MSc Program

Portfolio Sold-Out Problem in Numbers for DeFi Lending Protocols

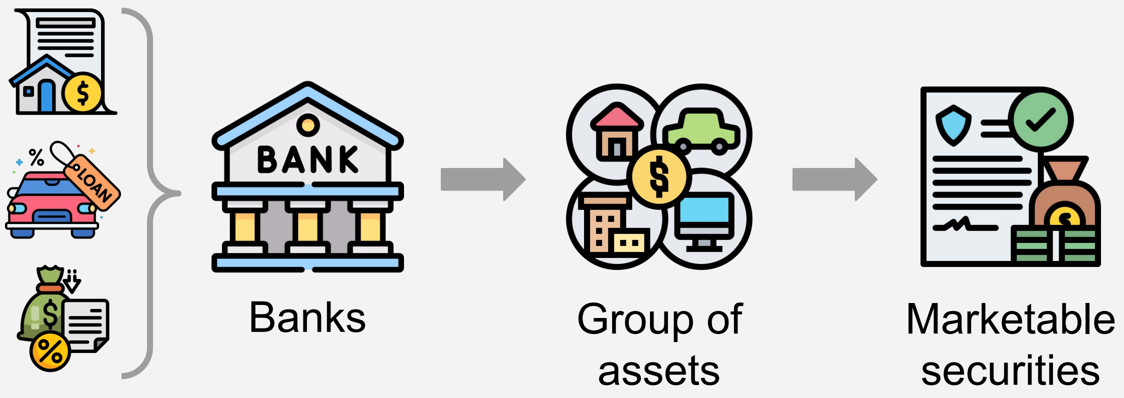
Student: *Waralak Pariwatphan*
Research Advisor: *Yury Yanovich*

Skoltech

October, 2023

Background : DeFi Meets Classic Finance

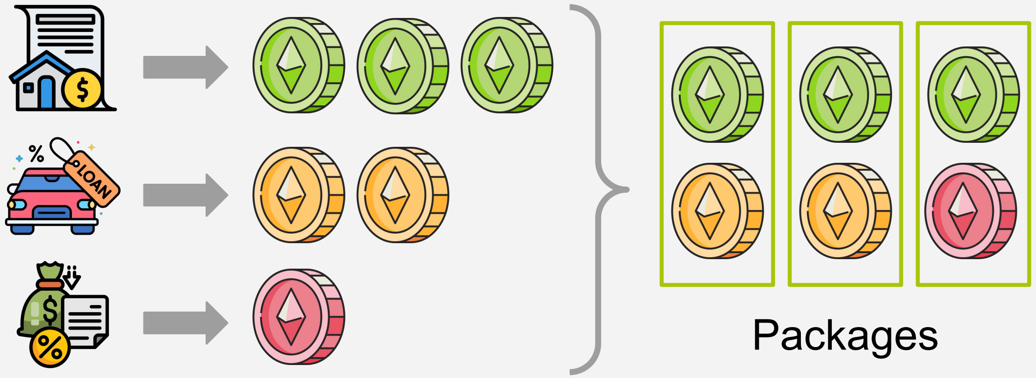
Securitization



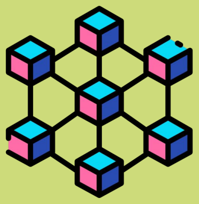
- Plenty of intermediaries between bank and investors
- Lack of transparency
- Regulatory and legal risk

Classic Finance

Tokenization

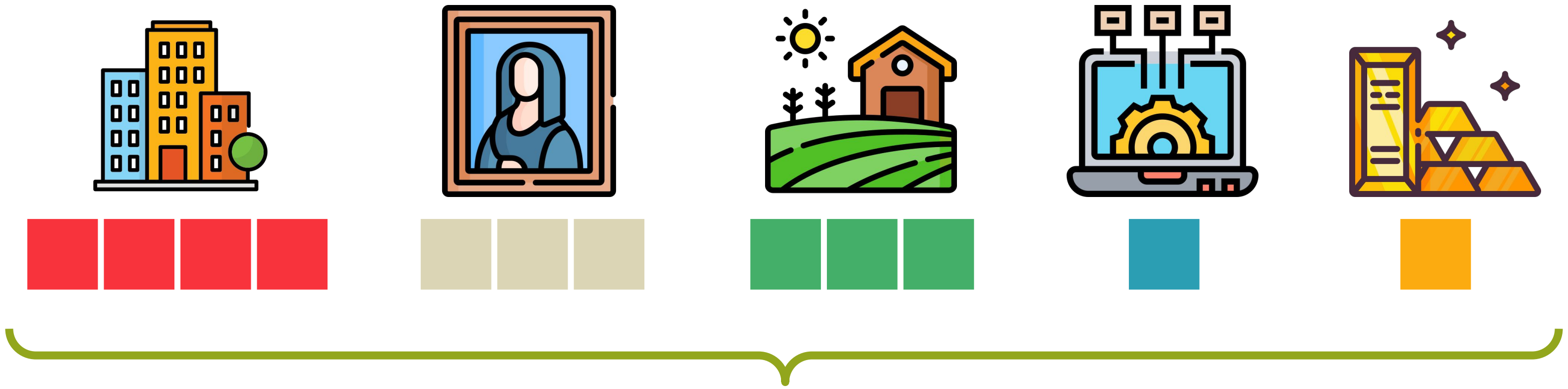


Solve



DeFi

Problem



How to construct as many packages as possible for a given token set under the specific risk?

Portfolio sold-out problem

Objectives

01

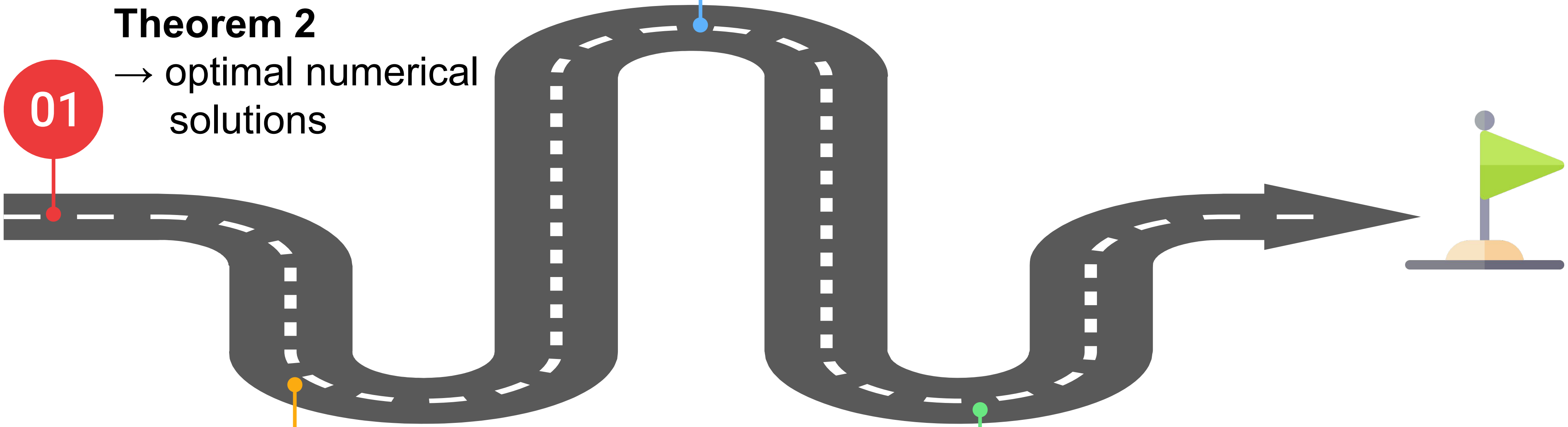
To design and develop optimization methods for obtaining the optimal portfolio sold-out in the following cases:

