Twilight: A Differentially Private Payment Channel Network

Maya Dotan, Saar Tochner, Aviv Zohar, and
Yossi Gilad, *The Hebrew University of Jerusalem*

https://www.usenix.org/conference/usenixsecurity22/presentation/dotan

This paper is included in the Proceedings of the
31st USENIX Security Symposium.

August 10–12, 2022 • Boston, MA, USA

978-1-939133-31-1

Open access to the Proceedings of the
31st USENIX Security Symposium is
sponsored by USENIX.

Twilight: A Differentially Private Payment Channel Network

Maya Dotan*, Saar Tochner*, Aviv Zohar, and Yossi Gilad

The Hebrew University of Jerusalem

Abstract

Payment channel networks (PCNs) provide a faster and cheaper alternative to transactions recorded on the blockchain. Clients can trustlessly establish payment channels with relays by locking coins and then send signed payments that shift coin balances over the network's channels. Although payments are never published, anyone can track a client's payment by monitoring changes in coin balances over the network's channels [23, 31]. We present Twilight, the first PCN that provides a rigorous differential privacy guarantee to its users. Relays in Twilight run a noisy payment processing mechanism that hides the payments they carry. This mechanism increases the relay's cost, so Twilight combats selfish relays that wish to avoid it using a trusted execution environment (TEE) that ensures they follow its protocol. The TEE does not store the channel's state, which minimizes the trusted computing base. Crucially, Twilight ensures that even if a relay breaks the TEE's security, it cannot break the integrity of the PCN. We analyze Twilight in terms of privacy and cost and study the trade-off between them. We implement Twilight using Intel's SGX framework and evaluate its performance using relays deployed on two continents. We show that a route consisting of 4 relays handles 820 payments/sec.

1 Introduction

Blockchain systems such as Bitcoin create a public ordered log of transactions without relying on a trusted party. Instead, these systems distribute trust among potentially many participants that connect in a peer-to-peer network. They allow for fair trading protocols [4] and can reduce friction in financial markets [17]. Users expect meaningful privacy guarantees from financial systems, but most blockchains only provide so-called "pseudo-anonymity", where actions are associated with pseudonym addresses rather than the users' identities. Several works show that often anyone can link users to their addresses by analyzing the public information posted on the blockchain [35, 37]. To counter such analysis, researchers proposed systems like ZCash [3], which rely on zero-knowledge proofs to provide excellent privacy. However, these systems bear significant performance costs due to their reliance on a system-wide consensus regarding the transactions' log.

A promising direction for achieving high-throughput and low-latency transactions with a meaningful privacy guarantee is connecting users through a payment channel network (PCN). In a PCN, users create payment channels with relays by locking coins in joint on-blockchain accounts and use the blockchain again only to commit the coin distribution when a channel closes. The relays are interconnected through payment channels as well. Alice pays an indirectly-connected Bob by finding a route of channels through the network's relays to him and then creating a payment that would move coins to the first relay, conditioned on a secret that only Bob knows. Each relay creates a similar payment for the next hop until Alice's payment reaches Bob, who reveals the secret. Intuitively, since users neither broadcast payments they make nor create payment channels with the users they transact, PCNs seem promising for simultaneously addressing the blockchain's sclability bottleneck and privacy challenge.

In practice, however, PCNs offer very little privacy [18, 24, 26, 31], much less than users would expect from traditional payment systems (e.g., governed by banks). Any user can test whether the liquidity (number of coins available to transact) on any channel is below a threshold that they choose – merely by asking to relay a payment on that channel for the threshold amount. As a result, anyone can learn about changes in each channel's liquidity, which leaks all the information about the network's payments and creates a privacy risk in practice [23].

We present Twilight, a PCN that hides a user's payment history from other users. Twilight provides a strong privacy guarantee supported by a differential privacy analysis. It models payments over a channel as queries, and the responses convey whether or not they can go through the channel (i.e., the channel has sufficient liquidity). Twilight provides payment-privacy by noising these responses. Instead of carrying a payment whenever sufficient liquidity exists, the relay rejects

*Both authors contributed equally

payments if they would not leave enough liquidity to cover a randomized "noise" payment. Through this noisy payment processing mechanism, Twilight ensures that only a small amount of statistical information about the payments it carries may leak to other users. The key to designing Twilight is structuring the noise such that the privacy guarantee holds even if the attacker uses many clients to send queries at a high rate and observes the responses over a long time. Since queries are cheap to execute, this property is crucial to providing a meaningful privacy guarantee. At the same time, it is important to avoid throttling the rate of queries, which hurts honest users and lowers the system's throughput even when there is no attack. To achieve this, Twilight leverages techniques from the differential privacy literature to reuse noise values when hiding the same set of payments [11,34]. This allows Twilight to group payments into subsets and mask them with noise, ensuring that new information about old payments does not leak to the attacker with every query.

Twilight's differential privacy approach comes at a cost: it might block some payments due to the noise that a vanilla PCN would carry. Relays in Twilight mitigate this issue by locking extra coins. Of course, a relay can always tear down the channel and get these coins back but locking coins still incurs an operational expense. We analyze this trade-off between privacy and cost, and evaluate a model where users cover the relay's expense with payment fees. We quantify the financial impairment from locking coins and estimate the payment volume a relay would need to process to cover this cost by charging fees for its privacy service. For example, a relay handling a payment volume of 79-coins/day and charges a 1% fee covers the cost of operating a noisy channel that hides 1-coin payments (for privacy level $\varepsilon = 0.15, \delta = 10^{-7}$, in differential privacy terms [12]).

Selfish relays may claim to their users that they do noisy payment processing but attempt to avoid it in practice to save costs. Twilight addresses this problem by running the noising logic in a trusted execution environment (TEE), which allows clients to verify the payment processing logic. We architect the code operating inside the TEE to avoid keeping the channels' state. In particular, it does not keep track of the channel's liquidity or payments already processed. Instead, the relay provides this state every time it calls the TEE to process a payment. Designing Twilight in this manner minimizes the trusted computing base inside the TEE but requires handling two key challenges. First, a relay might inform the TEE that it has very high liquidity, enough to cover any noise it might choose. Second, a relay might attempt to eliminate the effect of noising payments by repeatedly calling the TEE with the same payment until the TEE adds sufficiently-low noise to approve it. Twilight handles these challenges by noising payments on the relay's incoming channel and encrypting payments between TEEs until they reach the recipient. A relay cheating its TEE about the incoming channel's liquidity only risks losing coins by accepting payments that the previous hop cannot cover. Each TEE outputs its noisy payment processing result encrypted for the next TEE on the route, so all of a TEE's outputs are indistinguishable, and the relay cannot choose the response it likes. Only when a payment reaches the recipient, all relays en-route can decrypt the TEE outputs and update the channel liquidity. Importantly, Twilight does not rely on TEEs to secure funds. Even if a relay exploits a vulnerability in the TEE or Twilight's code that it executes to learn the TEE's secrets, all payments Twilight carries are valid, and the relay cannot steal coins from others.

We implement Twilight and test its performance using machines in America and Europe with the SGX TEE [9] in Azure. A route of 4 relays supports 820 payments/sec, and at this peak rate, the payment latency is 550ms above the network latency (which is about 510ms in our experiments).

In summary, our contributions are the following:

- A rigorous definition for privacy in PCNs based on differential privacy, and the design of Twilight, a PCN that meets this goal.
- The combination of noising payments inside a TEE to combat selfish relays.
- An analysis of Twilight's integrity, privacy and cost.
- An implementation and performance evaluation.

2 A Primer on PCNs

We overview the key concepts and mechanisms that comprise PCNs as background for describing Twilight.

Payment channels. Two participants establish a bidirectional payment channel by "locking" coins in a joint on-chain account that requires both parties to approve every spend. The amount that each participant locks is the channel's initial liquidity in the direction of their peer. The locked coins ensure that payments carried over the channel can always be redeemed. Alice pays Bob by signing a transaction that adjusts the split of their account balance (giving more coins to Bob than he deposited). Bob can counter-sign this transaction at any time and redeem his funds by posting it on the blockchain, which would also close the payment channel. Instead, Bob holds the transaction from Alice. He keeps the channel open to allow them to continue updating the balance split. A short appeal period begins when either user posts a transaction that closes the channel. During this time, posting a transaction with information that proves that the channel was closed with an obsolete state corrects the account split. In this manner, the appeal period protects users if their counterpart closed the channel using an outdated transaction.

Routing payments. A PCN comprises clients that issue payments and relays that carry these payments between clients for a fee. Relays and clients connect through payment channels, allowing Alice to pay Bob even if they are not directly connected. When two relays establish a channel, they start

announcing the channel through a peer-to-peer network. Alice's client finds a route to Bob and onion-encrypts it with the relays' public keys. Namely, it encrypts Bob's identity using the last relay's key, concatenates to the ciphertext that relay's identity, and encrypts again with the previous relay's key. It continues in this fashion for every hop. When a relay receives a payment, it subtracts a fee from the amount and decrypts the top layer to find the next hop.

