Extended version of the paper accepted at the 35th USENIX Security Symposium (USENIX Security 2026).

Crypto Wars in Secure Messaging: Covert Channels in Signal Despite Leaked Keys*

Rosario Giustolisi[1], Gabriele Lenzini[2], Chuanwei Lin[2,3],
Mohammadamin Rakeei[2], and Andy Rupp[2,4]

[1] IT University of Copenhagen
rosg@itu.dk
[2] University of Luxembourg
{gabriele.lenzini,andy.rupp}@uni.lu, m.amin.rakeei@gmail.com
[3] Karlsruhe Institute of Technology
lin@chuanwei.eu
[4] KASTEL Security Research Labs

Abstract. End-to-end encryption (E2EE) is the foundation of modern secure messaging, with the Signal protocol as the de facto standard in applications such as Signal, WhatsApp, Facebook Messenger and Google Messages. At the same time, the deployment of E2EE has led to growing pressure from authorities to decrypt user traffic under law enforcement. This raises a critical question: if an adversary can routinely decrypt Signal messages (for example via a mandated access or a leaked key), can users still communicate securely and covertly?

We address this question through the lens of anamorphic encryption, which enables hidden communication within seemingly legitimate ciphertexts, even against an adversary who can decrypt them. We design two constructions that embed covert channels into the existing Signal Double Ratchet protocol. Concretely, we show how to embed covert messages (i) into Diffie-Hellman keys used in the asymmetric ratchet, or (ii) into authentication tags produced in the symmetric ratchet. Our techniques are compatible with existing Signal-style deployments and require no changes by the service provider.

We formalize security in threat models that capture adversaries with decryption capabilities granted through lawful-access mechanisms, and prove that the resulting protocol transcripts are indistinguishable from those of standard Signal. We implement our constructions in the official Signal library and Android client, and show that they incur low overhead and are practical in real-world settings. Our results show that covert communication channels can persist even when conventional E2EE guarantees are compromised.

Keywords: Secure messaging, Signal, Double Ratchet, anamorphic encryption, covert channels

1 Introduction

End-to-end encryption (E2EE) has become essential infrastructure for private communication. The Signal protocol [PMS25,MP16] is the foundation of messaging for billions of users across Signal, WhatsApp, Facebook Messenger, and Google Messages. Its Double Ratchet algorithm provides forward secrecy and post-compromise security, properties that hold up under extensive formal analysis [KBB17,CJSV22,BFG+22,CGCD+20,ACD19,VGIK20,CRT24] and that users increasingly rely on for secure communication.

At the same time, E2EE is under growing political pressure. Proposals in the EU [Eur25], the UK [Val23], Sweden [Dig25], and Australia [Cro25] aim to require providers to enable decryption of E2EE traffic through key escrow, client-side scanning, or other technical mandates. Signal has stated it would leave certain markets rather than comply [Dig25,Cro25]. Other providers may comply, whether by choice or under legal pressure. For users who depend on E2EE, including journalists and sources, activists and organizers, lawyers and clients, and people in abusive situations, a practical question follows: if an adversary can decrypt your messages, is there still anything you can communicate privately?

* Authors are listed alphabetically by last name. Mohammadamin Rakeei and Chuanwei Lin contributed equally.

We explore this question using *anamorphic encryption* [PPY22]. Informally, anamorphic encryption lets a sender embed a second, covert message into a normal-looking ciphertext. An adversary who holds the regular decryption key can decrypt the visible message, yet cannot distinguish anamorphic ciphertexts from ordinary ones. Only someone with a separate *double key* can recover the hidden content. This contrasts with cover-media steganography, which hides data in objects such as images or audio and depends on access to suitable cover distributions. In practice, cover availability and robustness can be a bottleneck under sanitization, filtering, or censorship [KPP+23]. Anamorphic encryption sidesteps these constraints by operating inside cryptographic protocols. The adversary may decrypt the overt communication and inspect the resulting protocol values, so the covert channel must remain hidden under cryptographic inspection rather than rely on external cover objects. Prior work has developed the theory of anamorphic encryption [PPY22,BGH+24,CGM24,KPP+23,WCHY23,CGM25], but it has not been instantiated inside a widely deployed messaging protocol.

Applying anamorphic encryption to Signal raises protocol-specific challenges. Signal is not a single encryption call. The Double Ratchet evolves state over time, mixes asymmetric and symmetric steps, and must remain correct under message loss and out-of-order delivery. Protocol messages also carry multiple cryptographic values, such as DH public keys, ciphertexts, and authentication tags. A covert channel must blend into this full transcript, not just into one isolated primitive. It must also match the surveillance setting, since different disclosure patterns may reveal different pieces of key material.

We show that anamorphic communication can be realized within the Signal Double Ratchet with provable security and practical performance. We build two covert channels that hide data in different protocol components: one embeds covert messages in Diffie-Hellman (DH) public keys, and the other embeds covert messages in MAC authentication tags. In both cases, the resulting transcripts are indistinguishable from standard Signal executions for an adversary that can decrypt ordinary messages under the relevant key-disclosure assumptions. Empirically, the MAC-based channel incurs negligible overhead, and the optimized DH-based channel adds only a few milliseconds of overhead for short covert messages. Together, we show that covert communication can persist inside widely deployed E2EE protocols even when their primary confidentiality guarantees are weakened.

1.1 Contributions

We design, analyze, and implement covert channels for the Signal Double Ratchet. Our contributions are as follows.

Two practical ways to embed covert messages into Signal. We propose two families of constructions that hide data in different parts of the protocol and therefore apply under different key-disclosure patterns.

- **DH-based channel.** We embed covert messages into the Diffie-Hellman public keys generated during the asymmetric ratchet. The construction produces public keys that are computationally indistinguishable from honestly generated ones, even to an adversary who learns the Double Ratchet secrets that suffice to decrypt ordinary messages. We provide a symmetric double-key variant that requires pre-sharing the double key, and an asymmetric variant that allows covert initiation without prior coordination, assuming the receiver has published a corresponding public key.
- **MAC-based channel.** We embed covert messages into authentication tags in the symmetric ratchet. This construction requires no pre-shared double key, and applies in settings where surveillance yields only partial state. In particular, if the adversary learns either the root key or the DH private keys, but not both, then the derived MAC keys remain unknown to the adversary and can carry covert messages.

To our knowledge, these are the first anamorphic constructions realized within a deployed E2EE protocol.

Security definitions and proofs. We formalize *anamorphic security* via transcript indistinguishability: an adversary with access to the disclosed key material cannot distinguish anamorphic executions from standard Signal executions. We prove that our constructions satisfy this notion under standard cryptographic assumptions. We further extend both constructions with transcript binding to obtain IND-CCA2-style security against active adversaries who may inject, replay, or modify messages.

Concrete instantiation, optimization, and evaluation. We provide concrete instantiations and optimizations, and validate them in an end-to-end implementation.

Our DH-based channel is instantiated over X25519 using Edwards-Montgomery mappings. On the receiver side, naively extracting a covert message would require an exhaustive search over the message space, which quickly becomes impractical. We therefore introduce a lookup-table optimization that replaces exhaustive search with a single group operation followed by a table lookup, making a 32-bit covert message per DH ratchet step practical. The MAC-based channel supports messages up to the MAC length, which is 64 bits in our implementation.

We implement our constructions in a Signal-compatible Double Ratchet library and integrate them into an Android client, then evaluate their performance. In our measurements, the MAC-based channel incurs negligible overhead compared to unmodified Signal, while the optimized DH-based channel avoids the otherwise high extraction latency and achieves practical overhead for short covert messages. Both designs require no Signal server-side changes and remain compatible with existing Signal-style deployments. We also examine timing side channels, since decryption latency can leak whether a message carries covert content, and we discuss mitigations toward constant-time behavior.

1.2 Threat Model and Applicability

We consider an adversary who obtains Double Ratchet key material through a surveillance mechanism and can decrypt ordinary Signal messages. Different settings expose different subsets of the ratchet state, and our constructions are tailored to these disclosure patterns.

Full state disclosure. In the strongest setting, the adversary learns the complete Double Ratchet state used to decrypt and authenticate messages, including the root key, DH private keys, chain keys, and derived message keys. This can arise from broad key-disclosure requirements (as contemplated by recent EU and UK proposals [Eur25,Val23]), from taking a device and recovering secrets through forensic analysis, or from a court order that forces disclosure of device credentials and stored key material.

Our DH-based constructions address this setting. They remain secure even if all Double Ratchet secrets are disclosed, provided the parties have established a separate *double key* that is not obtained through the same surveillance mechanism.

Partial state disclosure. In some settings, surveillance reveals only part of the ratchet state while still enabling the adversary to decrypt the contents of normal Signal messages. Full state disclosure would require the authority to capture and maintain the entire Double Ratchet state for every monitored session, or to have direct control over the relevant infrastructure or devices.

In practice, many surveillance mechanisms fall short of this. Legal frameworks may authorize access to message contents without mandating disclosure of all cryptographic state, and forensic tools may extract some keys but not others. The cost of recording every ephemeral value for every session may exceed operational budgets. In these cases, the adversary obtains the per-message decryption key and perhaps some additional ratchet state, but not all of it. In particular, the adversary might learn root key material (or information sufficient to derive it) without obtaining the ephemeral DH private keys generated during later ratchet steps. Conversely, the adversary might obtain ephemeral DH private keys from runtime compromise or short-term access, without obtaining the earlier root key material. These patterns also capture operational realities such as incomplete device backups or selective legal authority that grants access to only some session secrets.

Our MAC-based construction exploits this separation. Signal derives the per-message MAC key from secrets that depend on both the root key and DH outputs. If the adversary learns only one side of this state, then the MAC key remains underivable. In that case, covert messages embedded in authentication tags remain confidential and indistinguishable, without requiring any pre-shared double key.

Establishing the double key. The DH-based constructions assume a double key unknown to the adversary. In our threat model, this key is not part of Signal's protocol state and is therefore outside the key material disclosed by the modeled lawful-access mechanism.

For the symmetric variant, both parties must share a double-key secret before covert communication begins. In practice, this can be done through in-person exchange (a passphrase, QR code, or NFC tap), derivation from shared history that predates the surveillance setting, or a one-time bootstrap over a channel outside the adversary's reach, such as physical mail or a trusted intermediary.

For the asymmetric variant, only the receiver holds the double secret key, while the corresponding public key can be published. A journalist, activist, or organization can post a double public key on a personal website, enabling any sender to initiate covert communication without prior coordination.

Unless refreshed out of band, we treat the double key as fixed for the corresponding covert context. Thus, the DH-based channel provides transcript undetectability only while the double key remains secret, and does not provide covert-channel forward secrecy or deniability once it is exposed: with the key and transcript, an adversary can recompute and verify past covert contents. A forward-secure variant could derive per-message double keys through a KDF chain analogous to the Double Ratchet, which we leave to future work.

The MAC-based construction requires no double key. It enables covert communication whenever surveillance is partial in the sense above, which makes it applicable even for users who cannot establish an out-of-band secret.

Scope and Non-Goals. Our constructions provide *covert content confidentiality*: the hidden message remains secret, and its presence is computationally undetectable from protocol transcripts under our assumptions.

We do not protect metadata, including who communicates with whom, timing, or frequency. Addressing metadata requires techniques such as anonymous routing or traffic shaping. Our base constructions target surveillance that is passive with respect to message delivery, while our transcript-binding variants address active adversaries who may inject, replay, modify, or reorder protocol messages. We do not address active man-in-the-middle attacks that break endpoint authentication or establish a separate compromised session with either party.

Covert bandwidth. The optimized DH-based channel embeds 32 bits per ratchet step, and the MAC-based channel embeds 64 bits per message. These rates suffice for short sensitive content (identifiers, coordinates, status flags) or for bootstrapping a separate secure channel (e.g., exchanging a session key), but not for bulk transfer.

1.3 Related Work

Secure messaging. End-to-end encrypted messaging aims to keep message contents available only to the communicating endpoints, even when intermediaries relay or store the traffic [KCKG24]. Modern secure messaging builds on Off-the-Record Messaging (OTR) [BGB04], which introduced forward secrecy and deniable authentication for interactive conversations. TextSecure [Mar13] brought ratcheted key evolution to mobile messaging and evolved into the Signal protocol, now widely deployed.

Related approaches pursue stronger deniability or concealment goals. Deniable encryption allows coerced users to explain a ciphertext transcript as encrypting alternative plaintexts [CDNO97,SW14]. Ephemeral messaging systems reduce retained local history through default deletion [XCW+16]. Network steganography hides communication within cover traffic or cover distributions [Sim84,KJGR21]. These techniques address important but different threat models: they do not directly study whether covert communication can persist inside the cryptographic values of a deployed E2EE protocol after ordinary message-decryption keys are disclosed.

Signal and the Double Ratchet. Signal combines the X3DH asynchronous key agreement protocol [MP16] with the Double Ratchet [PMS25]. The Double Ratchet has been analyzed extensively. Alwen et al. [ACD19] gave a modular security treatment, Poettering and Rösler [PR18] studied compromise resilience, and Cohn-Gordon et al. [CGCD+20] proved its guarantees in symbolic and computational models. Bienstock et al. [BFG+22] proposed a universally composable framework and introduced the Triple Ratchet for stronger adversarial settings.

Several recent works adapt Signal-style messaging to new threats. Teng and Rasmussen [TR24] studied in-protocol key verification, while Martiny et al. [MKA+21] analyzed Sealed Sender and showed statistical disclosure risks to sender anonymity. PQXDH [KS24] provides a post-quantum replacement for X3DH. Dodis et al. [DJK+25] proposed a Triple Ratchet that combines a post-quantum ratchet with erasure coding, and ML-KEM Braid [Sch25], building on sparse continuous key agreement [ADJ+25], supports post-quantum continuous key agreement compatible with the Double Ratchet. Dodis et al. [DMSDT25] studied subversion resistance using Reverse Firewalls to mitigate attacks that tamper with implementations.

Anamorphic encryption. Since the introduction of anamorphic encryption by Persiano et al. [PPY22], subsequent work has refined and extended the foundations. Wang et al. [WCHY23] formalized sender-

side anamorphic encryption for pseudorandom public-key schemes. Banerjee et al. [BPRS26] introduced a framework for public-key anamorphic multi-message extensions without relying on randomness recovery. Banfi et al. [BGH+24] revisited definitions and proposed stronger security notions. Catalano et al. [CGM24,CGM25] developed new constructions including homomorphic realizations. Kutyłowski et al. [KPP+23] argued that anamorphic channels are inherent to many encryption schemes, framing this as a "self-anti-censorship" property.

A parallel line of work studies *anamorphic-resistant encryption* (ARE), which aims to make covert channels impractical. Dodis and Goldin [DG25] and Carnemolla et al. [CCGM25] proposed ARE schemes with explicit resistance goals. Avitabile et al. [ABG+26] gave a practical construction without random oracles and limited covert bandwidth to logarithmically many bits. A more detailed survey appears in Appendix A.

Applications and gap. Anamorphic techniques have been explored beyond messaging, for example in electronic voting [GRL25] and in LLM-generated text channels [GR25]. Closely related, but not based on anamorphic encryption, Wink [CSS23] provides plausibly deniable messaging over the public channels of existing E2EE applications, including Signal. Wink uses TrustZone to inject encrypted hidden messages into ordinary random coins such as salts and IVs, preserving the application wire format under partial device compromise and compelled disclosure.

Our work is complementary: we instantiate anamorphic encryption directly inside the protocol logic of a deployed E2EE system, namely the Signal Double Ratchet, by embedding covert content in DH ratchet public keys and MAC tags. This gives protocol-native constructions under explicit ratchet-state disclosure models, rather than a deniable-messaging layer outside the ratchet.

