A Game-Theoretic Analysis of Cross-Chain Atomic Swaps with HTLCs

Jiahua Xu
UCL Centre for Blockchain Technologies
EPFL

Damien Ackerer
Covario

Alevtina Dubovitskaya
Lucerne University of Applied Sciences and Arts
Swisscom
UCL Centre for Blockchain Technologies

Abstract—To achieve interoperability between unconnected ledgers, hash time lock contracts (HTLCs) are commonly used for cross-chain asset exchange. The solution tolerates transaction failure, and can "make the best out of worst" by allowing transacting agents to at least keep their original assets in case of an abort. Nonetheless, as an undesired outcome, reoccurring transaction failures prompt a critical and analytical examination of the protocol. In this study, we propose a game-theoretic framework to study the strategic behaviors of agents taking part in cross-chain atomic swaps implemented with HTLCs. We study the success rate of the transaction as a function of the exchange rate of the swap, the token price and its volatility, among other variables. We demonstrate that in an attempt to maximize one's own utility as asset price changes, either agent might withdraw from the swap. An extension of our model confirms that collateral deposits can improve the transaction success rate, motivating further research towards collateralization without a trusted third party. A second model variation suggests that a swap is more likely to succeed when agents dynamically adjust the exchange rate in response to price fluctuations.

Index Terms—atomic swap, blockchain, game theory, stochastic price

I. INTRODUCTION

A. Background

An atomic swap is a coordination task where two parties are willing to exchange assets such that either both parties receive each other's original assets upon successful execution, or nothing in the event of failure [1]. Atomic swaps are easily achievable on a single ledger by implementing smart contracts such as automated market-making protocols [2].

For cross-chain asset swaps, the conventional approach is to use a centralized exchange, characterized by high efficiency and transaction speed. However, this requires intermediary fees and trust in the exchange (in terms of privacy and transparency of its matching mechanisms). In addition, centralized exchanges are vulnerable to different kinds of attacks [3], from wallet hacking [4], to DDoS attacks [5]. Over-the-counter (OTC) operations [6], [7] remain frequent for financial transactions. In the financial industry, it is common to use trusted third parties for OTC settlements, such as central clearing counterparties or broker-dealers, that have similar disadvantages as centralized exchanges.

To address some of the issues with centralized exchanges, and transactions with an intermediary in general, distributed exchanges (DEXs) have recently become a popular tool for cross-chain asset exchange [8]–[11]. In such a peer-to-peer

(P2P) environment, the transacting agents typically do not know each other and are thus exposed to malicious behaviors from their counterparty in DEXs that only provide match-making services. Therefore, the major challenge in such settings is to achieve atomicity of the cross-ledger transaction; that is, either the entirety of the transaction is executed, or, in case of failure, nothing occurs [1].

HTLCs,[1] first proposed on a Bitcoin forum by TierNolan [12], have been adopted by some DEXs [9], [10] to achieve atomicity of cross-chain transactions without direct communications between the ledgers. Studying and improving HTLCs has since been of high interest [13]–[15].

A hash time lock contract (HTLC) requires the two agents separately locking their assets on the respective blockchains, using the hash of a secret, generated by one of the users. The assets can then be unlocked upon revealing the preimage of a hash.

The users have accounts (wallets) on two disconnected ledgers ($Chain_a$ and $Chain_b$) executing smart contracts. Consider that Alice wants to send assets on $Chain_a$ to Bob, in exchange for assets on $Chain_b$ from Bob. At Step 1, Alice initiates the transaction by generating a secret (a key) that will be used to unlock the asset transfers later on. She then deploys a smart contract on $Chain_a$, that will lock her assets until time t_a. This contract will transfer to Bob the assets only if the secret generated by Alice is revealed and entered into the smart contract. To verify the secret, Alice reveals its hash as part of the smart contract (cf. Figure 1 Step 1). One important feature of this contract is that after time t_a, should the secret have not been revealed, the smart contract expires and Alice's assets will be unlocked and returned to her wallet.

Next, Bob can verify the contract deployed by Alice on $Chain_a$ (assets, delivery address, etc.) and use the hash submitted by Alice in order to deploy a similar contract on $Chain_b$ (cf. Figure 1 Step 2), specifying the amount he is willing to transfer to Alice and expiry time t_b. Until then Bob's assets are locked on $Chain_b$.

At Step 3, Alice can verify the contract deployed on $Chain_b$, unlock the assets, and initiate their transfer to her wallet by revealing the secret on $Chain_b$. As early as when the secret is revealed in the mempool of $Chain_b$ (even before Alice's
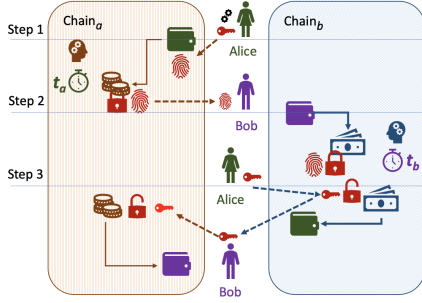

Fig. 1: Hash time lock contract (HTLC).

transfer is confirmed), Bob can use the secret to unlock the assets on $Chain_a$ and complete the cross-ledger transaction. In the best-case scenario, this mechanism enables atomic cross-ledger exchange of the assets without relying on a trusted party and without connection between ledgers. If Alice does not unlock the assets on $Chain_b$ before t_b, then the assets are transferred back to Bob, thus, she has no incentive to reveal the secret as that would allow Bob to execute the smart contract on $Chain_a$ and transfer the assets to his wallet while keeping his assets on $Chain_b$. In turn, once the secret is revealed, Bob's assets are transferred to Alice and he should execute the smart contract on $Chain_a$ immediately in order to complete the transaction, otherwise he transferred his assets without receiving Alice's assets.

B. Contributions

In this paper, we focus on the standard implementation of a cross-ledger atomic swap with hash time lock contracts, HTLCs between two agents who wish to exchange tokens. In our framework, we assume a stochastic token price and that the counterparties can choose to either *continue* or *stop* at any stage of the transaction.

We define a game-theoretic framework to study the agents' behaviors and the transaction outcome in atomic swaps. We focus on the standard protocol with HTLCs, yet the approach can be applied to different setups. The agents' utility functions depend on

1. the transaction outcome (success or failure),
2. the asset price variation (trading profits),
3. the duration of the transaction (locked in the game).

The agents' utility functions are symmetric, but the agents may have a different idiosyncratic willingness to complete the transaction, the so-called *success premium*. By backward induction, we derive the agents' optimal decisions, as well as the transaction success rate as a function of the agreed swap rate, actual token price and its volatility, among other variables. The standard setup has complete information symmetry, and we study the game with uncertainty in counterparties' success premium.

In an extension, we show that if agents would be required to post collateral, everything else being equal, the transaction success rate would be higher. In another extension, we show that if agents can adjust the amount of tokens that they lock in the HTLCs, everything else being equal, the transaction success rate would be higher.

To the best of our knowledge, our work is the first to perform a thorough step-by-step examination of HTLC agents' behavior through a game-theoretic model with numerical simulations.

II. RELATED WORK

Recent years have witnessed a plethora of cross-ledger transaction solutions besides HTLC. Wanchain [16] enables interfacing and asset conversion to the native Wanchain token (Wancoins) in order to perform a cross-ledger asset exchange subsequently. Wanchain implements a privacy protection mechanism through a ring signature scheme [17] and a one-time account mechanism via one-time use wallets created for each transaction. Interledger [18] uses Byzantine notaries to construct a payment chain from sender to recipient over multiple ledgers, and the STREAM Interledger Transport protocol [19] apply packetized payments [20]. Relays [21], sidechains [22], [23], off-chain payment channels [24]–[26], and solutions based on chain relays [27] require building interfaces to such systems, similar to the case of a blockchain-based medium like Wanchain.

Cross-ledger transaction protocols are actively studied by the distributed ledger community: Borkowski et al. [28] surveyed atomic swaps for distributed ledgers, Herlihy provided a first extensive analysis of the scheme and demonstrated that HTLCs are still vulnerable to attacks, such as DDoS or secret hack [29]. Moreover, asset price volatility and malicious behaviour from agents driven by an attempt to maximize financial profits could negatively affect the transaction counterpart. For instance, if Alice becomes inactive before completion of the transaction, then the assets will be blocked on both ledgers [27], [30]. While it can be tolerated by an agent, this can incur significant losses to a counterparty.

To reduce the risk of agents being exposed to adverse behaviour, collateral deposits or transaction fees can be used. In a recent work, Han et al. [30] view atomic swaps as American options (without premium), and discuss "optionality" as a risk imposed by the swap initiator. The initiator can choose at any moment before revealing the secret whether to proceed with the swap or to abort it. To reduce the risk of malicious behaviour by the swap initiator, the authors propose to implement a premium mechanism. In our work, we do not define honest or malicious actors explicitly. Instead, we assume that both actors act rationally in their attempts to maximize their utility, and may appear as either "honest" or "malicious" depending on the movement of the token price. One of the trading protocols that support atomic swaps between parties and exchanges is Arwen. Arwen leverages off-chain RFQ trades and uses an escrow-fee mechanism based on blockchain to incentivize a swap initiator to unlock the coins in a timely manner and address lockup griefing in HTLC [31].

