise may occur to any of the involved parties at some point in the future – they thus desire some forward secrecy guarantees. A number of approaches can be taken for transporting a (randomly chosen) file encryption key from the initiator to the recipients. The first option is public-key encryption – simply encrypting under each recipient's public key. This approach does not provide any forward secrecy, however if the initiator were to use puncturable encryption then this would provide a (currently inefficient) solution for acheiving forward secrecy. The users could also perform a (necessarily interactive) group key exchange protocol, however this requires all recipients to be online: a disqualifying criterion for many usage scenarios. The challenge of providing efficient key transport that allows asynchronous fetching by the recipients and simultaneously gives some forward secrecy guarantees appears to invoke trade-offs.

Recent work by Boyd et al. [10] (hereafter BDGJ) provided a solution that utilized the high availability of the storage provider. The initiator essentially performs key encapsulation, using an (public) encapsulation key belonging to the server, and sends an encapsulated value (out-of-band) to each recipient. Then, each recipient blinds this value in such a way that when it asks the server to decapsulate, the server does not learn anything about the underlying file encryption key, and the homomorphic properties of the scheme enable successful unblinding by the recipient. This encapsulation-and-blinding procedure was named by the authors as a *blinded KEM* (BKEM), and the complete protocol built from this was named as a cloud-assisted offline group key exchange (OAGKE). Forward secrecy is acheived if the recipients delete their ephemeral values after recovering the file encryption key, and if the server deletes its decapsulation key after all recipients have been online and recovered the file.

A conceptual overview of the construction, which can achieve all these security properties, is described in Figure 1, and we refer to their paper for full details [10]. In the protocol, the server runs the KG and Decap algorithms to help the initiator share file encryption key k. The blinding algorithm Blind, executed by the responder, should prohibit the server from learning any information about the file encryption key. After the server has decapsulated a blinded encapsulation, the responder can use the unblinding algorithm $\mathsf{Unblind}$ to retrieve the file encryption key.

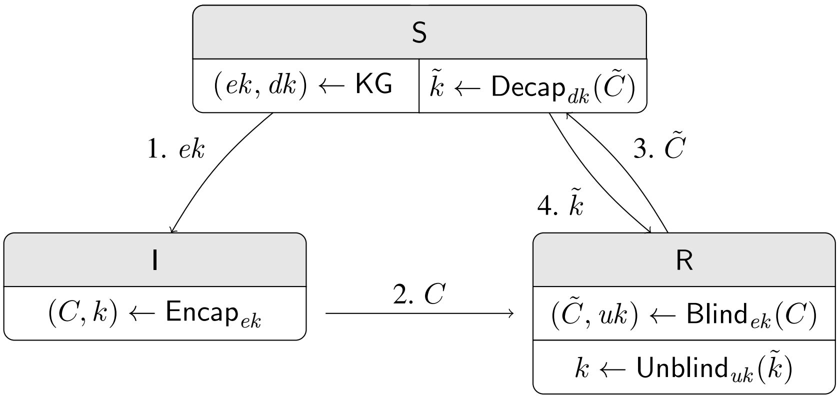
S
$(ek, dk) \leftarrow \mathsf{KG}$
$\tilde{k} \leftarrow \mathsf{Decap}_{dk}(\tilde{C})$
1. ek
3. $\tilde{C}$
4. $\tilde{k}$
I
$(C, k) \leftarrow \mathsf{Encap}_{ek}$
2. C
R
$(\tilde{C}, uk) \leftarrow \mathsf{Blind}_{ek}(C)$
$k \leftarrow \mathsf{Unblind}_{uk}(\tilde{k})$

Figure 1: A simplified overview of an OAGKE protocol [10] between an initiator I, server S and potentially many recipients R (one is given here for ease of exposition), built using a BKEM. File encryption key k is used by I to encrypt one or more files. The numbered arrows indicate the order in which operations occur.

While the approach appears promising, their two constructions – built from DDH and RSA, are somewhat ad hoc, and further do not resist attacks in the presence of quantum computers. In this work we wish to construct a post-quantum secure OAGKE protocol, where we need the individual components – a blinded KEM (parameterized by a homomorphic encryption scheme), a collision resistant hash function, a digital signature scheme, and a key derivation function – to all be post-quantum secure. Acheiving post-quantum security of all components except for the BKEM has been covered extensively in prior work, and thus we focus on finding post-quantum constructions of BKEMs. Much work has been done on constructing regular key encapsulation mechanisms (KEMs) [1,17,18,20,29,31] that are post-quantum secure [8,13,30,34,36], (the ongoing NIST standardization effort [39] specifically asks for KEMs) however BKEMs do not generalize KEMs, since decapsulation operates on blinded ciphertexts.

Providing post-quantum-secure BKEMs invokes a number of technical challenges. The Blind algorithm must modify the file encryption key by incorporating some randomness r, in such a way that after decapsulation (by the server) the recipient can strip off r to recover the file encryption key. In the DDH setting this is straightforward since the recipient can simply exponentiate the encapsulation, and apply the inverse on the received value from the server (the RSA setting is similarly straightforward), and, importantly, the encapsulation (with the underlying file encryption key) *and* multiple blinded samples (each with a value that is derived from the file encryption key) will all look like random group elements. In the security game for BKEMs (as provided by BDGJ), the adversary receives: an encapsulation of a 'real' key, a number of blinded versions of this encapsulation (blinded encapsulations), a number of blinded versions of the 'real' key (blinded keys), and either this 'real' key or a random key, and must decide which it has been given. If the blinded key samples (the $\tilde{k}$s) leak information about the file encryption key then the adversary's task in this game becomes much easier. For example, if the blinding algorithm alters the file encryption key such that the blinded keys are located close to it then exhaustive search becomes possible. We overcome this hurdle by using a big a blinding value to hide the file encryption key. Similarly the blinded encapsulation samples (the $\tilde{C}$s) can leak information about the blinding value used to hide the file encryption key, which can be used to recover the file encryption key. For example, if the blinded encapsulation is a linear combination of the original encapsulation, the blinding value, and some small error then the distance between the blinded encapsulation and the original encapsulation could reveal the blinding value, or a small interval containing it, and therefore the file encryption key. By making sure blinded encapsulations look fresh then all blinded encapsulation samples and the encapsulation looks independent of each other. We use these techniques to provide secure BKEMs built from (a variant of) NTRU [28,40] and ideas from Gentry's FHE scheme [23].

The second shortfall of the work of BDGJ lies in the non-generic nature of their constructions. The two provided schemes appear to have similar properties, yet do not immediately indicate how any further BKEM schemes could be constructed. We show how to generically build BKEMs from homomorphic encryption schemes with minimal properties. This allows us to more precisely cast the desirable properties of schemes used to build BKEMs, generalizing the way that the responder alters the content of an encapsulation (ciphertext) by adding an encrypted random value. Essentially, the resulting blinded ciphertext is an encryption of the sum of a file encryption key and the random value. The server can decrypt the blinded ciphertext to retrieve the blinded key, and then the responder can unblind by removing (subtracting) the random value.

1.1 Related work

Boyd et al. [10] formalized OAGKE and BKEMs, and they provided two BKEM constructions, based on Diffie-Hellman and RSA. To our knowledge these are the only BKEM constructions in the literature.

Many works focused on secure messaging have shown how to perform secure key transport in the presence of pre-keys of the recipients [16,37,41], we wish to avoid this assumption in our system architecture. Puncturable encryption has developed rapidly in recent years [6,21,26,27], however current constructions are still impractical or unsuitable for the cloud-based key transport scenario that we consider.

Gentry introduced the first fully homomorphic encryption (FHE) scheme, based on lattice problems, and gave a generic framework [23]. After Gentry's breakthrough several FHE schemes where constructed following his framework [11,15,22,25], where all of these schemes rely on the learning with errors (LWE) problem. Two FHE schemes based their security on an overstretched variant of the NTRU problem [9,32], however, subfield lattice attacks against this variant was subsequently found [2,14], and consequently these schemes are no longer secure. As a side note, our NTRU based BKEM construction relies on the hardness of the LWE problem.

To make a BKEM from existing post-quantum secure KEM schemes we need, for each individual scheme, a method for altering the encapsulations in a predictable way. Most of the post-quantum secure KEM schemes submitted to NIST are built from a PKE scheme, where we can use our techniques to make a BKEM if the PKE scheme supports one homomorphic operation. FrodoKEM is the only submission that advertises its additive homomorphic properties of its FrodoPKE scheme [3]. Other submissions based on lattices [33], LWE [4,5,19], or NTRU [7,12] are potential candidates for a BKEM construction. Note that the NTRU submission of Chen et al. [12] does not use the Gaussian distribution to sample their polynomials, and NTRU Prime of Bernstein et al. [7] uses a large Galois group to construct their polynomial field, instead of a cyclotomic polynomial. Furthermore, the NTRU contruction of Stehlé and Steinfeld chooses the distribution of the secret keys such that the public key looks uniformly random and they provide a security proof which relies on this.

1.2 Our contribution

Our aim in this work is to further the understanding of blinded KEMs and their possible instantiations, in order to deliver secure key transport protocols in cloud storage environments. Specifically, we provide:

- a generic homomorphic-based BKEM construction, and show that it meets the expected indistinguishability-based security property for BKEMs, under feasible requirements.
- two instantiations of our homomorphic-based BKEM, built from primitives with post-quantum security. The proof chain is as follows.

$$\text{Hard problems} \xrightarrow[\text{or Lyubashevsky et al. [35]}]{\text{Quantum, Gentry [23]}} \textsf{IND-CPA HE} \xrightarrow{\text{This work}} \begin{array}{c}\textsf{IND-secure}\\ \textsf{HE-BKEM}\end{array}$$

1.3 Organization

In Section 2 we provide the necessary background of ideal lattices and the discrete Gaussian Distribution. In Section 3 we formally define BKEM and their security. In Section 4 we construct a generic homomorphic BKEM schemes and analyze its security requirements. In Section 5 we provide two homomorphic-based BKEM constructions and prove that they are secure.

2 Preliminaries

This section introduces terminology and results from [23, 24, 38], and provides an introduction to our notation and building blocks for constructing post-quantum secure homomorphic encryption schemes. Towards the end of this section we detail two specific constructions of post-quantum secure homomorphic encryption schemes [23, 40].

2.1 Notation

Given n linearly independent vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \in \mathbb{R}^m$, the m-dimensional *lattice* L generated by the vectors is $L = \{\sum_{i=1}^n x_i \mathbf{b}_i \mid x_i \in \mathbb{Z}\}$. If $n = m$ then L is a *full-rank n-dimensional lattice*, we will always use full-rank lattices in this paper.

