מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Thesis for the degree
Master of Science

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

# Combinatorial Optimization Problems with Testing

*Chen Attias*

supervised by
Prof. Robert Krauthgamer
Prof. Retsef Levi (MIT)

June 9, 2016

**Abstract**

In stochastic versions of combinatorial optimization problems, the goal is to optimize an objective function under constraints that specify the feasible solutions, when some parameters involve uncertainty. We focus on maximizing a linear objective $\sum_{i=1}^{N} W_i x_i$, where $x_1, \ldots, x_N$ are the decision variables, $W_1, \ldots, W_N$ are mutually independent random coefficients, and the constraints are deterministic. This formulation captures many real-world problems and has been extensively studied in Operations Research. In most prior work, the uncertainty in the objective is modeled by unknown distributions and various mechanisms are used to acquire information on these distributions. In contrast, we assume that the distributions are known and use a mechanism to acquire information on realizations.

We consider problems where a decision-maker can test any desired coefficient $W_i$, which reveals its realization $w_i$ and incurs a fixed cost $c > 0$. Testing a coefficient reduces uncertainty, which can only improve the optimization, but it may or may not compensate for the testing cost. The decision-maker can test coefficients in a sequential and possibly adaptive manner, and then has to return a feasible solution based on the current information. A policy is a set of rules that determines the sequential decisions of the decision-maker in every possible situation. The goal is to find a policy that maximizes the expected profit, where profit is defined as the returned solution's objective value minus all the testing costs.

A policy that obtains optimal expected profit can usually be computed by dynamic programming. However, if the dimension of the problem is high, this method becomes intractable. As an alternative, we study policies that are myopic – a myopic policy decides whether to test coefficients based on the marginal profit from exactly one test, without considering possible future effects, i.e., it makes decisions based on a limited horizon of one test. Compared to dynamic programming, myopic policies are simpler and easier to compute, but more restrictive and thus possibly suboptimal.

We show that myopic policies can actually achieve optimal expected profit for a number of interesting problems:

1. Selection with testing, where the feasible solutions represent selecting exactly one of the $N$ coefficients and each decision variable $x_i$ indicates whether coefficient $W_i$ is selected;
2. Maximum spanning tree with testing, where the feasible solutions represent spanning trees in a given graph $G$ and each $x_i$ indicates whether the corresponding edge is included in the tree;
3. Linear optimization over a polymatroid with testing, where the feasible solutions are defined by a given submodular function.

For more general optimization problems with testing, we derive weaker results in the form of sufficient conditions under which a myopic rule for deciding whether to stop testing is optimal.

## Acknowledgements

First and foremost, I would like to wholeheartedly thank my wonderful advisors Professor Robert Krauthgamer and Professor Retsef Levi. I'm deeply grateful for the opportunity to work with them and benefit from their wise advice and guidance, as well as their support and patience. Selecting a thesis advisor, which is one of the most important decisions that a student has to make, involves overwhelming uncertainty, and yet somehow I've made the optimal decision.

I would like to express my gratitude to Professor Itai Benjamini for his support and counsel during the past two years.

In addition, I thank Yaron Shaposhnik for being an extraordinary brainstorming partner and friend.

I also thank the people of the department of Mathematics and Computer Science in the Weizmann Institute, for creating an amazing and supportive environment, especially Otniel Van Hendel, Anne Kenyon, Liran Szlak, Itay Glazer, Jonathan Rosenski, Anna Shtengel, Itay Safran and many others.

Finally, I dedicate this dissertation to my family – my mother Shula, my late father Ami, my brother Orr, and my partner Omri, who are my continuous source of inspiration.

# Contents

# 1 Introduction

Many real-world problems that involve uncertainty can be formulated in the framework of *sequential stochastic combinatorial optimization*, defined as follows. *Combinatorial optimization* is the task of finding an optimal solution in a finite collection of feasible solutions, i.e., a solution that maximizes (or minimizes) an objective function under specified constraints. For example, a solution could be a subset of edges or vertices in a graph problem, or an ordering in a scheduling problem, or a subset of items in a knapsack problem, and so forth. In the *stochastic* framework, the revenue (or cost) of a solution is modeled using random variables with known distributions, which captures an inherent uncertainty when selecting any solution. Throughout, we only consider cases where the randomness is restricted to the objective function and does not affect the constraints, i.e., the set of feasible solutions is deterministic. *Sequential* (or multi-stage) optimization models scenarios where information is revealed in stages, and decisions are required at each stage. In this case, a solution is described by a *policy*, which is a set of rules that determines a decision at each stage, given the currently available information. The goal is to find a policy that produces a feasible solution for all the possible data instances and maximizes the total expected revenue (or minimizes the total expected cost) of the selected solution.

In this thesis, we study stochastic combinatorial optimization problems in a setting where any random variable can be *tested* (in the sense of observing its specific realization) prior to returning a feasible solution. A test models acquiring information about a random variable – it reveals its realization, which reduces uncertainty and can only improve the optimization, but incurs a cost, and thus might have a negative effect overall. The goal is to find a policy that maximizes the expected *profit*, where profit is defined as the returned solution's objective value minus all the testing costs.

Levi et al. (2014) introduced the testing framework in the context of scheduling problems. The definition of profit balances between the costs and benefits of testing and thus captures an inherent tradeoff between spending resources to reduce uncertainty and perform the optimization. Our work follows their model of reducing uncertainty by testing random variables, but addresses a different and quite general set of optimization problems, as described next. We discuss comparisons with previous work in Section 1.2.

As an example, consider a manager who needs to select (exactly) one out of several competing projects. It will be useful later to view these projects as parallel edges between two vertices, where the edge weights represent project revenues. Different levels of uncertainty regarding the revenue of each project are modeled as different probability distributions, possibly estimated based on past experience or preliminary examination. To reduce this uncertainty, the manager can "test" any project, for example by performing market research to obtain the realization of its revenue, which incurs a cost. The goal of the manager is to maximize the expected profit, where profit in this example is the revenue of the selected project minus all the testing costs. If the testing cost is high, then the manager may not want to test many projects (and if it is extremely high, she may not wish to test any). So the first decision the manager must make is whether to perform any test at all, or to select a project based on the current information. If she decides that testing is worthwhile, the manager must also choose which project to test. Once the realization of the tested project is observed, she must make another decision based on her updated information – whether to test another project, and if so which one, and so on and so forth. These decisions define an adaptive policy for the manager, and our goal is to find such a policy with optimal expected profit. This example describes the *selection with testing* problem, which is the subject of Section 2.

As a second example, consider a contractor laying out a new telecommunication network between sites, modeled as a graph whose vertices represent sites and whose edges represent potential cable routes. Each edge has a weight that represents the cost of laying the corresponding cable. This cost may be affected by many factors, some of which are a-priori known (such as its length) and some not known (such as physical conditions along the route). To reduce the uncertainty, the contractor can test a certain route, for example by digging up small parts of it to examine the terrain. This test would reveal the installation cost, but incur a cost of its own. The contractor's goal is to minimize the expectation of the costs of installing the selected cables plus the testing costs, under the constraint that all the sites must be connected. In essence, the goal is to find a spanning tree in the network that minimizes the expectation of the combined costs of the edge weights and testing costs. It is a well known fact that finding a minimum spanning tree can be done by negating all edge weights and finding a maximum spanning tree, which is the variant we study in this work and refer to as *maximum spanning tree with testing*, or *MST with testing* for short. This problem generalizes selection with testing, because selection can be viewed as the MST problem on a graph with two vertices and parallel edges between them.

Even more generally, consider (deterministic) linear programs whose feasible region has the structure of a *polymatroid*, defined as follows. For a given ground set $N$ and a non-decreasing submodular set function $f : 2^N \to \mathbb{R}$, a polymatroid is a polyhedron in $\mathbb{R}^N$ consisting of all vectors $x$ with non-negative coordinates that satisfy $\sum_{s \in S} x(s) \leq f(S)$ for every subset $S \subseteq N$ (see Chapter 44 in Schrijver (2003) for more details). Although a polymatroid is defined by an exponential number of linear inequalities, Edmonds (1970) showed that optimizing a linear objective over a feasible region which is a polymatroid can be solved efficiently by a greedy algorithm. Another result by Edmonds (1971) shows that the maximum spanning tree problem can be formulated as a linear optimization problem where the feasible region is the forest polymatroid (see also Section 50.4 in Schrijver (2003)). In addition to selection and MST, many other problems fall under the polymatroid framework, including for example $K$-selection, the problem of selecting $K$ out of $N$ alternatives so as to maximize the total revenue. In *linear optimization over a polymatroid with testing* (LOPT for short), each *coefficient* $W_i$ in the linear objective $\sum_i W_i x_i$ is a random variable (for example, in MST with testing, the coefficients are random edge weights). We study the case of LOPTs in Section 3, and the special case of MST with testing in Section 3.5. We also discuss more general optimization problems with testing in Section 4.

**Stopping Rule and Testing-Order Rule.** In the aforementioned problems, the objective's coefficients (such as edge weights) are modeled as independent random variables with known distributions. At any stage in the execution of a policy, it can either *stop*, i.e., optimize based on the current information, or *test a coefficient*, i.e., observe its realization, which incurs a cost. Therefore, a policy is characterized by two rules that determine the decisions it makes at each stage. The first rule, referred to as a *stopping rule*, determines whether to stop or test. If the policy decides to test, then the second rule, referred to as a *testing-order rule*, determines which coefficient to test.

The most closely related literature is that of optimal stopping problems (see for example Ferguson (2012)), which concern choosing a time to take a particular action based on sequentially observed random variables in order to maximize an expected reward (or minimize an expected cost). A specific optimal stopping problem that is closely related to selection with testing is the famous house-selling problem, introduced by MacQueen and Miller Jr (1960). These problems are similar to our setting in the sense that both aim to find an optimal stopping rule. However, the difference is that in our setting a policy also consists of a testing-order rule, i.e., in what order to test the random variables. The number of possible testing orders is the number of permutations of the coefficients, i.e., factorial in the number of coefficients. In fact, the testing order may be determined adaptively as new information is gathered, and thus the testing order may need to be recomputed after each test. This implies that the testing framework is more general than the optimal stopping framework and has higher complexity. We discuss other related work in Section 1.2.

**Myopic Decisions.** Sequential decision problems such as those described above can generally be solved by dynamic programming, as discussed in Bellman (1954). However, if the dimension of the problem is high, then this method becomes intractable. A different approach used in such problems is to consider *myopic policies*, which choose between alternatives based on the marginal profit from only one action (test), without considering the future effects of their decision. In our framework, myopic policies choose the better of two alternatives: (1) to stop testing, which yields an expected profit based on the information gathered so far; (2) to test one more coefficient and then to stop testing, which yields an expected profit that also takes the new information and additional testing cost into account. Notice that a myopic policy determines both a stopping rule and a testing-order rule, because if it decides to test, it also determines which coefficient to test. The best myopic policy can be described alternatively as making an optimal decision when having a budget for at most one more test. Compared to dynamic programming, myopic policies are simpler and easier to compute, but more restrictive and thus potentially suboptimal. However, we show that for a number of interesting problems, myopic policies can achieve the same expected profit as dynamic programming, and can therefore achieve optimal expected profit.

## 1.1 Our Results[1]

We organize our results by problem complexity, ordered from the simplest to the most involved. Throughout, any random coefficient in the objective can be tested, which reveals its realization and incurs a fixed testing cost $c > 0$, which is the same for all coefficients and is known.

**Selection with Testing.** Selection with testing is the problem of selecting a candidate from $N$ alternatives (or parallel edges), whose revenues (weights) $W_i$ are mutually independent random variables, drawn from known distributions with the same finite mean (notice that the revenues need not be identically distributed). We provide for this problem a myopic policy that obtains optimal expected profit. Furthermore, the policy is threshold-based, i.e., it computes for each edge an a-priori threshold that depends only on its weight distribution and the testing cost (and therefore can be precomputed), and decides whether to stop testing by comparing the maximal realization revealed so far to this threshold. Moreover, the testing order of the edges is based on these thresholds, and is thus non-adaptive and can be precomputed as well.

**Maximum Spanning Tree with Testing.** Let $G = (V, E)$ be a connected graph with vertices $V = \{v_1, ..., v_n\}$ and edges $E = \{e_1, ..., e_m\}$. We prove that if the edge weights, denoted $W_i$, are drawn i.i.d. from a known distribution with a finite mean, then there exists a myopic policy for constructing a spanning tree that obtains optimal expected profit. In contrast to selection with testing, here the graph structure comes into play in the sense that information gathered on a certain edge can affect the desirability of testing other edges. In particular, the testing-order rule is adaptive.

[1]Some of this work was carried out jointly with Yaron Shaposhnik from MIT. In particular, the results in Sections 3 and 4 will appear also in his PhD thesis.

**Additional Results.** We also discuss more general stochastic optimization problems with testing. For linear optimization over a polymatroid with testing (as mentioned above, a polymatroid is a certain type of polytope defined by a submodular function), we show that under certain technical conditions (such as symmetry of the feasible region and a convex order between the random variables), there exist myopic policies that obtain optimal expected profit. This captures, for example, $K$-selection with testing, which is the problem of selecting $K$ out of $N$ alternatives (parallel edges) so as to maximize the expected profit.

For even more general optimization problems with testing, we derive weaker results in the form of sufficient conditions under which a myopic stopping rule is optimal.

For sake of simplicity, throughout we address only maximization problems, however similar results can be proved for minimization problems.

## 1.2 Related Work

In general, testing problems are stochastic multi-period optimization problems. Traditionally however, in these problems information gathering is an exogenous process. For example, the classical work of Scarf (1959) studies a finite horizon inventory model where at the beginning of every period, an order is placed and the demand which is a-priori random is realized. In this model decisions about inventory replenishment do not affect information collection. In other work on stochastic optimization, decisions could impact the belief state. For example, Dean et al. (2004) considered a stochastic knapsack problem where items have random sizes that are realized only when attempting to place these items in the knapsack. Thus, the item we choose affects what we know. Chen et al. (2009) studied a stochastic matching problem where edges need to be probed prior to their selection. Once an edge is probed and found fit, the edge is selected. In these two examples, item selection and information collection decisions are not disjoint, because once an item or an edge is probed, it is irrevocably selected. In contrast, in our model, information collection is separate from optimization, as the decision-maker first adaptively collects information, and only then selects a feasible solution. See also Gupta and Nagarajan (2013) and Adamczyk et al. (2013) for additional examples of stochastic optimization problems, in which decisions affect information collection.

Multiple areas of research are concerned with sequential information collection. Perhaps the most notable one is Ranking and Selection where information is gathered about a set of alternatives with the goal of choosing the best one, see, e.g., Swisher et al. (2003) for a survey. Different objectives were studied for this problem, including maximizing the probability of selecting the best alternative (e.g., the secretary problem), or minimizing the total cost of selecting the best alternative. The closely related multi-armed bandit problems (see, e.g., Gittins et al. (2011) and Bubeck and Cesa-Bianchi (2012)) is a framework for balancing rewards obtained from exploring different alternatives in the selection process, and committing to an alternative. Two known problems are especially close to the selection problem with testing that we study in Section 2. The first one is the famous house-selling problem, in which a seller meets buyers sequentially. Each potential buyer offers to buy the house at some price, and the seller needs to decide whether to sell the house at this price, or to continue meeting potential buyers. Meeting a customer is associated with some fixed cost, and the seller must therefore balance the meetings costs and improving the current offer. This is a stopping time problem with infinite number of buyers, where the offers are independent and identically distributed (i.i.d.). In contrast, in our setting there is a finite number of buyers and their offers are drawn from different distributions. Unlike the house-selling problem, in our setting the decision-maker needs to decide not only whether to accept or reject an offer, but also to choose the customer to meet next, which is important when customers are non-homogeneous. Guha et al. (2007) studied a problem that can be viewed as a selection with testing problem, where each alternative is distributed according to the Bernoulli distribution, and could incur different testing costs.

The area of Optimal Learning focuses on optimization problems with uncertain parameters that are described stochastically, and that can be learned. For example, Ryzhov and Powell (2012) studied a Linear Program with random cost coefficients, where there is a prior knowledge about the joint distribution of cost coefficients, and a budget for drawing samples from this distribution. The goal is to devise a learning strategy that maximizes the outcome of the optimization problem solved after the learning phase is completed. A similar setting in the context of a Ranking and Selection problem with correlated alternatives has been studied by Frazier et al. (2009). In general, Optimal Learning takes a Bayesian approach in which learning refers to sampling and updating the belief about unknown distributions. In contrast, we assume that the distributions are known and one can reduce uncertainty by observing the realization of random variables. This requires different methods and leads to results of a different nature. In particular, we prove optimality of certain myopic policies rather than asymptotic analysis or worst-case guarantees. Similarly, Ranking and Selection problems and Multi-armed bandit problems are often studied in Bayesian or adversarial settings. See Powell and Ryzhov (2012) for a comprehensive overview of recent work in the area of Optimal Learning.

Recently, Golovin and Krause (2011) introduced the concept of Adaptive Submodularity which generalizes the definition of submodularity. They showed that stochastic optimization problems that satisfy a few properties can be solved near-optimally by a simple greedy algorithm. In particular, various data collection problems satisfy these properties. For example, Javdani et al. (2014) showed that the problem of deciding adaptively on medical tests before deciding on an action is adaptively submodular. Two main differences between these approaches and ours are that, first, while the

Adaptive Submodularity framework provides general sufficient conditions under which a greedy algorithm is near-optimal, our work establishes the optimality of certain policies. Second, we do not focus on finding the best alternative (such as determining the exact medical condition), but rather on applications where the costs and benefits of testing must be balanced. See Section 13 in Golovin and Krause (2011) for a discussion about related problems and applications.

A related line of work deals with optimization problems with interval uncertainties. In these problems the value of some parameters is only known to lie within certain intervals. The goal in such problem is to decide on the set of parameters to be tested (also called queried or probed) that would guarantee finding an optimal or near-optimal values. Feder et al. (2000) studied the problem of minimizing the total cost of selecting the $k$-th smallest among $n$ uncertain values with interval uncertainties that guarantees a solution that is within a certain range. Feder et al. (2007) considered the problem of finding an optimal shortest path in a graph with uncertainty about its exact edge weights. Erlebach et al. (2008) and Megow et al. (2015) then studied a similar setting where the objective is to find a minimum spanning tree. In all of these problems, uncertainties are adversarial in nature and algorithms that guarantee worse-case bounds are developed (for adaptive and non-adaptive variants of these problems). In contrast, we assume a stochastic model of uncertainty and seek to optimize the expected values of some objectives. For further work in this area see Khanna and Tan (2001), Gupta et al. (2011) and Goerigk et al. (2015).

Multiple researchers explored probabilistic testing problems. Goel et al. (2006) studied non-adaptive testing problems where one needs to decide in advance about all future tests. Once testing decisions are made, the true realizations of the tested parameters are observed, and an optimization is performed based on the observed parameters. They assume that there is a finite budget and every coefficient is associated with a known testing cost. They show that in general these problems are NP-hard, and develop near-optimal algorithms for variants of the knapsack problem. Guha and Munagala (2007) studied the adaptive version of the above problem, and showed that for many problems there is a non-adaptive policy that achieves a constant factor approximation for the best adaptive problem (see also Guha and Munagala (2008) and Goel et al. (2010)). Whereas the above work considered a budget for testing, we assume that there is a testing cost that can be compared against the optimization value, and the goal is to maximize the optimization value minus the testing costs. For multiple problems we find not an approximately optimal, but rather an optimal policy.

The closest to our model and indeed the motivation for our work is a model studied by Sun et al. (2014) and more generally by Levi et al. (2014). The latter considered the problem of scheduling jobs with random processing times and weights by a single server with the objective of minimizing the expected weighted sum of completion times. They assumed that jobs can be tested, an activity that utilizes the server but reveals the exact processing times and weights of the tested job. They showed that the optimal policy admits the structure of an optimal stopping problem, and that under a certain condition, a myopic rule is optimal in determining when to stop testing. A significant difference from our work is that they consider an online setting in which testing and processing decisions are made interchangeably, that is, some jobs can be tested, other jobs can then be processed and so forth. In contrast, we study an offline setting, where we begin by testing to reduce uncertainty, and only then optimize. Moreover, we consider a much broader class of problems, for which the offline version of the above scheduling problem is a special case. We show that some of the structural properties of the optimal policy in the scheduling problems carry through to our generalized setting, including in particular, the main result, which is the optimality of a certain myopic policy. In addition, our analysis uncovers and explains a mathematical structure that results in the optimality of myopic policies.

# 2 Selection with Testing

In this section we study *selection with testing*, which is a basic example of combinatorial optimization problems with testing, and serves as a motivating example for the following sections. A straightforward dynamic programming formulation of the optimal policy has a large state space and therefore solving it may be intractable. In contrast, we design a simple, threshold-based, and easy to compute policy for the problem, and prove in Theorem 2.13 that if all edge weights have the same finite mean, then it obtains optimal expected profit.

## 2.1 Problem Formulation

Selection with testing is the problem of selecting a single candidate from $N$ candidates, which we model as $N$ parallel edges indexed by $1, \ldots, N$, each associated with a random weight $W_i$ drawn from a known distribution. We assume that all $W_i$ have the same finite mean, i.e., $\mathbb{E}[W_i] = \mu < \infty$ and that $W_1, \ldots, W_N$ are mutually independent. With slight abuse of notation, we refer to $W_i$ as the random weight of edge $i$, as well as its weight distribution and as the name of the respective edge.

In this model, a policy can test an edge at a cost $c > 0$ and obtain its exact realization. The objective is to maximize the expected *profit*, where profit is defined as the weight of the selected edge minus all testing costs (if any). At any moment during the execution, we denote by $w_{\max}$ the maximal weight among all tested edges (by convention, when there are no tested edges $w_{\max} = -\infty$), and by $\bar{W}$ the set of yet untested edges, hence the current state can be summarized as $\left(\bar{W}, w_{\max}\right)$. At any state, a policy can choose to either stop, i.e., optimize based on the currently available information (that is, with respect to $w_{\max}$), or to test some yet-untested edge with random weight $W_i$, which reveals its weight realization $w_i$ and changes the state to $(\bar{W} \setminus \{W_i\}, \max\{w_{\max}, w_i\})$. A policy should then choose between optimizing or testing another edge, and so forth. It is straightforward to verify that upon optimizing, the profit is maximized by choosing an edge with maximum expected weight.

