Skolkovo Institute of Science and Technology

MASTER'S THESIS

Portfolio Sold-Out Problem in Numbers for DeFi Lending Protocols

Master's Educational Program: Data Science

Student: ____________ Waralak Pariwatphan
signature

Research Advisor: ____________ Yury Yanovich
signature
PhD, Senior Research Scientist

Moscow 2023

Skolkovo Institute of Science and Technology

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Задача распродажи портфеля в цифрах для протоколов кредитования DeFi

Магистерская образовательная программа: Науки о данных

Студент: ____________ Вараалак Париватфан
подпись

Научный руководитель: ____________ Юрий Янович
подпись
д.ф.-м.н., старший научный сотрудник

Москва 2023

Contents

Chapter 1
Introduction

1.1 Motivation

Blockchain is a decentralized database with built-in auditability and a tamper-resistant log. It happened to reduce distrust in various domains, such as government, financial institutions, and commercial sections. On blockchain, no data or activities can be manipulated or erased. The data in blockchain is kept as transactions, which are then grouped into blocks and linked together using a chain. A smart contract is a computer technology designed to digitally facilitate, verify, or enforce the negotiation or performance of a contract on blockchain.

Decentralized Finance (DeFi) is a financial technology based on blockchain that eliminates the need for intermediaries and centralized financial institutions in financial transactions. MakerDao is the world's leading DeFi lending platform, allowing users to lock their reliable cryptocurrencies as collateral and generate a stable coin known as Dai. Borrowers must maintain their collateral ratio in order to avoid liquidation. If their loans are liquidated, the lending platform will auction their assets as collateral to recoup the loan and costs, including the system's stability fee and the penalty fee. All DeFi lending protocol transactions were saved in Ethereum data.

Portfolio optimization based on Markowitz's methodology is widely known in quantitative finance[9]. The fundamental concept is to diversify financial assets in a portfolio in order to reduce portfolio risk and select the most efficient portfolio. In the traditional investments, splitting financial assets is prohibited by the regulation due to a lack of audibility. Due to the vast size of invested assets, small private investors cannot participate in investment.

According to its main properties of decentralization, scalability, and security, blockchain can solve this issue. Blockchain technology is remarkable for permitting the spread of digital ledgers without the need for a central entity to keep its state safe. The main feature is tokenization, which can fractionize assets into tokens and also transform ownership asset rights into that digital token. Nevertheless, data in DeFi lending protocol transactions varies from traditional finance, and there are a few study topics regarding the portfolio sold-out problem in DeFi. This research will be beneficial for portfolio for digital asset in order to adapt blockchain to traditional finance.

1.2 Research Problem

Due to a lack of audibility in traditional finance, the regulation prohibits splitting financial assets but blockchain can solve that problem. Furthermore, there are a few studies on the portfolio sold-out problem in DeFi to serve as the foundation for developing DeFi lending protocols.

1.3 Goals and Objectives

This research has two primary objectives. First, this research will analyze the MakerDao dataset, including economic statistics such as Loss Given Default (LGD), Probability of Default (PD), and interest rate, as well as investigate the optimal portfolio sold-out by implementing tokenization algorithms and evaluating the tokenized asset fraction in the portfolio package for actual loan blockchain-based dataset.

Chapter 2
Literature Review

2.1 Blockchain and Smart Contracts

A blockchain is a decentralized database that keeps track of all transactions that occur on the blockchain. All transactions cannot be deleted or manipulated. The transactions are grouped into blocks, and each block contains the data from the preceding block. Each block is linked by a chain. The benefit of a blockchain is that it enables non-trusting people to securely communicate and trade assets without the need for a trusted third party. As a consequence, both the integrity and double-spending issues are alleviated [2].

Since the 1990s, smart contracts have been proposed as a digital transaction mechanism based on blockchain technology. Smart contracts are basically code containers that contain and copy real-world contract terms in the digital platform. Contracts are legally binding agreements between two or more parties that agree to fulfill their duties. Nevertheless, smart contracts can take the position of trusted third parties or intermediaries between contractual parties. They do it with the use of code execution, which is automatically distributed and validated by nodes in the network in a blockchain network [15].