Zamyatin et al. [27] suggested posting collateral at least equal to the assets locked on the blockchain for a trade. The

authors also proposed overcollateralization and a liquidation mechanism to mitigate extreme price fluctuations for both short and long term cross-ledger transactions. While such approach reduces economically rational agents' incentive to misbehave, it is disadvantageous in that if an agent would like to transfer all his assets of one kind, he will be obliged to execute multiple transactions, each with an amount (approximately) equal to half the amount of the assets he currently possesses. In the extension of our model (Section IV), we also propose that both agents place collateral on one of the chains. We then study the impact of the collateralization on the transaction success rate. This analysis allows us to determine the optimal level of collateral for both agents.

Zakhary et al. [32] highlight that even if both participants are honest, the atomicity of HTLC can be violated due to crash failures, preventing smart contract execution before the expiry time of the contract. To address this problem, the authors present all-or-nothing atomic cross-chain commitment protocols and discuss their implementations: the AC^3TW-atomic cross-chain commitment protocol with centralized trusted witness and the AC^3WN-atomic cross-chain commitment protocol that uses another blockchain as a witness network. This approach shares some similarities with the recently proposed notion of a so-called cross-chain deal [33], which, as a generalization of the atomic swap, aims to enhance its expressive power to support various types of commercial practices. Both cross-chain deals and the protocols proposed in [32] are based on exchange of "proofs" or "votes" instead of relying on hashed time locks.

Belotti et al. [34] are among the first to conduct a game-theoretical analysis of cross-chain swaps (including HTLC as presented in [29], [35] and commitment-based protocols AC^3TW and AC^3WN from [32]) and characterise their equilibria. In our work, we focus on the intuition behind the strategic behavior of the participants of HTLC and show that it heavily depends on parameters such as the token price trend and volatility, as well as transaction confirmation time on the employed blockchains.

III. A GAME-THEORETIC ANALYSIS

We seek to establish a model with a good balance of simplicity and fidelity. To this end, we apply reasonable and justifiable assumptions, and set realistic parameter values for numeric demonstration.

A. Basic setup

Alice, denoted by $\mathcal{A}$, wishes to trade some amount of $Token_a$ for 1 unit of $Token_b$, while Bob, denoted by $\mathcal{B}$, is willing to do the opposite. $Token_a$ and $Token_b$ are assets from two different blockchains, namely $Chain_a$ and $Chain_b$. $\mathcal{A}$ and $\mathcal{B}$ intend to swap assets with each other and agree on the exchange rate: $P_* Token_a = 1 Token_b$. Table I summarizes the expected balance change of $\mathcal{A}$'s and $\mathcal{B}$'s assets on the two chains through the swap.

For simplification, we make the following assumptions:

TABLE I: Agents' expected balance change by swap.

| Agent | Expected balance change by swap | |
	on $Chain_a$	on $Chain_b$
Alice ($\mathcal{A}$)	$-P_*\ Token_a$	$+1\ Token_b$
Bob ($\mathcal{B}$)	$+P_*\ Token_a$	$-1\ Token_b$

1. The time it takes for a transaction to be confirmed on $Chain_a$ or $Chain_b$ is constant, equal to τ_a and τ_b respectively.
2. Transaction fees are negligible relative to transaction volume.
3. $Token_a$ is the numéraire, in which $Token_b$ is priced and in which both transacting agents' utilities are measured.
4. $Token_b$'s price (denominated in $Token_a$), P_t, follows a geometric Brownian motion:

$$\ln \frac{P_{t+\tau}}{P_t} = \left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma\left(W_{t+\tau} - W_t\right) \quad (1)$$

where W follows a Wiener process with drift μ and infinitesimal variance σ^2.

5. $\mathcal{A}$ and $\mathcal{B}$ are fully rational, which means they always choose the option that maximizes their utility.
6. $\mathcal{A}$ and $\mathcal{B}$ have the same parametric utility function:

$$U_t^i = \mathbb{E}\left[\frac{(1+\alpha^i\ S)\ V_{t+T_t^i}}{e^{r^i\ T_t^i}}\right] \quad (2)$$

where
i: agent indicator, $i \in \{\mathcal{A}, \mathcal{B}\}$
t: time when the utility is assessed
V: asset value denominated in $Token_a$
T: time until end of game, when no further events directly connected to swap will occur
r: discount rate, $r > 0$
S: success indicator, 1 if the swap succeeds (i.e. agents balance change follows Table I), 0 if fails
α: success premium

7. $\mathcal{A}$ and $\mathcal{B}$ are aware of the value of each other's parameter set, i.e. $\mathcal{A}$ knows $(r^{\mathcal{B}}, \alpha^{\mathcal{B}})$, and $\mathcal{B}$ knows $(r^{\mathcal{A}}, \alpha^{\mathcal{A}})$.

According to Equation (1), given $Token_b$'s price at time t, P_t, the expectation (denoted by $\mathcal{E}$), probability density function (PDF, denoted by $\mathcal{P}$), and cumulative density function (CDF, denoted by $\mathcal{C}$) of its price at $t+\tau$ can be expressed as:

$$\mathcal{E}(P_t, \tau) := \mathbb{E}[P_{t+\tau} \mid P_t] = P_t e^{\mu\tau}$$

$$\mathcal{P}(x, P_t, \tau) := \mathbb{P}[P_{t+\tau} = x \mid P_t] = \frac{e^{-\frac{\left(\ln\frac{x}{P_t} - \left(\mu - \frac{\sigma^2}{2}\right)\tau\right)^2}{2\tau\sigma^2}}}{\sqrt{2\pi}\tau\sigma x}$$

$$\mathcal{C}(x, P_t, \tau) := \mathbb{P}[P_{t+\tau} \leq x \mid P_t] = \frac{\text{erfc}\left(\frac{\ln\frac{x}{P_t} - \left(\mu - \frac{\sigma^2}{2}\right)\tau}{\sqrt{2\tau}\sigma}\right)}{2}$$

where erfc is the complementary error function, and $x > 0$.

All actors act rationally to maximize their utility as defined in Equation (2). Intuitively, actors with a higher success premium α will act more "honestly", i.e. ceteris paribus, they are more likely to continue the game; on the other hand, actors with a lower success premium α may appear "malicious", since ceteris paribus, they are more likely to withdraw from

TABLE II: Notations summary.

Notation	Description
$\mathcal{A}, \mathcal{B}$	Alice, Bob
τ_a, τ_b	Transaction confirmation time on Chain_a, Chain_b
ε_b	Time for an initiated transaction to become discoverable in the mempool of Chain_b
t	Point in time
t_a, t_b	Points in time when the HTLCs on Chain_a, Chain_b expire
P	Price of Token_b denominated in Token_a
P_*	Agreed price of Token_b denominated in Token_a
U	Agent's utility denominated in Token_a
V	Asset value denominated in Token_a
r	Discount rate representing time preference
S	Indicator of whether the swap succeeds ($= 1$) or not ($= 0$)
α	Success premium
T	Time until end of game
μ	Wiener Process drift, see Equation (1)
σ^2	Wiener Process variance, see Equation (1)
$\mathcal{E}(P_t, \tau)$	Expectation of Token_b price at $t + \tau$ given its time-t price P_t
$\mathcal{P}(\cdot, P_t, \tau)$	PDF of Token_b price at $t + \tau$ given its time-t price P_t
$\mathcal{C}(\cdot, P_t, \tau)$	CDF of Token_b price at $t + \tau$ given its time-t price P_t

the game.

B. Decision timeline

Let ε_b denote the time needed to look up a transaction in the mempool of Chain_b after it has been initiated. This time is smaller than the transaction confirmation time on Chain_b, i.e.

$$\varepsilon_b < \tau_b \tag{3}$$

Let $\{t_n\}_{n \in \mathbb{N}}$ denote the points in time when agents have to make a decision, and t_a (t_b) denote the point in time when the HTLC on Chain_a (Chain_b) expires. See Table II for a comprehensive list of notations used in this paper.

According the HTLC protocol described in Section I-A, an atomic swap should work as follows:

1) Agreement and preparation:

t_0: $\mathcal{A}$ and $\mathcal{B}$ agree on the swap conditions, including exchange rate P_*, contract lock expiration time t_a and t_b etc. $\mathcal{A}$ generates a secret and its hash.

2) Action:

t_1: $\mathcal{A}$ uses the hash generated at t_0 to lock P_* Token_a on Chain_a through an HTLC that expires at t_a; thus,

$$t_1 \geq t_0 \tag{4}$$

t_2: $\mathcal{B}$ uses the same hash to lock 1 Token_b on Chain_b through another HTLC that expires at t_b.

$\mathcal{B}$ does so only after verifying that $\mathcal{A}$'s contract is in order and that its deployment has been confirmed on Chain_a; thus,

$$t_2 \geq t_1 + \tau_a \tag{5}$$

t_3: $\mathcal{A}$ uses the secret to unlock the 1 Token_b on Chain_b. $\mathcal{A}$ does so only after verifying that $\mathcal{B}$'s contract is in order and that its deployment has been confirmed on Chain_b; thus,
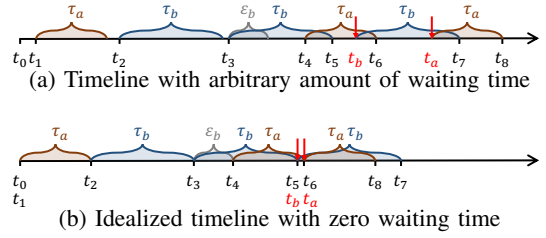
$$t_3 \geq t_2 + \tau_b \tag{6}$$

(a) Timeline with arbitrary amount of waiting time

(b) Idealized timeline with zero waiting time

Fig. 2: Swap timeline as described in Section III-B.