- continuous independent.
- continuous general.
- discrete independent.
- discrete general.

02

To apply optimization methods to the MakerDAO dataset

Main Contributions



01 Formulated and proved **Theorem 2** → optimal numerical solutions

02 Formulated and proved **Theorem 3** → NP-hard problem

03 Implemented tokenization on dataset

04 Proved market concentration as a reliable metric

Skoltech

Portfolio sold-out problem: An overview

Problem $\xrightarrow{\text{similar to}}$ Markowitz model

Different

small investors

Instead of seeking a single best portfolio, construct as many reasonably good portfolios as possible.

- One of subproblem is related to SOCP → optimal solution
- One of subproblem is related to Subset Sum Problem → NP-hard

Methodology

The portfolio is characterized by the number of assets $N \geq 1$ and a set of random variables $A_1 \cdot \xi_1, \ldots, A_N \cdot \xi_N$, where

- $A_1 \leq \cdots \leq A_N$ are deterministic positive numbers equal to the expected returns of each asset
- random variables $\xi_1, \ldots, \xi_\mathrm{N}$, describing the uncertainty per unit of return
  - $\mathbf{E}\xi_n = 1, n \in \overline{N}$
  - covariance $\operatorname{cov}(\xi_i, \xi_j) = K_{ij}, \quad i, j \in \overline{N}$
  - covariance matrix $\mathbf{K} = (K_{ij})_{i,j \in \overline{N}}$

A $\mathbf{package}$ composed of the portfolio $(\vec{A}, \vec{\xi})$ is a vector $\vec{c} \in \mathbb{R}^N$ such that

- $0 \leq \vec{c} \leq \vec{A}$ and
- $\mathbf{E}\vec{c}^{\,T}\vec{\xi} = 1.$

Skoltech

Methodology

The variance of the package $\vec{c}$ equals to $\mathrm{V}(\vec{c}) = \mathrm{Var}\vec{c}^T\vec{\xi} = \vec{c}^T\mathbf{K}\vec{c}$.

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

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

$$\mathrm{V}(\mathbf{C}_M) = \max_{m\in\overline{M}} \mathrm{V}(\vec{c}_m).$$

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

$$M \to \max_{M, \mathbf{C}_M : \mathrm{V}(\mathbf{C}_M) \leq \sigma^2}$$

Skoltech

Methodology : Portfolio sold-out special cases

By assets:

Categories	Covariance matrix types
Homogeneous	$\mathbf{K} = \sigma_0^2 \mathbb{I}^N$
Independent	$\mathrm{K}_{ij} = 0$ for $i \neq j$
General	any $\mathbf{K}$ is allowed

	Discrete	Continuous
Homogeneous	optimal solution	optimal solution
Independent	to be solved	to be solved
General	to be solved	numerical solution

By packages:

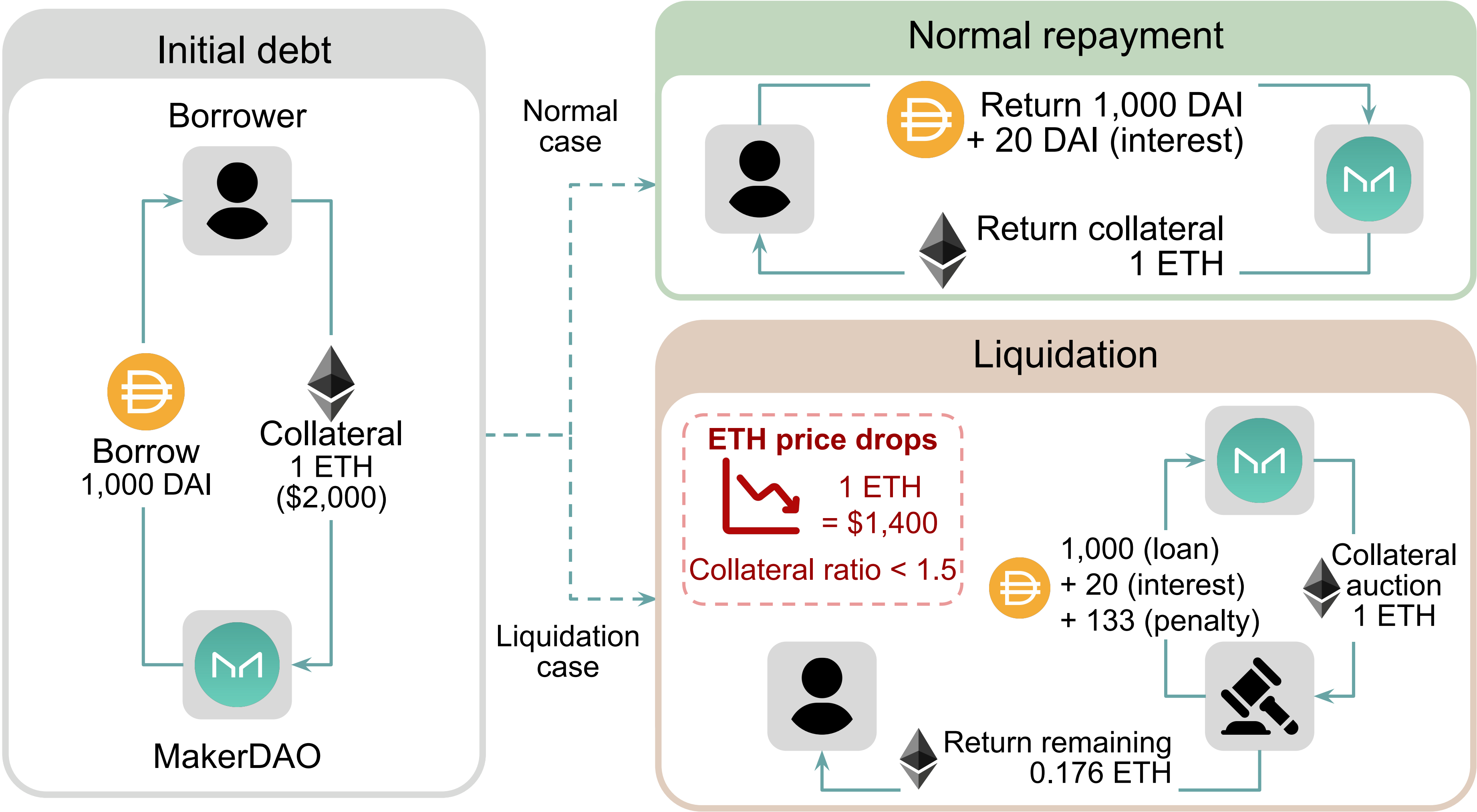
Categories	Package types
Discrete	$\mathbf{C}_M$ is boolean matrix
Continuous	$\mathbf{C}_M$ is real matrix