Hash-time-locked contracts (HTLCs). Hash-time-locked contracts are a mechanism for ensuring that every relay along a payment's route that sends coins to the next hop can recover the funds from the previous hop. Before Alice can pay Bob, his client chooses a random secret s and gives Alice's client $h(s)$, the cryptographic hash of s. Alice's payment moves coins to the first relay conditioned on the relay providing s before a deadline; this condition is called an HTLC. Each relay creates a similar payment conditioned on the preimage of $h(s)$, which it sends to the next hop in the route. As soon as Bob receives the payment, he can potentially post s on the blockchain and redeem the payment; this would also reveal s to all the other relays, allowing them to redeem their payments as well. Typically, however, Bob avoids the on-chain transaction: his client reveals s to the last relay and asks the relay to sign a transaction that updates the split in the channel between them – this time, without conditioning on s. That relay can then do the same, show s to the previous hop, and update the previous channel's split. This way, s propagates back through the route, updating all channel balances and eventually reflecting that Alice paid Bob.

3 Overview

Figure 1 shows a PCN connecting three users. The users connect to the PCN via private channels (illustrated by the dashed links in Figure 1). Their clients reject payments from unauthorized origins [36] which mitigates exposing information about the liquidity on these channels to other users. In this figure, Alice sends a payment to Bob via two inter-relay channels. In contrast to the user's private channels, relay-to-relay channels are public and any client can route payments over them. This allows an attacker to probe for changes in channel liquidity by asking the relay to route payments between his clients and observe whether the relay rejects his payments for exceeding the available liquidity. By probing for changes in liquidity across inter-relay channels, an attacker can track payments between users in the network [16, 18, 31]. Twilight is a PCN that hides its users' payments from attackers probing inter-relay channels by introducing noise to a relay's payment processing logic. It ensures users that relays add this noise by leveraging a trusted execution environment (TEE). We next formalize the threat model and Twilight's goals against it.

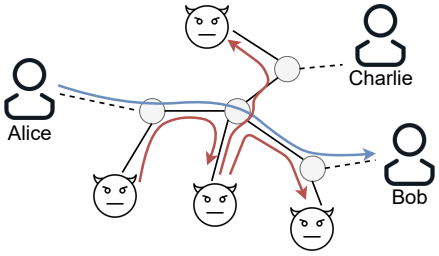

Figure 1: Users connect to relays through private channels (dashed black line), while inter-relay channels are public (solid black lines). An attacker sends payment requests between his clients (in red) to monitor changes in liquidity and infer about payments between users (in blue).

3.1 Threat model

Before delving into defining Twilight's threat model and goals, we discuss different types of potential attackers.

Off-path clients. Today's PCNs are vulnerable to attacks by *clients*: anyone on the Internet can probe payment channels and learn about others' payments [23, 31]. This constitutes a lower standard of privacy compared to traditional financial systems. Such attacks are easy to launch and difficult to block since the attacker may send probes via multiple clients and disguise probes as legitimate payment requests.

On-path selfish relays. Users should *know* their payment history is hidden from other clients. In particular, PCNs are designed as distributed systems that include relays operated by multiple autonomous organizations. A privacy solution that involves costs that intermediate relays can shirk (at the expense of users' privacy) should ensure that relays follow the protocol despite these costs.

On-path adversarial relays. Much like other PCNs, Twilight uses onion routing for payments (§2), which protects against network adversaries and malicious relays that only have one local visibility or presence in the route, so they can only view the previous or next hop but not the entire route. However, colluding relays located near the payer and payee may still track the payment by correlating its time, payout value, or locked contract's secret. Such attacks are harder to launch than those involving only off-path clients: as shown in [42], the number of relays on the route is small, and since the payer chooses which route to use, she can avoid relays that she considers less trustworthy (e.g., operated by less reputed organizations). We architect Twilight to provide similar usability to vanilla PCNs but with better privacy; it, therefore, does not handle such correlation attacks, since that would require delaying payments and somehow ensuring that Bob's balance changes regardless of Alice's balance changes (to break these correlations). Recent works allow decorrelating the locked contract's secret [31, 32]; we discuss it in related work. We believe that techniques from the private communi-

cation literature, such as Poisson mixing [33], where every relay delays payments using a randomized mechanism can address the timing challenge, but hiding correlated changes in balances remains a challenge. We leave this attacker model out of Twilight's scope.

Secrecy of closing balances. Payment channel teardown involves posting a blockchain transaction stating its closing balance split (§2). Since anyone watching the blockchain can observe the channel-close transaction, the privacy that Twilight can provide depends on the underlying blockchain. Twilight gives the most privacy when the PCN deploys over a blockchain that supports private transactions (e.g., using zero-knowledge proofs [25]). However, the closing transaction reveals only the aggregate sum of payments over the channel through its lifetime. In practice, payment channels are open for a long time [1], so we expect the closing balance to reveal very little about a user's payments. We design Twilight to be largely-independent of the blockchain choice (except for requiring smart contract support), and it can deploy over blockchains with or without support for private transactions.

3.1.1 Twilight's attacker

Twilight hides a user's payment history in the face of attackers who can run any number of the PCN's clients even if the user routes her payments through selfish relays. The attacker can route payments over any inter-relay channel to probe the available liquidity. These probes are extremely cheap: the attacker can abort the payments after seeing the response and avoid the associated fee.

3.2 Goals

Integrity. Twilight should enable private off-chain payments as a viable and trustless substitute to on-chain payments. Therefore, it must guarantee that the sequence of all payments that a channel carries can be committed to the blockchain. This property must hold even if all relays break the TEE security guarantee (i.e., they attest different code than what they execute and learn all the TEE's secrets).

Privacy. There are several ways to capture privacy in PCNs. For one, Twilight could target hiding just the payment amount. However, in many cases exposing that Alice pays Bob is sufficient to reveal sensitive information. For example, donating to a political party, no matter what amount, tells Alice's political views. Twilight could also ensure that Alice has at least one "cover story," making paying Bob appear similar to paying at least one other user. However, this approach might provide users with a cover story that does not make sense in practice (e.g., the alternative payee might be in another continent), and users might not be aware of the cover story the system then provides. Instead, Twilight sets an ambitious goal: ensuring that the attacker's view through the clients he operates is likely to be the same regardless of whether Alice makes her payment. In particular, this approach makes Alice paying Bob look similar to Alice paying *any other* user. (Since Alice paying Bob looks similar to not making a payment, which then looks similar to Alice paying any other user.)

More formally, consider the vector O of the attacker's observations (from routing payments between malicious clients) and two scenarios: one scenario where a user, call her Alice, pays another user, call him Bob, and the other scenario where Alice never makes this payment. Twilight's goal is to provide (ε, δ)-differential privacy [12] with respect to the two scenarios above. It guarantees that the following inequalities hold for a small $\varepsilon \geq 0$, except for a small error probability $\delta \geq 0$:

$$e^{-\varepsilon}\Pr[O|\mathsf{X}] \leq \Pr[O|\text{Alice} \rightarrow \text{Bob}] \leq e^{\varepsilon}\Pr[O|\mathsf{X}] \qquad (1)$$

The arrow denotes Alice paying Bob, and the X-mark denotes the case where she does not make this payment. The probability is over the coin-tosses in the nosing mechanism. Differential privacy quantifies the statistical information that leaks to the attacker. It provides a strong formal guarantee for the level of privacy users should expect. This privacy guarantee holds even for multi-hop payments (payments routed through several payment channels between Alice and Bob); we quantify impact of a payment's path length on its privacy in terms of ε and δ.

	Description
$\mu, \sigma > 0$	mean and standard deviation for Gaussian noise
T	number of time-slots (leaves in the tree)
$\mathcal{N}_t$	the minimal set of tree nodes covering $[0, t-1]$
N	$\lvert\mathcal{N}_t\rvert$ upper bound, $N = \lceil \log_b T \rceil + b - 2$
$b \geq 2$	children per node in tree

Table 1: Symbols of the noisy payment processing mechanism

4 Noisy Payment Processing

Twilight protects users against attackers probing inter-relay links (Figure 1). It provides a strong privacy guarantee, regardless of the attacker's probing rate, by adapting ideas for continuous release of information from the differential privacy literature [11, 12, 28] to the PCN context. The system splits time into short intervals, arranged in a tree, and models the attacker's probes as queries arriving in those intervals. Twilight cannot distinguish between the attacker's probes and real payments so it must respond to both. It protects users' privacy through noisy payment processing. Twilight noises the decision whether to carry or drop a payment (the relay's response) in a structured way, depending on the interval the query arrived in [34]. Table 1 summarizes the notations of the noisy payment processing mechanism described below.

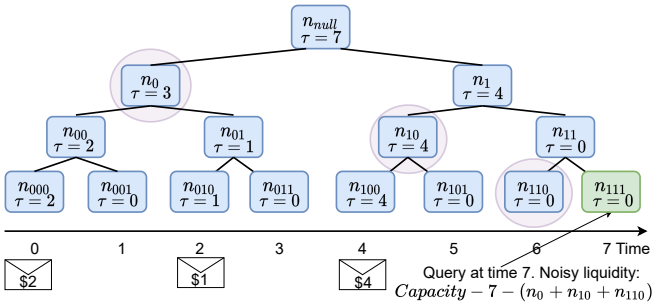

Figure 2: A channel's noise tree with $T = 2^3$ slots. For node i in the tree, n_i is the noise and τ_i the sum of payments in its subtree. The response mechanism for query at time 7 uses the nodes $\{0, 10, 110\} \in N_7$ (circled in red).

4.1 Response mechanism

Twilight divides the lifetime of a channel into T tiny time-slots. In every slot, at most one request to carry a payment may arrive. For example, dividing time into nanosecond slots seems reasonable since payments are over 100B-long (they include account addresses, a signature, etc.). Even with 100Gbps links, each request takes at least 7 time-slots to receive. Allocating 2^{64} nanosecond slots supports a channel's operation for over 500 years.