1.4 Paper Organization

Section 2 reviews the Double Ratchet and anamorphic encryption. Section 3 formalizes our system model and threat assumptions. Section 4 presents the DH-based constructions, and Section 5 presents the MAC-based construction. Section 6 reports performance measurements. Section 7 discusses limitations and future directions. Appendix A gives a detailed related-work survey, and Appendices B and C give formal security proofs.

2 Background

This work focuses on the two main components of the Double Ratchet protocol, namely the DH and symmetric ratchets. We therefore focus on the cryptographic primitives used in these components, and refer the reader to [PMS25] for additional details. Our description and Figures 1, 2, and 3 follow the Signal Double Ratchet specification [ACD19].

2.1 Cryptographic functions

The Double Ratchet protocol is primarily built upon ECDH key exchange and the HKDF key derivation function [Kra10]. In the following, we assume familiarity with standard cryptographic notions and outline the relevant cryptographic functions, specifying for each its interface, implementation, and key features.

- $\mathsf{Keygen}(\lambda) \to (\mathsf{sk}, \mathsf{pk})$: Given a security parameter λ, the key generation algorithm outputs a key pair $(\mathsf{sk}, \mathsf{pk})$. In the DH ratchet recalled below, it generates an ECDH key pair using either the X25519 or X448 function, denoted as $(\mathsf{sk}, \mathsf{pk} := \mathsf{sk} \cdot G)$, where G is the base point of the elliptic curve, $\mathsf{sk} \in \mathbb{Z}_p$ is a randomly chosen private key, and p is the prime order of the subgroup $\mathbb{G}$. We use additive notation for elliptic-curve groups: scalars lie in $\mathbb{Z}_p$, group elements lie in $\mathbb{G}$, and scalar multiplication of a point G by a scalar a is written $a \cdot G$. The default curve in the Signal protocol is Curve25519.
- $\mathsf{DH}(\mathsf{sk}_a, \mathsf{pk}_b) \to \mathsf{dh}_{\mathsf{out}}$: Given A's secret key sk_a and B's public key pk_b, the DH function computes the shared secret $\mathsf{dh}_{\mathsf{out}} := \mathsf{sk}_a \cdot \mathsf{pk}_b$. Note that $\mathsf{DH}(\mathsf{sk}_a, \mathsf{pk}_b) = \mathsf{DH}(\mathsf{sk}_b, \mathsf{pk}_a)$.
- $\mathsf{PRF}(\mathsf{prf}_{\mathsf{key}}, x) \to y$: Given a key $\mathsf{prf}_{\mathsf{key}}$ and an input x, the pseudorandom function (PRF) outputs an element y. It is required that the output be computationally indistinguishable from a value produced by a truly random function with the same domain and codomain.

- $\mathsf{KDF}(\mathsf{kdf}_{\mathsf{key}}, x) \rightarrow y$: Given a key $\mathsf{kdf}_{\mathsf{key}}$ and an input x, the key derivation function outputs an element y. It is required that the output be computationally indistinguishable from a value sampled uniformly at random from the same range. Moreover, when the KDF output is split into several parts, knowledge of one part should not leak information about the others.

  The KDFs for the Double Ratchet are HMAC-based. They consist of two components: first, the extract function takes the salt and input key material and produces a PRF key. Then, the expand function takes the PRF key, auxiliary information, and the desired output length, and outputs the derived key. We specify each KDF involved in the ratchet as follows.
- $\mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}, \mathsf{dh}_{\mathsf{out}}) \rightarrow (\mathsf{rk}_{\mathsf{new}}, \mathsf{ck})$: Given an old root key rk as salt and the shared secret $\mathsf{dh}_{\mathsf{out}}$ as input key material, the HMAC-based key derivation function (HKDF) generates a new root key $\mathsf{rk}_{\mathsf{new}}$ and a new chain key ck.
- $\mathsf{KDF}_{\mathsf{ck}}(\mathsf{ck}) \rightarrow (\mathsf{ck}_{\mathsf{new}}, \mathsf{mk})$: Given the old chain key ck, the HKDF generates a new chain key $\mathsf{ck}_{\mathsf{new}}$ and a message key mk. In this function, a constant value is used as the salt.
- $\mathsf{Enc}(\mathsf{mk}, \mathsf{pt}, \mathsf{ad}) \rightarrow (\mathsf{aes}_{\mathsf{key}}, \mathsf{mac}_{\mathsf{key}}, \mathsf{IV}, \mathsf{ct}, \mathsf{mac})$: Given a message key mk, a plaintext pt and the associated data ad, the encryption algorithm, implemented as Authenticated Encryption with Associated Data (AEAD) in the protocol, outputs an AES key $\mathsf{aes}_{\mathsf{key}}$, a MAC key $\mathsf{mac}_{\mathsf{key}}$, the initialization vector IV, a ciphertext ct and a MAC mac. We include the derived keys in the interface only for analyzing key-disclosure patterns. An ordinary Signal sender outputs only ct and mac. It proceeds in the following three steps, combining CBC and HMAC to ensure misuse resistance in cases where a key is inadvertently reused:
  1. $\mathsf{KDF}_{\mathsf{mk}}(\mathsf{mk}) \rightarrow (\mathsf{aes}_{\mathsf{key}}, \mathsf{mac}_{\mathsf{key}}, \mathsf{IV})$: Given mk as input key material, the HKDF outputs $\mathsf{aes}_{\mathsf{key}}$, $\mathsf{mac}_{\mathsf{key}}$ and IV. The HKDF salt is a zero-filled sequence that matches the output length of the hash function.
  2. $\mathsf{AES}_{\mathsf{CBC}}(\mathsf{aes}_{\mathsf{key}}, \mathsf{IV}, \mathsf{pt}) \rightarrow \mathsf{ct}$: Given $\mathsf{aes}_{\mathsf{key}}$, IV and pt, the encryption function, AES-256 in CBC mode with PKCS#7 padding, outputs a ciphertext ct.
  3. $\mathsf{MAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad}) \rightarrow \mathsf{mac}$: Given $\mathsf{mac}_{\mathsf{key}}$, ct and ad, the MAC function returns a mac.

2.2 The Double Ratchet protocol

Many messaging applications use Signal's Double Ratchet protocol to generate message keys in support of E2EE. This protocol ensures secure messaging by constantly updating encryption keys. It provides both forward secrecy and post-compromise secrecy. In a nutshell, every time a message is sent, a new key is generated, limiting the exposure of past or future messages if a key is compromised. The protocol bootstraps from a shared secret that sender and receiver have previously agreed on through a secure key agreement protocol. In Signal, this protocol is X3DH [MP16]. The Double Ratchet protocol consists of the symmetric ratchet and the DH ratchet.

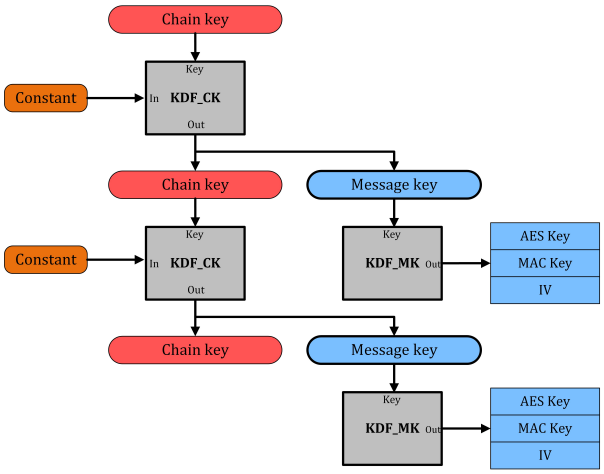

Fig. 1: The symmetric ratchet in [ACD19]

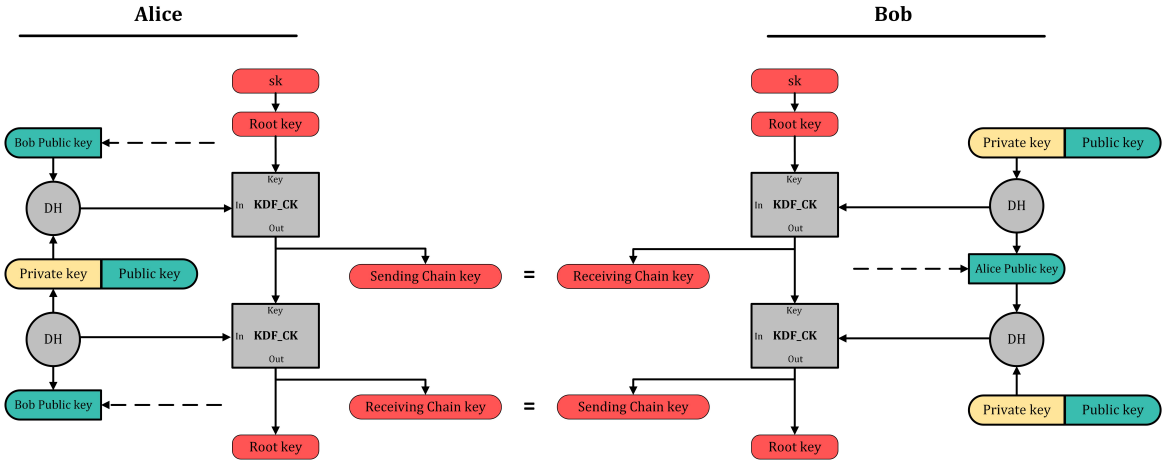

Fig. 2: The DH ratchet in [ACD19]

Symmetric ratchet. The goal of the symmetric ratchet is to generate a new message key for each message sent or received. It uses the functions $\mathsf{KDF}_{\mathsf{ck}}$ and Enc. The symmetric ratchet starts from an old chain key and derives a new chain key and a message key, as shown in Figure 1.

DH ratchet. The goal of the DH ratchet is to provide forward secrecy and post-compromise security against an attacker who gains access to a chain key. In the DH ratchet, each party generates a DH key pair, which serves as their current ratchet key pair $(\mathsf{sk}, \mathsf{pk})$. Every message is sent with a header that includes the sender's current public key pk. When one party receives a new public key from the other, they execute DH and replace their own key pair with a fresh one. This update triggers $\mathsf{KDF}_{\mathsf{ck}}$ to generate a new chain key, ensuring continuous key updates. The DH ratchet is illustrated in Figure 2.

In short, the general rules to run either ratchet can be summarized as follows:

- On sending or receiving a new message, a symmetric ratchet step updates the sending or receiving chain to generate a new message key.
- On receiving a new ratchet public key, a DH ratchet step replaces the chain keys before executing the symmetric ratchet.

Combining the DH ratchet and the symmetric ratchet forms the Double Ratchet protocol as depicted in Figure 3, which illustrates an example of a Double Ratchet scenario from Alice's point of view. In this flow, Alice receives an initial ratchet public key from Bob, then performs a DH ratchet step to obtain her first sending chain key. She then sends two messages within this chain. Subsequently, Bob replies with a new message that includes a new ratchet public key. Alice generates a new receiving chain key and derives the corresponding message keys for the messages received from Bob. The final part of this communication involves Alice sending a message along with a new ratchet public key. Note that Bob performs the same operations on his side, generating the corresponding message keys. Since only Alice and Bob share the initial root key generated via X3DH, no one else can compute the sending and receiving keys, and consequently, the message keys. This shared secret also ensures security against man-in-the-middle attacks by the server, which only relays the communication.

Figure 2 also illustrates that compromising a sending or receiving chain key does not affect the security of future keys once a later DH ratchet step mixes in a fresh $\mathsf{dh}_{\mathsf{out}}$ known only to Alice and Bob.

2.3 Anamorphic encryption

Anamorphic encryption [PPY22] protects communication even when a dictator restricts the permitted cryptographic primitives and holds the ordinary decryption key. It augments the standard encryption mode with an anamorphic mode: a sender can use a *double key* to embed a hidden message that only the intended receiver can recover. To the dictator, the keys and ciphertexts produced in standard and anamorphic modes are indistinguishable.

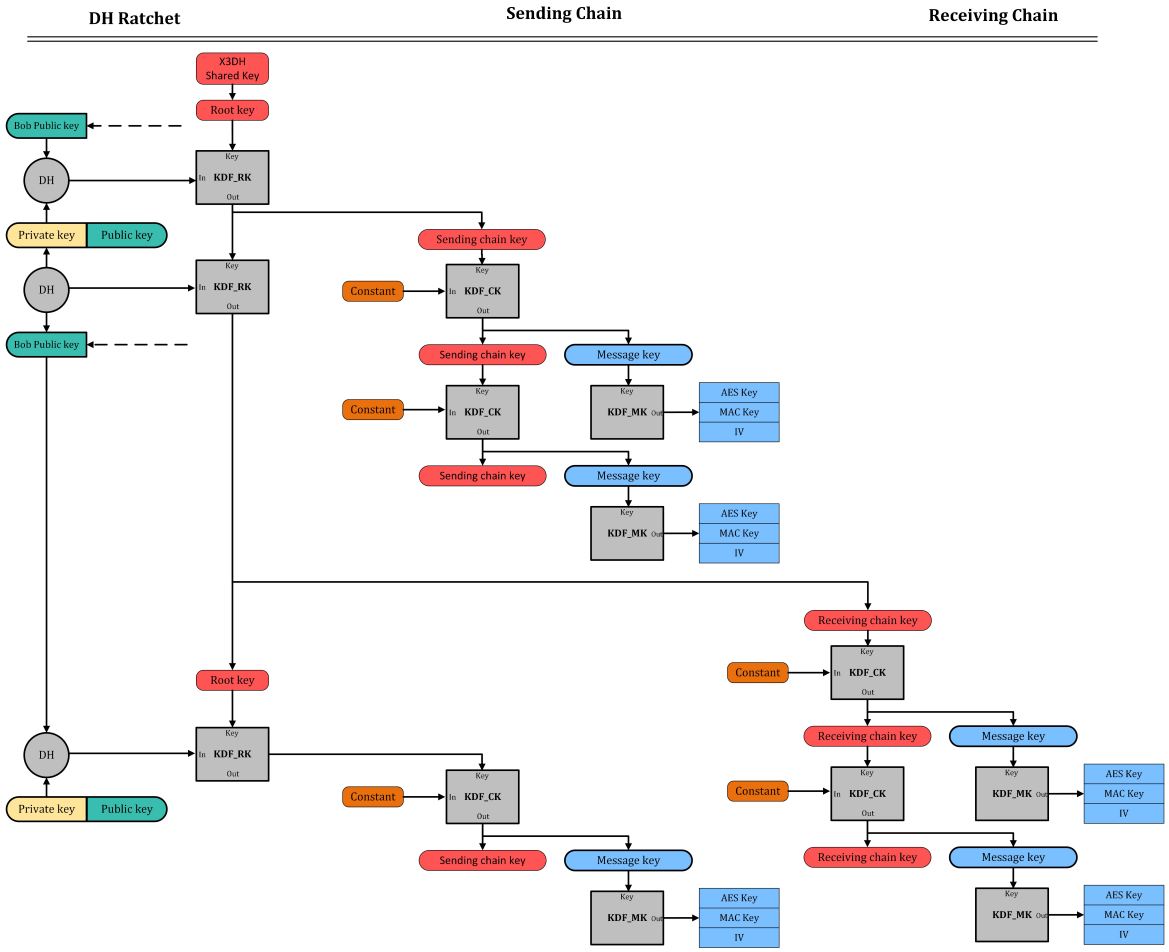

Fig. 3: An instance of the Double Ratchet protocol for Alice in [ACD19]

An anamorphic encryption scheme linked to a standard encryption scheme (Gen, Enc, Dec) consists of three algorithms (aGen, aEnc, aDec). In anamorphic mode, the aGen algorithm returns an anamorphic public key apk, an anamorphic secret key ask and a double key dk. The aEnc algorithm encrypts a normal message msg together with an anamorphic message amsg under dk, producing a ciphertext from which aDec can recover the hidden message using the double key.