The proportion of tokenized asset into the packages defines as

$$\text{Tokenized fraction} = \frac{\text{the total amount of tokenized asset}}{\text{the total amount of initial asset}}$$

Skoltech

Methodology : Dataset



MakerDAO loan system

Daily data

$\vec{A}$ is borrowers with debt amounts.

$\mathbf{K}$ is a covariance for their defaults.

Skoltech

Results : Portfolio sold-out problem

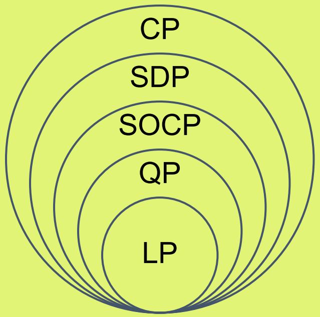
	Discrete	Continuous
Homogeneous	optimal explicit solution	optimal explicit solution
Independent	NPH	optimal numerical solution
General	NPH	optimal numerical solution

Solutions for portfolio sold-out special cases

Result : Continuous General Case

Theorem 2. Portfolio sold-out problem is polynomial reducible to second order cone programming.

→ optimal numerical solution.



LP: linear program,
QP: quadratic program,
SOCP second-order cone program,
SDP: semidefinite program,
CP: cone program.

Proof.

Denote $\overline{M} = \{1, \ldots, M\}$ and $M \in \mathbb{I}^+$.

If the matrix $\mathbf{C}_M \in \mathbb{R}^{N \times M}$ is the optimal solution to the continuous general problem,

then the matrix $\overline{\mathbf{C}}_M = \left(\bar{\vec{c}} | \ldots | \bar{\vec{c}}\right)$, where $\bar{\vec{c}} = \frac{1}{M} \sum_{m \in \overline{M}} \vec{c}_m$.

Consider $\vec{a} = M \cdot \bar{\vec{c}}$. The continuous general problem is equivalent to

$$\|\vec{a}\|_1 \to \max_{\vec{a}:} \begin{cases} \vec{a}^{\mathrm{T}} \mathbf{K} \vec{a} \leq \sigma^2 \|\vec{a}\|_1^2 \\ \vec{0} \leq \vec{a} \leq \vec{A} \end{cases}$$

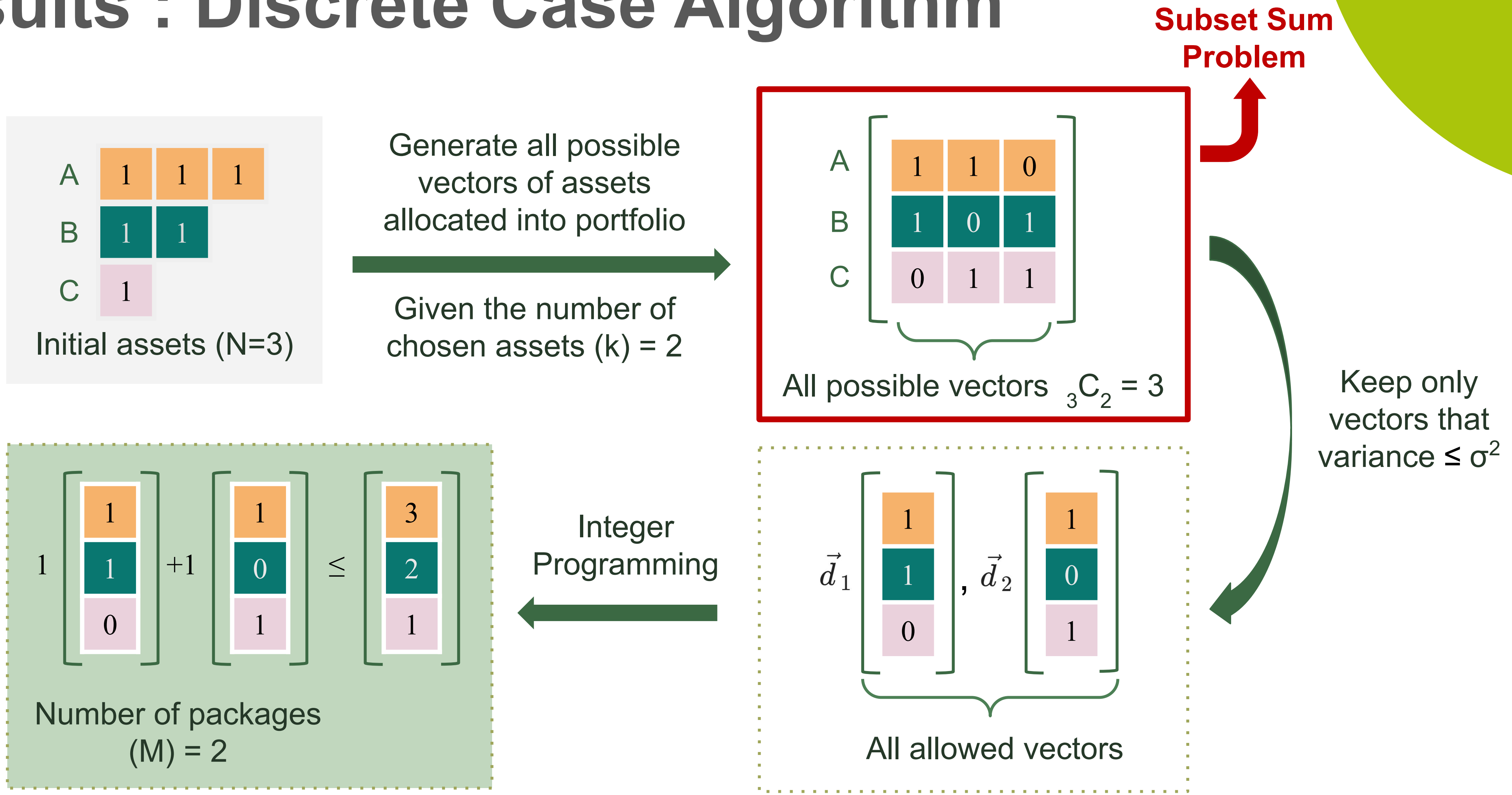
Result : Continuous Independent and General Cases