## 2.2 An Optimal Policy via Dynamic Programming

First consider a variant of the selection with testing problem in which a policy must fix a testing order in advance, i.e., the order in which edges are tested is decided before the adaptive decision of when to stop. In this case, there are $N!$ different testing orders to consider, which already describe a huge state space. Adding the option to dynamically decide which edge to test next leads to an even more elaborate computation, because all possible permutations of yet-untested edges must be reconsidered after each test. A less naive approach is to exploit the fact that we are looking for an edge with maximal expected weight, therefore all tested edges with weights lower than $w_{\max}$ can be discarded. Hence, for each state $\left(\bar{W}, w_{max}\right)$, the following dynamic program computes an optimal policy based on backward induction

$$J^{\text{opt}}\left(\bar{W}, w_{\max}\right) = \max \begin{cases} \max\left\{w_{\max}, \mu\right\} & // \text{ Optimize} \\ \max_{W_i \in \bar{W}}\left\{-c + \mathbb{E}\left[J^{\text{opt}}\left(\bar{W} \setminus \left\{W_i\right\}, \max\left\{w_{\max}, W_i\right\}\right)\right]\right\} & // \text{ Test an edge.} \end{cases} \tag{1}$$

**Observation 2.1.** *The policy computed by the Dynamic Program (1) achieves optimal expected profit.*

Since we need to consider all subsets $\bar{W} \subseteq \{W_1, \ldots, W_N\}$, the state space of the aforementioned DP is at least $2^N$ (even without considering all possible $w_{\max}$), thus selection with testing seems hard to solve. However, under the equal means assumption, we design a simple, easy to compute policy that obtains optimal expected profit in Theorem 2.13. We begin by addressing the base case, of only one yet-untested edge.

## 2.3 A Simple Optimal Policy for $N = 1$ Untested Edges

Consider the case of one untested edge with random weight $W_1$ and arbitrary $w_{\max}$ (which can model a state that is reached after some tests were made). There are merely two types of possible strategies. The first one is to *stop*, i.e., return (without further testing) an edge with maximal expected revenue, either a maximal tested edge or the single untested edge, and achieve an expected profit $\max\{w_{\max}, \mu\}$. The second strategy is to *test* the untested edge $W_1$ and then return the edge of largest weight out of all the tested edges, and achieve an expected profit $-c+\mathbb{E}\left[\max\left\{w_{\max}, W_1\right\}\right]$. An optimal policy can choose the strategy that has higher expected profit among the two, i.e., stops if and only if

$$\max\left\{w_{\max}, \mu\right\} \geq -c + \mathbb{E}\left[\max\left\{w_{\max}, W_1\right\}\right].$$

We examine how this decision whether to test depends on $w_{\max}$. Intuitively, if $w_{\max}$ has an extremely high value, then there is a low probability that an even higher value will be revealed by testing $W_1$, so it is better to select $w_{\max}$ and stop. Similarly, if $w_{\max}$ has an extremely low value, then it is better to select $W_1$ without testing it, because there is a low probability that an even lower value will be revealed by testing $W_1$. However, if $w_{\max}$ is in some intermediate range, then testing $W_1$, i.e., "buying" its observation, can be worthwhile, because there is a high uncertainty regarding which of $w_{\max}$ and $w_1$ is higher. Overall, testing is worthwhile when $w_{\max}$ is in a certain interval around $\mu$, which we determine by computing a lower testing threshold and an upper testing threshold in Definition 2.2.

Of course, if the cost of testing is extremely high, then selecting $\max\{w_{\max}, \mu\}$ without testing is the best option regardless of $w_{\max}$. Thus, the thresholds become irrelevant in this case. These intuitions are formalized in Lemma 2.5 and visualized in Figure 2.1.

**Definition 2.2.** *(Testing thresholds).*

*The* upper testing threshold *of a random variable* $W_1$ *is the unique* $\theta_1^+ \in \mathbb{R}$ *satisfying*

$$c = \mathbb{E}\left[\left(W_1 - \theta_1^+\right)^+\right]. \tag{2}$$

*The* lower testing threshold *of a random variable* $W_1$ *is the unique* $\theta_1^- \in \mathbb{R}$ *satisfying*

$$c = \mathbb{E}\left[\left(\theta_1^- - W_1\right)^+\right]. \tag{3}$$

We prove that the thresholds $\theta_1^+, \theta_1^-$ exist and that they are uniquely determined in Section 2.5.

**Lemma 2.3.** *There are only two possibilities for the order between the thresholds* $\theta_1^+, \theta_1^-$ *and* $\mu = \mathbb{E}[W_1]$*, specifically, if* $\mathbb{E}\left[(W_1-\mu)^+\right] \geq c$ *then* $\theta_1^- \leq \mu \leq \theta_1^+$*, otherwise* $\theta_1^+ < \mu < \theta_1^-$*.*

*Proof.* Observe that $(z)^+ - (-z)^+ = z$ for every $z \in \mathbb{R}$. It follows that for every $w_1 \in \mathbb{R}$,

$$(w_1 - \mu)^+ - (\mu - w_1)^+ = w_1 - \mu.$$

And thus by linearity of expectation,

$$\mathbb{E}\left[(W_1-\mu)^+\right] - \mathbb{E}\left[(\mu - W_1)^+\right] = \mathbb{E}\left[W_1 - \mu\right] = 0.$$

Define $\tilde{c}$ by

$$\tilde{c} = \mathbb{E}\left[(W_1-\mu)^+\right] = \mathbb{E}\left[(\mu - W_1)^+\right] \geq 0. \tag{4}$$

By Lemma 2.14 (see Section 2.5), the functions $b \to \mathbb{E}\left[(W_1 - b)^+\right]$ and $b \to \mathbb{E}\left[(b - W_1)^+\right]$ are strictly monotone in $b$ when they take positive values. Hence, comparing equation (4) to equations (2),(3) yields that in the case $0 < c \leq \tilde{c}$ we have

$$\theta_1^- \leq \mu \leq \theta_1^+,$$

and in the case $c > \tilde{c} \geq 0$ we have

$$\theta_1^+ < \mu < \theta_1^-.$$

□

**Fact 2.4.** *For every random variable* $X$ *and a constant* $b \in \mathbb{R}$*,*

$$\max\{b, X\} - b = \max\{0, X - b\} = (X - b)^+, \tag{5}$$

$$\max\{b, X\} - X = \max\{b - X, 0\} = (b - X)^+. \tag{6}$$

The next lemma identifies an optimal action for each value of $w_{\max}$, given additional parameters of the problem (such as $c$, $\mu$ and the distribution of $W_1$). Altogether, we obtain an optimal policy whose decision is almost completely determined by the maximum among $\{w_{\max}, \mu, \theta_1^+\}$.

**Lemma 2.5.** *For a state* $(W_1, w_{\max})$ *with one untested edge* $\{W_1\}$*,*

1. *If* $w_{\max} \geq \max\left\{\mu, \theta_1^+\right\}$ *then stopping and returning* $w_{\max}$ *(without testing any more edges) has optimal expected profit;*

2. *Else, if either*

   (a) $\mu > \max\left\{\theta_1^+, w_{\max}\right\}$*, or*

   (b) $\theta_1^+ > \max\{w_{\max}, \mu\}$ *and* $w_{\max} \leq \theta_1^-$*,*

   *then stopping and returning* $W_1$ *(without testing it) has optimal expected profit;*

3. *Otherwise, testing* $W_1$ *and returning an edge with weight* $\max\{w_1, w_{\max}\}$ *has optimal expected profit.*

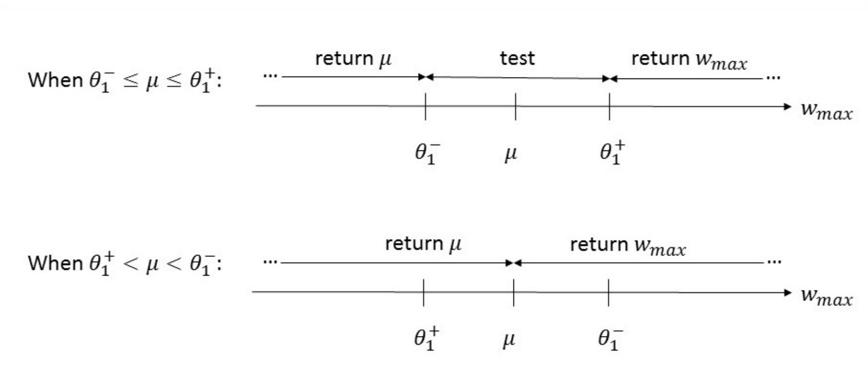


Figure 2.1: An optimal policy for one untested edge as a function of $w_{\max}$, as described in Lemma 2.5. The top figure depicts the case $\theta_1^- \leq \mu \leq \theta_1^+$ and the bottom figure depicts the other case $\theta_1^+ < \mu < \theta_1^-$ (see Lemma 2.3).

*Proof.* As mentioned in the beginning of Section 2.3, with only one untested edge, there are merely two types of possible strategies. The first one is to *stop*, i.e., return (without further testing) an edge with maximal expected revenue, either a maximal tested edge or the single untested edge, and achieve an expected profit $\max\{w_{\max}, \mu\}$. The second strategy is to *test* $W_1$ and then return the edge of largest weight among the tested edges, and achieve an expected profit $-c +$

$\mathbb{E}[\max\{w_{\max}, W_1\}]$. It is clearly an optimal policy to choose the strategy that has higher expected profit, i.e., stop if and only if

$$\max\{w_{\max}, \mu\} \geq -c + \mathbb{E}[\max\{w_{\max}, W_1\}].$$

Consider first the case $w_{\max} \geq \mu$. In this case, stopping and returning $w_{\max}$ is optimal if and only if $w_{\max} \geq -c + \mathbb{E}[\max\{w_{\max}, W_1\}]$, otherwise testing is optimal. The last inequality can be written as

$$c \geq \mathbb{E}[\max\{w_{\max}, W_1\} - w_{\max}],$$

or alternatively, by Definition 2.2 of the thresholds and by Fact 2.4, as

$$\mathbb{E}\left[\left(W_1 - \theta_1^+\right)^+\right] \geq \mathbb{E}\left[(W_1 - w_{\max})^+\right],$$

which in turn holds iff $w_{\max} \geq \theta_1^+$, by the monotonicity of $b \to \mathbb{E}\left[(W_1 - b)^+\right]$ (see Lemma 2.14 in Section 2.5). Therefore, when $w_{\max} \geq \max\left\{\mu, \theta_1^+\right\}$, stopping and returning $w_{\max}$ is optimal, which proves case 1. Otherwise, when $\theta_1^+ > w_{\max} \geq \mu$, testing is optimal, which is used to prove case 3.

Now consider the other case $w_{\max} < \mu$. In this case stopping and returning $W_1$ (without testing it) is optimal if $\mu \geq -c + \mathbb{E}[\max\{w_{\max}, W_1\}]$, otherwise testing is optimal. The last inequality can be written as

$$c \geq \mathbb{E}[\max\{w_{\max}, W_1\} - W_1],$$

or alternatively, by Definition 2.2 and Fact 2.4, as

$$\mathbb{E}\left[\left(\theta_1^- - W_1\right)^+\right] \geq \mathbb{E}\left[(w_{\max} - W_1)^+\right],$$

which holds iff $w_{\max} \leq \theta_1^-$, by the monotonicity of $b \to \mathbb{E}\left[(b - W_1)^+\right]$ (see Lemma 2.14 in Section 2.5). Therefore, when $w_{\max} < \mu$ and $w_{\max} \leq \theta_1^-$, stopping and returning $W_1$ without testing is optimal, and when $\theta_1^- < w_{\max} < \mu$, testing is optimal. By Lemma 2.3, either $\theta_1^+ < \mu < \theta_1^-$ holds, or $\theta_1^- \leq \mu \leq \theta_1^+$ holds. In the first case, $\theta_1^+ < \mu < \theta_1^-$, the testing region is empty, so altogether it is optimal to return $W_1$ without testing it if $\mu > \max\{\theta_1^+, w_{\max}\}$, which proves case 2(a). In the second case, $\theta_1^- \leq \mu \leq \theta_1^+$, the testing region is $\theta_1^- < w_{\max} < \mu$, and so altogether it is optimal to return $W_1$ without testing it when $\theta_1^+ > \max\{w_{\max}, \mu\}$ and $w_{\max} \leq \theta_1^-$, which proves case 2(b). Since there are only two types of alternatives (stop or test), in the remaining cases it is optimal to test, which proves case 3 and the lemma. The different cases are visualized in Figure 2.1. □

## 2.4 A Simple Optimal Policy for $N$ Untested Edges with Equal Means

Now consider having $N > 1$ untested edges $W_1, \ldots, W_N$ of equal finite mean $\mathbb{E}[W_i] = \mu < \infty$. The thresholds $\theta_i^-, \theta_i^+$ for each edge $W_i$ are defined below similarly to Definition 2.2, by considering $W_i$

as a single untested edge and comparing between a policy that stops and a policy that myopically tests $W_i$, i.e., tests only edge $W_i$ and stops. Notice that although the decision whether to stop depends on $w_{\max}$, the thresholds themselves only depend on $c$ and the distribution of $W_i$.

Clearly, if some edge $W_i$ is "worth testing" myopically, then stopping is suboptimal. The other direction is true as well, but is not as trivial. We will prove in Lemma 2.11 in Section 2.4.2 that if none of the edges is "worth testing" myopically, then stopping is optimal. This implies that one comparison of $\max\{w_{\max}, \mu\}$ to the highest upper testing threshold $\max_i \theta_i^+$ (among untested edges) is sufficient to determine whether to stop or to continue testing.

If stopping is suboptimal, then an optimal policy must now determine which edge to test. It may seem that the entire distribution of each untested edge $W_i$, and specifically the conditional distribution $W_i|W_i > w_{\max}$ might affect this decision, but perhaps surprisingly, we prove in Lemma 2.12 in Section 2.4.3 that it is always optimal to test the edge with the highest upper testing threshold $\theta_i^+$. This implies that the testing order is fixed and can be precomputed before any of the tests, and the only decision to make at a state is merely whether to stop or to continue testing.

Finally, the stopping rule and testing-order rule are combined into Policy 1, which achieves optimal expected profit, as we assert in Theorem 2.13 in Section 2.4.4.

### 2.4.1 Preliminaries

**Definition 2.6.** *(Testing thresholds).*

*The upper testing threshold of a random variable $W_i$ is the unique $\theta_i^+ \in \mathbb{R}$ satisfying*

$$c = \mathbb{E}\left[\left(W_i - \theta_i^+\right)^+\right]. \tag{7}$$

*The lower testing threshold of a random variable $W_i$ is the unique $\theta_i^- \in \mathbb{R}$ satisfying*

$$c = \mathbb{E}\left[\left(\theta_i^- - W_i\right)^+\right]. \tag{8}$$

*Without loss of generality we assume the edges are numbered such that*

$$\theta_1^+ \geq \theta_2^+ \geq \cdots \geq \theta_N^+. \tag{9}$$

We prove that the thresholds $\theta_i^+, \theta_i^-$ exist and that they are uniquely determined in Section 2.5.

**Definition 2.7.** *When a policy is at a state $\left(\bar{W}, w_{\min}\right)$, denote by $k = k\left(\bar{W}, w_{\min}\right)$ the index of an edge with maximal upper threshold among the yet-untested edges, i.e., $\theta_k^+ = \max_{W_i \in \bar{W}} \theta_i^+$.*

**Lemma 2.8.** *There are only two possibilities for the order between the thresholds $\theta_i^+, \theta_i^-$ and $\mu = \mathbb{E}[W_i]$, specifically, if $\mathbb{E}\left[(W_i - \mu)^+\right] \geq c$ then $\theta_i^- \leq \mu \leq \theta_i^+$, otherwise $\theta_i^+ < \mu < \theta_i^-$.*

The proof is identical to that of Lemma 2.3.

**Definition 2.9.** *(Policies)*

*For state* $(\bar{W}, w_{\max})$, *denote an optimal policy by* $\Pi^{\text{opt}}(\bar{W}, w_{\max})$ *and let* $J^{\text{opt}}(\bar{W}, w_{\max})$ *be its expected profit, as defined in section 2.2.*

*For state* $(\bar{W}, w_{\max})$, *denote by* $\Pi^{\text{stop}}(\bar{W}, w_{\max})$ *a policy that stops and returns* $\max\{w_{\max}, \mu\}$ *(where* $\mu$ *represents returning an untested edge). Its expected profit is*

$$J^{\text{stop}}(\bar{W}, w_{\max}) = \max\{w_{\max}, \mu\}. \tag{10}$$

*For state* $(\bar{W}, w_{\max})$ *with* $N > 1$ *untested edges, denote by* $\Pi_i^{\text{my}}(\bar{W}, w_{\max})$ *a myopic policy that tests* $W_i$ *and then stops and returns* $\max\{w_{\max}, \mu, w_i\}$. *Its expected profit is*

$$J_i^{\text{my}}(\bar{W}, w_{\max}) = -c + \mathbb{E}[\max\{w_{\max}, \mu, W_i\}]. \tag{11}$$

### 2.4.2 Stopping-Rule Analysis

In this section, we show in Lemma 2.11 that a myopic stopping rule with a simple, threshold-based structure, is optimal. First, we will need the following technical lemma that compares between a myopic policy $\Pi_i^{\text{my}}$ and stopping $\Pi^{\text{stop}}$.

**Lemma 2.10.** *For each state* $(\bar{W}, w_{\max})$ *with* $N > 1$ *untested edges, and for each untested edge* $W_i$, *policy* $\Pi^{\text{stop}}$ *is at least as good as the myopic policy* $\Pi_i^{\text{my}}$, *i.e.,*

$$J^{\text{stop}}(\bar{W}, w_{\min}) \geq J_i^{\text{my}}(\bar{W}, w_{\min})$$

*if and only if*

$$\max\{w_{\max}, \mu\} \geq \theta_i^+.$$

*Proof.* By equations (10),(11), Fact 2.4 and Definition 2.6,

$$\begin{aligned} J^{\text{stop}}(\bar{W}, w_{\min}) - J_i^{\text{my}}(\bar{W}, w_{\min}) &= \max\{w_{\max}, \mu\} + c - \mathbb{E}[\max\{w_{\max}, \mu, W_i\}] \\ &= c - \mathbb{E}[\max\{w_{\max}, \mu, W_i\} - \max\{w_{\max}, \mu\}] \\ &= \mathbb{E}\left[(W_i - \theta_i^+)^+\right] - \mathbb{E}\left[(W_i - \max\{w_{\max}, \mu\})^+\right]. \end{aligned}$$

By the monotonicity property proved in Lemma 2.14 (see Section 2.5),

$$\mathbb{E}\left[(W_i - \theta_i^+)^+\right] - \mathbb{E}\left[(W_i - \max\{w_{\max}, \mu\})^+\right] \geq 0$$

iff $\max\{w_{\max}, \mu\} \geq \theta_i^+$, and this concludes the proof. ∎

**Lemma 2.11.** *(Stopping Rule) For every state* $(\bar{W}, w_{\max})$ *and* $N > 1$ *untested edges, stopping is optimal, i.e.,* $J^{\text{opt}}(\bar{W}, w_{\max}) = J^{\text{stop}}(\bar{W}, w_{\max})$ *iff*

$$\max\{w_{\max}, \mu\} \geq \theta_k^+.$$

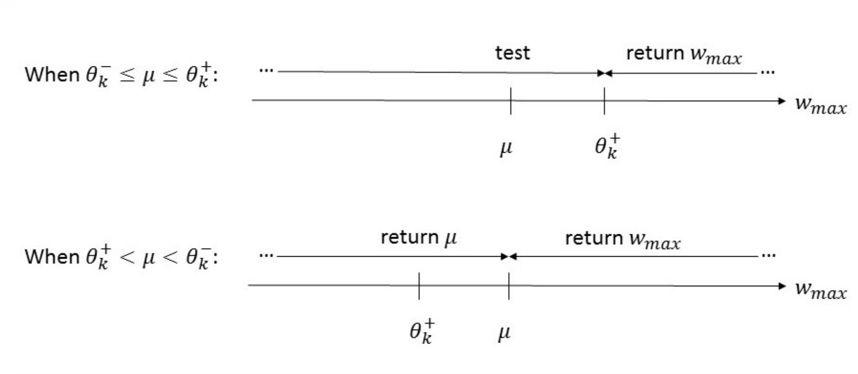


Figure 2.2: An optimal policy for $N > 1$ untested edges as a function of $w_{\max}$, as described in Lemma 2.11. Recall that $W_k$ is an edge with maximal upper testing threshold among the yet-untested edges. The top figure depicts the case $\theta_k^- \leq \mu \leq \theta_k^+$ and the bottom figure depicts the other case $\theta_k^+ < \mu < \theta_k^-$ (see Lemma 2.8).

*Proof.* Recall that by Definition 2.7, $w_k$ is an edge with maximal upper threshold among the yet-untested edges. First, assume that $\max\{w_{\max}, \mu\} < \theta_k^+$. To show that stopping is sub-optimal, it suffices to show a policy $\Pi'$ such that $J^{\text{stop}}(\bar{W}, w_{\max}) < J^{\Pi'}(\bar{W}, w_{\max})$. Taking $\Pi'$ to be $\Pi_k^{\text{my}}$ gives the desired result by Lemma 2.10, and this concludes the forward direction.

We prove the other direction by induction on the number of untested edges. Assume that $\max\{w_{\max}, \mu\} \geq \theta_k^+$ (thus $\max\{w_{\max}, \mu\} \geq \theta_i^+$ for every yet untested edge $W_i$). We shall prove this direction for all $N \geq 1$ edges, and the base case $N = 1$ was already proved in Lemma 2.5, cases 1 and 2(a).

For $N > 1$ untested edges, by the induction hypothesis we know that stopping is optimal for any subset of $N - 1$ untested edges out of $\bar{W}$. Assume that an optimal policy $\Pi^{\text{opt}}$ does not stop, and hence tests some edge $W_i$, at which point the yet-untested edges $\bar{W} \setminus \{W_i\}$ are a subset of $\bar{W}$ of size $N - 1$. Since $w_{\max}$ can only increase after the test, it is optimal to stop. Therefore, an optimal policy will only perform one test and then stop, i.e., it is the myopic policy $\Pi_i^{\text{my}}$. By Lemma 2.10,

$$J^{\text{stop}}(\bar{W}, w_{\max}) \geq J_i^{my}(\bar{W}, w_{\max}) = J^{\text{opt}}(\bar{W}, w_{\max}).$$

Therefore stopping is optimal for $N$ untested edges, which proves the inductive step. The testing and stopping regions are visualized in Figure 2.2. □

### 2.4.3 Testing-Order Analysis

In this section, we show in Lemma 2.12 that there exists an optimal policy that tests edges in a predetermined testing-order, which depends only on their upper testing thresholds $\theta_i^+$.

