**Professor Florentino DOMINGUEZ, MsC**
**E-mail: fdominguezv@esfm.ipn.mx**
**Instituto Politécnico Nacional**
**Professor Julio B. CLEMPNER, PhD**
**E-mail: julio@clempner.name**
**Instituto Politécnico Nacional**

# MULTIPERIOD MEAN-VARIANCE CUSTOMER CONSTRAINED PORTFOLIO OPTIMIZATION FOR FINITE DISCRETE-TIME MARKOV CHAINS

***Abstract.*** *The multi-period formulation aims at selecting an optimal investment strategy in a time-horizon able to maximize the final wealth while minimize the risk and determine the exit time. This paper is dedicated to solve the multi-period mean-variance customer constrained Markowitz's portfolio optimization problem employing the extraproximal method restricted to a finite discrete time, ergodic and controllable Markov chains for finite time horizon. The extraproximal method can be considered as a natural generalization of the convex programming approximation methods that largely simplifies the mathematical analysis and the economic interpretation of such model settings. We show that the multi-period mean-variance optimal portfolio can be decomposed in terms of coupled nonlinear programming problems implementing the Lagrange principle, each having a clear economic interpretation. This decomposition is a multi-period representation of single-period mean variance customer portfolio which naturally extends the basic economic intuition of the static Markowitz's model (where the investment horizon is practically never known at the beginning of initial investment decisions). This implies that the corresponding multi-period mean-variance customer portfolio is determined for a system of equations in proximal format. Each equation in this system is an optimization mean-variance problem which is solved using an iterating projection gradient method. Iterating these steps, we obtain a new quick procedure which leads to a simple and logically justified computational realization: at each iteration of the extraproximal method the functional of the mean-variance portfolio converges to an equilibrium point. We provide conditions for the existence of a unique solution to the portfolio problem by employing a regularized Lagrange function. We present the convergence proof of the method and all the details needed to implement the extraproximal method in an efficient and numerically stable way. Empirical results are finally provided to illustrate the suitability and practical performance of the model and the derived explicit portfolio strategy.*

***Keywords:****Multi-period portfolio extraproximal method Markov Chains optimization regularization.*

**JEL Classification: G11, C61, C69**



**DOI: 10.24818/18423264/53.1.19.03**

## 1.Introduction

### 1.1.Brief review

Portfolio selection is the problem of allocating given assets to securities drawn from a designated pool of securities for the purpose of maximizing portfolio return. The mean-variance analysis founded by [15] provides the fundamental basis for dealing with portfolio selection in single periods: a) mean refers to the endeavor to maximize the expected return of the portfolio return random variable, and b) variance, which is Markowitz's measure for risk, refers to the endeavor to minimize the variance of the portfolio return random variable. Some of its drawbacks are well known: the original Markowitz's formulation aims at selecting a single portfolio. This approach presents serious difficulties. However, the multi-period portfolio goal is to find an optimal investment strategy in a time-horizon $T$ selecting a set of intermediate portfolios instead of just one as originally proposed by Markowitz. The multi-period portfolio problem typically involves the expected return $U$, a negative quadratic risk term $V$ (variance) and constraints at time period$n = 1,2,\dots,T$. Let $c_n \in C_{adm} \subset \mathbb{R}^n$ denote the holdings of a portfolio of $N$ assets at time period$n = 1,2,\dots,T$. The entry $(c_n)_i$ denotes the amount of asset $i$ held in period $n$, considered as a real number indicating a short position. The initial portfolio $c_0$ is given, then the multi-period portfolio goal is to find $c_1,\dots,c_T$ maximize the final wealth while minimize the risk and determine the exit time. The variables in the multi-period portfolio problem are the sequence of positions $c_1,\dots,c_T$. We may also include additional constraints, such as "budget" constraint where $c_n$ represents the portfolio fractions invested in each asset. We refer to the solution of the problem

$$\Phi = E\left\{U(c_n) - \frac{\xi}{2}V(c_n)\right\} \to max\, c, \text{ subject to } c \geq 0, 1^T c = 1$$

where $\xi > 0$ is the risk aversion factor.

Merton's seminal work [16] proposed an analytical expression of the mean-variance efficient frontier in single-period portfolio selection. As a result, the mean-variance formulation motivates the development of extensions leading to several streams of research. One stream is on extending the original mean-variance single period model to a multi-period approach. Li and Ng [14] were the first in considering an analytical optimal solution to the mean-variance formulation in multi-period portfolio selection and proposed an algorithm for finding an optimal portfolio policy. Guo and Hu [12] studied the multi-period mean-variance portfolio optimization problem when exit time is uncertain. Zhu et al. [25] proposed a multi-period mean-variance model by incorporating a control of the probability of bankruptcy at each period. For different approaches to tackle the problem (see [13, 20])

There are many studies reported in the literature which try to efficiently or analytically solve the multi-period portfolio problem for Markov chains. Cakmak and

Ozekici [4] considered a multi-period mean-variance model where the model parameters change according to a stochastic market employing a similar solution technique used in [14] and [25] to obtain an explicit expression for the portfolio efficient frontier. They suggested that the mean vector and covariance matrix of the random returns of risky assets all depend on the state of the market during any period where the market process is assumed to follow a Markov chain. This work was improved in [5] considering several multi-period portfolio optimization models where the market consists of a riskless asset and several risky assets. An important feature of the improved model is that the stochastic evolution of the market is described by a Markov chain with perfectly observable states. Wei and Ye [21] suggested a multi-period mean-variance portfolio selection model imposed by a bankruptcy constraint in a stochastic market where the random returns of risky assets all depend on the state of the stochastic market, which is assumed to follow a Markov chain. Costa and Araujo [11] proposed a generalized multi-period mean-variance portfolio selection problem with the market parameters subject to Markov random regime switching to better respond to drastic movements of the security market. They suggested the necessary and sufficient conditions for obtaining an optimal control policy for this Markovian generalized multi-period mean-variance model, based on a recursive procedure. Wu and Li [22] investigated a multi-period mean-variance portfolio selection with regime switching and uncertain exit time where the returns of assets all depend on the states of the stochastic market which are assumed to follow a finite homogeneous Markov chain. Wu and Li [23] studied a non-self-financing portfolio optimization problem under the framework of multi-period mean–variance with Markov regime switching and a stochastic cash flow. Yao et al. [24] considered an uncertain exit time multi-period mean–variance portfolio selection problem with endogenous liabilities in a Markov jump market, where assets and liabilities of the balance sheet are simultaneously optimized. Chen et al. [6] develop an investment model with time consistency in Markovian markets involving a nested conditional expectation mapping and examined the differences of the investment policies with a riskless asset from those without a riskless asset. Bannister et al. [3] studied a multi-period portfolio selection problem in which a single period mean-standard-deviation criterion is used to construct a separable multi-period selection criterion obtaining a closed form optimal strategy which depends on selection schemes of investor's risk preference. Clempner and Poznyak [8] considered the subject of penalty regularized expected utilities and investigates the applicability of the method for computing the mean-variance Markowitz customer portfolio optimization problem.