Independent	$\mathrm{K}_{ij} = 0$ for $i \neq j$	less complex
General	any $\mathbf{K}$ is allowed	

	Discrete	Continuous
Homogeneous	optimal explicit solution	optimal explicit solution
Independent	to be solved	optimal numerical solution
General	to be solved	optimal numerical solution

Skoltech

Results : Discrete Case Algorithm



Results : Discrete Independent Case

Theorem 3. Discrete independent optimal portfolio sold-out problem is **NP-Hard**.

Proof.

The partition problem is NP-complete.

Given a set of positive integers $a_1, \ldots, a_N$, find out whether there is a subset of indexes $I \subset \{1, \ldots, N\}$ such that

$$\sum_{n \in I} a_n = \frac{S}{2}; \quad S \equiv \sum_{n=1}^{N} a_n.$$

We construct a tuple $(\vec{A}, k, K, \sigma^2)$ as the input:

- $\vec{A} = (1, \cdots, 1)^T \quad \text{where } \dim\left(\vec{A}\right) = 2N$
- $k = N$
- K is diagonal matrix with $(a_1, \ldots, a_N, 0, \ldots, 0)$ on the diagonal
- $\sigma^2 = \frac{S}{2}.$

The original problem can be reduces to

$$(M, D_M) = \arg\max_{M, D_M} M,$$

Results : Discrete Independent Case

$\mathrm{D}_M = \left(\vec{d}_1 | \ldots | \vec{d}_M\right) \in \{0,1\}^{2N \times M}$ and the number of packages $M \in \{0,1,2\}$,

The constraint is $\sum_{m=1}^{M} \vec{d}_m \leq \vec{A}$:

$$\sum_{i=1}^{N} d_{m,i}^2 \cdot a_i = \sum_{i=1}^{N} d_{m,i} \cdot a_i \leq \frac{S}{2}, \qquad [1]$$

If $M = 2$, then $\vec{A}$ is divided into 2 groups, $\vec{d}_1$ and $\vec{d}_2$:

- $\vec{d}_1$ and $\vec{d}_2$ meet the constraint [1]
- $\vec{A} = \vec{d}_1 + \vec{d}_2$
- $\vec{d}_1, \vec{d}_2 \in \{0,1\}^{2N}$.

$$\sum_{i=1}^{N} d_{1,i} \cdot a_i + \sum_{i=1}^{N} d_{2,i} \cdot a_i = \sum_{i=1}^{N} a_i = S.$$

Results : Discrete Independent Case

For example, given a solution subset of indexes $I \subset \{1, \dots, N\}$, we can construct

$$\vec{d}_1 = ([1 \in I], \dots, [N \in I], \underbrace{0, \cdots, 0}_{|I|}, \underbrace{1, \cdots, 1}_{N-|I|})^T$$

$$\vec{d}_2 = ([1 \notin I], \dots, [N \notin I], \underbrace{1, \cdots, 1}_{|I|}, \underbrace{0, \cdots, 0}_{N-|I|})^T,$$

For solution $M = 2$,

- as the solution:

$$\begin{cases} \sum_{i=1}^{N} d_{1,i} \cdot a_i = S/2 \\ \sum_{i=1}^{N} d_{2,i} \cdot a_i = S/2. \end{cases}$$

- $\vec{d}_1 + \vec{d}_2 = (1, \cdots, 1)^T = \vec{A}$
- $\vec{d}_1, \vec{d}_2 \in \{0,1\}^{2N}$.

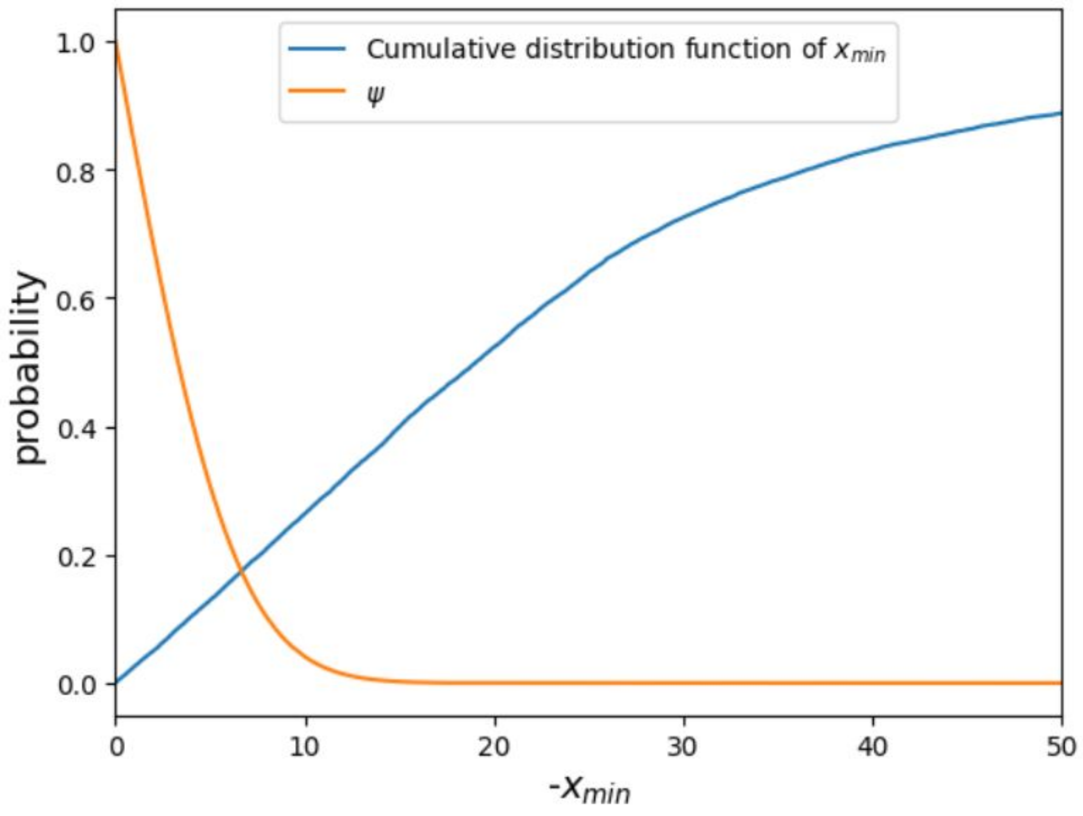
Skoltech

Result : Discrete Independent and General Cases

Independent	$\mathrm{K}_{ij} = 0$ for $i \neq j$	more complex
General	any $\mathbf{K}$ is allowed	

	Discrete	Continuous
Homogeneous	optimal explicit solution	optimal explicit solution
Independent	NPH	optimal numerical solution
General	NPH	optimal numerical solution

Results : MakerDAO dataset



Probability and level of default

Daily data

$\vec{A}$ is borrowers with debt amounts.