Suppose $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n\}$ is a basis of I, let $\mathcal{P}(\mathbf{B}) = \{\sum_{i=1}^n x_i \mathbf{b}_i \mid x_i \in [-1/2, 1/2), \mathbf{b}_i \in \mathbf{B}\}$ be the half-open parallelepiped associated to the basis $\mathbf{B}$.

Let $R = \mathbb{Z}[x]/(f(x))$ be a polynomial ring, where $f(x)$ is a monic polynomial of degree n. Any ideal $I \subseteq R$ yields a corresponding integer sublattice called *ideal lattice* of the polynomial ring. For convenience, we identify all ideals of R with its ideal lattice.

Let $\|\mathbf{v}\|$ be the Euclidean norm of a vector $\mathbf{v}$. Define the *norm of a basis* $\mathbf{B}$ to be the Euclidian norm of its longest column vector, that is, $\|\mathbf{B}\| = \max_{1 \le i \le n}(\|\mathbf{b}_i\|)$.

For a full-rank n-dimensional lattice L, let $L^* = \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}, \forall \mathbf{y} \in L\}$ denote its *dual lattice*. If $\mathbf{B}$ is a basis for the full-rank lattice L, then $(\mathbf{B}^{-1})^T$ is a basis of L^*. Let $\gamma_\times(R) = \max_{\mathbf{x},\mathbf{y} \in R} \frac{\|\mathbf{x} \cdot \mathbf{y}\|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$ be the *multiplicative expansion factor*.

For $\mathbf{r} \in R$, define $\mathbf{r} \mod \mathbf{B}$ to be the unique vector $\mathbf{r}' \in \mathcal{P}(\mathbf{B})$ such that $\mathbf{r} - \mathbf{r}' \in I$. We call $\mathbf{r} \mod \mathbf{B}$ to be the *distinguished representative* of the coset $\mathbf{r} + I$. Denote $R \mod \mathbf{B} = \{\mathbf{r} \mod \mathbf{B} \mid \mathbf{r} \in R\}$ to be the set of all distinguished representatives in R, this set can be chosen to be the same as the half-open parallelepiped $\mathcal{P}(\mathbf{B})$ associated to the basis $\mathbf{B}$. For convinience we treat $R \mod \mathbf{B}$ and $\mathcal{P}(\mathbf{B})$ as the same set.

Let $\mathcal{B}_{\mathbf{c}}(r)$ denote the ball centered at $\mathbf{c}$ with radius r, for $\mathbf{c} = \mathbf{0}$ we write $\mathcal{B}(r)$. For any n-dimensional lattice L and $i = 1, \ldots, n$, let the *ith successive minimum* $\lambda_i(L)$ be the smallest radius r such that $\mathcal{B}(r)$ contains i linearly independent lattice vectors.

The *statistical distance* between two discrete distributions D_1 and D_2 over a set S is $\Delta(D_1, D_2) = \frac{1}{2} \sum_{s \in S} |\mathbf{Pr}[D_1 = s] - \mathbf{Pr}[D_2 = s]|$.

2.2 Discrete Gaussian Distributions over Lattices

Definition 1 (Discrete Gaussian Distribution)**.** Let $L \subseteq \mathbb{R}^n$ be a lattice, $s \in \mathbb{R}^+$, $\mathbf{c} \in \mathbb{R}^n$. For all $\mathbf{x} \in L$ let $\rho_{s,\mathbf{c}}(\mathbf{x}) = \exp(-\pi \left\| \mathbf{x} - \mathbf{c} \right\|^2 / s^2)$. For a set S let $\rho_{s,\mathbf{c}}(S) = \sum_{\mathbf{x} \in S} \exp(-\pi \left\| \mathbf{x} - \mathbf{c} \right\|^2 / s^2)$. Define the *discrete Gaussian distribution* over L centered at $\mathbf{c}$ with standard deviation s to be the probability distribution

$$D_{L,s,\mathbf{c}}(\mathbf{x}) = \frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{\rho_{s,\mathbf{c}}(L)},$$

for all $\mathbf{x} \in L$.

If the standard deviation of a discrete Gaussian distribution is larger than the smoothing parameter, defined below, then there are known, useful, results of discrete Gaussian distributions that we will use the in this paper.

Definition 2 (Smoothing parameter)**.** For any lattice L and real value $\epsilon > 0$, let the *smoothing parameter* $\eta_\epsilon(L)$ denote the smallest s such that $\rho_{1/s}(L^* \setminus \{\mathbf{0}\}) \le \epsilon$. We say that "$s$ exceeds the smoothing parameter" if $s \ge \eta_\epsilon(L)$ for negligible ϵ.

Below we show that the discrete Gaussian distribution is spherical if its standard deviation is larger than the smoothing parameter.

Lemma 1 (Micciancio and Regev [38])**.** *Let L be any full-rank n-dimensional lattice. For any $\mathbf{c} \in \mathbb{R}^n$, real $\epsilon \in (0,1)$, and $s \ge \eta_\epsilon(L)$ we have*

$$\mathbf{Pr}[\|\mathbf{x} - \mathbf{c}\| > s \cdot \sqrt{n} \mid \mathbf{x} \leftarrow D_{L,s,\mathbf{c}}] \le \frac{1+\epsilon}{1-\epsilon} \cdot 2^{-n}$$

Lemma 2 (Micciancio and Regev [38])**.** *Let L be any full-rank n-dimensional lattice. For any $s \ge \eta_\epsilon(L)$, $\epsilon \in (0,1)$, and any $\mathbf{c} \in \mathbb{R}^n$, we have $\rho_{s,\mathbf{c}}(L) \in \left[\frac{1-\epsilon}{1+\epsilon}, 1\right] \cdot \rho_{s,\mathbf{0}}(L)$.*

For a discrete Gaussian distribution over L centered at $\mathbf{0}$, with standard deviation s, $D_{L,s,\mathbf{0}}$ we let the translated discrete Gaussian distribution over L centered at any $\mathbf{c}$, with standard deviation s, be $D_{L,s,\mathbf{c}}$. Using Lemma 1 and Lemma 2, we show that the statistical distance between the original discrete Gaussian distribution and its translated discrete Gaussian distribution is negligible when $\|\mathbf{c}\|$ is small.

Corollary 1. *Let $\epsilon > 0$ negligible, $s \ge \eta_\epsilon(L)$. If $\|\mathbf{c}\| \le \frac{\epsilon}{6\pi\sqrt{n}} \cdot s$ then the statistical distance between $D_{L,s,\mathbf{0}}$ and $D_{L,s,\mathbf{c}}$ is at most 3ϵ.*

Proof. Suppose $\epsilon = 2^{-(n-1)}$. As in Lemma 1 we have $\mathbf{Pr}[\mathbf{x} \notin \mathcal{B}(s\sqrt{n}) \mid \mathbf{x} \leftarrow D_{L,s,\mathbf{0}}] \le \epsilon$ and $\mathbf{Pr}[\mathbf{x} \notin \mathcal{B}_\mathbf{c}(s\sqrt{n}) \mid \mathbf{x} \leftarrow D_{L,s,\mathbf{c}}] \le \epsilon$. To show that the statistical distance between the two distributions is small we partition the lattice into two sets. First, we look at all $\mathbf{x} \in L$ which is *not* in the union $\mathcal{B}_\mathbf{c}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n})$.

$$\begin{aligned}\sum_{\mathbf{x} \in L \setminus (\mathcal{B}_\mathbf{c}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n}))} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| &\le \sum_{\mathbf{x} \in L \setminus \mathcal{B}_\mathbf{c}(s\sqrt{n})} D_{L,s,\mathbf{c}}(\mathbf{x}) + \sum_{\mathbf{x} \in L \setminus \mathcal{B}(s\sqrt{n})} D_{L,s,\mathbf{0}}(\mathbf{x}) \\ &\le 2\epsilon,\end{aligned}$$

which follows from Lemma 1. Second, we look at all $\mathbf{x}$ in the union $\mathcal{B}_\mathbf{c}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n})$.

$$\begin{aligned}\sum_{\mathbf{x} \in L \cap (\mathcal{B}_\mathbf{c}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n}))} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| &\le \sum_{\mathbf{x} \in L \cap \mathcal{B}(2s\sqrt{n})} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| \\ &\le 4\epsilon.\end{aligned}$$

The first inequality is straight forward. For the last inequality, we claim that for all $\mathbf{x} \in L \cap \mathcal{B}(2s\sqrt{n})$ we have $|D_{s,\mathbf{c}}(\mathbf{x}) - D_{s,\mathbf{0}}(\mathbf{x})| \leq 4\epsilon D_{s,\mathbf{0}}(\mathbf{x})$. Note that

$$\begin{aligned}
|D_{s,\mathbf{c}}(\mathbf{x}) - D_{s,\mathbf{0}}(\mathbf{x})| &\overset{\text{Lemma 2}}{\leq} \frac{1}{\rho_{s,\mathbf{0}}(L)} \left(\frac{1+\epsilon}{1-\epsilon} |\rho_{s,\mathbf{c}}(\mathbf{x}) - \rho_{s,\mathbf{0}}(\mathbf{x})| + \frac{2\epsilon}{1-\epsilon} \rho_{s,\mathbf{0}}(\mathbf{x}) \right) \\
&= \frac{1}{\rho_{s,\mathbf{0}}(L)} \left(\frac{1+\epsilon}{1-\epsilon} \left| e^{-\frac{\pi}{s^2}(\|\mathbf{x}-\mathbf{c}\|^2 - \|\mathbf{x}\|^2)} - 1 \right| \cdot \rho_{s,\mathbf{0}}(\mathbf{x}) + \frac{2\epsilon}{1-\epsilon} \rho_{s,\mathbf{0}}(\mathbf{x}) \right) \\
&\leq D_{s,\mathbf{0}}(\mathbf{x}) \left(\frac{1+\epsilon}{1-\epsilon} \cdot \frac{\pi}{s^2} \cdot \left| \|\mathbf{x}-\mathbf{c}\|^2 - \|\mathbf{x}\|^2 \right| + \frac{2\epsilon}{1-\epsilon} \right) \\
&\leq D_{s,\mathbf{0}}(\mathbf{x}) \left(\frac{1+\epsilon}{1-\epsilon} \cdot \frac{\pi}{s^2} \cdot \|\mathbf{c}\| \left(\|\mathbf{x}\| + \|\mathbf{x}-\mathbf{c}\| \right) + \frac{2\epsilon}{1-\epsilon} \right) \\
&\leq D_{s,\mathbf{0}}(\mathbf{x}) \left(\frac{1+\epsilon}{1-\epsilon} \cdot \frac{\pi}{s^2} \cdot \|\mathbf{c}\| \left(2s\sqrt{n} + 4s\sqrt{n} \right) + \frac{2\epsilon}{1-\epsilon} \right) \\
&\leq 4\epsilon D_{s,\mathbf{0}}(\mathbf{x}),
\end{aligned}$$

and we get

$$\sum_{\mathbf{x} \in L \cap \mathcal{B}(2s\sqrt{n})} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| \leq \sum_{\mathbf{x} \in L \cap \mathcal{B}(2s\sqrt{n})} 4\epsilon D_{s,\mathbf{0}}(\mathbf{x}) \leq 4\epsilon \sum_{\mathbf{x} \in L} D_{s,\mathbf{0}}(\mathbf{x}) = 4\epsilon.$$