### 1.2. Main results

We consider the modeling and solution of the multi-period mean-variance customer constrained Markowitz's portfolio optimization problem in Markov chains. The proposed multi-period mean-variance model pioneered by [19, 18, 1] provides a solution method able to find an optimal investment strategy in a finite time-horizon which to maximize the final wealth while minimize the risk and determine the exit time. For solving the problem, we present a two-step iterated procedure for the extraproximal method: a) *the first step* (the extra-proximal step) consists of a "prediction" which calculate the preliminary position approximation to the equilibrium point, and b) *the second step* is designed to find a "basic adjustment" of the previous prediction. The first step computes the direction of the future evolution at a given point and, the second step makes the proximal step from the same point along the predicted direction. In this sense, the prediction step can be conceptualized as a feed-back for the simplest proximal approach. Each equation in this system is an optimization mean-variance problem which is solved using an iterating projected gradient method. This research presents the following main contributions:

- We employ the extraproximal method to find analytically the optimal policy of the multi-period mean-variance customer constrained Markowitz's portfolio problem formulation for finite, homogeneous, ergodic and controllable Markov chains. The states of Markov chain are interpreted as the states of an economy.
- The extraproximal method simplifies the mathematical analysis and the economic interpretation of the Makovitz's model: the multi-period mean-variance optimal portfolio formulation can be decomposed in terms of coupled nonlinear programming problems implementing the Lagrange principle, each having a clear economic interpretation.
- The multi-period representation naturally extends the basic economic intuition of the static Markowitz's model for a single-period mean variance customer portfolio.
- We present a two-step iterated procedure for solving the extraproximal method: a) *the first step* (the extra-proximal step) consists of a "prediction" which calculate the preliminary position approximation to the equilibrium point, and b) *the second step* is designed to find a "basic adjustment" of the previous prediction.
- Each equation in this system is an optimization mean-variance problem which is solved using an iterating projection gradient method. The algorithm finds the optimal multi-period portfolio policy that maximize the expected utility value and the variance of the terminal wealth.
- We also present the convergence proof of the method.
- We provide conditions for the existence of a unique solution to the portfolio problem by employing a regularized Lagrange function.
- We present all the details needed to implement the extraproximal method.
- A numerical example illustrates the usefulness of our proposed method for multi-period mean-variance customer portfolio.

### 1.3.Organization of the paper

The rest of the paper is organized as follows. Section 2 introduces the main concepts related to homogeneous Markov chains needed to understand the rest of the paper. Section 3 derives an explicit solution for the multi-period Markowitz's portfolio optimization presenting the extraproximal method in terms of a coupled nonlinear programming problems implemented using the Lagrange principle. Section 4 presents the convergence analysis of the portfolio method. Section 5 illustrates the usefulness of our proposed method for multi-period mean-variance customer portfolio by discussing specific issued in a numerical example. Section 6 concludes and presents future work.

## 2.Homogeneous Markov chains model

Let us define a probability space$(\Omega,\mathcal{F},P)$ where $\Omega$ is a set of elementary events,$\mathcal{F}$ is the minimal $\sigma-$algebra of the subsets of $\Omega$, and P is a given probabilistic measure defined for any$\mathcal{A}\in\mathcal{F}$. Let us also consider the natural sequence $n=1,2,\ldots$ as a time argument. Let $S$ be a finite set consisting of states $\{s_1,\ldots,s_n\}$, $n\in\mathbb{N}$, called the *state space*. A *Stationary Markov chain* [17, 7] is a sequence of $S$-valued random variables $s(n),n\in\mathbb{N}$, satisfying the *Markov condition*:

$$P(s(n+1)=s_j|s(n)=s_{i_n},\ldots,s(1)=s_{i_1})=P\big(s(n+1)=s_j|s(n)=s_i\big)=:\pi_{ij}.$$

The random variables $s(n)$ are defined on the probability space$(\Omega,\mathcal{F},P)$ and take values in $S$. The Markov chain can be represented by a complete graph whose nodes are the states, where each edge $(s_i,s_j)\in S^2$ is labeled by the transition probability. The matrix $\Pi=(\pi_{ij})_{(s_i,s_j)\in S}\in[0,1]^{N\times N}$ determines the evolution of the chain: for each $n\in\mathbb{N}$, the power $\Pi^n$ has in each entry $(s_i,s_j)$ the probability of going from state $s_i$ to state $s_j$ in exactly$n$ steps.

**Definition 1.***A controllable finite homogeneous Markov chain ([17]) is a 4-tuple*

$$MC=\{S,A,\mathcal{K},\Pi\} \tag{1}$$

where:

- $S$ is a finite set of states.
- $A$ is the set of actions, which is a metric space. For each $s\in S, A(s)\subset A$ is the non-empty set of admissible actions at state $s\in S$. Without loss of generality we may take $A=\cup_{s\in S}A(s)$;
- $\mathcal{K}=\{(s,a)|s\in S,a\in A(s)\}$ is the set of admissible state-action pairs, which is a finite subset of $S\times A$;
- $\Pi=\left[\pi_{j|ik}\right]$ is a controlled transition matrix, where

$$\pi_{j|ik} \equiv P(s(n+1) = s_j | s(n) = s_i, a(n) = a_k)$$

represents the probability associated with the transition from state $s_i$ to state $s_j$ under an action $a_k \in A(s_i), k = 1, \dots, M,\ M \in \mathbb{N}$.

**Definition 2.** *A Markov Decision Process is a pair*

$$MDP = \{MC, U\} \tag{2}$$

where:

- MC is a controllable Markov chain (1)
- $U: S \times \mathcal{K} \to \mathbb{R}$ is a utility function, associating to each state a real value.

The *Markov property* of the decision process in (2) is said to be fulfilled if $P(s(n+1)|(s(1), \dots, s(n-1)), s(n), a(n)) = P(s(n+1) = s_j | s(n) = s_i, a(n) = a_k)$.

A sequence of random stochastic matrices $D = \{d_{k|i}(n)\}_{k=\overline{1,M}, i=\overline{1,N}}$ is said to be a *randomized control strategy* if:a) it is causal (independent on the future), that is, $d_{k|i}(n) = [d_{k|i}(n)]_{k=\overline{1,M}; i=\overline{1,N}}$ is $\mathcal{F}_{n-1}$-measurable where $\mathcal{F}_{n-1} :=$ $\sigma(s(1), a(1), d(1); \dots; s(n-1), a(n-1), d(n-1))$is the $\sigma$ −algebra generated by $(s(1), a(1), d(1); \dots; s(n-1), a(n-1), d(n-1))$; b) the random variables $(a(1), \dots, a(n-1))$ represent the "realizations" of the control actions, taking values on the finite set $A = \{a_1, \dots, a_M\}$ and satisfy the following property

$$d_{k|i}(n) = P(a(n) = a_k | s(n) = s_i \wedge \mathcal{F}_{n-1} \tag{3}$$

which represents the probability measure associated with the occurrence of an action $a(n)$ from state $s(n) = s_i$.