$\mathbf{K}$ is a covariance for their defaults.

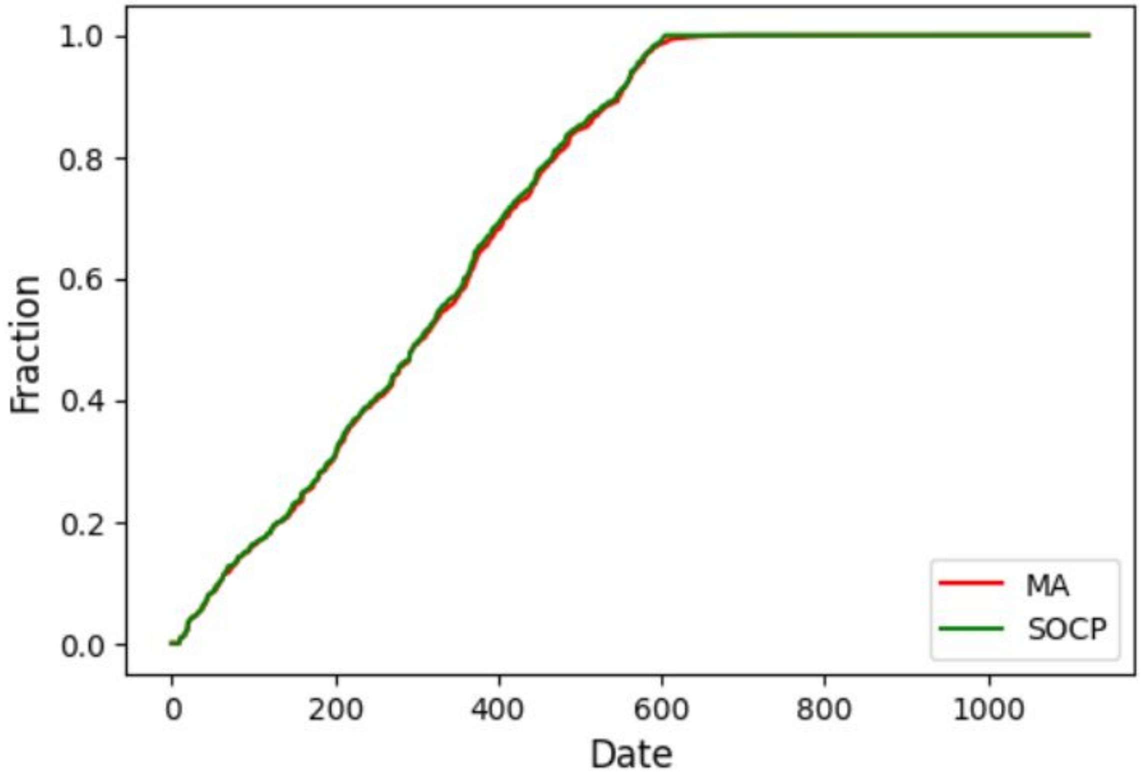
General case: Any $\mathbf{K}$ is allowed

Probability of default

$$\psi\left(x_{min}\right) = \int_0^T \frac{|x_{min} + fs|}{\sqrt{2\pi s^3}} e^{-\frac{(x_{min}+fs)^2}{2s}} ds$$

Skoltech

Result : Continuous General Case



Sorted tokenized fractions for continuous general case

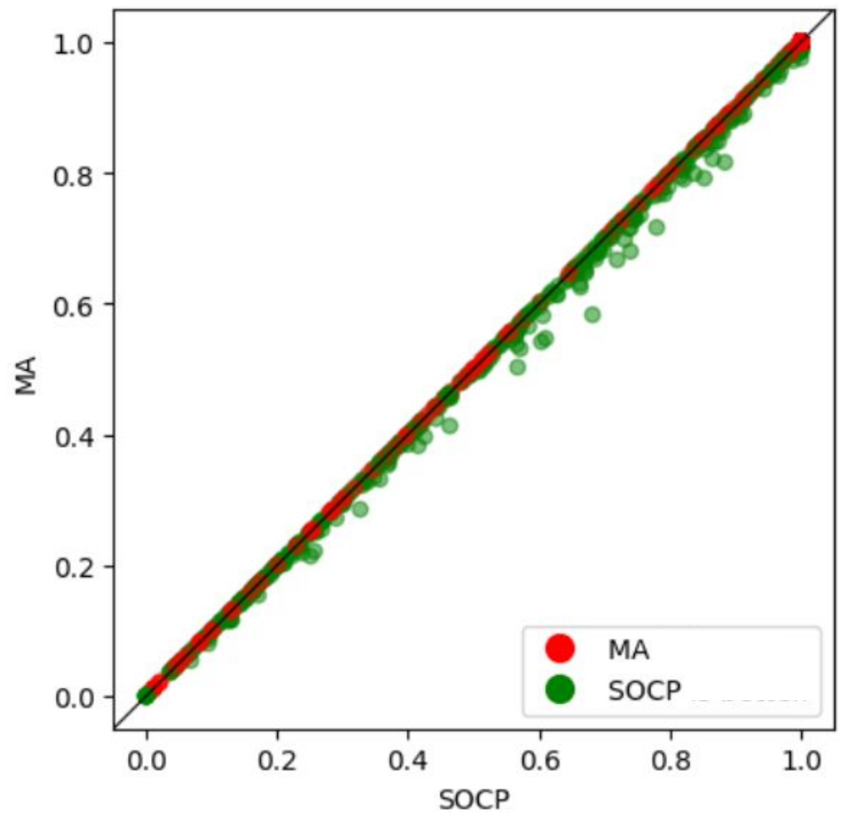
$$\text{Tokenized fraction} = \frac{\text{the total amount of tokenized asset}}{\text{the total amount of initial asset}}$$

Method	Tokenized Fraction
Baseline : Metaheuristic Algorithm (MEALPY)	65.20 %
Second Order Cone Program (CVXPY)	**65.69 %**