At t_4: $\mathcal{B}$ uses the same secret to unlock the P_* Token_a on Chain_a. $\mathcal{B}$ does so only after seeing the secret been revealed by $\mathcal{A}$ in the mempool of Chain_b; thus

$$t_4 \geq t_3 + \tau_\varepsilon \tag{7}$$

3) Receipt:

Which token an agent receives, and when, depends on the outcome of the swap. If both $\mathcal{A}$ and $\mathcal{B}$ hold on to their agreement by following the steps during the action phase as described in Section III-B2, then the swap succeeds and $\mathcal{A}$ and $\mathcal{B}$ receive tokens at the following two points in time respectively:

t_5: $\mathcal{A}$ receives the 1 Token_b after her transaction is confirmed on Chain_b, and this must take place before the lock contract expires at t_b; thus

$$t_5 = t_3 + \tau_b \leq t_b \tag{8}$$

t_6: $\mathcal{B}$ receives the P_* Token_a after his transaction is confirmed on Chain_a, and this must take place before the lock contract expires at t_a; thus

$$t_6 = t_4 + \tau_a \leq t_a \tag{9}$$

If, however, $\mathcal{A}$ or $\mathcal{B}$ withdraws at any point during the action phase as described in III-B2, then the swap fails and $\mathcal{A}$ and $\mathcal{B}$ have their original tokens returned to them at the following two points in time respectively:

t_7: the HTLC on Chain_b returns $\mathcal{B}$'s original 1 Token_b to him when the time lock expires at t_b, and $\mathcal{B}$ receives the 1 Token_b at t_7; thus

$$t_7 = t_b + \tau_b \tag{10}$$

t_8: the HTLC on Chain_a returns $\mathcal{A}$'s original 1 Token_a to her when the time lock expires at t_a, and $\mathcal{A}$ receives the P_* Token_a at t_8; thus

$$t_8 = t_a + \tau_a \tag{11}$$

Combining (3)–(11), we get

$$\begin{cases} t_0 \leq t_1 < t_1 + \tau_a \leq t_2 < t_2 + \tau_b \leq t_3 < t_3 + \varepsilon_b \\ t_3 + \varepsilon_b < t_3 + \tau_b = t_5 \leq t_b < t_b + \tau_b = t_7 \\ t_3 + \varepsilon_b \leq t_4 < t_4 + \tau_a = t_6 \leq t_a < t_a + \tau_a = t_8 \end{cases} \tag{12}$$

The relationship between different points in time can thus be illustrated as Figure 2a.

C. Zero waiting time

An idealized decision-making time of 0 allows the game to be characterized as a discrete one, where actions can only be taken at a specified, finite set of points in time. In this way, we can express the relationships between the critical points in time as (Figure 2b):

$$
\begin{cases}
t_5 = \overbrace{\underbrace{t_1}_{t_2} + \tau_a}^{t_0} + \tau_b + \tau_b = t_b \quad t_7 = t_b + \tau_b \\
t_6 = \underbrace{\overbrace{t_2 + \tau_b}^{t_3} + \varepsilon_b}_{t_4} + \tau_a = t_a \qquad t_8 = t_a + \tau_a
\end{cases}
\tag{13}
$$

This model simplification can be justified for multiple reasons. Firstly, at the outset of the swap, it should be of both agents' interest to agree on the terms such that the swap can be carried out in a swift manner. From a game-theoretical perspective, lengthening the decision-making time increases optionality for $\mathcal{A}$ at t_3, thus reducing the expected utility for $\mathcal{B}$. $\mathcal{B}$ in turn would postpone his decision at t_2 until as late as possible, to maximize his own optionality and minimize $\mathcal{A}$'s future optionality. $\mathcal{A}$ in turn would wait as long as possible at t_1 to kick off the swap in order to minimize $\mathcal{B}$'s future optionality and maximize her own. In turn at t_0, $\mathcal{B}$ is incentivised to only agree to the shortest time possible to reduce $\mathcal{A}$'s optionality at t_1.

In addition, a long lock time can reduce liquidity for $\mathcal{A}$ and $\mathcal{B}$ collectively. Formally, the negative impact of lock time on agents' utility is captured by their positive discount rate r, as shown in (2). Therefore, the agents would set the contract expiration time as early as possible, to reduce the time of assets being locked, in two ways:

1. An agent can receive his/her counterparty's original asset earlier rather than later in case the swap eventually succeeds. This is because an agent is forced to choose whether to continue or to withdraw immediately each time it is his/her turn to take an action; no immediate action (i.e. waiting) is equivalent to withdrawal since it precludes timely completion of necessary transactions before the contract expiration time, and thus lead to failure of the swap.
2. An agent can get back his/her original asset as soon as the counterparty withdraws from the swap, i.e. when the swap eventually fails.

D. Default parameters

We use backward induction to derive the optimal strategy for each agent. We solve for agents' best strategy numerically and graphically, as we show later in Section III-E that it quickly becomes non-trivial to analytically derive a closed-form expression. To this end, we set the default value of certain parameters as in Table III. We additionally specify the unit of parameters to put the model into perspective. In the following, we discuss the plausibility of selected values for blockchain-specific parameters.

TABLE III: Default value of parameters.

$\alpha^{\mathcal{A}} = 0.3$	$r^{\mathcal{A}} = 0.01$ (/hour)	$\tau_a = 3$ (hour)	$\varepsilon_b = 1$ (hour)	$\mu = 0.002$ (/hour)
$\alpha^{\mathcal{B}} = 0.3$	$r^{\mathcal{B}} = 0.01$ (/hour)	$\tau_b = 4$ (hour)	$P_{t_0} = 2$ (Token$_a$)	$\sigma = 0.1$ (/$\sqrt{\text{hour}}$)

Transaction confirmation time τ: We set the confirmation time on both chains to be in the order of hours. Confirmation time in this paper refers to the time needed to reach transaction finality with a high probability, which typically equals a multiple of the block time. Given the wide adoption of the computationally heavy consensus mechanism—Proof of Work [36], it is to date still common for a blockchain to have an hour-long confirmation time.[2]

Price trend μ: The default value of a positive price trend suggests the deflationary nature of Token$_b$, for example caused by higher levels of token buyback and burn compared to Token$_a$ [37]. In Section III-F, we explore the possibility when Token$_b$ is inflationary, i.e. $\mu < 0$, and when $\mu = 0$.

Volatility σ: The default hourly volatility value of 10% aligns with empirical evidence [38].

In Section III-F, we inspect modelling results with an array of different values for each parameter.

As per assumption, the values of all the parameters displayed in Table III are *common knowledge* (i.e. $\mathcal{A}$ knows, $\mathcal{B}$ knows, and $\mathcal{A}$ knows that $\mathcal{B}$ knows etc).

E. Backward induction

With backward induction, we start from t_4, the last possible action point. We then move backward to an earlier action point each time, assuming the swap is still ongoing (i.e. nobody has withdrawn up until that point). Recall that we use the idealized framework where decision-making time is reduced to a point in time. That is, at each decision-making point in time, agents choose an action from the two-element action set $\{cont, stop\}$.

1) t_4: $\mathcal{B}$ decides whether to unlock Token$_a$ ($cont$) or not ($stop$). Once $\mathcal{A}$ has unlocked Token$_b$ with her pre-generated secret, it does not make sense for $\mathcal{B}$ to withdraw and forgo the locked Token$_a$ (which yields to zero utility). Therefore, as soon as $\mathcal{B}$ sees the secret revealed through $\mathcal{A}$'s transaction from Chain$_b$'s mempool, $\mathcal{B}$ uses the secret to unlock Token$_a$. Thus, $\mathcal{B}$ chooses to continue *with certainty*.

2) t_3: $\mathcal{A}$ decides whether to unlock Token$_b$ ($cont$) or not ($stop$).

cont: $\mathcal{A}$ unlocks the 1 Token$_b$, and receives it at $t_5 = t_3 + \tau_b$, in which case $\mathcal{B}$ gets P_* Token$_a$ at $t_6 = t_3 + (\varepsilon_b + \tau_a)$.

$$
U_{t_3}^{\mathcal{A}}(cont) = \frac{(1+\alpha^{\mathcal{A}})\mathcal{E}(P_{t_3}, \tau_b)}{e^{r^{\mathcal{A}}\tau_b}}
\tag{14}
$$

$$
U_{t_3}^{\mathcal{B}}(cont) = \frac{(1+\alpha^{\mathcal{B}})P_*}{e^{r^{\mathcal{B}}(\varepsilon_b + \tau_a)}}
\tag{15}
$$

---

[2]See e.g. https://support.kraken.com/hc/en-us/articles/203325283-Cryptocurrency-deposit-processing-times
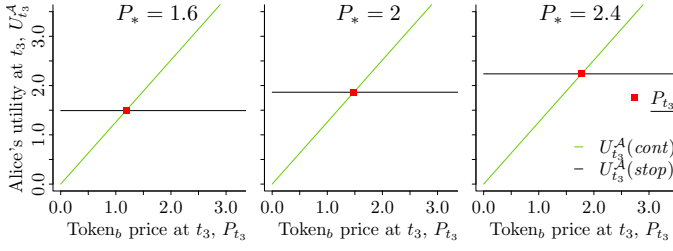
Fig. 3: Alice's utility at t_3, ($U_{t_3}^{\mathcal{A}}$) with different strategies (*cont, stop*), Token_b price values (P_{t_3}), and exchange rate values (P_*).

