

OCTERACT LIMITED
CASE STUDY

**Deterministic global optimization of
cardinality-constrained and risk-constrained
portfolios**

Nikos KAZAZAKIS
Gabriel LAU

MAY, 2018

1 Introduction

In 2016 the Financial Times reported that 99% of actively managed US equity funds have failed to beat their benchmark since 2006 [2]. Thus, active financial institutions are understandably under pressure to research and develop strategies that will allow them to consistently beat the market and ebb the huge shift of institutional money towards index funds.

In search of better active investment methods to construct portfolios, the finance community has placed great emphasis on the quality of data. Vast amounts of money are spent today to buy “good” input data, typically associated with buzz words such as AI, alpha signals, or NLP.

Although input data is undoubtedly important in any kind of scientific modeling - such as that of assets, it is only one part of what is necessary to build highly profitable portfolios. In this case study, we discuss a part that is typically overlooked: the accuracy of the models themselves, i.e., how well the mathematics describe the market, and the importance in being able to solve these mathematics correctly.

Our motivation is highly intuitive, because it is straightforward to realise that regardless of the quality of the input data to a mathematical model, it is meaningless to worry about said data unless (i) the model describes nature well, and (ii) we are able to solve the mathematics correctly in the first place.

The mathematical models in question are portfolio optimization models. These optimization models are used to identify the optimal allocation of assets for an investor’s portfolio. This kind of modeling is common across all major financial institutions: top asset managers with trillions of dollars under management, hedge funds, and investment arms of the largest banks to name a few, and for good reason. Given the magnitude of investment value under consideration, it is vitally important for these institutions to efficiently control their exposure to risk while maximizing their return.

These models trace back to the seminal work of Harry Markowitz in 1952 [1], for which he was awarded the Nobel prize in economics. Markowitz provided a general mathematical framework to determine how to distribute an initial capital across a risky selection of securities, to create an efficient portfolio - one that has the optimal trade-off between risk exposure and return. Despite being introduced almost 60 years ago, the Markowitz Mean-Variance framework is still the de facto industry standard for portfolio construction, with advancements mainly in introducing further diversification constraints (e.g. to avoid concentration in countries, company sizes, sectors, etc.), or in improving the approximation of the input data (e.g. covariance matrices, alpha signals, the Black-Litterman model, etc.).

The models in use today are admittedly more complicated than in 1952. Institutional investors are generally required to construct portfolios that have additional constraints other than diversification. This typically involves limiting the number of unique assets in the portfolio, or how many assets can be added outside a specified benchmark. The accuracy of these additional constraints can be game-changing, as they control the costs incurred when the assets are traded or held, as well as allowing the creation of portfolios concentrated enough to outperform the market.

In this case study, we look at two extensions to the traditional Markowitz portfolio optimization problem, and their impact to the investor’s decision-making, using real market data. The extensions involve (i) placing cardinality constraints on a portfolio, which limit the number of distinct assets to a predefined range, and (ii) tracking error constraints, which allow tight control on the exposure to risk and how closely a portfolio follows its benchmark. These extensions transform the basic Markowitz problem - otherwise solvable using any quadratic programming solver - to a much more

difficult non-convex optimization problem that cannot be solved to guaranteed global optimality by any solver currently on the market.

Using the Octeract deterministic global optimization solver for large-scale non-convex portfolio optimization problems, we demonstrate that (i) non-convex extensions are necessary to make models more accurate, and (ii) when people attempt to solve them using traditional trial-and-error approaches the results can be off by orders of magnitude.

2 Problem

The portfolio is assumed to only consist of long positions (asset proportions are always positive) and fully invested at all times (proportions sum up to 100%). Diversification is enforced through constraints which (i) ensure that the proportion of assets span enough sectors and market capitalization (company sizes), and (ii) controlling the sensitivity to the equity market (Beta). Extensions to the basic model are implemented via constraints that limit the total number of stocks that can be invested and impose lower and upper bounds to the tracking error of the portfolio. These additional constraints alter the mathematical formulation of the problem considerably, and as such, the extended problem can no longer be solved with well established convex optimization techniques. Nevertheless, this model reflects the challenges and decisions an equity portfolio manager would face on a daily basis much more accurately than simpler models do.

2.1 Mathematical formulation of the extended Markowitz model

The portfolio selection problem is defined as follows:

$$\min_{\mathbf{w}} (\mathbf{w} - \mathbf{w}^{bench})^T \mathbf{\Omega} (\mathbf{w} - \mathbf{w}^{bench}) - \lambda \boldsymbol{\alpha}^T (\mathbf{w} - \mathbf{w}^{bench}) \quad (1)$$

s.t.

$$w_i \geq 0 \quad \forall i \quad (2)$$

$$\sum_i w_i = 1 \quad (3)$$

$$-0.05 \leq w_i - w_i^{bench} \leq 0.05 \quad \forall i \quad (4)$$

$$-0.1 \leq \sum_{i \in \text{sector } j} w_i - w_i^{bench} \leq 0.1 \quad \forall j \quad (5)$$

$$-0.1 \leq \sum_{i \in MCAPQk} w_i - w_i^{bench} \leq 0.1 \quad \forall k \quad (6)$$

$$-0.1 \leq \sum_i \beta_i (w_i - w_i^{bench}) \leq 0.1 \quad (7)$$

$$K^L \leq \text{cardinality}(w_i \neq 0) \leq K^U \quad (8)$$

$$TE^L \leq \sqrt{(\mathbf{w} - \mathbf{w}^{bench})^T \mathbf{\Omega} (\mathbf{w} - \mathbf{w}^{bench})} \leq TE^U \quad (9)$$

Variables and parameters are defined as follows:

\mathbf{w} : the vector of portfolio weights (proportions).