Denote by $\Sigma$ the class of all randomized strategies $D$, that is,

$$\Sigma = \{d_{k|i}(n)\}$$

Considering Eq. (3) for any fixed strategy $D = \{d_{k|i}(n)\} \in \Sigma$ the conditional transition probability matrix $\Pi(d_{k|i}(n))$ can be defined as follows $\Pi(d_{k|i}(n)) = P(s(n+1) = s_j | s(n) = s_i) = \sum_{k=1}^{M} P(s(n+1) = s_j | s(n) = s_i, a(n) = a_k)\, d_{k|i}(n)$which represents the probability to move from the states $s_j$ to the state $s_i$ under the applied mixed strategy $d_{k|i}(n)$.

In the case of complete information on the payoff and transition matrices, *the dynamics of the Markov chain* is described as follows. The dynamics begins at the initial state $s(0)$ which (as well as the states further realized by the process) is assumed to be completely measurable. Each strategy is allowed to randomize, with distribution $d_{k|i}(n)$, over the pure action choices $a_k \in A(s_i), i = \overline{1,N}$ and $k = \overline{1,M}$. These choices induce immediate utilities $U$. The system tries to maximize the corresponding one-step utility. Next, the system moves to new state $s(n+1) = s_j$ according to the transition probabilities $\Pi(d_{k|i}(n))$. Based on the obtained utility, the systems adapt it mixed strategy computing $d_{k|i}(n+1)$ for the next selection of the control actions. In addition, these choices induce the state distribution dynamics as follows

$$P\big(s(n+1) = s_j\big) = \sum_{i=1}^{N}\left(\sum_{k=1}^{M} \pi_{j|ik} d_{k|i}\right) P(s(n) = s_i).$$

Considering ergodic Markov chains [7], for any fixed collection of stationary strategies $d_{k|i}(n) = d_{k|i}$ we have $P\big(s(n+1) = s_j\big) \to P\big(s = s_j\big), n \to \infty$, that is, in the ergodic case when the Markov chain is ergodic for any stationary strategy $d_{k|i}$ the distributions $P\big(s(n+1) = s_j\big)$ exponentially fast converge to their limits $P(s_i)$ satisfying

$$P(s_j) = \sum_{i=1}^{N}\left(\sum_{k=1}^{M} \pi_{j|ik} d_{k|i}\right) P(s_i). \tag{4}$$

Obviously, $P(s_j)$ is a function of $d_{k|i}$. The utility function is given by the values $W_{ik}$, so that the "average utility function" $\mathbf{U}$ in the stationary regime can be expressed as

$$\boldsymbol{U}(d_{(k|i)}) := \sum_{i,k} W_{ik} d_{k|i} P(s_i) \tag{5}$$

where$W_{ik} = \sum_j U_{ijk} \pi_{j|ik}$.

The change of variable, suggested below, significantly simplifies the representation of the payoff functions converting implicit nonlinear function $\boldsymbol{U}(d_{k|i})$ in to a polylinear one. To do so, let us introduce a matrix of elements $c := \big[c_{i|k}\big]_{i=\overline{1,N};k=\overline{1,M}}$ according to the following formula:

$$c_{i|k} = d_{k|i} P(s_i) \tag{6}$$

The admissible strategies ($c_{i|k} \in C_{adm}$) will be limited by the following requirements:

- each matrix $c_{i|k}$ represents a stationary mixed strategy, and, hence, belongs to the simplex $\Delta^{MN}$ defined by

$$\Delta^{MN} = \left\{ c_{i|k} \in \mathbb{R}^{MN} \,\middle|\, \sum_{i,k} c_{i|k} = 1, c_{i|k} \geq 0 \right\} \tag{7}$$

- the joint strategy variable $c_{i|k}$ satisfies the "ergodicity constraints" (4) and, hence, belongs to the convex, closed and bounded set

$$\mathcal{E} = \left\{ c_{i|k} \in \mathbb{R}^{MN} | \sum_{k} c_{j|k} = \sum_{i,k} \pi_{j|ik} c_{i|k} \right\}, \sum_{k} c_{i|k} > 0 \tag{8}$$

Then, $c_{i|k} \in \mathcal{C}_{adm} \coloneqq \Delta^{MN} \times \mathcal{E}$. Notice that by (6) it follows that

$$P(s_i) = \sum_{k} c_{i|k} \qquad d_{(k|i)} = \frac{c_{i|k}}{\sum_k c_{i|k}}. \tag{9}$$

In terms of $c$-variables the utility function $\boldsymbol{U}$ becomes: $\boldsymbol{U}(c_{i|k}) = \sum_{i=1}^{N} \sum_{k=1}^{M} W_{ik} c_{i|k}$.

## 3.Multi-period Markowitz's portfolio optimization

The original Markowitz's formulation aims at selecting a single portfolio where the investment horizon is practically never known at the beginning of initial investment decisions. The multi-period portfolio goal is to find an optimal investment strategy in a time-horizon able to maximize the final wealth while minimize the risk and determine the exit time.

### 3.1.Single period optimization

One may formally state Markowitz's decision model for mean-variance customer portfolio as follows [19, 18, 1]. We define $U(\alpha, c)$ as the wealth available for investment

$$\boldsymbol{U}(\alpha, c) := \sum_{i,k} W_{ik} \alpha_{k|i} c_{k|i} \to max_{\alpha \in A_{adm}, c \in \mathcal{C}_{adm}}$$

where $W_{ik} := \sum_{j=1}^{N} U_{ijk} \pi_{j|ik}$ and $c_{i|k} := d_{k|i} P(s_i)$.and the *variance* ($\mathrm{Var}(\alpha, c)$) as

$$\mathrm{Var}(\alpha, c) := \sum_{i,k} \left[ \alpha_{k|i} W_{ik} - \boldsymbol{U}(\alpha, c) \right]^2 c_{k|i} = \sum_{i,k} {\alpha_{k|i}}^2 {W_{ik}}^2 c_{k|i} - \boldsymbol{U}(\alpha, c)^2$$
$$\to max_{\alpha \in A_{adm}, c \in \mathcal{C}_{adm}}$$

The resulting *customer portfolio optimization* problem includes a model-user's tolerance for risk, and it is represented by the following expression:

$$\Phi(\alpha, c) \coloneqq \boldsymbol{U}(\alpha, c) - \frac{\xi}{2} \mathrm{Var}(\alpha, c) \to max_{\alpha \in A_{adm}, c \in \mathcal{C}_{adm}} \tag{10}$$

and $\frac{\xi}{2}$ is *the risk-aversion* parameter.

### 3.2.Multi-period portfolio

We define $\boldsymbol{U}(d_{(k|i)}(n))$ as the wealth available for investment and $d_{k|i}(n)$ as the amounts invested in the risky assets at time $n$ $(n = 1,2,\ldots,T)$. A policy $\{d(n)\}_{n \geq 0}$ maximizes the conditional mathematical expectation of the utility function $U(d_{k|i}(n))$ under the condition that the history of the process

$$\mathcal{F}_n \coloneqq \left\{ D_0, P\{s_0 = s_j\}_{j=\overline{1,N}}; \ldots; D_{n-1}, P\{s_n = s_j\}_{j=\overline{1,N}} \right\}$$

is fixed and cannot be changed hereafter, i.e., it realizes the conditional optimization rule

$$d(n) = argmax_{d(n)\in D} E\left\{\boldsymbol{U}\left(d_{(k|i)}\right)|\mathcal{F}_n\right\} \tag{11}$$

where $\boldsymbol{U}(d_{k|i}(n))$ is the utility function at the state $s_{(n)}$ and where $E\{\cdot\}$ denotes the expectation of "$\cdot$". Then, the wealth dynamics is given by the values $\boldsymbol{U}(d_{k|i}(n))$, so that the "average utility function" $\boldsymbol{U}$ for the multi-period portfolio can be expressed as

$$\boldsymbol{U}(d_{k|i}(n)) := \sum_{i,k} W_{ik} d_{k|i}(n) P(s_i) \tag{12}$$

where $W_{ik} = \sum_j U_{ijk}\pi_{j|ik}$.