2.2 DeFi Lending Protocols

Decentralized finance (DeFi) is a financial infrastructure that is developed on blockchain technology and smart contracts. DeFi has the advantages of being permissionless, trust - free, transparent, and networked. DeFi can provide a variety of financial services, such as a lending protocol, a Decentralized Exchange (DEX), and a payment platform. DeFi lending protocols let users to deposit tokens as collateral, lend another token, and give liquidity to lending pools. Liquidation, on the other hand, assists procedures in reducing debt exposure when collateral prices decline [14].

MakerDao is one of the peer-to-peer lending protocols that distributes Dai. Dai is a stable coin whose value is pegged to the US dollar. As the demand for Dai changes, the demand curve shifts owing to market conditions, Dai holders' confidence, or other causes, the supply curve is adjusted via a permissionless credit factory on Ethereum. As a result, the algorithm can keep the Dai price as near to 1 US dollar as possible.

Users can borrow Dai by securing their cryptocurrency assets as collateral. To reclaim the collateral, they must return Dai plus a charge. As long as their collateralized vault value exceeds their collateral ratio, their loan status will be secure. If the collateral value (in US dollars) goes too low, the protocol auctions off portion of the collateral to cover the outstanding debt and penalty cost. The protocol will burn Dai to reduce supply, and the vault owner will get any leftover collateral.

The system has numerous modules, such as the Dai Module, which contains the Dai Token Contract, the Vault Core Module, which contains the Vat and Spot contracts, and the Collateral Module, which contains the Join and Clip contracts [10].

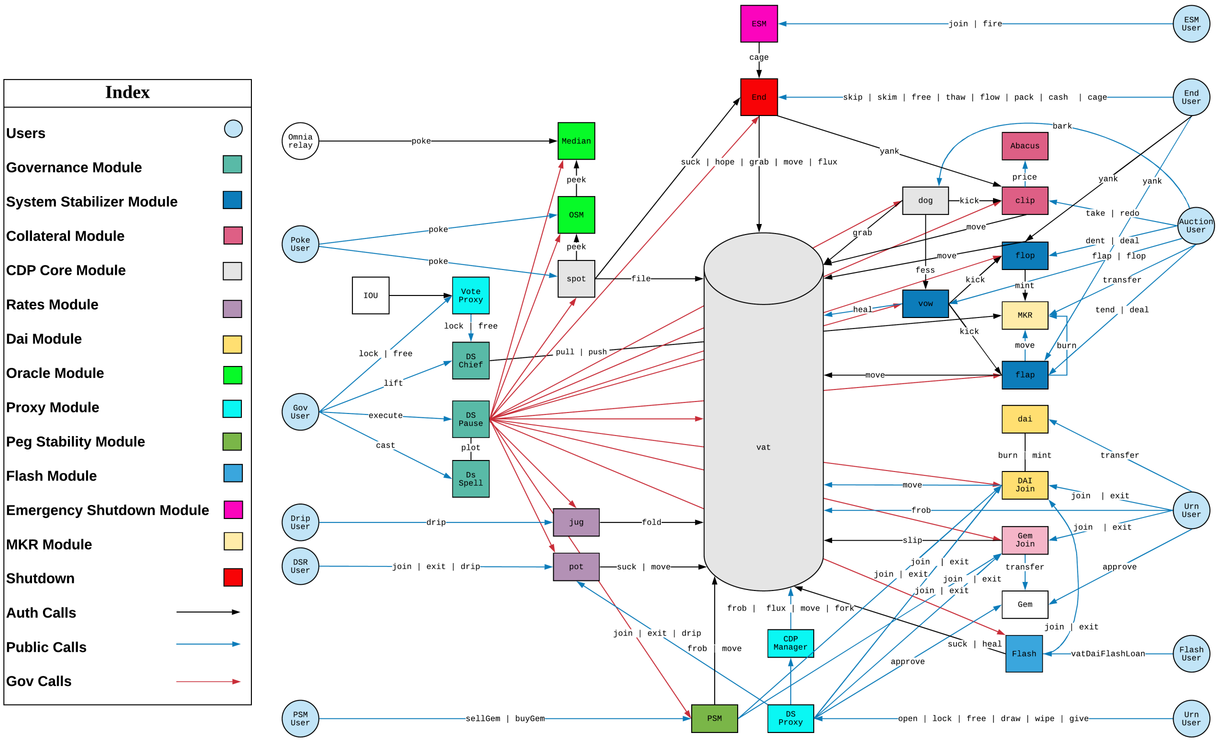

Figure 2.1: MakerDao module entities [4]