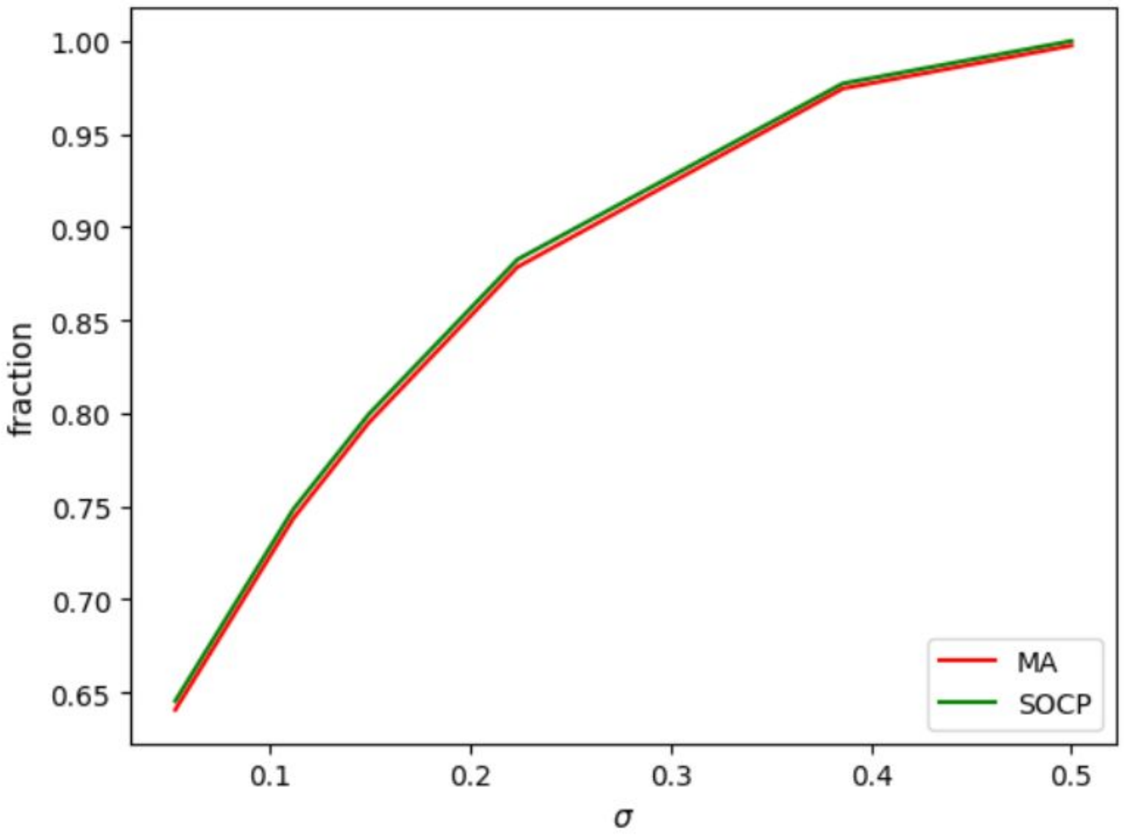
Tokenized fractions of the continuous general case with different optimization methods