**Lemma 2.12.** *(Testing Order) There exists an optimal policy that tests edges only in non-increasing order of their upper thresholds* $\theta_i^+$*.*

*Proof.* Proceed by induction on the number $N \geq 1$ of untested edges. The base case $N = 1$ trivially holds as there is only one possible order. For $N > 1$ untested edges, assume by induction that for any subset of $N-1$ untested edges out of $\bar{W}$, whenever stopping is sub-optimal, testing an edge with the largest upper testing threshold is optimal. Fix an optimal policy $\Pi^{\text{opt}}$. If $\Pi^{\text{opt}}$ stops or tests $W_k$ (recall by Definition 2.7 that $\theta_k^-$ is the largest upper testing threshold) then we are done by the induction hypothesis. Otherwise, $\Pi^{\text{opt}}$ tests $W_i$ for some $i \neq k$. By Lemma 2.11, not stopping implies $\max\{w_{\max}, \mu\} < \theta_k^+$. Notice that out of $w_{\max}$, $\mu$ and $\theta_k^+$, testing an edge $i \neq k$ can only change $w_{\max}$. According to Lemma 2.11, after testing $W_i$ there are 2 cases: the first is $W_i \geq \theta_k^+$, and then $\Pi^{\text{opt}}$ stops and returns $W_i$ (by Lemma 2.11); the second is $W_i < \theta_k^+$, and then $\Pi^{\text{opt}}$ does not stop (by the induction hypothesis) and the next edge to be tested will be $W_k$ (notice that $\theta_k^+$ remains minimal until $W_k$ is tested). Now if $W_k \geq \theta_k^+$, then $\Pi^{\text{opt}}$ returns $W_k$ and stops, otherwise it continues according to an optimal policy for that state and achieves expected profit

$$\mathbb{E}\left[J_{-i-k}^{\text{opt}} | W_k < \theta_k^+, W_i < \theta_k^+\right] = \mathbb{E}\left[J^{\text{opt}}\left(\bar{W} \setminus \{W_k, W_i\}, \max\{w_{\max}, W_k, W_i\}\right) | W_k < \theta_k^+, W_i < \theta_k^+\right].$$

The expected profit of policy $\Pi^{\text{opt}}$ can be written as

$$\begin{aligned} J^{\text{opt}}\left(\bar{W}, w_{\max}\right) \quad &= \quad -c + p_i \mathbb{E}\left[W_i | W_i \geq \theta_k^+\right] \\ &\quad + (1-p_i)\left(-c + p_k \mathbb{E}\left[W_k | W_k \geq \theta_k^+\right] + (1-p_k)\,\mathbb{E}\left[J_{-i-k}^{\text{opt}} | W_k < \theta_k^+, W_i < \theta_k^+\right]\right), \end{aligned} \tag{12}$$

where $p_i = \mathbb{P}\left(W_i \geq \theta_k^+\right)$, and similarly $p_k = \mathbb{P}\left(W_k \geq \theta_k^+\right)$.

We now consider an alternative policy $\Pi^{\text{alt}}$ as follows: policy $\Pi^{\text{alt}}$ starts by testing $W_k$. If $W_k \geq \theta_k^+$, then $\Pi^{\text{alt}}$ returns $W_k$ and stops. Otherwise $\Pi^{\text{alt}}$ tests edge $W_i$. If $W_i \geq \theta_k^+$, then $\Pi^{\text{alt}}$ stops and returns $W_i$. Otherwise it tests $W_i$ and then imitates policy $\Pi^{\text{opt}}$ in state $\left(\bar{W} \setminus \{W_k, W_i\}, \max\{w_{\max}, W_k, W_i\}\right)$. The expected profit of policy $\Pi^{\text{alt}}$ can be written as

$$\begin{aligned} J^{\text{alt}}\left(\bar{W}, w_{\max}\right) \quad &= \quad -c + p_k \mathbb{E}\left[W_k | W_k \geq \theta_k^+\right] \\ &\quad + (1-p_k)\left(-c + p_i \mathbb{E}\left[W_i | W_i \geq \theta_k^+\right] + (1-p_i)\,\mathbb{E}\left[J_{-i-k}^{\text{opt}} | W_k < \theta_k^+, W_i < \theta_k^+\right]\right). \end{aligned} \tag{13}$$

Notice that a policy that tests $W_k$ first and then follows the optimal policy is even better than Policy $\Pi^{\text{alt}}$. For example, after testing $W_k$ the next largest testing threshold $\theta_{k'}^+$ may be strictly smaller than $\theta_k^+$, i.e., $\theta_{k'}^+ < \theta_k^+$. According to Lemma 2.11 it is optimal to stop after testing $W_k$ iff $W_i \geq \theta_{k'}^+$. However policy $\Pi^{\text{alt}}$ tests $W_i$ even if $\theta_k^+ > W_i > \theta_{k'}^+$. Hence, to complete the proof, it suffices to show that for every state $\left(\bar{W}, w_{\max}\right)$, the inequality $J^{\text{alt}}\left(\bar{W}, w_{\max}\right) \geq J^{\text{opt}}\left(\bar{W}, w_{\max}\right)$ holds. To this end, using equations (12),(13) and careful manipulation, we obtain

$$\begin{aligned}
J^{\text{alt}}\left(\bar{W}, w_{\max}\right) - J^{\text{opt}}\left(\bar{W}, w_{\max}\right) &= p_k p_i \mathbb{E}\left[W_k | W_k \geq \theta_k^+\right] - p_k p_i \mathbb{E}\left[W_i | W_i \geq \theta_k^+\right] + p_k c - p_i c \\
&= p_k p_i \left(\mathbb{E}\left[W_k | W_k \geq \theta_k^+\right] - \theta_k^+ + \theta_k^+ - \mathbb{E}\left[W_i | W_i \geq \theta_k^+\right]\right) + p_k c - p_i c \\
&= p_k p_i \left(\mathbb{E}\left[W_k - \theta_k^+ | W_k \geq \theta_k^+\right] - \mathbb{E}\left[W_i - \theta_k^+ | W_i \geq \theta_k^+\right]\right) + p_k c - p_i c \\
&= p_i \mathbb{E}\left[\left(W_k - \theta_k^+\right)^+\right] - p_k \mathbb{E}\left[\left(W_i - \theta_k^+\right)^+\right] + p_k c - p_i c \\
&= p_i c - p_k \mathbb{E}\left[\left(W_i - \theta_k^+\right)^+\right] + p_k c - p_i c \\
&= p_k \left(\mathbb{E}\left[\left(W_i - \theta_i^+\right)^+\right] - \mathbb{E}\left[\left(W_i - \theta_k^+\right)^+\right]\right) \\
&\geq 0
\end{aligned}$$

where the last inequality is because the function $b \to \mathbb{E}\left[(W_i - b)^+\right]$ is strictly monotone in $b$ when it take positive values by Lemma 2.14 (see Section 2.5), and this concludes the proof of Lemma 2.12. □

### 2.4.4 A Simple Optimal Policy for Selection with Testing

We can now define Policy 1, and prove that it achieves optimal profit. Policy 1 is threshold-based, and its decision whether to test depends on which of the three $w_{\max}, \mu, \theta_k^+$ is largest (except for the last untested edge). In addition, it tests edges in a predetermined order of non-increasing upper thresholds $\theta_i^+$. These two properties make Policy 1 very easy to implement – the testing order can be computed in advance, and the decision whether to test or stop at any state is made by comparing the three values $w_{\max}, \mu$ and $\theta_k^+$.

**Theorem 2.13.** *For $N \geq 1$ untested edges with equal finite mean $\mu$, Policy 1 achieves optimal expected profit.*

**Policy 1**

1: **for** $l = 1, \ldots, N$ **do** // assuming $\theta_1^+ \geq \theta_2^+ \geq \cdots \geq \theta_N^+$ and $w_{\max} = -\infty$
2:   **if** $w_{\max} \geq \max\left\{\theta_l^+, \mu\right\}$ **then** // $w_{\max}$ is the largest of $\left\{w_{\max}, \mu, \theta_l^+\right\}$
3:     **return** an edge associated with $w_{\max}$
4:   **else if** $\mu \geq \max\left\{\theta_l^+, w_{\max}\right\}$ **then** // $\mu$ is the largest of $\left\{w_{\max}, \mu, \theta_l^+\right\}$
5:     **return** an arbitrary untested edge (without testing)
6:   **else** // $\theta_l^+$ is the largest of $\left\{w_{\max}, \mu, \theta_l^+\right\}$
7:     **if** $l = N$ and $w_{\max} \leq \theta_N^-$ **then**
8:       **return** edge $W_N$ (without testing)
9:     **else**
10:       test $W_l$ to reveal its realization $w_l$ // pay testing cost $c$
11:       set $w_{\max} = \max\{w_{\max}, w_l\}$
12:     **end if**
13:   **end if**
14: **end for**

*Proof.* Notice that lines 2-5 in Policy 1 are equivalent to the following: if $\max\{w_{\max}, \mu\} \geq \theta_l^+$, then stop and return $\max\{w_{\max}, \mu\}$ (in the sense that taking an arbitrary untested edge has expected cost of $\mu$). Hence, for the case $l < N$, i.e., more than one untested edge remaining, Theorem 2.13 is equivalent to claiming that the following is an optimal policy:

1. stop testing iff $\max\{w_{\max}, \mu\} \geq \theta_k^+$.

2. otherwise, test edges in non-increasing order of their upper testing thresholds $\theta_i^+$.

When there is more than one untested edge, Lemma 2.12 implies that testing edges according to their upper testing thresholds $\theta_i^+$ in non-increasing order is an optimal strategy, and Lemma 2.11 indicates when it is optimal to stop. When $l = N$, i.e., there is only one edge left the optimal policy is described in Lemma 2.5. Policy 1 implements the three lemmas and therefore is optimal. The theorem follows. □

## 2.5 Existence and Uniqueness of the Thresholds $\theta_i^-, \theta_i^+$

In this section, we show that for a random variable $W_i$ with a finite mean $\mathbb{E}[W_i] = \mu < \infty$, the thresholds $\theta_i^-, \theta_i^+$ from Definitions 2.2 and 2.6 exist and are uniquely determined. First, we show that the functions $b \to \mathbb{E}\left[(b - X)^+\right]$ and $b \to \mathbb{E}\left[(X - b)^+\right]$ are strictly monotone whenever they receive positive values.

**Lemma 2.14.** *Let $X$ be a random variable and let $I = [l, h]$ be the smallest interval such that $\mathbb{P}(X \in I) = 1$. Then the function $b \to \mathbb{E}[(b - X)^+]$ is strictly monotonically increasing for $b > l$, and is clearly constant $0$ for $b \le l$. Similarly, the function $b \to \mathbb{E}[(X - b)^+]$ is strictly monotonically decreasing for $b < h$, and is clearly constant $0$ for $b \ge h$.*

*Proof.* By definition, for all $b > l$

$$\mathbb{E}\left[(b - X)^+\right] = \int_{-\infty}^{\infty} \max\{b - x, 0\}\, p_x(x)\, dx = \int_l^b (b - x)\, p_x(x)\, dx. \tag{14}$$

Notice that $\int_l^b p_x(x)\, dx > 0$ for all $b > l$, thus for any $\delta > 0$,

$$\begin{aligned}
\mathbb{E}\left[(b + \delta - X)^+\right] &= \int_l^{b+\delta} (b + \delta - x)\, p_x(x)\, dx \\
&\ge \int_l^b (b + \delta - x)\, p_x(x)\, dx \\
&= \int_l^b (b - x)\, p_x(x)\, dx + \delta \underbrace{\int_l^b p_x(x)\, dx}_{>0} \\
&> \mathbb{E}\left[(b - X)^+\right],
\end{aligned}$$

and therefore $\mathbb{E}[(b - X)^+]$ is strictly monotonically increasing in $b$ when $b > l$. For $b \le l$, clearly $\mathbb{E}[(b - X)^+] = 0$.

The second part of the lemma is proved in a similar manner. □

We can now use Lemma 2.14 to prove the existence and uniqueness of the testing thresholds.

**Corollary 2.15.** *If a random variable $W_i$ has a finite mean $\mathbb{E}[W_i] < \infty$, then the lower and upper testing thresholds $\theta_i^-$ and $\theta_i^+$ exist and are uniquely determined.*

*Proof.* By Lemma 2.14, the functions $b \to \mathbb{E}\left[(b - W_i)^+\right]$ and $b \to \mathbb{E}\left[(W_i - b)^+\right]$ are strictly monotonic in $b$ whenever they are positive and thus attain any positive value $c > 0$ exactly once ($\mathbb{E}\left[(b - W_i)^+\right] \to \infty$ when $b \to \infty$ and $\mathbb{E}\left[(W_i - b)^+\right] \to \infty$ when $b \to -\infty$). Therefore, there must exist unique solutions for equations (7) and (8) (and in particular to equations(2) and (3)). □

# 3 Linear Optimization over a Polymatroid with Testing

In this section, we study a class of problems that generalize the selection problem discussed in Section 2. We call these *linear optimization over a polymatroid with testing* (LOPT) problems. In *linear optimization over a polymatroid* (LOP), the objective is to maximize a linear function when the constraint polyhedron has the structure of a polymatroid (Schrijver (2003)), a well-known notion in combinatorial optimization that is defined through an exponential number of constraints. In particular, for each subset of the decision variables there is a constraint asserting that the sum of these variables is smaller than a given submodular set function evaluated on this subset. In LOPTs, each coefficient of the linear objective is a random variable that can be tested to obtain its realization. There is a fixed cost associated with testing a coefficient, and the goal is to maximize the expected value of the linear program minus the testing costs.

We start by reviewing known results about polymatroid optimization problems in Section 3.1, and provide a few representative problems that can be modeled as polymatroids in Section 3.2. We then formally formulate the LOPT problem in Section 3.3 and prove that when the random coefficients have equal means, the myopic stopping rule for the selection problem (Lemma 2.11) is also optimal in deciding when to stop testing in LOPTs (Section 3.4). We then show that in some interesting special cases it is optimal to decide on the testing order myopically, based on the expected improvement to the objective with a limited budget of a single test. These cases include the Maximum Spanning Tree problem when edge weights are independent and identically distributed (Section 3.5), and a family of symmetric LOPTs, which implies that underlying optimization problem is symmetric under permutations of the coefficients (Section 3.6). In the latter, the random coefficients are drawn independently from distributions that satisfy convex order, which is a partial order between probability distributions that have equal means but different magnitudes of uncertainty (the convex order is discussed in detail in Section 3.6).

## 3.1 Submodular Set Functions and Polymatroids

We now review some basic definitions and properties of submodular set functions and polymatroids, as described in detail by Yao and Zhang (1997) and by Schrijver (2003).

**Submodular set functions**

Given a finite ground set $N$, a set function $f : 2^N \to \mathbb{R}$ is called *submodular* if for all subsets $A, B \subseteq N$,

$$f(A) + f(B) \geq f(A \cap B) + f(A \cup B). \tag{15}$$

Equivalently, $f$ is submodular if for all subsets $A \subseteq B \subseteq N$ and for all elements $e$ such that $e \in N \setminus B$,

$$f(A+e) - f(A) \geq f(B+e) - f(B). \tag{16}$$

Similarly to concavity, submodularity is a property of diminishing marginal returns. The above definition asserts that when an element $e$ is added to a set $A$, the value of $f$ increases at by the same or larger amount than when the same element is added to a larger set $B \supseteq A$. In fact, in the special case that $f(A)$ depends only on the cardinality of $A$, the function $f$ must be a concave function of $|A|$.

Supermodular function are defined analogously. $f$ is called *supermodular* if $-f$ is submodular, i.e., $f$ satisfies equations (15) and (16) with the inequality sign reversed. $f$ is *modular* if $f$ is both submodular and supermodular, i.e., if $f$ satisfies (15) and (16) with an equality. Note that the relation between submodular and supermodular functions is similar to the relation between concave and convex functions (for example, in terms minimizing one versus maximizing the other).

A set function $f$ on $N$ is called *non-decreasing* if $f(A) \leq f(B)$ whenever $A \subseteq B \subseteq N$, and *non-increasing* if $f(A) \geq f(B)$ whenever $A \subseteq B \subseteq N$. We say that $f$ is normalized if $f(\emptyset) = 0$.

**Polymatroids**

For each subset $S \subseteq N$, define

$$x(S) := \sum_{e \in S} x_e. \tag{17}$$

Using this notation, define a polyhedron $P_f$ associated with a set function $f$ on $2^N$

$$P_f = \left\{ x \in \mathbb{R}^N \,|\, x \geq \mathbf{0} \text{ and } x(S) \leq f(S) \;\forall S \subseteq N \right\}. \tag{18}$$

By Yao and Zhang (1997), $P_f$ is called the *polymatroid associated with* $f$ if $f$ is normalized, non-decreasing and submodular. Observe that every polymatroid is bounded (since $0 \leq x_e \leq f(\{e\})$ for each $e \in N$), and hence it is a polytope.

**Optimization over a Polymatroid using the Greedy Algorithm**

Let $f : 2^N \to \mathbb{R}$ be a normalized, non-decreasing, and submodular function, given by a *value giving oracle*, that is, by an oracle that returns $f(S)$ for any $S \subseteq N$. Given $f$ and a vector of positive coefficients $w = (w_1, \ldots, w_N) \in \mathbb{R}_+^N$, the maximization problem

$$\varphi(w) = \max_{x \in P_f} \sum_{e \in N} w_e x_e \tag{19}$$

is called a *Linear Optimization over a Polymatroid* (LOP) problem. Despite the fact that polymatroids are defined using exponentially many constraints, LOPs can be solved rather efficiently using the following greedy algorithm (see Yao and Zhang (1997)):

1. Define a non-increasing order on the elements by a permutation $\sigma : [N] \to [N]$ with respect to their coefficients, i.e., $w_{\sigma(1)} \geq w_{\sigma(2)} \geq ... \geq w_{\sigma(N)}$. Assume ties are broken in a consistent manner.

2. For each $i \in N$ let

$$x_{\sigma(i)} = f\left(\{\sigma(1), ..., \sigma(i-1), \sigma(i)\}\right) - f\left(\{\sigma(1), ..., \sigma(i-1)\}\right). \tag{20}$$

The greedy algorithm is a strongly polynomial-time algorithm since it requires a linear number of calls to the value oracle. Observe that if $f$ is integer valued, then the solution $x$ given by Equation (20) is integral.

Next, we provide examples of combinatorial optimization problems that can be written as LOPs.

## 3.2 Examples of LOPs

Linear optimization over a polymatroid captures many important problems in graph theory, linear algebra, and other branches of mathematics and computer science. We now review several representative examples.

**K-Selection**

In Section 2, we studied the problem of selecting an element from a given set so as to maximize its value. We now consider the more general problem of selecting $K$ elements from a given set so as to maximize their combined value. Without testing, this problem can be written as the linear program

$$\begin{array}{ll} \max & w^{\mathrm{T}} x \\ \text{s.t.} & \sum\limits_{i=1}^{N} x_i \leq K \\ & 0 \leq x_i \leq 1 \quad \forall i \in N. \end{array} \tag{21}$$

Interestingly, the problem can be also written as an LOP using the submodular function

$$f(S) = \begin{cases} 0 & S = \emptyset \\ \min(K, |S|) & \text{otherwise.} \end{cases} \tag{22}$$

To see this, we divide the constraints based on the cardinality of the set $S$:

1. $|S| = 1$: the resulting constraints are $x_i \leq 1$ for each $i \in N$,

2. $|S| = N$: the resulting constraint is $\sum\limits_{i=1}^{N} x_i \leq K$,

3. $|S| \leq K$: the resulting constraint can be written as $\sum\limits_{i \in S} x_i \leq |S|$, which is redundant by 1,

4. $|S| > K$: the resulting constraint can be written as $\sum_{i \in S} x_i \leq K$ and is redundant by 2.

**Matroid Optimization**

A matroid is a combinatorial structure that captures and generalizes the notion of independent sets in vector spaces. It is of particular interest as many subset selection problems can be formulated using matroids (either directly or using composition), and due to the algorithms for solving the corresponding optimization problems.

Formally, a matroid is a pair $(\mathcal{S}, \mathcal{I})$, where $\mathcal{S}$ is a finite set of elements (called the ground set), and $\mathcal{I}$ is a non-empty collection of subsets of $\mathcal{S}$ (called the independent sets), which satisfies the following properties (Schrijver (2003)):

1. if $I \in \mathcal{I}$, and $J \subset I$, then $J \in \mathcal{I}$ (closure to taking subsets),
2. if $I, J \in \mathcal{I}$ and $|I| < |J|$, then $I + z \in \mathcal{I}$ for some $z \in J \setminus I$ (all inclusion-wise maximal independent sets have the same cardinality).

We define the *rank function* of a set $U \subset \mathcal{S}$ to be the cardinality of the largest independent subset of $U$:

$$r(U) := \max\{|Z| : Z \in \mathcal{I}, Z \subseteq U\}.$$

Edmonds (1970) showed that given a matroid and a weight function on the ground set $w : \mathcal{S} \rightarrow \mathbb{R}_+$, finding an independent set whose total weight is maximal can be formulated as an LOP with the submodular function $f$ being the rank function $r$:

$$\begin{array}{ll} \max & \sum_{e \in \mathcal{S}} w(e) x_e \\ \text{s.t.} & x(U) \leq r(U) \quad \forall U \subseteq \mathcal{S} \\ & x_e \geq 0 \end{array} , \tag{23}$$

where $x_e$ is a decision variable that indicates if element $e \in \mathcal{S}$ is in the resulting set. Note that the rank of any singleton $\{e\}$ is equal to 1, and therefore $0 \leq x_e \leq 1$ for each $e \in \mathcal{S}$. Since $r$ is an integral submodular function, using the greedy algorithm we obtain that $x_e$ are binary variables.

**Graphic Matroids and the Maximum Spanning Tree (MST)**

One particularly interesting class of matroids is the *Graphic Matroid* (see Schrijver (2003)), defined as follows. Let $G = (V, E)$ be a connected graph with vertices $V = \{v_1, ..., v_n\}$ and edges $E = \{e_1, ..., e_m\}$. We denote by $V(E') \subset V$ the subset of vertices covered by $E'$:

$$V(E') = \left\{v_1 \in V : \exists v_2 : (v_1, v_2) \in E'\right\}.$$

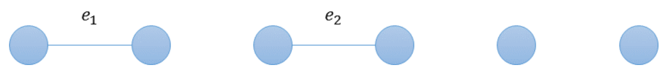


Figure 3.3: Example of subset of edges in a graphic matroid.