The investor wants to find an optimal investment strategy in a finite time-horizon to maximize his/her final wealth while minimize his/her risk. Then, we formulate the multi-period portfolio selection problem as follows. We consider diversification with respect to the number of customers chosen in the portfolio problem. Then, $\alpha_{i|k}(n)$ is the *number of customers* at state $i$ applying action $k$, $0 \le \alpha_{i|k}(n)$. The *utility* $\boldsymbol{U}$ of a customer portfolio is calculated as the sum of the weighted net presenting the value $u$ as follows

$$\boldsymbol{U}(\alpha(n), c(n)) := \sum_{i,k} W_{ik}\alpha_{k|i}(n) c_{k|i}(n) \to max_{\alpha\in A_{adm}, c\in\mathcal{C}_{adm}}$$

where $W_{ik} := \sum_{j=1}^{N} U_{ijk}\pi_{j|ik}$ and $c_{i|k}(n) := d_{k|i}(n)P(s_i)$.

The customer portfolio optimization problem [15] attempts to maximize the *mean value* ($U(\alpha(n), c(n))$) generated by all the customers while minimizing the *variance* ($\mathrm{Var}(U(\alpha(n), c(n)))$)

$$\begin{aligned} \mathrm{Var}(\alpha(n), c(n)) &: \\ &= \sum_{i,k}\left[\alpha_{k|i}(n)W_{ik} - \boldsymbol{U}(\alpha(n), c(n))\right]^2 c_{k|i} \\ &= \sum_{i,k} {\alpha_{k|i}}^2(n){W_{ik}}^2 c_{k|i}(n) - \boldsymbol{U}(\alpha(n), c(n))^2 \to max_{\alpha\in A_{adm}, c\in\mathcal{C}_{adm}} \end{aligned}$$

For practical purposes, the resulting *customer portfolio optimization* problem includes a model-user's tolerance for risk, and it is represented as follows:

$$\Phi(\alpha(n), c(n)) \coloneqq \boldsymbol{U}(\alpha(n), c(n)) - \frac{\xi}{2}\mathrm{Var}(\alpha(n), c(n)) \to max_{\alpha\in A_{adm}, c\in\mathcal{C}_{adm}} \tag{13}$$

where

$$A_{adm} = \begin{cases} \alpha = [\alpha_{i|k}(n)]_{i=\overline{1,N},k=\overline{1,M}} : \sum_{i=1}^{N}\sum_{k=1}^{M}\alpha_{i|k}(n) \leq \alpha^{+}, \\ \alpha_{i|k}(n) \in [\varepsilon, \alpha^{+}], \quad \varepsilon > 0 \end{cases} \tag{14}$$

The variable $\varepsilon > 0$ is introduced for avoiding divisions by 0. $\alpha^{+}$ denotes de maximum number of clients.

Because, the purpose is to obtain a higher mean value return ($\boldsymbol{U}(\alpha(n), c(n))$) also the corresponding risk level ($\mathrm{Var}(\alpha(n), c(n))$ increases. Here the goal is to find the values of $\alpha(n)$ and $c(n)$ that maximize the objective function in Eq. (13) subject to the following constrains:

$$\sum_{i=1}^{N}\sum_{k=1}^{M}\alpha_{i|k}(n)c_{i|k}(n)\eta_{i|k} \leq b_{ineq} \tag{15}$$

where $\eta_{i|k}$ are the resources destined for carrying out in state $i$ a promotion *k* and the admissible sets are as in Eq. (7), Eq. (8). The following optimization properties are the key to find efficient portfolios: a) the Markowitz model in Eq. (13) is a quadratic optimization problem (quadratic objective function and linear constraints in Eq. (15), Eqs. (7)-(8) and Eq. (14)), b) the feasibility set $C_{adm}$ is convex since it is the intersection of hyperplanes, c) the factor 1/2 of *the risk-aversion* parameter ξ is chosen for notational convenience and, d) the parameter $b_{ineq}$ is endogenously given (the budget is chosen by the decision maker in the respective model). The mean-variance Markowitz's portfolio is given by:

$$\begin{aligned}
&\mathcal{L}_{\theta,\delta}\left(c(n),\alpha(n),\lambda\left(\lambda_0,\lambda_{N+1},\lambda_1\right)\right) := \theta\Big[\textstyle\sum_{i=1}^{N}\sum_{k=1}^{M} W_{ik}\alpha_{i|k}(n)c_{i|k}(n) + \\
&\tfrac{\xi}{2}\textstyle\sum_{i=1}^{N}\sum_{k=1}^{M} W_{ik}\alpha_{i|k}(n)c_{i|k}(n)\sum_{\hat{\imath}=1}^{N}\sum_{\hat{k}=1}^{M} W_{\hat{\imath}\hat{k}}\alpha_{i|\hat{k}}(n)c_{\hat{\imath}|\hat{k}}(n) - \tfrac{\xi}{2}\sum_{i=1}^{N}\sum_{k=1}^{M} \Big] \\
&\textstyle\sum_{j=1}^{N}\lambda_{0,j}\left[\left(\sum_{i=1}^{N}\sum_{k=1}^{M}\pi_{j|ik}c_{i|k}(n) - \sum_{k=1}^{M}c_{j|k}(n)\right) - b_{eq,j}\right] - \\
&\lambda_{N+1}\left(\sum_{i=1}^{N}\sum_{k=1}^{M}c_{i|k}(n) - b_{eq,N+1}\right) - \lambda_1\left(\sum_{i=1}^{N}\sum_{k=1}^{M}\alpha_{i|k}(n)c_{i|k}(n)\eta_{i|k} - b_{ineq}\right) \\
&+\frac{\delta}{2}\left(-\|c(n)\|^2 - \|\alpha(n)\|^2 + \|\lambda_0\|^2 + \lambda_{N+1}^2 + \|\lambda_1\|^2\right)
\end{aligned} \tag{16}$$

The proximal regularization terms $\theta$ and $\delta$, for the Markowitz's *Regularized Portfolio Lagrange function (RPLF)*[10, 9], can be viewed as an additional quadratic risk and expected return for each asset that ensures the convergence to a unique portfolio.