Skoltech

Result : Continuous General Case

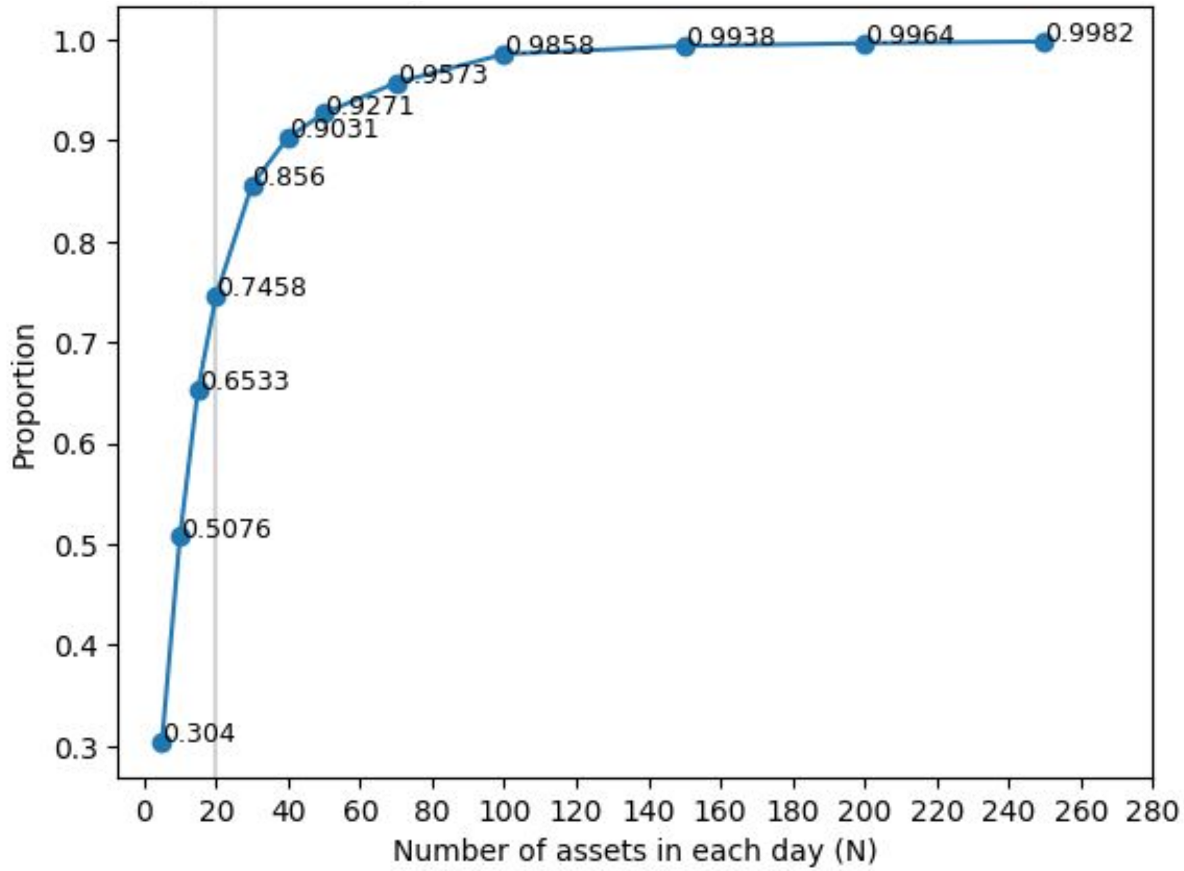


Comparing tokenized fractions with different optimization methods

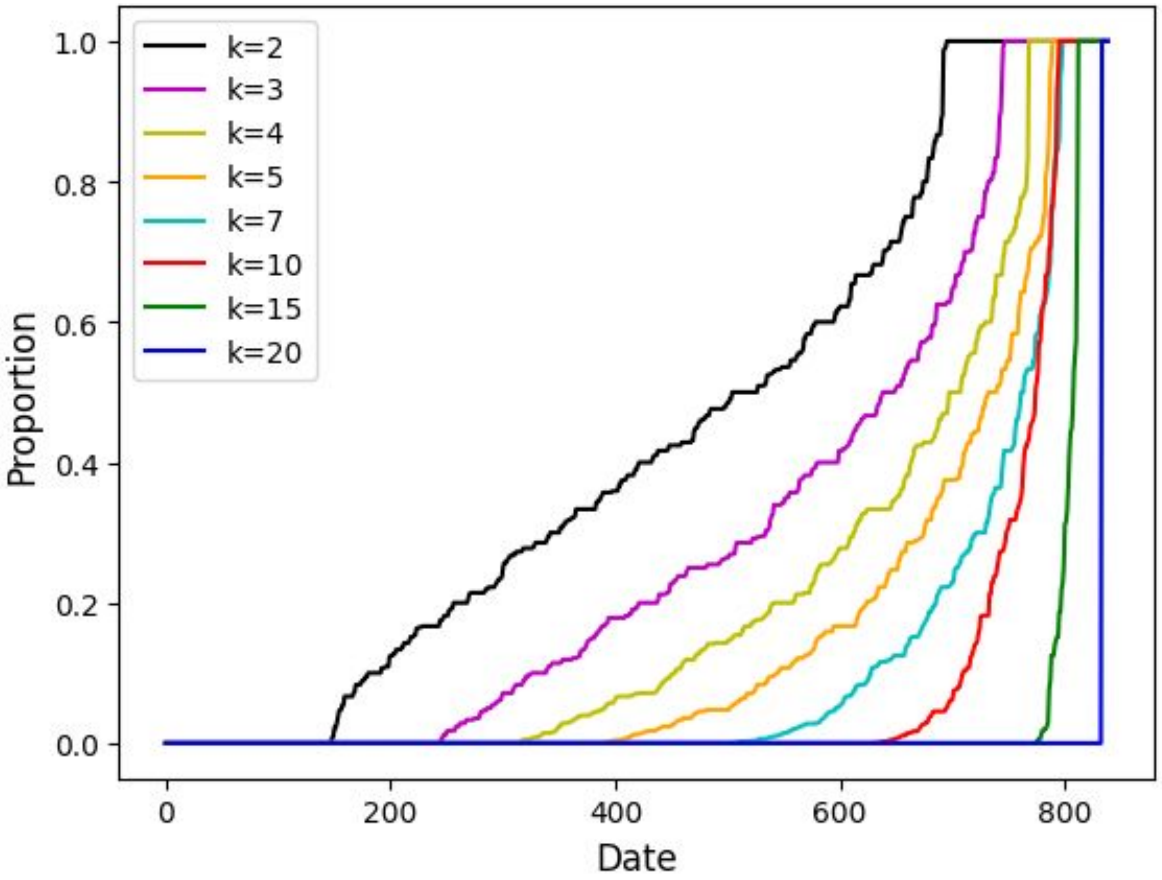


Comparing tokenized fractions with different levels of σ

Results : Discrete General Case

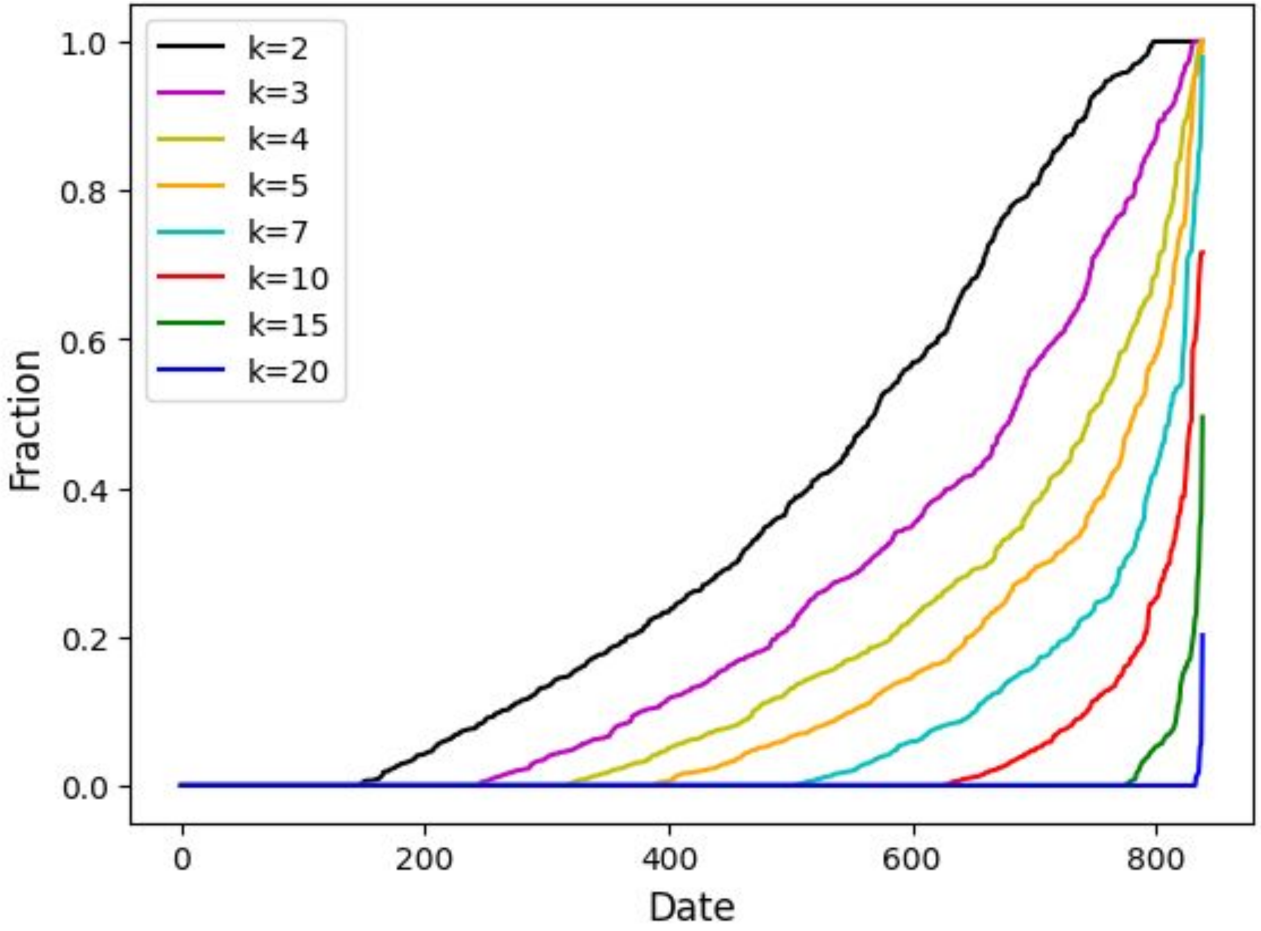


The proportion of data that have less than or equal to N assets to the total amount of data.