In the graphic matroid $(\mathcal{S}, \mathcal{I})$ defined over $G$, the ground set $\mathcal{S}$ contains all edges (that is, $\mathcal{S} = E$), and the collection of independent sets $\mathcal{I}$ contains the subsets $E' \subseteq E$ that form a forest (i.e., there are no cycles in $E'$). The rank function $r$ for the graphic matroid can then be written as follows. For a subset $E' \subseteq E$, let $\kappa(V, E')$ denote the number of components in the graph $G(V, E')$. Then, it is easily verified that

$$r(E') = |V(E')| - \kappa(V, E').$$

An interesting interpretation of the resulting LOP is that for any non-empty subset of edges $E' \subseteq E$ such that $G(V, E')$ has only one connected component, this constraint implies that $x(E') \leq |V(E')| - 1$, and is hence a spanning tree of the subgraph $G(V, E')$. Moreover, all the other constraints (associated with non-empty subsets $E' \subseteq E$ such that $G(V, E')$ has more than one connected component) are redundant and can be expressed using constraints on non-empty subsets where the induced subgraph has one connected component. Figure 3.3 illustrates a subset of two edges $e_1, e_2$, in a graphic matroid that includes 6 vertices. The rank function associated with the sets $\{e_1\}$,$\{e_2\}$, and $\{e_1, e_2\}$ is:

$$r(\{e_1\}) = 6 - 5 = 1, \quad r(\{e_2\}) = 6 - 5 = 1, \quad r(\{e_1, e_2\}) = 6 - 4 = 2.$$

The resulting constraints are:

$$x_1 \leq 1, \quad x_2 \leq 1, \quad x_1 + x_2 \leq 2,$$

respectively. The edges in of the set $\{e_1, e_2\}$ connect two components and the constraint associated with the set $\{e_1, e_2\}$ is indeed redundant.

Given edge weights $w_1, .., w_m$, the problem of finding a spanning tree of the graph $G$ (i.e, a subset of edges that connect all the vertices in the graph and does not contain a cycle) with maximal total weight is therefore a matroid optimization problem and can be solved using the following LOP:

$$\begin{array}{rl} \max & \sum_{e \in E} w_e x_e \\ \text{s.t.} & x(E') \leq r(E') \quad \forall E' \subseteq E \\ & x_e \geq 0 \end{array} \tag{24}$$

where $x_e$ represents the decision to include edge $e \in E$ in the maximum spanning tree.

Observe that the selection problem of a single element can be viewed as a special case of the maximum spanning tree problem, where $G$ has two vertices and parallel edges between them.

### 3.3 The LOPT Problem Formulation

Linear Optimization over a Polymatroid with Testing (LOPTs) are LOPs, in which the objective coefficients $\bar{W}$ are independent random variables from known distributions. Without testing, the vector $x$ that maximizes the expected objective value is the same vector $x$ that solves the LOP when the objective coefficients are $\mathbb{E}[\bar{W}]$:

$$\max_x \mathbb{E}\left[\bar{W}^{\mathrm{T}} x\right] = \max_x \mathbb{E}\left[\bar{W}\right]^{\mathrm{T}} x, \tag{25}$$

which follows directly from of the linearity of the expectation operator.

In LOPTs we assume that a decision-maker can test the coefficients prior to returning a feasible solution, i.e., solving the corresponding LOP based on the current information. Testing an objective coefficient $W_i$ reveals its realization, denoted subsequently $w_i$, but incurs a cost $c > 0$. After testing, a decision-maker can either test a yet-untested coefficients, or stop testing and simply optimize with respect to the tested coefficients and expected values of the untested coefficients, similarly to (25). Our goal is to develop an adaptive policy that maximizes the expected profit, where profit is defined as the returned solution's objective value minus all the testing costs. A policy for an LOPT problem decides adaptively whether to continue testing or optimize the LOP with respect to the expected values of the untested objective coefficients and known values of the tested objective coefficients. Whenever a policy decides to test, it must also decide which yet-untested coefficient to test, and this decision could depend on the values of already tested coefficients. To put things in context of the examples of Section 3.2, in the K-Selection problem, the coefficients $\bar{W}$ represent the random valuation of each element of the ground set. In the MST problem, the objective represent the random edges weights.

**Dynamic Programming Formulation of LOPTs**

The system state of an LOPT can be described as a tuple $\left(\bar{W}, \bar{w}\right)$, where $\bar{W}$ denotes the vector of random coefficients, and $\bar{w}$ denotes the vector of tested coefficients. After testing coefficient $W_i \in \bar{W}$, we transition to state $\left(\bar{W} - W_i, \bar{w} + W_i\right)$, where we use '-' and '+' to denote exclusion and inclusion of elements from a set, i.e., $\bar{W} - W_i \equiv \bar{W} \setminus \{W_i\}$ and $\bar{w} + W_i \equiv \bar{w} \cup \{W_i\}$. With a slight abuse of notation, we use $\varphi(\mathbb{E}[\bar{W}], \bar{w})$ to denote the expected value of an optimal solution to the LOP without future testing. This is based on the observation of Equation (25), in which the optimal solution to the LOP with random objective coefficients can be obtained from the LOP by replacing the untested coefficients with their expected values $\mathbb{E}[\bar{W}]$. Denote the value function of the optimal policy at state $\left(\bar{W}, \bar{w}\right)$ by $J^{\mathrm{opt}}\left(\bar{W}, \bar{w}\right)$. The dynamic programming formulation of the

LOPT problem can be then written as:

$$J^{\mathrm{opt}}\left(\bar{W},\bar{w}\right) = \max\begin{cases}\varphi(\mathbb{E}[\bar{W}],\bar{w}) & \text{stop}\\ -c+\mathbb{E}_{W_i}\left[J^{\mathrm{opt}}\left(\bar{W}-W_i,\bar{w}+W_i\right)\right] & \text{test } W_i\end{cases} \tag{26}$$

To motivate why testing could be valuable observe that $\varphi(\mathbb{E}[\bar{W}],\bar{w})$ computes a solution that is optimal with respect to the mean of the untested coefficients. However, this solution is not necessarily optimal with respect to each realization, and revealing the exact values through testing could therefore increase the total expected objective value function. In the extreme case, we can test all the coefficients and ensure obtaining the optimal solution for any specific objective vector realization, however, the testing costs might be too costly. The goal is to find a policy that optimally balances the benefit from reducing the uncertainty associated with the objective vector and the cost incurred by testing.

Observe that a state $\left(\bar{W},\bar{w}\right)$ corresponds to the product of a subset of untested coefficients and all possible realizations of the tested coefficients. This results in a huge state space, which we cannot hope to solve optimally without characterizing the optimal policy. To do so, we examine the structural properties of LOPs.

**Properties of LOPs and LOPTs**

Recall that by equation (19), for a given polymatroid $P_f$, the function $\varphi(w)$ denotes the optimal value of an LOP with a coefficients vector $w \in \mathbb{R}_+^N$. We now show what happens to the value of an LOP when one of the coefficients changes and the rest remain fixed. This will be useful for showing the optimality of a myopic stopping rule in Section 3.4.

**Lemma 3.1.** *Without loss of generality, consider coefficient $w_1$ and assume the remaining coefficients are fixed and sorted in non-increasing order $w_2 \geq \cdots \geq w_N$. Define the function $\tilde{\varphi}(w_1) = \varphi(w)$ of equation (19). The function $\tilde{\varphi}(w_1)$ satisfies the following properties:*

1. *$\tilde{\varphi}(w_1)$ is continuous, convex, and piecewise linear;*
2. *The separation points between linear segments in $\tilde{\varphi}(w_1)$ are the values of the remaining coefficients $w_2,\ldots,w_N$. In other words, $\tilde{\varphi}(w_1)$ is linear in each of the intervals $(w_i,w_{i+1})$ for all $i = 2,\ldots,N-1$.*
3. *The derivative of $\tilde{\varphi}(w_1)$ in any linear segment $(w_i,w_{i+1})$ is constant and equals to*

$$\frac{\partial}{\partial w_1}\varphi(w) = \frac{d}{dw_1}\tilde{\varphi}(w_1) = f(\{1,\ldots,i\}) - f(\{2,\ldots,i\}).$$

*Proof.* The first property is a known result for polytopes (see Theorem 5.3 in page 217 of Bertsimas and Tsitsiklis (1997)), and the other properties follow directly from the same theorem, which states that the derivative with respect to an objective coefficient is equal to the value of the respective decision variable. For LOPs, this decision variable is given by equation (20) of the greedy algorithm, which proves property 3. □

One can interpret the dynamic programming formulation of the LOPT as a composition of LOPs. As a consequence, some of the properties of LOPs carry through to LOPTs as given by the following lemma.

**Lemma 3.2.** *The value function $J^{\text{opt}}\left(\bar{W}, \bar{w}\right)$ is continuous, convex, and piecewise linear in every tested coefficient $w_t \in \bar{w}$.*

*Proof.* Proceed by induction on $N$. When $N = 0$, then $\bar{W} = \emptyset$ and the value function of $\Pi^{\text{opt}}$ is

$$J^{\text{opt}}\left(\emptyset, \bar{w}\right) = \varphi\left(\emptyset, \bar{w}\right)$$

which is continuous, convex, and piecewise linear in $w_t \in \bar{w}$ by Lemma 3.1.

For $N \geq 1$, the value function of $\Pi^{\text{opt}}$ is given by equation (26). By the induction hypothesis, $J^{\text{opt}}\left(\bar{W} - W_i, \bar{w} + W_i\right)$ is continuous, convex, and piecewise linear in $w_t$ for any random coefficient $W_i \in \bar{W}$ and any realization of $W_i$. Since that continuity, convexity, and the piecewise linearity are being preserved through summation and the maximum operator, the function $J^{\text{opt}}\left(\bar{W}, \bar{w}\right)$ is continuous, convex, and piecewise linear for every $w_t \in \bar{w}$. □

### Myopic Policies

We define a few terms that will be useful in the discussion of myopic policies:

**Definition 3.3.** *$\Pi^{\text{stop}}$ is the policy that at every state $\left(\bar{W}, \bar{w}\right)$ stops testing and solves the optimization problem using the expected values of the untested parameters. The value function $J^{\text{stop}}$ of policy $\Pi^{\text{stop}}$ satisfies:*

$$J^{\text{stop}}\left(\bar{W}, \bar{w}\right) = \varphi\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right). \tag{27}$$

**Definition 3.4.** *$\Pi_i^{\text{test}}$ is the policy that every state $\left(\bar{W}, \bar{w}\right)$ tests parameter $W_i \in \bar{W}$ and continues according to the optimal policy $\Pi^{\text{opt}}$. The value function $J_i^{\text{test}}$ of policy $\Pi_i^{\text{test}}$ satisfies:*

$$J_i^{\text{test}}\left(\bar{W}, \bar{w}\right) = -c + \mathbb{E}_{W_i}\left[J^{\text{opt}}\left(\bar{W} - W_i, \bar{w} + W_i\right)\right]. \tag{28}$$

**Definition 3.5.** *$\Pi_i^{\text{my}}$ is the policy that at every state $\left(\bar{W}, \bar{w}\right)$ tests parameter $W_i \in \bar{W}$ and stops. The value function $J_i^{\text{my}}$ of policy $\Pi_i^{\text{my}}$ satisfies:*

$$J_i^{\text{my}}\left(\bar{W}, \bar{w}\right) = -c + \mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W} - W_i\right], \bar{w} + W_i\right)\right]. \tag{29}$$

**Definition 3.6.** *For state* $(\bar{W}, \bar{w})$ *and untested parameter* $W_i \in \bar{W}$ *denote by* $\Delta_i(\bar{W}, \bar{w})$ *the myopic gain from testing* $W_i$*,*

$$\Delta_i(\bar{W}, \bar{w}) = J_i^{\text{my}}(\bar{W}, \bar{w}) - J^{\text{stop}}(\bar{W}, \bar{w}). \quad (30)$$

**Remark 3.7.** *Since the myopic policy* $\Pi_i^{\text{my}}(\bar{W}, \bar{w})$ *only observes one more realization and the stopping policy obtains no new information,* $\Delta_i$ *is only sensitive to* $W_i$*, and treats all other untested coefficients as their expectations. Hence,* $\Delta_i(\bar{W}, \bar{w}) = \Delta_i(\bar{W} - W_l, \bar{w} + \mathbb{E}[W_l])$ *for all* $W_l \in \{\bar{W} - W_i\}$*.*

## 3.4 An Optimal Stopping Rule for LOPT with Coefficients with Equal Means

We now prove that a myopic rule is optimal in determining when to stop testing coefficients in LOPTs in which the untested coefficients have equal means. We start by proving that a certain monotonicity property holds, which we then use to show that the myopic stopping rule is indeed optimal.

**Lemma 3.8.** *At each state* $(\bar{W}, \bar{w})$*, for each untested parameter* $W_i \in \bar{W}$ *and tested parameter* $w_t \in \bar{w}$*, the function* $\Delta_i(\bar{W}, \bar{w})$ *of an LOPT with equal mean* $\mu$ *of all untested coefficients is unimodal in* $w_t$ *and maximized at* $w_t = \mu$*.*

*Proof.* To prove the above, we first show that the function $\Delta_i$ is continuous and piecewise linear in $w_t \in \bar{w}$. We then show that in every linear segment where $w_t < \mu$, the derivative of $\Delta_i$ with respect to $w_t$ is non-negative, and that in every linear segment where $w_t > \mu$ the derivative is non-positive. The continuity of $\Delta_i$ would then imply that $\Delta_i$ is unimodal in $w_t$ and achieves its maximal value at $w_t = \mu$.

By definition (Equation (30)) the function $\Delta_i(\bar{W}, \bar{w})$ can be written as follows:

$$\begin{aligned} \Delta_i(\bar{W}, \bar{w}) &= J_i^{\text{my}}(\bar{W}, \bar{w}) - J^{\text{stop}}(\bar{W}, \bar{w}) \\ &= -c + \mathbb{E}_{W_i}\left[\varphi(\bar{W} - W_i, \bar{w} + W_i)\right] - \varphi(\bar{W}, \bar{w}) \\ &= -c + \sum_j Prob(W_i = w'_j)\varphi(\bar{W} - W_i, \bar{w} + w'_j) - \varphi(\bar{W} - W_i, \bar{w} + \mu), \end{aligned} \quad (31)$$

where $w'_j$ denotes specific realizations of the random coefficient $W_i$. In the last equality we use the fact that under the function $\varphi$, random coefficients are replaced by their expected value (which is why we replaced $\varphi(\bar{W}, \bar{w})$ with $\varphi(\bar{W} - W_i, \bar{w} + \mu)$).

In Equation (31), we see that the function $\Delta_i$ is the weighted summation of the function $\varphi$ applied to different states that share the same untested and tested coefficients, with the exception of having different values for the tested parameter $w_i$. By Lemma 3.1 (while changing to the names of the variables), the function $\varphi$ is continuous in $w_t \in \bar{w}$ and therefore so is the function $\Delta_i$.

Moreover, Lemma 3.1 asserts that the function $\varphi$ is piecewise linear with the non-smooth points being the set of coefficients $\bar{w}\setminus\{w_t\}\cup\{\mu\}$. Therefore, the summation in Equation (31) is a piecewise linear function of $w_t$ where the non-smooth points are the union of all coefficients values:

$$\mathcal{S} = w_t \setminus \{w_t\} \cup \{\mu\} \cup_j \{w'_j\}.$$

This implies that the function $\Delta_i$ is not only continuous, but also piecewise linear in $w_t$. Therefore, the function $\Delta_i$ is linear in every segment defined by two consecutive points in $\mathcal{S}$.

We now take derivative of $\Delta_i\left(\bar{W},\bar{w}\right)$ with respect to $w_t$ to show that the function $\Delta_i\left(\bar{W},\bar{w}\right)$ is non-decreasing in $w_t$ in segments where $w_t < \mu$, and that it is non-increasing in $w_t$ in segments where $w_t > \mu$ (by definition $\mu \in \mathcal{S}$ and therefore all segments are either below or above $\mu$). The derivative of $\Delta_i$ with respect to $w_t$ can be written as follow:

$$\begin{aligned}
\frac{\partial}{\partial w_t}\Delta_i\left(\bar{W},\bar{w}\right) &= \frac{\partial}{\partial w_t}\left(-c+\sum_j Prob(W_i=w'_j)\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right)-\varphi\left(\bar{W}-W_i,\bar{w}+\mu\right)\right)\\
&= \sum_j Prob(W_i=w'_j)\frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right)-\frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+\mu\right)\\
&= \sum_{j:w'_j<w_t}\left(Prob(W_i=w'_j)\frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right)\right)\\
&\quad+\sum_{j:w'_j>w_t}\left(Prob(W_i=w'_j)\frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right)\right)\\
&\quad-\frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+\mu\right), \qquad (32)
\end{aligned}$$

where the second equality follows from the linearity of the derivative operator, and because $c$ is a constant. In the third equality we split the support of $W_i$ to values that are higher and lower than $w_t$. Since that the support of $W_i$ is included in $\mathcal{S}$, there is no value $w'_j$ where $w'_j = w_t$.

Let $w_j^-$ be a realization of $W_i$ that satisfies $w_j^- < w_t$, and let $x_t^-$ denote the derivative of $\varphi\left(\bar{W}-W_i,\bar{w}+w_j^-\right)$ with respect to $w_t$:

$$x_t^- = \frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w_j^-\right).$$

By Lemma 3.1 (property (3)), $x^-$ is insensitive to the value of $w'_j$ as long as it remains below $w_t$:

$$x_t^- = \frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right), \text{ for all } w'_j < w_t.$$

Similarly, we denote by $x_t^+$ the derivative of $\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right)$ with respect to $w_t$ for any realization $w'_j > w_t$:

$$x_t^+ = \frac{\partial}{\partial w_t}\varphi\left(\bar{W}-W_i,\bar{w}+w'_j\right), \text{ for all } w'_j > w_t.$$

We can then write Equation (32) as follows:

$$\begin{aligned}
\frac{\partial}{\partial w_t}\Delta_i\left(\bar{W},\bar{w}\right) &= \sum_{j:w'_j<w_t} Prob(W_i=w'_j)x_t^- + \sum_{j:w'_j>w_t} Prob(W_i=w'_j)x_t^+ - x_t^{stop} \\
&= \sum_{j:w'_j<w_t} Prob(W_i=w'_j)\left(x_t^- - x_t^{stop}\right) + \sum_{j:w'_j>w_t} Prob(W_i=w'_j)\left(x_t^+ - x_t^{stop}\right) \\
&= Prob(W_i<w_t)\left(x_t^- - x_t^{stop}\right) + Prob(W_i>w_t)\left(x_t^+ - x_t^{stop}\right) \qquad (33)
\end{aligned}$$

where $x_t^{stop}$ denotes the derivative of $\varphi\left(\bar{W}-W_i,\bar{w}+\mu\right)$ with respect to $w_t$. Recall that we compute the derivative for each segment independently and that by construction $w_t \neq w'_j$ and $w_t \neq \mu$. Moreover, in every segment $x_t^-$, $x_t^+$, and $x_t^{stop}$ are constants.

Now that we have an expression for the derivative of $\Delta_i\left(\bar{W},\bar{w}\right)$ with respect to $w_t \in \bar{w}$, we show that it non-negative when $w_t<\mu$, and that it is non-positive when $w_t>\mu$.

Consider first the case where $w_t>\mu$. By Lemma 3.1, when $w_t>\mu$, the value of the derivative of $\Delta_i\left(\bar{W}-W_i,\bar{w}+w'_j\right)$ with respect to $w_t\in\bar{w}$ is the same for all realizations $w'_j$ of $W_i$ that are smaller than $w_t$, including $w_t=\mu$, which is why $x_t^{stop}=x_t^-$. We can then write Equation (33) as follows:

$$\begin{aligned}
\frac{\partial}{\partial w_t}\Delta_i\left(\bar{W},\bar{w}\right) &= Prob(W_i<w_t)\left(x_t^- - x_t^{stop}\right) + Prob(W_i>w_t)\left(x_t^+ - x_t^{stop}\right) \\
&= Prob(W_i<w_t)\left(0\right) + Prob(W_i>w_t)x_t^+ - Prob(W_i>w_t)x_t^{stop} \\
&= Prob(W_i>w_t)\left(f(\{j:w_j>w_t\}\cup\{t,i\}) - f(\{j:w_j>w_t\}\cup\{i\})\right) \\
&\quad -Prob(W_i>w_t)\left(f(\{j:w_j>w_t\}\cup\{t\}) - f(\{j:w_j>w_t\})\right) \\
&\leq 0,
\end{aligned}$$

where the third equality follows from Lemma 3.1, and the fact that $w_t>\mu$, and because $x_t^+$ corresponds to realizations of $W_i$ that are higher than $w_t$. The inequaity results from the submodularity of the function $f$.

Similarly, when $w_t<\mu$, the derivative of $\Delta_i\left(\bar{W}-W_i,\bar{w}+w'_j\right)$ with respect to $w_t$ is equal for all the realizations of $W_i$ that are larger than $w_t$ including $\mu$. This implies that $x_t^{stop}=x_t^+$, and Equation (33) can be written as follows:

$$\begin{aligned}
\frac{\partial}{\partial w_t}\Delta_i\left(\bar{W},\bar{w}\right) &= Prob(W_i<w_t)\left(x_t^- - x_t^{stop}\right) + Prob(W_i>w_t)\left(x_t^+ - x_t^{stop}\right) \\
&= Prob(W_i<w_t)x_t^- - Prob(W_i<w_t)x_t^{stop} + 0 \\
&= Prob(W_i<w_t)\left(f(\{j:w_j>w_t\}\cup\bar{W}\cup\{t\}) - f(\{j:w_j>w_t\}\cup\bar{W})\right) \\
&\quad -Prob(W_i<w_t)\left(f(\{j:w_j>w_t\}\cup\bar{W}\cup\{t,i\}) - f(\{j:w_j>w_t\}\cup\bar{W}\cup\{i\})\right) \\
&\geq 0,
\end{aligned}$$

which holds for similar reasons.

The function $\Delta_i$ is decreasing when $w_t > \mu$, and increasing when $w_t < \mu$, and is therefore unimodal and obtains its maximal value in $w_t = \mu$. □

Intuitively, Lemma 3.8 implies that the myopic gain of testing is monotonically decreasing when we test, as coefficients drift from their expected value. We formalize this in the following corollary.

**Corollary 3.9.** *At each state $\left(\bar{W}, \bar{w}\right)$ and for each untested parameter $W_j \in \bar{W}$, the myopic gain $\Delta_j\left(\bar{W}, \bar{w}\right)$ of an LOPT with equal mean $\mu$ does not increase by testing another untested coefficient $W_i \in \bar{W} - W_j$:*

$$\Delta_j\left(\bar{W}, \bar{w}\right) \geq \Delta_j\left(\bar{W} - W_i, \bar{w} + w_i\right), \text{ for every coefficient } W_i \in \bar{W} - W_j, \text{ and realization } w_i.$$

*Proof.* Immediate by Lemma 3.8. □

Corollary 3.9 implies that once all myopic gains are non-positive, they will remain non-positive in future steps. We use this fact to prove that the myopic stopping rule is optimal.

**Definition 3.10.** *We say that a policy adheres to the myopic stopping rule if at every state $\left(\bar{W}, \bar{w}\right)$ the policy stops, if and only if, all myopic gains are non-positive, that is:*

$$\forall W_i \in \bar{W} : \Delta_i\left(\bar{W}, \bar{w}\right) \leq 0.$$

**Theorem 3.11.** *The myopic stopping rule is optimal for LOPTs with untested coefficients that have equal means.*

*Proof.* One direction is straightforward since it is readily verified that if there exists an untested parameter $W_i \in \bar{W}$, such that $\Delta_i\left(\bar{W}, \bar{w}\right) > 0$ clearly stopping is not optimal since that a single test of coefficient $W_i$ outperforms stopping.