Combining the above results we have

$$\begin{aligned}
\Delta(D_{L,s,\mathbf{c}}(\mathbf{x}), D_{L,s,\mathbf{0}}(\mathbf{x})) =& \frac{1}{2} \sum_{\mathbf{x} \in L} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| \\
=& \frac{1}{2} \sum_{\mathbf{x} \in L \setminus (\mathcal{B}_{\mathbf{c}}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n}))} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| \quad + \\
& \frac{1}{2} \sum_{\mathbf{x} \in L \cap (\mathcal{B}_{\mathbf{c}}(s\sqrt{n}) \cup \mathcal{B}(s\sqrt{n}))} |D_{L,s,\mathbf{c}}(\mathbf{x}) - D_{L,s,\mathbf{0}}(\mathbf{x})| \\
\leq& 3\epsilon.
\end{aligned}$$

□

2.3 Gentry's homomorphic encryption scheme

Let $\mathsf{GHE} = (\mathsf{KG}_{\mathsf{GHE}}, \mathsf{Enc}_{\mathsf{GHE}}, \mathsf{Dec}_{\mathsf{GHE}}, \mathsf{Add}_{\mathsf{GHE}})$ be an (additively) Homomorphic encryption scheme derived from ideal lattices, with algrotihms as defined in Figure 2. The scheme is similar to Gentry's somewhat-homomorphic scheme [23]. The parameters of the GHE scheme are chosen as follows.

- Choose a polynomial ring $R = \mathbb{Z}[x]/(f(x))$ according to a security parameter λ;
- Choose a basis $\mathbf{B}_I$ of the ideal $I \subseteq R$;
- IdealGen is an algorithm which takes $(R, \mathbf{B}_I)$ as input and outputs public and secret bases $\mathbf{B}_J^{pk}$ and $\mathbf{B}_J^{sk}$ of some ideal J, where I and J are relatively prime;
- Samp is an algorithm which takes $(\mathbf{B}_I, \mathbf{x} \in R, s)$ as input and outputs a sample from the coset $\mathbf{x} + I$ according to a discrete Gaussian distribution with standard deviation s. In our construction we use the following two distributions.
  - $\mathsf{Samp}_1(\mathbf{B}_I, \mathbf{x}, s) = \mathbf{x} + D_{I,s,-\mathbf{x}}$;
  - $\mathsf{Samp}_2(\mathbf{B}_I, \mathbf{x}, s) = \mathbf{x} + D_{I,s,\mathbf{0}}$;
- The plaintext space $\mathcal{P} = R \mod \mathbf{B}_I$ is the set of distinguished representatives of cosets of I with respect to the basis $\mathbf{B}_I$.

```
$\mathsf{KG}_{\mathsf{GHE}}(R, \mathbf{B}_I)$ :
    $(\mathbf{B}_J^{pk}, \mathbf{B}_J^{sk}) \xleftarrow{\$} \mathsf{IdealGen}(R, \mathbf{B}_I)$
    $\mathsf{pk} = (R, \mathbf{B}_I, \mathbf{B}_J^{pk}, \mathsf{Samp}), \mathsf{sk} = \mathbf{B}_J^{sk}$
    return pk, sk

$\mathsf{Dec}_{\mathsf{GHE}}(\mathsf{sk}, \psi)$ :
    $\pi \leftarrow (\psi \mod \mathbf{B}_J^{sk}) \mod \mathbf{B}_I$
    return $\pi$

$\mathsf{Enc}_{\mathsf{GHE}}(\mathsf{pk}, s, \pi \in \mathcal{P})$ :
    $\psi' \leftarrow \mathsf{Samp}(\mathbf{B}_I, \pi, s)$
    $\psi \leftarrow \psi' \mod \mathbf{B}_J^{pk}$
    return $\psi$

$\mathsf{Add}_{\mathsf{GHE}}(\mathsf{pk}, \psi_1, \psi_2)$ :
    $\psi \leftarrow \psi_1 + \psi_2 \mod \mathbf{B}_J^{pk}$
    return $\psi$
```

Figure 2: The algorithms of the GHE homomorphic encryption scheme, which is similar to Gentry's somewhat homomorphic encryption scheme [23].

Correctness. Let X_{Enc} denote the image of Samp and X_{Dec} denote $R \mod \mathbf{B}_J^{sk} = \mathcal{P}(\mathbf{B}_J^{sk})$. Notice that all ciphertexts are in $X_{\mathsf{Enc}} + J$, since X_{Dec} is the set of distinguished representatives with respect to $\mathbf{B}_J^{sk}$. The correctness requirement of this encryption scheme is $X_{\mathsf{Enc}} \subseteq X_{\mathsf{Dec}}$. Furthermore, for the addition algorithm $\mathsf{Add}_{\mathsf{GHE}}$ to output valid ciphertexts we require that $X_{\mathsf{Enc}} + X_{\mathsf{Enc}} \subseteq X_{\mathsf{Dec}}$.

Let r_{Enc} be the smallest value such that $X_{\mathsf{Enc}} \subseteq \mathcal{B}(r_{\mathsf{Enc}})$ and let r_{Dec} be the largest value such that $X_{\mathsf{Dec}} \supseteq \mathcal{B}(r_{\mathsf{Dec}})$. By the spherical property of discrete Gaussian distribution (Lemma. 1) we know that, for Samp_1 as above, X_{Enc} is located inside the ball $\mathcal{B}(s\sqrt{n})$ with high probability and $r_{\mathsf{Enc}} = s\sqrt{n}$. For a general Samp algorithm, which is located in $\mathcal{B}(l_{\mathsf{Samp}})$, we have that $r_{\mathsf{Enc}} \leq (n + \sqrt{n} l_{\mathsf{Samp}}) \|\mathbf{B}_I\|$ [23]. For r_{Dec} we know that $r_{\mathsf{Dec}} = 1/(2 \cdot \|((\mathbf{B}_J^{sk})^{-1})^T\|)$ [23].

Obviously, if $r_{\mathsf{Enc}} \leq r_{\mathsf{Dec}}$ then the encryption scheme is correct. For GHE, if $r_{\mathsf{Enc}} \leq r_{\mathsf{Dec}}$, the probability of decryption error is less than $\frac{1+\epsilon}{1-\epsilon} \cdot 2^{-n}$, which is negligible.

2.4 The revised NTRU encryption scheme

The NTRU encryption scheme variant by Stehlé and Steinfeld [40], which relies on the LWE problem, has the similar structure as Gentry's homomorphic encryption scheme. We modify the NTRU scheme to use a discrete Gaussian distribution as the noise distribution instead of an elliptic Gaussian. Choose the parameters of the scheme as follows.

- $R = \mathbb{Z}[x]/(x^n + 1)$, where $n \geq 8$ is a power of 2;
- q is a prime, $5 \leq q \leq Poly(n)$, $R_q = R/q$;
- $p \in R_q^{\times}, I = (p)$;
- the plaintext space $\mathcal{P} = R/p$;
- set the noise distribution to be $D_{\mathbb{Z}^n,s,\mathbf{0}}$.

The algorithms of the scheme are given in Figure 3.

```
$\mathsf{KG}_{\mathsf{NTRU}}(n, q \in \mathbb{Z}, p \in R_q^{\times}, s > 0)$ :
    while $(f \mod q) \notin R_q^{\times}$ do
        $f' \leftarrow D_{\mathbb{Z}^n,s,\mathbf{0}}$
        $f = p \cdot f' + 1$
    while $(g \mod q) \notin R_q^{\times}$ do
        $g \leftarrow D_{\mathbb{Z}^n,s,\mathbf{0}}$
    $h = pg/f \in R_q$
    $(\mathsf{pk}, \mathsf{sk}) \leftarrow (h, f)$
    return $(\mathsf{pk}, \mathsf{sk})$

$\mathsf{Enc}_{\mathsf{NTRU}}(\mathsf{pk} = h, s, \pi \in \mathcal{P})$ :
    $e_1, e_2 \leftarrow D_{\mathbb{Z}^n,s,\mathbf{0}}$
    $\psi \leftarrow \pi + pe_1 + he_2 \in R_q$
    return $\psi$

$\mathsf{Dec}_{\mathsf{NTRU}}(\mathsf{sk} = f, \psi)$ :
    $\psi' = f \cdot \psi \in R_q$
    $\pi \leftarrow \psi' \mod p$
    return $\pi$

$\mathsf{Add}_{\mathsf{NTRU}}(\psi_1, \psi_2)$ :
    $\psi \leftarrow \psi_1 + \psi_2 \in R_q$
    return $\psi$
```

Figure 3: The algorithms of the revised NTRU encryption scheme [40].

Correctness Let $\psi' = f\pi + p(fe_1 + ge_2) \in R_q$ and $\psi'' = f\pi + p(fe_1 + ge_2) \in R$ (not modulo q), if $\|\psi''\|_\infty \leq q/2$ then the decryption algorithm will output π (see Lemma 12 of [40]). We will perform a single homomorphic addition and want to find a bound on the sum of two ciphertexts. Discrete Gaussian samples are bounded by $s\sqrt{n}$ with high probability (Lemma 1) and the message space parameter p is a polynomial with small coefficients, where we let p_i denote the largest coefficient of p. We have

$$\begin{aligned}\|f(\psi_1 + \psi_2)\|_\infty &= \left\|f(\pi_1 + \pi_2) + p_i(f(e_1 + e_1') + g(e_2 + e_2'))\right\|_\infty \\ &\leq 2(p_i^2(s\sqrt{n})^2 + p_i^2 s\sqrt{n} + p_i s\sqrt{n} + p_i + (s\sqrt{n})^2) \\ &\leq 8p_i^2 s^2 n.\end{aligned}$$

The standard deviation s is greater or equal to $\eta_\epsilon(\mathbb{Z}^n)$ and has to satisfy $\eta_\epsilon(\mathbb{Z}^n) \leq s$ and $8p_i^2 s^2 n < q/2$ for the decryption to be correct, with high probability.