Time slots are numbered and arranged in the leaves of a tree; each node in the tree has a label. Consider a binary tree, for now. The root has the empty label, and every node's left/right child appends '0'/'1' to the parent's label. Figure 2 illustrates this tree structure. The relay assigns every node i in the tree an independently drawn Gaussian noise, $n_i \in_R Gauss(\mu, \sigma^2)$. In addition, node i stores τ_i, the sum of all payment amounts over the channel (positive or negative, depending on the direction) in the time-slots beneath it.

When a payment requesting to move m coins to the channel's other endpoint arrives at the relay at time slot t, the relay finds $\mathcal{N}_t$, the minimal set of nodes in the tree that precisely covers (i.e., includes the ancestors of) the time slots in the interval $[0, t-1]$. (We give the algorithm to find the minimal covering and prove its correctness in an online technical report [10].) The relay agrees to carry the payment if Equation 2 holds:

$$channel.Capacity - \sum_{i \in \mathcal{N}_t} (\tau_i + n_i) \geq m \qquad (2)$$

Figure 2 illustrates the response mechanism for a request at time t = 7. In this example, $\mathcal{N}_7 = \{0, 10, 110\}$ is the minimal covering set of nodes for the time slots preceding the query (slots 0–6). Since $\mathcal{N}_t$ is the minimal set of tree-nodes that covers the interval $[0, t-1]$, $\sum_{i \in \mathcal{N}_t} \tau_i$ is the sum of all payment amounts before time t, so $channel.Capacity - \sum_{i \in \mathcal{N}_t} \tau_i$ is the current liquidity. The response in Equation 2 obfuscates it by subtracting noise, i.e., the aggregate of Gaussians $\sum_{i \in \mathcal{N}_t} n_i$.

The tree structure from the differential privacy literature [11] ensures that for any user payment at time $t' < t$, there is only one ancestor of the leaf node t' in the covering $\mathcal{N}_t$. Only that ancestor might leak information about user's payment (by contributing a different value to the response mechanism if the user never makes this payment). Since there are only $\log T$ ancestors for any slot, there are only a few tree nodes that might leak information about a user's payment over time (even if the attacker probes often).

Intuitively, the more nodes in $\mathcal{N}_t$, the more noise the payment processing mechanism subtracts from the channel's capacity in Equation 2, and the less accurate the response becomes. Qardaji et al. [34] optimize privacy and accuracy using a tree with a higher branching factor. When every node in the tree branches to b children (Table 1), the size of the covering set $|\mathcal{N}_t| \leq \lceil \log_b T \rceil + b - 2$ (one node from each level in the tree, except the last level where there may be up to $b-1$ leaves). We denote this upper bound by N (Table 1). With $T = 2^{64}$ time slots, using $b = 8$ results in the minimal number of nodes in the covering, $N = 28$ nodes (rather than 64).

If the covering set has less than the maximal N nodes, the relay "pads" it with dummy nodes, i.e., nodes where $\tau = 0$ and fresh noise is drawn. So, every time a relay responds it uses the aggregate of N Gaussian random variables as noise. The reason for padding is that the aggregate is much more likely to be non-negative than the noise of a single node. We later use this property to prove Twilight's integrity (§6).

4.2 Stateless noising

Storing the tree illustrated in Figure 2 is impractical for a large number of time-slots (e.g., 2^{64}). Every node i in the tree contains two elements: the sum of payment amounts in the time-slots of its subtree (τ_i) and an independently drawn noise (n_i). The response mechanism in Equation 2 only uses the sum of all payment amounts before the request's time slot. A relay, therefore, only keeps track of this sum. It also avoids storing the noise values and instead recomputes them when they are needed. Specifically, the relay keeps one global secret s, and for a payment-channel *chanID*, sets the secret value of tree node i to be deterministically drawn from the noise distribution seeded by the hash value $h(s, chanID, i)$.

5 Countering Selfish Relays using TEEs

Users rely on Twilight's relays for privacy, but how can a payer trust that the relays she chooses for her payment's route indeed perform noisy payment processing (§4)? Moreover, she has to be convinced that the relays *continue* to perform noisy payment processing since the attacker's future queries may leak information about her payment. Providing users with this guarantee is important since relays may try to circumvent the noising mechanism to avoid rejecting payments they would otherwise carry and losing the associated fees.

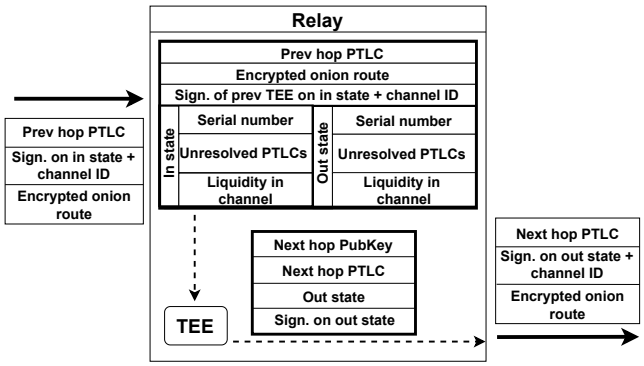

Figure 3: Protocol messages and relay-to-TEE interface. We abbreviate incoming/outgoing channel state by in/out state.

Twilight addresses this concern using a trusted execution environment (TEE) on its relays. TEEs ship with commodity hardware (e.g., SGX is now standard in Intel's i9 10th generation processors [19]); their remote attestation allows clients to verify the payment processing logic, which relays run inside an enclave [9]. Each channel is associated with a state that includes the current balance split between peers, all current pending payments, and a message serial number (for dispute resolution). Our design does not require the TEE to keep track of this state, which minimizes the trusted computing base and simplifies Twilight's implementation but introduces challenges as we describe next. Crucially, Twilight ensures that even if a rogue relay breaks its TEE and learns its secrets, the PCN maintains integrity: all payments over the channel are valid, and the relay cannot steal coins from others.

Challenges. Twilight must force a relay to involve the TEE in processing payments to ensure it performs noisy payment processing throughout the channel's lifetime. Figure 3 illustrates how a relay interacts with its TEE. The enclave running inside the TEE does not store the channel's state but receives it from the relay with every payment processing request to decide whether to approve a payment. This leads to two key challenges: First, the relay may provide the enclave with incorrect or outdated states. Specifically, ones in which a high amount of liquidity is available. Such manipulation can ensure that the noisy liquidity check in Equation 2 always passes. The relay may do this selectively, whenever it has sufficient liquidity to forward the payment and wishes to pass the noisy check with greater certainty. Second, the TEE cannot remember which payments it processed, so the relay may invoke the TEE multiple times to process the same payment until it adds sufficiently low noise to accept it.

Architecture. Twilight solves these challenges by hiding payment amounts and the TEE's responses from the relay, encrypting them for the TEE at the next hop. The enclave running inside the TEE performs noisy payment processing on the incoming channel when the relay receives a new payment request. Namely, the enclave (noisily) checks that the previous hop has enough liquidity on the channel to cover the current payment. To do so, the enclave gets from the relay the incoming channel's state, which consists of the channel's liquidity and a list of "unresolved" payments (for which the secret that completes the coin transfer has yet to be revealed, see §2), and processes the payment as in §4. The reasoning behind noising payments on the incoming channel (rather than the outgoing channel) is to mitigate a relay's incentive to cheat. Suppose the relay informs its enclave of false high liquidity on the incoming channel; this nullifies the check that the previous hop has sufficient liquidity to cover the payment, so the relay risks moving coins to the next hop without any return. There is no privacy risk if the relay reports false liquidity: since the TEE outputs are hidden from the relay (encrypted for the next hop), it cannot choose a response it likes. Therefore, it cannot adjust the noise distribution to the payment; instead, it has to submit the noisy response to the next hop. Privacy holds even if the relay colludes with the relay in the next hop, who also cannot access the information, which is encrypted for its TEE. We next explain how the TEEs establish public keys that facilitate this encryption and how relays interact with their TEEs to process payments. We provide a fully detailed version of the protocol in an online technical report [10].

5.1 Establishing public keys and channel IDs

The TEE creates a secret key, and the relay binds the corresponding public key to its payment channel. When two parties establish a new payment channel, they deposit coins to an account managed through a smart contract. Each party gives one public key to govern this account; a relay uses the public key corresponding to its TEE's secret key. In Twilight, clients contact the relays and check: (1) that the key was created in the TEE, and using remote attestation, (2) that the logic inside the TEE never exports the keys it creates and performs noisy payment processing. By registering the TEE's public key in the smart contract, the relay ensures its users that the TEE must be processing their payments on this channel. The smart contract's account address serves as the channel ID; signed messages about the channel include this ID, so they cannot be confused with other channels. The smart contract allows flexibility in designing dispute resolution when a channel closes. Specifically, Twilight uses the smart contract functionality to allow an offended party to reveal an unsettled encrypted payment (and correct the closing balance, if needed), as we describe in §5.3.