We prove the other direction by induction on the number of untested coefficients $k$. When $k = 1$ the myopic stopping rule is optimal by definition. We prove the step $k > 1$ by contradiction. Suppose that at state $\left(\bar{W}, \bar{w}\right)$ the myopic gains are non-positive ($\forall W_i \in \bar{W} : \Delta_i\left(\bar{W}, \bar{w}\right) \leq 0$) and that the optimal policy $\pi$ tests the yet-untested coefficient $W_j$. Using Lemma 3.8, in the next state $\left(\bar{W} - W_j, \bar{w} + W_j\right)$ the myopic gains remain non-positive regardless of the realization of $W_j$, that is:

$$\forall W_i \in \bar{W} - W_j : \Delta_i\left(\bar{W} - W_j, \bar{w} + W_j\right) \leq 0.$$

Using the induction hypothesis, it is therefore optimal to stop in state $\left(\bar{W} - W_j, \bar{w} + W_j\right)$. This implies that policy $\pi$ tests exactly once, and that the following holds:

$$J^{\pi}\left(\bar{W}, \bar{w}\right) = J_j^{MY}\left(\bar{W}, \bar{w}\right) \leq J^{STOP}\left(\bar{W}, \bar{w}\right)$$

where the equality holds since that policy $\pi$ tests exactly once, and the inequality holds since that it is equivalent to $\Delta_j\left(\bar{W},\bar{w}\right)\leq 0$. This is a contradiction to the suboptimality of stopping in state $\left(\bar{W},\bar{w}\right)$. ☐

An interesting consequence of Lemma 3.8 is the following. Suppose that the myopic gain for testing coefficient $i$ is positive. If we test coefficient $j$ and the realization of $W_j$ happens to be close to $\mu$, then the myopic gain of testing coefficient $i$ remains positive. That is, we can think about untested coefficients as being a-priori at the mean value of their respective distributions; once we test, the coefficients drift away from the mean values. The closer the realizations are to the mean, the more likely we are to test again. On the other hand, if the realizations are farther away from the mean, then we are less likely to test again. That is, for every state, there exists an interval around the mean value where testing is optimal, as given by the following corollary.

**Corollary 3.12.** *For an LOPT with untested coefficients that have the same mean and each given state* $\left(\bar{W},\bar{w}+v\right)$*, it is either optimal to stop regardless of the value* $v$*, or there exists* $v_1, v_2$*, such that it is optimal to test if and only if* $v_1\leq v\leq v_2$*.*

*Proof.* By Lemma 3.8, the function $\Delta_i\left(\bar{W},\bar{w}+v\right)$ is unimodal in $v$ and maximal in $v=\mu$. Thus if for each $i$ it holds that $\Delta_i\left(\bar{W},\bar{w}+\mu\right)\leq 0$, then by Theorem 3.11 it is not optimal to test for any value of $v$. If on the other hand, there exists an untested coefficient $W_i$ such that $\Delta_i\left(\bar{W},\bar{w}+\mu\right)>0$, then we define $v_1$, and $v_2$ as follows:

$$v_1=\inf\left\{v:\exists j \text{ s.t. } \Delta_j\left(\bar{W},\bar{w}+v\right)>0\right\},$$

and,

$$v_2=\sup\left\{v:\exists j \text{ s.t. } \Delta_j\left(\bar{W},\bar{w}+v\right)>0\right\}.$$

By Theorem 3.11, it is optimal to test, if and only if, $v_1\leq v\leq v_2$. ☐

The monotonic decrease in the myopic gain when we test also implies that if at a given state stopping is optimal, and we still decide to test, it will always be optimal to stop in the next state, for any choice of coefficient to be tested regardless of its realization. As we shall see, this will be useful in later sections when analyzing certain suboptimal policies.

**Corollary 3.13.** *For an LOPT with untested coefficients that have the same mean, if stopping at state* $\left(\bar{W}+W_i,\bar{w}\right)$ *is optimal, then stopping is also optimal at state* $\left(\bar{W},\bar{w}+v\right)$*, where* $v$ *is a realization of* $W_i$*.*

*Proof.* Immediate by Corollary 3.9 and Theorem 3.11. ☐

We completely characterized the decision to stop testing for LOPTs with identical mean value of untested coefficients. In the next two sections, we address the issue of deciding which coefficient to test, when stopping is not optimal.

### 3.5 A Simple Optimal Policy for MST with Testing

We now focus on a special LOPT, *Maximum Spanning Tree with Testing*, or *MST with Testing* for short. The formulation of the MST problem as an LOP is described in Section 3.2. Theorem 3.11 implies that when the cost coefficients (edge weights in MST) have the same mean, the myopic stopping rule is optimal. In this section, we show that when the edge weights are also identically distributed, then a myopic policy that consists of a myopic stopping rule and a myopic testing-order rule, i.e., the order in which edges are tested is determined myopically, achieves optimal expected profit. It is a well known fact that finding a minimum spanning tree can be done by negating all edge weights and finding a maximum spanning tree, so we will only address the maximization problem.

Formally, let $G = (V, E)$ be a connected graph with vertices $V = \{v_1, ..., v_n\}$ and edges $E = \{e_1, ..., e_m\}$. The edge weights denoted $W_{e_i}$, or $W_i$ for short, are drawn i.i.d. from a known distribution $W_i \sim W$, with a finite mean $\mathbb{E}[W_i] = \mu < \infty$. The cost of testing an edge is $c > 0$. The goal is to obtain an optimal policy that finds a spanning tree of the graph $G$, such that it maximizes the expected profit, where profit in this case is the weight of the selected spanning tree minus testing costs.

**Remark 3.14** (Tie breaking)**.** *Throughout this section, we will use Kruskal's algorithm to construct MSTs. Kruskal's algorithm orders the edges according to their weights in non-increasing order and adds them to the MST if they do not close a cycle with existing tree edges. We will assume that ties between edge weights are broken in a consistent manner, i.e., if there are several possible MSTs, the algorithm will always return the same MST for the same weights. Furthermore, for simplicity, we will assume that if an edge weight changes, such that it is tied with an existing edge weight, the tie will be broken, such that the new value will be ordered after the old value in Kruskal's algorithm.*

When a policy stops testing, it optimizes based on tested edge weights and expectations of untested edge weights $(\bar{w}, \mathbb{E}[\bar{W}])$, so in fact the MST is computed on a deterministic graph, i.e., a graph with deterministic edge weights. Therefore, we start by considering a deterministic graph and defining the important notion of a substituting edge, that will be useful for following proofs. Intuitively, the substituting edge is the best candidate for swapping with $e_i$ into (or out of) the tree after testing it.

**Definition 3.15.** *Let $T$ be a maximum spanning tree constructed by Kruskal's algorithm in a deterministic graph $H = (V, E)$. For each edge $e_i \in E$, the unique* substituting edge $e_{sub(i)}$, *and the unique cycle* $\text{Cycle}_i$ *that contains both $e_i$ and $e_{sub(i)}$, are defined as follows:*

- *If the edge $e_i \notin T$, then* $\text{Cycle}_i$ *is the unique cycle that is created by adding $e_i$ to the tree. The substituting edge $e_{sub(i)}$ is defined as the tree edge with smallest weight in* $\text{Cycle}_i \setminus \{e_i\}$, *i.e., appears latest in the ordering of Kruskal's algorithm.*

- *If the edge $e_i \in T$, then consider all the non-tree edges that close with the tree a cycle that contains $e_i$. The substituting edge $e_{sub(i)}$ is the edge that has maximal weight among them, i.e., appears first in the ordering of Kruskal's algorithm. Denote by* $\text{Cycle}_i$ *the unique cycle that is created by adding $e_{sub(i)}$ to the tree.*

We can now prove a lemma on the effect of changing one edge weight in a deterministic graph.

**Lemma 3.16.** *Let $T$ be a maximum spanning tree constructed by Kruskal's algorithm in a deterministic graph $H$. Consider changing the weight of edge $e_i$ from $w_i$ to $w'_i$ and let the resulting MST be $T'$. Then the tree $T'$ is one out of three possible trees, the same tree $T' = T$, or $T' = T - e_i + e_{sub(i)}$, or $T' = T + e_i - e_{sub(i)}$.*

*Proof.* We use the following well-known "Cycle Properties" of the MST:

1. For an edge $e_l \notin T$, and for every edge $e_k \in \text{Cycle}_l$, it holds that $w_k \geq w_l$.

2. For every cycle $C$ in $H$, if there exists $e_l \in C$ such that $w_k > w_l$ for every other edge $e_k \in C$, then edge $e_l \notin T$. Notice that by our tie breaking assumption, there is a unique such $e_l$ even if the former strong inequality is replaced by its weak version $w_k \geq w_l$.

Consider first the case $e_i \notin T$. If $w'_i \leq \min\{w_k : e_k \in \text{Cycle}_i\}$, then by the first cycle property also $e_i \notin T'$. Since $e_i \notin T$ and $e_i \notin T'$, then the only cycle $e_i$ can close with $T'$ is $\text{Cycle}_i$. Hence, the two executions of Kruskal's algorithm make the same decisions (also) for all other edges and therefore $T' = T$. Otherwise, $w'_i > w_{sub(i)}$. Hence, $e_i$ is before $e_{sub(i)}$ in the ordering and Kruskal's algorithm will place $e_i \in T'$ and $e_{sub(i)} \notin T'$. Notice that no other edge can be affected by this change because there can only be one cycle formed by adding $e_i$ to the tree $T$. Therefore $T' = T + e_i - e_{sub(i)}$.

Now consider the case $e_i \in T$. By the first cycle property, $w_k \geq w_{sub(i)}$ for all $e_k \in \text{Cycle}_i$. If $w'_i > w_{sub(i)}$ their relative order in Kruskal's algorithm remains the same, and by the second cycle property $e_{sub(i)} \notin T'$ and $e_i \in T'$. Because $e_{sub(i)}$ is the first in the ordering among all other non-tree edges that form a cycle that contained $e_i$ with $T$, there is no cycle containing $e_i$ in which relative orders switch, hence Kruskal's algorithm will return the same tree $T' = T$. Otherwise, $w'_i \leq w_{sub(i)}$, the relative orders of $e_i$ and $e_{sub(i)}$ switch (and possibly relative orders of $e_i$ and other non-tree edges that form a cycle that contained $e_i$ with $T$ also switch). Recall that $e_{sub(i)}$ is ordered after all the (tree) edges in $\text{Cycle}_i \setminus \{e_i, e_{sub(i)}\}$, and before all (non-tree) edges that form a cycle that contained $e_i$ with $T$. Thus $e_{sub(i)}$ does not form a cycle with edges that are before it in the ordering and are in $T'$, so $e_{sub(i)} \in T'$ and $e_i$ does form a cycle with tree edges since $\text{Cycle}_i \setminus \{e_i\} \subset T'$, so $e_i \notin T'$. Also, no other edge is affected by the change because every other non-tree edge that formed a cycle that contained $e_i$ with $T$, now forms a cycle that contains $e_{sub(i)}$ with $T'$ (by transitivity). Hence $T' = T - e_i + e_{sub(i)}$.

In conclusion, $T$ and $T'$ differ by at most one swap, and this swap can only be of $e_i$ and its substituting edge $e_{sub(i)}$. □

Going forward, we return to the setting of MST with testing, in which some of the edge weights are random. Recall that we denote the random edge weights by $\bar{W}$ and the tested edge weights by $\bar{w}$. Recall also Definitions 3.3-3.5 of policies $\Pi^{\text{stop}}\left(\bar{W}, \bar{w}\right)$, $\Pi_i^{\text{test}}\left(\bar{W}, \bar{w}\right)$, and $\Pi_i^{\text{my}}\left(\bar{W}, \bar{w}\right)$. For each state $\left(\bar{W}, \bar{w}\right)$, if a policy chooses to stop, it returns a MST of the graph $G$, denoted by $T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$, which is constructed by Kruskal's algorithm using expectations to replace random variables. For an untested edge $e_i$, the substituting edge $e_{sub(i)}\left(\bar{W}, \bar{w}\right)$ and $\text{Cycle}_i\left(\bar{W}, \bar{w}\right)$ are defined similarly to $e_{sub(i)}$ and $\text{Cycle}_i$ in Definition 3.15, with regard to the MST $T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$. Finally, recall definition 3.6 of the myopic gain of testing edge $e_i$, denoted by $\Delta_i\left(\bar{W}, \bar{w}\right)$. For simplicity, when the state is clear from the context, we omit it from the notation.

We can now state the main result of this section. We show a myopic policy for MST with testing when edge weights are i.i.d, that achieves optimal expected profit.

**Theorem 3.17.** *If all the edge weights $W_i$ are independent and identically distributed with mean $\mu$, then at each state $\left(\bar{W}, \bar{w}\right)$ the following policy is optimal: test an edge with the highest positive myopic gain $\Delta_i\left(\bar{W}, \bar{w}\right)$, and stop when all myopic gains are non-positive.*

In Section 3.2, the MST with testing problem was formulated as an LOPT and therefore, by Section 3.4, it is optimal to stop testing when all myopic gains are non-positive. It remains to show that a policy that tests in non-increasing order of myopic gains is optimal. We do so by showing that testing an edge can only affect edges that have the same myopic gain value. This implies that if a myopic gain value is positive at some state, it remains positive as long as no other edge with the same myopic value is tested. We then use this to show that testing by descending myopic gains is optimal. First, we prove a couple of useful technical results.

**Lemma 3.18.** *At every state* $\left(\bar{W},\bar{w}\right)$*, testing a yet-untested edge* $e_i$ *can cause at most one difference between* $T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ *and* $T^{\text{stop}}\left(\bar{W}-W_i,\bar{w}+w_i\right)$*, which is swapping between* $e_i$ *and its substituting edge* $e_{sub(i)}\left(\bar{W},\bar{w}\right)$*.*

*Proof.* $T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ and $T^{\text{stop}}\left(\bar{W}-W_i,\bar{w}+w_i\right)$ are constructed using the deterministic values $(\mathbb{E}[\bar{W}],\bar{w})$ and $(\mathbb{E}[\bar{W}-W_i],\bar{w}+w_i)$ respectively. The only difference between these two executions of Kruskal's algorithm is the weight of $e_i$ that changes from $\mathbb{E}[W_i]$ to $w_i$. Hence, by Lemma 3.16, the MST can change by at most one edge, and that change is exactly swapping between $e_i$ and $e_{sub(i)}\left(\bar{W},\bar{w}\right)$. □

We can now use the previous lemma to simplify the expression for the myopic gain of edge $e_i$. For consistency we use the notation $\mathbb{E}\left[W_{sub(i)}\right]$ to denote the weight of edge $e_{sub(i)}$, although is can potentially be a deterministic value $w_{sub(i)}$ if the edge $e_{sub(i)}$ was already tested.

**Corollary 3.19.** *At each state* $\left(\bar{W},\bar{w}\right)$ *and for each untested edge* $e_i$*:*

1. *If* $e_i\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$*, then* $\mathbb{E}\left[W_{sub(i)}\right]\geq\mu$ *and the myopic gain of* $e_i$ *is*

$$\Delta_i\left(\bar{W},\bar{w}\right)=\mathbb{E}\left[\left(W_i-\mathbb{E}\left[W_{sub(i)}\right]\right)^+\right]-c. \tag{34}$$

2. *If* $e_i\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$*, then* $\mathbb{E}\left[W_{sub(i)}\right]\leq\mu$ *and the myopic gain of* $e_i$ *is*

$$\Delta_i\left(\bar{W},\bar{w}\right)=\mathbb{E}\left[\left(\mathbb{E}\left[W_{sub(i)}\right]-W_i\right)^+\right]-c. \tag{35}$$

*Proof.* If $e_i\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$, then $e_{sub(i)}\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ by Definition 3.15. By the cycle properties $\mathbb{E}[W_j]\geq\mathbb{E}[W_i]=\mu$ for all $e_j\in\text{Cycle}_i$. In particular, also $\mathbb{E}\left[W_{sub(i)}\right]\geq\mu$. The myopic gain will be, by Definition 3.6 and Lemma 3.18,

$$\begin{aligned}
\Delta_i\left(\bar{W},\bar{w}\right) &= J_i^{\text{my}}\left(\bar{W},\bar{w}\right)-J^{\text{stop}}\left(\bar{W},\bar{w}\right)\\
&= -c+\mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W}-W_i\right],\bar{w}+W_i\right)\right]-\varphi\left(\mathbb{E}\left[\bar{W}\right],\bar{w}\right)\\
&= \mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W}\right],\bar{w}\right)+W_i-\mathbb{E}\left[W_{sub(i)}\right]|W_i>\mathbb{E}\left[W_{sub(i)}\right]\right]\mathbb{P}\left(W_i>\mathbb{E}\left[W_{sub(i)}\right]\right)\\
&\quad+\mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W}\right],\bar{w}\right)|W_i\leq\mathbb{E}\left[W_{sub(i)}\right]\right]\mathbb{P}\left(W_i\leq\mathbb{E}\left[W_{sub(i)}\right]\right)-\varphi\left(\mathbb{E}\left[\bar{W}\right],\bar{w}\right)-c\\
&= \mathbb{E}_{W_i}\left[W_i-\mathbb{E}\left[W_{sub(i)}\right]|W_i>\mathbb{E}\left[W_{sub(i)}\right]\right]\mathbb{P}\left(W_i>\mathbb{E}\left[W_{sub(i)}\right]\right)-c\\
&= \mathbb{E}\left[\left(W_i-\mathbb{E}\left[W_{sub(i)}\right]\right)^+\right]-c.
\end{aligned}$$

Similarly, if $e_i\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$, then $e_{sub(i)}\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ by Definition 3.15. By the cycle properties $\mathbb{E}\left[W_{sub(i)}\right] \leq \mu$. The myopic gain will be, by Definition 3.6 and Lemma 3.18,

$$\begin{aligned}
\Delta_i\left(\bar{W}, \bar{w}\right) &= J_i^{\text{my}}\left(\bar{W}, \bar{w}\right) - J^{\text{stop}}\left(\bar{W}, \bar{w}\right) \\
&= -c + \mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W} - W_i\right], \bar{w} + W_i\right)\right] - \varphi\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right) \\
&= \mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right) + \mathbb{E}\left[W_{sub(i)}\right] - W_i | W_i < \mathbb{E}\left[W_{sub(i)}\right]\right] \mathbb{P}\left(W_i < \mathbb{E}\left[W_{sub(i)}\right]\right) \\
&\quad + \mathbb{E}_{W_i}\left[\varphi\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right) | W_i \geq \mathbb{E}\left[W_{sub(i)}\right]\right] \mathbb{P}\left(W_i \geq \mathbb{E}\left[W_{sub(i)}\right]\right) - \varphi\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right) - c \\
&= \mathbb{E}_{W_i}\left[\mathbb{E}\left[W_{sub(i)}\right] - W_i | W_i < \mathbb{E}\left[W_{sub(i)}\right]\right] \mathbb{P}\left(W_i < \mathbb{E}\left[W_{sub(i)}\right]\right) - c \\
&= \mathbb{E}\left[\left(\mathbb{E}\left[W_{sub(i)}\right] - W_i\right)^+\right] - c.
\end{aligned}$$

□

We can now show the independence between testing edges with different myopic gains.

**Lemma 3.20.** *Assume that at state $\left(\bar{W}, \bar{w}\right)$ the two untested edges $e_i, e_j$ have different myopic gain values $\Delta_i\left(\bar{W}, \bar{w}\right) \neq \Delta_j\left(\bar{W}, \bar{w}\right)$. Then for each realization $w_j$ of $W_j$, we have $\Delta_i\left(\bar{W} - W_j, \bar{w} + w_j\right) = \Delta_i\left(\bar{W}, \bar{w}\right)$. Similarly, for each realization $w_i$ of $W_i$, we have $\Delta_j\left(\bar{W} - W_i, \bar{w} + w_i\right) = \Delta_j\left(\bar{W}, \bar{w}\right)$.*

*Proof.* Assume without loss of generality that at state $\left(\bar{W}, \bar{w}\right)$ a policy tests $e_j$ and reveals its realization $w_j$. Assume towards contradiction that $\Delta_i\left(\bar{W} - W_j, \bar{w} + w_j\right) \neq \Delta_i\left(\bar{W}, \bar{w}\right)$. Recall that by Corollary 3.19, if $e_i \notin T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$ then $\Delta_i\left(\bar{W}, \bar{w}\right) = \mathbb{E}\left[\left(W_i - \mathbb{E}\left[W_{sub(i)}\left(\bar{W}, \bar{w}\right)\right]\right)^+\right] - c$, and otherwise $\Delta_i\left(\bar{W}, \bar{w}\right) = \mathbb{E}\left[\left(\mathbb{E}\left[W_{sub(i)}\left(\bar{W}, \bar{w}\right)\right] - W_i\right)^+\right] - c$, and that by Lemma 3.18, the only change that can happen to the MST is that $e_j$ and its substituting edge $e_{sub(j)}\left(\bar{W}, \bar{w}\right)$ swap. Therefore, the only cases where testing $e_j$ can affect $\Delta_i$ are:

1. If the substituting edge of $e_i$ does not change, but the value $\mathbb{E}\left[W_{sub(i)}\left(\bar{W}, \bar{w}\right)\right] \neq w_{sub(i)}\left(\bar{W} - W_j, \bar{w} + w_j\right)$, which implies that $e_{sub(i)}\left(\bar{W}, \bar{w}\right) = e_j$;

2. Else if the tree $T^{\text{stop}}$ changes such that if $e_i \notin T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$ then $e_i \in T^{\text{stop}}\left(\bar{W} - W_j, \bar{w} + w_j\right)$ or the other way around, which implies that $e_i = e_{sub(j)}\left(\bar{W}, \bar{w}\right)$; or

3. Otherwise, if the substituting edge of $e_i$ changes such that $e_{sub(i)}\left(\bar{W}, \bar{w}\right) \neq e_{sub(i)}\left(\bar{W} - W_j, \bar{w} + w_j\right)$.