2.5 Hard lattice problems

The following lattice problems, assumed to be hard, are used in the paper.

Definition 3 (Shortest Vector Problem (SVP)). Given a basis $\mathbf{B}$ for a n-dimensional lattice L, output a nonzero vector $\mathbf{v} \in L$ of length at most $\lambda_1(L)$.

Definition 4 (Ideal Shortest Independent Vector Problem (SIVP)). Fix the following parameters; a polynomial ring R, and a positive real $\gamma \geq 1$. Let $\mathbf{B}_I$ be a basis for an ideal lattice I of R. Given $\mathbf{B}_I$, and the parameters, output a basis $\mathbf{B}_I'$ of I with $\|\mathbf{B}_I'\| \leq \gamma \cdot \lambda_n(I)$.

Reduce Hard problems to the semantic security of Gentry's encryption scheme The following two theorems describe Gentry's reduction from worst-case SIVP to the semantic security of the encryption scheme GHE, via the ideal independent vector improvement problem (IVIP).

Theorem 1 (Gentry [23] (Corollary 14.7.1), reduce IVIP to semantic security). *Suppose that* $s_{\mathsf{IVIP}} < (\sqrt{2}s\epsilon - 4n^2(\max\{\|\mathbf{B}_I\|\})^2)/(n^4\gamma_\times(R)\,\|f\|\max\{\|\mathbf{B}_I\|\})$*, where s is the Gaussian deviation parameter in the encryption scheme* GHE*. Also suppose that* $s/2$ *exceeds the smoothing parameter of* I*, that* IdealGen *always outputs an ideal* J *with* $s \cdot \sqrt{n} < \lambda_1(J)$*, and that* $[R : I]$ *is prime. Finally, suppose that there is an algorithm* $\mathcal{A}$ *that breaks the semantic security of* GHE *with advantage* ϵ*. Then there is a quantum algorithm that solves* s_{IVIP}-IVIP *for an* $\epsilon/4$ *(up to negligible factors) weight fraction of bases output by* IdealGen.

Theorem 2 (Gentry [23] (Theorem 19.2.3 and Corollary 19.2.5), reduce SIVP to IVIP). *Suppose* $d_{\mathsf{SIVP}} = (3 \cdot e)^{1/n} \cdot d_{\mathsf{IVIP}}$*, where e is Euler's constant. Suppose that there is an algorithm* $\mathcal{A}$ *that solves* s_{IVIP}-IVIP *for parameter* $s_{\mathsf{IVIP}} > 16 \cdot \gamma_\times(R)^2 \cdot n^5 \cdot \|f\| \cdot g(n)$ *for some* $g(n)$ *that is* $\omega(\sqrt{\log n})$*, whenever the given ideal has* $\det(J) \in [a, b]$*, where* $[a, b] = [d_{\mathsf{IVIP}}^n, 2 \cdot d_{\mathsf{IVIP}}^n]$*. Assume that invertible prime ideals with norms in* $[a, b]$ *are not negligibly sparse. Then, there is an algorithm* $\mathcal{B}$ *that solves worst-case* d_{SIVP}-SIVP.

In summary we have the following informal result, which we will use to prove that our GHE-BKEM (see Section 5.4) is post quantum secure.

Theorem 3 (Gentry [23]). *If there exists an algorithm that breaks the semantic security of* GHE *with parameters chosen as in Theorem 1 and Theorem 2, then there exists a quantum algorithm that solves worst-case* SIVP.

Reduce Hard problems to the semantic security of the revised NTRU encryption scheme We define the ring learning with error problem as follows. For $s \in R_q$, an error distribution D over R_q, define $A_{s,D}$ to be a distribution that outputs tuples of the form $(a, as+e)$, where a is sampled uniformly at random from R_q and e is sampled from D. The problem is to distinguish between tuples sampled from $A_{s,D}$ and uniformly random ones.

Definition 5 (Ring-LWE)**.** Let $\mathcal{D}$ be a distribution over a family of distributions, each over R_q. The *Ring Learning With Errors Problem* with parameters q, and $\mathcal{D}$ ($\mathsf{R\text{-}LWE}_{q,\mathcal{D}}$) is as follows. Let D be sampled from $\mathcal{D}$ and s be sampled uniformly at random from R_q. Given access to an oracle $\mathcal{O}$ that produces samples in R_q^2, distinguish whether $\mathcal{O}$ outputs samples from the distribution $A_{s,D}$ or $U(R_q^2)$. The distinguishing advantage should be non-negligible.

Lyubashevsky et al. [35] proposed a reduction from SIVP or SVP (both are thought to be hard problems) to R-LWE.

Theorem 4 (Lyubashevsky et al. [35])**.** *Let $\alpha < \sqrt{logn/n}$ and $q = 1 \mod 2n$ be a poly(n)-bounded prime such that $\alpha q \geq \omega(\sqrt{\log n})$. Then there is a polynomial-time quantum reduction from $O(\sqrt{n}/\alpha)$-approximate* SIVP *(or* SVP*) on ideal lattices to* $\mathsf{R\text{-}LWE}_{q,D_s}$ *given only $l(\geq 1)$ samples, where $s = \alpha \cdot (nl/\log(nl))^{1/4}$.*

We will consider a different variant of the R-LWE problem, namely $\mathsf{R\text{-}LWE}^{\times}_{\mathsf{HNF}}$, which is the same as $\mathsf{R\text{-}LWE}_{q,\mathcal{D}}$ except for the oracle $\mathcal{O}$ that outputs samples from the distribution $A^{\times}_{s,D}$ or $U(R_q^2)$, where $A^{\times}_{s,D}$ outputs $(a, as + e)$ with $a \in R_q^{\times}, s \in D$. The analysis in the end of Section 2 of Stehlé and Steinfeld [40] shows that when $q = \Omega(n)$, $\mathsf{R\text{-}LWE}^{\times}_{\mathsf{HNF}}$ remains hard.

The security proof of NTRU encryption scheme is similar to the security proof of Lemma 3.8 provided by Stehlé and Steinfeld [40]. The proof technique relies on the uniformity of public key and $p \in R_q^{\times}$. However, we chose a slightly different error distribution for our construction in Section 5.5, but the adaption to our setting is straightforward.

Lemma 3. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into n irreducible factors modulo prime $q \geq 5$. Let $0 < \epsilon < 1/3$, $p \in R_q^{\times}$ and $s \geq 2n\sqrt{\ln(8nq)} \cdot q^{1/2+\epsilon}$. For any* IND-CPA *adversary $\mathcal{A}$ against* NTRU *encryption scheme, there exists an adversary $\mathcal{B}$ solving* $\mathsf{R\text{-}LWE}^{\times}_{\mathsf{HNF}}$ *such that*

$$\mathbf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{NTRU}}(\mathcal{A}) \leq \mathbf{Adv}_{\mathsf{R\text{-}LWE}^{\times}_{\mathsf{HNF}}}(\mathcal{B}) + q^{-\Omega(n)}.$$

3 Blinded KEM

The blinded KEM primitive is the most important building block that BDGJ used to construct their key transport protocol [10] – also required are a signature scheme, a public-key encryption scheme, a hash function and a key derivation function. In this paper we only focus on blinded KEMs.

A *blinded KEM* scheme BKEM is parameterized by a key encapsulation mechanism KEM = (KG, Encap, Decap), a blinding algorithm Blind and an unblinding algorithm Unblind; put together we have that BKEM = (KG, Encap, Blind, Decap, Unblind).

The *key generation* algorithm KG outputs an encapsulation key $ek \in \mathcal{K}_E$ and a decapsulation key $dk \in \mathcal{K}_D$. The *encapsulation* algorithm Encap takes as input an encapsulation key and outputs a (file encryption) key $k \in \mathcal{K}_F$ together with an encapsulation $C \in \mathcal{C}$ of that key. The *blinding* algorithm takes as input an encapsulation key and an encapsulation and outputs a blinded encapsulation $\tilde{C} \in \mathcal{C}$ and an unblinding key $uk \in \mathcal{K}_U$. The *decapsulation* algorithm Decap takes a decapsulation key and a (blinded) encapsulation as input and outputs a (blinded) key $\tilde{k} \in \mathcal{K}_B$. The *unblinding* algorithm takes as input an unblinding key and a blinded key and outputs a key.

Definition 6 (Correctness of a BKEM)**.** Scheme BKEM has *correctness* if:

$\mathsf{Unblind}_{uk}(\tilde{k}) = k$, when $(ek, dk) \leftarrow \mathsf{KG}$, $(C, k) \leftarrow \mathsf{Encap}_{ek}$, $(\tilde{C}, uk) \leftarrow \mathsf{Blind}_{ek}(C)$ and $\tilde{k} \leftarrow \mathsf{Decap}_{dk}(\tilde{C})$.

(The KEM scheme has *correctness* if $\mathsf{Decap}_{dk}(C) = k$, when $(ek, dk) \leftarrow \mathsf{KG}$ and $(C, k) \leftarrow \mathsf{Encap}_{ek}$.)

We parameterize all BKEM schemes by a public key encryption scheme (PKE), since any PKE scheme can trivially be turned into a KEM. We modify the above definition to be a PKE-based BKEM, where the KEM algorithms are described in Figure 4.

Definition 7 (PKE-based BKEM)**.** Let BKEM be a blinded KEM, where the underlying scheme $\mathsf{KEM} = (\mathsf{KG}, \mathsf{Encap}, \mathsf{Decap})$ is parameterized by a PKE scheme $\mathsf{PKE} = (\mathsf{KG}_{\mathsf{PKE}}, \mathsf{Enc}, \mathsf{Dec})$ as in Figure 4. We call such a BKEM a *PKE-based BKEM*.

```
$\mathsf{KG}(\lambda)$ :
    $\mathsf{pk}, \mathsf{sk} \leftarrow \mathsf{KG}_{\mathsf{PKE}}(\lambda)$
    $(ek, dk) \leftarrow (\mathsf{pk}, \mathsf{sk})$
    return $ek, dk$
```

```
$\mathsf{Encap}_{ek}$ :
    $k \xleftarrow{\$} \mathcal{M}$
    $C \leftarrow \mathsf{Enc}_{ek}(k)$
    return $C, k$
```

```
$\mathsf{Decap}_{dk}(\tilde{C})$ :
    $\tilde{k} \leftarrow \mathsf{Dec}_{dk}(\tilde{C})$
    return $\tilde{k}$
```

Figure 4: KEM algorithms parameterized by a PKE scheme $\mathsf{PKE} = (\mathsf{KG}_{\mathsf{PKE}}, \mathsf{Enc}, \mathsf{Dec})$.

3.1 Security

We define indistinguishability under chosen-plaintext attack (IND-CPA) for public-key encryption and indistinguishability (IND) for blinded KEMs, respectively.