5.2 Processing payments inside TEEs

In Twilight, relays receive and forward encrypted payments. As a result, relays do not know the outcome of noisy payment processing, so payment messages always reach the payee,

even if a TEE on some relay along the route rejects the payments. Twilight does not keep payments hidden forever. A relay must be able to claim a resolved payment by posting a transaction on the blockchain with the payee's secret (which conditions the coin transfer across the route) in case of dispute with its partner. To support this functionality, the payee chooses a fresh private/public encryption key-pair for the payment's locked contract. The public key is unique to that payment's locked contract and allows hiding its amount. Instead of indicating "failure" when there is insufficient liquidity, the relay outputs a locked contract for the next hop with a hidden amount of zero (indicating no payment). The relay can decrypt the payment amount and redeem it when learning the corresponding private key from the payee, which also allows claiming the payment and is analogous to the HTLC secret from vanilla PCNs (§2). We term these locked contracts *"Private Time-Locked Contracts" (PTLCs)* rather than HTLCs (hash time-locked contracts). PTLCs generalize HTLCs; similarly to HTLCs, they specify an expiration time (e.g., by block depth on the blockchain) and signed by the previous hop (see §2). The main difference between them is that while HTLC payouts are always legible (even if not redeemable without the secret), the payment amounts in PTLCs cannot be read without the secret. This allows providing the PTLC contents specifically for a relay's TEE when the payment propagates to the payee, and preventing the relay from learning the result of its TEE's noisy payment processing from the PTLC it outputs for the next relay.

Consider a payment from Alice to Bob and the private key that Bob chooses for conditioning this payment. Bob informs Alice of the corresponding public key, which she includes in her payment's PTLC. On the route between Alice and Bob, the enclave running in each relay's TEE must learn the payment amount when it processes Alice's payment, i.e., before Bob confirms it and reveals the secret. To support this functionality, Alice uses non-malleable encryption with a fresh ephemeral symmetric key to hide the payment amount in the PTLC. PTLCs include two ciphertexts of that key: (1) under the next hop's TEE public key, and (2) under the PTLC's public key that Bob chose. The TEE decrypts the amount encrypted under its key and checks for consistency with the second ciphertext by encrypting the ephemeral key it recovers with the PTLC's public key. The second encryption is deterministic to allow this check; it is safe since the ephemeral symmetric keys are never reused.

Alice submits her payment for Bob to the first relay in the route she chooses. We next describe how relays process payments and their interaction with their TEEs, following Figure 3.

TEE input. When a relay calls its TEE to process a payment inside an enclave, it provides the payment message from the previous hop, the corresponding incoming channel ID and its liquidity in the direction of the payment, and any PTLCs for unresolved payments from the previous-hop (these are payments that have not been completed yet, but represent commitments from the previous hop).

TEE processing. The enclave decrypts the amount and performs the consistency check for the two ciphertexts above. It subtracts from this amount the relay's fee. It then computes "uncommitted liquidity", which is the incoming liquidity minus the coins in the relay's unresolved PTLCs. Lastly, it performs noisy payment processing (§4) and evaluates Equation 2 to determine whether to accept the payment.

The noise that a relay induces depends on the current time (§4). However, it is unsafe for the relay to provide the time directly to the enclave. Otherwise, a selfish relay could identify some slot in the tree where the nodes in its covering-set ($\mathcal{N}_t$, see Table 1) sum to very little noise, and then replay this slot's timestamp for all payments to essentially circumvent the noising mechanism. Fortunately, TEEs allow enclaves to read the number of cycles since boot, and guarantee its authenticity, which serves as time [20]. The TEE also allows detecting when a relay reboots and the time initializes to zero: on boot, the TEE draws a random number which Twilight's enclave echos to the relay with every response. The relay must provide this value to the TEE when asking to process a payment. The enclave compares the given value by reading the random number again; if they match, the relay has consistent time since the last response (and recursively since boot) and processes the payment as usual. Otherwise, the relay refuses to process the payment. This ensures users that the relay noises payments correctly throughout the channel's lifetime.

When the relay reboots, the random number changes and the TEE refuses to process more payments. The relay then has to close the channel on-chain by posting the last closing balance message from its partner. We expect relays to have high up-time, to collect fees from all payments that need to route through them; so the cost of the blockchain transaction in such "forced" closures is amortized over a long time. A relay may also use another computer with TEE as backup, which keeps track of time and allows recovery in case of a crash without closing the channel. We describe the details of this extension in a technical report [10].

TEE output. The TEE decrypts the top layer of the onion route and creates a new PTLC for the next relay's TEE with the remaining amount (encrypted under a fresh symmetric key for the next hop's TEE and the PTLC public key from Bob). The TEE then signs a message combining the channel ID, its liquidity, all pre-existing unresolved PTLCs and the new PTLC along with an incremented serial number. The next hop checks the TEE's signature on this combination before it continues processing the payment (this signature both ensures that the contents were created by the TEE and hence noised, and that the next hop will be able to claim funds on the blockchain). When the enclave rejects a payment, it uses zero for the amount in the output PTLC. This ensures that the following relays on the route also use zero amounts,

meaning that the payment will not eventually take place.

Payment confirmation and coin transfers. When Bob receives the final PTLC, his client checks that the amount from Alice is sufficient. In this case, Bob reveals the PTLC secret key to the previous relay, which serves as a receipt. It allows the relay to decrypt the PTLC's amount and post the PTLC, if necessary, to the blockchain to ensure the relay gets paid. Each hop then propagates the secret backward, which enables coin transfers across the route (as described in §2).

5.2.1 Side channels

Several exploits in the past illustrate that TEEs can be vulnerable to side channels (see survey on Intel's SGX platform [13]). We architect Twilight to maintain PCN integrity even if attackers completely break the TEE's security promises, but given such a vulnerability, a relay may circumvent Twilight's noisy payment processing and jeopardize its users' privacy. In particular, attackers with physical access to the TEE (the hosting relay in Twilight's case) may exploit channels such as temperature and power consumption readings to learn sensitive information like secret keys. However, TEE manufacturers issued countermeasures in response to known attacks and, using remote attestation, allowed users to learn whether the TEEs they contact run these countermeasures. Users in Twilight, thus, can avoid relays with known TEE vulnerabilities.

Another important type of side channel attack exploits application-specific vulnerabilities within the enclave. Twilight's code inside the enclave must be hardened against such attacks. This is typically achieved by ensuring data-oblivious computation [2, 43] that results in constant processing time, low variance power consumption, etc.

5.3 Channel teardown

When a user leaves the PCN, or a relay needs to replenish liquidity in a channel with another relay, they close their channels (and open another one, in the relay-to-relay case). Closing a channel involves an on-chain transaction to the smart contract, splitting the locked coins between the channel endpoints. Similar to vanilla PCNs, parties sign and exchange closing balance messages after each payment (§2). A relay's TEE signs these messages, as the TEE's secret key is authoritative for the relay's payment channels (§5.1). Each party stores the latest message from their counterpart, so they can post it on-chain to close the channel (even without interacting with their TEE at close-time). These messages have a serial number, local to the channel, and reference unresolved PTLCs on the channel by their hashed values (see Figure 3). The smart contract managing the locked coins in the channel's account allows a short appeal period, where parties may post closure messages with higher serial number, and then allows parties to claim the PTLCs referenced by the last message posting them along with their secrets.

To hide the channel's closing balance split (§3.1), Twilight can deploy over a blockchain that allows for private transactions in its smart contracts. For example, Ethereum supports a rich enough language that allows implementing "shielded transactions" in its smart contracts using zero-knowledge proofs (see [25] for implementation). Such a deployment allows fast off-chain payments with a differential privacy guarantee and avoids exposing information on channel tear-downs. We, therefore, believe this combination makes an attractive privacy-performance trade-off.

6 Analysis

We analyze Twilight's design against its goals from §3.

6.1 Integrity

All payments accepted by a relay must be valid, so that the resulting balances when closing channels may be committed to the blockchain. This property must hold even if all relays break their TEEs' security guarantees (§5), i.e., even in this extreme case, no relay can steal coins from others.

To achieve integrity, it is sufficient to ensure that the relay subtracts non-negative noise from its channel's liquidity (Equation 2). Thus, any payment that the relay accepts is of at most the channel's liquidity amount and does not overspend (so it can be committed to the blockchain). The noise that the relay subtracts is the sum of at most N Gaussian random variables (§4.1). Theorem 1 shows that setting $\mu >> \frac{\sigma}{\sqrt{N}}$ ensures that the chance for negative noise is extremely small.

Theorem 1. *A relay accepts a payment that overspends the channel's liquidity with probability* $\leq GaussCDF(\mu N, \sigma^2 N; 0)$.

Where GaussCDF is the noise cumulative distribution function (CDF) with mean μN and variance $\sigma^2 N$, evaluated at 0. And μ, σ, N are the noise parameters, summarized in Table 1.

Proof. Let us compute the probability that the relay adds negative noise for some time slot. Noisy payment processing adds N random variables to the liquidity, each distributing $Gauss(\mu, \sigma^2)$. Thus, a relay's noise distributes $Gauss(N\mu, N\sigma^2)$. (The sum of Gaussians is also a Gaussian.) The chance for negative noise is $GaussCDF(N\mu, N\sigma^2; 0)$. Namely, the noise CDF evaluated at 0. □

When $N\mu >> \sqrt{N}\sigma$, i.e., the noise's mean is much greater than its standard deviation, the chance of obtaining negative noise is extremely small. It decays proportionally to $e^{-(\frac{\mu\sqrt{N}}{\sigma})^2}$. For example, for $\mu = 10\frac{\sigma}{\sqrt{N}}$ the chance for negative noise error is about 2^{-100}. This property holds regardless of the TEE's security, i.e., even in case a new vulnerability allows the relays to circumvent TEE protections; we provide the

proof for integrity under under insecure TEEs in an online technical report [10].