### 3.3. Multi-period portfolio optimization method

The initial portfolio $c_0$ and initial number of clients $\alpha_0$ are given. The goal is to find the portfolio $c^*$ which maximizes the utility and minimizes the variance and the optimal number of $\alpha^*$. We also include additional constraints related to the Markov restrictions and the number of clients of the mean-variance portfolio. The Lagrangian of mean-variance Markowitz portfolio is given in Eq. (16) by $\mathcal{L}(c,\alpha,\lambda)$. For computing the multi-period portfolio optimization problem, we employ the extraproximal method following iterative formulas [2] given by:

**First half-step** (prediction)

$$\begin{aligned}\bar{\lambda}_n &= argmin_{\lambda\geq 0}\left\{\frac{1}{2}\|\lambda-\lambda_n\|^2+\gamma\mathcal{L}(c_n,\alpha_n,\lambda)\right\}\\ \bar{c}_n &= argmax_{c\in\mathcal{C}_{adm}}\left\{-\frac{1}{2}\|c-c_n\|^2+\gamma\mathcal{L}(c,\alpha_n,\bar{\lambda}_n)\right\}\\ \bar{\alpha}_n &= argmax_{\alpha\in A_{adm}}\left\{-\frac{1}{2}\|\alpha-\alpha_n\|^2+\gamma\mathcal{L}(c_n,\alpha,\bar{\lambda}_n)\right\}\end{aligned} \tag{17}$$

**Second half-step** (approximation):

$$\begin{aligned}\lambda_{n+1} &= argmin_{\lambda\geq 0}\left\{\frac{1}{2}\|\lambda-\lambda_n\|^2+\gamma\mathcal{L}(\bar{c}_n,\bar{\alpha}_n,\lambda)\right\}\\ c_{n+1} &= argmax_{c\in\mathcal{C}_{adm}}\left\{-\frac{1}{2}\|c-c_n\|^2+\gamma\mathcal{L}(c,\bar{\alpha}_n,\bar{\lambda}_n)\right\}\\ \alpha_{n+1} &= argmax_{\alpha\in A_{adm}}\left\{-\frac{1}{2}\|\alpha-\alpha_n\|^2+\gamma\mathcal{L}(\bar{c}_n,\alpha,\bar{\lambda}_n)\right\}\end{aligned} \tag{18}$$

Each iteration of formulas presented in Eq. (17) and Eq. (18) has a natural interpretation and involves three nonlinear equations, corresponding to evaluation of the three extraproximal operators. Evaluating the extraproximal Eq. (17) and Eq. (18) of the objective involves solving three related optimization problems, one for each possible sequence of outcomes. The first step computes the direction of the future evolution of the portfolio at a given point and, the second step makes the proximal step from the same point along the predicted direction of the portfolio. In this sense, the prediction step can be conceptualized as a feed-back for the simplest proximal approach of the portfolio.

**Remark 3.** *The regularization term can be viewed as an additional quadratic risk and expected return for each asset.*

**4.Convergence analysis of the portfolio method**

The convergence of any process and the estimation of the rate of convergence depend on the behavior of the objective function given in Eq.(13) in the neighborhood of the solution of the problem.

Given $c=(x,w)$ let us define the extended variables $\tilde{x}:=(x)\in\tilde{X}:=X, \tilde{y}:=(w,\lambda)^T\in\tilde{Y}:=W\times\mathbb{R}^+, \widetilde{w}=(\widetilde{w}_1,\widetilde{w}_2)^T\in\tilde{X}\times\tilde{Y}, \tilde{v}=(\tilde{v}_1,\tilde{v}_2)^T\in\tilde{X}\times\tilde{Y}$ and the functions, the functions

$$\tilde{\mathcal{L}}(\tilde{x},\tilde{y}):=f(c,\lambda), \phi(\widetilde{w},\tilde{v}):=\tilde{\mathcal{L}}(\widetilde{w}_1,\tilde{v}_2)-\tilde{\mathcal{L}}(\tilde{v}_1,\widetilde{w}_2).$$

For $\widetilde{w}_1=\tilde{x}, \widetilde{w}_2=\tilde{y}, \tilde{v}_1=\tilde{v}_1^*=\tilde{x}^*$ and $\tilde{v}_2=\tilde{v}_2^*=\tilde{y}^*$ we have

$$\phi(\widetilde{w},\tilde{v}):=\tilde{\mathcal{L}}(\tilde{x},\tilde{y}^*)-\tilde{\mathcal{L}}(\tilde{x}^*,\tilde{y})$$

In these variables the extraproximal approach can be represented as

$$\tilde{v}^*=arg\min_{v\in\tilde{X}\times\tilde{Y}}\left\{\frac{1}{2}\|\widetilde{w}-\tilde{v}^*\|^2+\gamma\phi(\widetilde{w},\tilde{v}^*)\right\} \tag{19}$$

*Let $f(z)$ be a convex function defined on a convex set $Z$. If $z^*$ is a minimizer of function*

$$\Phi(z)=\frac{1}{2}\|z-x\|^2+\alpha f(z) \tag{20}$$

on $Z$ with fixed $x$, then $f(z)$ satisfies the inequality

$$\frac{1}{2}\|z^*-x\|^2+\alpha f(z^*)\le\frac{1}{2}\|z-x\|^2+\alpha f(z)-\frac{1}{2}\|z-z^*\|^2 \tag{21}$$

*If all partial derivative of $\tilde{\mathcal{L}}(\tilde{x},\tilde{y})$ satisfy the Lipschitz condition with positive constant $C$, then the following Lipschitz-type condition holds:*

$$\|[\phi(\widetilde{w}+h,\tilde{v}+g)-\phi(\widetilde{w},\tilde{v}+g)]-[\phi(\widetilde{w}+h,\tilde{v})-\phi(\widetilde{w},\tilde{v})]\| \le C\|h\|\|g\| \tag{22}$$

valid for any $\widetilde{w},h,\tilde{v},g\in\tilde{X}\times\tilde{Y}$.

The following theorem presents the convergence conditions of (17) - (18) and gives the estimate of its rate of convergence.

**Theorem 6.** *Assume that the portfolio given in Eq. (13) has a solution. The function $\tilde{\mathcal{L}}(\tilde{x},\tilde{y})$ is differentiable in $\tilde{x}$ and $\tilde{y}$, whose partial derivative with respect to $\tilde{y}$ satisfies the Lipschitz condition. For any $\delta\in(0,1)$, there exists a small-enough $0<\gamma_0<\frac{1}{\sqrt{2}C}$ such that, for any $0<\gamma\le\gamma_0$. Then, the sequence $\{\tilde{v}_n\}$ generated by the extraproximal procedure (17) - (18), monotonically converges in norm with geometric progression rate $\rho\in(0,1)$ to a portfolio $\dot{\tilde{v}}$, i.e., $\tilde{v}_n \underset{n\to\infty}{\rightarrow} \dot{\tilde{v}}$.*

*Proof.* See the appendix

**Corollary 7.** *Assume that the portfolio given in Eq. (13) has a solution. The function $\tilde{\mathcal{L}}(\tilde{x},\tilde{y})$ is differentiable in $\tilde{x}$ and $\tilde{y}$, whose partial derivative with respect to $\tilde{y}$ satisfies the Lipschitz condition. For any $\delta\in(0,1)$, there exists a small-enough $0<$*

*$\gamma_0 < \frac{1}{\sqrt{2}C}$ such that, for any $0 < \gamma \le \gamma_0$. Then, we have that $\lambda_n \underset{n\to\infty}{\to} \lambda^*$ and $\alpha_n \underset{n\to\infty}{\to} \alpha^*$*

