GearBox: Optimal-size Shard Committees by Leveraging the Safety-Liveness Dichotomy

Bernardo David
ITU
Copenhagen, Denmark
bernardo@bmdavid.com

Bernardo Magri
The University of Manchester
Manchester, UK
bernardo.magri@manchester.ac.uk

Christian Matt
Concordium
Zurich, Switzerland
cm@concordium.com

Jesper Buus Nielsen
Concordium Blockchain Research Center, Aarhus University
Aarhus, Denmark
jbn@cs.au.dk

Daniel Tschudi
Concordium
Zurich, Switzerland
dt@concordium.com

ABSTRACT

Sharding is an emerging technique to overcome scalability issues on blockchain based public ledgers. Without sharding, every node in the network has to listen to and process all ledger protocol messages. The basic idea of sharding is to parallelize the ledger protocol: the nodes are divided into smaller subsets that each take care of a fraction of the original load by executing lighter instances of the ledger protocol, also called shards. The smaller the shards, the higher the efficiency, as by increasing parallelism there is less overhead in the shard consensus.

In this vein, we propose a novel approach that leverages the sharding safety-liveness dichotomy. We separate the liveness and safety in shard consensus, allowing us to dynamically tune shard parameters to achieve essentially optimal efficiency for the current corruption ratio of the system. We start by sampling a relatively small shard (possibly with a small honesty ratio), and we carefully trade-off safety for liveness in the consensus mechanism to tolerate small honesty without losing safety. However, for a shard to be live, a higher honesty ratio is required in the worst case. To detect liveness failures, we use a so-called control chain that is always live and safe. Shards that are detected to be not live are resampled with increased shard size and liveness tolerance until they are live, ensuring that all shards are always safe and run with optimal efficiency. As a concrete example, considering a population of 10K parties with at most 30% corruption and 60-bit security, previous designs required over 5800 parties in each shard to guarantee security. Our design requires only 1713 parties in the worst case with maximal corruption, and in the optimistic case works with only 35 parties without compromising security.

Moreover, in this highly concurrent execution setting, it is paramount to guarantee that both the sharded ledger protocol and its sub protocols (i.e., the shards) are secure under composition. To prove the security of our approach, we present ideal functionalities capturing a sharded ledger as well as ideal functionalities capturing the control chain and individual shard consensus, which needs adjustable liveness. We further formalize our protocols and prove that they securely realize the sharded ledger functionality in the UC framework.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CCS '22, November 7–11, 2022, Los Angeles, CA, USA

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9450-5/22/11…$15.00
https://doi.org/10.1145/3548606.3559375

CCS CONCEPTS

• **Theory of computation → Cryptographic protocols**.

KEYWORDS

blockchain, sharding

ACM Reference Format:
Bernardo David, Bernardo Magri, Christian Matt, Jesper Buus Nielsen, and Daniel Tschudi. 2022. GearBox: Optimal-size Shard Committees by Leveraging the Safety-Liveness Dichotomy. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22), November 7–11, 2022, Los Angeles, CA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3548606.3559375

1 INTRODUCTION

Since the introduction of Bitcoin [25], there has been an explosion of interest in blockchains, both in research and practice. One of the biggest practical obstacles for the large-scale adoption of public blockchain systems is the low throughput of transactions in systems such as Bitcoin. As a solution to overcome this limited scalability, a method called *sharding* has been proposed.

The basic idea of sharding is to parallelize the execution by dividing the network into smaller components, called shards. The smaller the shards are the higher the efficiency is, due to increasing parallelism and less overhead in the shard consensus. However, the security of the shards requires its size to be big since small shards have lower security. For example, assuming at most 30% corruption[1] overall with a total of 10K parties, the minimal shard size that guarantees at most 33% corruption with 60-bit security[2] in a randomly selected shard is 5886. As we show in the full version,

[1] In blockchains, the corruption bounds are typically weighted by some resource, e.g., computing power for proof-of-work systems, or stake in proof-of-stake systems. To simplify the presentation, we ignore this weighting in the introduction. We stress, however, that our results are not limited to this simplified setting and can indeed be used in a weighted setting. See Section 4.1 for a further discussion.

[2] A security level of 60-bit means that security holds with probability at least $1 - 2^{-60}$.

the bounds are almost perfectly linear in the security parameter, so for 30-bit security the size would have to be about 3000. Thus, to get to a shard size in the hundreds, unsatisfying security levels would have to be adopted.

The security of a blockchain system consists of two main properties: (1) *Liveness* says that the blockchain will eventually output new messages to all peers, and (2) *Safety* says that the peers agree on the sequence of messages being output. The liveness threshold L and the safety threshold S are the levels of corruption under which liveness and safety are guaranteed, respectively. Existing sharding solutions [21, 23, 34] need to guarantee security of the shards (both liveness and safety) in the worst-case corruption scenario, thus forcing them to consider equal bounds for liveness and safety, severely constraining the size of shards. In this work, we overcome this apparent barrier by leveraging the *shard safety-liveness dichotomy* and choosing different safety and liveness bounds. We devise a protocol where shards are always safe but can eventually lose liveness; when that happens the shards are respawned with adjusted liveness and safety parameters. This allows for significantly smaller shards.

The shard safety-liveness dichotomy defines the possible tuples (L, S) of liveness and safety thresholds for which a (shard) consensus protocol provides security. For partially synchronous protocols, the safety-liveness dichotomy says that $2L + S < 100\%$. Hence in the case of $S = L$ it must be that $L, S < 33\%$. For synchronous protocols the safety-liveness dichotomy says that $L + S < 100\%$, hence whenever $S = L$ it must be that $L, S < 50\%$. These dichotomies can be derived from the following observations. For the synchronous dichotomy, we use the crucial fact that external parties can post on and read from the shards. We observe that if a reader can be convinced about the state of a shard while a subset of (relative) size L is crashed, then it means that a subset of size $1 - L$ can convince the reader. Thus if $S = 1 - L$, this means that a reader can be convinced after only talking to a set of potentially corrupt servers. For the partially synchronous dichotomy we use the fact that if the unknown network delay is large enough, a reader cannot distinguish a corrupt subset of (relative) size L from a slow and honest subset. Moreover, if a subset of size L is crashed, a reader must by definition be able to make a decision. Hence, a reader that makes a decision (say, on which message it saw on the shard first) should be able do this without having heard from a fraction L of the honest parties. This means that from the fraction $L + S$ of the parties it heard, a fraction L or a fraction S might be corrupted. The shard dichotomies discussed here follow from standard arguments, and for completeness we provide formal proofs in the full version.

1.1 Our Contributions

We present a novel sharding approach that leverages the safety-liveness dichotomy to get the smallest possible shard committees without sacrificing safety. Our sharding design has security against a fraction of $t < 1/3$ corrupt shard committee members in the partially synchronous setting.[3]

The shards will, in the optimistic model, start to run with a low liveness threshold and a high safety threshold, e.g., $S = 89\%$ and $L = 5\%$. Being safe against up to 89% corruption allows for sampling much smaller committees, however being live only against up to 5% corruption makes it a lot more likely for a shard to deadlock. For this, we use an approach where an independent ledger, that we call control-chain (CC), manages the shards by constantly monitoring them for liveness. This is done by letting the shards post "heart beat" transactions on the CC. The CC can then "take down" a deadlocked shard and spin up a new shard with a new random committee and a higher liveness threshold (and a lower safety threshold), leading to bigger shards. This can be iterated until a shard is found which gives liveness (and safety). On the other hand, if no deadlocks are detected within a certain period, shards can be dynamically scaled down, leading to optimal shard sizes for the current corruption ratio. Crucially, at no point safety is compromised.

Our design has for each shard what we call a "gearbox" of consensus protocols: shards at the top are larger (therefore slower) but have robust liveness, while shards at the bottom are small (therefore faster) but have a lower liveness. The CC changes gear upwards in the gearbox when deadlocks are detected switching to a larger shard, and can over time change gear downwards when there is no signs of an attack. At the top of the gearbox the gear cannot change upwards, so a deadlocked shard is just restarted with the same parameters and a new random committee. The only requirement to get eventual liveness is that, when the top consensus algorithm is instantiated with a random committee, it happens with constant probability that the corruption threshold is low enough to get liveness. This approach allows us to select the best committee size given the *unknown* actual corruption threshold, albeit at the cost of resampling the committee until liveness is achieved.

Next, we describe two ways to instantiate our framework.

Partially synchronous. One can run with a partially synchronous CC and partially synchronous shards. At the bottom gear one could have $L = 0\%$ and $S = 99\%$. For a population of 10K peers and assuming an overall corruption level of at most 30%, this would give a committee size of 35 guaranteeing that safety is not violated except with probability 2^{-60}. At the top gear one could use a consensus protocol with $L = 30\%$ and $S = 39\%$. This would give a committee size of 1713 guaranteeing not more than 39% corruption (with 60-bit security). Note that we sample from a ground population with corruption at most 30% and need a committee with corruption at most 30% for liveness in the top gear. It is easy to see that we get at most 30% corruption with a constant probability, which gives eventual liveness. This already gives a significant improvement over existing designs that require that liveness only fails with negligible probability. Moreover, 30% corruption is a worst-case assumption and in typical executions, there will be much less corruption. In the other extreme with 0 corruption, the shards will already be live in the lowest gear with 35 parties. Even for more realistic 20% actual corruption, 207 parties per shard are sufficient to obtain a live shard with constant probability, dramatically improving the state of the art. See Section 4.3 for more details on required committee sizes.

Mixed. One can run a synchronous CC tolerating 49% corruption, and at the bottom of the gearbox we again start with a partially synchronous shard with $L = 0\%$ and $S = 99\%$. We run partially synchronous up until $L = 25\%$ and $S = 49\%$. After that, we then switch to a synchronous shard with $S = L = 49\%$ corruption.

[3] Although our techniques can be used to get security against a fraction of $t < 1/2$ corrupt committee members in the synchronous setting, we focus on the more desirable (and challenging) partially synchronous case.

This allows a design tolerating 49% overall corruption, but running partially synchronous small shards until 25% corruption. This is interesting since partially synchronous protocols can achieve higher throughput in good network conditions by avoiding waiting for the end of rounds, as synchronous protocols do.

In the rest of the paper we focus on the partially synchronous setting, and therefore we stick with the $2L + S < 100\%$ dichotomy. Thus we need less than 33% corruption in the ground population to get safety and liveness of the CC.

Static vs. adaptive security. Protocols that rely on the honest majority of long living committees for security clearly fail if an adaptive adversary can instantly corrupt half of the committee. This is also true for our sharding protocol of Section 5. We point out that this is not a weakness of our design but to the best of our knowledge applies to all sharding protocols based on randomly chosen committees. While Free2Shard [27] tolerates adaptive corruptions, it assumes a different security model, and thus cannot directly be compared to our protocol. If we assume the adversary can adaptively corrupt parties, but only after some delay [24], one can obtain security by periodically resampling committees, see Section 5.3.

Cross-Shard communication. Cross-shard transactions have extensively been studied in the literature [1, 30, 35]. For our sharding approach, special care needs to be taken since shards could lose liveness during a cross-shard transaction. We show in Section 5.4 on the example of Atomix [21], how existing protocols can be adapted to work in our setting. The basic idea is to leverage the control chain to finalize messages on the shards. This can be done without an additional overhead by including Merkle tree hashes of the relevant transactions in the heartbeats that are regularly posted on the control chain. This ensures that the amount of data posted on the control chain is independent of the number of cross-shard transactions, and short Merkle proofs can be used to prove finality of transactions on shards that holds even under restarting shards.