6.2 Privacy

Before we delve into the analysis of Twilight's differential privacy guarantee, we argue that as long as a relay's TEE is secure, that relay performs noisy payment processing. Consider a relay with a TEE that attests to running noisy payment processing on a channel. Any client that routes a payment through the channel ensures that the secret keys for the channel's on-chain account are created in, and never exported from, the TEE (§5.1). When a relay processes a payment's PTLC, it must sign the PTLC for the next hop using the TEE's secret. Hence, the relay must call the TEE, with the PTLC from the previous hop, which executes noisy payment processing and creates the next PTLC according to the result of Equation 2.

The privacy analysis for noisy payment processing proceeds as follows. We first analyze the differential privacy guarantee when the attacker probes a channel just once (Theorem 2). We then reason about hiding a payment under many probes, traversing a route with several channels, and the privacy amplification from choosing one of several available routes (Theorem 3). Lastly, we discuss differential privacy for multiple payments.

Privacy against a single probe. Consider a payment channel's noise tree (Figure 2) and Alice's payment arriving at the relay at time slot t. The only nodes in the tree affected by her payment are the $\log_b T$ ancestors of slot t (see parameters in Table 1). When the attacker issues a probe at time t', only one of these ancestors affects the relay's response (Equation 2), call it node i. Thus, for a single probe, it is sufficient to analyze the privacy loss from the information in a single tree node. For example, the payment at slot 2 in Figure 2 only affects the amounts (denoted by the τ's) that nodes $010, 01, 0,$ and the root record. Node i contributes to the response the sum of its noise and transacted amounts (i.e., $n_i + \tau_i$). Theorem 2 captures, in (ε, δ)-differential privacy terms, the difference in the chance that i contributes the same value whether Alice makes a payment of m coins.

Theorem 2. *Consider the time slot where Alice may make a payment to Bob for m coins. Let i be a node in a channel's tree that is an ancestor of that time slot. Then, except with probability* δ*, the following inequalities hold:*

$$e^{-\varepsilon} \leq \frac{Pr[\tau_i + n_i | Alice \to Bob]}{Pr[\tau_i + n_i | X]} \leq e^{\varepsilon}$$

Where $\varepsilon = \frac{mc}{\sigma}, \delta = 2 \cdot GaussCDF(\mu, \sigma^2; \mu - c\sigma)$.

Informally, δ bounds the probability of drawing extreme noise values (over c standard deviations below the mean). The parameter $c > 0$ allows to trade a larger ε for a smaller δ.

Proof. Alice pays m coins to Bob. So, for any value η that node i might contribute to the calculation of Equation 2 (i.e., the sum $\tau_i + n_i$), we need to bound the ratio:

$$\frac{\Pr[\tau_i + n_i = \eta | Alice \to Bob]}{\Pr[\tau_i + n_i = \eta | X]} = \frac{\Pr[n_i = \eta - \tau_i - m]}{\Pr[n_i = \eta - \tau_i]}$$

Namely, for node i to contribute the same value in both scenarios (Alice pays m coins to Bob vs. Alice does not make this payment), the noise in case Alice pays Bob should be m less than the noise in the case she does not pay him. The noise n_i distributes $Gauss(\mu, \sigma^2)$. For convenience, let us substitute $x_i = \eta - \tau_i$, which is the noise value that the relay should draw if there is no payment from Alice to Bob (s.t. i contributes η). The Gauss distribution PDF gives that the above term equals:

$$= \frac{e^{-\frac{(x_i - m)^2 - 2x_i\mu + 2m\mu + \mu^2}{2\sigma^2}}}{e^{-\frac{x_i^2 - 2x_i\mu + \mu^2}{2\sigma^2}}} = \frac{e^{-\frac{x_i^2 - 2mx_i + m^2 - 2x_i\mu + 2m\mu + \mu^2}{2\sigma^2}}}{e^{-\frac{x_i^2 - 2x_i\mu + \mu^2}{2\sigma^2}}} = e^{-\frac{-2mx_i + m^2 + 2m\mu}{2\sigma^2}} = e^{\frac{2mx_i - m^2 - 2m\mu}{2\sigma^2}} \quad (3)$$

The chance that the relay draws an "extreme" noise value for x_i is low. More precisely, except with probability $\delta_{right} = 1 - GaussCDF(\mu, \sigma^2; \mu + c\sigma) = GaussCDF(\mu, \sigma^2; \mu - c\sigma)$, it holds that $x_i \leq \mu + c\sigma$. Substituting x_i in Equation 3 with this upper-bound, we get that except with probability δ_{right}:

$$\leq e^{\frac{2m(\mu + c\sigma) - m^2 - 2m\mu}{2\sigma^2}} \leq e^{\frac{cm}{\sigma}} = e^{\varepsilon}$$

The computation for the inequality in the other direction, showing that $e^{-\varepsilon} \leq \frac{\Pr[\tau_i + n_i | Alice \to Bob]}{\Pr[\tau_i + n_i | X]}$, is similar. (We derive ε by substituting $x_i \geq \mu - c\sigma$, which holds except with probability $\delta_{left} = GaussCDF(\mu, \sigma^2; \mu - c\sigma)$.) Overall, using the union bound, $\delta = \delta_{left} + \delta_{right} = 2GaussCDF(\mu, \sigma^2; \mu - c\sigma)$. □

Privacy under many probes. Attackers may probe a channel many times; however, there are only $\log_b T$ ancestors in the tree for Alice's payment time slot. Thus, only those nodes might contribute different values to the relay's response (and, therefore, leak information about Alice's payment). We compose, in §6.2.2, the privacy guarantee from Theorem 2 over $\log_b T$ different observations that the attacker might get. This gives the differential privacy guarantee that a single channel provides (even if the adversary continuously probes it).

Privacy over a payment's route. The attacker may probe any inter-relay payment channel on Alice's payment route. Therefore, the level of privacy reduces with the number of inter-relay payment channels that it traverses. Although minimizing the number of hops benefits privacy, the costs that current PCN implementations opt to minimize might cause clients to prefer longer routes. For instance, in Lightnings' most common implementations, routing is not done by choosing the shortest-length route. In the C-Lightning implementation, the route's length is one of the parameters determining which route a payer chooses but in LND and Eclair (the other two most popular implementations), minimizing the route length is not explicit.

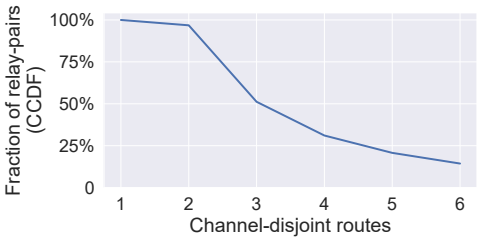

Figure 4: Shortest length channel disjoint routes between relays in the Lightning network topology (from Nov. 7th, 2021).

Usually, however, these costs decrease with the number of hops: having more hops typically results in less preferred routes, e.g., requiring clients to lock coins for longer time or increasing the total fee the sender would end up paying the relays. Indeed, measurements show that for all the implementations listed above, at least 80% of routes chosen by the routing algorithm in the Lightning Network (the largest PCN today) have no more than three inter-relay channels [42].

6.2.1 Privacy amplification by random route selection

In practice, multiple channel-disjoint routes of the shortest length often exist. For example, we connected a client to the Lightning Network and retrieved its topology (the snapshot was taken Nov. 7th, 2021). We found that there are usually 2 – 5 channel-disjoint routes of the shortest length between two relays (nodes with more than one channel); see Figure 4. Twilight leverages this insight about PCN topology to amplify users' privacy. Clients choose one of the shortest channel-disjoint routes uniformly at random. Intuitively, randomized route selection improves privacy since the attacker does not know which of the channels he probes are on the payment route. If Twilight is to be deployed on Lightning clients, it would need to relax the clients' path selection strategies (described above) to consider paths that are "close enough" to the minimal cost. This relaxation would accommodate their current route selection strategies and allow the client to choose from multiple routes to amplify privacy.

6.2.2 Payment privacy quantification

Theorem 3 captures the composition of the arguments made in this subsection and quantifies them in differential privacy terms. It shows that a random selection of 1-out-of-r possible routes of length l in Twilight improves the differential privacy's ε proportionally to $\sqrt{r}$, and that ε impairs proportionally to $\sqrt{\log_b T}$ and $\sqrt{l}$.

Theorem 3. *Let r be the number of channel-disjoint routes with l inter-relay channels between Alice and Bob. Alice's m-coin payment to Bob is (ε,δ) differentially-private where:*

$$\varepsilon = \ln\left(1+\frac{mc\sqrt{l\log_b T}}{\sigma\sqrt{r}}+\frac{l\log_b T m^2}{2\sigma^2}\right),$$

$$\delta = 2\cdot GaussCDF(\tilde{\mu},\tilde{\sigma}^2;\ \tilde{\mu}-c\tilde{\sigma})$$
$$And,\ \tilde{\mu}=rl\log_b T\mu,\ \tilde{\sigma}^2=rl\log_b T\sigma^2$$

Proof. Given in technical report [10] □

To get a fair degree of privacy, $\sigma >> m$ (the variance in the noise hides a payment that Alice might make). In this case, ε is very close to $\frac{mc\sqrt{l\log_b T}}{\sigma\sqrt{r}}$ (i.e., Theorem 3 generalizes Theorem 2; it gives a similar result when $T,r,l=1$).