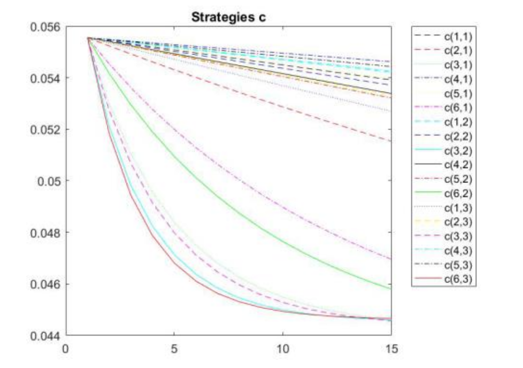


**Figure 1: Multi-period portfolio - strategies c**

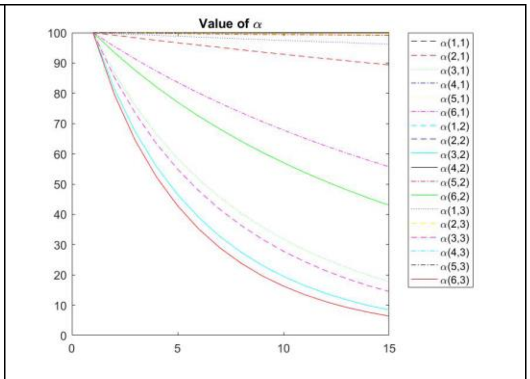


**Figure 2: Multi-period portfolio - number of customers**

**5.Numerical example**

We conclude this section by providing a numerical example to illustrate the practical performance of the extraproximal method in the context of mean-variance customer portfolio. The extraproximal method developed in Eq. (17) and Eq. (18) cannot be directly applied to solve the portfolio choice problem. In this section, we show how one can approximate the solution to such a problem solving a sequence of portfolio choice problems. We consider the static portfolio choice problem with a strictly concave utility function given by the regularized Lagrangian given by $\mathcal{L}(c,\alpha,\lambda)$ (see the appendix). We consider the problem of maximizing the expected utility $\mathrm{E}\{\Phi(\alpha,c)\}$ where $\Phi(\cdot,\cdot)$has a risk aversion $\frac{\xi}{2}$. Finally, we illustrate that, as $n \to \infty$, the sequence $\{\tilde{v}_n\}$ converges to the optimal solution$\dot{\tilde{v}}$ of the portfolio choice problem described in Eq. (17) and Eq. (18).

The resulting multi-period portfolio (see Figure 1 and Figure 2) generated by the recurrent method are given, for instance, by

$$d^*_{i|k}(n=2) = \begin{bmatrix} 0.3387 & 0.3375 & 0.3238 \\ 0.3353 & 0.3350 & 0.3296 \\ 0.3402 & 0.3400 & 0.3198 \\ 0.3358 & 0.3358 & 0.3284 \\ 0.3385 & 0.3388 & 0.3226 \\ 0.3407 & 0.3408 & 0.3185 \end{bmatrix} \qquad \alpha^*_{i|k}(n=2) = \begin{bmatrix} 100.0000 & 99.1434 & 86.7730 \\ 100.0000 & 100.0000 & 95.4821 \\ 100.0000 & 100.0000 & 81.5175 \\ 99.9876 & 99.9175 & 93.4628 \\ 99.6989 & 99.9570 & 85.3000 \\ 100.0000 & 100.0000 & 79.6686 \end{bmatrix}$$

$$d^*_{i|k}(n=7) = \begin{bmatrix} 0.3530 & 0.3458 & 0.3012 \\ 0.3433 & 0.3415 & 0.3152 \\ 0.3534 & 0.3520 & 0.2946 \\ 0.3445 & 0.3440 & 0.3115 \\ 0.3496 & 0.3514 & 0.2991 \\ 0.3532 & 0.3537 & 0.2931 \end{bmatrix} \quad \alpha^*_{i|k}(n=7) = \begin{bmatrix} 100.0000 & 95.0994 & 45.5780 \\ 100.0000 & 100.0000 & 76.7207 \\ 100.0000 & 100.0000 & 32.6517 \\ 99.9721 & 99.5537 & 68.1221 \\ 98.2687 & 99.7829 & 41.5909 \\ 100.0000 & 100.0000 & 28.8877 \end{bmatrix}$$

$$d^*_{i|k}(n=15) = \begin{bmatrix} 0.3593 & 0.3434 & 0.2973 \\ 0.3512 & 0.3469 & 0.3019 \\ 0.3554 & 0.3521 & 0.2925 \\ 0.3503 & 0.3492 & 0.3005 \\ 0.3500 & 0.3540 & 0.2960 \\ 0.3538 & 0.3550 & 0.2912 \end{bmatrix} \quad \alpha^*_{i|k}(n=15) = \begin{bmatrix} 100.0000 & 89.3542 & 17.7208 \\ 100.0000 & 100.0000 & 55.6880 \\ 100.0000 & 100.0000 & 8.3732 \\ 100.0000 & 99.1186 & 42.9569 \\ 96.2191 & 99.6327 & 14.4388 \\ 100.0000 & 100.0000 & 6.3474 \end{bmatrix}$$

## 6.Conclusion and future work

This paper developed a solution approach for the multi-period mean-variance customer constrained Markowitz's portfolio optimization problem based-on the extraproximal method which maximize the final wealth while minimize the risk and determine the exit time in finite discrete time, ergodic and controllable Markov chains for a finite time horizon. The method simplifies the mathematical analysis of the Markowitz model and highlights it economic structure. We showed that multi-period mean-variance customer optimal portfolio can be decomposed in terms of coupled nonlinear programming problems implementing the Lagrange principle, each having a clear economic interpretation. The extraproximal method is two-step iterated procedure where a) *the first step* consists of a "prediction" which calculate the preliminary position approximation to the equilibrium point, and b) *the second step* is designed to find a " basic adjustment" of the previous prediction of the portfolio. We also presented the convergence proof of the method for the multi-period portfolio. We provided conditions for the existence of a unique solution to the portfolio problem by employing a regularized Lagrange function. We concluded this paper by providing a numerical example to show the practical performance of the extraproximal method in the context of multi-period mean-variance customer portfolio. Future research aims at incorporating in the model the corresponding multi-period mean-variance frontiers, the impact of taking transitions costs into account and the consideration of a penalty approach on the implied Markowitz model.