Fig. 4: Bob's utility at t_2, ($U_{t_2}^{\mathcal{B}}$) with different strategies (*cont, stop*), Token_b price values (P_{t_2}), and exchange rate values (P_*).

stop: $\mathcal{A}$ waives the contract and has the P_* Token_a returned to her at $t_8 = t_3 + (\varepsilon_b + 2\tau_a)$, in which case $\mathcal{B}$ gets 1 Token_b at $t_7 = t_3 + 2\tau_b$.

$$U_{t_3}^{\mathcal{A}}(stop) = \frac{P_*}{e^{r^{\mathcal{A}}(\varepsilon_b + 2\tau_a)}} \tag{16}$$

$$U_{t_3}^{\mathcal{B}}(stop) = \frac{\mathcal{E}(P_{t_3}, 2\tau_b)}{e^{r^{\mathcal{B}} 2\tau_b}} \tag{17}$$

By assumption, $\mathcal{A}$ chooses the option that maximizes her utility. Intuitively, when current Token_b price P_{t_3} is sufficiently large, $\mathcal{A}$ chooses to stop so that she receives Token_b in the end; when P_{t_3} is sufficiently small, $\mathcal{A}$ chooses to continue so that she receives Token_a in the end.

Let $\underline{P_{t_3}}$ denote the cut-off price that equates $U_{t_3}^{\mathcal{A}}(cont)$ and $U_{t_3}^{\mathcal{A}}(stop)$, i.e.

$$\frac{(1+\alpha^{\mathcal{A}})\mathcal{E}(P_{t_3}, \tau_b)}{e^{r^{\mathcal{A}}\tau_b}} = \frac{P_*}{e^{r^{\mathcal{A}}(\varepsilon_b + 2\tau_a)}}$$

$$\underline{P_{t_3}} = \frac{e^{(r^{\mathcal{A}}-\mu)\tau_b - r^{\mathcal{A}}(\varepsilon_b+2\tau_a)}P_*}{1+\alpha^{\mathcal{A}}} \tag{18}$$

Clearly, $\underline{P_{t_3}}$ increases with P_* (see also Figure 3). This is because higher P_* makes the option stop more attractive for $\mathcal{A}$, driving the threshold price $\underline{P_{t_3}}$ higher.

$\mathcal{A}$'s strategy at t_3 can be summarized as:

$$\begin{cases} cont, & P_{t_3} > \underline{P_{t_3}} \\ stop, & P_{t_3} \le \underline{P_{t_3}} \end{cases} \tag{19}$$

3) t_2: $\mathcal{B}$ decides whether to write an HTLC on Chain_b (*cont*) or not (*stop*).

cont: Even when $\mathcal{B}$ chooses to continue, whether the swap eventually succeeds depends on how Token_b price evolves until t_3. Therefore, the utility of $\mathcal{A}$ and $\mathcal{B}$ at t_2 can be expressed with their time-discounted, expected utility at t_3:

$$U_{t_2}^{\mathcal{A}}(cont) = \frac{\int_{\underline{P_{t_3}}}^{\infty} \mathcal{P}(x, P_{t_2}, \tau_b) U_{t_3}^{\mathcal{A}}(cont)dx + \mathcal{C}(\underline{P_{t_3}}, P_{t_2}, \tau_b) U_{t_3}^{\mathcal{A}}(stop)}{e^{r^{\mathcal{A}}\tau_b}} \tag{20}$$

$$U_{t_2}^{\mathcal{B}}(cont) = \frac{[1-\mathcal{C}(\underline{P_{t_3}}, P_{t_2}, \tau_b)]U_{t_3}^{\mathcal{B}}(cont) + \int_0^{\underline{P_{t_3}}} \mathcal{P}(x, P_{t_2}, \tau_b) U_{t_3}^{\mathcal{B}}(stop)dx}{e^{r^{\mathcal{B}}\tau_b}} \tag{21}$$

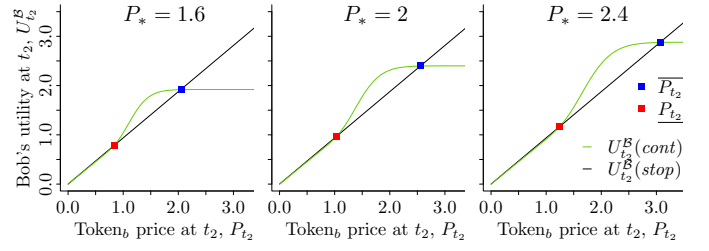*stop:* $\mathcal{B}$ withdraws from the deal and keeps 1 Token_b at t_2. Since $\mathcal{A}$ already locked P_* Token_a, she will receive her original Token_a back at $t_8 = t_2 + (\tau_b + \varepsilon_b + 2\tau_a)$.

$$U_{t_2}^{\mathcal{A}}(stop) = \frac{P_*}{e^{r^{\mathcal{A}}(\tau_b + \varepsilon_b + 2\tau_a)}} \tag{22}$$

$$U_{t_2}^{\mathcal{B}}(stop) = P_{t_2} \tag{23}$$

By assumption, $\mathcal{B}$ also chooses the option that maximizes his utility. In Figure 3, $U_{t_3}^{\mathcal{B}}(cont)$ and $U_{t_3}^{\mathcal{B}}(stop)$ are plotted as a function of P_{t_3}. Intuitively, if current Token_b price P_{t_2} is very low, then although $\mathcal{B}$ would want to swap to receive Token_a, he can safely assume that Alice would not honor the agreement at t_3 anyway, so $\mathcal{B}$ will not bother to continue; if P_{t_2} is very high, then $\mathcal{B}$ would like to keep Token_b and won't swap at all.

Therefore, Bob chooses *cont* when P_{t_2} falls in a feasible range, denoted by $(\underline{P_{t_2}}, \overline{P_{t_2}})$. Figure 4 shows that this range expands and shifts to the higher end with larger P_*.

$\mathcal{B}$'s strategy at t_2 can be summarized as:

$$\begin{cases} cont, & \underline{P_{t_2}} < P_{t_2} \le \overline{P_{t_2}} \\ stop, & P_{t_2} \le \underline{P_{t_2}} \text{ or } P_{t_2} > \overline{P_{t_2}} \end{cases} \tag{24}$$

Note that depending on the value of parameters, the two utility curves might not always have two intersections (disregarding the origin). For example, the lower $\alpha^{\mathcal{B}}$ is, the narrower the feasible range of P_{t_2} is, because $\mathcal{B}$ is less desperate to swap. When $\alpha^{\mathcal{B}}$ is sufficiently small, $U_{t_2}^{\mathcal{B}}(cont) < U_{t_2}^{\mathcal{B}}(stop), \forall P_{t_2} > 0$, and the swap always fails. This will be further discussed in Section III-F.

4) t_1: $\mathcal{A}$ decides whether to initiate the swap by writing an HTLC on Chain_a (*cont*) or not (*stop*).

cont: The utility of $\mathcal{A}$ and $\mathcal{B}$ at t_1 can be expressed by time-discounting their expected utility at $t_2 = t_1 + \tau_a$:

$$U_{t_1}^{\mathcal{A}}(cont) = \frac{\int_{\underline{P_{t_2}}}^{\overline{P_{t_2}}} \mathcal{P}(P_{t_2}, P_{t_1}, \tau_a) U_{t_2}^{\mathcal{A}}(cont)dP_{t_2} + \left[1-\mathcal{C}(\overline{P_{t_2}}, P_{t_1}, \tau_a)+\mathcal{C}(\underline{P_{t_2}}, P_{t_1}, \tau_a)\right] U_{t_2}^{\mathcal{A}}(stop)}{e^{r^{\mathcal{A}}\tau_a}} \tag{25}$$

$$U_{t_1}^{\mathcal{B}}(cont) = \frac{\int_{\underline{P_{t_2}}}^{\overline{P_{t_2}}} \mathcal{P}(P_{t_2}, P_{t_1}, \tau_a) U_{t_2}^{\mathcal{B}}(cont)dP_{t_2} + \int_0^{\underline{P_{t_2}}} \mathcal{P}(P_{t_2}, P_{t_1}, \tau_a) U_{t_2}^{\mathcal{B}}(stop)dP_{t_2}}{e^{r^{\mathcal{B}}\tau_a}} \tag{26}$$
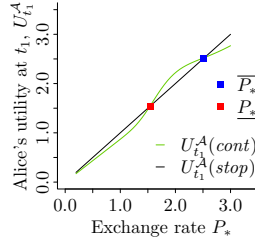
Fig. 5: Alice's utility at t_1, $(U_{t_1}^{\mathcal{A}})$ with different strategies (*cont*, *stop*), and exchange rate values (P_*).
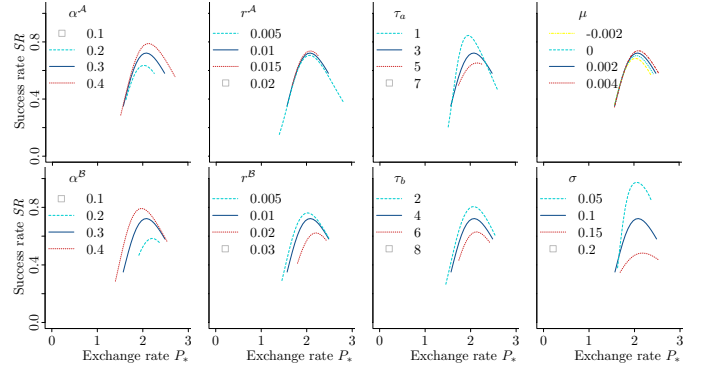