6.2.3 Multiple payments

Until now, we analyzed differential privacy for one payment. However, if Alice makes multiple sensitive payments, the attacker may try to learn about any of them. This scenario is known as composition in the differential privacy literature. Fundamentally, differential privacy deteriorates with the number of payments Alice wishes to hide, but the composed result remains differentially private. Maintaining the rigorous differential privacy guarantee, albeit with higher ε,δ, can be crucial. For example, a court that should be convinced "beyond a reasonable doubt" requires a very high degree of certainty, making even relatively high ε,δ guarantee valuable.

The literature also provides theorems for computing the ε,δ guarantee of such composition. This quantification is important. It allows users to avoid sensitive payments when exceeding a "privacy budget" (i.e., when ε,δ reflect a risk they deem too high). The most general result states that composing k invocations of an (ε_i,δ_i)-differentially private mechanism results in $(\sum_{i=1}^{k}\varepsilon_i,\sum_{i=1}^{k}\delta_i)$-differential privacy [12, Thm. 3.16]. Another, more powerful, composition theorem [22] holds when the noise in the composed invocations is independent, this result states that ε grows with $\sqrt{k}$. In our case, when multiple payments traverse the same route, different invocations of the noisy payment processing mechanism within the same channel are not independent (due Twilight's use of the tree), and thus only guarantee linear growth of ε. However, when payments traverse disjoint routes, the more advanced composition theorem holds, and ε degrades slower. This implies that Alice can increase her privacy level under multiple payments by opening channels to more relays. Intuitively, these channels allow Alice to use more disjoint paths, which helps her in two ways: First, randomizing over more paths provides better ε,δ to begin with (see §7.1 for example ε,δ values for different route lengths and number of available disjoint routes). Second, when paths are disjoint, the stronger composition results hold. The online technical report analyzes this scenario [10].

7 The Cost of Privacy

The noise that a relay induces artificially reduces its channel's effective capacity: the relay might deny payments that spend amounts close to the channel's liquidity and lose the

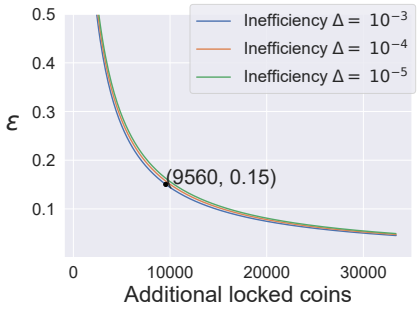

Figure 5: Privacy-efficiency trade-off for a channel with 2^{64} time-slots. Fixing $\delta = 10^{-7}$ and 1-coin payment. Locking more coins allows lower ε.

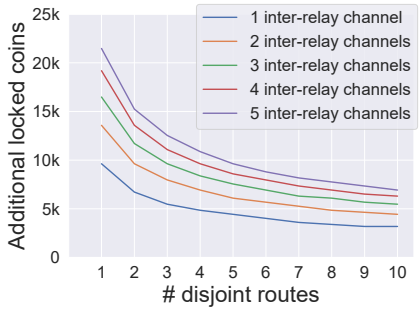

Figure 6: Extra locked coins needed to achieve $\varepsilon = 0.15, \delta = 10^{-7}$ differential privacy, with $\Delta = 10^{-3}$ inefficiency as a function of the number of disjoint routes.

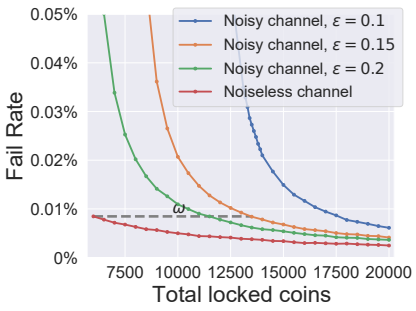

Figure 7: Payment failure rate on channels that provide different level of privacy.

associated fees. In this section, we study this cost. We present a metric for a channel's ability to process a sequence of payments, illustrate the trade-off between privacy and efficiency, and quantify the cost of operating a Twilight relay.

Definition 1. *A noisy payment channel $\mathcal{C}$ is Δ-inefficient compared to a noiseless channel $\mathcal{C}'$ if any payment accepted through channel $\mathcal{C}'$, is rejected in $\mathcal{C}$ with probability at most Δ.*

We use the inefficiency metric to measure the cost of noisy channels. Theorem 4 quantifies the extra coins that a relay should lock to make the noisy channel Δ-inefficient compared to the noiseless alternative. Let $Gauss(\mu, \sigma^2)$ be the noise distribution for nodes in a relay's noise tree, and N the maximal number of nodes in the covering set for any time slot (Table 1).

Theorem 4. *If the noisy channel $\mathcal{C}$ is initialized with ω more coins in both directions than the noiseless channel $\mathcal{C}'$, then it is Δ-inefficient compared to $\mathcal{C}'$, where:*
$\omega = \mu N + t$, *and* $\Delta = GaussCDF(\mu N, \sigma^2 N;\ \mu N - t)$.

The knob value t allows trading a higher ω for a lower Δ.

Proof. Given in technical report [10] □

The Gauss distribution CDF at $\mu N - t$ falls very quickly with t (proportionally to e^{-t^2}). For example, achieving $\Delta = 0.1\%$ (with $\delta = 10^{-7}, \varepsilon = 0.15$ privacy) requires $\omega = 95608$; see Figure 5. Achieving $\Delta = 10^{-10}$ requires locking only 15% more coins (for the same privacy level, $\omega = 11000$).

Multi-payment sequences and multi-hop routes. The inefficiency metric composes for a sequence of n payments routed over l channels: Consider two sequences of l noisy and noiseless channels, where noisy channel i is Δ_i-inefficient compared to the i^{th} channel in the noiseless sequence. Using the union bound, we find that a sequence of payments that is accepted via the noiseless route might be rejected from the noisy route with a probability of at most $n\sum_{i=1}^{l}\Delta_i$.

7.1 Privacy-efficiency trade-off

Comparing Theorem 3 and Theorem 4, which summarize Twilight's privacy and efficiency properties, allows reasoning about the number of coins that a relay should lock in the channel's smart contract to support a certain level of privacy. Figure 5 illustrates the privacy-efficiency trade-off for one channel, trading higher ω for lower ε. Figure 6 then focuses on a particular privacy setting ($\varepsilon = 0.15, \delta = 10^{-7}$) and explores the trade-off for scenarios where clients have multiple options for disjoint routes, and these routes are of multiple hops. Since clients boost privacy by uniformly selecting routes (§6.2.1), a network that offers many disjoint routes improves the efficiency (requires locking fewer coins for similar privacy level).

Payment success ratio. Our analysis thus far refers to the chance that *any* payment fails. However, in practice, a channel may process payments back and forth. Some payments may even fail for over-spending on a noiseless channel, while they would succeed on the noisy channel due to the extra liquidity (ω in Theorem 4) and sufficiently low noise. We use simulation to compare the failure rate on a long sequence of payments. Our simulation, in Figure 7, focuses on a single channel, where we send 1-coin payments left or right with a uniform distribution. We run this simulation for three levels of privacy (values of ε, fixing $\delta = 10^{-7}$). Each data point is the failure rate on a sequence of 10^8 payments. In this simulation, payments route over a single payment channel without concurrency. We observe that as the channel's capacity grows, the noised channel's success ratio converges to that of the noiseless channel. We also see that the extra locked coins needed for achieving $\Delta = 10^{-3}$ on the noisy channel with $\varepsilon = 0.15$ and $\delta = 10^{-7}$ is around $7k$ over the noiseless channel (see dashed horizontal line in Figure 7). This is an improvement over the theoretical bound from Theorem 4 (which is illustrated in Figure 5).

7.2 Quantifying the relay's cost and incentives

Twilight requires relays to run code in TEEs and induce noise when processing payments. Modern commodity processors ship with TEEs (e.g., see Intel's i9 processor spec [19]). Thus, we believe that deploying a TEE should not incur a high cost on the relay's operator and focus on the cost of operating noisy payment processing, which increases a relay's locked coins.

Although Twilight's relays need to lock more coins than the noiseless alternative (say, $9.5k$ additional coins per Figure 5), these coins *return to the relays* when the channel closes, so relay operators only lose any potential interest that they could have gotten for these extra coins. Moreover, the financial impairment rate from locking coins is fixed, and *amortizes over all payments* carried during the channel's lifetime. Given these observations, we can compute the increase in relay fee that would cover the cost of operating a Twilight channel depending on the number of payments it carries. For example, consider the privacy-efficiency trade-off point highlighted in Figure 5 ($\varepsilon = 0.15, \delta = 10^{-7}, \Delta = 10^{-3}$); we find that given a yearly interest rate of 3%,a relay that handles 79 payments per day can cover its operational cost by charging 1% fee from each payment . Previous works on differential privacy (in other contexts) statistically modeled users' actions, treating them as noise (e.g., [6]). This approach reduces the number of coins a relay should lock, thus reducing Twilight's operational cost. However, it requires strong additional assumptions about users that Twilight avoids: that many users are honest and submit payments according to a known distribution.