UC formalization. To prove the security of our approach, we formalize an ideal functionality capturing a sharded ledger as well as functionalities capturing the consensus guarantees we require from the control chain and from the shards, which need to have adjustable liveness in our approach. We build on these functionalities to construct our sharded ledger protocol, which we prove to UC-realize the sharded ledger functionality. To the best of our knowledge, ours is the first sharded ledger protocol to achieve security under arbitrary composition, which is an extremely important property in settings where a number of protocols are executed in parallel (e.g., blockchains). Moreover, we introduce and model the concept of timed ledgers, which go beyond guaranteeing that messages recorded on the ledger remain ordered in a certain way, also allowing parties to obtain explicit timestamps for messages.

1.2 Technical Overview

The main building blocks of our sharding protocol are a Control Chain and Shard ledgers, which we discuss below and describe in detail in Section 3.

Control chain. The control chain (CC) is used to orchestrate the sharding protocol, storing shard management metadata. The CC is modeled as a timed ledger functionality that orders and timestamps incoming messages. More formally, the control-chain is a totally-ordered broadcast with persistency and a timestamp guarantee. The CC is executed by all parties but it only stores metadata of size *independent* from the contents of the shards. We show in Section 6.1 how one can realize a control chain using different existing blockchain protocols.

Shards. Shards are again modelled as a ledger functionality parameterized by the size of the shard committee and an adversary structure for the liveness guarantee. This allows us to instantiate shards with committees of different sizes and with different liveness guarantees. A shard consensus protocol is only executed by a small shard committee. As we propose a general approach to sharding, our shard functionality can be readily instantiated by standard blockchain or permissioned consensus protocols. In Section 4.2, we show how to leverage specific properties of our approach to achieve a particularly efficient instantiation of the shard functionality using a simple BFT-style consensus protocol with a leader that proposes blocks. Our protocol does not guarantee liveness if the leader is corrupted as we can always recover using the control-chain. This comes at the price of possibly having to resample committees more often and is in contrast to existing protocols, that come with complex mechanisms to resolve an unresponsive leader. To minimize resamplings, one can also use existing protocols that include leader replacement, as discussed in Section 4.2.

The protocol. The goal of our sharding protocol in Section 5 is to achieve a sharded ledger. Essentially, the sharded ledger consists of multiple copies of the ledger type used as control-chain. Our protocol starts by initializing all shards with the smallest possible committees, i.e., using the smallest liveness threshold. The initialization (and later restarting) of shards is done by posting a command on the CC. Once the shard committee responsible for that shard sees the command on the CC, it starts executing that shard.

At the core of our sharding protocols are the heartbeats that allow to asses the liveness of a shard. They are implemented by asking the shards to periodically post hashes of blocks on the CC as a way to timestamp and "finalize" the blocks. A block is considered valid if it is accompanied by a proof that at least one honest party agrees with it (e.g., signatures on the block by a sufficiently large number of parties in the shard committee). Each time a valid block is posted, a timeout is set. If the next valid block does not arrive before the timeout, the shard is considered deadlocked. The timestamp is used to uniquely determine whether a shard made the time out.

If a shard is considered deadlocked, it is restarted with a bigger committee guaranteeing a larger liveness threshold, thus increasing the chance that it will have liveness. This way, the protocol ensures that a committee of close to optimal size is eventually selected for each shard. In other words, the size of shards is dynamically adjusted to match the actual corruption ratio in the network.

Note that such a sharded ledger can be trivially achieved (in UC) by simply instantiating multiple copies of the control-chain. We therefore emphasize that our sharding protocol is designed for the sake of efficiency. In the trivial solution every party needs to participate in each ledger copy, while our solution allows to select small committees that maintain a shard each. The selection of committees is discussed in Section 4.1.

1.3 Related Work

In the last few years, many shard-based blockchain protocols have been proposed by the scientific community and by the industry in the form of whitepapers. Most of the proposals by the industry, despite many containing nice ideas and innovations, follow an heuristic approach, where no formal security guarantees are proposed or formally proven. Thus, in this section we only discuss a few of the most well-known (peer-reviewed) sharding protocol proposals and refer the interested reader to the survey of Wang et al. [30] that gives a nice overview of the state-of-the-art in sharding protocols. Finally, we point out a common issue with all the proposals that hinders their practical usage.

There exists other approaches, such as Prism [4] and OHIE [33], to overcome blockchain scalability problems. While these two are orthogonal approaches and limited to the proof-of-work setting, our work focuses on more generic sharding solutions. For a general overview of blockchain scalability solutions, see [28].

Sharding protocols. To the best of our knowledge, Elastico [23] is the first sharding protocol proposed for public blockchains. The protocol is synchronous and runs in "epochs"; in every epoch each party solves a PoW puzzle based on randomness obtained from the previous epoch. The PoW's least-significant bits are used to form the committees that will run each shard and process the transactions. Even though the authors of [23] advocate for a small committee size per shard (around 100 parties), the probability of a shard being unsafe gets very high, close to 97%, after only six epochs, as shown in [21]. This renders the protocol completely insecure when used with small committees.

Building upon Elastico's ideas, and improving it in many ways, OmniLedger [21] is a sharding protocol that generates identities and assigns participants to shard committees using a synchronous PoW independent identity-blockchain. However, like Elastico, OmniLedger can only efficiently[4] tolerate up to $t < n/4$ corruptions on the total number of parties in the system. During each epoch, new randomness is generated for a leader election lottery. The protocol can achieve low latency for the confirmation of transactions whenever $t < n/8$.

RapidChain [34] is a synchronous sharding protocol that tolerates up to $n/3$ corrupt parties out of the total number of participants. The protocol is bootstrapped by a committee election protocol that initially selects a reference committee of size $m = O(\log n)$. At the end of every epoch, the reference committee is responsible for generating fresh randomness that will be used to select the committees for all the shards at the end of the *first* epoch, and to reconfigure the committees of existing shards in subsequent epochs.

Using an approach closely related to sharding, Monoxide [31] proposes a scale-out blockchain that contains many independent chains (called zones) running in parallel that divides the work-load of the entire system; communication, computation and storage is shared among the different zones, making the burden of maintaining the entire system shared among the nodes running each zone. When a "cross-zone" transaction happens, an eventual atomicity technique is used in order to keep consistency among the different zones.

[4]It can handle up to 33%, but with bad performance. See Footnote 2 in [21].

The work of Avarikioti et al. [1] proposes a framework with security properties tailored for sharded ledger protocols, building upon the Bitcoin backbone model of Garay et al. [15]. More specifically, the authors propose the novel notions of *consistency* and *scalability* for sharded ledgers that intuitively says that, cross-shard transactions must preserve safety and sharded systems must gain some speed-up in comparison to a non-sharded system, respectively. Moreover, the authors analyze many existing sharded ledger protocols in their model and prove if the protocol satisfy the proposed definition or not. Unfortunately, the model proposed in [1] is not composable, making it difficult to argue security of a sharded ledger protocol when combining it with a larger system.

Common issues. A common factor in all the previously described sharding protocols is that, for a robust security parameter, the size of the shard's committee needs to be large in order to guarantee the safety properties for each shard. In Section 4.3 we present some concrete numbers for the smallest size of committees needed to guarantee different honesty levels considering 60-bits of security. For example, with a total population of 10000 parties and at most 30% overall corruption, an honest supermajority (33% corruption) can only be guaranteed with 5886 parties per committee. Considering only 2000 parties in the total population requires 1716 parties per committee. And even reducing the assumption on the overall corruption to only 20% with 2000 parties in the total population still requires 540 parties in every shard.

Moreover, none of these previous works consider (or prove) security of sharding protocols under composability. This is a major shortcoming since sharding protocols (and blockchains in general) are mostly used as building blocks of larger systems (e.g., cryptocurrencies and smart contracts), which requires them to retain their security under composability.

2 PRELIMINARIES

We denote by P a party in the party set $\mathcal{P}$. We denote by $\mathsf{Honest} \subseteq \mathcal{P}$ the set of honest parties during the protocol execution. We denote by $H\colon \{0,1\}^* \to \{0,1\}^\kappa$ a collision-resistant hash function.

2.1 Security Model

Since our protocols make essential use of time, we need a notion of UC security for (partial) synchronous protocols. We thus need to assume that parties have access to a reliable network functionality with bounded delay Δ_{NET}, similar to the functionality $\mathcal{F}_{\text{N-MC}}^{\Delta_{\text{NET}}}$ in [3]. We further need a notion of time and access to clocks [18], and we assume an idealized signature functionality [2, 7]. To keep the presentation simple, we do not model all these functionalities and refer to the cited papers for UC-related details.

Time. The functionality $\mathcal{F}_{\text{CLOCK}}$ essentially amounts to assuming perfectly synchronized discrete clocks. We use *time slot* to denote the time period between two ticks of the clock $\mathcal{F}_{\text{CLOCK}}$. We use slot length to denote the length of time slots and we assume it to be fixed. In a slight abuse of notation, we also sometimes call a time slot a tick. By tick r we mean the time slot starting after the clock ticked r times. We assume time starts with a clock tick, so the first tick is tick 1. The execution proceeds in a way such that if honest party P_i is in tick r_i and honest party P_j is in tick r_j, then $|r_i - r_j| \leq 1$. We assume that

each party has enough computational power to complete arbitrary polynomial time computations in each time slot. Note that this may not be a good model of reality, but by assuming a known upper bound on the clock drift in "real life", and designing our protocols such that a bounded number of computation is required in each round, setting the slot length long enough, and assuming an upper bound on the message delivery, the model can clearly be realized. We stress that our goal is to communicate our sharding mechanism, which we believe is best done in a simple model. Implementing the protocol securely in practice is a highly non-trivial but largely orthogonal task.

Network. The network $\mathcal{F}_{\text{N-MC}}^{\Delta_{\text{NET}}}$ allows parties to multi-cast messages. The adversary determines when messages are output at honest parties, but the delay can be at most Δ_{NET} ticks. The bound Δ_{NET} is not known to honest parties making this a partially synchronous communication model.

Static vs. adaptive adversaries. As discussed in the introduction, we consider the set of corrupted parties to be static. We discuss possible extensions for adaptive corruptions in Section 5.3.

3 LEDGER FUNCTIONALITIES

3.1 (Sharded) Timed ledger

Timed ledger. A timed ledger $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ allows all parties to input messages, which will then be totally ordered and provided with a timestamp. More formally, every party P_i can input messages to the ledger, and retrieve a list TO_i of already ordered messages and their timestamps. The ledger provides the following guarantees:

Persistence: For message lists TO_i and TO_j seen by honest parties P_i and P_j, respectively, TO_i is a prefix of TO_j or vice versa.

Liveness and bounded timestamps: For any message m input at time t by an honest party, there is a time $t' \leq t + \Delta$ such that by time $t' + \Delta$, (m, t') is in TO_i for any honest P_i. The maximal delay Δ is fixed, but unknown to honest parties.

The timestamp property ensures that messages get timestamps on which the parties agree. We call these the *ledger arrival time*. They are important in defining precisely whether a message made it before a timeout. This is particularly important for parties that view messages on the ledger long after they were added.