Fig. 6: Swap success rate SR as a function of exchange rate P_* with different parameter values. Default parameter values are set in Table III. Line plots illustrate SR curves with viable parameter values; non-viable values are marked with □.

stop: $\mathcal{A}$ does not initiate the swap and keeps P_* Token$_a$ at t_1. $\mathcal{B}$ also keeps his 1 Token$_b$.

$$U_{t_1}^{\mathcal{A}}(stop) = P_* \tag{27}$$
$$U_{t_1}^{\mathcal{B}}(stop) = P_{t_1} \tag{28}$$

Recall that in our idealized swap game, $\mathcal{A}$ must initiate the swap *immediately* after the terms of swap are agreed upon, as allowing for waiting time only reduces agents' utility. Therefore, $\mathcal{A}$ and $\mathcal{B}$ must agree on a rate P_* that makes $\mathcal{A}$ willing to take the first step. Intuitively, if P_* is too high, then $\mathcal{A}$ would not want to swap; if P_* is too low, then $\mathcal{A}$ understands the high likelihood of fail because $\mathcal{B}$ would not want to continue at t_2, so $\mathcal{A}$ would not start the swap either.

Thus, the exchange rate P_* must lie within a range $(\underline{P_*}, \overline{P_*})$ to ensure the start of the swap (see Figure 5). Using the values from Table III we numerically solve the feasible range as:[3]

$$(\underline{P_*}, \overline{P_*}) = (1.5, 2.5) \tag{29}$$

$\mathcal{A}$'s strategy at t_1 can be summarized as:

$$\begin{cases} cont, & \underline{P_*} < P_* \leq \overline{P_*} \\ stop, & P_* \leq \underline{P_*} \text{ or } P_* > \overline{P_*} \end{cases} \tag{30}$$

F. Success rate

We define the success rate (SR) of a swap to be the likelihood of completion of the swap *after* it has been initiated, i.e. after $\mathcal{A}$ has made the first move at t_1. With the values of all other parameters being fixed (Table III), SR is a function of P_*, and can be expressed as:

$$SR(P_*) = \int_{\underline{P_{t_2}}(P_*)}^{\overline{P_{t_2}}(P_*)} \mathcal{P}\left(x, P_{t_1}, \tau_a\right) \left[1 - \mathcal{C}(\underline{P_{t_3}}(P_*), x, \tau_b)\right] dx,$$
$$\underline{P_*} < P_* \leq \overline{P_*} \tag{31}$$

In Figure 6, we show how success rate SR changes with the exchange rate P_*. SR curves with the default parameter setting (Table III) are plotted in blue line (—), which are compared with SR curves with different parameter values. Irrespective of the parameter values, the $SR(P_*)$ curve is always concave, with the SR-maximizing point residing between $\underline{P_*}$ and $\overline{P_*}$. As suggested in Section III-E, this is because overly low P_* reduces the likelihood of continuation at t_3 and t_2, while overly high P_* reduces the likelihood of continuation at t_2.

---

[3]Note that $P_{t_1} = P_{t_0}$ since $t_1 = t_0$ is assumed in (13).

Next, we discuss how the value setting of other parameters affects the success rate.

1) Success premium α: The parameter success premium describes the excess utility that an agent receives when the swap succeeds. The parameter captures not only the excess utility an agent gains from possessing the counterparty's token over his/her own token, but also the utility of guarding his/her reputation. That is to say, the more an agent cares about honoring an agreement, the higher α will be. As shown in Figure 6, ceteris paribus, higher α leads to higher SR. This is true for both $\alpha^{\mathcal{A}}$ and $\alpha^{\mathcal{B}}$. In addition, higher α renders a bigger feasible range of P_*. Note that when α is too small (either with $\mathcal{A}$ or $\mathcal{B}$), the swap would never be initiated.

2) Time preference r: The parameter time preference describes an agent's *impatience* level. Larger r suggests a higher degree of impatience, i.e. possessing an asset *right now* is more valuable for the agent than obtaining it *later*. As an HTLC swap requires an asset-locking period, a certain degree of *patience* is needed for both $\mathcal{A}$ and $\mathcal{B}$ to enter the agreement. Thus, as shown in Figure 6, larger r results in a narrower viable range of values for P_*; exceedingly high r renders any P_* value infeasible, i.e. the swap would never be initiated.

3) Transaction confirmation time τ: With the presence of time preference ($r^{\mathcal{A}}, r^{\mathcal{B}} > 0$), longer transaction confirmation time, either on Chain$_a$ or Chain$_b$, reduces agents' utility in engaging in a swap. Therefore, higher τ_a or τ_b shrinks the viable range of P_*. When P_* is always chosen optimally (as to maximize SR), lower τ_a or τ_b increases SR.

4) Price trend μ and volatility σ: Figure 6 shows that, ceteris paribus, higher degree of upward price trend of Token$_b$ increases SR. In contrast, higher volatility reduces max SR.

IV. MODEL EXTENSION

In this section, we expand on our basic model described in Section III and discuss two derivations.

A. HTLC with collateral

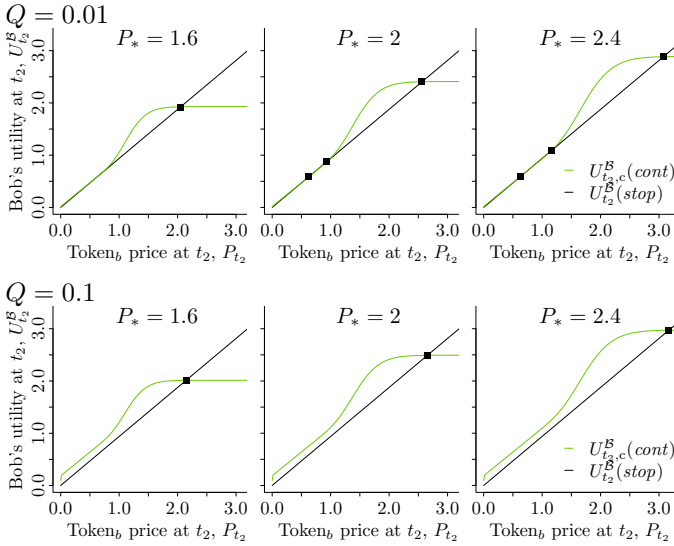In this section, we discuss an HTLC game where both agents place collateral into a smart contract before the actual

Fig. 7: Bob's utility at t_2, $(U_{t_2}^{\mathcal{B}})$. Indifference points between *cont* and *stop* are marked with ■.

.

Fig. 8: Alice's and Bob's utility at t_1 $(U_{t_1}^{\mathcal{A}}, U_{t_1}^{\mathcal{B}})$. Each agent's indifference points between *cont* and *stop* are marked with ■.

swap. All assumptions from Section III-A with the exception of Assumption 6. remain unchanged. We assume additionally:

1) $\mathcal{A}$ and $\mathcal{B}$ move an allowance to a trusted smart contract on Chain$_a$ in order to charge each of them simultaneously the same amount of collateral, Q Token$_a$, before the swap;

2) the smart contract is connected to an Oracle which observes the transaction outcomes on Chain$_a$ and Chain$_b$;

3) if the swap succeeds, the Oracle transfers to each agent their original collateral; if an agent chooses *stop* at any point during the swap, the other agent receives both agents' collateral from the Oracle;

4) agent's utility function is as follows:

$$U_{t,\mathrm{c}}^i = \mathbb{E}\left[\frac{(1+\alpha^i\,S)\,V_{t+T_t^i}}{e^{r^i\,T_t^i}} + \frac{C_t^i}{e^{r^i\,t_c}}\right] \qquad (32)$$

where

C: value of collateral to be received back,

t_c: time until receiving the collateral.

We use subscript "c" only when an expression differs from the one in the basic setup (Section III-E).

This setup is theoretical as there is presently no Oracle service that would be able to monitor the actions as described, to the best of our knowledge. Yet, the Bisq framework is similar in spirit with the key difference that a human arbitrator replaces the Oracle. The goal of this section is to study the impact of collateralization on the agents' behaviors and thus on the transaction outcome. A new atomic swap protocol with collateral will be discussed in a follow-up work.