Definition 8. Let $\mathsf{PKE} = (\mathsf{KG}_{\mathsf{PKE}}, \mathsf{Enc}, \mathsf{Dec})$ be a public key encryption scheme. The IND-CPA advantage of any adversary $\mathcal{A}$ against PKE is

$$\mathbf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A}) = 2\left|\Pr[\mathbf{Exp}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A}) = 1] - 1/2\right|,$$

where the experiment $\mathbf{Exp}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A})$ is given in Figure 5 (left). We say that PKE is IND-CPA-secure if $\mathbf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A})$ is negligible for any adversary $\mathcal{A}$.

```
$\mathbf{Exp}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A})$ :
    $b \xleftarrow{\$} \{0,1\}$
    $(\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{KG}_{\mathsf{PKE}}$
    $(\mathsf{m}_0, \mathsf{m}_1) \xleftarrow{\$} \mathcal{A}$
    $C_b \leftarrow \mathsf{Enc}_{\mathsf{pk}}(\mathsf{m}_b)$
    $b' \leftarrow \mathcal{A}(\mathsf{pk}, C_b)$
    return $b' \overset{?}{=} b$
```

```
$\mathbf{Exp}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r)$ :
    $b \xleftarrow{\$} \{0,1\}$
    $(ek, dk) \leftarrow \mathsf{KG}$
    $(C, k_1) \leftarrow \mathsf{Encap}_{ek}$
    $k_0 \xleftarrow{\$} \mathcal{K}_F$
    for $j \in \{1, \ldots, r\}$ do
        $(\tilde{C}_j, uk_j) \leftarrow \mathsf{Blind}_{ek}(C)$
        $\tilde{k}_j \leftarrow \mathsf{Decap}_{dk}(\tilde{C}_j)$
    $b' \leftarrow \mathcal{A}(ek, C, k_b, \{(\tilde{C}_j, \tilde{k}_j)\}_{1 \le j \le r})$
    return $b' \overset{?}{=} b$
```

Figure 5: IND-CPA experiment $\mathbf{Exp}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKE}}(\mathcal{A})$ for PKE scheme PKE (left). Indistinguishability experiment $\mathbf{Exp}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r)$ for a BKEM scheme BKEM (right).

Definition 9. Let $\mathsf{BKEM} = (\mathsf{KG}, \mathsf{Encap}, \mathsf{Blind}, \mathsf{Decap}, \mathsf{Unblind})$ be a blinded KEM. The *distinguishing advantage* of any adversary $\mathcal{A}$ against BKEM getting r blinded encapsulations and their blinded decapsulation tuples is

$$\mathbf{Adv}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r) = 2\left|\Pr[\mathbf{Exp}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r) = 1] - 1/2\right|,$$

where the experiment $\mathbf{Exp}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r)$ is given in Figure 5 (right). We say that BKEM is IND-secure if $\mathbf{Adv}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r)$ is negligible for any adversary $\mathcal{A}$.

4 Homomorphic-based BKEM

We now show how to turn a homomorphic encryption scheme with certain properties into a BKEM, and analyze the security requirements of such a BKEM.

4.1 Generic homomorphic-based BKEM

We look for PKE schemes with the following homomorphic property: suppose C and C' are ciphertexts of k and k', resp., then $\mathsf{Dec}_{\mathsf{sk}}(C \oplus_1 C') = k \oplus_2 k'$, where $\oplus_1$ and $\oplus_2$ denote two group operations. We see two reasons to look at such PKE schemes.

The first reason is that in a BKEM scheme we want the blinding algorithm to alter the file encryption key k. Having a homomorphic encryption (HE) scheme makes this possible and we can construct a blinding algorithm. The second reason is that we want k' to hide k such that the adversary is unable to gain information about k even with knowledge of $\tilde{k} = k \oplus_2 k'$. With a homomorphic encryption scheme we can combine two independently random ciphertexts and make a third one.

We can construct blinding and unblinding algorithms, using this homomorphic property, to create a BKEM with correctness. To blind an encapsulation C (with corresponding file encryption key k) the Blind algorithm creates a fresh encapsulation C' (with corresponding blinding value k') using the Encap_{ek} algorithm, the blinded encapsulation $\tilde{C}$ is computed as $\tilde{C} \leftarrow C \oplus_1 C'$. The unblinding key uk is the inverse element of k' with respect to $\oplus_2$, that is, $uk \leftarrow k'^{-1}$. The blinding algorithms outputs $\tilde{C}$ and uk. The decapsulation algorithm can evaluate the blinded encapsulation because of the homomorphic property. The blinded key $\tilde{k}$ is the output of the decapsulation algorithm, that is, $\tilde{k} \leftarrow \mathsf{Decap}_{dk}(\tilde{C})$. To unblind $\tilde{k}$ the unblinding algorithm outputs $\tilde{k} \oplus_2 uk$, which is $(k \oplus_2 k') \oplus_2 (k'^{-1}) = k$, and so the BKEM scheme has correctness. Formally, we define the BKEM scheme constructed above as follows.

Definition 10 (Homomorphic-based BKEM)**.** Let BKEM be a PKE-based BKEM, as in Definition 7. Suppose the underlying public key encryption scheme is a homomorphic encryption scheme $\mathsf{HE} = (\mathsf{KG}_{\mathsf{HE}}, \mathsf{Enc}, \mathsf{Dec})$ such that for any $k, k' \in \mathcal{M}$ and any key pair $(\mathsf{sk}, \mathsf{pk}) \xleftarrow{\$} \mathsf{KG}_{\mathsf{HE}}$ it holds that

$$\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{pk}}(k) \oplus_1 \mathsf{Enc}_{\mathsf{pk}}(k')) = k \oplus_2 k'$$

where $(\mathcal{M}, \oplus_2)$ is the plaintext group and $(\mathcal{C}, \oplus_1)$ is the ciphertext group. Furthermore, let the blinding and unblinding algorithms operate according to Figure 6. We call such a scheme BKEM a *homomorphic-based BKEM*.

```
$\mathsf{Blind}_{ek}(C)$ :
    $(C', k') \leftarrow \mathsf{Encap}_{ek}$
    $\tilde{C} \leftarrow C \oplus_1 C'$
    $uk \leftarrow k'^{-1}$
    return $\tilde{C}, uk$
```

```
$\mathsf{Unblind}_{uk}(\tilde{k})$ :
    $k \leftarrow \tilde{k} \oplus_2 uk$
    return $k$
```

Figure 6: Blinding and unblinding algorithms of the homomorphic based BKEM.

All BKEM schemes considered in the rest of this paper are homomorphic-based BKEMs.

The homomorphic encryption scheme HE does not need to be fully homomorphic, since we only need one operation in the blinding algorithm: a partially homomorphic encryption scheme is sufficient.

4.2 Security requirements

In the indistinguishability game IND for BKEMs the adversary $\mathcal{A}$ has r blinded samples, which are the following two sets: $\{\tilde{k}_i = k \oplus_2 k'_i\}_{i=1\ldots r}$ and $\{\tilde{C}_i = C \oplus_1 C'_i\}_{1,\ldots,r}$, in addition to an encapsulation C of the real file encryption key. We want the blinded samples and the encapsulation to be random looking such that the combination of all these values does not reveal any information about the underlying file encryption key k that is being transported.

First, we show how to choose the blinding values k'_i to make the blinded keys $\tilde{k}_i$ look random. Then, we show how to make the blinded encapsulations $\tilde{C}_i$ look like a fresh output of the encapsulation algorithm, similar to circuit privacy [23]. Finally, we show how an $\mathsf{IND\text{-}CPA}$-secure HE scheme ensures that the encapsulation does not reveal any information about the file encryption key.

Eventually, we provide the main theorem in this paper stating how to achieve an IND secure BKEM scheme. Particularly, if the underlying HE scheme is post-quantum $\mathsf{IND\text{-}CPA}$ secure then the corresponding homomorphic-based BKEM scheme is post-quantum IND secure.

4.2.1 Random-looking blinded keys

We want the blinded key to look like a random element of the space containing blinded keys. In the IND game the adversary will be given several blinded keys of the form $\tilde{k} = k \oplus_2 k'$, where k is the file encryption key and k' is a blinding value, and wishes to gain information about k.

Let k be sampled uniformly at random from the file encryption key set $\mathcal{K}_F$ and let k' be sampled uniformly at random from the blinding value set $\mathcal{K}_R$. We would like that the size of $\mathcal{K}_F$ is large enough to prevent a brute force attacker from guessing the key k, say $|\mathcal{K}_F| = 2^\lambda$ for some security parameter λ. If $\mathcal{K}_R$ is a small set then the value of any blinded key $\tilde{k} = k \oplus_2 k'$ will be located within a short distance around k, so the adversary can successfully guess k with high probability. We always assume that $\mathcal{K}_R$ is at least as large as $\mathcal{K}_F$.

If a given blinded key $\tilde{k}$ can be expressed as a result of any file encryption key k and a blinding value k', with respect to an operation, then our goal is to ensure that the adversary cannot get any information of the true file encryption key hidden in $\tilde{k}$, and ideally we wish it to be indistinguishable from a random element.

Definition 11 (ϵ-blinded blinded key)**.** Let BKEM be a blinded KEM with blinded key set $\mathcal{K}_B$. Let k be sampled uniformly random from the file encryption key set $\mathcal{K}_F$ and let k' be sampled uniformly random from the blinding value set $\mathcal{K}_R$. We define a *ϵ-blinded blinded key set* $\mathsf{S} := \{\tilde{k} \in \mathcal{K}_B \mid \forall k \in \mathcal{K}_F, \exists 1 k' \in \mathcal{K}_R \text{ such that } \tilde{k} = k \oplus_2 k'\}$: we say that BKEM has ϵ-blinded blinded keys if

$$\mathbf{Pr}\left[\tilde{k} = k \oplus_2 k' \in \mathsf{S} \mid k \xleftarrow{\$} \mathcal{K}_F, k' \xleftarrow{\$} \mathcal{K}_R\right] = 1 - \epsilon.$$

Suppose the adversary is given any number of ϵ-blinded blinded keys from S with the same underlying file encryption key k. By the definition of the ϵ-blinded blinded set the file encryption key k can be any value in $\mathcal{K}_F$ and all values are equally probable. In other words, guessing k, given ϵ-blinded blinded keys, is the same as guessing a random value from $\mathcal{K}_F$. To prevent giving the adversary a better chance at guessing the key k we wish the blinded keys to be located inside the ϵ-blinded blinded key set S with high probability, which means we want ϵ to be small.