2.3 Geometric Brownian motion

Geometric Brownian motion is a stochastic process used to represent random motions. It is similar to a random walk. It is commonly applied in the study of financial markets, such as stock prices, foreign exchange, financial assets, and so on. A Wiener stochastic process is the standard Brownian motion [13]. Let $W(t)$ is the value in $\mathbf{R}$ when $t \in [0, \infty)$, the properties of Wiener process are

1. $W(0) = 0$.
2. for all $0 \leq t_1 < t_2, W(t_2) - W(t_1) \sim N(0, t_2 - t_1)$.
3. $W(t)$ has independent increments. For all $0 \leq t_1 < t_2 < \cdots < t_n$, the random variables $W(t_2) - W(t_1), W(t_3) - W(t_2), \cdots, W(t_n) - W(t_{n-1})$ are independent.
4. $W(t)$ paths are continuous.

First passage times of levels [1] are applied in a various applications. The time when the process hits an arbitrary threshold a for the first time is a random variable, according to Brownian Motion theory. It is possible to assume that the process begins at a certain position $x_0 > 0$ and is bounded to the positive semi-axis by an absorbing barrier x = 0. Assume that $a > 0$ and let $T_a = \inf\{t \geq 0 : B_t = a\}$. Then we will have

$$\mathrm{P}(T_a < t) = \mathrm{P}\left(\sup_{s \leq t} B_s > a\right). \tag{2.1}$$

Applying the reflection principle of D. Andre, it follows that

$$\mathrm{P}(T_a < t) = 2\mathrm{P}(B_t \geq a). \tag{2.2}$$

Since

$$\mathrm{P}\left(B_t \geq a\right) = \frac{1}{\sqrt{2\pi t}} \int_a^{\infty} e^{-\frac{x^2}{2t}} dx, \tag{2.3}$$

So,

$$P\left(T_a < t\right) = \int_0^t \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{a^2}{2s}} ds \tag{2.4}$$

2.4 Probability of Default (PD)

Probability of default (PD) for each single debt over period time T can be formulated as:

$$\begin{aligned} \psi\left(x_{min}\right) &= P\left(T_{x_{min},f} < T\right) \\ &= \int_0^T \frac{-x_{min} - ft}{\sqrt{2\pi s^3}} e^{-\frac{(-x_{min}+ft)^2}{2s}} ds \end{aligned} \tag{2.5}$$

$$x_{min}(t) = \frac{1}{\sigma} ln\left(\frac{d_0 \cdot r_{min}}{a_0 \cdot e_0}\right) + ft \tag{2.6}$$

where f is stability fee, x is level in Brownian motion, x_{min} is level of default, r_{min} is liquidation ratio, a is collateral assets, and e is exchange rate [7].

2.5 Tokenization

The OECD has proposed two methods of asset tokenization as following [12]:

Tokenization of real assets that exist off-the-chain is the process of representing an existing asset on a distributed ledger in digital platform [**?**]. Real assets can be tokenized on blockchain using smart contracts and used separately as tokens and vaults. Tokens on the blockchain can represent economic value and asset rights.

Tokenization of real assets that exist off-the-chain is the method of representing native tokens that are generated directly on-chain and distributed ledger, such as tokens issued in initial coin offerings (ICOs) that are created on the blockchain. ICOs are comprised of start-up enterprises creating digital tokens and distributing them to investors in return for funding and fundraising.

