Fork-Resilient Cross-Blockchain Transactions through Algebraic Topology

Dongfang Zhao
University of Nevada, United States
dzhao@unr.edu

Abstract
The cross-blockchain transaction (CBT) serves as a cornerstone for the next-generation, blockchain-based data management systems. However, state-of-the-art CBT models do not address the effect of the possible local fork suspension that might invalidate the entire CBT. This paper takes an algebraic-topological approach to abstract the blockchains and their transactions into simplicial complexes and shows that CBTs cannot complete in either a committed or an aborted status by a $t$-resilient message-passing protocol. This result implies that a more sophisticated model is in need to support CBTs and, thus, sheds light on the future blockchain designs.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Blockchains, distributed transactions, algebraic topology, solvability

Funding Dongfang Zhao: U.S. Department of Energy, contract No. DE-SC0020455

1 Introduction
The cross-blockchain transaction (CBT) serve as a cornerstone for the next-generation, blockchain-based data management systems: the inter-blockchain operations would enable the interoperability among distinct, potentially heterogeneous, blockchains. The state-of-the-art blockchain implementation can only support two-party transactions between two distinct blockchains through the sidechain protocol [4], incurring considerable latency in terms of hours and yet acceptable for the targeting cryptocurrency applications [1]. One recent work by Herlihy [2] notably studied how to support general cross-blockchain operations among an arbitrary number of distinct blockchains through serialized hash locks implemented in smart contracts, assuming a relaxed semantics on the atomicity of operations. Later, Zakhary et al. [5] proposed a 2PC-based protocol to support both parallelism and atomicity. One denominator of these recent works is that they did not consider the possible forks commonly seen in blockchain implementations: a cross-blockchain transaction can still be invalidated if part of its “local” changes on some blockchains is committed but then suspended due to the fork competition within a blockchain. To this end, Zhao [7] proposed a point-set-topological approach to map the fork-induced topological space to the transaction’s topological space—providing a powerful tool to study a fork-resilient CBT through topological equivalence, i.e., homeomorphism.

This paper takes into account the possible local fork suspension and analyzes the completeness of CBTs: whether a CBT can proceed to a completed or aborted final status. Our assumption of the underlying computation model is as follows: a message-passing communication model, an asynchronous timing model, and a crash-failure adversary model with $t$-resilience, $2t < (n + 1)$ where $n \in \mathbb{Z}_+$ and $(n + 1)$ is the total number of nodes. We take an algebraic-topological approach to abstract the blockchains and their transactions into simplicial complexes, and show that CBTs cannot complete in either a committed or an aborted status by a $t$-resilient message-passing protocol. This result, thus, implies that a more sophisticated model is in need to support CBTs in the face of local fork suspension.
2 Fork-Resilient Cross-Blockchain Transactions through Algebraic Topology

We denote the set of distinct blockchains \( C \), whose cardinality is at least two: \( |C| \geq 2 \). Each blockchain is an element \( C_i \in C \), where \( 0 \leq i \leq n = |C| - 1 \). Each blockchain is a list of blocks, each of which is identified by its index \( j \): \( C_i = (v_i^0, \ldots, v_i^l) \). Of note, \( v_i^0 \) is also called the genesis block of \( C_i \) in the literature of blockchains. An \((n + 1)\)-party transaction carried out on \( C \) touches one and only one block at each blockchain. Specifically, an \((n + 1)\)-party global transaction \( T \) can be represented as a set of \( n + 1 \) local transactions \( t_i \), an element in block \( v_i^l \). The granularity of blockchain growth (and suspension when forks occur) is a block, and if we assume our interest is in a single global transaction, we can represent the transaction with the set of involved blocks: \( T = \bigcup_{0 \leq i \leq n} v_i^l \).

We use \( \dim \) and \( \skel^k \) as the function operators of a simplex’s dimension and \( k \)-skeleton, respectively. We use \(|\sigma|\) to denote the geometric realization, i.e. the polygon, of (abstract) simplex \( \sigma \). The \( N \)-time Barycentric and Chromatic subdivisions are denoted \( \text{Bary}^N \) and \( \text{Ch}^N \), respectively. A complete list of notations and definitions in combinatorial topology can be found in [3]. We assume an asynchronous, message-passing communication model among blockchains. We only consider crash failures in this preliminary study and assume the number of faulty nodes \( t \) is less than 50%: \( t < \frac{n+1}{2} \) in blockchains.