8 Implementation

There are three components to our prototype of Twilight. The first is the smart contract for managing on-chain channel accounts between two parties, which we implement for Ethereum in Solidty [14] (68 lines of code). A limitation of our smart contract implementation is that it does not use shielded on-chain transactions (e.g., implemented using Solidity in [25]). It, therefore, exposes a channel's closing balance split (i.e., the aggregate amount of payments on the channel through its lifetime). We argue in §3.1 that for long-lived channels this is typically safe in practice. The second component is the enclave running in Intel's SGX [9] as a TEE. We implement the noisy payment processing logic inside the enclave in C++17 (965 lines of code). The last component is the relay, which calls the enclave and implements the networking logic for carrying payments across the route. We implement the relay in Python3.8 (886 lines of code). The clients use the same networking logic as the relays (but do not run noisy payment processing). Our implementation uses ChaCha20 [27] for symmetric encryption, elliptic curve `secp256r1` [39] for public key operations, and SHA3-Keccak [5] as hash function.

Inside the enclave, we generate the TEE's secret key using `sgx_ecc256_create_key_pair` and run the randomized response mechanism (§4.1). Running this mechanism requires the enclave to read the current time; we call `rdtsc` to get the number of CPU cycles since boot [20] which serves as a high-resolution clock. We detect reboots by reading the nonce that `sgx_get_trusted_time` returns, which is initialized at boot time (as discussed in §5.2). The channel's noise tree (Figure 2) has $T = 2^{64}$ time slots to support long-lived channels.

9 Evaluation

We use Twilight's prototype to evaluate the throughput and latency, and to measure the cost of resolving disputes on-chain (§9.1). We use simulations to evaluate the privacy benefits of noisy payment processing under partial adoption (§9.2).

9.1 Performance and cost

Setup. We deploy Twilight's implementation over a `Standard_DC1s_v2` machine type in Azure [8], which has one CPU and 4 GB RAM. This machine has a single `Intel SGX-1` TEE. We deploy Twilight on machines in two Azure regions, across two continents, `eastus` and `northeurope`. To ensure our tests experiment real Internet latencies, the routes for the payments we evaluate alternate between relays in both regions. We measure the average round-trip latency between machines in the two regions to be 84.89ms (with 0.58 standard deviation). In the following experiment results, each data point is an average of 40 repetitions. We use error bars to show the standard deviation from the mean.

Throughput. In Figure 8 we measure the rate of completed payments as a function of the rate of issued payments for different route lengths. The throughput continues to grow until around 820 resolved payments/sec, which is over twice of a relay's throughput in the Lightning PCN [21] (about 358 payments/sec with one relay). We attribute this performance improvement to batch-processing payments (cf., in Lightning, each payment requires expensive invocations of an underlying Bitcoin client). When the payment issuance rate exceeds 800 payments/sec, the relays' backlog grows and eventually causes the throughput to start decreasing. We, therefore, cap the relay's backlog at 3000 payments to allow handling bursty payments and avoid congestion collapse. Since PCNs are horizontally scalable, the more relays join Twilight, the more disjoint routes are available (proportionally growing Twilight's throughput).

Latency across a route. Figure 9 evaluates payment latency in Twilight by sending payments across paths of different lengths and using different payment issuance rates. After a minute, when the system is in steady-state, we measure the latency for completing payments. Before backlogs start forming in the relays, the latency is under 1.1s even with 4 relays (e.g., when 800/sec payments are issued). We attribute about half of this latency to the network RTT time (510ms from Alice to Bob via the 4 relays). The rest is due to Twilight's

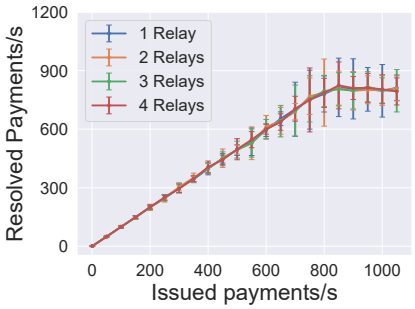

Figure 8: Throughput by payment issuance rate for different route lengths.

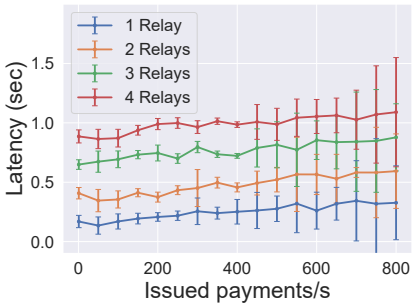

Figure 9: Latency by payment issuance rate for different route lengths.

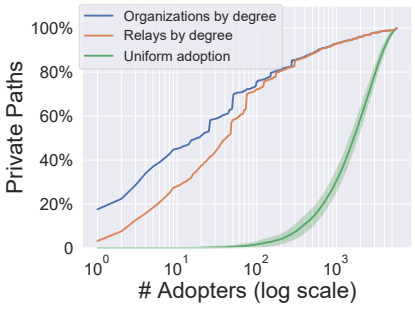

Figure 10: Privacy gain under partial noisy payment processing adoption.

processing costs inside the TEE. From this, we conclude that Twilight does not substantially increase the latency over a vanilla PCN (that routes payments via a similar number of hops). Even when payments are issued at a relatively high rate (e.g., 800 payments/sec) Twilight's processing does not dominate the latency over the network RTT.

Smart contract gas cost. We use Ganache [15] to evaluate the gas price of running the smart contract managing the channel's account on Ethereum. We gather the gas prices using "Web3.estimateGas" (an Ethereum API). The total gas cost for handling each disputed PTLC is less than 2.4k gas. The major pieces comprising this cost are: (1) 700 gas to validity check of the dispute, (2) 800 gas to decrypt the disputed PTLC, and (3) 900 gas to parse the plaintext and update the balance.

Each PTLC is encoded into 28B, which users can post to the smart contract to dispute the channel's closing balance split (along with the partner's signature over that PTLC and a matching secret key). This encoding comprises 8B for the amount, 4B for the PTLC's timeout (in blocks), and 16B for the authentication code that ensures the secret key is correct. Storing the PTLC on the blockchain costs 15k gas and dominates its processing cost.

9.2 Partial noisy processing adoption

Performing noisy payment processing involves costs for the relay (§7). We consider the effect of weakening Twilight's requirements through a compatibility mode where relays can participate in the network without noising payments in TEEs. Clients can tell which relays do noisy payment processing in TEEs using remote attestation and prefer routes where all relays do so. We evaluate the fraction of payer-payee pairs of nodes with a fully adopting path. As an example network, we use the Lightning network topology (snapshot from Nov. 7^{th}, 2021), the largest PCN today. We that assume the payer and payee run Twilight's client, and evaluate three scenarios regarding Twilight's adoption (in its full form) in the network: (1) relays adopt by descending order in the number of channels they have, (2) organizations (companies operating relays) adopt across all their relays by descending order in the number of channels they have (we use relay lists from [7,30]), and (3) uniform random adoption across the network.

Figure 10 shows that if the largest 5 organizations in the network adopt noisy payment processing, then 47% of pairs of nodes in the network will have a connecting route with full adoption. The largest organization, "LN-BIG" [30] alone connects around 19% of the pairs. More generally, we see that adoption on a small number of relays or organizations at the core of the network can protect routes between a large portion of the network's nodes, giving a tangible path to significant privacy improvement. Uniform adoption is less effective, and several hundreds of relays should adopt before the benefit becomes significant.

10 Related Work

Several works point out that today's PCNs do not offer much privacy [16, 18, 31], and Kappos et al. evaluate such attacks in practice [23]. Bolt [16] gives a strong privacy guarantee by establishing payment channels over ZCash, but its architecture is restricted to just one relay. Namely, a hub that connects everyone in the network, which limits the scalability of the design and risks availability in the case that hub goes offline.

Malavolta et al. [31, 32] propose a privacy-preserving HTLC for PCNs. Instead of using the same HTLC secret for every hop, secrets across the route are cryptographically linked but appear random. Thus, malicious relays along the route cannot link payments through their HTLC secrets. This construction is compatible with Twilight's PTLCs and allows avoiding such "secret-correlation" attacks. SpeedyMurmurs [38] modifies the client's routing algorithm to split payments across several paths. In this manner, an attacker that does not control a relay on all routes cannot learn the exact payment amount or uniquely identify the payer and payee, but he can learn information about the users' "direction." As discussed in [32], these solutions [31, 32, 38] are only partial to the privacy problem in PCNs. In particular, since channels cannot transfer more funds than their liquidity, the attacker can measure channels' liquidity on inter-relay links by request-

ing relays to carry payments. He can then correlate liquidity changes across channels to track the users' payments [23], the problem that Twilight tackles. Quantifying and bounding statistical information leakage to an active adversary that probes channels to deduce payment routes is challenging, which Twilight achieves through differential privacy.

Joancomarti et al. [18] show how malicious clients can monitor changes in channel liquidity. Tang et al. consider a PCN that continuously advertises channel liquidity with fresh noise [40]. They do not specify a noise mechanism or means to enforce it but show that the adversary can quickly learn the liquidity on every channel. Twilight addresses this problem by utilizing ideas from the differential privacy literature; it can avoid continuously publishing channel-liquidity with fresh noise since it is safe to reuse noise values when hiding the same payments set. This allows Twilight to provide a rigorous differential privacy guarantee for its users.

TEEchain uses TEEs to build a high-throughput PCN [29] and relies on the security of its relays' TEEs for integrity. In particular, there are no disputes since TEEs are trusted to close channels at the correct balance. Moreover, TEEchain does not protect against attacks on its users' privacy. In contrast, Twilight is focused on privacy and uses TEEs only to ensure relays perform randomized response to hide users' payments.