4.2.2 Fresh-looking blinded encapsulations

Blinded encapsulations are constructed from two encapsulations, one containing the file encryption key and one containing a blinded value, where we want it to look like a fresh encapsulation, containing the result of the two values with respect to $\oplus_2$. In the IND game for BKEMs the adversary $\mathcal{A}$ gets r blinded samples and has knowledge of the set $\{\tilde{C}_i = C \oplus_1 C'_i\}_{1,\ldots,r}$, where C is an encapsulation of a file encryption key k and C'_i is an encapsulation of a blinding value. We want this set to be indistinguishable from the output set of the encapsulation algorithm.

Definition 12 (ϵ-blinded blinded encapsulation)**.** Let $\mathsf{HE\text{-}BKEM}$ be a homomorphic-based BKEM. Let ek be any encapsulation key and C_0 be an encapsulation with the underlying file encryption key k_0. We say that $\mathsf{HE\text{-}BKEM}$ has *ϵ-blinded blinded encapsulation* if the statistical distance between the following distributions is at most ϵ:

$$X = \{C_0 \oplus_1 C' \mid k' \xleftarrow{\$} \mathcal{K}_R, C' \leftarrow \mathsf{Enc}_{ek}(k')\},$$
$$Y = \{C \mid k' \xleftarrow{\$} \mathcal{K}_R, C \leftarrow \mathsf{Enc}_{ek}(k_0 \oplus_2 k')\}.$$

The above property ensures that the output of the blinding algorithm looks like a fresh encapsulation expect for probability ϵ. Note that the BKEM constructions of Boyd et al. [10], $\mathsf{DH\text{-}BKEM}$ and $\mathsf{RSA\text{-}BKEM}$, both have 0-blinded blinded encapsulation.

It is well known that in a fully homomorphic encryption scheme the product of two ciphertexts is much larger compared to the sum of two ciphertexts, hence, it is easier to achieve ϵ-blinded blinded encapsulation for one addition compared to one multiplication. In our constructions we will use addition.

4.2.3 Indistinguishability of BKEM

Furthermore, if we want to achieve indistinguishability of blinded KEM. We require the underlying homomorphic encryption scheme have some kind of semantic security to protect the message (the file encryption key) in the ciphertext (the encapsulation).

Theorem 5 (Main Theorem). *Let* BKEM *be a homomorphic based BKEM designed as in Definition 10 from a homomorphic encryption scheme* HE*. Let the file encryption key k and the blinding value k' be sampled uniformly random from the large sets $\mathcal{K}_F$ and $\mathcal{K}_R$, respectively. Suppose* BKEM *has ϵ_1-blinded blinded encapsulations and ϵ_2-blinded blinded keys. For any adversary $\mathcal{A}$ against* BKEM *getting r blinded encapsulations and their blinded decapsulation samples, there exists an* $\mathsf{IND\text{-}CPA}$ *adversary $\mathcal{B}$ against* HE *such that*

$$\mathbf{Adv}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r) \leq 2r(\epsilon_1 + \epsilon_2) + \mathbf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{HE}}(\mathcal{B})$$

Proof. The proof of the theorem consists of a sequence of games. □

Game 0

The first game is the experiment $\mathbf{Exp}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r)$, given in Figure 5 (right). From Definition 9 we have that

$$\mathbf{Adv}^{\mathsf{IND}}_{\mathsf{BKEM}}(\mathcal{A}, r) = 2\left|\Pr[E_0] - 1/2\right| .$$

Game 1

Modify the game such that the $\mathsf{Unblind}$ algorithm outputs a random ϵ-blinded blinded key (Definition 11), $\tilde{k} \xleftarrow{\$} \mathsf{S}$, and the Blind algorithm outputs an encapsulation of this random key, $\tilde{C} \leftarrow \mathsf{Enc}_{ek}(\tilde{k})$.

We first prove that a real pair of blinded key and blinded encapsulation output in Game 0 is $(\epsilon_1 + \epsilon_2)$ statically close to the modified values output in Game 1.

Suppose $k_0 \in \mathcal{K}_F$ is the file encryption key and $C_0 \leftarrow \mathsf{Enc}_{ek}(k_0)$ is the encapsulation with k_0, let $X = \{(k_0 \oplus_2 k', C_0 \oplus_1 C') \mid k' \xleftarrow{\$} \mathcal{K}_R, C' \leftarrow \mathsf{Enc}_{ek}(k')\}$ be the statistical distribution of the real pair of blinded key and blinded encapsulation output in Game 0, and $Y = \{(\tilde{k}, \tilde{C}) \mid \tilde{k} \xleftarrow{\$} \mathsf{S}, \tilde{C} \leftarrow \mathsf{Enc}_{ek}(\tilde{k})\}$ be the statistical distribution of the modified values output in Game 1. We define a middle distribution $Z = \{(k_0 \oplus_2 k', C) \mid k' \xleftarrow{\$} \mathcal{K}_R, C \leftarrow \mathsf{Enc}_{ek}(k_0 \oplus_2 k')\}$. We compute the statistical distance between X and Y as follows.

$$\begin{aligned}
\Delta(X,Y) &= \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathcal{K}_B\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[X=(\tilde{k},\tilde{C})] - \mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big|)\\
&= \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathcal{K}_B\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[X=(\tilde{k},\tilde{C})] - \mathbf{Pr}[Z=(\tilde{k},\tilde{C})] + \mathbf{Pr}[Z=(\tilde{k},\tilde{C})] - \mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big|)\\
&\leq \Delta(X,Z) + \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathcal{K}_B\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[Z=(\tilde{k},\tilde{C})] - \mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big|)\\
&\leq \epsilon_1 + \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathcal{K}_B\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[Z=(\tilde{k},\tilde{C}) \mid \tilde{k}\in\mathsf{S}]\cdot\mathbf{Pr}[\tilde{k}\in\mathsf{S}]\\
&\qquad + \mathbf{Pr}[Z=(\tilde{k},\tilde{C}) \mid \tilde{k}\notin\mathsf{S}]\cdot\mathbf{Pr}[\tilde{k}\notin\mathsf{S}] - \mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big|)\\
&\leq \epsilon_1 + \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathsf{S}\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[Z=(\tilde{k},\tilde{C})]\cdot(1-\epsilon_2) - \mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big| + \sum_{\substack{\tilde{k}\notin\mathsf{S}\\ \tilde{C}\in\mathcal{C}}} \Big|\mathbf{Pr}[Z=(\tilde{k},\tilde{C})]\Big|)\\
&\leq \epsilon_1 + \frac{1}{2}(\sum_{\substack{\tilde{k}\in\mathsf{S}\\ \tilde{C}\in\mathcal{C}}} \Big|\epsilon_2\cdot\mathbf{Pr}[Y=(\tilde{k},\tilde{C})]\Big| + \epsilon_2)\\
&\leq \epsilon_1 + \epsilon_2
\end{aligned}$$

For r samples we get

$$\Big|\Pr[E_1] - \Pr[E_0]\Big| \leq r(\epsilon_1+\epsilon_2).$$

Next, we claim that there exists an adversary $\mathcal{B}$ against IND-CPA security of HE such that

$$2\Big|\Pr[E_1] - \frac{1}{2}\Big| = \mathbf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{HE}}(\mathcal{B}).$$

We construct a reduction $\mathcal{B}$ that plays the IND-CPA game by running $\mathcal{A}$, it simulates the responses of Game 1 to $\mathcal{A}$ as follows.

1. $\mathcal{B}$ flips a coin $b \xleftarrow{\$} \{0,1\}$,
2. $\mathcal{B}$ simulates the key generation algorithm KG by running its own $\mathsf{KG}_{\mathsf{HE}}$,
3. $\mathcal{B}$ simulates the encapsulation by randomly choosing two group key k_0, k_1, sends challenge query with input (k_0, k_1) to its IND-CPA challenger, and forwards the response C to $\mathcal{A}$,
4. $\mathcal{B}$ simulates the outputs of the Blind and Unblind algorithm by running Encap algorithms $\tilde{k} \xleftarrow{\$} \mathsf{S}, \tilde{C} \leftarrow \mathsf{Enc}_{ek}(\tilde{k})$, and outputs $\tilde{C}$ as the blinded encapsulation and $\tilde{k}$ as the blinded key.
5. When $\mathcal{A}$ asks for a challenge, $\mathcal{B}$ sends k_b to $\mathcal{A}$.
6. After $\mathcal{A}$ sends back a guess b', $\mathcal{B}$ sends b to the challenger if $b' = 1$ and $1-b$ if $b' = 0$.

If $\mathcal{B}$ interacts with $\mathbf{Exp}^{\mathsf{ind\text{-}cpa\text{-}}b}_{\mathsf{HE}}(\mathcal{A})$ then $\mathcal{B}$ perfectly simulates the inputs of $\mathcal{A}$ in Game 1 when the output of the key is a real key. Otherwise ($\mathcal{B}$ interacts with $\mathbf{Exp}^{\mathsf{ind\text{-}cpa\text{-}}(1\text{-}b)}_{\mathsf{HE}}(\mathcal{A})$), k_b is a random key to $\mathcal{A}$ and $\mathcal{B}$ perfectly simulate the inputs of $\mathcal{A}$ in Game 1 when the output of the key is a random key.

Remark 1. *As a specific case of Theorem 5, the* DH-BKEM *construction of BDGJ has 0-blinded blinded encapsulations and 0-blinded blinded keys, and the indistinguishability of* DH-BKEM *is upper bounded by* DDH *advantage (defined in the real-or-random sense instead of left-or-right).*

$$\mathbf{Adv}^{\mathsf{IND}}_{\mathsf{DH\text{-}BKEM}}(\mathcal{A}, r) \leq \mathbf{Adv}^{\mathsf{DDH}}(\mathcal{B})$$

5 Instantiating Homomorphic-based BKEMs

We provide two specific homomorphic-based BKEM constructions, based on Gentry's homomorphic encryption scheme (see Section 2.3) and the NTRU variant by Stehlé and Steinfeld (see Section 2.4). We also prove that the BKEMs that result from these HE schemes are post-quantum secure by reducing to hard lattice problems.