We again employ backward induction to derive agents' utility-maximizing strategy.

1) t_4: At this point, if $\mathcal{A}$ has released the secret, the Oracle will determine that $\mathcal{A}$ has fulfilled all her obligations and releases her collateral Q Token$_a$ at t_4. Thus, $\mathcal{A}$ will receive
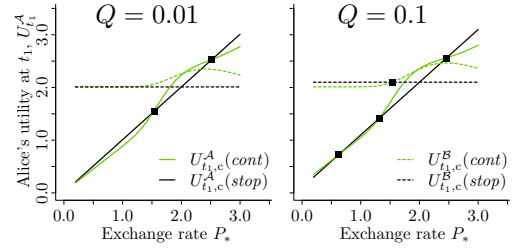
Q Token$_a$ at $t_4 + \tau_a$. Same as in the basic scenario described in Section III-E1, $\mathcal{B}$ chooses to unlock Token$_a$ (*cont*) with certainty. If $\mathcal{A}$ has not released the secret, the Oracle will transfer $\mathcal{A}$'s collateral to $\mathcal{B}$.

2) t_3: At this point, $\mathcal{B}$ has written an HTLC on Chain$_b$ as agreed, and hence has no further chance for foul play. The Oracle thus return's $\mathcal{B}$'s collateral, and $\mathcal{B}$ receives Q Token$_a$ at $\tau_3 + \tau_\alpha$. If now $\mathcal{A}$ waives the contract, then her utility equals $U_{t_3}^{\mathcal{A}}(stop)$ as described in Equation (16) and her collateral will be transferred to be in the next step. If $\mathcal{A}$ chooses to unlocks the 1 Token$_b$, the swap succeeds and she receives Token$_b$ at t_5 plus her collateral Q Token$_a$ at $\tau_4 + \tau_\alpha$.

Therefore, $\mathcal{A}$ chooses *cont* over *stop* if:

$$U_{t_3}^{\mathcal{A}}(stop) < U_{t_3}^{\mathcal{A}}(cont) + \frac{Q}{e^{r^{\mathcal{A}}(\varepsilon_b+\tau_a)}}$$
$$P_{t_3} > \left(\frac{P_*}{e^{r^{\mathcal{A}}(\tau_\epsilon+2\tau_a)}} - \frac{Q}{e^{r^{\mathcal{A}}(\varepsilon_b+\tau_a)}}\right)\frac{e^{(r^{\mathcal{A}}-\mu)\tau_b}}{1+\alpha^{\mathcal{A}}} \qquad (33)$$

Since $P_{t_3} \geq 0$, we express P_{t_3}'s lower bound as:

$$\underline{P_{t_3,\mathrm{c}}} = \frac{e^{(r^{\mathcal{A}}-\mu)\tau_b}}{1+\alpha^{\mathcal{A}}} \cdot \max\left(\frac{P_*}{e^{r^{\mathcal{A}}(\tau_\epsilon+2\tau_a)}} - \frac{Q}{e^{r^{\mathcal{A}}(\varepsilon_b+\tau_a)}}, 0\right)$$

3) t_2: $\mathcal{B}$ decides whether to write an HTLC on Chain$_b$ (*cont*) or not (*stop*).

cont: As discussed in Section IV-A2, if $\mathcal{B}$ chooses *cont* at this point, the Oracle will determine at t_3 that $\mathcal{B}$ has fulfilled his obligations, and will return his collateral at that time.

In addition, $\mathcal{B}$ expects that at t_4, $\mathcal{A}$ will honour the deal when $P_{t_3} > \underline{P_{t_3,\mathrm{c}}}$ in which case $\mathcal{B}$ gets P_* Token$_a$ at t_6, and waive the deal otherwise in which case Bob gets 1 Token$_b$ at t_7 plus $\mathcal{A}$'s collateral Q at $t_3 + \tau_a$.

Therefore, the utility of $\mathcal{A}$ and $\mathcal{B}$ utility at t_2 is:

$$U_{t_2,\mathrm{c}}^{\mathcal{A}}(cont) = \frac{\begin{array}{c}\int_{\underline{P_{t_3,\mathrm{c}}}}^{\infty}\mathcal{P}\left(x,P_{t_2},\tau_b\right)\left(U_{t_3}^{\mathcal{A}}(cont)+\frac{Q}{e^{r^{\mathcal{A}}(\varepsilon_b+\tau_a)}}\right)dx \\ +\mathcal{C}\left(\underline{P_{t_3,\mathrm{c}}},P_{t_2},\tau_b\right)U_{t_3}^{\mathcal{A}}(stop)\end{array}}{e^{r^{\mathcal{A}}\tau_b}} \qquad (34)$$

$$U_{t_2,\mathrm{c}}^{\mathcal{B}}(cont) = \frac{\begin{array}{c}\frac{Q}{e^{r^{\mathcal{B}}\tau_a}}+\left[1-\mathcal{C}(\underline{P_{t_3,\mathrm{c}}},P_{t_2},\tau_b)\right]U_{t_3}^{\mathcal{B}}(cont) \\ +\int_{0}^{\underline{P_{t_3}}}\mathcal{P}(x,P_{t_2},\tau_b)\left(U_{t_3}^{\mathcal{B}}(stop)+\frac{Q}{e^{r^{\mathcal{B}}(\varepsilon_b+\tau_a)}}\right)dx\end{array}}{e^{r^{\mathcal{B}}\tau_b}} \qquad (35)$$

stop: $\mathcal{B}$ withdraws from the deal and keep 1 Token$_b$. $\mathcal{B}$'s utility is the same as Equation (23). The swap stops due to $\mathcal{B}$'s foul play. The Oracle thus releases both agents collateral, $2Q$ in total, to $\mathcal{A}$ at t_3, who will receive the fund at $\tau_3 + \tau_a$.

Intuitively, if P_{t_2} is too high, then $\mathcal{B}$ would like to keep the valuable Token$_b$ and wouldn't want to swap, and $\mathcal{B}$ would therefore choose *stop*; if P_{t_2} is too low (say, close to zero),
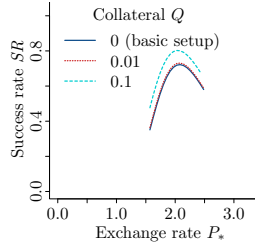
Fig. 9: Swap success rate SR as a function of exchange rate P_* with different collateral values Q.

(a) Optimal Token$_b$ amount X_* for Bob to lock based on P_{t_2} and Token$_a$ amount P_* locked at t_1.

(b) Alice's utility at t_1, $U_{t_1,x}^{\mathcal{A}}$ as a function of the amount of Token$_a$ to lock P_*.

Fig. 10: Agents' decision making with uncertain swap rate.

forfeiting the valuable collateral to keep the worthless Token$_B$ would not be sensible—even if $\mathcal{A}$ chooses not to reveal secret in the next step, $\mathcal{B}$ would at least be able to receive the collateral; therefore, $\mathcal{B}$ would choose $cont$.

Thus, Equation $U_{t_2,c}^{\mathcal{B}}(cont) = U_{t_2}^{\mathcal{B}}(stop)$ has an *odd* number of roots. Figure 7 shows that there can be 1 or 3 intercepts between curve $U_{t_2,c}^{\mathcal{B}}(cont)$ and $U_{t_2}^{\mathcal{B}}(stop)$, depending on the value of Q and P_*. Define set $\mathfrak{P}_{t_2}$ such that:

$$P_{t_2} \in \mathfrak{P}_{t_2} \iff U_{t_2,c}^{\mathcal{B}}(cont) > U_{t_2}^{\mathcal{B}}(stop)$$

Hence, $\mathcal{B}$ would choose $cont$ if and only if P_{t_2} falls in $\mathfrak{P}_{t_2}$.

4) t_1: $\mathcal{A}$ and $\mathcal{B}$ simultaneously make the decision on whether to engage in the swap ($cont$) or not ($stop$).

cont: The utility of $\mathcal{A}$ and $\mathcal{B}$ at t_2 can be expressed with their time-discounted, expected utility at $t_2 = t_1 + \tau_a$:

$$U_{t_1,c}^{\mathcal{A}}(cont) = \frac{\int_{x \in \mathfrak{P}_{t_2}} \mathcal{P}(x,P_{t_1},\tau_a) U_{t_2,c}^{\mathcal{A}}(cont) dx + \left(\int_{x \in \mathbb{R}^+ \setminus \mathfrak{P}_{t_2}} \mathcal{P}(x,P_{t_1},\tau_a) dx \right) \left(U_{t_2}^{\mathcal{A}}(stop) + \frac{2Q}{e^{r^{\mathcal{A}}(\tau_b + \tau_a)}} \right)}{e^{r^{\mathcal{A}} \tau_a}} \quad (36)$$

$$U_{t_1,c}^{\mathcal{B}}(cont) = \frac{\int_{x \in \mathfrak{P}_{t_2}} \mathcal{P}(x,P_{t_1},\tau_a) U_{t_2,c}^{\mathcal{B}}(cont) dx + \int_{x \in \mathbb{R}^+ \setminus \mathfrak{P}_{t_2}} \mathcal{P}(x,P_{t_1},\tau_a) U_{t_2}^{\mathcal{B}}(stop) dx}{e^{r^{\mathcal{A}} \tau_a}} \quad (37)$$

stop: $\mathcal{A}$ and $\mathcal{B}$ decide not to engage in the swap so that they can keep their original token and the collateral. Thus:

$$U_{t_1,c}^{\mathcal{A}}(stop) = U_{t_1}^{\mathcal{A}}(stop) + Q = P_* + Q \quad (38)$$
$$U_{t_1,c}^{\mathcal{B}}(stop) = U_{t_1}^{\mathcal{B}}(stop) + Q = P_{t_1} + Q \quad (39)$$