2.1 Models

We denote the set of distinct blockchains \( C \), whose cardinality is at least two: \( |C| \geq 2 \). Each blockchain is an element \( C_i \in C \), where \( 0 \leq i \leq n = |C| - 1 \). Each blockchain is a list of blocks, each of which is identified by its index \( j \): \( C_i = (v_i^0, \ldots, v_i^l) \). Of note, \( v_i^0 \) is also called the genesis block of \( C_i \) in the literature of blockchains. An \((n + 1)\)-party transaction carried out on \( C \) touches one and only one block at each blockchain. Specifically, an \((n + 1)\)-party global transaction \( T \) can be represented as a set of \( n + 1 \) local transactions \( t_i \), an element in block \( v_i^l \). The granularity of blockchain growth (and suspension when forks occur) is a block, and if we assume our interest is in a single global transaction, we can represent the transaction with the set of involved blocks: \( T = \bigcup_{0 \leq i \leq n} v_i^l \).

We use \( \dim \) and \( \skel^k \) as the function operators of a simplex’s dimension and \( k \)-skeleton, respectively. We use \(|\sigma|\) to denote the geometric realization, i.e. the polygon, of (abstract) simplex \( \sigma \). The \( N \)-time Barycentric and Chromatic subdivisions are denoted \( \text{Bary}^N \) and \( \text{Ch}^N \), respectively. A complete list of notations and definitions in combinatorial topology can be found in [3]. We assume an asynchronous, message-passing communication model among blockchains. We only consider crash failures in this preliminary study and assume the number of faulty nodes \( t \) is less than 50%: \( t < \frac{n+1}{2} \) in blockchains.

2.2 Task

A task of CBT is represented by a triple \((I, O, \Delta)\), where \( I \) is the input simplical complex, \( O \) is the output simplical complex, and \( \Delta \) is the carrier map \( \Delta : I \rightarrow 2^O \).

Each vertex, i.e., 0-simplex, in \( I \) is a tuple in the form of \((v_i^0, \text{val}_i)\), where \( v_i^0 \) is defined in §2.1 is block-\( i \) at blockchain-\( j \) and \( \text{val}_i \in \{0, 1, \bot\} \). The meaning in the input set is as follows, 0: local transaction not committed, 1: local transaction committed, and \( \bot \): the branch where this block resides is suspended. There is an edge, i.e., 1-simplex, between every pair of vertices in \( I \) except that both vertices are the same block. In general, an \( l \)-simplex in \( I \) comprises a set of distinct \( l + 1 \) blocks as vertices and the higher-dimensional \( k \)-skeletions, \( 1 \leq k \leq l \). Overall, for a \((n + 1)\)-blockchain transaction, the input complex \( I \) comprises \( 3(n + 1) \) vertices and simplices of dimension up to \( n \), i.e., \( \dim(I) = n \).

Each vertex in \( O \) is a tuple \((v_i^l, \text{val}_{\text{out}})\), where \( v_i^l \) is, again, a specific block and \( \text{val}_{\text{out}} \in \{1, 0\} \) with the same semantics defined for \( \text{val}_i \). Indeed, all of local transactions in \( T \) should only end up with either committed \( 1 \) or aborted \( 0 \), respecting the atomicity requirement. The 1-simplices of \( O \) are all the edges connecting vertices whose \( \text{val}_{\text{out}} \)’s are equal, either 0 or 1, among all blocks. Therefore, by definition, the output simplicial complex is disconnected and has two path-connected components: the global transaction is either (i) successfully committed, or (ii) aborted without partial changes.

We now construct the carrier map \( \Delta \), which maps each simplex from \( I \) to a subcomplex of \( O \). Without loss of generality, pick any \( l \)-simplex \( \sigma \in I \), \( 0 \leq l \leq n \), and \( \Delta \) specifies:

- If all the \( \text{val}_i \)’s in \( \sigma \) are 1, then \( \skel^0 \Delta(\sigma) = \{(v, 1) : v \in \skel^0 \sigma\} \).
- If any of the \( \text{val}_i \)’s in \( \sigma \) is \( \bot \), \( \skel^0 \Delta(\sigma) = \{(v, 0) : v \in \skel^0 \sigma\} \).
- For other cases, \( \skel^0 \Delta(\sigma) = \{(v, 0), (v, 1) : v \in \skel^0 \sigma\} \).
- Any \( k \)-face \( \tau \in \sigma \), \( 0 \leq k \leq l \), is similarly mapped.

Note that, by definition, \( \Delta \) is rigid: In any of the above three cases, for any \( l \)-simplex \( \sigma \in I \), \( \dim(\Delta(\sigma)) = l \). Evidently, \( \Delta \) is monotonic: adding new simplices into \( \sigma \) can only
enlarge the mapped subcomplex in $\mathcal{O}$. Furthermore, $\Delta$ is name-preserving as constructed. Therefore, $\Delta$ is a well-defined carrier map from $\mathcal{I}$ to $2^\mathcal{O}$.