For the Double Ratchet, we instantiate the same three conceptual components. Specifically, there is a double-key generation algorithm aGen that samples a double key or double-key pair. Using this double key material, the anamorphic message can be embedded and recovered. We omit the explicit invocation of the double-key generation algorithm and assume that, when necessary, the double key material is generated separately.

3 System Model and Threat Assumptions

This section formalizes the system model, adversary capabilities, and security goals for our anamorphic constructions. We define the key-disclosure patterns under which each construction operates and the security properties we achieve.

3.1 System Model

We consider two parties, Alice and Bob, communicating via the Signal Double Ratchet protocol through a centralized server. The server relays all messages and observes the full protocol transcript, including message headers, ciphertexts, and authentication tags. We assume the underlying cryptographic primitives are instantiated as specified in the Signal documentation [PMS25].

Our constructions augment the standard Double Ratchet with an *anamorphic mode* that embeds covert messages into protocol values. The mechanism for enabling this mode varies by construction and is detailed in subsequent sections.

3.2 Adversary Model

We model the adversary $\mathcal{A}$ as controlling the Signal server and obtaining enough cryptographic state to decrypt ordinary Signal messages through a key-disclosure mechanism. Formally, $\mathcal{A}$ receives the complete protocol transcript between Alice and Bob, including all message headers and ciphertexts, together with a subset of Double Ratchet state determined by the *key-disclosure pattern* summarized in Table 1.

We consider two adversary classes.

Passive adversary. The adversary observes the transcript and disclosed keys but does not modify, inject, or reorder messages. This models surveillance that decrypts communications without interfering with delivery.

Active adversary. The adversary may additionally inject, replay, modify, or reorder messages. Our base constructions achieve security against passive adversaries. We present transcript-binding upgrades that achieve security against active adversaries.

3.3 Key-Disclosure Patterns

Different surveillance mechanisms expose different subsets of Double Ratchet state. Revealing some keys may enable derivation of others. For instance, exposing both the root key and the DH private keys lets the adversary reconstruct chain and message keys. We identify two principal disclosure patterns corresponding to our two construction families. Throughout, we assume $\mathsf{aes}_{\mathsf{key}}$ is disclosed, so $\mathcal{A}$ can always decrypt ordinary message plaintexts.

Full state disclosure. The adversary learns the complete Double Ratchet state, including the root key, DH private keys, and all derived chain, message, MAC, and AES keys. This pattern arises from comprehensive key-escrow requirements or forensic device compromise.

Our DH-based constructions (Sym-DH and Asym-DH) address this setting. They require a *double key* established outside the Double Ratchet and unknown to $\mathcal{A}$. Covert messages are embedded into DH public keys, which remain indistinguishable from honestly generated keys even when all other state is disclosed. Sym-DH requires pre-sharing the double key between endpoints, while Asym-DH allows the receiver to publish a double public key, enabling covert initiation without prior coordination.

Partial state disclosure. The adversary learns either the root key or the DH private keys, but not both simultaneously. Since chain keys are derived via $\mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}, \mathsf{DH}(\mathsf{sk}_A, \mathsf{pk}_B))$, partial disclosure prevents the adversary from deriving chain, message, and MAC keys through the standard KDF chain. We assume the adversary still obtains $\mathsf{aes}_{\mathsf{key}}$. This enables decryption of ordinary plaintexts while leaving $\mathsf{mac}_{\mathsf{key}}$ unknown.

This model does not cover full device compromise or disclosure of the whole message key mk. If mk or $\mathsf{mac}_{\mathsf{key}}$ is disclosed, the MAC-based channel is no longer hidden, as captured by the full-state setting. Instead, the model captures surveillance mechanisms whose purpose is content recovery. The authority obtains the value needed to decrypt the visible message, or the resulting plaintext, without receiving the authentication key or all intermediate ratchet secrets. Such a design is sufficient for reading ordinary messages and avoids storing the full evolving Double Ratchet state for every monitored session. It also captures incomplete forensic or backup-based extraction, where one side of the root-key/DH computation is available but the other is not. Under this scoped leakage, $\mathsf{aes}_{\mathsf{key}}$ can be available for ordinary decryption while $\mathsf{mac}_{\mathsf{key}}$ remains unknown.

This pattern captures scenarios such as root key disclosure without access to ephemeral DH private keys (for example, backup-based surveillance that captures long-term state but not runtime ephemeral values), or DH private key disclosure without root key access (for example, short-term device compromise during an active session).

Our MAC-based construction (Sym-MAC) builds on this separation. It requires no pre-shared double key. Covert messages are embedded into authentication tags using the MAC key that $\mathcal{A}$ cannot derive.

Table 1 summarizes the key-disclosure tolerance of each construction. A checkmark indicates that the corresponding key may be disclosed while preserving anamorphic security. The diamond (◊) for Sym-MAC indicates that *either* the root key *or* the DH private key may be leaked, but not both.

Table 1: Anamorphic constructions under different key-disclosure patterns. ◊ allows either the root key or the DH private key to be leaked.

Construction	Leak					No PSK
	Root key	DH private key	Chain/Message key	MAC key	AES key	required
Sym-DH	✓	✓	✓	✓	✓	×
Asym-DH	✓	✓	✓	✓	✓	✓
Sym-MAC	◊	◊	×	×	✓	✓

3.4 Scope and Assumptions

We consider key-disclosure mechanisms that enable decryption without changing the message format or protocol logic. Surveillance designs that modify other aspects of the Double Ratchet, such as introducing additional protocol messages or altering the ratchet structure, are out of scope because their security impact requires separate analysis.

Our constructions do not change Signal's ordinary out-of-order delivery handling. Covert delivery is tied only to protocol values already carried by delivered Signal messages, namely DH public keys or MAC tags. A reordered or delayed Signal message therefore delays the corresponding covert carrier, but ordinary Signal message processing remains unchanged. For the block-based CCA variants, covert extraction waits until all carriers in the block have been received.

We omit the double key generation process and assume that double keys are established through a separate channel outside of Signal. For Sym-DH, this requires a pre-shared secret. For Asym-DH, the receiver publishes a double public key. For Sym-MAC, no double key is needed.

3.5 Security Goals

For our constructions, we define and prove *anamorphic security*, which ensures that an adversary cannot distinguish anamorphic mode from standard mode. That is, the adversary cannot detect whether the observed protocol components contain an anamorphic message. We formalize this as follows.

Definition 1 (Anamorphic Security). *A construction achieves anamorphic security with respect to a target component T if for every PPT adversary $\mathcal{A}$,*

$$\left|\Pr[\mathcal{A}(T^{\mathsf{anam}}, \mathsf{aux}) = 1] - \Pr[\mathcal{A}(T^{\mathsf{std}}, \mathsf{aux}) = 1]\right| \leq \mathsf{negl}(\lambda),$$

where T^{anam} denotes the target component generated in anamorphic mode (embedding a covert message), T^{std} denotes the same component generated in standard mode, and aux *captures any additional information available to $\mathcal{A}$ under the relevant key-disclosure pattern. The instantiation of T and* aux *is specified per construction. For DH-based constructions, T consists of DH public keys. For the MAC-based construction, T consists of authentication tags.*

Anamorphic security yields IND-CPA-style privacy for the covert content. If it is not possible to detect the presence of a covert message, it is also not possible to decide which of two candidate messages was embedded. Moreover, we present transcript-binding upgrades that attain a stronger, CCA-style guarantee against active adversaries. The definition here is an overview. Construction-specific security games and full proofs are provided in Appendices B and C.

4 Covert Channel in DH Ratchet

To embed covert messages into the DH public keys used in the DH ratchet, we construct an anamorphic channel with a hidden double key. For now, we assume that all keys other than this double key may be leaked. Later in this section, we explain how this assumption is instantiated in different scenarios.

4.1 Base Construction

Recall that in the DH ratchet, fresh DH key pairs are generated regularly. We show how to create *fake* DH key pairs that embed covert messages while remaining indistinguishable from honest ones. We present two algorithms: one generates the fake DH key pairs and the other extracts the hidden messages.

Here, we provide an abstract description over an additive prime-order group $\mathbb{G}$ with generator G. A standard DH key generation algorithm samples $\mathsf{sk} \leftarrow\!\$\, \mathbb{Z}_p$ and outputs $\mathsf{pk} = \mathsf{sk} \cdot G \in \mathbb{G}$.

Algorithm 1 generates $\ell + 1$ key pairs that embed an anamorphic message vector $\overrightarrow{\mathsf{amsg}} \in \mathcal{M}^\ell$ using the double key $\mathsf{dk} \in \{0,1\}^\lambda$, and Algorithm 2 recovers $\overrightarrow{\mathsf{amsg}}$ from PK consisting of $\ell + 1$ public keys pk_i. Here, the pseudorandom function PRF with key space $\{0,1\}^\lambda$ maps elements of $\mathbb{G}$ into $\mathbb{Z}_p$.

Algorithm 1 $\mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}})$

parse $\overrightarrow{\mathsf{amsg}} = \{\mathsf{amsg}_i\}_{i=1}^{\ell}$
$(\mathsf{sk}_1, \mathsf{pk}_1) \leftarrow \mathsf{Keygen}(\lambda)$
for $i \in [2, \ell+1]$ **do**
  $\mathsf{sk}_i = \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) + \mathsf{amsg}_{i-1} \bmod p$
  $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$
end for
return $\{(\mathsf{sk}_i, \mathsf{pk}_i)\}_{i=1}^{\ell+1}$

Algorithm 2 $\mathsf{aDec}(\mathsf{dk}, \mathsf{PK})$

parse $\mathsf{PK} = (\mathsf{pk}_i)_{i\in[1,\ell+1]}$
for $i \in [2, \ell+1]$ **do**
  $\Delta_i = \mathsf{pk}_i - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G$
  $\mathsf{amsg}_{i-1} = \mathsf{dlog}_G(\Delta_i)$
end for
return $\{\mathsf{amsg}_i\}_{i=1}^{\ell}$

Since recovering anamorphic messages requires solving a discrete logarithm over the admissible message space, $\mathcal{M}$ must be chosen significantly smaller than the group order p. We interpret dlog_G in Algorithm 2 as a bounded search over $\mathcal{M}$: it returns amsg if the residual point equals $\mathsf{amsg} \cdot G$ for some $\mathsf{amsg} \in \mathcal{M}$, and returns $\perp$ otherwise. Under this condition, the construction is robust: decrypting an anamorphic ciphertext under the correct double key yields the embedded message, while an invalid residual point is rejected except with probability at most $|\mathcal{M}|/p$. For Curve25519 and a 32-bit covert payload, this false-accept probability is at most $2^{32}/2^{252} = 2^{-220}$. As we show later, 32-bit payloads are practical with precomputation.

To integrate this into the full Double Ratchet protocol, Alice and Bob perform the following steps to transmit ℓ anamorphic messages within the DH ratchet using $\ell + 1$ DH keys. Without loss of generality, we present our construction in the setting where Alice sends covert messages to Bob, and assume for now that Alice and Bob share the same private double key dk.

1. Alice runs a fake key generation algorithm $\mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}})$ as described in Algorithm 1 with her private double key dk.
2. Alice sends $\ell + 1$ Signal messages to Bob. Before sending each message, she waits for a successful delivery notification[5].
3. Bob collects the last $\ell + 1$ ratchet public keys $\mathsf{PK} = \{\mathsf{pk}_1, \cdots, \mathsf{pk}_{\ell+1}\}$ sent by Alice and runs $\mathsf{aDec}(\mathsf{dk}, \mathsf{PK})$ to retrieve $\overrightarrow{\mathsf{amsg}}$ described in Algorithm 2.

We now argue correctness and security of the construction. Intuitively, correctness requires that, given the correct double key dk, the anamorphic messages should be recovered from the fake DH public keys produced by our algorithms. Anamorphic security requires that detecting the covert channel be infeasible: an adversary should not be able to distinguish honestly generated fresh key pairs from key pairs that embed anamorphic messages. The formal definitions and proofs are provided in Appendix B.

Proof sketch. Correctness follows because Bob recomputes the same $\mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1})$ value as Alice and subtracts it from pk_i, leaving $\mathsf{amsg}_{i-1} \cdot G$ for discrete-log recovery over the chosen anamorphic message space. For security, replace each $\mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1})$ output by an independent uniform value in $\mathbb{Z}_p$, which changes the adversary's view only by PRF security. Then each secret key $\mathsf{sk}_i = z_i + \mathsf{amsg}_{i-1} \bmod p$ is uniform whenever z_i is uniform, so the corresponding public key $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$ has the honest DH key distribution and hides amsg_{i-1}. The remaining loss comes from possible repeated inputs in the lazy-sampling argument and is bounded by roughly $(\ell Q)^2/p$ over Q oracle queries, which is negligible.

[5] This is necessary to avoid raising suspicion from the server. If Alice sends all $\ell+1$ messages at once, it may appear suspicious that she is not using the same sending chain key for some of them. When Bob sends a Signal message, including a successful delivery notification, he also includes a new DH ratchet public key.

4.2 Stronger Security Guarantee: IND-CCA2

We now strengthen our security guarantee. Analogous to IND-CCA2 security for an encryption scheme, we require that security holds even against an adversary that obtains ciphertexts generated in anamorphic mode and actively interacts with the protocol. In particular, such an adversary must neither be able to distinguish anamorphic executions from honest ones nor learn any information about the embedded anamorphic messages.

This CCA-style variant is block-based. Bob can run the anamorphic decoding and the transcript check only after receiving all DH public keys in the block: the $\ell + 1$ keys carrying ℓ covert messages and the additional integrity-check key. Until then, ordinary Signal processing can proceed, but CCA-protected covert extraction is deferred.

Intuitively, the updated fake key generation algorithm binds the covert encoding to the current transcript and adds an extra key pair used as an integrity check. The decryption algorithm is modified accordingly to verify this binding before extracting the anamorphic message. This yields active security at the cost of only one additional key pair.

We update the fake key generation algorithm (Algorithm 1) and the anamorphic message decryption algorithm (Algorithm 2) to Algorithm 3 and Algorithm 4, respectively. We add an auxiliary PRF^* (same key space) that maps the transcript to the secret key, binding keys to the transcript.

Algorithm 3 $\mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}})$

parse $\overrightarrow{\mathsf{amsg}} = \{\mathsf{amsg}_i\}_{i=1}^{\ell}$
$(\mathsf{sk}_1, \mathsf{pk}_1) \leftarrow \mathsf{Keygen}(\lambda)$
for $i \in [2, \ell + 1]$ **do**
  $\mathsf{sk}_i = \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) + \mathsf{amsg}_{i-1} \bmod p$
  $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$
end for
$T = \mathsf{pk}_1 \| \cdots \| \mathsf{pk}_{\ell+1}$
$\mathsf{sk}_{\ell+2} = \mathsf{PRF}^*(\mathsf{dk}, T)$
$\mathsf{pk}_{\ell+2} = \mathsf{sk}_{\ell+2} \cdot G$
return $\{(\mathsf{sk}_i, \mathsf{pk}_i)\}_{i=1}^{\ell+2}$

Algorithm 4 $\mathsf{aDec}(\mathsf{dk}, \mathsf{PK})$

parse $\mathsf{PK} = (\mathsf{pk}_i)_{i \in [1,\ell+2]}$
$T = \mathsf{pk}_1 \| \cdots \| \mathsf{pk}_{\ell+1}$
if $\mathsf{pk}_{\ell+2} \neq \mathsf{PRF}^*(\mathsf{dk}, T) \cdot G$ **then**
  return $\perp$
end if
for $i \in [2, \ell + 1]$ **do**
  $\Delta_i = \mathsf{pk}_i - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G$
  $\mathsf{amsg}_{i-1} = \mathsf{dlog}_G(\Delta_i)$
end for
return $\overrightarrow{\mathsf{amsg}}$

Security claim. Assuming PRF and PRF^* are secure PRFs, the transcript-binding DH variant achieves the CCA-style security goal from Section 3. The first $\ell + 1$ public keys are distributed as in the base construction, so they remain indistinguishable from honest DH public keys. The extra public key binds the block to $T = \mathsf{pk}_1 \| \cdots \| \mathsf{pk}_{\ell+1}$ through $\mathsf{PRF}^*(\mathsf{dk}, T)$. Therefore, any fresh block that an active adversary modifies, injects, or replays is accepted only if its final public key equals $\mathsf{PRF}^*(\mathsf{dk}, T) \cdot G$, where T is the transcript derived from the received block. Without dk, producing such a value for a new transcript is a PRF forgery, which succeeds only with negligible probability. The formal game-based definition and proof are given in Appendix B.