5.1 Two Homomorphic-based BKEM

Let HE $= (\mathsf{KG}_{\mathsf{HE}}, \mathsf{Enc}_{\mathsf{HE}}, \mathsf{Dec}_{\mathsf{HE}})$ be a homomorphic encryption scheme described in Section 2.3 or Section 2.4 . Let L be any full-rank n-dimensional lattice, for any $\epsilon \in (0,1)$, $s \geq \eta_\epsilon(L)$, and $r \geq \frac{6\pi sn}{\epsilon}$. The abstract construction of HE-BKEM is in Figure 7. Suppose HE-BKEM has ϵ_2–blinded blinded keys, a detailed description of these designs follows in Section 5.2.

```
KG(λ) :
    $\mathsf{pk}, \mathsf{sk} \leftarrow \mathsf{KG}_{\mathsf{HE}}(\lambda)$
    $(ek, dk) \leftarrow (\mathsf{pk}, \mathsf{sk})$
    return $ek, dk$

Encap_ek :
    $k \xleftarrow{\$} \mathcal{K}_F$
    $C \leftarrow \mathsf{Enc}_{\mathsf{HE}}(ek, s, k)$
    return $C, k$

Blind_ek(C) :
    $k' \xleftarrow{\$} \mathcal{K}_R$
    $C' \leftarrow \mathsf{Enc}_{\mathsf{HE}}(ek, r, k')$
    $\tilde{C} \leftarrow \mathsf{Add}_{\mathsf{HE}}(C, C')$
    $uk \leftarrow -k' \mod \mathbf{B}$
    return $\tilde{C}, uk$

Decap_dk(C̃) :
    $\tilde{k} \leftarrow \mathsf{Dec}_{\mathsf{HE}}(dk, \tilde{C})$
    return $\tilde{k}$

Unblind_uk(k̃) :
    $k \leftarrow \tilde{k} + uk \mod \mathbf{B}$
    return $k$
```

Figure 7: HE-BKEM, where $\mathbf{B}$ is the basis of the plaintext space $\mathcal{P}$.

5.2 Constructions of random-looking blinded keys

We want the blinded keys to be in the ϵ-blinded blinded key set S with high probability, and we analyze the requirements of the blinding values. We provide two constructions of the ϵ-blinded blinded keys set S as follows.

Construction I. A file encryption key of HE-BKEM is a random element located in a subspace of the underlying HE scheme's message space $\mathcal{M}$. We want to take a small file encryption key k and add a large blinding value k' to produce a slightly larger blinded key $\tilde{k}$, hence, the corresponding key sets should satisfy $\mathcal{K}_F \subseteq \mathcal{K}_R \subseteq \mathcal{K}_B \subseteq \mathcal{M}$.

Suppose $\mathcal{M}$ is HE scheme's message space with generators $1, x, \ldots, x^{n-1}$ and order q, i.e. $\mathcal{M} = \{d_0 + d_1 x + \cdots + d_{n-1}x^{n-1} \mid d_i \in \mathbb{F}_q\}$. The addition of two elements in $\mathcal{M}$ is defined as follows

$$(a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}) + (b_0 + b_1 x + \cdots + b_{n-1}x^{n-1})$$
$$= (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_{n-1} + b_{n-1})x^{n-1}$$

Suppose $\mathcal{K}_F = \{d_0 + d_1 x + \cdots + d_{n-1}x^{n-1} \mid d_i \in \mathbb{Z}_{\lfloor\sqrt{q/2}\rfloor}\}$ and $\mathcal{K}_R = \{d_0 + d_1 x + \cdots + d_{n-1}x^{n-1} \mid d_i \in \mathbb{Z}_{\lfloor q/2 \rfloor}\}$. Notice that for any $c_i \in \{\lfloor\sqrt{q/2}\rfloor, \ldots, \lfloor q/2 \rfloor\}$ and any $a_i \in \mathbb{Z}_{\lfloor\sqrt{q/2}\rfloor}$ there exists a unique $b_i = c_i - a_i \in \mathbb{Z}_{\lfloor q/2 \rfloor}$. In other words, for these restricted $c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$ and for any $a_0 + a_1 x + \cdots a_{n-1}x^{n-1} \in \mathcal{K}_F$ there exists a unique $b_0 + b_1 x + \cdots b_{n-1}x^{n-1} \in \mathcal{K}_R$ such that $(a_0 + a_1 x + \cdots a_{n-1}x^{n-1}) + (b_0 + b_1 x + \cdots b_{n-1}x^{n-1}) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$. Then

$$\mathsf{S} = \{d_0 + d_1 x + \cdots + d_{n-1}x^{n-1} \mid d_i \in \{\lfloor\sqrt{q/2}\rfloor, \ldots, \lfloor q/2 \rfloor\}\}$$

Note that for any $i \in \{0, \ldots, n-1\}$

$$\mathbf{Pr}\left[a_i + b_i \in \{\lfloor\sqrt{q/2}\rfloor, \ldots, \lfloor q/2 \rfloor\} \mid a_i \xleftarrow{\$} \mathbb{Z}_{\lfloor\sqrt{q/2}\rfloor}, b_i \xleftarrow{\$} \mathbb{Z}_{\lfloor q/2\rfloor}\right] = 1 - \frac{\lfloor\sqrt{q/2}\rfloor - 1}{\lfloor q/2\rfloor}.$$

Hence, the probability of a blinded key locates in the ϵ-blinded blinded set is

$$\mathbf{Pr}\left[\tilde{k} = k + k' \in \mathsf{S} \mid k \xleftarrow{\$} \mathcal{K}_F, k' \xleftarrow{\$} \mathcal{K}_R\right] = \left(1 - \frac{\lfloor\sqrt{q/2}\rfloor - 1}{\lfloor q/2\rfloor}\right)^n \approx 1 - \frac{n}{\lfloor\sqrt{q/2}\rfloor}.$$

In this construction, HE-BKEM has ϵ-blinded blinded keys with $\epsilon = n/\lfloor\sqrt{q/2}\rfloor$. For suitably large q, the above ϵ can be made negligible.

Construction II. Let the file encryption key k be an element in a subset of $\mathcal{M}$, we want to add a random blinding value k' from the whole message space $\mathcal{M}$ to produce a random-looking blinded key $\tilde{k}$, hence, the corresponding key sets should satisfy $\mathcal{K}_F \subseteq \mathcal{K}_R = \mathcal{K}_B = \mathcal{M}$.

For any blinded key $\tilde{k} \in \mathcal{M}$ and any file encryption key $k \in \mathcal{K}_F$ there exists a unique random value $k' = \tilde{k} - k \mod \mathbf{B} \in \mathcal{M}$ such that $\tilde{k} = k + k' \mod \mathbf{B}$, thus the ϵ-blinded blinded set S is $\mathcal{M}$ and we have

$$\mathbf{Pr}\left[\tilde{k} = k + k' \mod \mathbf{B} \in \mathsf{S} \mid k \xleftarrow{\$} \mathcal{K}_F, k' \xleftarrow{\$} \mathcal{M}\right] = 1.$$

In this construction, HE-BKEM has ϵ-blinded blinded keys with $\epsilon = 0$.

Remark 2. *Both of these constructions can be applied to our* HE-BKEM *schemes.*

5.3 Construction of fresh-looking blinded encapsulations

We claim that the above constructed HE-BKEM has 4ϵ-blinded blinded encapsulations. The idea is to take the small constant ciphertext and add a ciphertext with big errors and the resulting ciphertext should look like the big error ciphertext. The details are showed in the following lemma.

Lemma 4. *Let* HE-BKEM *be a homomorphic based BKEM with the underlying homomorphic encryption scheme, described in Section 2.3 or Section 2.4 . Let ek be any encapsulation key, recall that the encryption algorithm* $\mathsf{Enc}_{\mathsf{HE}}(ek, s, \cdot)$ *uses the discrete Gaussian distribution $D_{L,s,\mathbf{0}}$ as the error distribution. Suppose $C_0 = \mathsf{Enc}_{\mathsf{HE}}(ek, s, k_0)$ is an encapsulation of the underlying file encryption key k_0. For any $\epsilon \in (0,1)$, $s \geq \eta_\epsilon(L)$, and $r \geq \frac{6\pi s n}{\epsilon}$ the statistical distance of the following distributions is at most 4ϵ*

$$X = \{C_0 \oplus_1 C' \mid k' \xleftarrow{\$} \mathcal{K}_R, C' \leftarrow \mathsf{Enc}_{\mathsf{HE}}(ek, r, k')\}$$
$$Y = \{C \mid k' \xleftarrow{\$} \mathcal{K}_R, C \leftarrow \mathsf{Enc}_{\mathsf{HE}}(ek, r, k_0 \oplus_2 k')\},$$

which means HE-BKEM *has 4ϵ-blinded blinded encapsulation.*

Proof. As in the proof of Corollary 1, assume $\epsilon = 2^{-(n-1)}$. From Lemma 1 we have $\mathbf{Pr}[\mathbf{x} \notin \mathcal{B}(s\sqrt{n}) \mid \mathbf{x} \leftarrow D_{L,s,\mathbf{0}}] \leq \epsilon$, which means the size of the error outputted by the distribution $D_{L,s,\mathbf{0}}$ is upper bounded by $s\sqrt{n}$ expect for negligible probability ϵ.

For Gentry's scheme, suppose $C_0 = k_0 + e_0$, where $e_0 \leftarrow D_{L,s,\mathbf{0}}$. From Corollary 1 we know that for a small error $\|e_0\| \leq s\sqrt{n}$ and big randomness $r \geq \|e_0\| \frac{6\pi\sqrt{n}}{\epsilon}$ the statistical distance between $D_{L,r,\mathbf{0}}$ and $D_{L,r,\mathbf{e_0}}$ is at most 3ϵ. So the following approximation holds

$$C_0 \oplus_1 \mathsf{Enc}_{\mathsf{HE}}(ek, r, k') = k_0 + e_0 + k' + D_{L,r,\mathbf{0}} \approx k_0 + k' + D_{L,r,\mathbf{0}} = \mathsf{Enc}_{\mathsf{HE}}(ek, r, k_0 \oplus_2 k').$$

The above result can be easily adapted to NTRU encryption scheme. □

5.4 Indistinguishability of GHE-BKEM

The following result says GHE-BKEM is an IND-secure BKEM with post-quantum security.

Corollary 2. *Let* GHE-BKEM *be a homomorphic-based BKEM described in Section 5.1. For negligible* $\epsilon_1 = \epsilon, \epsilon_2$*, choose parameters as in Lemma 4, Theorem 1 and Theorem 2. Suppose* GHE-BKEM *has* ϵ_2*-blinded blinded keys. Then* GHE-BKEM *has* $4\epsilon_1$*-blind blinded encapsulation. Furthermore, if there is an algorithm that breaks the indistinguishability of* GHE-BKEM*, i.e. the distinguishing advantage of this algorithm against* GHE-BKEM *getting* r *blinded encapsulation and their blinded decapsulation tuples is non-negligible, then there exists a quantum algorithm that solves worst-case* SIVP.

Proof. By Lemma 4 we know GHE-BKEM has $4\epsilon_1$-blinded blinded encapsulation.

Theorem 5 states that if there is an algorithm that breaks the indistinguishability of GHE-BKEM then there exists an algorithm breaks IND-CPA security of GHE and by Theorem 3 we have a quantum algorithm that solves worst-case SIVP.