Define set $\mathfrak{P}_*^{\mathcal{A}}$ and $\mathfrak{P}_*^{\mathcal{B}}$ such that:

$$P_{t_2} \in \mathfrak{P}_{t_2}^{\mathcal{A}} \iff U_{t_2,c}^{\mathcal{A}}(cont) > U_{t_2,c}^{\mathcal{A}}(stop)$$
$$P_{t_2} \in \mathfrak{P}_{t_2}^{\mathcal{B}} \iff U_{t_2,c}^{\mathcal{B}}(cont) > U_{t_2,c}^{\mathcal{B}}(stop)$$

Hence, the exchange rate P_* must be in $\mathfrak{P}_* = \mathfrak{P}_{t_2}^{\mathcal{A}} \cup \mathfrak{P}_{t_2}^{\mathcal{B}}$, since otherwise agents' external utility ($stop$) exceeds the expected utility from the swap ($cont$) and the swap would not be initiated (see Figure 8).

The swap's success rate SR can thus be expressed as:

$$SR(P_*) = \int_{x \in \mathfrak{P}_{t_2}} \mathcal{P}(x,P_{t_1},\tau_a) \left[1 - \mathcal{C}(\underline{P_{t_3,c}}(P_*), x, \tau_b) \right] dx,$$
$$P_* \in \mathfrak{P}_* \quad (40)$$

Figure 9 shows SR increases with collateral amount Q. This is because higher Q allows for larger price movement,

by expanding the feasible Token$_b$ price range at both t_2 (see Figure 7) and t_1 (see Equation (33)).

B. Uncertain exchange rate

In this model variation, we discuss an HTLC game in which agents not only choose between $cont$ and $stop$, but also the exact amount of funds to lock in: P_* Token$_a$ at t_1 and X Token$_b$ t_2, respectively. This renders the actual exchange rate uncertain at the outset of the game. We apply subscript "$_x$" for expressions different from their counterpart in the baseline model.

1) t_4: Same as Section III-E1

2) t_3: Similar to Section III-E2, there exists a minimum feasible Token$_b$ price, expressed as $\underline{P_{t_3,x}}(X)$ since it depends on X, the actual transferred amount at t_2. It can be easily derived that

$$\underline{P_{t_3,x}}(X) = \underline{P_{t_3}}/X, \text{ with } \underline{P_{t_3,x}}(0) = \infty \quad (41)$$

where $\underline{P_{t_3}}$ is expressed in Equation (18).

3) t_2: $\mathcal{B}$ determines the value of $X \geq 0$ that maximizes his excess utility, namely the utility he obtains by proceeding with locking X less the utility he keeps by retaining X Token$_b$.

$$U_{t_2,x}^{\mathcal{A}}(X) = \frac{X \int_{\underline{P_{t_3,x}}(X)}^{\infty} \mathcal{P}\left(x,P_{t_2},\tau_b\right) U_{t_3}^{\mathcal{A}}(cont) dx + \mathcal{C}(\underline{P_{t_3,x}}(X),P_{t_2},\tau_b) U_{t_3}^{\mathcal{A}}(stop)}{e^{r^{\mathcal{A}} \tau_b}} \quad (42)$$

$$U_{t_2,x}^{\mathcal{B}}(X) = \frac{\left[1 - \mathcal{C}(\underline{P_{t_3,x}}(X),P_{t_2},\tau_b) \right] U_{t_3}^{\mathcal{B}}(cont) + X \int_{0}^{\underline{P_{t_3,x}}(X)} \mathcal{P}\left(x,P_{t_2},\tau_b\right) U_{t_3}^{\mathcal{B}}(stop) dx}{e^{r^{\mathcal{B}} \tau_b}} - X P_{t_2} \quad (43)$$

Denote the optimal X as $X_*(P_{t_2})$, a function of P_{t_2}. Thus,

$$X_*(P_{t_2}) = \underset{X \geq 0}{\operatorname{argmax}} U_{t_2,x}^{\mathcal{B}}(X) \quad (44)$$

Figure 10a shows at any given amount of Token$_a$ that $\mathcal{A}$ has locked at t_1, P_*, the optimal amount of Token$_b$ that $\mathcal{B}$ should lock first increases and then decreases with P_{t_2}.

This aligns with the intuition that when P_{t_2} is too low, i.e. Token$_b$ is valueless, then the likelihood of $\mathcal{A}$'s withdrawal at t_3 is high, making $\mathcal{B}$ unwilling to lock in big funds at t_2; when P_{t_2} is too high, i.e. Token$_b$ is very valuable, then the amount of Token$_b$ that $\mathcal{B}$ needs to commit also becomes low to make the deal worthwhile for himself. Generally at a given P_{t_2}, X_* increases with P_*, reflecting a degree of fairness of the game.
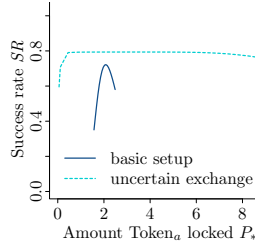
Fig. 11: Swap success rate SR as a function of initial amount of Token_a locked in by Alice P_*.

4) t_1: $\mathcal{A}$ takes into account the fact that $\mathcal{B}$ will choose an amount of Token_b to lock in at t_2 that maximizes his own utility, based on Token_b price at t_2 and the amount of Token_a that she commits, P_*.

$\mathcal{A}$ thus chooses the value of P_* that maximizes her excess utility, namely the utility she obtains by entering the swap with P_* Token_a in excess of the utility she keeps by retaining P_* Token_a. Therefore,

$$U_{t_1,\text{x}}^{\mathcal{A}}(P_*) = \frac{\int_0^\infty \mathcal{P}(P_{t_2}, P_{t_1}, \tau_a) U_{t_2,\text{x}}^{\mathcal{A}}(X_*(P_{t_2})) dP_{t_2}}{e^{r^{\mathcal{A}} \tau_a}} - P_* \tag{45}$$

Similar to Figure 5, Figure 10b shows that the excess utility first increases and then decreases with P_*. $\overline{P_*}$ represents the amount that maximizes $U_{t_1,\text{x}}^{\mathcal{A}}$. Nevertheless, $\mathcal{A}$ might only be able to afford a lesser amount due to a possible budget constraint on her side. $\underline{P_*}$ represents the lowest possible amount that $\mathcal{A}$ needs to enter for a non-negative excess utility.

We can express the swap's success rate SR under uncertain exchange rate as:

$$SR(P_*) = \int_0^\infty \mathcal{P}(x, P_{t_1}, \tau_a) \left[1 - \mathcal{C}(\underline{P_{t_3,\text{x}}}(X(x)), x, \tau_b) \right] dx, \tag{46}$$

Figure 11 compares the success rate between the basic setup and the scenario with uncertain exchange rate. Interestingly, absence of pre-determined interest rate boosts the success rate.

V. DISCUSSION

A. Interpretation of findings

Multiple findings can be drawn from our analysis that are relevant for real-world applications. It has previously been mentioned that the agent completing the transaction receives a free *American option*, meaning that she has the choice to complete the transaction, or not, based on whether the asset price changes at her advantage. However, in this work, we show that the other agent (not only the swap initiator) may also leave the game midway, incentivized by a potentially higher financial gain. This scenario has thus far been neglected in the literature.

Our analysis also suggests that the collateral deposits can be dynamically adjusted depending on the terms of the swap (e.g. exchange rate) and optimization goal (e.g. maximizing utility, or maximizing success rate).

In the last model extension we show that the likelihood of completing the transaction is higher when the agents dynamically adjust the exchange rate to account for token price fluctuations. This is because an arbitrarily fixed exchange rate takes away the flexibility for agents to adjust their commitment based on the latest market condition (price change in our case).

B. Limitations and future work

Our work motivates multiple future research directions.

Firstly, simulation studies can be performed based on our model framework and its derivation using real market data.

Secondly, trustless protocols supporting collateral deposit without a third party can be designed. To date, collaterals are typically deposited at a trusted third party in practice which is then responsible for resolving payment disputes. For instance, Bisq [8], an information platform for quotes and P2P transactions with arbitrators, uses a postage of collateral with possible intervention of an arbitrator, thus providing only a limited level of "distributiveness" and still requiring some trust in the arbitration system. Note that transactions executed via Bisq require a collateral deposit and an arbitrator fee. Discussion with community members revealed that 3-5% of transactions fail and go to arbitration, and that this percentage increases during periods of higher market volatility.

Thirdly, HTCL protocols can be further improved. HTLCs have known limitations [27], [39], including strong assumptions required to maintain security, interactiveness, exclusiveness to public blockchains (for public mempools), and the need for synchronizing clocks between blockchains and temporal locking of assets.