4.3 Instantiation of Double Key dk

The DH-based construction assumes Alice and Bob hold a double key dk unknown to the adversary. We now describe two instantiations corresponding to different key-establishment settings.

Symmetric variant (Sym-DH). Both parties share a pre-shared secret dk established out-of-band. Alice uses dk to generate fake DH keys encoding covert messages, and Bob uses dk to recover them. The adversary, who observes the full Double Ratchet state but not dk, cannot detect or decode the covert channel.

Asymmetric variant (Asym-DH). The receiver Bob holds a double key pair $(\mathsf{dsk}_B, \mathsf{dpk}_B)$ and publishes the public component dpk_B. Alice also holds a double key pair and publishes dpk_A. Alice embeds covert messages using dpk_B, and Bob recovers them using dsk_B, as both parties can compute a common "double key" via $\mathsf{dk} = \mathsf{DH}(\mathsf{dsk}_B, \mathsf{dpk}_A) = \mathsf{DH}(\mathsf{dsk}_A, \mathsf{dpk}_B)$. This removes the need for prior coordination and suits settings where publishing a long-term public key is feasible, such as a journalist or organization posting a double public key on their website.

4.4 Optimization

As discussed before, naively recovering anamorphic messages requires an exhaustive search over the anamorphic message space, so $\mathcal{M}$ must be small. We now present optimized variants of Algorithms 1

and 2 that use a precomputed lookup table, as in our implementation. Intuitively, the optimized algorithms replace exhaustive search with a single group operation and a table lookup. We successfully tested this optimization for values up to 32 bits. It can be extended further, since the main computational effort lies in precomputing the lookup table only once. The same lookup-table technique extends to the construction with a stronger security guarantee in a similar way. The following is a concrete instantiation over Curve25519.

Table 2: Notation used in Algorithms 5 and 6

Symbol	Description
G_E	Base point of the Edwards curve E.
P_E and Q_E	Points on the Edwards curve E.
$\mathsf{f}_{\mathrm{E}\to\mathrm{M}}$	Map from Edwards curve to Montgomery representation. Note that this function provides the extra bit of information sign needed for inversion, since the Montgomery model does not retain sign information.
$\mathsf{f}_{\mathrm{M}\to\mathrm{E}}$	Map from Montgomery to Edwards representation, using the supplied choice of sign for the Edwards point.
$\mathsf{lsb}(\mathsf{pk}_i)$	Least significant bit of pk_i.
$\mathsf{lookup}(\mathsf{Q}_E)$	Lookup table mapping an Edwards curve point to a message.

In Algorithm 5, each fake secret key is chosen as a PRF-derived scalar (keyed by dk and the previous public key) plus the anamorphic message. The corresponding public key is then selected so that its least significant bit determines the correct sign to invert back to the unique Edwards point. This ensures that the public keys are distributed as in the honest scheme, while encoding a small message-dependent offset in a reversible way. Algorithm 6 uses dk to recompute and subtract the PRF-derived scalar from the recovered Edwards point, yielding a point of the form $\mathsf{amsg} \cdot G_E$. Since we precompute a lookup table mapping all candidate points $\mathsf{amsg} \cdot G_E$ to their corresponding messages, decryption of each anamorphic block requires only one scalar multiplication and a table lookup, without iterating over all possible messages. This makes, for example, a 32-bit anamorphic message space practical in our implementation. The correctness and security guarantees of the original construction remain unchanged.

Algorithm 5 $\mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}}, \ell)$

$(\mathsf{sk}_1, \mathsf{pk}_1) = \mathsf{Keygen}(\lambda)$
for $i \in [2, \ell+1]$ **do**
  $\mathsf{ctr} = 0$
  repeat
    $x = \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}, \mathsf{ctr})$
    $\mathsf{sk}_i = x + \mathsf{amsg}_{i-1} \bmod p$
    $\mathsf{P}_E = \mathsf{sk}_i \cdot G_E$
    $(\mathsf{pk}_i, \mathsf{sign}) = \mathsf{f}_{\mathrm{E}\to\mathrm{M}}(\mathsf{P}_E)$
    $\mathsf{ctr} = \mathsf{ctr} + 1$
  until $\mathsf{lsb}(\mathsf{pk}_i) = \mathsf{sign}$
end for
return $\{(\mathsf{sk}_i, \mathsf{pk}_i)\}_{i=1}^{\ell+1}$

Algorithm 6 $\mathsf{aDec}(\mathsf{dk}, \mathsf{PK})$

$\mathsf{PK} = (\mathsf{pk}_i)_{i\in[1,\ell+1]}$
for $i \in [2, \ell+1]$ **do**
  $\mathsf{ctr} = 0$
  repeat
    $x = \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}, \mathsf{ctr})$
    $\mathsf{sign} = \mathsf{lsb}(\mathsf{pk}_i)$
    $\mathsf{P}_E = \mathsf{f}_{\mathrm{M}\to\mathrm{E}}(\mathsf{pk}_i, \mathsf{sign})$
    $\mathsf{Q}_E = \mathsf{P}_E - x \cdot G_E$
    $\mathsf{amsg}_{i-1} = \mathsf{lookup}(\mathsf{Q}_E)$
    $\mathsf{ctr} = \mathsf{ctr} + 1$
  until $\mathsf{amsg}_{i-1} \neq \perp$
end for
return $\{\mathsf{amsg}_i\}_{i=1}^{\ell}$

Precomputation technique. The lookup table is precomputed by calculating $\mathsf{Q}_E = \mathsf{amsg} \cdot G_E$ for each amsg in the anamorphic message space and storing the mapping $\mathsf{Q}_E \mapsto \mathsf{amsg}$. This speeds up repeated point-to-index queries by performing the expensive elliptic-curve computations in advance. The deterministic lookup table is built over a fixed range of candidate values (amsg), storing for each value the corresponding curve point in compressed form Q_E together with its index. The table is then sorted by the compressed point encoding, so that a query only requires compressing the target point and performing a binary search to recover the matching index. This replaces repeated full-range curve computations with a one-time preprocessing step and fast file-based lookups, making the method both

efficient and easy to reproduce across systems. The robustness bound above follows from the lookup table covering only the admissible message space rather than the full group.

5 Covert Channel in Symmetric Ratchet

To embed hidden messages into the symmetric ratchet, we create an anamorphic channel based on the MAC key that remains unknown to the server. Only the intended receiver, who holds this key, can recover the hidden messages as part of integrity verification. This construction, *Sym-MAC*, targets scenarios where *either* the root key *or* the private DH ratchet key is leaked to the server, but not both, so that the resulting chain, message, and MAC keys are never fully exposed. From the adversary's view, the MACs carrying hidden messages are indistinguishable from ordinary MACs. No pre-shared key or public double key is required, as the anamorphic messages are embedded using MAC keys derived within the Double Ratchet.

5.1 Base Construction

We first describe two algorithms that specify how to generate *fake* MACs and how to extract hidden messages from them.

Algorithm 7 embeds an anamorphic message $\mathsf{amsg} \in \mathcal{M}$ into a MAC tag under key $\mathsf{mac}_{\mathsf{key}}$ by XORing amsg with an honestly computed MAC. Algorithm 8 recovers the anamorphic message $\mathsf{amsg} \in \mathcal{M}$ by recomputing the honest MAC tag under $\mathsf{mac}_{\mathsf{key}}$ and extracting amsg via XOR.

Algorithm 7 $\mathsf{fakeMAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad}, \mathsf{amsg})$

$t = \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad})$
$\mathsf{mac} = t \oplus \mathsf{amsg}$
return mac

Algorithm 8 $\mathsf{aDec}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad}, \mathsf{mac})$

$t = \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad})$
if $\mathsf{mac} = t$ **then**
    skip ▷ Not anamorphic
else
    $\mathsf{amsg} = t \oplus \mathsf{mac}$
    return amsg
end if

This construction eliminates the small message-space constraint required by the DH-based channel. For the base construction, the anamorphic message space is $\mathcal{M} = \{0,1\}^\lambda \setminus \{0^\lambda\}$, where 0^λ is reserved to denote no covert payload because it yields the honest MAC tag. In our implementation, $\lambda = 64$, matching the MAC size specified in the Signal Double Ratchet documentation. This restriction applies only to the base construction. The active variant below can encode an all-zero payload block because its final binding tag identifies the block as anamorphic.

MAC verification and rejection. A receiver that runs the unmodified Signal verification procedure would reject an anamorphic tag, because $\mathsf{mac} \neq \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad})$ whenever $\mathsf{amsg} \neq 0$. In our construction, the intended receiver runs aDec with the same $\mathsf{mac}_{\mathsf{key}}$ before treating the mismatch as an ordinary authentication failure: equality means the tag is ordinary, while a mismatch is interpreted as an anamorphic payload. This behavior is invisible to the adversary in the partial-disclosure model, because $\mathsf{mac}_{\mathsf{key}}$ is not disclosed. If $\mathsf{mac}_{\mathsf{key}}$ is disclosed, the adversary can recompute the honest tag and detect the modification, so Sym-MAC does not apply. In the active variant below, the final tag authenticates the block and causes modified non-anamorphic tags to be rejected.

To integrate this into the full Double Ratchet protocol, Alice and Bob perform the following steps to transmit one anamorphic message amsg along with one normal Signal message m.

1. Alice encrypts a normal Signal message m into $\mathsf{ct} = \mathsf{AES}_{\mathsf{CBC}}(\mathsf{aes}_{\mathsf{key}}, \mathsf{IV}, m)$ and computes a tag as described in Algorithm 7 instead of the original mac.
2. Bob detects and recovers the anamorphic message amsg by running aDec as shown in Algorithm 8.

We now argue correctness and security of the construction. Correctness requires that, given the MAC key $\mathsf{mac}_{\mathsf{key}}$, any valid hidden message can be recovered. This follows directly from the invertibility of the XOR operation.

Anamorphic security requires that MACs carrying hidden messages be indistinguishable from honestly generated MACs. This holds as long as $\mathsf{mac}_{\mathsf{key}}$ remains secret and the underlying MAC produces pseudorandom tags, since XOR with a fixed message preserves indistinguishability from

random. As illustrated in Figure 1, secrecy of $\mathsf{mac}_{\mathsf{key}}$ implies secrecy of the corresponding message key and, in turn, of the chain key. Because deriving the chain key requires both the DH output and the root key, the construction remains secure if the adversary compromises either the private ratchet keys or the root key, as long as it does not obtain both simultaneously. The formal definition and proof are provided in Appendix C.

Proof sketch. Correctness follows directly from XOR. Bob recomputes the honest tag $t = \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}}, \mathsf{ct}, \mathsf{ad})$ and obtains $\mathsf{amsg} = \mathsf{mac} \oplus t$. For security, under an undisclosed $\mathsf{mac}_{\mathsf{key}}$, the honest tag t is pseudorandom to the adversary. The anamorphic tag $t \oplus \mathsf{amsg}$ is therefore also pseudorandom for any fixed covert message amsg, so it is indistinguishable from an honest MAC tag. In the Double Ratchet instantiation, the partial-disclosure condition preserves this argument because the adversary cannot derive $\mathsf{mac}_{\mathsf{key}}$ unless it obtains both the root key and the DH secret needed by the KDF chain.

5.2 Stronger Security Guarantee: IND-CCA2

As for the DH-based construction, we require that security holds even against an adversary that obtains ciphertexts generated in anamorphic mode and actively interacts with the protocol.

This CCA-style variant is block-based. Bob can run the anamorphic decoding and the transcript check only after receiving all $\ell + 1$ MAC tags in the block. Until then, ordinary Signal message processing can proceed, but CCA-protected covert extraction is deferred.

Intuitively, Alice embeds ℓ anamorphic messages, each of length λ, within $\ell + 1$ normal Signal messages. The $\ell+1$ MAC tags are generated so that the first ℓ tags are used in the symmetric ratchet to carry the corresponding covert payloads, and the final tag authenticates the preceding transcript to provide an integrity check.

We update the fake MAC algorithm (Algorithm 7) and the anamorphic message decryption algorithm (Algorithm 8) to Algorithm 9 and Algorithm 10, respectively. We add an auxiliary PRF with the same key space as the MAC key that maps the transcript to the tag.

Algorithm 9 $\mathsf{fakeMAC}(\overrightarrow{\mathsf{mac}_{\mathsf{key}}}, \overrightarrow{\mathsf{ct}}, \overrightarrow{\mathsf{ad}}, \overrightarrow{\mathsf{amsg}})$

for $i \in [1, \ell]$ **do**
  $t_i = \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}_i}, \mathsf{ct}_i, \mathsf{ad}_i)$
  $\mathsf{mac}_i = t_i \oplus \mathsf{amsg}_i$
end for
$T = \mathsf{mac}_1 \| \ldots \| \mathsf{mac}_\ell$
$\mathsf{mac}_{\ell+1} = \mathsf{PRF}(\mathsf{mac}_{\mathsf{key}_{\ell+1}}, T)$
return $\overrightarrow{\mathsf{mac}}$

Algorithm 10 $\mathsf{aDec}(\overrightarrow{\mathsf{mac}_{\mathsf{key}}}, \overrightarrow{\mathsf{ct}}, \overrightarrow{\mathsf{ad}}, \overrightarrow{\mathsf{mac}})$

for $i \in [1, \ell + 1]$ **do**
  $t_i = \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}_i}, \mathsf{ct}_i, \mathsf{ad}_i)$
end for
if $\mathsf{mac}_i = t_i$ for $i \in [1, \ell + 1]$ **then**
  skip ▷ Not anamorphic
else
  for $i \in [1, \ell]$ **do**
    $\mathsf{amsg}_i = \mathsf{mac}_i \oplus t_i$
  end for
  $T = \mathsf{mac}_1 \| \ldots \| \mathsf{mac}_\ell$
  $\sigma = \mathsf{PRF}(\mathsf{mac}_{\mathsf{key}_{\ell+1}}, T)$
  if $\sigma = \mathsf{mac}_{\ell+1}$ **then**
    return $\overrightarrow{\mathsf{amsg}}$
  else
    return $\perp$
  end if
end if

Alice and Bob perform the following steps to transmit ℓ anamorphic messages $\overrightarrow{\mathsf{amsg}}$ within $\ell + 1$ normal Signal messages $\overrightarrow{m}$.

1. Alice encrypts $\ell+1$ normal Signal messages $\overrightarrow{m}$ into $\overrightarrow{\mathsf{ct}} = \mathsf{AES}_{\mathsf{CBC}}(\overrightarrow{\mathsf{aes}_{\mathsf{key}}}, \overrightarrow{\mathsf{IV}}, \overrightarrow{m})$, where $\overrightarrow{\mathsf{aes}_{\mathsf{key}}}$ and $\overrightarrow{\mathsf{IV}}$ are generated honestly along the KDF chain.
2. Alice then computes the tags as described in Algorithm 9 and uses them as MACs corresponding to the messages instead of the original $\overrightarrow{\mathsf{mac}}$.
3. Bob detects and recovers the anamorphic messages $\overrightarrow{\mathsf{amsg}}$ by running aDec as shown in Algorithm 10.