Moreover, the ledger arrival time may differ from the time the message was input (by an honest party) or the time the message is delivered locally to an honest party. However, the bounded delay property ensures that those differences in time are bounded by Δ (similar to the delay assumed in the underlying network). Formally, we define the functionality $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ as a special case of a sharded timed ledger with a single shard, i.e., $\mathcal{F}_{\text{BD-TL}}^{\Delta} := \mathcal{F}_{\text{BD-STL}}^{\Delta,1}$. See below for the definition of $\mathcal{F}_{\text{BD-STL}}^{\Delta,\nu}$.

Realizing a timed ledger. The timed functionality can be implemented by a BFT consensus such as HotStuff [32], Algorand [11] or Tendermint [5, 22], or by a Nakamoto-style blockchain such as Bitcoin [25]. In case a Nakamoto-style blockchain is used, combining it with a finality layer, e.g., Casper the Friendly Finality Gadget [6] or Afgjort [14], is advisable to improve the finality time. See Section 6.1 for more details on possible implementations of $\mathcal{F}_{\text{BD-TL}}^{\Delta}$.

Sharded timed ledger. A sharded timed ledger is simply a collection of multiple timed ledgers running in parallel, where each individual ledger is refereed to as a shard. For simplicity, we assume the number of shards ν is fixed. In practice, it is desirable to have the ability to start new shards. This can be done easily with our techniques (similarly to how deadlocked shards are restarted), but we omit this to simplify the modelling.

Every shard is identified by a shard identifier $\mathtt{sid} \in \{1, \ldots, \nu\}$. Parties can send message to different shards using the shard identifier. Furthermore, every party P_i can retrieve a totally ordered list of messages in a shard $\mathtt{sid}$, denoted by $\mathsf{TO}_i^{\mathtt{sid}}$. Within each shard, the same guarantees as described above for the timed ledgers apply. Using the timestamps on each shard, parties can also merge the total orders of every shard to obtain a global total order. Inter-shard communication can also be implemented on top of a sharded timed ledger as discussed in Section 5.4.

Next, we formally define the $\mathcal{F}_{\text{BD-STL}}^{\Delta,\nu}$ functionality.

Functionality Sharded Timed Ledger $\mathcal{F}_{\text{BD-STL}}^{\Delta,\nu}$

Initialization

1: **for** $P_i \in \mathcal{P}$ and all $\mathtt{sid} \in \{1, \ldots, \nu\}$ **do**
2: $\quad \mathsf{TO}_i^{\mathtt{sid}} := ()$ *// Shard Total Order of shard* sid *for* P_i
3: $\quad \mathsf{Sent}^{\mathtt{sid}} := \emptyset$ *// Messages sent to Shard* sid

Interface for party $P_i \in \mathcal{P}$

Input: $(\textsc{Send}, \mathtt{sid}, m)$ *// Assume* $(\mathtt{sid}, m)$ *is sent at most once per honest* P_i.
1: Output $(P_i, \mathtt{sid}, m)$ to the adversary.
2: Add (P_i, m, t^{Now}) to $\mathsf{Sent}^{\mathtt{sid}}$.

Input: $(\textsc{Get}, \mathtt{sid})$.
3: Output $(\textsc{Get}, P_i)$ to the adversary. *// Inform adversary.*
4: Return $\mathsf{TO}_i^{\mathtt{sid}}$.

Interface for adversary

Input: $(\textsc{Add}, \mathtt{sid}, m, t)$
1: Append (m, t) to $\mathsf{TO}_i^{\mathtt{sid}}$. *// Unless it violates the below restrictions*

At any time, $\mathcal{F}_{\text{BD-STL}}^{\Delta,\nu}$ **automatically enforces these restrictions:**

Persistence: If P_i and P_j are honest, then either $\mathsf{TO}_i^{\mathtt{sid}}$ is a prefix of $\mathsf{TO}_j^{\mathtt{sid}}$ or $\mathsf{TO}_j^{\mathtt{sid}}$ is a prefix of $\mathsf{TO}_i^{\mathtt{sid}}$.

Liveness and bounded timestamps: For all messages $(\mathtt{sid}, m, t) \in \mathsf{Sent}^{\mathtt{sid}}$, there is a $t' \leq t + \Delta$ such that by time $t' + \Delta$, we have $(m, t') \in \mathsf{TO}_i^{\mathtt{sid}}$ for all honest P_i.

Notes about model and UC.. For the reader used to the formalization of a blockchain via a chain of blocks and properties like common prefix, chain quality and chain growth as in [16], the formalization via $\mathcal{F}_{\text{BD-STL}}$ might look overly simplified. That is meant as a feature. We only formalize the features relevant for our sharding protocol and abstract away other details. This simplifies the presentation and allows $\mathcal{F}_{\text{BD-STL}}$ to be instantiated using many different consensus protocols, as discussed in Section 6.1.

Our formalization is also closer to a property-based definition than a typical UC definition. We believe the functionality and its guarantees are easier to understand like this. Enforcing the restrictions guaranteeing persistence and liveness is done by dropping Add commands violating any restriction, and by executing extra Add commands when required to meet any restriction. We omit a

description of how exactly these enforcements are handled by the functionality since it is irrelevant for our protocol.

Also note that whenever we write that the functionality outputs something to the adversary, this is meant to leak to the adversary that the corresponding action has occurred. It it not supposed to hand over the activation token to the adversary. Technically, this can be understood as leaving the message in the functionality for the adversary and next time the adversary is activated, it can query the functionality to fetch the message. We again omit the details from the definition to simplify the presentation and focus on our main ideas rather than technicalities.

3.2 Shards

In essence, a shard is just a ledger. For our sharding protocol, however, we need to be explicit about the fact that we want to instantiate shards of different sizes and that a shard may not be live if it is instantiated with a committee that has a too high corruption ratio. Our sharding protocol therefore does not use $\mathcal{F}_{\text{BD-TL}}$ for the shards, but the functionality $\mathcal{F}_{\text{SHARD}}^{s,\mathcal{L},\Delta}$, which we introduce next. The protocol then realizes $\mathcal{F}_{\text{BD-STL}}$ by restarting deadlocked shards internally, cf. Section 5.

The functionality $\mathcal{F}_{\text{SHARD}}^{s,\mathcal{L},\Delta}$ is, in addition to the upper bound Δ on the delay, parameterized by the size s of the committee of parties that will be in charge of maintaining the shard, and the set $\mathcal{L}$ representing the liveness adversary structure. Upon initialization, a committee C of size s is sampled uniformly at random from all parties, and the resulting list is sent to all parties. We represent the committee as a vector $C = (P_1, \ldots, P_s)$ of parties, which allows giving special roles to specific parties, e.g., using the first committee member P_1 as a leader. We will sometimes abuse notation and refer to C as a set. The adversary structure $\mathcal{L}$ means that the shard functionality must maintain its liveness when the list $(i_1, \ldots, i_t)$ of the indices of corrupted parties $P_{i_1}, \ldots, P_{i_t}$ is in $\mathcal{L}$.

The parties $P_i \in C$ can interact with the shard functionality by sending transactions to the shard through the Send command, and retrieve the ledger through the Get command. The parties can also "close" the shard by issuing the Close command. Looking ahead, this is useful when the sharded ledger protocol (Section 5.2) requests parties to shut down a shard in order to start a new shard with different parameters and parties. Moreover, we allow *all* parties (including external parties not in C) to verify that C is indeed the correct committee. Additionally, all parties (including external ones) can request "finality proofs" from the functionality through the GetFinProof command. This proof can then be verified by *any* party. Our functionality offers two ways to verify such proofs: Using VerifyFinProof, one can verify a proof relative to a message vector $\vec{m}$, i.e., it can verify that the messages in $\vec{m}$ are finalized in the ledger. Alternatively, external parties can use VerifyFinLength to verify a proof relative to an integer ℓ, i.e., to simply verify that *at least* ℓ messages have been finalized so far. The latter allows to check liveness by ensuring a growing ℓ without needing to know the actual messages.

As the timed ledger, we model the guarantees of the ledger in the shards by letting the adversary control how messages are added to the ledger and imposing some restrictions on the adversary in the form of properties of the shard functionality. The persistence property is the standard property that one expects from a ledger, i.e., intuitively all honest parties will maintain ledgers that are prefixes of each other. We formalize this by considering a global FTO (finalized total order) and guaranteeing that the ledgers of all honest parties are prefixes of FTO. The liveness property is also standard and says that any message sent by an honest party will make it into the ledger of all honest parties after at most Δ time.[5] What is special about liveness of shards is that the property only needs to hold if the corrupted parties are in $\mathcal{L}$.

The novel properties that we require for our shard functionality are called *proof soundness* and *censorship resilience*. Proof soundness intuitively says that valid finality proofs can only be produced for correct statements. Censorship resilience prevents an adversary from excluding specific messages from the ledger, while including others. Note that liveness already guarantees that *all* messages will be added to the ledger within time Δ. Censorship resilience is thus a guarantee in case the ledger is not live. We formalize this as the guarantee that when a party is sending a message for inclusion, it will be included at most two ledger updates later.[6] In other words, either the message gets included, or the ledger stops completely. This is useful for our sharding protocol, because we need to detect liveness failures and in that case restart the shard. Censorship resilience now ensures that either the ledger is live, or it stops completely, which can be detected from the outside (in contrast to some undetectable censorship).

We formally define the shard functionality next and in Section 4.2 we show a protocol that UC-realizes this functionality.

Functionality $\mathcal{F}_{\text{SHARD}}^{s,\mathcal{L},\Delta}$

Interface for party $P_i \in \mathcal{P}$

Input: (InitShard, sid) *// Initialize shard with ID* sid
1: Output (InitShard, P_i, sid) to the adversary.
2: */* Select shard committee of size s */*
3: Upon receiving (InitShard, sid) (with the same sid) from all honest parties in $\mathcal{P}$, sample a sequence C uniformly among all sequences of length s from $\mathcal{P}$
4: Set FTO = () and $\text{TO}_i := ()$ for all $P_i \in C$
5: Send (sid, C) to all parties in $\mathcal{P}$

Interface for party $P_i \in C$ *// After InitShard*

Input: (Send, m)
1: Send (Send, P_i, m) to the adversary.

Input: Get
2: Send (Get, P_i) to the adversary.
3: **return** TO_i

Input: Close
4: Send (Close, P_i) to the adversary.

Input: GetFinProof
5: Send (GetFinProof, P_i) to the adversary, who immediately sends back a proof π such that no record $(\text{TO}_i, \pi, 0)$ has been stored.
6: Store the record $(\text{TO}_i, \pi, 1)$
7: **return** (TO_i, π)

[5] Note that the delay Δ is a parameter of the functionality, but it may not be known to the honest parties. This is in particular the case when considering the partially synchronous model.

[6] When the ledger is realized using blocks, this means the block after the next block must include this message. One could more generally also allow for larger delays than two blocks, but we here avoid the extra parameter.

Interface for adversary

Input: (Add, $\vec{m}, i$)
1: Append $\vec{m}$ to TO_i.

Input: (AddFinal, $\vec{m}$)
2: Append $\vec{m}$ to FTO.

Public interface *// Any (even "outside") party can use this interface.*

Input: (VerifyCommittee, sid, C)
1: If (sid, C) has been sent to all parties in $\mathcal{P}$, return 1, otherwise, return 0.