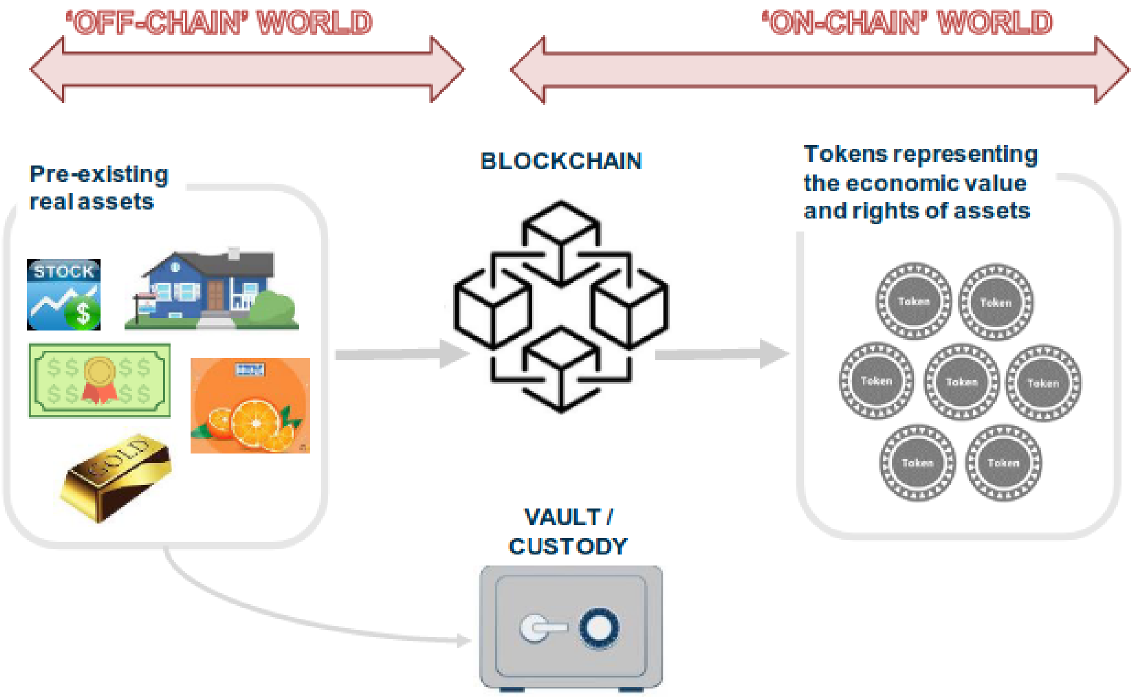

Figure 2.2: Tokenization of real assets that exist off-the-chain [12]

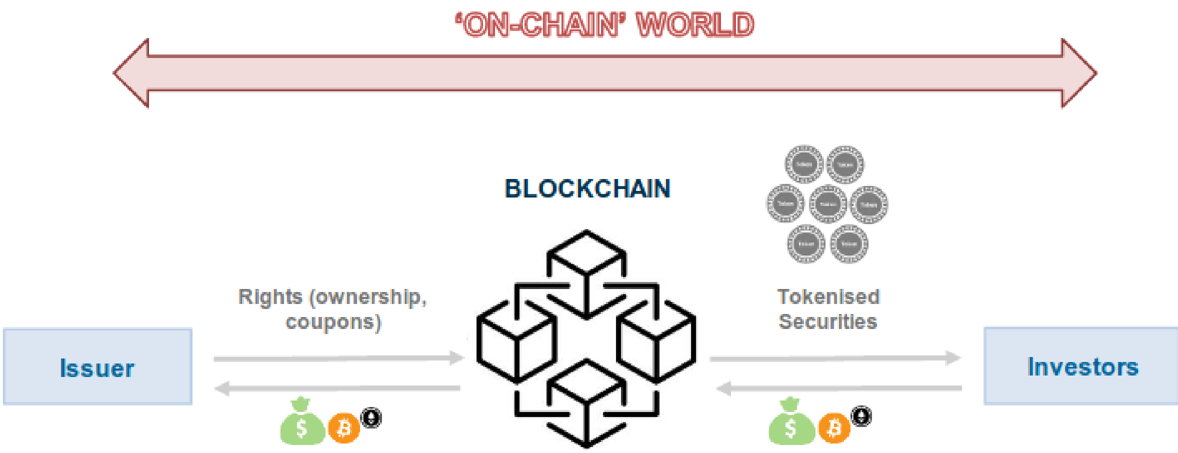

Figure 2.3: Tokenization of assets "native" to the blockchain [12]

Chapter 3
Methods

3.1 Portfolio sold-out problem

According to the portfolio sold-out problem [6], the variance of the package $\vec{c}$ equals

$$V(\vec{c}) = \mathrm{Var}\, \vec{c}^{\,T}\vec{\xi} = \vec{c}^{\,T}\mathbf{K}\vec{c}. \tag{3.1}$$

A set of M packages $\mathbf{C}_M = (\vec{c}_1|\dots|\vec{c}_M) \in R^{N\times M}$ is the tokenization of the portfolio $(\vec{A}, \vec{\xi})$ if $\sum_{m=1}^{M} \vec{c}_m \leq \vec{A}$.