Security claim. Assuming the MAC and PRF outputs are pseudorandom under an undisclosed $\mathsf{mac}_{\mathsf{key}}$, the transcript-binding MAC variant achieves the CCA-style security goal from Section 3. The

first ℓ tags remain indistinguishable from honest MAC tags by the base construction. The final tag is computed as $\mathsf{PRF}(\mathsf{mac}_{\mathsf{key}_{\ell+1}}, T)$ on the received transcript value. Thus, an active adversary who modifies, injects, or replays a block must produce the correct final tag for the resulting transcript before the block is accepted. Since $\mathsf{mac}_{\mathsf{key}_{\ell+1}}$ is hidden under the partial-disclosure condition, this reduces to forging a fresh PRF output and succeeds only with negligible probability. The formal game-based definition and proof are given in Appendix C.

6 Performance

We implement prototypes of our anamorphic messaging constructions [GLL+26] in the official Signal application [Sig25c] using Rust, Java, and Kotlin. We empirically evaluate their impact across different covert bit-rate settings, both in the official Signal-Android client [Sig25b] and in libSignal [Sig25a], the platform-agnostic library used by official Signal clients and mainly written in Rust.

We run our experiments on an Android device running Android 14 with a Qualcomm Snapdragon 888 processor and 8 GB of RAM, and on Android Studio emulators hosted on a 2023 MacBook Pro with an M2 Pro processor and 16 GB of RAM.

For the optimized DH-based construction, we compute the lookup table using an Intel Xeon Gold 6248R 3 GHz with 96 cores and 800 GB of RAM. The 16-bit lookup table is generated in a few seconds and takes 2.3 MB of space, while the 32-bit lookup table is generated in around 50 minutes and takes 144 GB of space. Both time and space complexity are exponential. However, the lookup table needs to be computed only once and can be exposed through a remote query service if local storage is an issue. Our optimized implementation supports remote lookup-table queries, allowing clients to avoid local table storage.

Table 3: Performance in milliseconds on a real Android device (top) and Android emulator (bottom).

Construction	**Anamorphic length**	**Signal-Android**				**libSignal**		
		Decryption		**Round-trip**		**Encryption**	**Decryption**	
		Avg	Worst	Avg	Worst	Avg	Avg	Worst
Android real device								
Original	None	1.56	—	20.5	—	1.53	1.42	—
Base (A)Sym-DH	16-bit	822.8	1622.2	841.48	1673.28	1.71	811.34	1617.41
Sym-MAC	64-bit	1.57	—	20.6	—	1.53	1.40	—
Android emulator								
Original	None	1.50	—	120.10	—	1.11	1.11	—
Base (A)Sym-DH	16-bit	438.2	888.2	567.3	1010.2	1.81	440.3	883.2
Optimized (A)Sym-DH	16-bit	7.01	14.12	149.96	150.14	1.13	5.77	13.20
	32-bit	7.22	65.10	150.32	200.18	1.32	5.43	65.88
Sym-MAC	64-bit	1.51	—	120.15	—	1.11	1.12	—

Table 3 presents performance results for experiments conducted on a real Android device and an Android emulator. We report the average encryption and decryption times for random messages using each anamorphic messaging construction, alongside the original Signal protocol.

We also evaluate the worst-case decryption time, which occurs when a message does *not* contain an anamorphic payload. In addition, we measure the round-trip time on the recipient side, defined as the duration between receiving the message from the server and sending the corresponding delivery receipt back. This metric is important for assessing potential timing side channels, as the time taken to generate and send the receipt could reveal to the server whether the message was anamorphic.[6]

[6] One could avoid sending the "delivery" message at all. In such a case, the DH ratchet would not be updated, since the delivery message carries a fresh public key, and the sender would need to wait for a message reply from the receiver before sending a new anamorphic message in the Sym-DH and Asym-DH constructions. In addition, not sending a delivery message might appear suspicious to the server, as sending a delivery message is standard in Signal.

As decryption is the most computationally intensive part of our constructions and a key performance indicator, we evaluate decryption time in both Signal-Android and libSignal. To evaluate aDec in our constructions (Sym-DH, Asym-DH, and Sym-MAC), we implemented the base variants. We also implemented the optimized (A)Sym-DH variants on the Android emulator to show their performance advantage. Note that the PRFs in Sym-DH and Asym-DH are instantiated using HMAC-SHA256.

In general, message encryption times are not affected by the generation of anamorphic messages. Looking at libsignal on the real device, the original Signal protocol takes approximately 1.53 ms to encrypt a message, while all anamorphic constructions require nearly the same time to encrypt an anamorphic message, regardless of its size. Similar results can be observed on the Android emulator, as shown in Table 3.

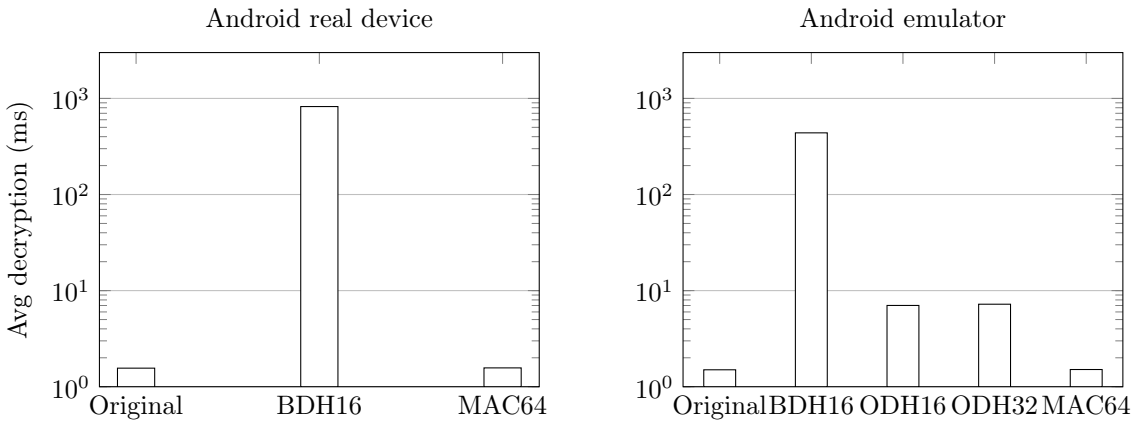

Fig. 4: Average decryption latency measured in the Signal-Android library on the Android emulator and real Android device, grouped by construction and anamorphic bit-length (log scale). BDH16, MAC64, ODH16, and ODH32 stand for Base (A)Sym-DH 16-bit, Sym-MAC 64-bit, Optimized (A)Sym-DH 16-bit, and Optimized (A)Sym-DH 32-bit, respectively.

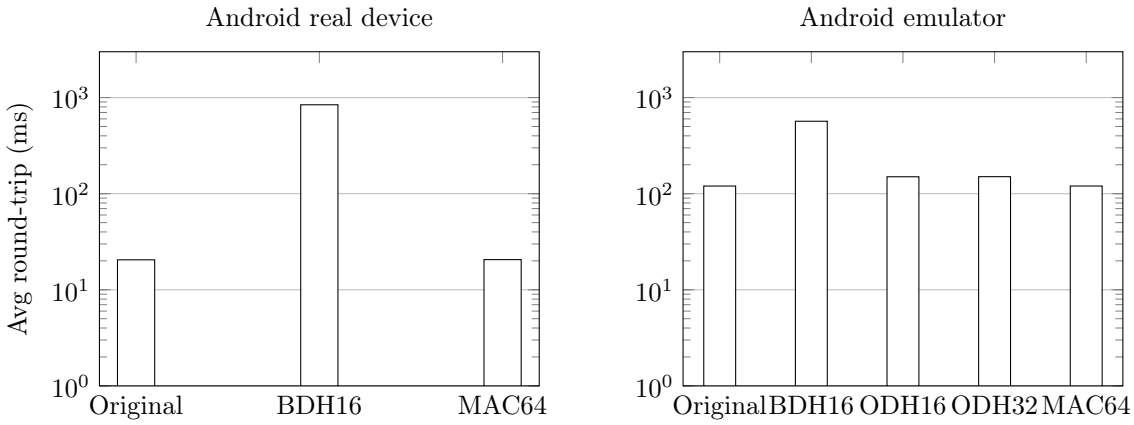

Fig. 5: Average round-trip latency measured in the Signal-Android library on the Android emulator and real Android device, grouped by construction and anamorphic bit-length (log scale).

Decryption time depends on the construction. This is expected, as decryption may involve iterative operations, particularly the discrete-logarithm computation required to recover the anamorphic message in the Sym-DH and Asym-DH constructions. However, there is no significant average difference within the Optimized (A)Sym-DH construction, nor between decryption in the core libSignal library and in the Signal-Android client. In particular, the optimized (A)Sym-DH construction avoids

exhaustive on-the-fly discrete-logarithm search and has only 7x decryption overhead. The user-visible round-trip time in the optimized version stays close to baseline Signal.

The round-trip time, as measured in the Signal-Android code, varies considerably only in the base (A)Sym-DH construction, while it stays close to the original Signal implementation in both the Sym-MAC and optimized (A)Sym-DH constructions.

Therefore, the Sym-MAC construction is the most efficient according to all indicators, but it supports the weakest threat model. The optimized (A)Sym-DH construction supports the strongest threat model and performs within the same order of magnitude as the original Signal construction. The base (A)Sym-DH construction takes significantly longer than the original Signal protocol, as the receiver must search over up to 16 bits of anamorphic information using multiple scalar multiplications, as shown in Figures 4 and 5. This construction is therefore suitable only when a lookup table cannot be maintained. In this case, to mitigate side-channel risks, the receiver can send the delivery receipt after a fixed delay matching the average round-trip time in the original implementation, rather than waiting for successful decryption. However, the sender should transmit each message using an interval longer than the worst-case round-trip time, approximately 1600 ms on the real device and 1000 ms on the emulator.

7 Conclusions

In this work, we presented two constructions for anamorphic channels for the Signal Double Ratchet protocol: a DH-based channel that embeds payloads in DH key pairs and a MAC-based channel that embeds payloads in authentication tags of the symmetric ratchet. Our designs remain compatible with existing Signal-style deployments, require no server changes, and operate even when one or more Double Ratchet keys are disclosed, for example as a result of a surveillance mechanism. Our analysis and implementations show that the DH-based and MAC-based channels can efficiently carry at least 32 and 64 hidden bits per Signal message, respectively.

To our knowledge, this is the first practical instantiation of anamorphic encryption within a deployed E2EE protocol. The Signal Double Ratchet underpins communication for billions of users, many of whom, including journalists, activists, lawyers, and people in vulnerable situations, depend on its confidentiality guarantees in contexts where the consequences of exposure are severe.

Several authorities have pushed Signal and other E2EE messaging platforms to enable surveillance by breaking E2EE confidentiality to snoop on messages. Our results indicate that even if decryption keys are exposed, covert channels that preserve normal protocol behavior can still exist.

Forward secrecy for the covert channel is outside the scope of our constructions. If the DH-channel double key is later exposed, an adversary with the transcript can recompute and verify prior covert contents. A natural direction is to derive per-message double keys through a KDF chain analogous to the Double Ratchet. Anamorphic authenticated key exchange [WHL24] provides one possible starting point.

A key direction for future work concerns the potential deployment of Anamorphic-Resistant Encryption schemes. If ARE systems become efficient and widely used, it will be important to understand whether some form of anamorphism can still be maintained, and under what leakage or functionality constraints. In particular, studying ARE-style defenses in the context of the Signal protocol would clarify the minimal conditions under which covert channels remain possible.

Finally, our approach is worth investigating in protocols that reuse Double Ratchet–style building blocks, including recent post-quantum proposals such as Triple Ratchet [DJK+25] and Apple's PQ3 [App24]. Beyond messaging, similar embeddings may apply to other E2EE services, including file sharing and real-time audio and video, offering a clearer path to supporting E2EE even under key-disclosure pressure.

Acknowledgements

Research at the University of Luxembourg was supported by the project Security and Verifiability of Electronic Testing and Assessment Systems (SEVERITAS, `INTER_ANR_20_14926102_SEVERITAS`). This work was supported by funding from the topic Engineering Secure Systems of the Helmholtz Association (HGF) and by KASTEL Security Research Labs.

References

ABG+26. G. Avitabile, V. Botta, E. Giunta, M. Mielniczuk, and F. Migliaro. The malice of elfs: practical anamorphic-resistant encryption without random oracles. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 303–333. Springer, 2026.

ACD19. J. Alwen, S. Coretti, and Y. Dodis. The double ratchet: security notions, proofs, and modularization for the signal protocol. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 129–158. Springer, 2019.

ADJ+25. B. Auerbach, Y. Dodis, D. Jost, S. Katsumata, and R. Schmidt. How to compare bandwidth constrained two-party secure messaging protocols: A quest for a more efficient and secure post-quantum protocol. In *Proceedings of the 34th USENIX Security Symposium*, 2025.

App24. Apple Security Engineering and Architecture. iMessage with PQ3: The new state of the art in quantum-secure messaging at scale. `https://security.apple.com/blog/imessage-pq3/`, 2024. Apple Security Research, 21 February 2024. Accessed: 2026-06-01.

BFG+22. A. Bienstock, J. Fairoze, S. Garg, P. Mukherjee, and S. Raghuraman. A more complete analysis of the signal double ratchet algorithm. In *Annual International Cryptology Conference*, pages 784–813. Springer, 2022.

BGB04. N. Borisov, I. Goldberg, and E. Brewer. Off-the-record communication, or, why not to use pgp. In *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society*, pages 77–84, 2004.

BGH+24. F. Banfi, K. Gegier, M. Hirt, U. Maurer, and G. Rito. Anamorphic encryption, revisited. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 3–32. Springer, 2024.

BPRS26. S. Banerjee, T. Pal, A. Rupp, and D. Slamanig. Simple asymmetric anamorphic encryption and signature using multi-message extensions. In *Annual International Cryptology Conference*. Springer, 2026. To appear.

CCGM25. D. Carnemolla, D. Catalano, E. Giunta, and F. Migliaro. Anamorphic resistant encryption: the good, the bad and the ugly. In *Annual International Cryptology Conference*, pages 472–503. Springer, 2025.

CDNO97. R. Canetti, C. Dwork, M. Naor, and R. Ostrovsky. Deniable encryption. In *Annual International Cryptology Conference*, pages 90–104. Springer, 1997.

CGCD+20. K. Cohn-Gordon, C. Cremers, B. Dowling, L. Garratt, and D. Stebila. A formal security analysis of the signal messaging protocol. *Journal of Cryptology*, 33(4):1914–1983, 2020.

CGM24. D. Catalano, E. Giunta, and F. Migliaro. Anamorphic encryption: New constructions and homomorphic realizations. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 33–62. Springer, 2024.

CGM25. D. Catalano, E. Giunta, and F. Migliaro. Generic anamorphic encryption, revisited: New limitations and constructions. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 275–303. Springer, 2025.

CJSV22. R. Canetti, P. Jain, M. Swanberg, and M. Varia. Universally composable end-to-end secure messaging. In *Annual International Cryptology Conference*, pages 3–33. Springer, 2022.

Cro25. D. Croft. Signal threatens to withdraw from Australia as government pushes for backdoor data access. `https://www.cyberdaily.au/security/12437-signal-threatens-to-withdraw-from-australia-as-government-pushes-for-backdoor-data-access`, 2025. Cyber Daily, 30 July 2025. Accessed: 2026-06-01.