Input: (VerifyFinProof, $\vec{m}, \pi$)
2: **if** record $(\vec{m}, \pi, b)$ for some $b \in \{0,1\}$ exists **then return** b
3: **else**
4: Send (VerifyFinProof, $\vec{m}, \pi$) to adversary, who immediately replies with $b \in \{0,1\}$.
5: Store the record $(\vec{m}, \pi, b)$.
6: **return** b

Input: (VerifyFinLength, ℓ, π) *// Only verify length ℓ of message vector*
7: **if** record $(\vec{m}, \pi, b)$ for some $\vec{m}$ with $|\vec{m}| = \ell$ and $b \in \{0,1\}$ exists **then**
8: **return** b
9: **else**
10: Send (VerifyFinLength, ℓ, π) to the adversary.
11: Adversary immediately replies with $b \in \{0,1\}$ and $\vec{m}$ with $|\vec{m}| = \ell$.
12: Store the record $(\vec{m}, \pi, b)$.
13: **return** b

At any time, the functionality automatically enforces the following:

Let $A \subseteq \{1, \ldots, s\}$ be the indices of corrupted parties in C. We call the ledger *live* if $A \in \mathcal{L}$. Call the ledger *weakly closed* if some honest party input Close.

Persistence: For all $P_i \in C \setminus A$, TO_i is a prefix of FTO.

Liveness: (*If the ledger is live and not weakly closed*) After a message m was input (via (Send, m)) by an honest party for the first time,we have $m \in \mathsf{TO}_i$ for all $P_i \in C \setminus A$ at most Δ time later.

Censorship Resilience: After a message m was input (via (Send, m)) by an honest party at time t and TO_i for a honest P_i was updated twice after time $t + \Delta$, we have $m \in \mathsf{TO}_i$.

Proof Soundness: If (VerifyFinProof, $\vec{m}, \pi$) returns 1, then $\vec{m}$ is a prefix of FTO. If (VerifyFinLength, ℓ, π) returns 1, then $|\mathsf{FTO}| \geq \ell$.

4 COMMITTEE SELECTION AND SHARD CONSENSUS

4.1 Committee Selection

In this section, we describe a committee-selection functionality, which is a core part of our sharding solution. We then discuss how to realize it given a randomness beacon and analyze the required committee sizes.

Committee selection functionality. The functionality $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ is parameterized by the set $\mathcal{P}$ of parties executing the committee selection, and a set $\mathcal{U}$ of parties from which the committee gets selected. It allows parties to request uniformly distributed sequences over $\mathcal{U}$ of a given length. This corresponds to the committee selection step in the shard functionality, see Section 3.2. The functionality is formally defined as follows:

Functionality $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$

Interface for party $P_i \in \mathcal{P}$

Input: (SelectCom, cid, s) *// Select committee with ID* cid *of size s*
1: Output (SelectCom, P_i, cid, s) to the adversary.
2: Upon receiving (SelectCom, cid, s) (with the same cid and s) from all honest parties in $\mathcal{P}$, sample a sequence C uniformly among all sequences of length s from $\mathcal{U}$, and send (cid, C) to all parties in $\mathcal{P}$.

Public interface *// Any (even "outside") party can use this interface.*

Input: (VerifyCommittee, cid, C)
1: If (cid, C) has been sent to all parties in $\mathcal{P}$, return 1, otherwise, return 0.

Selecting parties proportional to their resources. The functionality $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ selects parties from a set $\mathcal{U}$ such that each party is selected as a committee member with equal probability. However, in permissionless blockchain protocols, corruption thresholds are typically expressed in terms of the amount of a restricted resource controlled by a party (e.g., the amount of relative stake in proof-of-stake based blockchains or the amount of computational power in proof-of-work based blockchains). Hence, it is necessary to map the parties executing the underlying blockchain protocol into (virtual) parties in such a set $\mathcal{U}$ according to the resources they control. In the setting of proof-of-stake based blockchains, such mapping can be achieved by the techniques commonly known as "follow-the-satoshi" [8, 20], "weighing by stake" [13], and "cryptographic sortition" [11, 12, 17]. In the setting of Proof-of-Work based blockchains, committee selection has also been studied [26]. We can thus assume that individual parties are mapped into users in the set $\mathcal{U}$ proportional to the relevant resources they control, and refer the interested readers to the aforementioned results on committee selection on blockchains.

Realizing $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ from a randomness beacon. Given access to a randomness beacon, it is straightforward to realize $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$. Such a randomness beacon gives all parties access to uniformly sampled randomness and allows parties to verify the randomness. A formal definition is provided in [9], where it is also shown how this functionality can be efficiently UC-realized both based on the DDH assumption (with UC zero knowledge as setup) or on the CDH assumption (with a global random oracle as setup). Moreover the protocols proposed in [9] require a public bulletin board that guarantees that posted messages become immutable and accessible to all honest parties. Note that such a bulletin board can easily be realized by a ledger $\mathcal{F}_{\text{BD-TL}}$. See Section 6.3 for more details on how to instantiate such a randomness beacon.

A straightforward way to realize $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ assuming a randomness beacon works as follows. Given a committee size s, use randomness from the beacon to sample a uniformly random sequence C from $\mathcal{U}$ with $|C| = s$. Note that since the beacons also provide the same randomness to all parties, the parties agree on the selected committees. This also directly allows parties to verify selected committees. It is easy to see that this realizes $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$.

Remark on our committee selection. Looking ahead, we always instantiate the committee selection functionality $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ with the entire population $\mathcal{U}$, i.e., every time $\mathcal{F}_{\text{ComSel}}^{\mathcal{P},\mathcal{U}}$ is executed it returns a random party from the entire population $\mathcal{U}$. Crucially, this means that it is possible that the same party can be part of more than one shard at the same time. This greatly simplifies our analysis and allows us to focus on the techniques to reduce the size of sharding committees. We stress that for practical purposes a more complex committee selection should be used, minimizing the number of shards a single party can be assigned to.

Alternative ways to realize committee selection. In the setting of proof-of-work based blockchains, committee selection based on the proof-of-work mechanism itself (without randomness beacons) has been constructed in [26]. In previous works on proof-of-stake based blockchain consensus protocols [8, 11–13, 17, 20], several methods have been proposed for selecting committees in a publicly verifiable way using randomness beacons. These methods can be classified in two main categories according to the underlying randomness beacon: (1) uniformly random committee selection using randomness beacons based on coin tossing with guaranteed output delivery [8, 20]; and (2) biased committee selection using randomness beacons based on verifiable random functions [11–13, 17]. The simple protocol we have described above falls into category 1. While the methods in category 2 allow an adversary to bias committee selection in ways not possible in category 1, they are more efficient. To keep the presentation simple, our formalization and the derived bounds assume uniform committee selection. However, our results can be extended to also work with biased committees.

4.2 Shard Consensus

In this section, we present a simple protocol using the committee selection functionality from Section 4.1 to implement the shard functionality described in Section 3.2.

Our protocol is parameterized by the committee size s and the maximal number t_{L} of corrupted parties in the committee that can be tolerated without losing liveness. The parameter t_{L} can be any number in $\{0, \ldots, \lfloor \frac{s-1}{2} \rfloor\}$. Based on the safety-liveness dichotomy, the protocol then sets $t_{\mathrm{S}} := s - 2t_{\mathrm{L}} - 1$, where t_{S} is the maximal number of corrupted committee members that can be tolerated without breaking safety. For example, with $s = 100$, one can set $t_{\mathrm{L}} = 33$ to obtain $t_{\mathrm{S}} = 33$, which are the classical bounds for partially synchronous Byzantine agreement. One can also set $t_{\mathrm{L}} = 0$ to obtain $t_{\mathrm{S}} = n - 1$, i.e., if full honesty is required for liveness, all but one party can be corrupted without breaking safety. Consequently, the protocol can only be instantiated securely with committee sizes s that guarantee at most t_{S} corruptions except with negligible probability. See Section 4.3 for how to compute these minimal committee sizes.

The protocol idea is simple. Upon Initialization, the functionality $\mathcal{F}_{\mathrm{CT}}^{\mathcal{P},\mathcal{D}}$ is invoked to obtain a uniformly chosen committee of size s. The first member of the committee is designated a special leader role we call "sequencer". The sequencer periodically proposes a new block containing new messages, which is then signed by the other parties. A block is considered final if at least $s - t_{\mathrm{L}}$ parties have signed it. This is the same basic idea underlying many BFT consensus algorithms [10, 32]. One difference is that our protocol also considers values of t_{L} smaller than the maximal $\lfloor \frac{s-1}{2} \rfloor$. Furthermore, typical BFT consensus protocols include (usually involved) mechanisms for replacing a corrupted leader [10, 32]. Since our sharding protocol already has an external mechanism (using the control chain) to replace committees, we do not need a leader replacement mechanism in our shard consensus, greatly simplifying it. This means our shard will be live if at most t_{L} parties are corrupted and the leader is honest. That is, our protocol realizes $\mathcal{F}_{\mathrm{SHARD}}^{s,\mathcal{L},\Delta}$ for

$$\mathcal{L} = \{A \subseteq \{1, \ldots, s\} \mid |A| \leq t_{\mathrm{L}} \ \wedge \ 1 \notin A\}.$$

The "special features" of $\mathcal{F}_{\mathrm{SHARD}}$ are straightforward to achieve: Committee verification is provided by the committee selection functionality. Finalization proofs are simply the signatures from the committee members, which can be publicly verified. To guarantee censorship resilience, every party sends a list of "old" messages that have not yet been included in the ledger together with the signature on the proposed block. If the sequencer does not include these messages in the next block, the party refuses to sign that block. Thus, if the sequencer tries to censor a message the honest committee members want to have included, the shard will lose liveness, and in the overall sharding protocol, the committee (including the malicious sequencer) will be replaced.

For a formal description and a security proof, see the full version.

Leader replacement within the shard consensus. We have described a particularly simple protocol without leader replacement since our sharding protocol detects corrupted leaders and consequently resamples the committee. The advantage of this is that the shard consensus is extremely simple and efficient with honest leaders. The downside is that in case of a corrupted leader, the whole committee gets resampled, which may be less efficient than an optimized leader-replacement mechanism of state-of-the-art BFT consensus protocols such as HotStuff [32].

Alternatively, one can use a BFT consensus with builtin leader replacement in the shards. In that case, one can realize $\mathcal{F}_{\mathrm{SHARD}}^{s,\mathcal{L},\Delta}$ for

$$\mathcal{L} = \{A \subseteq \{1, \ldots, s\} \mid |A| \leq t_{\mathrm{L}}\}.$$

That is, the additional restriction of an honest leader is not needed and fewer committee resamplings in the shard protocol may suffice to reach liveness.

These protocols are typically only specified for the special case $t_{\mathrm{L}} = t_{\mathrm{S}} = \lfloor \frac{s-1}{3} \rfloor$. It is straightforward to generalize, e.g., HotStuff [32] to also work with smaller t_{L} and $t_{\mathrm{S}} = s - 2t_{\mathrm{L}} - 1$ by only accepting blocks that have been signed by at least $s - t_{\mathrm{L}}$ committee members. The "special features" of $\mathcal{F}_{\mathrm{SHARD}}$ can be achieved similarly as in our protocol described above and liveness and safety follow can be proven analogously.