Lastly, a more realistic and sophisticated setup can be brought into our framework. For example, future models may incorporate different risk-free rates for the two exchanged tokens, which resembles the settings of the Garman Kohlhagen model. In addition, blockchain transaction fees or coin stacking (similar to earning dividends or interest on a locked-in asset) may have an impact on agents' actions. Our model can also be extended to consider repeated games, stochastic individual utility, success premium as a random variable, etc.

VI. CONCLUSION

We introduce a game-theoretic approach to model agent behaviors in cross-ledger transactions. This allows us to study the viability and sensitivity of different protocols with respect to environment variables: agents' knowledge and utility functions, and the price dynamics. We study in-depth the atomic swap as implemented with hash time lock contracts, for which we derived the success rate of a transaction. In particular, we showed that both transacting counterparties can rationally decide to walk away from the transaction, and at different times. This is a more realistic setup that relaxes the assumptions from previous works [30], [40] that only the swap initiator can benefit from price variation and act upon it. A sensitivity analysis reveals that price volatility significantly affects the success rate of the transaction.

We extend our basic setup along two directions. First, we show that introducing collateral deposit, in a purely theoretical way, increases the success rate of the transaction. Second, we sketch out the scenario where agents are uncertain about the amount of funds their counterparty intends to commit.

Two important conclusions are therefore that cross-ledger atomic swap trustless protocols could benefit from the use of disciplinary mechanisms, such as collateral deposit, and that allowing agents to dynamically adjust the swap amount can increase the success rate.

REFERENCES

[1] H. Garcia-Molina, "Using semantic knowledge for transaction processing in a distributed database," *ACM Transactions on Database Systems (TODS)*, vol. 8, no. 2, pp. 186–213, 1983.

[2] J. Xu, N. Vavryk, K. Paruch, and S. Cousaert, "SoK: Decentralized Exchanges (DEX) with Automated Market Maker (AMM) protocols," 3 2021. [Online]. Available: http://arxiv.org/abs/2103.12732

[3] T. Moore and N. Christin, "Beware the middleman: Empirical analysis of Bitcoin-exchange risk," in *International Conference on Financial Cryptography and Data Security*. Springer, 2013, pp. 25–33.

[4] "Details of $5 Million Bitstamp Hack Revealed." [Online]. Available: https://www.coindesk.com/unconfirmed-report-5-million-bitstamp-bitcoin-exchange

[5] "Crypto Exchange Bitfinex Bounces Back after a DDoS Attack." [Online]. Available: https://www.ccn.com/crypto-exchange-bitfinex-bounces-back-after-a-ddos-attack

[6] "itBit." [Online]. Available: https://www.itbit.com/otc

[7] "HiveEx-Large Volume Cryptocurrency OTC Brokerage." [Online]. Available: https://www.hiveex.com

[8] "Bisq White Paper." [Online]. Available: https://docs.bisq.network/exchange/whitepaper.html

[9] "Decred-compatible cross-chain atomic swapping," 2018. [Online]. Available: https://github.com/decred/atomicswap/

[10] Komodo, "Komodo's Atomic-Swap Powered, Decentralized Exchange: Barterdex," 2021. [Online]. Available: https://docs.komodoplatform.com/whitepaper/chapter6.html

[11] "0x White paper." [Online]. Available: https://0x.org/pdfs/0x_white_paper.pdf

[12] "Bitcoin Wiki: Atomic cross-chain trading." [Online]. Available: https://en.bitcoin.it/wiki/

[13] J. Kirsten and H. Davarpanah, "Anonymous Atomic Swaps Using Homomorphic Hashing," *Available at SSRN 3235955*, 2018.

[14] G. Zyskind, C. Kisagun, and C. Fromknecht, "Enigma Catalyst: A machine-based investing platform and infrastructure for crypto-assets," 2018.

[15] J. A. Liu, "Atomic Swaptions: Cryptocurrency Derivatives," *arXiv preprint arXiv:1807.08644*, 2018.

[16] "Building Super Financial Markets for the New Economy." [Online]. Available: https://wanchain.org/files/Wanchain-Whitepaper-EN-version.pdf

[17] R. L. Rivest, A. Shamir, and Y. Tauman, "How to leak a secret: Theory and applications of ring signatures," in *Theoretical Computer Science*. Springer, 2006, pp. 164–186.

[18] S. Thomas and E. Schwartz, "A protocol for interledger payments," 2015. [Online]. Available: https://interledger.org/interledger.pdf

[19] Interledger, "STREAM: A Multiplexed Money and Data Transport for ILP," 2020. [Online]. Available: https://github.com/interledger/rfcs/blob/master/0029-stream/0029-stream.md

[20] A. Dubovitskaya, D. Ackerer, and J. Xu, "A Game-Theoretic Analysis of Cross-Ledger Swaps with Packetized Payments," in *Workshop Proceedings of Financial Cryptography and Data Security*, 2021. [Online]. Available: http://arxiv.org/abs/2103.02056

[21] "Btc relay." [Online]. Available: https://github.com/ethereum/btcrelay

[22] A. Back, M. Corallo, L. Dashjr, M. Friedenbach, G. Maxwell, A. Miller, A. Poelstra, J. Timón, and P. Wuille, "Enabling blockchain innovations with pegged sidechains," *URL: http://www.opensciencereview.com/papers/123/enablingblockchain-innovations-with-pegged-sidechains*, 2014.

[23] S. Johnson, P. Robinson, and J. Brainard, "Sidechains and interoperability," *arXiv preprint arXiv:1903.04077*, 2019.

[24] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016.

[25] L. Luu, V. Narayanan, C. Zheng, K. Baweja, S. Gilbert, and P. Saxena, "A secure sharding protocol for open blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 17–30.

[26] A. Miller, I. Bentov, R. Kumaresan, and P. McCorry, "Sprites: Payment channels that go faster than lightning," *arXiv preprint arXiv:1702.05812*, 2017.

[27] A. Zamyatin, D. Harz, J. Lind, P. Panayiotou, A. Gervais, and W. Knottenbelt, "XCLAIM: Trustless, interoperable, cryptocurrency-backed assets," in *IEEE Symposium on Security and Privacy*. IEEE, 5 2019, pp. 193–210. [Online]. Available: https://ieeexplore.ieee.org/document/8835387/

[28] M. Borkowski, D. McDonald, C. Ritzer, and S. Schulte, "Towards Atomic Cross-Chain Token Transfers: State of the Art and Open Questions within TAST," *Distributed Syste ms Group, TU Wien (Technische Universität Wien), Vienna, Austria, Tech. Rep*, 2018.

[29] M. Herlihy, "Atomic cross-chain swaps," in *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*, ACM. Association for Computing Machinery, 7 2018, pp. 245–254.

[30] R. Han, H. Lin, and J. Yu, "On the optionality and fairness of Atomic Swaps," in *AFT 2019 - Proceedings of the 1st ACM Conference on Advances in Financial Technologies*. New York, New York, USA: ACM Press, 2019, pp. 62–75. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3318041.3355460

[31] E. Heilman, S. Lipmann, and S. Goldberg, "The Arwen Trading Protocols (Full Version)." [Online]. Available: https://eprint.iacr.org/2020/024.pdf

[32] V. Zakhary, D. Agrawal, and A. E. Abbadi, "Atomic commitment across blockchains," *arXiv preprint arXiv:1905.02847*, 2019.

[33] M. Herlihy, B. Liskov, and L. Shrira, "Cross-chain Deals and Adversarial Commerce," *Proceedings of the VLDB Endowment*, vol. 13, no. 2, pp. 100–113, 2019. [Online]. Available: http://www.vldb.org/pvldb/vol13/p100-herlihy.pdf

[34] M. Belotti, S. Moretti, M. Potop-Butucaru, and S. Secci, "Game theoretical analysis of Atomic Cross-Chain Swaps," 2020. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02414356

[35] TierNolan, "Atomic swaps using cut and choose," 2016. [Online]. Available: https://bitcointalk.org/index.php?topic=1364951

[36] M. S. Ferdous, M. J. M. Chowdhury, M. A. Hoque, and A. Colman, "Blockchain Consensus Algorithms: A Survey," 1 2020. [Online]. Available: http://arxiv.org/abs/2001.07091

[37] S. Tang and S. S. Chow, "Systematic market control of cryptocurrency inflations: Work-in-progress," in *BCC 2018 - Proceedings of the 2nd ACM Workshop on Blockchains, Cryptocurrencies, and Contracts, Co-located with ASIA CCS 2018*. New York, New York, USA: ACM Press, 2018, pp. 61–63. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3205230.3205240

[38] Digiconomist, "Guide to Cryptocurrency Volatility," 2014. [Online]. Available: https://digiconomist.net/guide_to_cryptocurrency_volatility/

[39] T. Koensa and E. Polla, "Assessing Interoperability Solutions for Distributed Ledgers." [Online]. Available: https://www.ingwb.com/media/2667864/assessing-interoperability-solutions-for-distributed-ledgers.pdf

[40] T. Eizinger, L. Fournier, and P. Hoenisch, "The state of atomic swaps," 2018. [Online]. Available: http://diyhpl.us/wiki/transcripts/scalingbitcoin/tokyo-2018/atomic-swaps/