The variance V of tokenization $\mathbf{C}_\mathrm{M}$ is the maximum variance of its packages:

$$V(\mathbf{C}_\mathrm{M}) = \max_{\mathrm{m}\in\overline{\mathrm{M}}} V\left(\vec{c}_\mathrm{m}\right). \tag{3.2}$$

For a given portfolio $(\vec{\mathrm{A}}, \vec{\xi})$ and a variance threshold $\sigma^2 > 0$, the portfolio sold-out problem is

$$\mathrm{M} \to \max_{\mathrm{M}, \mathbf{C}_\mathrm{M}: V(\mathbf{C}_\mathrm{M}) \leq \sigma^2}. \tag{3.3}$$

Portfolio Sold-Out Problem: Special Cases

Consider by assets and packages		
Cases	Discrete	Continuous
Homogeneous	$\mathbf{K} = \sigma_0^2\mathbf{I}_N$ $\mathbf{C}_M$ is Boolean matrix.	$\mathbf{K} = \sigma_0^2\mathbf{I}_N$ $\mathbf{C}_M$ is real matrix.
Independent	$K_{ij} = 0$ for $i \neq j$ $\mathbf{C}_M$ is Boolean matrix.	$K_{ij} = 0$ for $i \neq j$ $\mathbf{C}_M$ is real matrix.
General	any K is allowed $\mathbf{C}_M$ is Boolean matrix.	any K is allowed $\mathbf{C}_M$ is real matrix.

3.2 Tokenization algorithms

Homogeneous systems have useful properties for analysis and design, such as stabilization, convergence rates, and trajectory scalability [5]. The algorithms' outputs are the number of created tokens M and the composition matrix $C_M = (\vec{c}_1|\dots|\vec{c}_M)$ with M columns and N rows. The amount of assets that make up into the packages $\bar{a} = \sum_{m=1}^{M} \vec{c}_m$ [6].

3.2.1 Discrete algorithm

For discrete homogeneous tokenization, there are the set of M packages:

$$C_M = (\vec{c}_1|\dots|\vec{c}_M) \in \left\{0, \frac{1}{k}\right\}^{N\times M} \tag{3.4}$$

with the variance reduction parameter k of the portfolio $(\vec{A}, K)$ if:

$$\sum_{m=1}^{M} \vec{c}_m \leq \vec{A} \tag{3.5}$$

and

$$\forall m \in \bar{M} : \|\vec{c}_m\|_1 = 1. \tag{3.6}$$

3.2.2 Continuous algorithm

For continuous homogeneous tokenization, there are the set of M packages:

$$C_M = (\vec{c}_1 | \ldots | \vec{c}_M) \in \mathbf{R}^{N \times M} \tag{3.7}$$

with the variance reduction parameter k of the portfolio $(\vec{A},$ K) if :

$$\forall m \in \bar{M} : \|\vec{c}_m\|_2 \leq \frac{1}{k} \wedge \|\vec{c}_m\|_1 = 1 \tag{3.8}$$

3.2.3 Compact continuous algorithm

The portfolio sold-out problem formulation is

$$M \rightarrow \max_{M, \mathrm{C}_M} . \tag{3.9}$$

The constraints are:

$$\begin{gathered} \vec{c}^T \,\mathrm{K} \vec{c}_m \leq \sigma^2, \\ \sum_{m=1}^{M} \vec{c}_m \leq \vec{A}, \\ \|\vec{c}\|_1 = 1. \end{gathered} \tag{3.10}$$

There is a proof that an optimal matrix $\bar{C}_M$ with $\vec{c_1} = \vec{c_2} = ... = \vec{c_M} = constant$, where $\vec{c} = \frac{1}{M} \sum_{m=1}^{M} \vec{c}_m$. Thus, original portfolio sold-out problem can be reformulated as

$$M \rightarrow \max_{\mathrm{C}_M} \tag{3.11}$$

with the constraints

$$\begin{gathered} \vec{a}^T K \vec{a} \leq \sigma^2 \cdot \|\vec{a}\|_1^2, \\ \vec{0} \leq \vec{a} \leq \vec{A}. \end{gathered} \tag{3.12}$$

MEALPY [16] and SOCP are two examples for optimization methods that may be used to solve compact continuous algorithms.