**Appendix.**

**Proof of Theorem 6**

*Proof.*1) Taking in (21) $\alpha = \gamma$
and $z = \widetilde{w}, x = \tilde{v}_n, z^* = \hat{v}_n f(z) = \phi(\widetilde{w}, \tilde{v}_n), f(z^*) = \phi(\hat{v}_n, \tilde{v}_n)$ we obtain

$$\frac{1}{2}\|\hat{v}_n - \tilde{v}_n\|^2 + \gamma\phi(\hat{v}_n, \tilde{v}_n) \leq \frac{1}{2}\|\widetilde{w} - \tilde{v}_n\|^2 + \gamma\phi(\widetilde{w}, \tilde{v}_n) - \frac{1}{2}\|\widetilde{w} - \hat{v}_n\|^2 \qquad (23)$$

Again putting in (21) $\alpha = \gamma$ and

$$z = \widetilde{w}, x = \tilde{v}_n, z^* = \tilde{v}_{n+1} f(z) = \phi(\widetilde{w}, \hat{v}_n), f(z^*) = \phi(\tilde{v}_{n+1}, \hat{v}_n)$$

we get

$$\frac{1}{2}\|\tilde{v}_{n+1} - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}_{n+1}, \hat{v}_n) \leq \frac{1}{2}\|\widetilde{w} - \tilde{v}_n\|^2 + \gamma\phi(\widetilde{w}, \hat{v}_n) - \frac{1}{2}\|\widetilde{w} - \tilde{v}_{n+1}\|^2 \qquad (24)$$

Selecting $\widetilde{w} = \tilde{v}_{n+1}$ in (23) and $\widetilde{w} = \hat{v}_n$ in (24) we obtain

$$\frac{1}{2}\|\hat{v}_n - \tilde{v}_n\|^2 + \gamma\phi(\hat{v}_n, \tilde{v}_n) \leq \frac{1}{2}\|\tilde{v}_{n+1} - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}_{n+1}, \tilde{v}_n) - \frac{1}{2}\|\tilde{v}_{n+1} - \hat{v}_n\|^2 \qquad (25)$$

$$\frac{1}{2}\|\tilde{v}_{n+1} - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}_{n+1}, \hat{v}_n) \leq \frac{1}{2}\|\hat{v}_n - \tilde{v}_n\|^2 + \gamma\phi(\hat{v}_n, \hat{v}_n) - \frac{1}{2}\|\hat{v}_n - \tilde{v}_{n+1}\|^2 \qquad (26)$$

Adding (25) with (26) and using (22) for

$$\widetilde{w} + h = \tilde{v}_{n+1}, \widetilde{w} = \hat{v}_n, \tilde{v} + g = \tilde{v}_n, \tilde{v} = \hat{v}_n, h = \tilde{v}_{n+1} - \hat{v}_n, g = \tilde{v}_n - \hat{v}_n$$

we finally conclude

$\|\tilde{v}_{n+1} - \hat{v}_n\|^2 \leq \gamma\phi(\tilde{v}_{n+1}, \tilde{v}_n) - \phi(\hat{v}_n, \tilde{v}_n)] - \gamma\phi(\tilde{v}_{n+1}, \hat{v}_n) - \phi(\hat{v}_n, \hat{v}_n)] \leq \gamma C\|\tilde{v}_{n+1} - \hat{v}_n\|\|\tilde{v}_n - \hat{v}_n\|$ which implies

$$\|\tilde{v}_{n+1} - \hat{v}_n\| \leq \gamma C\|\tilde{v}_n - \hat{v}_n\| \qquad (27)$$

2) Now, taking $\widetilde{w} = \tilde{v}_{n+1}$ in (23) and $\widetilde{w} = \tilde{v}^*$ in (24) we get

$$\frac{1}{2}\|\hat{v}_n - \tilde{v}_n\|^2 + \gamma\phi(\hat{v}_n, \tilde{v}_n) \leq \frac{1}{2}\|\tilde{v}_{n+1} - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}_{n+1}, \tilde{v}_n) - \frac{1}{2}\|\tilde{v}_{n+1} - \hat{v}_n\|^2$$

$$\frac{1}{2}\|\tilde{v}_{n+1} - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}_{n+1}, \hat{v}_n) \leq \frac{1}{2}\|\tilde{v}^* - \tilde{v}_n\|^2 + \gamma\phi(\tilde{v}^*, \hat{v}_n) - \frac{1}{2}\|\tilde{v}^* - \tilde{v}_{n+1}\|^2$$

Adding these two inequalities and multiplying by two yields

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + \|\hat{v}_n - \tilde{v}_n\|^2 - 2\gamma\phi(\tilde{v}^*, \hat{v}_n) + 2\gamma\phi(\tilde{v}_{n+1}, \hat{v}_n) + \phi(\hat{v}_n, \tilde{v}_n) - \phi(\tilde{v}_{n+1}, \tilde{v}_n)] \leq \|\tilde{v}^* - \tilde{v}_n\|^2$$

Adding and subtracting the term $\varphi(\hat{v}_n, \hat{v}_n)$ we have

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + \|\hat{v}_n - \tilde{v}_n\|^2 + 2\gamma[\phi(\hat{v}_n, \hat{v}_n) - \phi(\tilde{v}^*, \hat{v}_n)] + 2\gamma[\phi(\tilde{v}_{n+1}, \hat{v}_n) - \phi(\hat{v}_n, \hat{v}_n) + \phi(\hat{v}_n, \tilde{v}_n) - \phi(\tilde{v}_{n+1}, \tilde{v}_n)] \leq \|\tilde{v}^* - \tilde{v}_n\|^2$$

Using (22) with $\tilde{w} + h = \tilde{v}_{n+1}$, $\tilde{w} = \hat{v}_n$, $\tilde{v} + k = \tilde{v}_n$ and $\tilde{v} = \hat{v}_n$ we have $h = \tilde{v}_{n+1} - \hat{v}_n$ and $k = \tilde{v}_n - \hat{v}_n$, and the inequality above becomes

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + \|\hat{v}_n - \tilde{v}_n\|^2 + 2\gamma[\phi(\hat{v}_n, \hat{v}_n) - \phi(\tilde{v}^*, \hat{v}_n)] - 2\gamma C\|\tilde{v}_{n+1} - \hat{v}_n\|\|\tilde{v}_n - \hat{v}_n\| \leq \|\tilde{v}^* - \tilde{v}_n\|^2.$$

Applying (27) to the last term in the left-hand side and in view of the strict convexity property of $\phi$ given by $\phi(\hat{v}_n, \hat{v}_n) - \phi(\tilde{v}^*, \hat{v}_n) \geq \delta\|\hat{v}_n - \tilde{v}^*\|^2$we get

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + 2\gamma\delta\|\hat{v}_n - \tilde{v}^*\|^2 + (1 - 2\gamma^2C^2)\|\tilde{v}_n - \hat{v}_n\|^2 \leq \|\tilde{v}^* - \tilde{v}_n\|^2$$

Applying the identity $2\langle a - c, c - b\rangle = \|a - b\|^2 - \|a - c\|^2 - \|c - b\|^2$ with $a = \hat{v}_n$, $b = \tilde{v}^*$ and $c = \tilde{v}_n$, to the left-hand side of the last inequality we have

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + (1 - 2\gamma^2C^2)\|\tilde{v}_n - \hat{v}_n\|^2 + 2\gamma\delta[2\langle\hat{v}_n - \tilde{v}_n, \tilde{v}_n - \tilde{v}^*\rangle + \|\tilde{v}_n - \hat{v}_n\|^2 + \|\tilde{v}_n - \tilde{v}^*\|^2$$
$$= \|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + (1 + 2\gamma\delta - 2\gamma^2C^2)\|\tilde{v}_n - \hat{v}_n\|^2 + 4\gamma\delta\langle\hat{v}_n - \tilde{v}_n, \tilde{v}_n - \tilde{v}^*\rangle + 2\gamma\delta\|\tilde{v}_n - \tilde{v}^*\|^2 \leq \|\tilde{v}^* - \tilde{v}_n\|^2$$