CRT24. D. Collins, D. Riepel, and S. A. O. Tran. On the tight security of the double ratchet. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*, pages 4747–4761, 2024.

CSS23. A. Chakraborti, D. Suciu, and R. Sion. Wink: deniable secure messaging. In *32nd USENIX security symposium (USENIX Security 23)*, pages 1271–1288, 2023.

DG25. Y. Dodis and E. Goldin. Anamorphic-resistant encryption; or why the encryption debate is still alive. In *Annual International Cryptology Conference*, pages 440–471. Springer, 2025.

Dig25. Digital Watch Observatory. Sweden considers law requiring encrypted messaging backdoors, Signal threatens to exit. `https://dig.watch/updates/sweden-considers-law-requiring-encrypted-messaging-backdoors-signal-threatens-to-exit`, 2025. 26 February 2025. Accessed: 2026-06-01.

DJK+25. Y. Dodis, D. Jost, S. Katsumata, T. Prest, and R. Schmidt. Triple ratchet: A bandwidth efficient hybrid-secure signal protocol. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 302–331. Springer, 2025.

DMSDT25. Y. Dodis, B. Magri, N. Stephens-Davidowitz, and Y. Tselekounis. Guarding the signal: secure messaging with reverse firewalls. In *Annual International Cryptology Conference*, pages 39–74. Springer, 2025.

Eur25. European Commission. ProtectEU: A European Internal Security Strategy. COM(2025) 148 final, CELEX:52025DC0148, 2025. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 1 April 2025.

GLL+26. R. Giustolisi, G. Lenzini, C. Lin, M. Rakeei, and A. Rupp. Crypto Wars in Secure Messaging: Covert Channels in Signal Despite Leaked Keys. Zenodo, https://doi.org/10.5281/zenodo.20425141, May 2026. Software artifact.

GR25. D. Gligoroski and M. Raikwar. An llm framework for cryptography over chat channels. In *Proceedings of the Workshop on Privacy in Large Language Models (LLM) and Natural Language Processing (NLP) 2025*, pages 32–43, 2025.

GRL25. R. Giustolisi, M. Rakeei, and G. Lenzini. Anamorphic voting: Ballot freedom against dishonest authorities. In *International Conference on Financial Cryptography and Data Security. FC 2025 International Workshops*, pages 266–280. Springer, 2025.

KBB17. N. Kobeissi, K. Bhargavan, and B. Blanchet. Automated verification for secure messaging protocols and their implementations: A symbolic and computational approach. In *2017 IEEE European symposium on security and privacy (EuroS&P)*, pages 435–450. IEEE, 2017.

KCKG24. M. Knodel, S. Celi, O. Kolkman, and G. Grover. Definition of end-to-end encryption. Cryptology ePrint Archive, Paper 2024/2085, 2024.

KJGR21. G. Kaptchuk, T. M. Jois, M. Green, and A. D. Rubin. Meteor: Cryptographically secure steganography for realistic distributions. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 1529–1548, 2021.

KPP+23. M. Kutylowski, G. Persiano, D. H. Phan, M. Yung, and M. Zawada. The self-anti-censorship nature of encryption: On the prevalence of anamorphic cryptography. *Proceedings on Privacy Enhancing Technologies*, 2023(4):170–183, 2023.

Kra10. H. Krawczyk. Cryptographic extraction and key derivation: The hkdf scheme. In *Annual Cryptology Conference*, pages 631–648. Springer, 2010.

KS24. E. Kret and R. Schmidt. The PQXDH Key Agreement Protocol. https://signal.org/docs/specifications/pqxdh/, 2024. Revision 3, 2023-05-24, last updated 2024-01-23.

Mar13. M. Marlinspike. Advanced cryptographic ratcheting. https://signal.org/blog/advanced-ratcheting/, 2013. Signal Blog, 26 November 2013. Accessed: 2026-06-01.

MKA+21. I. Martiny, G. Kaptchuk, A. J. Aviv, D. S. Roche, and E. Wustrow. Improving signal's sealed sender. In *NDSS*, 2021.

MP16. M. Marlinspike and T. Perrin. The X3DH Key Agreement Protocol. https://signal.org/docs/specifications/x3dh/, 2016. Revision 1, 2016-11-04.

PMS25. T. Perrin, M. Marlinspike, and R. Schmidt. The Double Ratchet Algorithm. https://signal.org/docs/specifications/doubleratchet/, 2025. Revision 4, 2025-11-04.

PPY22. G. Persiano, D. H. Phan, and M. Yung. Anamorphic encryption: private communication against a dictator. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 34–63. Springer, 2022.

PR18. B. Poettering and P. Rösler. Towards bidirectional ratcheted key exchange. In *Annual International Cryptology Conference*, pages 3–32. Springer, 2018.

Sch25. R. Schmidt. The ML-KEM Braid Protocol. https://signal.org/docs/specifications/mlkembraid/, 2025. Revision 1, 2025-02-21, last updated 2025-09-26.

Sig25a. Signal. libsignal. https://github.com/signalapp/libsignal, 2025. Github repository, Accessed: May 2025.

Sig25b. Signal. Signal-android. https://github.com/signalapp/Signal-Android, 2025. Github repository, Accessed: May 2025.

Sig25c. Signal. signalapp. https://github.com/signalapp, 2025. Github repository, Accessed: May 2025.

Sim84. G. J. Simmons. The prisoners' problem and the subliminal channel. In *Advances in Cryptology: Proceedings of Crypto 83*, pages 51–67. Springer, 1984.

SW14. A. Sahai and B. Waters. How to use indistinguishability obfuscation: deniable encryption, and more. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 475–484, 2014.

TR24. W. L. Teng and K. Rasmussen. Defending against attack on the cloned: In-band active man-in-the-middle detection for the signal protocol. *arXiv preprint arXiv:2410.16098*, 2024.

Val23. C. Vallance. Signal would "walk" from UK if Online Safety Bill undermined encryption. https://www.bbc.com/news/technology-64584001, 2023. BBC News, 24 February 2023. Accessed: 2026-06-01.

VGIK20. N. Vatandas, R. Gennaro, B. Ithurburn, and H. Krawczyk. On the cryptographic deniability of the signal protocol. In *International Conference on Applied Cryptography and Network Security*, pages 188–209. Springer, 2020.

WCHY23. Y. Wang, R. Chen, X. Huang, and M. Yung. Sender-anamorphic encryption reformulated: Achieving robust and generic constructions. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 135–167. Springer, 2023.

WHL24. W. Wang, S. Han, and S. Liu. Anamorphic authenticated key exchange: Double key distribution under surveillance. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 168–200. Springer, 2024.

XCW+16. B. Xu, P. Chang, C. L. Welker, N. N. Bazarova, and D. Cosley. Automatic archiving versus default deletion: what snapchat tells us about ephemerality in design. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pages 1662–1675, 2016.

A Additional Related Work

Secure messaging. The foundations of modern secure messaging go back to Off-the-Record Messaging (OTR) [BGB04], which introduced forward secrecy and deniability. Later, TextSecure [Mar13] extended these ideas with a structured cryptographic framework that evolved into the Signal protocol [PMS25], now considered the standard for private messaging. At the heart of Signal is the Double Ratchet algorithm, which has been formally studied in multiple works. Alwen et al. [ACD19] presented a modular analysis of its components; Poettering and Rösler [PR18] investigated its compromise-resilience; and Cohn-Gordon et al. [CGCD+20] provided symbolic and computational proofs of its E2EE guarantees. Bienstock et al. [BFG+22] proposed a more complete Universally Composable framework, introducing a minimal extension called the Triple Ratchet, which captures stronger adversarial models with no added cost.

More recently, Teng and Rasmussen [TR24] showed that key verification in Signal can be automated within a modified protocol, removing the need for out-of-band user checks such as safety-number comparison or QR scanning and strengthening resistance to man-in-the-middle attacks. Martiny et al. [MKA+21] analyzed the Sealed Sender mechanism in Signal and demonstrated how statistical disclosure attacks could compromise its anonymity guarantees.

In parallel, the Signal protocol has been extended to resist emerging threats such as quantum adversaries. PQXDH [KS24] replaces the X3DH handshake with a post-quantum alternative, while Dodis et al. [DJK+25] introduced a Triple Ratchet protocol that incorporates a more efficient PQ-ratchet and erasure coding, addressing bandwidth limitations in hybrid-secure settings. In a related development within quantum-secure Signal, the ML-KEM Braid protocol [Sch25] was proposed based on the ideas of Auerbach et al. [ADJ+25]. It is a Sparse Continuous Key Agreement (SCKA) built on ML-KEM that lets two parties derive a sequence of post-quantum shared secrets with forward secrecy and post-compromise security. The Classical Continuous Key Agreement in the Double Ratchet emits a fresh secret each round trip, whereas SCKA handles bandwidth constraints by sending large KEM data in pieces so that new secrets may arrive over multiple round trips, and the resulting keys can feed a Double Ratchet for secure messaging. Additionally, secure messaging protocols such as Signal's X3DH and Double Ratchet can be enhanced with subversion resilience through techniques such as *Reverse Firewalls*, which protect against attacks that modify cryptographic algorithms and help users detect tampering even with compromised implementations [DMSDT25].

The robustness of secure messaging systems, particularly Signal, lies in its ability to guarantee security and privacy even in adversarial environments. However, as the landscape of cryptographic security evolves, new threats such as deployment-level compromises emerge, challenging the reliability of E2EE. In light of these challenges, anamorphic encryption offers a promising approach to preserve security in the presence of surveillance that compromises E2EE.

Anamorphic encryption. Anamorphic encryption was first introduced by Persiano et al. [PPY22] as a method for enabling covert communication when a surveillance authority can decrypt messages, but a cooperating sender and receiver can still exchange hidden content embedded inside normal-looking ciphertexts. Since its introduction, several works have proposed stronger definitions [CGM24,CGM25]. Other works adapted the idea to current cryptographic primitives [KPP+23], [BGH+24], and [WCHY23].

One of the early developments was by Wang et al. [WCHY23], who introduced the notion of ℓ-sender-anamorphic encryption, formalizing secure constructions for pseudo-random public key encryption (PKE) schemes. Following this work, Banerjee et al. [BPRS26] introduced a framework for

public-key anamorphic μ-message extensions, embedding covert messages within multiple ciphertexts of CPA-secure PKE schemes without relying on randomness recovery.

Applications and resistance. Anamorphic-resistant encryption (ARE) studies the opposite goal. Dodis and Goldin [DG25] and Carnemolla et al. [CCGM25] propose encryption schemes that specifically block covert communication, opposing the core goal of anamorphic encryption. ARE schemes aim to make covert channels either slow or entirely impractical, thereby undermining anamorphic encryption's ability to enable hidden communication in the presence of surveillance that breaks E2EE. Avitabile et al. [ABG+26] present a practical Anamorphic-Resistant Encryption (ARE) scheme that addresses key limitations of previous constructions [DG25,CCGM25], notably by eliminating the reliance on random oracles and indistinguishability obfuscation. Their scheme, based on Extremely Lossy Functions (ELFs) and Unique Non-Interactive Zero-Knowledge proofs (UNIZKs), ensures that covert communication is limited to a logarithmic number of bits, effectively hindering the viability of anamorphic encryption for transmitting long anamorphic messages.

In recent years, research on anamorphism in application-level cryptographic protocols has gained considerable momentum. Giustolisi et al. [GRL25] adapted the ElGamal-based construction from Wang et al. [WCHY23] to electronic voting systems, enabling voters to embed hidden preferences within ballots that are indistinguishable from legitimate ones, even against malicious authorities.

More recently, Gligoroski et al. [GR25] explored the application of anamorphic encryption in conversational settings, embedding covert messages within text exchanged over public chat channels. Their framework introduced an LLM-based embedding function, enabling both symmetric and public key encryption schemes (e.g., password-authenticated encryption and elliptic curve DH key exchange with ephemeral keys) to operate over natural language outputs. This work opened new directions for covert communication through LLM-generated text.

Despite its theoretical foundations, anamorphic encryption has seen limited exploration in practical applications. Secure messaging systems like Signal are often targeted by adversaries or regulatory efforts seeking to undermine E2EE. Legislative proposals for lawful access to encrypted data [Eur25] highlight the need for secure messaging protocols that remain functional under E2EE compromise. Our work builds on the principles of anamorphic encryption, applying it to the Signal Double Ratchet protocol to enable covert communication even if the confidentiality of E2EE is intentionally broken.

B Security Proofs for DH-Based Constructions

B.1 Correctness and Anamorphic Security for the Base Construction in Section 4.1

Correctness. Each anamorphic message amsg_i is correctly recovered since, for all $i \in [2, \ell+1]$, it holds that

$$\begin{aligned}
\mathsf{amsg}_{i-1} &= \mathsf{dlog}_G\left(\mathsf{pk}_i - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G\right) \\
&= \mathsf{dlog}_G\left(\mathsf{sk}_i \cdot G - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G\right) \\
&= \mathsf{dlog}_G((\mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) + \mathsf{amsg}_{i-1}) \cdot G \\
&\qquad - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G) \\
&= \mathsf{dlog}_G(\mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G + \mathsf{amsg}_{i-1} \cdot G \\
&\qquad - \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) \cdot G) \\
&= \mathsf{dlog}_G\left(\mathsf{amsg}_{i-1} \cdot G\right) \\
&= \mathsf{amsg}_{i-1}.
\end{aligned}$$

Anamorphic Security. The construction achieves anamorphic security if for every probabilistic polynomial-time (PPT) adversary $\mathcal{D}$ it holds that

$$\left|\Pr[\mathsf{GReal}^{\mathcal{D}}_{\mathsf{Kgen}}(\lambda) = 1] - \Pr[\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{fKgen}}(\lambda) = 1]\right| \leq \mathsf{negl}(\lambda)$$

where the security games are defined as follows:

$\underline{\mathsf{GReal}^{\mathcal{D}}_{\mathsf{Kgen}}(\lambda):}$

1. $\mathsf{dk} \leftarrow \mathsf{aGen}(\lambda)$
2. **return** $\mathcal{D}^{\mathcal{O}_{\mathsf{Kgen}}(\cdot)}(\lambda)$ where $\mathcal{O}_{\mathsf{Kgen}}(\overrightarrow{\mathsf{amsg}}) = \{\mathsf{Keygen}(\lambda)\}_{i=1}^{\ell+1}$

$\underline{\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{fKgen}}(\lambda):}$

1. $\mathsf{dk} \leftarrow \mathsf{aGen}(\lambda)$
2. **return** $\mathcal{D}^{\mathcal{O}_{\mathsf{fKgen}}(\cdot)}(\lambda)$ where $\mathcal{O}_{\mathsf{fKgen}}(\overrightarrow{\mathsf{amsg}}) = \mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}}, \ell)$

In the games defined for this construction, we explicitly include the generation of the double key dk, which was omitted in the original construction. This step ensures that dk has the same distribution in both games. The value dk is consumed only by the anamorphic oracle and is not revealed to the distinguisher $\mathcal{D}$; the real oracle ignores it. Moreover, $\mathcal{D}$ does not require access to other keys, as they do not influence the generation of DH ratchet key pairs.