The proportion of the number of allowed vectors to the total number of possible vectors.

Results : Discrete General Case

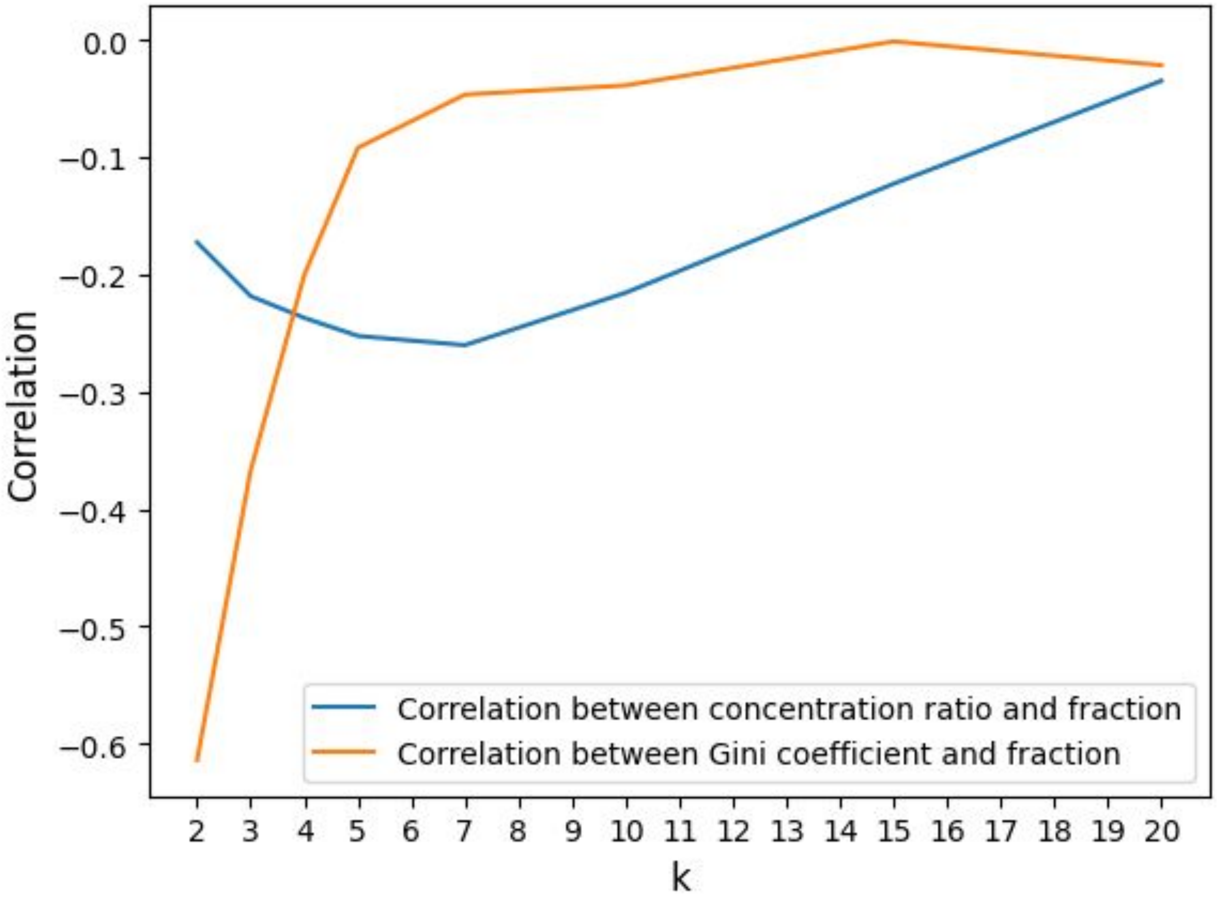


Sorted tokenized fractions for discrete general case

The number of chosen assets (k)	Tokenized Fraction
k=2	28.03 %
k=3	20.40 %
k=4	14.93 %
k=5	11.73 %
k=7	7.17 %
k=10	3.45 %
k=15	0.78 %
k=20	0.04 %

Tokenized fractions of the discrete general case vary with the amount of chosen assets

Results : Discrete General Case



Comparing correlation between market concentration and fraction and correlation between Gini coefficient and fraction with different number of chosen assets (k)

Discussion

	Discrete	Continuous
Homogeneous	optimal explicit solution	optimal explicit solution
Independent	NPH	optimal numerical solution
General	NPH	optimal numerical solution

computational complexity

Previous works
- Davydov and Yanovich 2020
- Chaleenutthawut et al. 2021

SOCP → upper bound
- Davidsson 2011

Research limitation: Based on single-collateral type

Discussion

	Discrete	Continuous
Homogeneous	optimal explicit solution	optimal explicit solution
Independent	NPH	optimal numerical solution
General	NPH	optimal numerical solution

computational complexity

Previous works
- Davydov and Yanovich 2020
- Chaleenutthawut et al. 2021

SOCP → upper bound
- Davidsson 2011

Research limitation: Based on single-collateral type

Skoltech

Conclusions

- Portfolio sold-out problem:
  – continuous independent case
  – continuous general case
  – discrete independent case
  – discrete general case

  Solved by SOCP

  NP-hard problem

- For continuous cases: SOCP → upper bound comparisons with MA

- For discrete cases:
  – k influences tokenized fractions.
  – market concentration → slightly impact on tokenized fractions.
  – Gini coefficient → significantly affects tokenized fractions for small k.

Skoltech

Scientific novelty

- Most researchers prioritize primary market portfolio optimization.
- Use blockchain to tokenize loans and create secondary market commodities.
- Blockchain offers new competitive advantages.

Skoltech

Innovation

- Emerging ventures, with a focus on DeFi lending protocols.
- Financial institutions and banks exploring cutting-edge solutions for loan portfolio management.

Skoltech

Outcomes

Paper: Currently, we are working on paper for publication.

Skoltech

Outlook

- Research on multi-collateral
- Apply tokenization to other datasets
- Cooperate with QuickToken.Ltd to make the software product

Acknowledgements

I would like to express my heartfelt gratitude to Yury Yanovich for the invaluable support and thoughtful suggestions provided.

Skoltech

Thank you for your attention

Q&A

Skoltech