□

5.5 Indistinguishability of NTRU-BKEM

The following result says NTRU-BKEM is an IND-secure BKEM with post-quantum security.

Corollary 3. *Let* NTRU-BKEM *be a homomorphic based BKEM constructed in Section 5.1. For negligible* $\epsilon_1 = \epsilon, \epsilon_2$*, choose parameters as in Lemma 4, Lemma 3, and Theorem 4. Suppose* NTRU-BKEM *has* ϵ_2*-blinded blinded keys. Then* NTRU-BKEM *has* $4\epsilon_1$*-blinded blinded encapsulation. Furthermore, if there is an algorithm that breaks the indistinguishability of* NTRU-BKEM*, then there exists a quantum algorithm that solves* $O(\sqrt{n}/\alpha)$*-approximate* SIVP *(or* SVP*) on ideal lattices.*

Proof. By Lemma 4 we know NTRU-BKEM has $4\epsilon_1$-blinded blinded encapsulation.

Theorem 5 states that if there is an algorithm that breaks the indistinguishability of NTRU-BKEM then there exists an algorithm that breaks IND-CPA security of NTRU. By Lemma 3 there exists an adversary solving $\text{R-LWE}^{\times}_{\text{HNF}}$ and by Theorem 4 there exists a quantum algorithm that solves SIVP. □

References

[1] Masayuki Abe, Rosario Gennaro, Kaoru Kurosawa, and Victor Shoup. Tag-kem/dem: A new framework for hybrid encryption and a new analysis of kurosawa-desmedt kem. In Ronald Cramer, editor, *Advances in Cryptology – EUROCRYPT 2005*, pages 128–146, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[2] Martin Albrecht, Shi Bai, and Léo Ducas. A subfield lattice attack on overstretched ntru assumptions. In *Proceedings, Part I, of the 36th Annual International Cryptology Conference on Advances in Cryptology — CRYPTO 2016 - Volume 9814*, pages 153–178, Berlin, Heidelberg, 2016. Springer-Verlag.

[3] Erdem Alkim, Joppe W. Bos, Léo Ducas, Karen Easterbrook, Brian LaMacchia, Patrick Longa, Ilya Mironov, Valeria Nikolaenko, Chris Peikert, Ananth Raghunathan, and Douglas Stebila. FrodoKEM: Learning With Errors Key Encapsulation. https://frodokem.org/files/FrodoKEM-specification-20190330.pdf. Submission to the NIST Post-Quantum Standardization project, round 2.

[4] Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange—a new hope. In *25th USENIX Security Symposium (USENIX Security 16)*, pages 327–343, Austin, TX, August 2016. USENIX Association.

[5] Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS-Kyber (version 2.0). https://pq-crystals.org/kyber/data/kyber-specification-round2.pdf. Submission to the NIST Post-Quantum Standardization project, round 2.

[6] Nimrod Aviram, Kai Gellert, and Tibor Jager. Session resumption protocols and efficient forward security for TLS 1.3 0-rtt. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part II*, volume 11477 of *Lecture Notes in Computer Science*, pages 117–150. Springer, 2019.

[7] Daniel J. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, and Christine van Vredendaal. NTRU Prime: reducing attack surface at low cost. https://ntruprime.cr.yp.to/papers.html.

[8] Joppe W. Bos, Craig Costello, Léo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take off the ring! practical, quantum-secure key exchange from LWE. In Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi, editors, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, pages 1006–1018. ACM, 2016.

[9] Joppe W. Bos, Kristin Lauter, Jake Loftus, and Michael Naehrig. Improved security for a ring-based fully homomorphic encryption scheme. In *Proceedings of the 14th IMA International Conference on Cryptography and Coding - Volume 8308*, IMACC 2013, pages 45–64, Berlin, Heidelberg, 2013. Springer-Verlag.

[10] Colin Boyd, Gareth T. Davies, Kristian Gjøsteen, and Yao Jiang. Offline assisted group key exchange. In Liqun Chen, Mark Manulis, and Steve Schneider, editors, *Information Security - 21st International Conference, ISC 2018, Guildford, UK, September 9-12, 2018, Proceedings*, volume 11060 of *Lecture Notes in Computer Science*, pages 268–285. Springer, 2018.

[11] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ITCS '12, pages 309–325, New York, NY, USA, 2012. ACM.

[12] Cong Chen, Oussama Danba, Jeffrey Hoffstein, Andreas Hülsing, Joost Rijneveld, John M. Schanck, Peter Schwabe, William Whyte, and Zhenfei Zhang. NTRU). https://ntru.org/f/ntru-20190330.pdf. Submission to the NIST Post-Quantum Standardization project, round 2.

[13] Jung Hee Cheon, Kyoohyung Han, Jinsu Kim, Changmin Lee, and Yongha Son. A practical post-quantum public-key cryptosystem based on \textsf splwe. In Seokhie Hong and Jong Hwan Park, editors, *Information Security and Cryptology - ICISC 2016 - 19th International Conference, Seoul, South Korea, November 30 - December 2, 2016, Revised Selected Papers*, volume 10157 of *Lecture Notes in Computer Science*, pages 51–74, 2016.

[14] Jung Hee Cheon, Jinhyuck Jeong, and Changmin Lee. An algorithm for ntru problems and cryptanalysis of the ggh multilinear map without a low-level encoding of zero. *LMS Journal of Computation and Mathematics*, 19(A):255–266, 2016.

[15] Jung Hee Cheon, Andrey Kim, Miran Kim, and Yongsoo Song. Homomorphic encryption for arithmetic of approximate numbers. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology – ASIACRYPT 2017*, pages 409–437, Cham, 2017. Springer International Publishing.

[16] Katriel Cohn-Gordon, Cas Cremers, Luke Garratt, Jon Millican, and Kevin Milner. On ends-to-ends encryption: Asynchronous group messaging with strong security guarantees. In David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang, editors, *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, pages 1802–1819. ACM, 2018.

[17] Ronald Cramer and Victor Shoup. Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack. Cryptology ePrint Archive, Report 2001/108, 2001. https://eprint.iacr.org/2001/108.

[18] Ronald Cramer and Victor Shoup. Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack. *SIAM J. Comput.*, 33(1):167–226, January 2004.

[19] Jan-Pieter D'Anvers, Angshuman Karmakar, Sujoy Sinha Roy, and Frederik Vercauteren. Saber: Module-lwr based key exchange, cpa-secure encryption and cca-secure kem. In Antoine Joux, Abderrahmane Nitaj, and Tajjeeddine Rachidi, editors, *Progress in Cryptology – AFRICACRYPT 2018*, pages 282–305, Cham, 2018. Springer International Publishing.

[20] Alexander W. Dent. A designer's guide to kems. In Kenneth G. Paterson, editor, *Cryptography and Coding*, pages 133–151, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[21] David Derler, Tibor Jager, Daniel Slamanig, and Christoph Striecks. Bloom filter encryption and applications to efficient forward-secret 0-rtt key exchange. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part III*, volume 10822 of *Lecture Notes in Computer Science*, pages 425–455. Springer, 2018.

[22] Junfeng Fan and Frederik Vercauteren. Somewhat practical fully homomorphic encryption. Cryptology ePrint Archive, Report 2012/144, 2012. https://eprint.iacr.org/2012/144.

[23] Craig Gentry. *A Fully Homomorphic Encryption Scheme*. PhD thesis, Stanford, CA, USA, 2009. AAI3382729.

[24] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 197–206, New York, NY, USA, 2008. ACM.

[25] Craig Gentry, Amit Sahai, and Brent Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013*, pages 75–92, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[26] Matthew D. Green and Ian Miers. Forward secure asynchronous messaging from puncturable encryption. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, pages 305–320. IEEE Computer Society, 2015.

[27] Felix Günther, Britta Hale, Tibor Jager, and Sebastian Lauer. 0-rtt key exchange with full forward secrecy. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part III*, volume 10212 of *Lecture Notes in Computer Science*, pages 519–548, 2017.

[28] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In Joe Buhler, editor, *Algorithmic Number Theory, Third International Symposium, ANTS-III, Portland, Oregon, USA, June 21-25, 1998, Proceedings*, volume 1423 of *Lecture Notes in Computer Science*, pages 267–288. Springer, 1998.

[29] Dennis Hofheinz and Eike Kiltz. Secure hybrid encryption from weakened key encapsulation. In Alfred Menezes, editor, *Advances in Cryptology - CRYPTO 2007*, pages 553–571, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[30] Andreas Hülsing, Joost Rijneveld, John Schanck, and Peter Schwabe. High-speed key encapsulation from ntru. In *International Conference on Cryptographic Hardware and Embedded Systems (CHES)*, pages 232–252. Springer, 2017.

[31] Kaoru Kurosawa and Yvo Desmedt. A new paradigm of hybrid encryption scheme. In Matt Franklin, editor, *Advances in Cryptology – CRYPTO 2004*, pages 426–442, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[32] Adriana López-Alt, Eran Tromer, and Vinod Vaikuntanathan. On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 1219–1234, New York, NY, USA, 2012. ACM.

[33] Xianhui Lu, Yamin Liu, Dingding Jia, Haiyang Xue, Jingnan He, Zhenfei Zhang, Zhe Liu, Hao Yang, Bao Li, and Kunpeng Wang. LAC Lattice-based Cryptosystems. Submission to the NIST Post-Quantum Standardization project, round 2.

[34] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Monaco / French Riviera, May 30 - June 3, 2010. Proceedings*, volume 6110 of *Lecture Notes in Computer Science*, pages 1–23. Springer, 2010.

[35] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. *J. ACM*, 60(6):43:1–43:35, November 2013.

[36] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-lwe cryptography. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings*, volume 7881 of *Lecture Notes in Computer Science*, pages 35–54. Springer, 2013.

[37] Moxie Marlinspike and Trevor Perrin. The X3DH key agreement protocol. https://signal.org/docs/specifications/x3dh/, November 2016.

[38] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM J. Comput.*, 37(1):267–302, April 2007.

[39] NIST Post-Quantum Cryptography Standardization. https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization. Accessed: 2019-11-15.

[40] Damien Stehlé and Ron Steinfeld. Making ntru as secure as worst-case problems over ideal lattices. In *Proceedings of the 30th Annual International Conference on Theory and Applications of Cryptographic Techniques: Advances in Cryptology*, EUROCRYPT'11, pages 27–47, Berlin, Heidelberg, 2011. Springer-Verlag.

[41] The messaging layer security (MLS) protocol. Internet draft, in progress. https://datatracker.ietf.org/wg/mls/about. Accessed: 2019-11-25.