Theorem 2. *The construction presented in Section 4.1 satisfies anamorphic security if the* PRF *used in Algorithm 1 is a secure PRF, as introduced in Section 2.1.*

Proof. In the following sequence of hybrid games, let $\mathcal{D}$ be a PPT adversary and Q its total number of oracle queries.
Game 0**:** This is the anamorphic game $\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{fKgen}}(\lambda)$ in the security definition, where the oracle generates $(\mathsf{pk}_i, \mathsf{sk}_i)$ as described in Algorithm 1, $\mathsf{sk}_i = \mathsf{PRF}(\mathsf{dk}, \mathsf{pk}_{i-1}) + \mathsf{amsg}_{i-1}$, and $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$.
Game 1**:** In the second game, we replace all applications of $\mathsf{PRF}(\mathsf{dk}, \cdot)$ by the application of a uniformly chosen function $\mathcal{F} : \mathbb{G} \to \mathbb{Z}_p$.
Game 2**:** In the third game, we define $\mathcal{F}$ iteratively by lazy sampling: each time $\mathcal{F}$ is applied, we check whether a mapping for the given input a has already been established, i.e., whether $a \mapsto z$ is contained in some set $\mathcal{M}_{\mathcal{F}}$. If so, z is the output. Otherwise, a fresh $z \leftarrow \mathbb{Z}_p$ is sampled uniformly and the mapping $a \mapsto z$ is added to $\mathcal{M}_{\mathcal{F}}$.
Game $(3, j)$ **for** $j \in [0, \ell Q]$**:** In this series of games, we transition from Game 2 to Game 4 by modifying how each of the ℓQ applications of $\mathcal{F}$ is computed. We define Game $(3, 0)$ to be identical to Game 2. In Game $(3, j)$, the challenger handles the first j applications of $\mathcal{F}$ as follows: instead of checking whether a mapping for the corresponding input is already defined in $\mathcal{M}_{\mathcal{F}}$, it samples a uniform output and adds the corresponding mapping to $\mathcal{M}_{\mathcal{F}}$. Note that in Game $(3, \ell Q)$ all applications of $\mathcal{F}$ are replaced by uniformly sampling elements from $\mathbb{Z}_p$.
Game 4**:** This is the real game $\mathsf{GReal}^{\mathcal{D}}_{\mathsf{Kgen}}(\lambda)$ in the anamorphic security definition, where the oracle generates $(\mathsf{pk}_i, \mathsf{sk}_i)$ by choosing $\mathsf{sk}_i \leftarrow \mathbb{Z}_p$ uniformly at random and setting $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$.

Game 0 and Game 1 are computationally indistinguishable, as a distinguisher for them could be turned into one for the pseudorandomness of PRF. The games Game 1 and Game 2 are perfectly indistinguishable, as Game 2 uses an equivalent way to define $\mathcal{F}$. Game $(3, 0)$ coincides with Game 2 by definition. The difference between games Game $(3, j-1)$ and Game $(3, j)$ consists in how the output of the j-th application of $\mathcal{F}$ is computed. In Game $(3, j-1)$ we check whether the input pk_{j-1} occurred before in $\mathcal{M}_{\mathcal{F}}$, while we do not perform this check in Game $(3, j)$. Each PRF input pk_{i-1} is uniform in $\mathbb{G}$: pk_1 is honestly generated, and once $\mathcal{F}$ is freshened, $\mathsf{sk}_i = z_i + \mathsf{amsg}_{i-1}$ is uniform in $\mathbb{Z}_p$, so $\mathsf{pk}_i = \mathsf{sk}_i \cdot G$ is uniform. Hence the probability that the j-th input collides with an earlier one is at most $\frac{j-1}{p}$. Thus, in total, the probability that games Game $(3, 0)$ and Game $(3, \ell Q)$ differ is upper bounded by $\frac{(\ell Q)^2}{2p}$, which is negligible. Finally, note that Game $(3, \ell Q)$ and Game 4 are identical because $\mathsf{sk}_i = z_i + \mathsf{amsg}_{i-1}$ for $z_i \leftarrow \mathbb{Z}_p$ yields a uniformly random sk_i in Game $(3, \ell Q)$. □

B.2 IND-CCA2 Security for Upgraded Construction in Section 4.2

Definition 3 (IND-CCA **for the Covert Channel in the DH Ratchet).** *Let $\mathcal{A}$ be a PPT algorithm. The security experiment for indistinguishability under adaptive chosen ciphertext attack is* $\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}$ *shown in Figure 6. The advantage of $\mathcal{A}$ in the experiment is*

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A}) = \left| \Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A}) \to 1\right] - \frac{1}{2} \right|$$

$\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A})$

$\mathsf{dk} \leftarrow\!\$\ \{0,1\}^{\lambda}$
$S = \emptyset$
$b \leftarrow\!\$\ \{0,1\}$
$b' \leftarrow \mathcal{A}^{\mathcal{O}_{\mathsf{fKGen}_b}(\cdot,\cdot),\mathcal{O}_{\mathsf{aDec\text{-}DH}}(\cdot)}(1^{\lambda}, 1^{\ell})$
if $b' = b$ **then**
  return 1
else
  return 0

$\mathcal{O}_{\mathsf{fKGen}_b}(\overrightarrow{\mathsf{amsg}}_0, \overrightarrow{\mathsf{amsg}}_1)$

$\{(\mathsf{sk}_i, \mathsf{pk}_i)\}_{i=1}^{\ell+2} \leftarrow \mathsf{fakeKeygen}(\mathsf{dk}, \overrightarrow{\mathsf{amsg}}_b)$
$C \leftarrow (\mathsf{pk}_1, \ldots, \mathsf{pk}_{\ell+2})$
$S = S \cup \{C\}$
return $((\mathsf{sk}_1, \ldots, \mathsf{sk}_{\ell+2}), C)$

$\mathcal{O}_{\mathsf{aDec\text{-}DH}}(C^*)$

if $C^* \in S$
  return $\perp$
else
  parse $C^* = (\mathsf{pk}^*_1, \ldots, \mathsf{pk}^*_{\ell+2})$
  return $\mathsf{aDec}(\mathsf{dk}, C^*)$

Fig. 6: Security experiment for IND-CCA security of the covert channel in the DH ratchet

Theorem 4. *If the* PRF*s used in Appendix B are secure PRFs, then for every PPT adversary* $\mathcal{A}$*, it holds that* $\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A}) \leq \mathsf{negl}(\lambda)$.

Proof. Let $\Pr[\mathsf{Game}\ i]$ be the adversary's success probability in game $\mathsf{Game}\ i$, and let Q be the total number of oracle queries.
Game 0**:** This game is exactly the IND-CCA experiment and oracle. Hence,

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A}) = \left|\Pr[\mathsf{Game}\ 0] - \frac{1}{2}\right|$$

Game 1**:** In this game, we replace every use of PRF^* in the challenge oracle and in the decryption algorithm aDec with a random function $\mathcal{F} : \{0,1\}^* \mapsto \mathbb{Z}_p$.
Let $\mathcal{B}_1$ be a PPT adversary against PRF security. By PRF security, we have

$$|\Pr[\mathsf{Game}\ 0] - \Pr[\mathsf{Game}\ 1]| \leq \mathsf{Adv}^{\mathsf{prf}}(\mathcal{B}_1)$$

Game 2**:** In this game, we define $\mathcal{F}$ iteratively by lazy sampling: we maintain a table for $\mathcal{F}$, and on input T: if T is already in the table, return the stored scalar s, otherwise sample $s \leftarrow\!\$\ \mathbb{Z}_p$ uniformly, store and return it. Game 1 and Game 2 use equivalent methods to define $\mathcal{F}$, and thus

$$\Pr[\mathsf{Game}\ 1] = \Pr[\mathsf{Game}\ 2]$$

Game 3**:** We first define a Forge event: the aDec algorithm accepts a *fresh* ciphertext. In Game 2, challenge queries reveal pairs $(T_h, \mathcal{F}(T_h))$ through the leaked secret key $\mathsf{sk}_{\ell+2}$, but these values do not determine $\mathcal{F}(T^*)$ for a fresh transcript $T^* = (\mathsf{pk}^*_1 \| \cdots \| \mathsf{pk}^*_{\ell+1})$ outside the challenge set. For such a fresh transcript, the scalar $s = \mathcal{F}(T^*)$ is uniform and hidden. Acceptance requires guessing the final public key $s \cdot G$. In a group of prime order p, this succeeds with probability $\frac{1}{p}$ per query, and thus

$$\Pr[\mathsf{Forge}] \leq \frac{Q}{p}.$$

This game is identical to Game 2, except that the decryption oracle aborts on the first accepting fresh decryption query. Then,

$$|\Pr[\mathsf{Game}\ 2] - \Pr[\mathsf{Game}\ 3]| \leq \Pr[\mathsf{Forge}]$$

Game 4**:** In this game, we replace $\{(\mathsf{sk}_i, \mathsf{pk}_i)\}_{i=1}^{\ell+1}$ with fresh honestly generated DH key pairs.
Let $\mathcal{B}_2$ be a PPT adversary against the anamorphic security defined above. Then, we have

$$|\Pr[\mathsf{Game}\ 3] - \Pr[\mathsf{Game}\ 4]| \leq \mathsf{Adv}^{\mathsf{ana}}_{\mathsf{DH}}(\mathcal{B}_2).$$

Game 5**:** In the final game, the challenge blocks contain honestly generated DH key pairs and the decryption oracle rejects every fresh accepting query by the previous hop. The adversary's view no longer depends on the hidden bit b. Therefore, we have

$$\Pr[\text{Game } 5] = \frac{1}{2}.$$

Combining the hops, we derive

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{DH}}(\mathcal{A}) \leq \mathsf{Adv}^{\mathsf{prf}}(\mathcal{B}_1) + \mathsf{Adv}^{\mathsf{ana}}_{\mathsf{DH}}(\mathcal{B}_2) + \frac{Q}{p}$$

which is negligible. □

C Security Proofs for MAC-Based Construction

C.1 Anamorphic Security for Base Construction in Section 5.1

Anamorphic Security. The construction achieves anamorphic security if for every PPT adversary $\mathcal{D}$ it holds that

$$\left|\Pr[\mathsf{GReal}^{\mathcal{D}}_{\mathsf{MAC}}(\lambda) = 1] - \Pr[\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{aMAC}}(\lambda) = 1]\right| \leq \mathsf{negl}(\lambda)$$

where the security games are defined as follows.

In the security games, the distinguisher receives the public DH ratchet keys and either rk_0 or $\mathsf{dh}_{\mathsf{out}}$, according to the partial-disclosure model. The oracle is stateful: after one DH-ratchet KDF step, each oracle query advances the sending chain and therefore uses a fresh message key and MAC key. This models consecutive messages in the same symmetric ratchet chain.

$\underline{\mathsf{GReal}^{\mathcal{D}}_{\mathsf{MAC}}(\lambda):}$

1. $\mathsf{rk}_0 \leftarrow \{0,1\}^\lambda$
2. $(\mathsf{sk}_A, \mathsf{pk}_A) \leftarrow \mathsf{Keygen}(\lambda)$
3. $(\mathsf{sk}_B, \mathsf{pk}_B) \leftarrow \mathsf{Keygen}(\lambda)$
4. $\mathsf{dh}_{\mathsf{out}} \leftarrow \mathsf{DH}(\mathsf{sk}_A, \mathsf{pk}_B)$
5. $(\mathsf{rk}_1, \mathsf{ck}_0) \leftarrow \mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}})$
6. $j \leftarrow 0$
7. **return** $\mathcal{D}^{\mathcal{O}_{\mathsf{MAC}}(\cdot,\cdot,\cdot)}(\lambda, \mathsf{info}, \mathsf{pk}_A, \mathsf{pk}_B)$ where $\mathcal{O}_{\mathsf{MAC}}(\mathsf{amsg}, \mathsf{ad}, m)$:
   - $j \leftarrow j + 1$
   - $(\mathsf{ck}_j, \mathsf{mk}_j) \leftarrow \mathsf{KDF}_{\mathsf{ck}}(\mathsf{ck}_{j-1})$
   - $(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{mac}_{\mathsf{key}\,j}, \mathsf{IV}_j) \leftarrow \mathsf{KDF}_{\mathsf{mk}}(\mathsf{mk}_j)$
   - $\mathsf{ct}_j \leftarrow \mathsf{AES}_{\mathsf{CBC}}(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{IV}_j, m)$
   - **return** $\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{IV}_j, \mathsf{MAC}(\mathsf{mac}_{\mathsf{key}\,j}, \mathsf{ct}_j, \mathsf{ad})$

$\underline{\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{aMAC}}(\lambda):}$

1. $\mathsf{rk}_0 \leftarrow \{0,1\}^\lambda$
2. $(\mathsf{sk}_A, \mathsf{pk}_A) \leftarrow \mathsf{Keygen}(\lambda)$
3. $(\mathsf{sk}_B, \mathsf{pk}_B) \leftarrow \mathsf{Keygen}(\lambda)$
4. $\mathsf{dh}_{\mathsf{out}} \leftarrow \mathsf{DH}(\mathsf{sk}_A, \mathsf{pk}_B)$
5. $(\mathsf{rk}_1, \mathsf{ck}_0) \leftarrow \mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}})$
6. $j \leftarrow 0$
7. **return** $\mathcal{D}^{\mathcal{O}_{\mathsf{aMAC}}(\cdot,\cdot,\cdot)}(\lambda, \mathsf{info}, \mathsf{pk}_A, \mathsf{pk}_B)$ where $\mathcal{O}_{\mathsf{aMAC}}(\mathsf{amsg}, \mathsf{ad}, m)$:
   - $j \leftarrow j + 1$
   - $(\mathsf{ck}_j, \mathsf{mk}_j) \leftarrow \mathsf{KDF}_{\mathsf{ck}}(\mathsf{ck}_{j-1})$
   - $(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{mac}_{\mathsf{key}\,j}, \mathsf{IV}_j) \leftarrow \mathsf{KDF}_{\mathsf{mk}}(\mathsf{mk}_j)$
   - $\mathsf{ct}_j \leftarrow \mathsf{AES}_{\mathsf{CBC}}(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{IV}_j, m)$
   - **return** $\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{IV}_j, \mathsf{fakeMAC}(\mathsf{mac}_{\mathsf{key}\,j}, \mathsf{ct}_j, \mathsf{ad}, \mathsf{amsg})$

In both games, $\mathsf{info} = \mathsf{dh}_{\mathsf{out}}$ or rk_0

Theorem 5. *The construction Sym-MAC defined in Section 5.1 satisfies anamorphic security if the Decisional Diffie–Hellman problem is hard in $\mathbb{G}$, the* MAC *used in Algorithm 7 is modeled as a secure PRF,* HMAC *used in HKDF-Extract is secure both as a PRF under a hidden key and as a dual-PRF/computational extractor on a hidden DH output, and all remaining* KDF*s involved are secure KDFs, as introduced in Section 2.1.*

Proof. In the following sequence of hybrid games, let $\mathcal{D}$ be a PPT adversary and let Q be its total number of oracle queries. All random functions used in the hybrids are evaluated consistently, and the stateful oracle advances the chain once per query.

Game 0**:** This is the anamorphic game $\mathsf{GAnam}^{\mathcal{D}}_{\mathsf{aMAC}}(\lambda)$ in the security definition. Query j derives $(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{mac}_{\mathsf{key}\,j}, \mathsf{IV}_j)$ from the current chain key and returns $\mathsf{MAC}(\mathsf{mac}_{\mathsf{key}\,j}, \mathsf{ct}_j, \mathsf{ad}) \oplus \mathsf{amsg}$.

Game 1**:** In this game, the sequence $\{(\mathsf{aes}_{\mathsf{key}\,j}, \mathsf{mac}_{\mathsf{key}\,j}, \mathsf{IV}_j)\}_{j\in[1,Q]}$ generated by the KDF chain is replaced by Q independent uniform triples.

Game 2: In this game, for every query j, the value $\mathsf{MAC}(\mathsf{mac}_{\mathsf{key}_j}, \mathsf{ct}_j, \mathsf{ad})$ is replaced by a fresh uniform value $r_j \leftarrow \{0,1\}^\lambda$. The oracle returns $(\mathsf{aes}_{\mathsf{key}_j}, \mathsf{IV}_j, r_j \oplus \mathsf{amsg})$.
Game 3: In this game, each value $r_j \oplus \mathsf{amsg}$ is replaced by an independent uniform value $u_j \leftarrow \{0,1\}^\lambda$.