Defining $d = 1 + 2\gamma\delta - 2\gamma^2C^2$ and completing the square form of the third and fourth terms yields

$$\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + d\|\tilde{v}_n - \hat{v}_n\|^2 + 4\gamma\delta\langle\hat{v}_n - \tilde{v}_n, \tilde{v}_n - \tilde{v}^*\rangle + \frac{(2\gamma\delta)^2}{d}\|\tilde{v}_n - \tilde{v}^*\|^2 - \frac{(2\gamma\delta)^2}{d}\|\tilde{v}_n - \tilde{v}^*\|^2 + 2\gamma\delta\|\tilde{v}_n - \tilde{v}^*\|^2 \leq \|\tilde{v}^* - \tilde{v}_n\|^2$$

and $c\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 + \|\tilde{v}_{n+1} - \hat{v}_n\|^2 + \left\|\sqrt{d}(\tilde{v}_n - \hat{v}_n) + \frac{2\gamma\delta}{\sqrt{d}}(\tilde{v}_n - \tilde{v}^*)\right\|^2 \leq \left(1 - 2\gamma\delta + \frac{(2\gamma\delta)^2}{d}\right)\|\tilde{v}^* - \tilde{v}_n\|^2$

finally implying $\|\tilde{v}^* - \tilde{v}_{n+1}\|^2 \leq \rho\|\tilde{v}^* - \tilde{v}_n\|^2 \leq \rho^{n+1}\|\tilde{v}^* - \tilde{v}_0\|^2 \underset{n\to\infty}{\to} 0$ with $\rho = 1 - 2\gamma\delta + \frac{(2\gamma\delta)^2}{d} \in (0,1)$. The Theorem is proven.

## REFERENCES

[1] **E.Asiain,J. B. Clempner and A. S. Poznyak(2018),*A Reinforcement Learning Approach for Solving the Mean Variance Customer Portfolio in Partially Observable Models***. *International Journal on Artificial Intelligence Tools*, 27(08), 1850034;

[2] **A. S. Antipin(2005),*An Extraproximal Method for Solving Equilibrium Programming Problems and Games***. *Computational Mathematics and Mathematical Physics*, 45(11):1893–1914;

[3] **H. Bannister, B. Goldys, S. Penev, and W. Wu(2016),*Multiperiod Mean-Standard-Deviation Time Consistent Portfolio Selection***.*Automatica*, 73:15–26;

[4] **U.Cakmak and S. Ozekici(2006),*Portfolio Optimization in Stochastic Markets***. *Mathematical Methods of Operations Research*, 63:151–168;

[5] **U.Cakmak and S. Ozekici(2007),*Multi-period Portfolio Optimization Models in Stochastic Markets Using the Mean-Variance Approach***.*European Journal of Operational Research*, 179:186–202;

[6] **Z. Chen, G. Li, and Y. Zhao(2014),*Time-consistent Investment Policies in Markovian Markets: A Case of Mean-variance Analysis***.*Journal of Economic Dynamics and Control*, 40:293–316;

[7] **J. B. Clempner and A. S. Poznyak(2014),*Simple Computing of the Customer Lifetime Value: A Fixed Local-Optimal Policy Approach***. *Journal of Systems Science and Systems Engineering*, 23(4):439–459;

[8] **J. B. Clempner and A. S. Poznyak(2018),*Sparse Mean-Variance Customer Markowitz Portfolio Selection for Markov Chains: A Tikhonov's Regularization Penalty Approach***. *Optimization and Engineering*. To be published;

[9] **J. B. Clempner and A. S. Poznyak (2018),*A Tikhonov Regularization Parameter Approach for Solving Lagrange Constrained Optimization Problems***.*Engineering Optimization*. To be published;

[10] **J. B. Clempner and A. S. Poznyak(2018),*A Tikhonov Regularized Penalty Function Approach for Solving Polylinear Programming Problems*.***Journal of Computational and Applied Mathematics*,328:267-286;

[11] **O. L. V. Costa and M. V. Araujo(2008),*A Generalized Multi-Periodmean-Variance Portfolio Optimization with Markov Switching Parameters***. *Automatica*,44(10):2487–2497;

[12] **W. J. Guo and Q. Y. Hu(2005),*Multi-period Portfolio Optimization when Exit Time ss Uncertain***. *Journal of Management Science and Engineering*, 8:14–19;

[13] **B. Li and S. C. H. Hoi(2014),*Online Portfolio Selection: A Survey***.*ACM Computing Surveys*, 46(3):1–35;

[14] **D. Li and W. L. Ng(2000),*Optimal Dynamic Portfolio Selection: Multiperiodmean-Variance Formulation***. *Mathematical Finance*, 10:387–406;

[15] **H. Markowitz(1952),Portfolio Selection. *Journal of Finance***, 7:77–91;

[16] **R. C. Merton (1972),*An Analytical Derivation of the Efficient Portfolio Frontier*.** *Journal of Financial and Quantitative Analysis*, 7(4):1851–1872;

[17] **A. S. Poznyak, K. Najim and E. Gomez-Ramirez(2000),*Self-learning Control of Finite Markov Chains***. Marcel Dekker, New York;

[18] **E. M. Sánchez, J. B. Clempner and A. S. Poznyak(2015),*A Priori-Knowledge/Actor-Critic Reinforcement Learning Architecture for Computing the Mean'variance Customer Portfolio: The Case of Bank Marketing Campaigns*.** *Engineering Applications of Artificial Intelligence*, 46, Part A:82–92;
[19] **E. M. Sánchez, J. B. Clempner and A. S. Poznyak(2015),*Solving the Mean-Variance Customer Portfolio in Markov Chains Using Iterated Quadratic/Lagrange Programming: A Credit-Card Customer-Credit Limits Approach.*** *Expert Systems with Applications*, 42(12):5315–5327;
[20] **M. C. Steinbach(2001),*Markowitz Revisited: Mean-variance Models in Financial Portfolio Analysis*.** *Journal of the Society for Industrial and AppliedMathematics*43:31–85;
[21] **S. Z. Wei and Z. X. Ye(2007), *Multi-Period Optimization Portfolio with Bankruptcy Control in Stochastic Market*.** *Applied Mathematics and Computation*, 186:414–425;
[22] **H. Wu and Z. Li(2011),*Multi-Period Mean-Variance Portfolio Selection with Markov Regime Switching and Uncertain Time-Horizon*.** *Journal of Systems Science and Complexity.*, 24:140–155;
[23] **H. Wu and Z. Li(2012),*Multi-period Mean-Variance Portfolio Selection with Regime Switching and a Stochastic Cash Flow*.***Insurance: Mathematics and Economics*, 50(3):371–384;
[24] **H. Yao, Y. Lai and Z. Hao (2013), *Uncertain Exit Time Multi-Period Mean-Variance Portfolio Selection with Endogenous Liabilities and Markov Jumps*.** *Automatica*, 49:3258–3269;
[25] **S. S. Zhu, D. Li and S. Y. Wang,(2004),*Risk Control over Bankruptcy in Dynamic Portfolio Selection: A Generalized Mean-Variance Formulation*.** *IEEE Transactions on Automatic Control*, 49:447–457.