For case 1, assume towards contradiction that $e_j = e_{sub(i)}\left(\bar{W}, \bar{w}\right)$. Then $e_i$ and $e_j$ are on the same cycle, one of them is in the tree and the other is out of the tree. Assume first that $e_i \in T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$, which implies that $e_j \notin T^{\text{stop}}\left(\bar{W}, \bar{w}\right)$. Then by Corollary 3.19, $\Delta_i\left(\bar{W}, \bar{w}\right) = \mathbb{E}\left[\left(\mu - W_i\right)^+\right] - c$. Since $e_j$ is outside the tree with expected weight $\mu$, then every edge in $\text{Cycle}_j \setminus \{e_j\}$ must have weight greater or equal to $\mu$, specifically $\mathbb{E}[W_{sub(j)}] \geq \mu$. On the other hand, the substituting edge is by Definition 3.15 the smallest edge in the $\text{Cycle}_j \setminus \{e_j\}$, which contains $e_i$. Therefore, $\mathbb{E}[W_{sub(j)}] \leq \mu$, and hence $\mathbb{E}[W_{sub(j)}] = \mu$. So the myopic gain of edge $e_j$ is

$\Delta_j\left(\bar{W},\bar{w}\right)=\mathbb{E}\left[(W_j-\mu)^+\right]-c$. Using the fact that $W_i$ and $W_j$ are independent and identically distributed and a simple variation on Equation (4), $\Delta_i\left(\bar{W},\bar{w}\right)=\Delta_j\left(\bar{W},\bar{w}\right)$, in contradiction to our assumption. Now assume that $e_i\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$, which implies that $e_j\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$. Then by Corollary 3.19, $\Delta_i\left(\bar{W},\bar{w}\right)=\mathbb{E}\left[(W_i-\mu)^+\right]-c$. Since $e_j$ is in the tree with expected weight $\mu$, then $e_{sub(j)}\left(\bar{W},\bar{w}\right)\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ and it is the minimal such edge that closes a cycle with $e_j$. By similar reasons as before it must be that $\mathbb{E}[W_{sub(j)}]=\mu$. So the myopic gain of edge $e_j$ is $\Delta_j\left(\bar{W},\bar{w}\right)=\mathbb{E}\left[(\mu-W_j)^+\right]-c$. Similarly, $\Delta_i\left(\bar{W},\bar{w}\right)=\Delta_j\left(\bar{W},\bar{w}\right)$, in contradiction to our assumption.

Case 2 is symmetric to case 1, and thus also leads to contradiction.

For case 3, since $e_j$ is the edge whose weight changes, then either $e_{sub(i)}\left(\bar{W},\bar{w}\right)=e_j$ or $e_{sub(i)}\left(\bar{W}-W_j,\bar{w}+w_j\right)=e_j$. The former case was covered by case 1, so it remains to prove the latter. Assume towards contradiction that $e_{sub(i)}\left(\bar{W}-W_j,\bar{w}+w_j\right)=e_j$. First consider the case $e_i\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$. Edge $e_{sub(i)}\left(\bar{W},\bar{w}\right)\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ is the minimal edge in $\text{Cycle}_i\left(\bar{W},\bar{w}\right)\setminus\{e_i\}$, so $e_i$ can only replace it as substituting edge if $w_j<\mathbb{E}[W_{sub(i)}\left(\bar{W},\bar{w}\right)]$. However, for $e_j$ to enter the tree instead of $e_{sub(j)}\left(\bar{W},\bar{w}\right)$, then $\mathbb{E}\left[W_{sub(j)}\left(\bar{W},\bar{w}\right)\right]<w_j$. Together we get $\mathbb{E}\left[W_{sub(j)}\left(\bar{W},\bar{w}\right)\right]<\mathbb{E}[W_{sub(i)}\left(\bar{W},\bar{w}\right)]$, in contradiction to the minimality of $e_{sub(i)}\left(\bar{W},\bar{w}\right)$. Now consider the case $e_i\in T^{\text{stop}}\left(\bar{W},\bar{w}\right)$. In this case, assume towards contradiction that $e_{sub(i)}\left(\bar{W}-W_j,\bar{w}+w_j\right)=e_j$. This implies that adding $e_j$ to the tree $T^{\text{stop}}\left(\bar{W}-W_j,\bar{w}+w_j\right)$ will close a cycle with $e_i$. Since $\mathbb{E}[W_{sub(i)}\left(\bar{W},\bar{w}\right)]\neq\mathbb{E}\left[W_{sub(j)}\left(\bar{W},\bar{w}\right)\right]$, then the edges $e_{sub(j)}\left(\bar{W},\bar{w}\right)\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ and $e_{sub(i)}\left(\bar{W},\bar{w}\right)\notin T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ cannot form the same cycle with the tree (if they did then only the heavier of them would have been the substituting edge for both $e_j$ and $e_i$). However, this implies that the tree $T^{\text{stop}}\left(\bar{W},\bar{w}\right)$ had a cycle containing $e_j$ and $e_i$, in contraction to it being a tree.

In conclusion, testing $e_j$ cannot have any affect on $\Delta_i$.

□

**Proof of Theorem 3.17.** Proceed by induction on $N$, the number of untested edges. For $N=1$, the base case, the myopic policy is trivially an optimal policy. For the inductive step, consider $N>1$ untested edges. Assume by the inductive hypothesis that for any subset of $\bar{W}$ of size less or equal to $N-1$, an optimal policy tests according to descending myopic gains. We shall show that it is also optimal for $N$ untested edges. Order the untested edges such that $\Delta_1\left(\bar{W},\bar{w}\right)\geq\Delta_2\left(\bar{W},\bar{w}\right)\geq\ldots\geq\Delta_N\left(\bar{W},\bar{w}\right)$. Denote by $\Pi^{\text{opt}}$ an optimal policy. Clearly if $\Pi^{\text{opt}}$ stops then our policy does the same by Theorem 3.11 and the proof follows. Also, if there exists a single edge with positive myopic gain, then this edge is clearly $e_1$ and $\Pi^{\text{opt}}$ tests it, so the optimal policy tests in descending order of non-negative myopic gains (because it is the only order for one edge), which completes the proof. Thus, assume there is more than one edge with positive myopic gain. Let $e_i$ be the first edge that $\Pi^{\text{opt}}$ tests. Clearly, if $\Pi^{\text{opt}}$ tests $e_i$ such that $\Delta_i\left(\bar{W}, \bar{w}\right) = \Delta_1\left(\bar{W}, \bar{w}\right)$, then by the induction hypothesis, the rest of the testing order of the optimal policy is in descending order of myopic gains and the proof follows. We remain with the case that $\Pi^{\text{opt}}$ tests some edge $e_i$ such that $\Delta_1\left(\bar{W}, \bar{w}\right) > \Delta_i\left(\bar{W}, \bar{w}\right) \geq 0$. In policy $\Pi^{\text{opt}}$, testing $e_i$ will possibly lead to some change in the optimal tree, but by Lemma 3.20 and Corollary 3.9, no matter what this change is, the myopic gain $\Delta_1$ will remain the highest non-negative myopic gain. Hence, an edge obtaining a myopic gain that equals $\Delta_1$ has to be tested next by the induction hypothesis. Without loss of generality, assume it is $e_1$. Denote by $\Pi^{\text{my}}$ a myopic policy that tests according to descending myopic gains, and tests $e_1$ first. Assume towards contradiction that $\Pi^{\text{my}}$ is strictly suboptimal, i.e., $J^{\text{opt}}\left(\bar{W}, \bar{w}\right) > J^{\text{my}}\left(\bar{W}, \bar{w}\right)$.

We define a suboptimal policy $\Pi^{\text{sub}}$ that tests $e_1$ first, $e_i$ second, and then mimics the optimal policy. Clearly, $J^{\text{my}}\left(\bar{W}, \bar{w}\right) \geq J^{\text{sub}}\left(\bar{W}, \bar{w}\right)$ because they perform the same first test and, according to the induction hypothesis, $\Pi^{\text{my}}$ performs optimally after the first test.

By Lemma 3.20, testing $e_i$ does not influence $\Delta_1$ and vice versa. Since both myopic gains are non-negative, both policies $\Pi^{\text{sub}}$ and $\Pi^{\text{opt}}$ always test both edges $e_1$ and $e_i$. After conducting both tests the sample path will be the same for both policies by the induction hypothesis. Hence, $J^{\text{sub}}\left(\bar{W}, \bar{w}\right) = J^{\text{opt}}\left(\bar{W}, \bar{w}\right)$, which concludes the proof. □

## 3.6 An Optimal Testing Rule for Symmetric LOPTs

In this section, we study another class of LOPTs for which a myopic policy, i.e., a policy composed of both a myopic stopping rule and a myopic testing-order rule, obtains optimal expected profit. Specifically, we study LOPT problems that are *symmetric*, i.e., where the polymatroid $P_f$ is invariant to permutations of the decision variables, and in which there is a partial order between the untested coefficients known as a *convex order*. This case is different from the MST problem which in general is a-symmetric (as will be shown below), and more general than the MST in the sense that the objective coefficients are in convex order and need not be identically distributed.

We start in Section 3.6.1 by defining symmetric problems, and discuss the symmetry of the selection problem and the a-symmetry of the MST problem. In Section 3.6.2 we define the convex order, and present some of its properties. Finally, in Section 3.6.3 we define the myopic policy and show that it is optimal for the class of symmetric LOPTs where untested coefficients satisfy convex order.

### 3.6.1 Symmetric LOPTs

Intuitively, we say that an optimization problem is symmetric if the only defining characteristic of an unknown parameter is its value. For example, the selection problem presented in Section 2 is symmetric, as we can permute the value of parameters and obtain an equivalent problem. In contrast, the MST problem is not symmetric, because every parameter is associated with an edge on a graph, and permuting the parameters values can result in a considerably different optimal value. Before proving these statements, we start with a definition of symmetric LOPTs:

**Definition 3.21.** *An LOPT is symmetric if the associated polymatroid $P_f$ is invariant to permutations of the decision variables $x_1, \ldots, x_N$.*

This definition leads to the following symmetry property of LOPTs.

**Corollary 3.22.** *An LOPT is symmetric if the associated function $\varphi(w)$ of equation (19) is a symmetric function for all $w$.*

*Proof.* By definition 3.21, the constraint polytope $P_f$ of $\varphi(w)$ is invariant to permutations of the decision variables and by equation (19) the objective is linear, therefore swapping two coefficients does not change the optimization. □

The next lemma characterizes an interesting family of symmetric problems. It includes the Selection problem discussed in Section 2.

**Lemma 3.23.** *Every LOPT that is defined by a submodular function $f(S)$ that only depends on the cardinality of the set $S$ defines a symmetric LOPT problem.*

*Proof.* Let $P$ denote an LOP with objective coefficients $\bar{w}$, and let $P_r$ denote an identical LOP with a permutation $\bar{w}_r$ over the coefficients. We need to show that for every permutation $\bar{w}_r$ of the coefficients $\bar{w}$, the resulting value of the optimal solution for the two LOPs is identical, that is $\varphi(\bar{w}) = \varphi(\bar{w}_r)$. We do this in two steps:

1. Show that the feasible region of the two problems is identical, and that every permutation $\bar{x}_r$ of a solution $\bar{x}$ is also feasible;

2. Show that for every solution $\bar{x}$ to $P$, there exists a permutation $\bar{x}_r$ that achieves the same objective value for the problem $P_r$.

To see (1), we look at all the constraints associated with some set cardinality $l$, which can be written as:

$$\forall U \subseteq S \text{ s.t. } |U| = l : \sum_{i \in U} x_i \leq f(l).$$

These constraints are symmetric with respect to the order of the permutation, for every value of $l$. Therefore for every solution $\bar{x}$, and a permutation $r$, the solution $\bar{x}_r$ is also feasible.

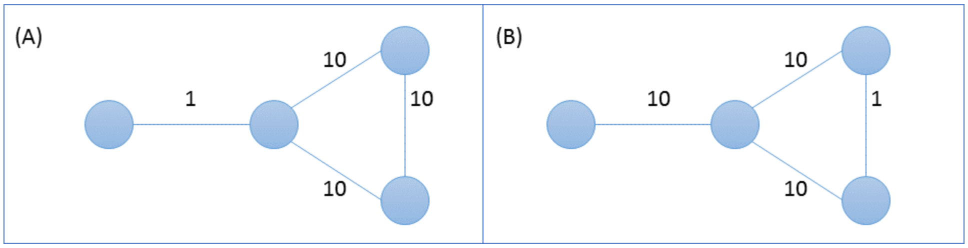


Figure 3.4: A non-symmetric MST.

To see why (2) holds, observe that for every permutation $r$, the following holds:

$$\bar{w}_r^T \bar{x}_r = \bar{w}^T \bar{x}.$$

□

**Corollary 3.24.** *The K-Selection problem is a symmetric LOPT.*

**Lemma 3.25.** *There exists an MST that is a non-symmetric LOPT.*

*Proof.* Figure 3.4 illustrates a non-symmetric MST. We see that for the same graph (an LOP instance) there are different optimal solution values for different permutations of edge weights (objective coefficients). In Figure 3.4A the value of the MST is equal to 21, whereas in Figure 3.4B the value of the MST is equal to 30. □

### 3.6.2 The Convex Order

We now review the definition and main properties of the convex order. In particular, we describe the mean preserving local spread and show its relation to the convex order. These concepts and properties are key to the analysis in Section 3.6.3.

The convex order is a mathematical relation between probability distributions, that is often used to model difference in risk profiles (see Müller and Stoyan (2002)). Intuitively, distributions that are higher in convex order are more variable or spread around their mean values. Formally, we say that $X \leq_{cx} Y$ (X is smaller in convex order than Y), if the expected value of any convex function $u$ on the two random variables results in a lower value for $X$, that is

$$X \leq_{cx} Y \Leftrightarrow \mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)], \text{for any convex function } u \text{ for which expectations exist.}$$

Figure 3.5 illustrates the convex order in discrete and continuous distributions. In the left side of Figure 3.5 is an example of the convex order relation in discrete distributions. In the figure are three discrete distributions that have the same mean value, but are different otherwise. In particular,

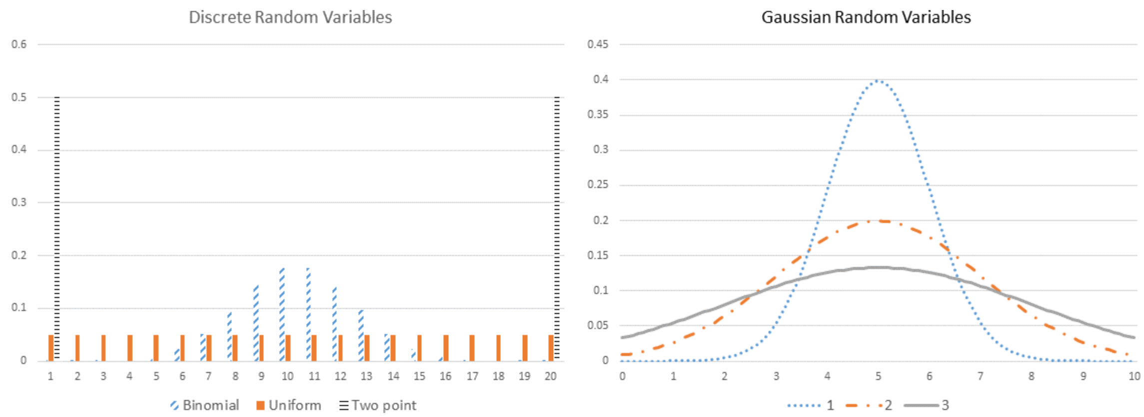


Figure 3.5: Examples of discrete (left) and continuous (right) distributions in convex order.

the two-point distribution is higher in convex order than the uniform distribution, which is higher in convex order than the binomial distribution. The right side of Figure 3.5 depicts three normally distributed random variables with equal means and different variance. These distributions are monotonically increasing in the convex order where a higher variance corresponds to a higher convex order (e.g., the distribution denoted by 3 is higher both in variance and convex order than the distributions denoted by 1 and 2). This is not a coincidence as variance is a convex function, hence convex order implies monotonicity in variance.

### The Mean Preserving Local Spread

Perhaps the simplest type of convex order is the mean preserving local spread as given by the following definition:

**Definition 3.26.** *(Müller and Stoyan (2002)) Let $F$ and $G$ be distribution functions of discrete distribution whose common support is a finite set of points $x_1 < x_2 < ... < x_n$ with probability mass function $f$ and $g$ respectively. Then $G$ is said to differ from $F$ by a local spread, if there exists some $i \in 2, 3, ..., n-1$ such that $0 = g(x_i) \leq f(x_i), g(x_{i-1}) \geq f(x_{i-1}), g(x_{i+1}) \geq f(x_{i+1})$, and $g(x_j) = f(x_j)$ for all $j \notin i-1, i, i+1$. A local spread is said to be mean preserving if $F$ and $G$ have the same mean. Write $F \leq_{LS} G$ if $G$ is a mean preserving local spread of $F$.*

Intuitively, if distribution $G$ can be obtained from a distribution $F$ by shifting some of the probability mass in one point of the support (which we call the *focal point*) to two adjacent points while preserving the mean, then $G$ is a mean preserving local spread of $F$. Figure 3.6 illustrates two distributions $F$ and $G$ defined over the support $\{1, 2, ..., 7\}$. Observe that the conditions for the local spread are satisfied:

- for $i \in 1, 5, 6, 7 : g_i = f_i$

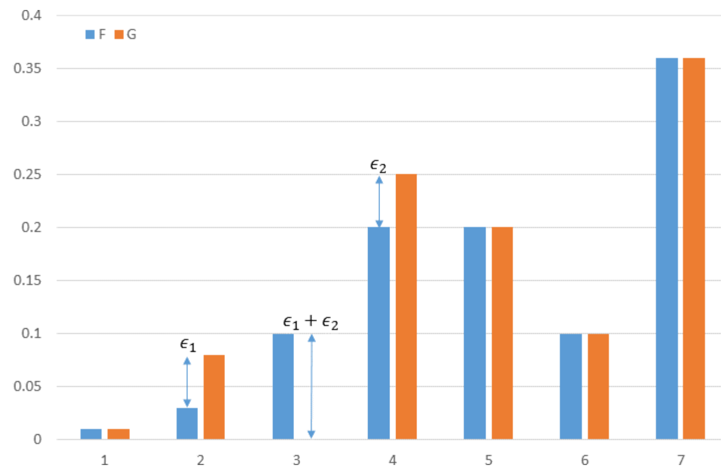


Figure 3.6: An example of the mean preserving local spread.

- $g_2 = f_2 + \epsilon_1$
- $g_3 = 0$, $f_3 = \epsilon_1 + \epsilon_2$ ($i = 3$ is the focal point)
- $g_4 = f_4 + \epsilon_2$

We argue that comparing random variables that are ordered in local spread is significantly simpler than analyzing arbitrary distributions that are in convex order. For example, comparing the expected values of some function of two random variables that are in local spread is easier than comparing the expected values of arbitrary distributions, because many of the terms cancel out.

It is easy to show that local spread implies a convex order between distributions. Interestingly, the reverse direction holds as well. That is, if two distributions are in convex order, then there exists a series of distributions leading from one distribution to another, where every two consecutive distributions are in local spread. This will be critical to the analysis where we essentially reduce the model with general convex order relations to a model where there is only one pair of distributions that are in local spread. Formally:

**Theorem 3.27.** *(Müller and Stoyan (2002)) Let $F$ and $G$ be distribution functions of discrete distribution with finite support. Then $F \leq_{cx} G$ holds if and only if there is a finite sequence $F_1, ..., F_k$ with $F_1 = F$ and $F_k = G$, such that $F_i \leq_{LS} F_{i+1}$ for $i = 1, ..., k-1$.*

### 3.6.3 A Simple Optimal Policy for Symmetric LOPTs under Convex Order

Similarly to Section 3.5, we define the myopic policy as the policy that stops when all myopic gains are non-positive, and otherwise tests the coefficient that achieves the highest myopic gain. We prove that this policy is in fact optimal in determining not only in deciding when to stop testing (Theorem 3.11), but also and in deciding on which coefficient to test.

Formally, let $i^*(\bar{W}, \bar{w})$ denote the index of the untested coefficient that obtains the highest myopic gain at state $(\bar{W}, \bar{w})$ (in case of a tie we choose the highest index):

$$i^*(\bar{W}, \bar{w}) = argmax_{i \in [K]} \Delta_i(\bar{W}, \bar{w}).$$

We define the myopic policy as follows:

**Definition 3.28.** *At any state $(\bar{W}, \bar{w})$ the myopic policy tests coefficient $i^*(\bar{W}, \bar{w})$ if $\Delta_i^* > 0$, and otherwise stops.*

Observe that the myopic policy stops testing according to the myopic stopping rule (Definition 3.10), which means that the decision to stop testing is optimal for any LOPT with coefficients that have the same mean value (Theorem 3.11).

Another interesting property of the myopic policy is that for symmetric LOPTs, when the untested coefficients are in convex order, the myopic policy may only choose to test the coefficient of the highest order:

**Lemma 3.29.** *For symmetric LOPTs with untested coefficients that are in convex order $W_1 \leq_{cx} W_2 ... \leq_{cx} W_k$, testing coefficient $W_k$ obtains the highest myopic gain:*

$$k = argmax_{i \in [K]} \Delta_i(\bar{W}, \bar{w}), \ \textit{for all states} \ (\bar{W}, \bar{w}).$$

*Proof.* Using Definition 3.6:

$$\Delta_k \left(\bar{W}, \bar{w}\right) = J_k^{\text{my}} \left(\bar{W}, \bar{w}\right) - J^{\text{stop}} \left(\bar{W}, \bar{w}\right) \geq J_i^{\text{my}} \left(\bar{W}, \bar{w}\right) - J^{\text{stop}} \left(\bar{W}, \bar{w}\right) = \Delta_i \left(\bar{W}, \bar{w}\right)$$

where the inequality follows from the definition of the convex order and the fact that $J_i^{\text{my}} \left(\bar{W}, \bar{w}\right)$ can be written as the expectation of the following expression:

$$-c + \varphi \left(\mathbb{E}\left[\bar{W} - W_i\right], \bar{w} + W_i\right),$$

which is convex function of $W_i$ (Lemma 3.2). □

We are now ready to present the main result of this section, which is that for symmetric LOPTs with coefficients that are in convex order, the myopic policy is optimal. Intuitively, this tells us that when the only distinguishing factor between two untested parameters is their distributions (as is the case in symmetric problems), we should favor testing the coefficient that is “more variable” than others (e.g., the parameter that is highest in convex order). This is the parameter that gives us most information, and is more likely to improve the optimization outcome. As an extreme example, consider the constant $\mu$ which is a trivial distribution that is smaller in convex order than any other distribution with mean $\mu$. Testing this distribution benefits less than testing any other distribution (and in fact, it is never optimal to test it). Note that in the absence of symmetry (such as in MSTs), the specific structure of the problem must also be taken into account when deciding on which parameter to test, in addition to uncertainty reduction.

**Theorem 3.30.** *For symmetric LOPTs with untested coefficients that are in convex order $W_1 \leq_{cx} W_2 ... \leq_{cx} W_k$, the myopic policy is optimal.*

*Proof.* We prove the theorem by induction on $k$. When $k = 1$ there is a single yet-untested coefficient and the theorem trivially holds. For the case $k > 1$, recall that the myopic gain from testing coefficient $k$ is the highest (Lemma 3.29). This implies that if $\Delta_k \leq 0$, then for every $j$ the myopic gain is non-positive ($\Delta_j \leq 0$). According to Theorem 3.11 stopping is optimal, in which case, the myopic policy is also optimal. We therefore only need to consider the case where $\Delta_k > 0$, in which testing is optimal. We show that there exists an optimal policy that tests coefficient $k$.