4.3 Determining the Committee Size

Our sharding protocol needs to find the smallest committee size s_{min} that guarantees that the ratio of corrupted parties in the selected committee is below some given threshold with overwhelming probability. Consider a scenario with a total population of n parties $\mathcal{P}$ such that at most t parties are corrupt, and a committee C with size s sampled uniformly at random from $\mathcal{P}$. We denote by $\mathsf{FAIL}_{t',s}^{t,n}$ the event where the committee C contains more than t' corrupt parties. The probability of the event $\mathsf{FAIL}_{t',s}^{t,n}$ happening can be expressed as the cumulative hypergeometric probability mass function:

$$\Pr\left[\mathsf{FAIL}_{t',s}^{t,n}\right] = \sum_{i=t'+1}^{i=s} \frac{\binom{t}{i}\binom{n-t}{s-i}}{\binom{n}{s}}.$$

Given a maximal admissible corruption ratio $\frac{t'}{s}$ of the sampled committee C of size s one can find the smallest size s_{min} for which $\Pr\left[\mathsf{FAIL}_{t',s}^{t,n}\right] \leq 2^{-\kappa}$, for some security parameter κ. In the full version, we derive an analytical bound on the required committee size and provide Python code that computes these sizes precisely.

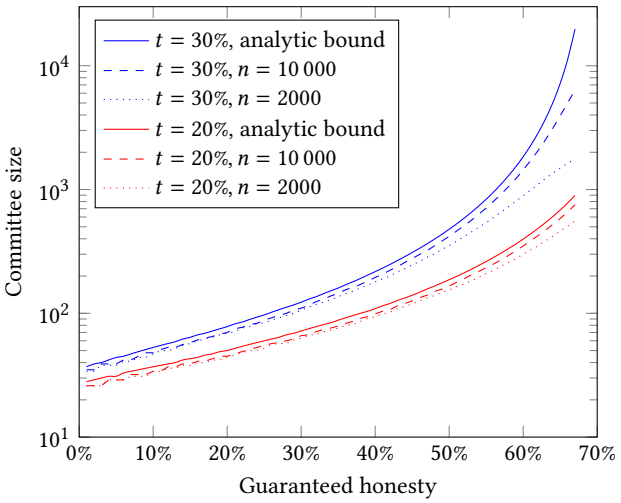

Figure 1: Required committee sizes (on a log scale) to guarantee different honesty levels in the committee with probability $1 - 2^{-60}$, assuming a total population of $n \in \{2000, 10\,000\}$ parties with $t \in \{20\%, 30\%\}$ corruption ratio.

Figure 1 shows the relation between the minimum committee sizes and the level of required guaranteed honesty in the committees for different settings. As the graphs show, the required committee size grows exponentially with the required honesty level. Hence, requiring a smaller guaranteed honesty level than the usual 1/3 significantly improves performance.

Recall that our shard consensus from Section 4.2 can be instantiated with different liveness and safety thresholds L and S, respectively, as long as $S + 2L < 100\%$. Since we want our shards to always be safe, we need to sample committees such that the corruption ratio in the committee is at most S with probability at least $1 - 2^{-\kappa}$ for security parameter κ. Table 1 shows minimal committee sizes for different safety thresholds in different settings.

Note that if the actual corruption ratio in the total population is close to the liveness threshold, there is a good probability for the shard being live. That means if we assume an overall corruption ratio of at most 30%, we never need to use liveness thresholds above 30%, corresponding to safety thresholds of 39%. This already gives a significant performance improvement over prior work, which needed a threshold of 1/3. In the optimistic case, where the actual corruption ratio is below the worst case assumption of 30%, we can run with even smaller committees.

5 CONSTRUCTING A SHARDED LEDGER

5.1 Overview

To realize $\mathcal{F}_{\text{BD-STL}}^{\Delta',\nu}$, we use a timed ledger $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ as a control chain. In order to keep the shards live and monitor their liveness, the parties in $\mathcal{P}$ will follow instructions based on messages posted on $\mathcal{F}_{\text{BD-TL}}^{\Delta}$. For the sake of simplicity, we model $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ as a local functionality. However, our proof does not crucially rely on this modeling choice as it is not necessary for the simulator to program $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ or interfere in its behavior in any way. Moreover, we use a repository $\mathcal{F}_{\text{Repo}}$ to store finalized parts of a shard and previous shards' states, which parties can later obtain when reading from a shard or joining a shard committee. This use of $\mathcal{F}_{\text{Repo}}$ captures the fact that it is necessary for at least one party to always store the state of each shard. See the full version for more details on $\mathcal{F}_{\text{Repo}}$.

In order to execute shards, we use a "gearbox" of shard functionalities $\mathcal{F}_{\text{Shard}}^{s_1,\mathcal{L},\Delta}, \ldots, \mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ with increasing committee sizes $s_1 \leq s_2 \leq \cdots \leq s_\ell$ to handle shard consensus. In principle, each shard could use its own gearbox, with a different progression of shard functionalities for each shard (e.g., mixing different shard consensus protocols). Moreover, each shard could operate with a different liveness structure $\mathcal{L}$ and a different delay Δ. For the sake of presentation, we use a gearbox with the same shard functionalities with fixed liveness structures and delay parameters for all shards. There is a statistical security parameter κ, and a liveness guarantee $\gamma > 0$, e.g., $\gamma = \frac{1}{2}$. The gearbox has the following properties.

Always safe: For any committee size s_i, $\mathcal{F}_{\text{Shard}}^{s_i,\mathcal{L},\Delta}$ is safe except with probability $2^{-\kappa}$.

Eventually live: $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ is live with probability at least γ.

We achieve safety and liveness at the same time by first running $\mathcal{F}_{\text{Shard}}^{s_1,\mathcal{L},\Delta}$ for a random committee. If it loses liveness we switch gears to $\mathcal{F}_{\text{Shard}}^{s_2,\mathcal{L},\Delta}$ and so on. We start with $\mathcal{F}_{\text{Shard}}^{s_1,\mathcal{L},\Delta}$ with poor liveness but a very small committee and consequently high efficiency. On the other hand, $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ has strong liveness at the cost of efficiency. The other $\mathcal{F}_{\text{Shard}}^{s_i,\mathcal{L},\Delta}$ act as intermediary points on the liveness versus efficiency scale, so some of them might be realized by the same consensus mechanism with the same parameters. For example, it may make sense to sample committees of the same size several times to try to hit one with high honesty and hence liveness. As long as we have the properties described above, we can switch to the next functionality in our gearbox every time we lose liveness. When we hit $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$, we simply switch to a new instance of $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ with a fresh committee, which guarantees that we will eventually get liveness since $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ is live with probability at least γ.

Optimistic committee sizes. As described above, the gearbox increases committee sizes until the shard is live. We now discuss how large committees need to be to get liveness. For concreteness, we again consider a total population of $10K$ nodes with at most 30% corruption, as in Section 4.1. Even though we consider a worst-case corruption of up to 30% in the total population, in an optimistic scenario the system can have a lot less corruption. Note that the corruption ratio in the sampled shard is close to the overall corruption ratio with high probability. Thus, when the liveness threshold of the current gear is equal to the actual total corruption threshold, there is a good probability that the shard is live. Since the gearbox ensures that shards are always safe, one can think of shifting up as increasing the liveness threshold L and adjusting the safety threshold S such that $S + 2L < 100\%$. As soon as the liveness threshold matches the unknown actual corruption ratio, the shard is live with good probability and the gearbox stops switching up. This means the protocol automatically finds the optimal shard size without knowing the actual corruption ratio. Note that even in the worst case with 30% overall corruption, we only need $L = 30\%$ and $S = 39\%$. Hence, we can sample a committee with a guaranteed

Table 1: Minimum committee sizes for different liveness and safety thresholds such that $S + 2L < 100\%$ with total populations $n = 2000$ and $n = 10\,000$ with 20% and 30% corruption. Minima are computed to guarantee safety except with probability 2^{-60}. Values for $n = \infty$ are the analytical bounds.

Liveness threshold:		0%	5%	10%	15%	20%	25%	30%	33.3333%
Safety threshold:		99%	89%	79%	69%	59%	49%	39%	33.3333%
n	t	**Minimal committee size**							
∞	30%	37	55	82	130	232	528	2264	16 037
10 000	30%	35	51	75	116	207	462	1713	5886
2000	30%	34	50	71	112	190	382	990	1716
∞	20%	28	38	52	75	115	199	438	854
10 000	20%	26	34	47	67	104	178	385	717
2000	20%	26	34	46	66	99	164	326	540
		our work							state of the art

corruption of at most 39%, instead of the 33% required by solutions not leveraging the safety-liveness dichotomy.

The numbers in Table 1 can thus be used to determine which committee sizes are required to achieve liveness with good probability for a given actual corruption ratio. For example, in the worst case with 30% overall corruption, liveness can be expected for liveness threshold 30%, corresponding to 1713 committee members. With 20% actual corruption, we only need committees of size 207.

5.2 The Sharded Ledger Protocol $\Pi_{\text{BD-STL}}$

In our protocol, parties in $\mathcal{P}$ continuously perform a number of maintenance actions to detect the potential loss of liveness in shards and ensure that the next shard functionalities in the gearboxes take over the operation of shards that lose liveness. These actions can be divided as follows: (1) Shard Management, which are actions performed by all parties in $\mathcal{P}$ in order to maintain all shards live; (2) Shard Operation, which are actions performed by the parties in the committee responsible for operating a given shard. Members of a shard committee only execute shard management commands related to their shard when these commands have been issued by a majority of parties in $\mathcal{P}$ and appear in a finalized state of the control chain. Moreover, when receiving inputs, the parties execute instructions that realize the interfaces in $\mathcal{F}_{\text{BD-STL}}^{\Delta',\nu}$. Notice that $\mathcal{F}_{\text{BD-STL}}^{\Delta',\nu}$ operates with a fixed number of shards and that start and stop operations are only performed in order to restart shards that have lost liveness but not to add new shards or remove existing ones. While this simplification is done for the sake of a clear exposition, it is straightforward to generalize the protocol to allow for adding and removing shards. Protocol $\Pi_{\text{BD-STL}}$ works as follows.

Protocol $\Pi_{\text{BD-STL}}$

Parties in $\mathcal{P}$ interact with each other and with functionalities $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, $\mathcal{F}_{\text{Repo}}$ and the gearbox of functionalities $\mathcal{F}_{\text{Shard}}^{s_1,\mathcal{L},\Delta}, \ldots, \mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$. All parties in $\mathcal{P}$ execute the Shard Management steps continuously and execute Shard Operation steps for a given shard when elected as a shard committee member. Each shard is identified by `sid` and has a gear parameter h, indicating the current committee size C_h and functionality $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ responsible for that shard, start time t, indicating when the shard execution with $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ started in terms of $\mathcal{F}_{\text{BD-TL}}^{\Delta}$'s ledger time, and finalization timeout t_{Timeout}.

Shard Management. When execution starts, parties $P_i \in \mathcal{P}$ execute the **Init** steps and then continuously perform **FinalizeCheck**.

Init: Execute Start steps for all shards with parameters $h = 1, t = 1, t_{\text{Timeout}} = \Delta_{init}$ for all `sid`, where $\Delta_{init} \in \mathbb{Z}, \Delta_{init} > 0$.

FinalizeCheck: Parties $P_i \in \mathcal{P}$ keep counters $L_{last}^{\texttt{sid}}$ initially set to 0 for each shard identified by `sid`. Parties $P_i \in \mathcal{P}$ continuously send (Get) to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, receiving TO_i. For every shard identified by `sid`, all $P_i \in \mathcal{P}$ perform the following steps to check that a shard has liveness:

1: For every message $((\text{Finalize}, \texttt{sid}, H, \pi, L), t) \in \text{TO}_i$ check that that $t < c_{Timeout}^{\texttt{sid}}$ and $L > L_{last}^{\texttt{sid}}$, send $(\text{VerifyFinLength}, L, \pi)$ to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$, obtaining b, and check $b = 1$.