Game 0 and Game 1 are computationally indistinguishable by Lemma 6, which applies to the full stateful sequence of derived keys. In Game 1, each $\mathsf{mac}_{\mathsf{key}_j}$ is uniform and used for one MAC evaluation. A standard hybrid over the Q oracle queries shows that Game 1 and Game 2 are computationally indistinguishable by the PRF security of MAC. The transition from Game 2 to Game 3 is perfect because XOR with a fixed value is a bijection on $\{0,1\}^\lambda$ and r_j is uniform.

The same argument, without the final XOR, transforms the real game $\mathsf{GReal}^{\mathcal{D}}_{\mathsf{MAC}}(\lambda)$ into the same ideal game Game 3. Hence the real and anamorphic games are computationally indistinguishable. □

Lemma 6. *Let an adversary receive the public ratchet keys $\mathsf{pk}_A, \mathsf{pk}_B$ and either rk_0 or $\mathsf{dh}_{\mathsf{out}} = \mathsf{DH}(\mathsf{sk}_A, \mathsf{pk}_B)$, but not both. For any polynomial Q, let*

$$\mathcal{K}_Q = \{(\mathsf{aes}_{\mathsf{key}_j}, \mathsf{mac}_{\mathsf{key}_j}, \mathsf{IV}_j)\}_{j\in[1,Q]}.$$

If $\mathcal{K}_Q$ is derived by one root KDF step followed by Q symmetric-chain steps, then it is computationally indistinguishable from Q independent uniform triples.

Proof. We first consider the root KDF step

$$(\mathsf{rk}_1, \mathsf{ck}_0) \leftarrow \mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}}),$$

where HKDF-Extract is instantiated as $\mathsf{HMAC}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}})$ followed by HKDF-Expand.

If $\mathsf{dh}_{\mathsf{out}}$ is disclosed and rk_0 is hidden, then rk_0 is a uniform secret HMAC key. The extract output is pseudorandom by the PRF security of HMAC, and HKDF-Expand maps this pseudorandom value to pseudorandom $(\mathsf{rk}_1, \mathsf{ck}_0)$.

If rk_0 is disclosed and $\mathsf{dh}_{\mathsf{out}}$ is hidden, then the public values $\mathsf{pk}_A, \mathsf{pk}_B$ allow a DDH reduction to replace $\mathsf{dh}_{\mathsf{out}}$ with a uniform group element. With a known salt rk_0 and a uniform hidden input, the extract output $\mathsf{HMAC}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}})$ is pseudorandom by the assumed dual-PRF/computational-extractor property of HKDF-Extract. HKDF-Expand then yields pseudorandom $(\mathsf{rk}_1, \mathsf{ck}_0)$.

Once ck_0 is indistinguishable from uniform, a hybrid over the Q calls to $\mathsf{KDF}_{\mathsf{ck}}$ replaces each pair $(\mathsf{ck}_j, \mathsf{mk}_j)$ with independent uniform values. A second hybrid over the Q calls to $\mathsf{KDF}_{\mathsf{mk}}$ replaces each triple $(\mathsf{aes}_{\mathsf{key}_j}, \mathsf{mac}_{\mathsf{key}_j}, \mathsf{IV}_j)$ with an independent uniform triple. These hybrids rely on the security of the corresponding KDFs and preserve consistency of repeated inputs. Therefore, the full stateful sequence of derived message keys is indistinguishable from independent uniform triples. □

C.2 IND-CCA2 Security for Upgraded Construction in Section 5.2

Definition 7 (IND-CCA for the Covert Channel in the Symmetric Ratchet). *Let $\mathcal{A}$ be a PPT algorithm. The security experiment for indistinguishability under adaptive chosen ciphertext attack is $\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}$ shown in Figure 7. The advantage of $\mathcal{A}$ in the experiment is*

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A}) = \left|\Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A}) \to 1\right] - \frac{1}{2}\right|.$$

Theorem 8. *If the assumptions of Theorem 5 hold and the transcript-binding* PRF *in Algorithm 9 is secure, then for every PPT adversary $\mathcal{A}$, it holds that $\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A}) \le \mathsf{negl}(\lambda)$.*

Proof. Let $\Pr[\mathsf{Game}\ i]$ be the adversary's success probability in game Game i, and let Q be the total number of oracle queries. The challenge oracle is stateful: every block consumes the next $\ell+1$ message keys in the same symmetric chain.
Game 0: This game is exactly the IND-CCA experiment and oracle. Hence,

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A}) = \left|\Pr[\mathsf{Game}\ 0] - \frac{1}{2}\right|.$$

Game 1: In this game, for each block h, we replace the keyed transcript-binding function used in the challenge oracle and in aDec with an independent random function $\mathcal{F}_h : \{0,1\}^* \mapsto \{0,1\}^\lambda$. Concretely, both uses of $\mathsf{PRF}(\mathsf{mac}_{\mathsf{key}_{h,\ell+1}}, \cdot)$ are replaced by $\mathcal{F}_h(\cdot)$.
Let $\mathcal{B}_1$ be a PPT adversary against PRF security. By PRF security and a hybrid over the blocks, we have

$$|\Pr[\mathsf{Game}\ 0] - \Pr[\mathsf{Game}\ 1]| \le \mathsf{Adv}^{\mathsf{prf}}(\mathcal{B}_1).$$

$\mathsf{Exp}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A})$

$\mathsf{rk}_0 \leftarrow \{0,1\}^\lambda$
$(\mathsf{sk}_A, \mathsf{pk}_A) \leftarrow \mathsf{Keygen}(\lambda)$
$(\mathsf{sk}_B, \mathsf{pk}_B) \leftarrow \mathsf{Keygen}(\lambda)$
$\mathsf{dh}_{\mathsf{out}} \leftarrow \mathsf{DH}(\mathsf{sk}_A, \mathsf{pk}_B)$
$(\mathsf{rk}_1, \mathsf{ck}_0) \leftarrow \mathsf{KDF}_{\mathsf{rk}}(\mathsf{rk}_0, \mathsf{dh}_{\mathsf{out}})$
$\mathsf{info} = \mathsf{dh}_{\mathsf{out}}$ or rk_0
$S = \emptyset;\ H = \emptyset$
$h = 0;\ j = 0$
$b \leftarrow\$ \{0,1\}$
$b' \leftarrow \mathcal{A}^{\mathcal{O}_{\mathsf{fMAC}_b}, \mathcal{O}_{\mathsf{aDec\text{-}MAC}}}(1^\lambda, 1^\ell, \mathsf{info}, \mathsf{pk}_A, \mathsf{pk}_B)$
if $b' = b$ **then**
  return 1
else
  return 0

$\mathcal{O}_{\mathsf{fMAC}_b}(\overrightarrow{\mathsf{amsg}}_0, \overrightarrow{\mathsf{amsg}}_1, \overrightarrow{m}, \overrightarrow{\mathsf{ad}})$

$h \leftarrow h + 1$
for $i \in [1, \ell + 1]$ **do**
  $j \leftarrow j + 1$
  $(\mathsf{ck}_j, \mathsf{mk}_j) \leftarrow \mathsf{KDF}_{\mathsf{ck}}(\mathsf{ck}_{j-1})$
  $(\mathsf{aes}_{\mathsf{key}\,h,i}, \mathsf{mac}_{\mathsf{key}\,h,i}, \mathsf{IV}_{h,i}) \leftarrow \mathsf{KDF}_{\mathsf{mk}}(\mathsf{mk}_j)$
  $\mathsf{ct}_{h,i} \leftarrow \mathsf{AES}_{\mathsf{CBC}}(\mathsf{aes}_{\mathsf{key}\,h,i}, \mathsf{IV}_{h,i}, m_i)$
$C \leftarrow \mathsf{fakeMAC}(\overrightarrow{\mathsf{mac}_{\mathsf{key}\,h}}, \overrightarrow{\mathsf{ct}}_h, \overrightarrow{\mathsf{ad}}, \overrightarrow{\mathsf{amsg}}_b)$
$H[h] \leftarrow (\overrightarrow{\mathsf{mac}_{\mathsf{key}\,h}}, \overrightarrow{\mathsf{ct}}_h, \overrightarrow{\mathsf{ad}})$
$S = S \cup \{(h, C)\}$
return $(h, \overrightarrow{\mathsf{aes}_{\mathsf{key}\,h}}, \overrightarrow{\mathsf{IV}}_h, \overrightarrow{\mathsf{ct}}_h, \overrightarrow{\mathsf{ad}}, C)$

$\mathcal{O}_{\mathsf{aDec\text{-}MAC}}(h, C^*)$

if $(h, C^*) \in S$
  return $\perp$
if $h \notin H$
  return $\perp$
else
  parse $H[h] = (\overrightarrow{\mathsf{mac}_{\mathsf{key}\,h}}, \overrightarrow{\mathsf{ct}}_h, \overrightarrow{\mathsf{ad}}_h)$
  return $\mathsf{aDec}(\overrightarrow{\mathsf{mac}_{\mathsf{key}\,h}}, \overrightarrow{\mathsf{ct}}_h, \overrightarrow{\mathsf{ad}}_h, C^*)$

Fig. 7: Security experiment for IND-CCA security of the covert channel in the symmetric ratchet

Game 2**:** In this game, each $\mathcal{F}_h$ is defined by lazy sampling: on input T, if T is already in the table for block h, return the stored value σ; otherwise, sample $\sigma \leftarrow\$ \{0,1\}^\lambda$, store it, and return it. Game 1 and Game 2 use equivalent methods to define the random functions, and thus

$$\Pr[\text{Game 1}] = \Pr[\text{Game 2}].$$

Game 3**:** We first define a Forge event: the aDec algorithm accepts a fresh tag block (h, C^*) that was not returned by the challenge oracle. Parse $C^* = (\mathsf{mac}^*_1, \ldots, \mathsf{mac}^*_{\ell+1})$ and let $T^* = \mathsf{mac}^*_1 \| \cdots \| \mathsf{mac}^*_\ell$. If T^* is fresh for block h, then $\mathcal{F}_h(T^*)$ is uniform and hidden, so acceptance requires guessing $\mathsf{mac}^*_{\ell+1} = \mathcal{F}_h(T^*)$. If T^* equals the challenge transcript for block h, then the only accepting final tag is the challenge final tag, which makes the whole block equal to the rejected challenge block. Therefore a fresh accepting query succeeds with probability at most $\frac{1}{2^\lambda}$ per query, and thus

$$\Pr[\mathsf{Forge}] \leq \frac{Q}{2^\lambda}.$$

This game is identical to Game 2, except that the decryption oracle aborts on the first accepting fresh decryption query. Then,

$$|\Pr[\text{Game 2}] - \Pr[\text{Game 3}]| \leq \Pr[\mathsf{Forge}].$$

Game 4**:** In this game, we replace the first ℓ challenge tags mac_i for every block with fresh random elements. By Lemma 6, all block keys are jointly indistinguishable from independent uniform keys. A hybrid over the $Q\ell$ one-time MAC keys in the first ℓ positions of each block, together with PRF security of MAC, replaces the corresponding MAC outputs by uniform values. The binding-position keys are not used in this hop, since their transcript checks were already replaced by the random functions $\mathcal{F}_h$ in Game 1.
Let $\mathcal{B}_2$ be a PPT adversary against the anamorphic security defined above. Then, we have

$$|\Pr[\text{Game 3}] - \Pr[\text{Game 4}]| \leq \mathsf{Adv}^{\mathsf{ana}}_{\mathsf{MAC}}(\mathcal{B}_2).$$

Game 5**:** In the final game, the challenge tag blocks are independent of the hidden bit b, and the decryption oracle rejects every fresh accepting query by the previous hop. Therefore, we have

$$\Pr[\mathsf{Game}\ 5] = \frac{1}{2}.$$

Combining the hops, we derive

$$\mathsf{Adv}^{\mathsf{ind\text{-}cca}}_{\mathsf{MAC}}(\mathcal{A}) \leq \mathsf{Adv}^{\mathsf{prf}}(\mathcal{B}_1) + \mathsf{Adv}^{\mathsf{ana}}_{\mathsf{MAC}}(\mathcal{B}_2) + \frac{Q}{2^\lambda}$$

which is negligible. □

D Ethical Considerations

Our work sits at the intersection of privacy technology and regulatory debate. We embed covert communication channels into the Signal Double Ratchet protocol, enabling users to exchange hidden messages even when an adversary can decrypt their regular communications. We present a stakeholder-based ethics analysis following the USENIX Security guidelines.

Stakeholders

Our work primarily concerns people who need private communication under surveillance. This includes journalists protecting sources, activists, and ordinary citizens in countries where encryption may be compromised by law. For these users, our constructions offer a layer of protection when traditional E2EE fails.

Developers and maintainers of messaging platforms gain a clearer understanding of where covert channels can persist inside widely deployed protocols, which informs threat modeling and security audits. However, platforms may also face pressure to deploy countermeasures, which could introduce new security issues through misinformed mitigations.

For policymakers and law enforcement, our work provides technical input to ongoing debates. We demonstrate that decryptability does not imply detectability, and that covert channels survive even under comprehensive key disclosure. This matters for realistic assessment of what surveillance mandates can achieve.

Finally, we recognize the dual-use concern directly. The same techniques protecting vulnerable users could conceal harmful communications by malicious actors. We address this below.

Human Subjects, Data, and Platform Interaction

This research involves no human subjects, no user studies, and no collection of private data. We do not access or monitor real user communications. All experiments run in controlled settings on our own Android devices and emulators, measuring computational overhead and timing metrics.

We do not probe or interact with live Signal infrastructure. Our implementation modifies client and library code only in a research context. We did not contact Signal for vulnerability disclosure because we found no vulnerability. Our constructions extend the protocol without weakening existing guarantees.

Risks and Potential Harms

The primary ethical concern is dual-use. Covert channels can be exploited by malicious actors to evade legitimate oversight. However, anamorphic encryption already exists in the cryptographic literature. Our contribution is applying these techniques to a deployed protocol, not inventing new capabilities.

More broadly, publishing practical techniques may provoke stronger surveillance countermeasures with negative societal consequences.

Mitigations

We are careful about what we claim. Our formal security proofs show indistinguishability under stated assumptions. Our experiments measure timing and overhead. We also specify which keys must remain secret for each construction. Users must assess which threat model fits their situation.

To reduce the risk of enabling misuse, our goal is to inform security research and policy rather than provide deployment-ready tools. We release code to support reproducibility of our measurements and figures, and we avoid packaging a drop-in covert messaging feature or operational guidance aimed at evading oversight for end-user deployment.

Rationale for Conducting the Research

We conducted this research because key-disclosure and lawful-access proposals often assume that decrypting overt messages is sufficient to eliminate private communication. Testing this assumption requires analyzing the protocol components where covert channels could persist. We limited the work to controlled experiments on our own devices and emulators, without interacting with live Signal infrastructure or collecting user data. We considered the dual-use risk before implementation and scoped the artifact toward reproducibility of the paper's claims rather than an end-user covert messaging tool.

Rationale for Publication

Publication is justified because it addresses a question with direct policy relevance: under key-disclosure models, does decrypting overt messages eliminate covert communication? Our answer is no. Covert channels persist in protocol elements like DH keys and MAC tags. This finding benefits implementers, auditors, and policymakers by enabling informed decisions.

We acknowledge residual dual-use risk. We believe the benefits to vulnerable populations, scientific understanding, and policy clarity outweigh this risk. The barrier to actual deployment is high, requiring pre-shared secrets, modified clients on both ends, and active cooperation between parties.

E Open Science

Our code is available at `https://doi.org/10.5281/zenodo.20425141`.