11 Conclusion

We presented Twilight, a new PCN that is focused on privacy. Twilight hides a user's payments from other users in the network using differential privacy. Relays convince users that they will hide their payments by leveraging TEEs. Our analysis shows that Twilight provides rigorous privacy and incurs moderate costs. We implemented Twilight and tested its performance across a route of relays in two continents and evaluated it under partial adoption using simulations. Our evaluation shows that it provides solid performance compared to Lightning, today's most popular PCN (providing no privacy), and gives a tangible path to payment privacy in PCNs.

Availability

Our code is available online along with instructions for reproducing the evaluation results, see link in [41].

Acknowledgments

We thank Adam D. Smith and Katrina Ligett for insightful discussions on differential privacy and its applications, and our shepherd Stefanie Roos. Yossi Gilad was partially supported by the Alon fellowship and the Hebrew University cyber security research center and a gift from Microsoft. Aviv Zohar, Maya Dotan, and Saar Tochner were partially supported by grants from the Israel Science Foundation (grants 1504/17 & 1443/21) and by a grant from the Hebrew University cyber security research center.

References

[1] 1ML. Lightning network statistics. https://1ml.com/statistics, 2020.

[2] Adil Ahmad, Kyungtae Kim, Muhammad Ihsanulhaq Sarfaraz, and Byoungyoung Lee. Obliviate: A data oblivious filesystem for intel sgx. In *NDSS*, 2018.

[3] Eli Ben-Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from Bitcoin. In *IEEE Symposium on Security and Privacy*, pages 459–474. IEEE Computer Society, 2014.

[4] Iddo Bentov and Ranjit Kumaresan. How to use Bitcoin to design fair protocols. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology - CRYPTO*, volume 8617 of *Lecture Notes in Computer Science*, pages 421–439. Springer, 2014.

[5] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. The keccak sha-3 submission. *NIST*, 6(7):16, 2011.

[6] Raghav Bhaskar, Abhishek Bhowmick, Vipul Goyal, Srivatsan Laxman, and Abhradeep Thakurta. Noiseless database privacy. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT*, volume 7073 of *LNCS*, pages 215–232. Springer, 2011.

[7] Bitfinex. The nodes of bitfinex. https://ln.bitfinex.com/, 2022.

[8] Marshall Copeland, Julian Soh, Anthony Puca, Mike Manning, and David Gollob. Microsoft azure. *Apress: New York, NY, USA*, 2015.

[9] Victor Costan and Srinivas Devadas. Intel SGX explained. Report 2016/086, Cryptology ePrint Archive, February 2016.

[10] Maya Dotan, Saar Tochner, Aviv Zohar, and Yossi Gilad. Twilight: A differentially private payment channel network. Cryptology ePrint Archive, Report 2022/136, 2022. https://ia.cr/2022/136.

[11] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N. Rothblum. Differential privacy under continual observation. In Leonard J. Schulman, editor, *STOC*, pages 715–724. ACM, 2010.

[12] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[13] Shufan Fei, Zheng Yan, Wenxiu Ding, and Haomeng Xie. Security vulnerabilities of sgx and countermeasures: A survey. *ACM Computing Surveys*, 54(6):1–36, 2021.

[14] Ethereum Foundation. Solidity programming language. https://docs.soliditylang.org/en/latest/.

[15] Ganache. https://www.trufflesuite.com/docs/ganache/overview.

[16] Matthew Green and Ian Miers. Bolt: Anonymous payment channels for decentralized currencies. In Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *CCS*, pages 473–489. ACM, 2017.

[17] Greg Brockman. Stellar. https://stripe.com/blog/stellar, 2014.

[18] Jordi Herrera-Joancomartí, Guillermo Navarro-Arribas, Alejandro Ranchal Pedrosa, Cristina Pérez-Solà, and Joaquin Garcia-Alfaro. On the difficulty of hiding the balance of lightning network channels. In Steven D. Galbraith, Giovanni Russello, Willy Susilo, Dieter Gollmann, Engin Kirda, and Zhenkai Liang, editors, *AsiaCCS*, pages 602–612. ACM, 2019.

[19] Intel product specifications. https://ark.intel.com/content/www/us/en/ark/products/199332/intel-core-i910900k-processor-20m-cache-up-to-5-30-ghz.html. Accessed: 2021-12-21.

[20] Intel. SGX software developer's manual. https://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-software-developer-vol-3d-part-4-manual.pdf.

[21] Joost Jager. Lightning node performance: Exploring the path to 1 000 tps. https://bottlepay.com/blog/bitcoin-lightning-benchmarking-performance/.

[22] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *International conference on machine learning*, pages 1376–1385. PMLR, 2015.

[23] George Kappos, Haaroon Yousaf, Ania Piotrowska, Sanket Kanjalkar, Sergi Delgado-Segura, Andrew Miller, and Sarah Meiklejohn. An empirical analysis of privacy in the lightning network. In *Financial Cryptography*, 2021.

[24] George Kappos, Haaroon Yousaf, Ania Piotrowska, Sanket Kanjalkar, Sergi Delgado-Segura, Andrew Miller, and Sarah Meiklejohn. An empirical analysis of privacy in the lightning network. In *International Conference on Financial Cryptography and Data Security*, pages 167–186. Springer, 2021.

[25] Chaitanya Konda, Michael Connor, Duncan Westland, Quentin Drouot, and Paul Brody. Nightfall protocols for private transactions on the ethereum blockchain using zk-snarks. https://github.com/EYBlockchain/nightfall.

[26] Satwik Prabhu Kumble, Dick Epema, and Stefanie Roos. How lightning's routing diminishes its anonymity. In *The 16th International Conference on Availability, Reliability and Security*, pages 1–10, 2021.

[27] Adam Langley, W Chang, Nikos Mavrogiannopoulos, Joachim Strombergson, and Simon Josefsson. Chacha20-poly1305 cipher suites for transport layer security (tls). *RFC 7905*, 2016.

[28] Ninghui Li, Min Lyu, Dong Su, and Weining Yang. *Differential Privacy: From Theory to Practice*. Synthesis Lectures on Information Security, Privacy, & Trust. Morgan & Claypool Publishers, 2016.

[29] Joshua Lind, Oded Naor, Ittay Eyal, Florian Kelbert, Emin Gün Sirer, and Peter R. Pietzuch. Teechain: A secure payment network with asynchronous blockchain access. In Tim Brecht and Carey Williamson, editors, *SOSP*, pages 63–79. ACM, 2019.

[30] LN-BIG. The nodes of ln-big. https://lnbig.com/#/our-nodes, 2022.

[31] Giulio Malavolta, Pedro Moreno-Sanchez, Aniket Kate, Matteo Maffei, and Srivatsan Ravi. Concurrency and privacy with payment-channel networks. In Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *CCS*, pages 455–471. ACM, 2017.

[32] Giulio Malavolta, Pedro Moreno-Sanchez, Clara Schneidewind, Aniket Kate, and Matteo Maffei. Anonymous multi-hop locks for blockchain scalability and interoperability. In *NDSS*. The Internet Society, 2019.

[33] Ania M. Piotrowska, Jamie Hayes, Tariq Elahi, Sebastian Meiser, and George Danezis. The loopix anonymity system. In Engin Kirda and Thomas Ristenpart, editors, *26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017*, pages 1199–1216. USENIX Association, 2017.

[34] Wahbeh H. Qardaji, Weining Yang, and Ninghui Li. Understanding hierarchical methods for differentially private histograms. *Proc. VLDB Endow*, 6(14):1954–1965, 2013.

[35] Fergal Reid and Martin Harrigan. An analysis of anonymity in the Bitcoin system. In *SocialCom/PASSAT*, pages 1318–1326. IEEE Computer Society, 2011.

[36] Torkel Rogstad. Lightning network 101: Privacy. Medium post, 2019.

[37] Dorit Ron and Adi Shamir. Quantitative analysis of the full Bitcoin transaction graph. In Ahmad-Reza Sadeghi, editor, *Financial Cryptography*, volume 7859 of *Lecture Notes in Computer Science*, pages 6–24. Springer, 2013.

[38] Stefanie Roos, Pedro Moreno-Sanchez, Aniket Kate, and Ian Goldberg. Settling payments fast and private: Efficient decentralized routing for path-based transactions. In *NDSS*. The Internet Society, 2018.

[39] secp256r1. https://neuromancer.sk/std/secg/secp256r1.

[40] Weizhao Tang, Weina Wang, Giulia C. Fanti, and Sewoong Oh. Privacy-utility tradeoffs in routing cryptocurrency over payment channel networks. *Measurement and Analysis of Computing Systems*, 4(2):29:1–29:39, 2020.

[41] Saar Tochner, Maya Dotan, Aviv Zohar, and Yossi Gilad. Twilight prototype implementation. https://github.com/saart/Twilight, 2022.

[42] Saar Tochner, Aviv Zohar, and Stefan Schmid. Route hijacking and DoS in off-chain networks. In *AFT*, pages 228–240. ACM, 2020.

[43] Jiyong Yu, Lucas Hsiung, Mohamad El Hajj, and Christopher W. Fletcher. Data oblivious isa extensions for side channel-resistant and high performance computing. Cryptology ePrint Archive, Report 2018/808, 2018. https://ia.cr/2018/808.