2: Let $((\text{Finalize}, \texttt{sid}, H, \pi, L_{max}), t) \in \text{TO}_i$ be the message with the maximum L for which the checks of Step 1 succeeded. Set $c_{Timeout}^{\texttt{sid}} = t + t_{\text{Timeout}}$ and set $L_{last} = L_{max}$.

3: If the checks did not succeed for any message $((\text{Finalize}, \texttt{sid}, H, \pi, L), t) \in \text{TO}_i$ (i.e., the shard has timed out), execute the Stop procedure for shard `sid` and, after the stop command from a majority of parties in $\mathcal{P}$ appears in a future TO_j finalized by $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, execute the Start procedure for shard `sid` with incremented parameters $h++, t', t_{\text{Timeout}}++$, where t' is a future ledger time w.r.t. $\mathcal{F}_{\text{BD-TL}}^{\Delta}$. If $h > \ell$, set $h = \ell$ and use a new instance of $\mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$.

Start: All parties $P_i \in \mathcal{P}$ proceed as follows to start a shard identified by `sid` with parameters h, t, t_{Timeout}:

1: Send $(\text{Send}, (\text{Start}, \texttt{sid}, h, t, t_{\text{Timeout}}))$ to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, i.e., a command to start shard `sid` with $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ at ledger time t w.r.t. $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ with finalization timeout t_{Timeout}.

2: Send $(\text{InitShard}, \texttt{sid})$ to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$.

3: Set a finalization timeout counter to $c_{Timeout}^{\texttt{sid}} = t + t_{\text{Timeout}}$.

Stop: All parties $P_i \in \mathcal{P}$ send $(\text{Send}, (\text{Stop}, \texttt{sid}))$ to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, instructing parties to stop executing shard `sid`.

Shard Operation. For every shard identified by `sid`, parties $P_i \in \mathcal{P}$ continuously send (Get) to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, receiving TO_i. A command for starting or stopping a shard is only considered valid if it has been posted to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ by a majority of the parties in $\mathcal{P}$. Parties $P_i \in \mathcal{P}$ execute the following Shard Operation instructions according to the messages in TO_i and the ledger time:

Start Shard: When there is a command $((\text{Start}, \texttt{sid}, h, t, t_{\text{Timeout}}), t') \in \text{TO}_i$ posted by a majority of the parties in $\mathcal{P}$, all $P_i \in \mathcal{P}$ wait for $(\texttt{sid}, C_{\texttt{sid}})$ from $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$. If $P_i \in C_{\texttt{sid}}$, it sets $c_{Timeout}^{\texttt{sid}} = t + t_{\text{Timeout}}$ and responds to inputs $(\text{Send}, \texttt{sid}, m)$ and $(\text{Get}, \texttt{sid})$.

Finalize: A system parameter Δ_e is estimated in order to ensure that finalization messages are finalized by $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ before the timeout. Every time a shard is restarted, Δ_e is incremented by all parties. At ledger time $c_{Timeout}^{\mathtt{sid}} - \Delta_e$, parties $P_i \in C_{\mathtt{sid}}$ for shard $\mathtt{sid}$ proceed as follows:

1: Send Get to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ for shard $\mathtt{sid}$, obtaining $\mathsf{TO}_i^{\mathtt{sid}}$.
2: Send GetFinProof to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ for shard $\mathtt{sid}$, obtaining the corresponding finalization proof π.
3: Send $(\text{Send}, \mathsf{TO}_i)$ to $\mathcal{F}_{\text{Repo}}$, receiving $H(\mathsf{TO}_i)$. Notice that if a prefix of TO_i has already been stored in $\mathcal{F}_{\text{Repo}}$, only the new messages in TO_i w.r.t. this prefix need to be sent to $\mathcal{F}_{\text{Repo}}$.
4: Send $(\text{Send}, (\text{Finalize}, \mathtt{sid}, H(\mathsf{TO}_i), \pi, |\mathsf{TO}_i|))$ to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$.

Stop Shard: When parties $P_i \in C_{\mathtt{sid}}$ see a command $((\text{Stop}, \mathtt{sid}), t) \in \mathsf{TO}_i$ from a majority of the parties in $\mathcal{P}$, they send Close to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ and stop executing further shard operation instructions for shard $\mathtt{sid}$.

Interfaces from $\mathcal{F}_{\text{BD-STL}}^{\Delta',\nu}$ for Parties $P_i \in \mathcal{P}$.

On input $(\text{Send}, \mathtt{sid}, m)$**:** P_i proceeds as follows:

1: Send Get to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, receiving TO_i.
2: Find the latest message $((\text{Start}, \mathtt{sid}, h, t, t_{\text{Timeout}}), t_1) \in \mathsf{TO}_i$ posted by a majority of parties in $\mathcal{P}$. If there exists a message $((\text{Stop}, \mathtt{sid}), t_2)$ for $t_2 > t_1$ posted by a majority of parties in $\mathcal{P}$, repeat this step until a new message $((\text{Start}, \mathtt{sid}, \mathsf{cid}, C_h, t, t_{\text{Timeout}}), t_3) \in \mathsf{TO}_i$ for $t_3 > t_2$ posted by a majority of parties in $\mathcal{P}$ appears.
3: If $P_i \in C_{\mathtt{sid}}$, when receiving m (as input or from another party), P_i sends (Send, m) to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ for shard $\mathtt{sid}$. Otherwise, Send m to all parties in $C_{\mathtt{sid}}$ (obtained from $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ for shard $\mathtt{sid}$).
4: Continuously send (Get) to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, receiving TO_i and checking that there is a message $((\text{Finalize}, \mathtt{sid}, H, \pi, L), t) \in \mathsf{TO}_i$ such that $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ returns 1 when queried with $(\text{VerifyFinProof}, \vec{m}, \pi)$, where $\vec{m}$ is in D obtained by sending (Get, H) to $\mathcal{F}_{\text{Repo}}$. If a message $((\text{Stop}, \mathtt{sid}), t)$ appears before these checks succeed, P_i goes to Step 1 and waits for a new message $((\text{Start}, \mathtt{sid}, h', t', t'_{\text{Timeout}}), t) \in \mathsf{TO}_i$ posted by a majority of parties in $\mathcal{P}$.

On input $(\text{Get}, \mathtt{sid})$**:** If $P_i \in C_{\mathtt{sid}}$ (obtained from $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ for shard $\mathtt{sid}$), P_i sends (Get) to $\mathcal{F}_{\text{BD-TL}}^{\Delta}$, receiving TO_i. Otherwise, P_i ignores the next steps. P_i determines $\mathsf{TO}_i^{\mathtt{sid}}$ by executing the following instructions starting from the largest value of L_j for each $((\text{Finalize}, \mathtt{sid}, H_j, \pi_j, L_j), t_j) \in \mathsf{TO}_i$:

1: Send $(\text{VerifyFinLength}, L_j, \pi_J)$ to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ (where h is determined by the last valid message $((\text{Start}, \mathtt{sid}, h, t, t_{\text{Timeout}}), t') \in \mathsf{TO}_i$), obtaining b.
2: Send (Get, H_j) to $\mathcal{F}_{\text{Repo}}$ to get D_j, and send $(\text{VerifyFinProof}, D_j, \pi_j)$ to $\mathcal{F}_{\text{Shard}}^{s_h,\mathcal{L},\Delta}$ to get b'.
3: If $b = 0$ or $b' = 0$, ignore the next step and proceed to the next message $((\text{Finalize}, \mathtt{sid}, H_{j+1}, \pi_{j+1}, L_{j+1}), t_{j+1}) \in \mathsf{TO}_i$ with $L_{j+1} < L_j$. Otherwise, Let D' be the new messages in D_j that are not contained in D_{j-1}. Append (D', t_j) to $\mathsf{TO}_i^{\mathtt{sid}}$.

Notice that if a previous version of $\mathsf{TO}_i^{\mathtt{sid}}$ has already been computed, P_i only needs to perform these steps for new messages $((\text{Finalize}, \mathtt{sid}, H_j, \pi_j, L_j), t_j) \in \mathsf{TO}_i$ such that $t_j > \hat{t}$, where $\hat{t}$ is the highest ledger time registered in the previous version of $\mathsf{TO}_i^{\mathtt{sid}}$. Finally, P_i outputs $\mathsf{TO}_i^{\mathtt{sid}}$.

Theorem 5.1. *Protocol $\Pi_{BD\text{-}STL}$ described above UC-realizes functionality $\mathcal{F}_{BD\text{-}STL}^{\Delta',\nu}$ in the $(\mathcal{F}_{BD\text{-}TL}^{\Delta}, \mathcal{F}_{Repo}, \mathcal{F}_{Shard}^{s_1,\mathcal{L},\Delta}, \ldots, \mathcal{F}_{Shard}^{s_\ell,\mathcal{L},\Delta})$-hybrid model in the partially synchronous model (i.e., where Δ, Δ' are unknown but finite) with security against active static adversaries.*

The proof of Theorem 5.1 is presented in the full version.

5.3 Extensions

While Protocol $\Pi_{\text{BD-STL}}$ realizes a sharded ledger $\mathcal{F}_{\text{BD-STL}}$, it can have its efficiency and security improved by extending the way it switches gears over functionalities $\mathcal{F}_{\text{Shard}}^{s_1,\mathcal{L},\Delta}, \ldots, \mathcal{F}_{\text{Shard}}^{s_\ell,\mathcal{L},\Delta}$ for each shard. Here we describe some of these extensions informally.

5.3.1 Damping. So far we only showed how to move up the gearbox of consensus algorithms. That is enough to prove eventual liveness. In practice one also wants to have a way to regain efficiency if the loss of liveness was due to some temporary event such as a burst error in the network. This can be achieved using some heuristic. The timestamps on the control chain can be used to determine the uptime of the shards, and if the uptime exceeds some heuristic threshold, one tries to move down the gearbox again. This will tend to find the optimal position in the gearbox producing only some acceptable downtime. Since safety is never violated, any heuristic can be used that works well in practice.

5.3.2 Adaptive and Mobile Corruptions. Our analysis so far has assumed static corruptions. An inherent problem with committee-based protocols is that an adaptive adversary can corrupt all parties in a committee to break security. A straightforward way to tolerate slowly adaptive corruptions (as recently formalized as δ-delayed corruptions in [24]), one can resample committees repeatedly. This can even be combined with the damping described above, i.e., one can close and reopen shards after some timeouts and possibly change gears when resampling. This is secure as long as the time it takes to corrupt a party is longer than the resampling timeout.

5.4 Inter-Shard Transactions and Communication

There are different approaches for inter-shard communication in the literature (for an overview of different solutions, see [1, 30, 35]). Describing such a solution in full detail in our model would require us to include a notion of transactions and accounts (or UTXO [25]), which we have not included in our formalism for clarity. A formal treatment of inter-shard communication is thus out of the scope of this paper, where we want to focus on how to minimize committee sizes using the safety-liveness dichotomy. Below, we sketch how existing solutions can be adapted to our setting.

Special care needs to be taken in our setting because it can happen that shards are not live and consequently get restarted with a new committee. Thus, one cannot immediately trust transactions appearing on one shard for intershard transactions. One can, however, leverage the heartbeats posted on the control chain, which is always safe and live, to ensure transaction finality.