3.3 Optimization methods

3.3.1 Second-order cone programming

The second-order cone programming (SOCP) problem is a convex optimization problem with the objective of minimizing the linear function. The function's constraint is defined as the intersection of an affine linear manifold with the Cartesian product of second-order (Lorentz) cones. The duality of this problem can be also regarded as a particular case of general duality theory for problems with non-negative constraints[11]. The problem of convex quadratic programming can be expressed as a SOCP problem for this research subject.

SOCP's standard form and linear programs are expressly comparable:

$$\begin{aligned} &\min \sum_{i=1}^{k} c_i x_i \\ &\text{s.t. } \sum_{i=1}^{k} x_i \mathbf{a_i} = \mathbf{b} \\ &\quad x_i \geq 0, \quad \text{for } i = 1, \cdots, k \end{aligned} \tag{3.13}$$

where $x_i \in \mathbf{R}, i = 1, ..., n$ and $c_i \in \mathbf{R}, i = 1, \cdots, n$ are scalars represented the objective function coefficients. The constraint $\mathbf{a_i} \in \mathbf{R}^m, i = 1, \cdots, n$ and $\mathbf{b} \in \mathbf{R}^m$ are vectors and $x_i \in \mathbf{R}, i = 1, ..., k$, are in spaces of dimension one.

Furthermore, linear programming is a subset of SOCP, and $\mathbf{R}^2$, $K = \{(x_0; x_1) \in \mathbf{R}^2 \mid x_0 \geq |x_1|\}$ is a rotation of the non-negative quadrant. As a consequence, SOCP can be transformed into a linear program formula if all second order cones are one or two dimensional.

In summary, SOCP is a convex programming problem because second-order cones are convex sets. Also, the dimension exceeds two and is not polyhedral. As a result of this characteristic, the feasible region is not polyhedral [3].

The package CVXPY can solve the SOCP problem. For instance, the issue may be addressed as follows:

$$\begin{aligned} &\min f^T x \\ &\text{s.t. } \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \cdots, m \\ &\quad Fx = g \end{aligned} \tag{3.14}$$

where $x \in \mathbf{R}^n$ is variables. The problem data are $f \in \mathbf{R}^n, A_i \in \mathbf{R}^{n_i \times n}, b_i \in \mathbf{R}_i^n, c_i, d_i \in \mathbf{R}, F \in \mathbf{R}^{p \times n}$, and $g \in \mathbf{R}^p$ [8].

Bibliography

[1] a. N. Shiryaev. *Essentials of stochastic finance*. 1999.

[2] Alharby, M., Aldweesh, A., and Moorsel, A. V. Blockchain-based smart contracts: A systematic mapping study of academic research (2018).

[3] Alizadeh, F., and Goldfarb, D. Second-order cone programming. *Mathematical Programming, Series B 95* (2003).

[4] Author, N. N. Vat - detailed documentation, the maker protocol's core accounting system, 2022.

[5] Chaleenutthawut, Y., Davydov, V., Kuzmin, A., and Yanovich, Y. Practical blockchain-based financial assets tokenization.

[6] Davydov, V., and Yanovich, Y. Optimal portfolio sold-out via blockchain tokenization.

[7] Davydov, V. A., Kruglik, S. A., and Yanovich, Y. A. Probability of the default-free state for tocken package from independent loans. *Journal of Communications Technology and Electronics 67* (6 2022), 778–786.

[8] Diamond, S., and Boyd, S. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research 17*, 83 (2016), 1–5.

[9] Herve M. Tenkam, J. C. M., and Mwambi, S. M. Optimization and diversification of cryptocurrency portfolios: A composite copula-based approach. *Appl. Sci. 2022, 12, 6408* (2022).

[10] Kenton, Wouter, Soren, Tom, and B, C. Maker protocol 101, 12 2020.

[11] Nesterov, Y., and Nemirovskii, A. *Interior-Point Polynomial Algorithms in Convex Programming*. 1994.

[12] OECD(2020). The tokenisation of assets and potential implications for financial markets, oecd blockchain policy series. Tech. rep., OECD, 2020.

[13] Pishro-Nik, H. Introduction to probability, statistics, and random processes. *Kappa Research LLC* (2014).

[14] Shuai, Y., and Wei, C. An evaluation system for defi lending protocols.

[15] Taherdoost, H. Smart contracts in blockchain technology: A critical review. *Information 14* (2 2023), 117.

[16] Thieu, N. V., and Mirjalili, S. MEALPY: a Framework of The State-of-The-Art Meta-Heuristic Algorithms in Python, June 2022.