Assume by contradiction that at some state testing coefficient $k$ is not optimal, and that there exists an optimal policy that tests coefficient $j$. Let $\left(\bar{W}' + W_j + W_k, \bar{w}\right)$ denote any state that has two or more untested coefficients, where $\bar{W}'$ is the set of untested coefficients that are neither $W_k$ nor $W_j$ ($\bar{W}'$ could be the empty set). Moreover, let $\pi_j$ denote the optimal policy that starts by testing coefficient $j$. We construct a policy $\pi_k$ which starts by testing coefficient $k$, and show that the value function under policy $\pi_k$ is equal to or higher than the value function under policy $\pi_j$. That is, we show that the following holds:

$$J_{\pi_j}\left(\bar{W}' + W_j + W_k, \bar{w}\right) - J_{\pi_k}\left(\bar{W}' + W_j + W_k, \bar{w}\right) \leq 0. \tag{36}$$

Before proving Equation (36), we briefly outline the remainder of the proof:

1. We define a sub-optimal policy denoted as policy Y.
2. We reformulate Equation (36) by adding and subtracting the term $\mathbb{E}J_Y\left(\bar{W}' + W_k, \bar{w} + W_k\right)$ (where $J_Y$ denotes the value function under policy Y). This simplifies the analysis by allowing us to compare expressions that are more similar to each other. The new expression is given by Equation (37).
3. We then use the mean preserving local spread (Theorem 3.27) to create an upper bound for Equation (37) using a sum of a telescopic series. Proving that every term in the telescopic series is non-positive is a sufficient condition for the proof to hold. This essentially reduces the problem from arbitrary convex orders to a mean preserving local spread. This term is given by Equation (39).

4. We explicitly write Equation (39). Since this term rather complex, we divide it to two parts denoted by 'L' and 'R'. First we derive the term 'L', then derive the term 'R', and then sum the two. We present this summation using a function we denote by $\phi$, and argue that in order to complete the proof it is sufficient to show that the function $\phi$ is convex.

5. We establish the convexity of the function $\phi$.

**Step 1: Policy "Y".**

We start by defining policy Y as follows:

**Definition 3.31.** *At any state* $\left(\bar{W}' + Y, \bar{w}\right)$ *(with* $Y$ *denoting any untested coefficient, such as* $W_j$*,* $W_k$*, or possibly a different random variable), policy* $Y$ *tests coefficient* $Y$ *at state* $\left(\bar{W}' + Y, \bar{w}\right)$*, if the optimal policy tests at state* $\left(\bar{W}' + W_k, \bar{w}\right)$*, and otherwise stops. If policy* $Y$ *tests, it later imitates the optimal policy.*

We make the following observations about policy Y:

1. Policy Y is defined with respect to the untested coefficients in $\bar{W}'$.

2. Using the induction hypothesis, at state $\left(\bar{W}' + W_k, \bar{w}\right)$ the optimal policy may test only $W_k$ as it highest in convex order.

3. Under policy Y the same control is selected at states: $\left(\bar{W}' + Y_1, \bar{w}\right)$ and $\left(\bar{W}' + Y_2, \bar{w}\right)$ (where $Y_1$ and $Y_2$ correspond to two different random variables). At both states the policy either stops, or tests coefficients $Y_1$ and $Y_2$, respectively. This means that when testing, the two states transition to states $\left(\bar{W}', \bar{w} + Y_1\right)$ and $\left(\bar{W}', \bar{w} + Y_2\right)$, respectively.

4. At state $\left(\bar{W}' + W_k, \bar{w}\right)$ policy Y is in fact the optimal policy.

We next use policy Y to reformulate the sufficient condition for the optimality of policy $\pi_k$ given by Equation (36).

**Step 2: Reformulating Equation (36).**

Using the DP formulation of Equation (26), we can write Equation (36) as follows:

$$\begin{aligned}
& J_{\pi_j}\left(\bar{W}'+W_j+W_k,\bar{w}\right)-J_{\pi_k}\left(\bar{W}'+W_j+W_k,\bar{w}\right)\\
= & -c+\mathbb{E}J\left(\bar{W}'+W_k,\bar{w}+W_j\right)-\left(-c+\mathbb{E}J\left(\bar{W}'+W_j,\bar{w}+W_k\right)\right)\\
= & \mathbb{E}J\left(\bar{W}'+W_k,\bar{w}+W_j\right)-\mathbb{E}J\left(\bar{W}'+W_j,\bar{w}+W_k\right)\\
= & \mathbb{E}J\left(\bar{W}'+W_k,\bar{w}+W_j\right)-\mathbb{E}J\left(\bar{W}'+W_j,\bar{w}+W_k\right)\\
& -\left(\mathbb{E}J_Y\left(\bar{W}'+W_k,\bar{w}+W_k\right)-\mathbb{E}J_Y\left(\bar{W}'+W_k,\bar{w}+W_k\right)\right)\\
= & \mathbb{E}J\left(\bar{W}'+W_k,\bar{w}+W_j\right)-\mathbb{E}J_Y\left(\bar{W}'+W_k,\bar{w}+W_k\right)\\
& +\mathbb{E}J_Y\left(\bar{W}'+W_k,\bar{w}+W_k\right)-\mathbb{E}J\left(\bar{W}'+W_j,\bar{w}+W_k\right), \qquad (37)
\end{aligned}$$

where the first equality follow directly from Equation (26), and in the second equality we cancel the common term $c$. In the third equality we introduce two terms that sum to zero, where the subscript $Y$ denotes that the value function is computed under policy Y. By rearranging the terms we obtain the last equality.

Observe that unlike Equation (36) in which we subtract the value function of two states that are different in both untested and tested coefficients, in Equation (37) we subtract terms that share either the same tested or untested coefficients (corresponding to the lower and upper rows in Equation (37)).

**Step 3: Bounding Equation (37) by a sum of a telescopic series.**

Using Theorem 3.27 and the convex order $W_j \leq_{cx} W_k$, there exists a series of random variables $Y_1, ..., Y_m$ such that: (1) $Y_1 \leq_{LS} ... \leq_{LS} Y_m$; (2) $Y_1 \overset{d}{=} W_j$; and (3) $Y_m \overset{d}{=} W_k$. We can then express

Equation (37) using the series $Y_1, ..., Y_m$:

$$\begin{aligned}
& \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + W_j\right) - \mathbb{E}J_Y\left(\bar{W}' + W_k, \bar{w} + W_k\right) \\
& + \mathbb{E}J_Y\left(\bar{W}' + W_k, \bar{w} + W_k\right) - \mathbb{E}J\left(\bar{W}' + W_j, \bar{w} + W_k\right) \\
= \; & \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_1\right) - \mathbb{E}J_Y\left(\bar{W}' + W_k, \bar{w} + Y_m\right) \\
& + \mathbb{E}J_Y\left(\bar{W}' + Y_m, \bar{w} + W_k\right) - \mathbb{E}J\left(\bar{W}' + Y_1, \bar{w} + W_k\right) \\
= \; & \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_1\right) - \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_m\right) \\
& + \mathbb{E}J_Y\left(\bar{W}' + Y_m, \bar{w} + W_k\right) - \mathbb{E}J\left(\bar{W}' + Y_1, \bar{w} + W_k\right) \\
\leq \; & \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_1\right) - \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_m\right) \\
& + \mathbb{E}J_Y\left(\bar{W}' + Y_m, \bar{w} + W_k\right) - \mathbb{E}J_Y\left(\bar{W}' + Y_1, \bar{w} + W_k\right) \\
= \; & \sum_{l=1}^{m-1}\left(\mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_l\right) - \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_{l+1}\right)\right) \\
& + \sum_{l=1}^{m-1}\left(\mathbb{E}J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + W_k\right) - \mathbb{E}J_Y\left(\bar{W}' + Y_l, \bar{w} + W_k\right)\right), \qquad (38)
\end{aligned}$$

where in the first equality we substitute $W_j$ and $W_k$ with $Y_1$ and $Y_m$, respectively. In the second equality, we remove the subscript Y because policy Y and $\pi_j$ coincide (by definition of policy Y). In the first inequality, we use the fact that policy Y is sub-optimal, which we then rewrite as a sum of a telescopic series to obtain Equation (38).

Therefore, a sufficient condition for equations 36 and (37) to hold is that every term in the telescopic series given by Equation 38 is non-positive. That is, for every $W_j \leq_{cx} Y_l \leq_{LS} Y_{l+1} \leq_{cx} W_k$ the following inequality holds:

$$\begin{aligned}
& \underbrace{\mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_l\right) - \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_{l+1}\right)}_{L} \\
+ & \underbrace{\mathbb{E}J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + W_k\right) - \mathbb{E}J_Y\left(\bar{W}' + Y_l, \bar{w} + W_k\right)}_{R} \leq 0, \qquad (39)
\end{aligned}$$

where L denotes the summation in the upper row, and R denotes the summation in the lower row.

**Step 4: Using local spread to simplify Equation (39).**

We now delve into Equation (39) and explicitly write the expectations using probability distributions. We use the fact that the random variables $Y_l$ and $Y_{l+1}$ are in local spread, which means that their distributions is identical in all but three points in the support of their distributions. Specifically, by Definition 3.26, there exists a support $y_1 < y_2 < ... < y_z$, a focal point $s$ $(1 < s < z)$, and probabilities $\epsilon_1$ and $\epsilon_2$, which satisfy:

- $Prob(Y_{l+1} = y_{s-1}) = Prob(Y_l = y_{s-1}) + \epsilon_1$
- $Prob(Y_{l+1} = y_{s+1}) = Prob(Y_l = y_{s+1}) + \epsilon_2$
- $Prob(Y_{l+1} = y_s) = Prob(Y_l = y_s) - \epsilon_1 - \epsilon_2$=0
- $Prob(Y_{l+1} = y_i) = Prob(Y_l = y_i)$, for $i \notin \{s-1, s, s+1\}$
- $\epsilon_1(y_s - y_{s-1}) = \epsilon_2(y_{s+1} - y_s)$, or equivalently:

$$y_s = \frac{\epsilon_1}{(\epsilon_1 + \epsilon_2)} y_{s-1} + \frac{\epsilon_2}{(\epsilon_1 + \epsilon_2)} y_{s+1}, \tag{40}$$

The reader is referred to Figure 3.6 for an intuitive illustration. In addition, we use $v_m$ to denote values from the support of the random variables $W_k$.

We can now express the expressions $L$ and $R$ using the support and probabilities of the distributions $Y_l$ and $Y_{l+1}$.

**Step 4.1: The term L.**

The term $L$ represents the difference in the expected value function of two states that have the same untested coefficients and all but one different tested coefficient. Using the definition of local spread, we can write the difference in expectations as follows:

$$\begin{aligned}
L &= \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_l\right) - \mathbb{E}J\left(\bar{W}' + W_k, \bar{w} + Y_{l+1}\right) \\
&= \sum_{m=1}^{z} Prob(Y_l = y_m) J\left(\bar{W}' + W_k, \bar{w} + y_m\right) \\
&\quad - \sum_{m=1}^{z} Prob(Y_{l+1} = y_m) J\left(\bar{W}' + W_k, \bar{w} + y_m\right) \\
&= \sum_{m=1}^{z} \left(Prob(Y_l = y_m) - Prob(Y_{l+1} = y_m)\right) J\left(\bar{W}' + W_k, \bar{w} + y_m\right) \\
&= -\epsilon_1 J\left(\bar{W}' + W_k, \bar{w} + y_{s-1}\right) \\
&\quad + (\epsilon_1 + \epsilon_2) J\left(\bar{W}' + W_k, \bar{w} + y_s\right) \\
&\quad - \epsilon_2 J\left(\bar{W}' + W_k, \bar{w} + y_{s+1}\right),
\end{aligned} \tag{41}$$

where the first equality is the definition of L, and in the second equality we explicitly write expectations. The third equality follows from simple arithmetics, and in the fourth equality, we eliminate identical terms using the local spread and the fact that for $i \notin \{s-1, s, s+1\}$: $Prob(Y_l = y_i) = Prob(Y_{l+1} = y_i)$, and for $i \in \{s-1, s, s+1\}$, we can express the difference in the probability mass function using $\epsilon_1$ and $\epsilon_2$.

**Step 4.2: The term $R$.**

The term $R$ is the difference of the value functions under policy Y applied to two very similar states that share the same tested coefficients, and are different only by a single untested coefficient. Under policy Y, at both states, the same control is chosen according to what policy $\pi_j$ does at state $\left(\bar{W}' + W_k, \bar{w} + W_k\right)$. We therefore divide the support of the random variable $W_k$ into two sets based on whether policy $\pi_j$ stops (the set $\mathcal{S}_1$) or tests (the set $\mathcal{S}_2$) at state $\left(\bar{W}' + W_k, \bar{w} + W_k\right)$:

$$\mathcal{S}_1 = \left\{m : J_{\pi_j}\left(\bar{W}' + W_k, \bar{w} + v_m\right) = \varphi_P\left(\bar{W}' + W_k, \bar{w} + v_m\right)\right\},$$

and

$$\mathcal{S}_2 = \left\{m : J_{\pi_j}\left(\bar{W}' + W_k, \bar{w} + v_m\right) = -c + \mathbb{E}_k\left[J_{\pi_j}\left(\bar{W}' + W_k, \bar{w} + v_m\right)\right]\right\}.$$

We can then rewrite the term $R$ as follows:

$$\begin{aligned}
R &= \mathbb{E}J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + W_k\right) - \mathbb{E}J_Y\left(\bar{W}' + Y_l, \bar{w} + W_k\right) \\
&\underset{(1)}{=} \sum_m Prob(W_k = v_m)\left(J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + v_m\right) - J_Y\left(\bar{W}' + Y_l, \bar{w} + v_m\right)\right) \\
&\underset{(2)}{=} \sum_{m \in S_1} Prob(W_k = v_m)\left(J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + v_m\right) - J_Y\left(\bar{W}' + Y_l, \bar{w} + v_m\right)\right) \\
&\quad + \sum_{m \in S_2} Prob(W_k = v_m)\left(J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + v_m\right) - J_Y\left(\bar{W}' + Y_l, \bar{w} + v_m\right)\right) \\
&\underset{(3)}{=} \sum_{m \in S_1} Prob(W_k = v_m)\left(\varphi\left(\mathbb{E}[\bar{W}' + Y_{l+1}], \bar{w} + v_m\right) - \varphi\left(\mathbb{E}[\bar{W}' + Y_l], \bar{w} + v_m\right)\right) \\
&\quad + \sum_{m \in S_2} Prob(W_k = v_m)\left(J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + v_m\right) - J_Y\left(\bar{W}' + Y_l, \bar{w} + v_m\right)\right) \\
&\underset{(4)}{=} \sum_{m \in S_1} Prob(W_k = v_m)\,(0) \\
&\quad + \sum_{m \in S_2} Prob(W_k = v_m)\left(J_Y\left(\bar{W}' + Y_{l+1}, \bar{w} + v_m\right) - J_Y\left(\bar{W}' + Y_l, \bar{w} + v_m\right)\right) \\
&\underset{(5)}{=} \sum_{m \in S_2} Prob(W_k = v_m)\left(-c + \mathbb{E}J\left(\bar{W}', \bar{w} + v_m + Y_{l+1}\right)\right) \\
&\quad + \sum_{m \in S_2} -Prob(W_k = v_m)\left(-c + \mathbb{E}J\left(\bar{W}', \bar{w} + v_m + Y_l\right)\right) \\
&\underset{(6)}{=} \sum_{m \in S_2} Prob(W_k = v_m)\left(\mathbb{E}J\left(\bar{W}', \bar{w} + v_m + Y_{l+1}\right) - \mathbb{E}J\left(\bar{W}', \bar{w} + v_m + Y_l\right)\right) \\
&\underset{(7)}{=} \sum_{m \in S_2} Prob(W_k = v_m)\epsilon_1 J\left(\bar{W}', \bar{w} + v_m + y_{s-1}\right) \\
&\quad - \sum_{m \in S_2} Prob(W_k = v_m)\left(\epsilon_1 + \epsilon_2\right) J\left(\bar{W}', \bar{w} + v_m + y_s\right) \\
&\quad + \sum_{m \in S_2} Prob(W_k = v_m)\epsilon_2 J\left(\bar{W}', \bar{w} + v_m + y_{s+1}\right)
\end{aligned} \tag{42}$$

where in equality (1) we explicitly write the expectation with respect to $W_k$, and in equality (2) we divide the support to the sets $\mathcal{S}_1$, and $\mathcal{S}_2$. By the definition of the set $\mathcal{S}_1$, policy Y stops when $v_m \in \mathcal{S}_1$, which allows us to replace the value function $J_Y$ with the stopping cost $\varphi$, and obtain equality (3). Using the symmetry of the function $\varphi$, and the convex order which implies that $\mathbb{E}[Y_l] = \mathbb{E}[Y_{l+1}]$, we obtain that the difference in the summation of over the set $\mathcal{S}_1$ is equal to zero, which brings us to equality (4). In equality (5), we use the fact that when $v_m \in \mathcal{S}_2$ testing is optimal, which we simplify to obtain equality (6). Finally, and similarly to the derivation of the term L, we can write the expectations using probability mass functions, and cancel identical terms using the similarity of the distributions of $Y_l$ and $Y_{l+1}$.

Now that we derived the terms for L and R, we can sum them to obtain Equation (39) which we try to prove is non-positive.

**Step 4.3: The term $L + R$.**

We can sum the derived terms for L and R (equations 41 and 42, respectively):

$$\begin{aligned}
L + R \;=\;& -\epsilon_1 J\left(\bar{W}' + W_k, \bar{w} + y_{s-1}\right) + (\epsilon_1 + \epsilon_2) J\left(\bar{W}' + W_k, \bar{w} + y_s\right) \\
& -\epsilon_2 J\left(\bar{W}' + W_k, \bar{w} + y_{s+1}\right) \\
& + \sum_{m \in S_2} Prob(W_k = v_m)\epsilon_1 J\left(\bar{W}', \bar{w} + v_m + y_{s-1}\right) \\
& - \sum_{m \in S_2} Prob(W_k = v_m)\left(\epsilon_1 + \epsilon_2\right) J\left(\bar{W}', \bar{w} + v_m + y_s\right) \\
& + \sum_{m \in S_2} Prob(W_k = v_m)\epsilon_2 J\left(\bar{W}', \bar{w} + v_m + y_{s+1}\right) \\
=\;& (\epsilon_1 + \epsilon_2)\left(J\left(\bar{W}' + W_k, \bar{w} + y_s\right) - \sum_{m \in S_2} Prob(W_k = v_m) J\left(\bar{W}', \bar{w} + v_m + y_s\right)\right) \\
& -\epsilon_1\left(J\left(\bar{W}' + W_k, \bar{w} + y_{s-1}\right) - \sum_{m \in S_2} Prob(W_k = v_m) J\left(\bar{W}', \bar{w} + v_m + y_{s-1}\right)\right) \\
& -\epsilon_2\left(J\left(\bar{W}' + W_k, \bar{w} + y_{s+1}\right) - \sum_{m \in S_3} Prob(W_k = v_m) J\left(\bar{W}', \bar{w} + v_m + y_{s+1}\right)\right) \\
=\;& (\epsilon_1 + \epsilon_2)\phi(y_s) - \epsilon_1\phi(y_{s-1}) - \epsilon_2\phi(y_{s+1}) \\
=\;& (\epsilon_1 + \epsilon_2)\left(\phi(y_s) - \frac{\epsilon_1}{\epsilon_1 + \epsilon_2}\phi(y_{s-1}) - \frac{\epsilon_2}{\epsilon_1 + \epsilon_2}\phi(y_{s+1})\right) \qquad (43)
\end{aligned}$$

where the function $\phi(y)$ is defined as follows:

$$\phi(y) = J\left(\bar{W}' + W_k, \bar{w} + y\right) - \sum_{m \in S_2} Prob(W_k = v_m) J\left(\bar{W}', \bar{w} + v_m + y\right).$$

This implies that the sufficient condition for the optimality of policy $\pi_k$ (Equation (39)), can be written as follows:

$$\phi(y_s) - \frac{\epsilon_1}{\epsilon_1 + \epsilon_2}\phi(y_{s-1}) - \frac{\epsilon_2}{\epsilon_1 + \epsilon_2}\phi(y_{s+1}) \leq 0. \tag{44}$$

By the definition of local spread and Equation 40, $y_s$ is a convex combination of the points $y_{s-1}$ and $y_{s+1}$:

$$y_s = \frac{\epsilon_1}{(\epsilon_1 + \epsilon_2)} y_{s-1} + \frac{\epsilon_2}{(\epsilon_1 + \epsilon_2)} y_{s+1},$$

and therefore to prove that Equation 44 holds, it is sufficient to show that the function $\phi(y)$ is convex in $y$. Observe that the convexity of $\phi(y)$ does not trivially holds since as it is the difference of two convex functions.

We next show that the function $\phi(y)$ is indeed convex and complete our proof.

**Step 5: Establishing the convexity of the function $\phi(y)$.**

To prove that the function $\phi(y)$ is convex, we calculate the derivative of $\phi(y)$ in the three consecutive intervals, denoted as intervals 1, 2, and 3. We show that the function $\phi(y)$ is continuous, piecewise linear, whose derivative is piecewise constant and increasing, and therefore the function $\phi(y)$ is convex.

We start by defining the three intervals. Corollary 3.12 implies that there are three possible ranges of values for $y$:

1. $y < v_1$ : it is optimal to stop at state $(\bar{W}' + W_k, \bar{w} + y)$
2. $v_1 \leq y \leq v_2$ : it is optimal to test at state $(\bar{W}' + W_k, \bar{w} + y)$
3. $v_2 < y$ : it is optimal to stop at state $(\bar{W}' + W_k, \bar{w} + y)$

Recall that the set $S_1$ is defined to be the values of the support $y_1, ..., y_z$ at which stopping is optimal, and therefore $S_1$ is equal to the union of intervals 1 and 3. Similarly, the set $S_2$ is the equal to interval 2.

We now calculate the derivative in each of the intervals.