2.3 Solvability

- **Definition 1 (Colorless CBT).** A colorless version of CBT, $(\mathcal{I}, \mathcal{O}', \Xi)$, is defined similarly as the general, “colored” CBT, $(\mathcal{I}, \mathcal{O}, \Delta)$, without the block identities on vertices in $\mathcal{O}'$. Also, no identity match is required for the carrier map $\Xi : \mathcal{I} \rightarrow 2^{\mathcal{O}'}$.

- **Lemma 2.** For colorless CBT $(\mathcal{I}, \mathcal{O}', \Xi)$, there does not exist a continuous map $f : |\text{skel}^t\mathcal{I}| \rightarrow |\mathcal{O}'|$ carried by $\Xi$, where $0 < t < \frac{n+1}{2}$.
  
  **Proof Sketch.** The condition $t < \frac{n+1}{2}$ is trivially satisfied by the assumption of crash failures, as the blockchains would have been hard forked otherwise. Since we assume at least one fork suspension would occur, we have $t > 0$. The input simplicial complex $\mathcal{I}$ is pure of dimension $n$ by construction, meaning that $|\text{skel}^t\mathcal{I}|$ is $(t-1)$-connected. Because $t > 0$, $|\text{skel}^t\mathcal{I}|$ is at least 0-connected (i.e., path-connected). As a result, the geometric realization $|\text{skel}^t\mathcal{I}|$ must be connected. However, we know that $\mathcal{O}'$ has two disjoint connected components; so $|\mathcal{O}'|$ is not connected. Therefore, a continuous map carried by $\Xi$ does not exist. 

- **Lemma 3.** Colorless CBT $(\mathcal{I}, \mathcal{O}', \Xi)$ does not have a $t$-resilient message-passing protocol.
  
  **Proof Sketch.** For contradiction, suppose a protocol solves task $(\mathcal{I}, \mathcal{O}', \Xi)$. Then we know that, after $N$ times of Barycentric subdivisions, the carrier map can be written in this form $\Xi(\sigma) = Bary^N \text{skel}^1 \sigma$, for $\sigma \in \mathcal{I}$. That is, there exists a carrier map $\Phi : Bary^N \text{skel}^1 \mathcal{I} \rightarrow 2^{\mathcal{O}'}$. Taking the geometric realizations, we thus have a continuous map $f = |\Phi| : |Bary^N \text{skel}^1 \mathcal{I}| \rightarrow |\mathcal{O}'|$. Note that a subdivision does not change the geometric realization: $|Bary^N \text{skel}^1 \mathcal{I}| = |\text{skel}^1 \mathcal{I}|$. Thus, we have $f : |\text{skel}^1 \mathcal{I}| \rightarrow |\mathcal{O}'|$, a contradiction to Lemma 2.

- **Lemma 4.** A model for colorless CBT $(\mathcal{I}, \mathcal{O}', \Xi)$ reduces to one for general CBT $(\mathcal{I}, \mathcal{O}, \Delta)$.
  
  **Proof Sketch.** Suppose a protocol $P$ solves $(\mathcal{I}, \mathcal{O}, \Delta)$, we simulate $P$ with a protocol $P'$ for $(\mathcal{I}, \mathcal{O}', \Xi)$ as follows. For any $l$-simplex in $\mathcal{O}$, we drop the prefix of the $l$ vertices with map $\varphi : \mathcal{Z} \times V \rightarrow V$ such that $(k, val_{out}) \mapsto (val_{out}) \in \mathcal{O}'$, $0 \leq k \leq l$. The carrier map in the colorless counterpart is $\Xi = \Delta \circ \varphi$, such that for $\sigma \in \mathcal{I}$, $\Xi(\sigma) = \Delta(\varphi(\sigma)) \subseteq \Delta(\sigma)$, i.e., $\Xi$ is carried by $\Delta$.

- **Proposition 5.** For $t < \frac{n+1}{2}$, $(\mathcal{I}, \mathcal{O}, \Delta)$ does not have a $t$-resilient message-passing protocol.
  
  **Proof.** The claim follows directly from Lemma 3 and Lemma 2.
Dongfang Zhao. Topological properties of multi-party blockchain transactions. *CoRR*, abs/2004.01045, 2020. URL: [https://arxiv.org/abs/2004.01045](https://arxiv.org/abs/2004.01045)