Adapting Atomix. We give a high-level overview of how Atomix[7], which is the intershard transaction protocol of OmniLedger [21], can be adapted to our setting. When a user wants to initiate a transaction from shard A to shard B using Atomix, the user's client first creates an "export" transaction on shard A, which effectively locks the user's funds. Once the transaction is included in shard A, the user's client posts a proof that the transaction was accepted by shard A on shard B, which makes the funds available on shard B.

If such a protocol is naively used with our sharding approach, it may happen that a lock transaction appears on shard A and the funds are unlocked on shard B, but shard A crashes before

[7] Note that the Atomix protocol allowed for a subtle replay attack, which was fixed in a subsequent work [29]. This issue is, however, not relevant for the high-level overview we give here.

a heartbeat can finalize the lock transaction. When shard A gets restarted with a fresh committee starting from the last heartbeat, the new committee may include other transactions on shard A first, invalidating the lock transaction and allowing to double spend.

To avoid this issue, the acceptance proof must not only prove that a transaction has happened on shard A, but also that it was finalized on the control chain. Recall that in our sharding protocol, all shards regularly post heartbeats on the control chain, where each heartbeat contains a hash of the latest block in the shard. For intershard transactions, we additionally assume that block headers contain the root of a Merkle tree containing (among other data) all the transactions in a block produced by the shard as well as hashes of the blockheader of the parent block. An acceptance proof now consists of a Merkle proof for the transaction in the block plus additional block-header hashes if the transaction was in a block between two heartbeats. These proofs can be further optimized by arranging the hashes of block headers in a tree.

Note that with this approach, even though all intershard messages are finalized through the control chain, only a single heartbeat is needed to finalize all messages of a given shard since the last heartbeat, independently of how many intershard transactions were initiated on that shard. Hence, the throughput of the shards is not bottlenecked by the control chain.

Shard-committee driven intershard transactions. Atomix is a client-driven approach in which the client directly sends the messages to the receiving shard. This puts an additional responsibility on the clients, which may be undesirable in practice. An alternative solution is to let the (leader of the) shard committee who prepares the heartbeats directly post the acceptance proofs of outgoing intershard transactions to all receiving shards. An inter-committee routing protocol such as the one from RapidChain [34] should then be used to minimize communication overheard. Nevertheless, this may put an undesirable overhead on the committee leaders.

Another possibility is to give rewards to the users including the acceptance proofs to the receiving shards. The users' clients and the committee leaders can still include those proofs, but if they fail to do so, there is an incentive for others to add the missing proofs. Exploring these options in detail is beyond the scope of this paper.

6 INSTANTIATIONS

Our treatment so far has been mostly at an abstract level, where we construct a sharded ledger from several building blocks that can be instantiated with different protocols. This makes our treatment more general and allows for modularity. In this section, we provide some concrete instantiations with data points to evaluate the efficiency of the approach. Note that these are simply examples of how one could instantiate the building blocks, and other options with potentially better performance can also be used.

6.1 Instantiation of Timed Ledger

We suggest to instantiate the timed ledger $\mathcal{F}_{\text{BD-TL}}^{\Delta} = \mathcal{F}_{\text{BD-STL}}^{\Delta,1}$ using HotStuff [32] . This directly provides a ledger via a sequence of blocks containing a list of messages. The blocks are proposed by a leader and validated by other parties signing them. HotStuff includes a mechanism to replace the leader in case the leader fails or is corrupted. To ensure a dishonest leader cannot censor messages, leaders should also be rotated regularly. As discussed in [32], this comes without a significant overhead.

The only feature required by $\mathcal{F}_{\text{BD-TL}}^{\Delta}$ that a basic HotStuff implementation does not provide are the timestamps. These can be added easily assuming (weakly) synchronized clocks: The leader adds the current time as a timestamp to the proposed block, and validators only accept a block if its timestamp is larger than that of the previous block and the timestamp is not in the future. Given that basic HotStuff has bounded latency, it is easy to see that this provides the "bounded timestamps" property.

While all parties in the system must be able to read data from the control chain and post messages there, not all of them are required to act as validators in the consensus. We need that the control chain is always safe and live, i.e., less than $1/3$ corruption can be tolerated on the control chain. As one can see from the data in Table 1, sampling a subset of size 16 037 is sufficient in the worst case to guarantee less than $1/3$ corruption in the subset. Hence, the control chain consensus never needs to be run with more than 16 037 validators.

6.2 Instantiation of Shard Consensus and GearBox

As discussed in Section 4.2, one can use our simple consensus protocol without leader replacement, or alternatively a protocol such as HotStuff [32] that includes leader rotation for the shard consensus. If a protocol without leader replacement is used, this will come at the cost of resampling the whole committee in case of a corrupted leader, but it will potentially provide better performance with an honest leader due to the simplicity of the protocol. In our analysis below, we will use performance data from HotStuff, but count the expected resamplings assuming no leader replacement happens in the protocol, providing a worst-case analysis.

For our sharding protocol, we need a "gearbox" of shard functionalities $\mathcal{F}_{\text{SHARD}}^{s_1,\mathcal{L},\Delta}, \ldots, \mathcal{F}_{\text{SHARD}}^{s_\ell,\mathcal{L},\Delta}$ with increasing committee sizes $s_1 \leq s_2 \leq \cdots \leq s_\ell$ (see Section 5). A decent choice is to have four gears with liveness thresholds 10%, 20%, 25%, and 30%. As one can see in Table 1, this requires committee sizes of 82, 232, 528, and 2264, respectively, in the worst case with up to 30% overall corruption and no upper limit on the total number of parties.

6.3 Instantiation of Randomness Beacon

It is observed in [1] that all currently known sharding protocols require a source of randomness. We can leverage the modular nature of our approach and use an external unbiasable randomness beacon such as DRAND (https://drand.love) in a practical implementation. We could also use the (publicly verifiable) randomness intrinsically generated in an external proof-of-stake (PoS) based blockchain (e.g., Algorand or Cardano), which also use such randomness for committee selection. In these cases, our control chain is not burdened by randomness generation at all.

In order to obtain a self-contained solution departing from a HotStuff based control chain, randomness beacons traditionally used by PoS blockchains can be executed on top of our control chain, since these protocols only rely on the chain for writing information in a black-box manner. One option is using a UC-secure unbiasable randomness beacon based on Publicly Verifiable Secret

Sharing (PVSS) from ALBATROSS [9], which would concretely cost 417120 modular exponentiations per party and about 21.2 megabytes written in total to the control chain in order to generate one unbiasable output. A more efficient solution can be obtained by running a UC-secure [13] Verifiable Random Function (VRF) based randomness beacon on top of the control chain. Concretely, this alternative would cost 2112 modular exponentiations per party and 0.065 megabytes of total communication. However, such VRF based beacons have a bounded bias. As shown in [13], such small bias can be tolerated for selecting committees by slightly increasing the committee sizes. See the full version for more details.

6.4 Efficiency Analysis of Overall Protocol

Number of sampled committees to obtain liveness. Using the numbers from Section 6.2, we have a GearBox with 4 gears, having liveness thresholds between 10% and 30%. If the actual corruption in the overall system is below 10%, the first sampled committee will already have a corruption ratio below the threshold with high probability. That is because the Chernoff bound implies that these corruption ratios are very close with high probability. To achieve liveness (if our protocol with a fixed leader is used), we additionally need that the uniform leader is honest, which is the case with probability at least 90%. Hence, the expected number of sampled committees is less than 2 in this case.

In the worst case, where the actual corruption in total population equals the maximal 30%, we most likely have to move to the last gear with 30% liveness threshold. In that gear, the probability to sample a committee with not more than the expected 30% corruption is roughly 1/2. Additionally, an honest leader is selected with probability at least 70%, i.e., the overall probability to sample a live committee is at least 35%. Hence, the expected total number of sampled committees is less than 6 (3 committees in the lower gears plus 3 committees in the highest gear on expectation).

Latency and throughput. We here give some rough estimates on the concrete performance based on experimental data for a prototype implementation from the HotStuff paper [32]. That paper only provides data for up to 128 nodes. We thus extrapolate that data to get some rough idea on the performance. The numbers here are thus not to be understood as precise data points, but only as rough estimates. The actual performance in a production system with code optimized for the particular setting may deviate significantly.

For the control chain, use the experimental data with 128 byte payload data. Since the control chain only needs to process heartbeat messages consisting of hash values and some metadata, 128 byte are sufficient for control-chain messages. The data available in the paper suggests that in that setting, the latency l in milliseconds of messages with committees of size s can be approximated by the linear function $l = 0.37s + 6$. The throughput t in messages per second is roughly inversely proportional to the latency and can be approximated by the function $t = \frac{2\,400\,000}{l}$. As discussed in Section 6.1, the committee size of the control chain will never have to exceed 16 037. Hence, we can estimate a latency of around 6 seconds and a throughput of about 404 messages per second.

For the shard consensus, we use the data for 1024 bytes of payload data because the shards contain the actual transactions, which may be more complex. As above, we approximate the latency in that setting by the function $l = 0.67s + 20$ and the throughput again by $t = \frac{2\,400\,000}{l}$. Thus, we obtain for the four gears with committee sizes of 82, 232, 528, and 2264, the latency values of 75, 175, 374, and 1537 milliseconds, and the throughput values of 32 000, 13 714, 6417, and 1561 transactions per second, respectively.

Note that transactions can only be considered fully final and intershard transactions can only be imported to the receiving shard once their hashes appear on the control chain. Thus the perceived latency (but not the throughput) is limited by the latency of the control chain. Nevertheless, the numbers show that our approach of minimizing committee sizes can vastly improve performance, even in the worst case with 30% corruption, were we can still operate with shards of size 2264 instead of 16 037 required by prior work.

Scalability and limits on the number of shards. To analyze the scalability of sharding proposals, Avarikioti et al. [1] consider the overhead in terms of communication, storage, and computation. In our protocol, the overhead is limited by ensuring that parties only need to communicate with parties in the same committee (of the shard or the control chain). This is even true for intershard communication when using the solution based on Atomix sketched in Section 5.4. Since HotStuff has linear communication complexity [32] and our committee sizes are bounded, our protocol scales with respect to communication complexity. With respect to space and computational overhead, note that nodes only need to store and validate transactions in the shards they are assigned to, including outgoing or incoming intershard transactions of these shards, plus the data on the control chain. Since the control chain only contains Merkle tree hashes of the transactions and is thus independent of the number of transactions (intra- or intershard), our protocol scales up to a large constant with the number of shards.

For unlimited scalability, the number shards should be able to scale with the size of the total population. This is not the case in our protocol since all shards need to post their heartbeats to the control chain, and thus the number of shards is limited by the throughput of the control chain and the frequency of the heartbeats. Assuming 400 messages per second throughput on the control chain as estimated above, this means that if shards post their heartbeats every 10 seconds, one can have up to 4000 shards, and if heartbeats are only posted once a minute, this allows for 24 000 shards. While this does not allow for unlimited scalability, we believe these numbers to be more than sufficient in practice in the foreseeable future.

If truly unlimited scalability is desired, an approach with multiple control chains, e.g., arranged in a tree, is necessary. We leave the exploration thereof as a direction for future work.