**Interval 1**. By definition of interval 1, it is optimal to stop at state $(\bar{W}' + W_k, \bar{w} + y)$, which implies that it is also optimal to stop at any state $(\bar{W}', \bar{w} + y + v_m)$ (Corollary 3.13). We can therefore write the function $\phi$ in interval 1 as follows:

$$\phi_1(y) \quad = \quad J_{Stop}\left(\bar{W}' + W_k, \bar{w} + y\right) - \sum_{m \in S_2} Prob(W_k = v_m) J_{Stop}\left(\bar{W}', \bar{w} + v_m + y\right). \tag{45}$$

We are interested in $\phi_1'(y)$ in the range $y < v_1$:

$$
\begin{aligned}
\phi_1'(y) &= \frac{d}{dy}\left(J_{Stop}\left(\bar{W}' + W_k, \bar{w} + y\right) - \sum_{m \in S_2} Prob(W_k = v_m) J_{Stop}\left(\bar{W}', \bar{w} + v_m + y\right)\right) \\
&\stackrel{(1)}{=} \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}' + W_k], \bar{w} + y\right) - \sum_{m \in S_2} Prob(W_k = v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_m + y\right)\right) \\
&\stackrel{(2)}{=} \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + y + \mu\right) - \sum_{m \in S_2} Prob(W_k = v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_m + y\right)\right) \\
&\stackrel{(3)}{=} \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + y + v_1\right) - \sum_{m \in S_2} Prob(W_k = v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\right) \\
&\stackrel{(4)}{=} \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\left(1 - \sum_{m \in S_2} Prob(W_k = v_m)\right)\right) \\
&\stackrel{(5)}{=} \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\left(\sum_{m \in S_1} Prob(W_k = v_m)\right) \\
&\stackrel{(6)}{=} \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right) \sum_{m: v_m < v_1} Prob(W_k = v_m) \\
&\quad + \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right) \sum_{m: v_m > v_2} Prob(W_k = v_m) \\
&\stackrel{(7)}{=} \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right) \sum_{m: v_m < v_1} Prob(W_k = v_m) \\
&\quad + \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_2 + y\right) \sum_{m: v_m > v_2} Prob(W_k = v_m)
\end{aligned}
\tag{46}
$$

where the first transition we use DP formulation (Equation 26), and in transition (2) we use symmetry. In transition (3) we use the fact that $\varphi$ is solved by the greedy algorithm, and that in the resulting solution, the value of the variable corresponding to coefficient $y$, is not affected by changes in coefficients with higher values. Since that in interval 1 $y < v_1$, changing any coefficient that is larger or equal to $v_1$ to $v_1$ does not change the value of the variable corresponding to $y$, and therefore does not affect the derivative of $\varphi$ with respect to $y$. This allows us to replace $\mu$ by $v_1$. Similarly, we may replace any coefficient $v_m$ in the set $\mathcal{S}_2$ (which values are greater than $v_1$, by definition) with $v_1$.

With simple arithmetic we obtain transitions (4) and (5). In transition (6) we split the values of $v_m$ in the set $\mathcal{S}_1$ to values that are smaller than $v_1$, and to values that are higher than $v_2$. Finally with similar arguments to transition (3) we obtain transition (7).

**Interval 2**. In this interval, it is optimal to perform a test while being at state $(\bar{W}'+W_k, \bar{w}+y)$. We can write the derivative of $\phi(y)$ in this interval as follows:

$$
\begin{aligned}
&\phi_2'(y) \\
\overset{(1)}{=}\ & \frac{d}{dy}\left(J_{Test}\left(\bar{W}'+W_k, \bar{w}+y\right) - \sum_{m\in S_2} Prob(W_k=v_m) J\left(\bar{W}', \bar{w}+v_m+y\right)\right) \\
\overset{(2)}{=}\ & \frac{d}{dy}\left(-c + \sum_{m\in S_1\cup S_2} Prob(W_k=v_m) J\left(\bar{W}', \bar{w}+v_m+y\right) - \sum_{m\in S_2} Prob(W_k=v_m) J\left(\bar{W}', \bar{w}+v_m+y\right)\right) \\
\overset{(3)}{=}\ & \frac{d}{dy}\left(\sum_{m\in S_1} Prob(W_k=v_m) J\left(\bar{W}', \bar{w}+v_m+y\right)\right) \\
\overset{(4)}{=}\ & \frac{d}{dy}\left(\sum_{m\in S_1} Prob(W_k=v_m) J_{Stop}\left(\bar{W}', \bar{w}+v_m+y\right)\right) \\
\overset{(5)}{=}\ & \frac{d}{dy}\left(\sum_{m\in S_1} Prob(W_k=v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_m+y\right)\right) \\
\overset{(6)}{=}\ & \frac{d}{dy}\left(\sum_{m:v_m<v_1} Prob(W_k=v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_m+y\right)\right) \\
& + \frac{d}{dy}\left(\sum_{m:v_m>v_2} Prob(W_k=v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_m+y\right)\right) \\
\overset{(7)}{=}\ & \frac{d}{dy}\left(\sum_{m:v_m<v_1} Prob(W_k=v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_1+y\right)\right) \\
& + \frac{d}{dy}\left(\sum_{m:v_m>v_2} Prob(W_k=v_m) \varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_2+y\right)\right) \\
\overset{(8)}{=}\ & \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_1+y\right) \sum_{m:v_m<v_1} Prob(W_k=v_m) \\
& + \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w}+v_2+y\right) \sum_{m:v_m>v_2} Prob(W_k=v_m)
\end{aligned}
\tag{47}
$$

where equality (1) follows from the definition of interval 2. In (2) we use the definition of expectation, and in (3) we eliminate similar terms and the constant $c$ which does not affect the derivative. In (4) we use the fact that when $v_m \in \mathcal{S}_1$ it is optimal to stop at state $(\bar{W}'+W_k, \bar{w}+v_m)$, which implies that it is also optimal to stop at state $(\bar{W}', \bar{w}+y+v_m)$ (Corollary 3.13). We can then rewrite the last expression using the function $\varphi$ (transition (5)). In (6) we split the summation to two, and in (7) we use Lemma 3.2 and the fact that $y$ is in interval 2 which allows us to change any coefficients that are outside of interval 2, that is, the values of $v_m$. In (8) we simply rearrange the terms.

**Interval 3**. Similarly to to interval 1, it is optimal to stop at states $(\bar{W}' + W_k, \bar{w} + y)$ and $(\bar{W}', \bar{w} + y + v_m)$ when $y > v_2$, and we can write the derivative of the function $\phi(y)$ in this interval as follows:

$$\begin{aligned}
\phi_3'(y) &= \frac{d}{dy}\left(J_{Stop}\left(\mathbb{E}[\bar{W}' + W_k], \bar{w} + y\right) - \sum_{m\in S_2} Prob(W_k = v_m) J_{Stop}\left(\mathbb{E}[\bar{W}'], \bar{w} + v_m + y\right)\right) \\
&= \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}' + W_k], \bar{w} + y\right) - \sum_{m\in S_2} Prob(W_k = v_m)\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_m + y\right)\right) \\
&= \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + y + \mu\right) - \sum_{m\in S_2} Prob(W_k = v_m)\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_m + y\right)\right) \\
&= \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + y + v_1\right) - \sum_{m\in S_2} Prob(W_k = v_m)\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\right) \\
&= \frac{d}{dy}\left(\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + y + v_1\right)\left(1 - \sum_{m\in S_2} Prob(W_k = v_m)\right)\right) \\
&= \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\left(\sum_{m\in S_1} Prob(W_k = v_m)\right) \\
&= \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_1 + y\right)\sum_{m:v_m<v_1} Prob(W_k = v_m) \\
&\quad + \frac{d}{dy}\varphi\left(\mathbb{E}[\bar{W}'], \bar{w} + v_2 + y\right)\sum_{m:v_m>v_2} Prob(W_k = v_m), \qquad (48)
\end{aligned}$$

where all the transitions follows for similar arguments as in interval 1.

We see that the derivative of $\phi(y)$ is the same in all three intervals (equations 46, 47, and 48 are identical). Moreover, this is the derivative of a non-decreasing piecewise linear and convex function (Lemma 3.2), which means that $\phi'(y)$ is piecewise constant and non-decreasing. The continuity of $\phi(y)$, and the fact that its derivative is piecewise constant, and non-decreasing, implies that $\phi(y)$ is convex in $y$.

The convexity of $\phi(y)$, implies that the inequalities in equations 44, 39, and 36 hold, and therefore policy $\pi_k$ is optimal, which contradicts the suboptimality of policy $\pi_k$. □

# 4 A General Testing Problem

In Section 3, we proved that a stopping rule that is based on myopic gains is optimal for certain LOPTs. In this section, we prove that the same stopping rule is optimal for a more general class of problems. Starting with the general problem formulation (Section 4.1), we then provide a sufficient condition under which a myopic stopping rule is optimal (Section 4.2).

## 4.1 Problem Formulation

Consider an optimization problem $P$ that depends on a vector of parameters $\bar{w}$ (more generally, it can be a matrix or a tensor, but we describe it as a vector for simplicity). For example, $\bar{w}$ can be the matrix and cost coefficients of a linear program, or the weights of edges in a set cover problem, or perhaps the cost parameters of an inventory control problem. For every vector $\bar{w}$, there is an optimal policy that achieves the objective value which we denote by $\varphi_P(\bar{w})$.

In a General Testing Problem (GTP), the untested parameters are mutually independent random variables, denoted by a vector $\bar{W}$. Moreover, the realization of these random variables can be observed by testing at the cost $c > 0$. At any point of time, we can either test one of the parameters, or to stop testing and return a feasible solution with respect to tested values and expected values of the untested parameters. Our goal is to find an optimal policy to decide adaptively on which parameters to test and when to stop.

As the testing problem evolves, more parameters are tested. Let $\bar{W}$ denote the set of untested parameters and $\bar{w}$ the set of tested (or realized) parameters. We can then describe the system state using the tuple $(\bar{W}, \bar{w})$.

**Assumption 4.1.** *The vector of variables that solves the expected maximization problem $P$ is the same as the vector of variables that solves the maximization problem $P$ over the expectations of coefficients, i.e.,*

$$x = \arg\max \mathbb{E}\left[\varphi_P\left(\bar{W}, \bar{w}\right)\right] = \arg\max \varphi_P\left(\mathbb{E}\left[\bar{W}\right], \bar{w}\right).$$

Notice that except for the case of a linear objective we encountered in the Section 3, there are many other functions for which Assumption 4.1 holds when the untested coefficients in $\bar{W}$ are mutually independent.

Using Assumption 4.1, the general testing problem can be written as the following Dynamic Programming problem:

$$J^{\text{opt}}\left(\bar{W}, \bar{w}\right) = \max \begin{cases} \varphi(\mathbb{E}[\bar{W}], \bar{w}) & \text{optimize} \\ -c + \mathbb{E}_{W_i}\left[J^{\text{opt}}\left(\bar{W} - W_i, \bar{w} + W_i\right)\right] & \text{test}_i \end{cases} \tag{49}$$

where $J^{\text{opt}}$ denotes the value function of the optimal policy at state $\left(\bar{W},\bar{w}\right)$. We use '-' and '+' to denote exclusion and inclusion of elements from a set ($\bar{W}-W_i \equiv \bar{W}\setminus\{W_i\}$ and $\bar{w}+W_i \equiv \bar{w}\cup\{W_i\}$).

Testing implies that the algorithm continues to learn and update the unknown parameters. Optimizing, on the other hand, implies that the algorithm stops learning and solves the optimization problem using the expected values of the remaining untested parameters. Optimizing problems using expectations is a common practice and often is optimal (as is the case for the problems reviewed in Section 3).

**Connection to Optimal Stopping Problems**

There is a similarity between problems in optimal stopping theory and GTPs. In both types of problems, the goal is to choose a stopping rule that maximizes the expected reward. However, in GTPs, the policy can also define the order in which tests are made. So, in this sense, GTPs are generalizations of optimal stopping problems in cases where the testing order does not matter. In other words, if the parameters are identically distributed and the function $\varphi$ is symmetric, then the corresponding GTP becomes an optimal stopping time problem.

## 4.2 An Optimal Stopping Rule for GTPs

Observe that under Assumption 4.1, the definitions in Section 3.3 hold for the general case as well.

**Assumption 4.2.** *At every state $\left(\bar{W},\bar{w}\right)$, for every untested parameter $W_i \in \bar{W}$ and for every tested parameter $w_t \in \bar{w}$, the function $\Delta_i\left(\bar{W},\bar{w}\right)$ is maximized at $w_t = \mathbb{E}[W_t]$.*

**Theorem 4.3.** *(Myopic Stopping Rule) Under assumptions 4.1 and 4.2, stopping is optimal, if and only if, stopping is at least as good as every myopic testing policy, i.e.,*

$$J^{\text{opt}}\left(\bar{W},\bar{w}\right) = J^{\text{stop}}\left(\bar{W},\bar{w}\right) \Leftrightarrow \forall W_i \in \bar{W}, J_i^{\text{my}}\left(\bar{W},\bar{w}\right) \leq J^{\text{stop}}\left(\bar{W},\bar{w}\right). \tag{50}$$

*Proof.* Proceed by induction on the number of untested parameter $k$. When $k=1$ myopic testing is the same as testing, so the lemma clearly holds.

When $k>1$, assume that there exists a coefficient $W_{i'} \in \bar{W}$ such that $J_{i'}^{\text{my}}\left(\bar{W},\bar{w}\right) > J^{\text{stop}}\left(\bar{W},\bar{w}\right)$. Then it is suboptimal to stop, and therefore $J^{\text{opt}}\left(\bar{W},\bar{w}\right) > J^{\text{stop}}\left(\bar{W},\bar{w}\right)$.

Now for the opposite direction, assume

$$\forall W_i \in \bar{W}, \Delta_i\left(\bar{W},\bar{w}\right) = J_i^{\text{my}}\left(\bar{W},\bar{w}\right) - J^{\text{stop}}\left(\bar{W},\bar{w}\right) \leq 0. \tag{51}$$

We need to show that $J^{\text{stop}}\left(\bar{W},\bar{w}\right) \geq J_i^{\text{test}}\left(\bar{W},\bar{w}\right)$ for all $W_i \in \bar{W}$. To do that, it suffices to show that is not optimal to test more than once, i.e., $J_i^{\text{test}}\left(\bar{W},\bar{w}\right) = J_i^{\text{my}}\left(\bar{W},\bar{w}\right)$. Let us test once.

Suppose we test coefficient $W_t \in \bar{W}$ and get the realization $w_t$, moving to state $\left(\bar{W} - W_t, \bar{w} + w_t\right)$. By Remark 3.7 and equation (51),

$$\forall W_i \in \left\{\bar{W} - W_t\right\}, \Delta_i\left(\bar{W} - W_t, \bar{w} + \mathbb{E}[W_t]\right) = \Delta_i\left(\bar{W}, \bar{w}\right) \leq 0. \tag{52}$$

By assumption 4.2, and equation (52), for all $w_t \in \mathbb{R}^+$

$$\forall W_i \in \left\{\bar{W} - W_t\right\}, \Delta_i\left(\bar{W} - W_t, \bar{w} + w_t\right) \leq \Delta_i\left(\bar{W} - W_t, \bar{w} + \mathbb{E}[W_t]\right) \leq 0. \tag{53}$$

Using the induction hypothesis, if we test any yet untested parameter, it is optimal to stop in the next iteration. This implies that $J_i^{\text{test}}\left(\bar{W}, \bar{w}\right) = J_i^{\text{my}}\left(\bar{W}, \bar{w}\right)$ for all $W_i \in \bar{W}$. Along with equation (51), this implies

$$\forall W_i \in \bar{W}, J_i^{\text{test}}\left(\bar{W}, \bar{w}\right) \leq J^{\text{stop}}\left(\bar{W}, \bar{w}\right).$$

Therefore, stopping is better than testing any edge, which implies stopping is optimal

$$J^{\text{opt}}\left(\bar{W}, \bar{w}\right) = J^{\text{stop}}\left(\bar{W}, \bar{w}\right).$$

□

Assumption 4.2 states that the myopic gain is monotonically decreasing with testing regardless of our choice for the tested parameters and the realizations of the tested parameters. This implies that once myopically it is not beneficial to test, it will remain so in the future. This explains why Theorem 4.3 holds and why the myopic stopping rule is optimal.

# 5 Concluding Remarks and Future Work

For many interesting cases of stochastic combinatorial optimization problems with testing that have real-world applications, we obtained an optimal policy that is described using myopic rules. We assumed throughout that the testing cost $c > 0$ is uniform and fixed, that the random coefficients $W_1, \ldots, W_N$ are mutually independent, and that they have equal finite means $\mathbb{E}[W_i] = \mu$.

These results suggest several directions for future work. First, it may be interesting to examine the problems mentioned in this thesis with less restrictive assumptions. For example, by examining non-uniform testing cost, i.e., each coefficient has a different testing cost, or even a case with random testing costs. In particular, relaxing the equal-means assumption for selection with testing would be especially intriguing, because it can lead to progress on other interesting problems, such as shortest path with testing.

For general optimization problems with testing, we found sufficient conditions under which a myopic stopping rule is optimal. It remains open to find sufficient conditions under which a myopic testing-order rule is optimal. One may also wonder about *necessary* conditions for the existence of a myopic policy that obtains optimal profit. We conjecture that a necessary condition for a myopic policy to obtain optimal for a testing problem is that the deterministic version of the problem can be solved efficiently by a greedy algorithm. Notice that this condition holds for all problems that were considered in this thesis.

So far we considered only solutions that achieve the optimum exactly. It may be interesting to consider also myopic policies that lead to approximately optimal solutions for problems whose deterministic version is NP-hard. A first step in this direction may be to examine NP-hard problems that are approximated well by a greedy algorithm, such as set cover (or more generally, problems that fall under the formulation of Wolsey (1982)).

# References

Adamczyk, Marek, Maxim Sviridenko, Justin Ward. 2013. Submodular stochastic probing on matroids. *arXiv preprint arXiv:1310.4415* .

Bellman, Richard. 1954. The theory of dynamic programming. Tech. rep., DTIC Document.

Bertsimas, Dimitris, John N Tsitsiklis. 1997. *Introduction to linear optimization*, vol. 6. Athena Scientific Belmont, MA.

Bubeck, Sébastien, Nicolo Cesa-Bianchi. 2012. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv:1204.5721* .

Chen, Ning, Nicole Immorlica, Anna R Karlin, Mohammad Mahdian, Atri Rudra. 2009. Approximating matches made in heaven. *Automata, Languages and Programming*. Springer, 266–278.

Dean, Brian C, Michel X Goemans, J Vondrdk. 2004. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*. IEEE, 208–217.

Edmonds, Jack. 1970. Submodular functions, matroids, and certain polyhedra. *Combinatorial structures and their applications* 69–87.

Edmonds, Jack. 1971. Matroids and the greedy algorithm. *Mathematical programming* **1**(1) 127–136.

Erlebach, Thomas, Michael Hoffmann, Danny Krizanc, Matús Mihal'Ák, Rajeev Raman. 2008. Computing minimum spanning trees with uncertainty. *arXiv preprint arXiv:0802.2855* .

Feder, Tomás, Rajeev Motwani, Liadan O'Callaghan, Chris Olston, Rina Panigrahy. 2007. Computing shortest paths with uncertainty. *Journal of Algorithms* **62**(1) 1–18.

Feder, Tomas, Rajeev Motwani, Rina Panigrahy, Chris Olston, Jennifer Widom. 2000. Computing the median with uncertainty. *Proceedings of the thirty-second annual ACM symposium on Theory of computing*. ACM, 602–607.

Ferguson, Thomas S. 2012. Optimal stopping and applications .

Frazier, Peter, Warren Powell, Savas Dayanik. 2009. The knowledge-gradient policy for correlated normal beliefs. *INFORMS journal on Computing* **21**(4) 599–613.

Gittins, John, Kevin Glazebrook, Richard Weber. 2011. *Multi-armed bandit allocation indices*. John Wiley & Sons.

Goel, Ashish, Sudipto Guha, Kamesh Munagala. 2006. Asking the right questions: Model-driven optimization using probes. *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 203–212.

Goel, Ashish, Sudipto Guha, Kamesh Munagala. 2010. How to probe for an extreme value. *ACM Transactions on Algorithms (TALG)* **7**(1) 12.

Goerigk, Marc, Manoj Gupta, Jonas Ide, Anita Schöbel, Sandeep Sen. 2015. The robust knapsack problem with queries. *Computers & Operations Research* **55** 12–22.

Golovin, Daniel, Andreas Krause. 2011. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research* 427–486.

Guha, Sudipto, Kamesh Munagala. 2007. Model-driven optimization using adaptive probes. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 308–317.

Guha, Sudipto, Kamesh Munagala. 2008. Adaptive uncertainty resolution in bayesian combinatorial optimization problems. *arXiv preprint arXiv:0812.1012* .

Guha, Sudipto, Kamesh Munagala, Saswati Sarkar. 2007. Information acquisition and exploitation in multichannel wireless systems. *IEEE Transactions on Information Theory*. Citeseer.

Gupta, Anupam, Viswanath Nagarajan. 2013. A stochastic probing problem with applications. *Integer Programming and Combinatorial Optimization*. Springer, 205–216.

Gupta, Manoj, Yogish Sabharwal, Sandeep Sen. 2011. The update complexity of selection and related problems. *arXiv preprint arXiv:1108.5525* .

Javdani, Shervin, Yuxin Chen, Amin Karbasi, Andreas Krause, J Andrew Bagnell, Siddhartha Srinivasa. 2014. Near optimal bayesian active learning for decision making. *arXiv preprint arXiv:1402.5886* .

Khanna, Sanjeev, Wang-Chiew Tan. 2001. On computing functions with uncertainty. *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 171–182.

Levi, Retsef, Thomas Magnanti, Yaron Shaposhnik. 2014. Scheduling with testing .

MacQueen, James, RG Miller Jr. 1960. Optimal persistence policies. *Operations Research* **8**(3) 362–380.

Megow, Nicole, Julie Meißner, Martin Skutella. 2015. Randomization helps computing a minimum spanning tree under uncertainty. *Algorithms-ESA 2015*. Springer, 878–890.

Müller, Alfred, Dietrich Stoyan. 2002. *Comparison methods for stochastic models and risks*, vol. 389. Wiley.

Powell, Warren B, Ilya O Ryzhov. 2012. *Optimal learning*, vol. 841. John Wiley & Sons.

Ryzhov, Ilya O, Warren B Powell. 2012. Information collection for linear programs with uncertain objective coefficients. *SIAM Journal on Optimization* **22**(4) 1344–1368.

Scarf, Herbert. 1959. The optimality of (s, s) policies in the dynamic inventory problem .

Schrijver, Alexander. 2003. *Combinatorial optimization: polyhedra and efficiency*, vol. 24. Springer Science & Business Media.

Sun, Zhankun, Nilay Tanık Argon, Serhan Ziya. 2014. Priority scheduling of jobs with unknown types .

Swisher, James R, Sheldon H Jacobson, Enver Yücesan. 2003. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **13**(2) 134–154.

Wolsey, Laurence A. 1982. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica* **2**(4) 385–393.

Yao, David D, Li Zhang. 1997. Stochastic scheduling via polymatroid optimization. *LECTURES IN APPLIED MATHEMATICS-AMERICAN MATHEMATICAL SOCIETY* **33** 333–364.