ACKNOWLEDGMENTS

Bernardo David was supported by a grant from Concordium Foundation and by Independent Research Fund Denmark grants number 9040-00399B (TrA^{2}C), number 9131-00075B (PUMA), and number 0165-00079B (Foundations of Privacy Preserving and Accountable Decentralized Protocols). Bernardo Magri performed part of the work while at Concordium Blockchain Research Center, Aarhus University. Jesper Buus Nielsen was partially funded by the Concordium Foundation, the Independent Research Fund Denmark under grant 8021-00366B (BETHE), and the Carlsberg Foundation under the Semper Ardens Research Project CF18-112 (BCM).

REFERENCES

[1] Georgia Avarikioti, Eleftherios Kokoris-Kogias, and Roger Wattenhofer. 2019. Divide and Scale: Formalization of Distributed Ledger Sharding Protocols. *CoRR* abs/1910.10434 (2019). arXiv:1910.10434 http://arxiv.org/abs/1910.10434

[2] Michael Backes and Dennis Hofheinz. 2004. How to Break and Repair a Universally Composable Signature Functionality. In *ISC 2004 (LNCS, Vol. 3225)*, Kan Zhang and Yuliang Zheng (Eds.). Springer, Heidelberg, 61–72.

[3] Christian Badertscher, Ueli Maurer, Daniel Tschudi, and Vassilis Zikas. 2017. Bitcoin as a Transaction Ledger: A Composable Treatment, See [19], 324–356. https://doi.org/10.1007/978-3-319-63688-7_11

[4] Vivek Kumar Bagaria, Sreeram Kannan, David Tse, Giulia C. Fanti, and Pramod Viswanath. 2019. Prism: Deconstructing the Blockchain to Approach Physical Limits. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 585–602. https://doi.org/10.1145/3319535.3363213

[5] Ethan Buchman. 2016. *Tendermint: Byzantine Fault Tolerance in the Age of Blockchains.* Master's thesis. The University of Guelph, Guelph, Ontario, Canada. http://hdl.handle.net/10214/9769

[6] Vitalik Buterin and Virgil Griffith. 2017. Casper the Friendly Finality Gadget. *CoRR* abs/1710.09437 (2017). arXiv:1710.09437

[7] Ran Canetti. 2004. Universally Composable Signature, Certification, and Authentication. In *17th IEEE Computer Security Foundations Workshop, (CSFW-17 2004), 28-30 June 2004, Pacific Grove, CA, USA*. IEEE Computer Society, 219. https://doi.org/10.1109/CSFW.2004.24

[8] Ignacio Cascudo and Bernardo David. 2017. SCRAPE: Scalable Randomness Attested by Public Entities. In *ACNS 17 (LNCS, Vol. 10355)*, Dieter Gollmann, Atsuko Miyaji, and Hiroaki Kikuchi (Eds.). Springer, Heidelberg, 537–556. https://doi.org/10.1007/978-3-319-61204-1_27

[9] Ignacio Cascudo and Bernardo David. 2020. ALBATROSS: Publicly AttestabLe BATched Randomness Based On Secret Sharing. In *ASIACRYPT 2020, Part III (LNCS, Vol. 12493)*, Shiho Moriai and Huaxiong Wang (Eds.). Springer, Heidelberg, 311–341. https://doi.org/10.1007/978-3-030-64840-4_11

[10] Miguel Castro and Barbara Liskov. 1999. Practical Byzantine Fault Tolerance. In *Proceedings of the Third Symposium on Operating Systems Design and Implementation* (New Orleans, Louisiana, USA) *(OSDI '99)*. USENIX Association, USA, 173–186.

[11] Jing Chen and Silvio Micali. 2019. Algorand: A secure and efficient distributed ledger. *Theor. Comput. Sci.* 777 (2019), 155–183. https://doi.org/10.1016/j.tcs.2019.02.001

[12] Phil Daian, Rafael Pass, and Elaine Shi. 2019. Snow White: Robustly Reconfigurable Consensus and Applications to Provably Secure Proof of Stake. In *FC 2019 (LNCS, Vol. 11598)*, Ian Goldberg and Tyler Moore (Eds.). Springer, Heidelberg, 23–41. https://doi.org/10.1007/978-3-030-32101-7_2

[13] Bernardo David, Peter Gazi, Aggelos Kiayias, and Alexander Russell. 2018. Ouroboros Praos: An Adaptively-Secure, Semi-synchronous Proof-of-Stake Blockchain. In *EUROCRYPT 2018, Part II (LNCS, Vol. 10821)*, Jesper Buus Nielsen and Vincent Rijmen (Eds.). Springer, Heidelberg, 66–98. https://doi.org/10.1007/978-3-319-78375-8_3

[14] Thomas Dinsdale-Young, Bernardo Magri, Christian Matt, Jesper Buus Nielsen, and Daniel Tschudi. 2020. Afgjort: A Partially Synchronous Finality Layer for Blockchains. In *SCN 20 (LNCS, Vol. 12238)*, Clemente Galdi and Vladimir Kolesnikov (Eds.). Springer, Heidelberg, 24–44. https://doi.org/10.1007/978-3-030-57990-6_2

[15] Juan A. Garay, Aggelos Kiayias, and Nikos Leonardos. 2015. The Bitcoin Backbone Protocol: Analysis and Applications. In *EUROCRYPT 2015, Part II (LNCS, Vol. 9057)*, Elisabeth Oswald and Marc Fischlin (Eds.). Springer, Heidelberg, 281–310. https://doi.org/10.1007/978-3-662-46803-6_10

[16] Juan A. Garay, Aggelos Kiayias, and Nikos Leonardos. 2017. The Bitcoin Backbone Protocol with Chains of Variable Difficulty, See [19], 291–323. https://doi.org/10.1007/978-3-319-63688-7_10

[17] Yossi Gilad, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nickolai Zeldovich. 2017. Algorand: Scaling Byzantine Agreements for Cryptocurrencies. In *Proceedings of the 26th Symposium on Operating Systems Principles, Shanghai, China, October 28-31, 2017*. ACM, 51–68. https://doi.org/10.1145/3132747.3132757

[18] Jonathan Katz, Ueli Maurer, Björn Tackmann, and Vassilis Zikas. 2013. Universally Composable Synchronous Computation. In *TCC 2013 (LNCS, Vol. 7785)*, Amit Sahai (Ed.). Springer, Heidelberg, 477–498. https://doi.org/10.1007/978-3-642-36594-2_27

[19] Jonathan Katz and Hovav Shacham (Eds.). 2017. *CRYPTO 2017, Part I*. LNCS, Vol. 10401. Springer, Heidelberg.

[20] Aggelos Kiayias, Alexander Russell, Bernardo David, and Roman Oliynykov. 2017. Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol, See [19], 357–388. https://doi.org/10.1007/978-3-319-63688-7_12

[21] Eleftherios Kokoris-Kogias, Philipp Jovanovic, Linus Gasser, Nicolas Gailly, Ewa Syta, and Bryan Ford. 2018. OmniLedger: A Secure, Scale-Out, Decentralized Ledger via Sharding. In *2018 IEEE Symposium on Security and Privacy*. IEEE Computer Society Press, 583–598. https://doi.org/10.1109/SP.2018.000-5

[22] Jae Kwon. 2014. Tendermint: Consensus without Mining. manuscript. https://tendermint.com/static/docs/tendermint.pdf.

[23] Loi Luu, Viswesh Narayanan, Chaodong Zheng, Kunal Baweja, Seth Gilbert, and Prateek Saxena. 2016. A Secure Sharding Protocol For Open Blockchains. In *ACM CCS 2016*, Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi (Eds.). ACM Press, 17–30. https://doi.org/10.1145/2976749.2978389

[24] Christian Matt, Jesper Buus Nielsen, and Søren Eller Thomsen. 2022. Formalizing Delayed Adaptive Corruptions and the Security of Flooding Networks. In *Advances in Cryptology – CRYPTO 2022*. Springer International Publishing, Cham. To appear.

[25] Satoshi Nakamoto. 2009. Bitcoin: A peer-to-peer electronic cash system. manuscript. http://www.bitcoin.org/bitcoin.pdf.

[26] Rafael Pass and Elaine Shi. 2017. Hybrid Consensus: Efficient Consensus in the Permissionless Model. In *31st International Symposium on Distributed Computing, DISC 2017, October 16-20, 2017, Vienna, Austria (LIPIcs, Vol. 91)*, Andréa W. Richa (Ed.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 39:1–39:16. https://doi.org/10.4230/LIPIcs.DISC.2017.39

[27] Ranvir Rana, Sreeram Kannan, David Tse, and Pramod Viswanath. 2022. Free2Shard: Adversary-resistant Distributed Resource Allocation for Blockchains. *Proc. ACM Meas. Anal. Comput. Syst.* 6, 1 (2022), 11:1–11:38. https://doi.org/10.1145/3508031

[28] Abdurrashid Ibrahim Sanka and Ray C.C. Cheung. 2021. A systematic review of blockchain scalability: Issues, solutions, analysis and future research. *Journal of Network and Computer Applications* 195 (2021), 103232. https://doi.org/10.1016/j.jnca.2021.103232

[29] Alberto Sonnino, Shehar Bano, Mustafa Al-Bassam, and George Danezis. 2020. Replay Attacks and Defenses Against Cross-shard Consensus in Sharded Distributed Ledgers. In *IEEE European Symposium on Security and Privacy, EuroS&P 2020, Genoa, Italy, September 7-11, 2020*. IEEE, 294–308. https://doi.org/10.1109/EuroSP48549.2020.00026

[30] Gang Wang, Zhijie Jerry Shi, Mark Nixon, and Song Han. 2019. SoK: Sharding on Blockchain. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies, AFT 2019, Zurich, Switzerland, October 21-23, 2019*. ACM, 41–61. https://doi.org/10.1145/3318041.3355457

[31] Jiaping Wang and Hao Wang. 2019. Monoxide: Scale out Blockchains with Asynchronous Consensus Zones. In *16th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2019*, Jay R. Lorch and Minlan Yu (Eds.). USENIX Association, 95–112.

[32] Maofan Yin, Dahlia Malkhi, Michael K. Reiter, Guy Golan-Gueta, and Ittai Abraham. 2019. HotStuff: BFT Consensus with Linearity and Responsiveness. In *38th ACM PODC*, Peter Robinson and Faith Ellen (Eds.). ACM, 347–356. https://doi.org/10.1145/3293611.3331591

[33] Haifeng Yu, Ivica Nikolic, Ruomu Hou, and Prateek Saxena. 2020. OHIE: Blockchain Scaling Made Simple. In *2020 IEEE Symposium on Security and Privacy, SP 2020, San Francisco, CA, USA, May 18-21, 2020*. IEEE, 90–105. https://doi.org/10.1109/SP40000.2020.00008

[34] Mahdi Zamani, Mahnush Movahedi, and Mariana Raykova. 2018. RapidChain: Scaling Blockchain via Full Sharding. In *ACM CCS 2018*, David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang (Eds.). ACM Press, 931–948. https://doi.org/10.1145/3243734.3243853

[35] Alexei Zamyatin, Mustafa Al-Bassam, Dionysis Zindros, Eleftherios Kokoris-Kogias, Pedro Moreno-Sanchez, Aggelos Kiayias, and William J. Knottenbelt. 2021. SoK: Communication Across Distributed Ledgers. In *Financial Cryptography and Data Security*, Nikita Borisov and Claudia Diaz (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 3–36. https://doi.org/10.1007/978-3-662-64331-0_1