\mathbf{w}^{bench} : the vector of benchmark weights.

$\mathbf{\Omega}$: the covariance matrix which encodes the pairwise correlations between asset returns.

$\boldsymbol{\alpha}$: the vector of expected return scores (higher scores equate to higher future expected returns).

λ : a parameter which scales the contribution of the expected return to the objective function.

w_i : the weight of the overall portfolio held in asset i .

w_i^{bench} : the weight of the benchmark portfolio held in asset i .

β_i : the beta of asset i , which is a measure of its volatility.

$cardinality(w_i \neq 0)$: the number of non-zero weighted stocks.

K^L, K^U : the lower and upper bounds on the cardinality respectively.

TE^L, TE^U : the lower and upper bounds on the tracking error respectively.

A description of the each function and constraint is provided below:

- (1) The objective function balancing risk (1st term) and excess return (2nd term).
- (2) Long-only constraint where asset weights can not be negative (no short-selling).
- (3) No leverage constraint where the entire portfolio is invested.
- (4) Constraints on the minimum/maximum active weight per asset.
- (5) Constraints on the active weights for each sector.
- (6) Constraints on the active weights for each market capitalization.
- (7) Constraints on the active weights taking into account each asset's volatility.
- (8) Constraint on the cardinality of the portfolio.
- (9) Constraint on the tracking error (squareroot of the risk term in (1)) of the portfolio.

In simple terms, the mathematical formulation describes finding the set of portfolio weights which minimizes the risk and maximizes the return relative to the benchmark, given a set of constraints which enforce:

1. Diversification
2. Investing in a target number of stocks
3. Taking a certain amount of risk.

2.2 Equity market data

For this study, we will be using a dataset extracted from the historical analysis of the S&P500, a standard benchmark index for the equities market. The dataset includes covariance matrices, alpha scores, and beta scores, extracted from real benchmark data.

A typical workflow for a financial institution would involve portfolio rebalancing on a four-weekly period with historical backtesting of their latest strategies over a 10-year period. The historical backtesting would use datasets very similar to the ones used in this case study.

3 Pareto frontiers for cardinality-constrained portfolios

Assume that we only want to invest in 50 assets. As simple as that might sound, in a set of 500 elements, there are $\frac{500!}{50!(500-50)!} = 2.3 \times 10^{69}$ different combinations of 50 elements - these are a *lot* of combinations. In terms of scale, they are 100 times the number of atoms in the Milky Way galaxy.

Because this is a very hard problem to solve, it is not uncommon for quantitative analysts to use various heuristics, which return, to the best of an analyst’s knowledge (because these problems can not be solved without global optimization), a good combination of assets.

In order to investigate the discrepancy between heuristic solutions and rigorous ones, we interviewed professional quantitative analysts regarding the types of heuristics they use on a day-to-day basis, and applied the heuristic method that is most commonly used by our interviewees on historical data. As we shall see later on, the difference between the two methods of identifying combinations of assets can be staggering.

This investigation is performed by generating the optimal (Pareto) frontiers for a range of fixed cardinality S&P 500 portfolios using the Oocteract solver, in order to quantify the difference between the globally optimal cardinalities (combinations of assets) and all others.

To provide some context, the Markowitz Pareto frontier is the set of globally optimal solutions which balance the trade-off between risk and expected return, i.e., it is the set of solutions which offer the maximum possible expected return for a given level of risk (or, equivalently, the solutions which provide the minimum possible risk for a given level of return). As we shall see, *an input dataset is associated with an optimal cardinality which produces a better frontier than all others*. This cardinality changes along with the dataset, and is impossible to detect unless global optimization is used. We should note that by “impossible” we mean: (i) the actual best number of assets, (ii) the specific combination, and (iii) the exact proportion of each asset in a portfolio.

3.1 Methodology

We investigate a range of S&P 500 portfolios with fixed cardinalities between $K = 26$ to $K = 492$ (Table 1) using the Oocteract solver. At the time of writing, this solver is the only commercially available software that can solve the portfolio problem specified in Section 2.1 to guaranteed global optimality. We do so by generating a representative sample of points lying on the the Pareto frontier for a given cardinality, by solving a series of fixed tracking error ($TE^L = TE^U$) optimization problems, allowing us to efficiently traverse the front. The range of tracking errors that are sampled for each cardinality is shown in Table 1.

3.2 Comparison heuristic

For comparison, we present solutions obtained using a heuristic trial and error method commonly used by our interviewees, to assign stocks for some fixed cardinality. It should be pointed out that this method is simply one out many, and is used in this case study purely as a means of making a realistic comparison. Nevertheless, because there is no heuristic method which can guarantee globally optimal cardinality, the discrepancy between the heuristic results and global optimization is very real.

The heuristic method dictates the following steps to obtain a portfolio with cardinality K :

1. Solve the optimization model described in Section 2.1 *without* constraints (8) and (9) to find the K assets with the highest weights.

2. Fix the weights of the assets other than the top K assets to zero.
3. Reoptimize the portfolio with the fixed zero-weighted assets to obtain a set of new weights for the top K assets.

Despite being commonly used by our interviewees, this method suffers from a number of limitations:

1. It can only be used to find a portfolio of fixed cardinality, but cannot be used to find the optimal cardinality between two bounds (Equation (8) where $K^L \neq K^U$) without enumerating over all cardinalities.
2. The combination it identifies is effectively random. This methodology in no way brings us closer to the globally optimal combination for a fixed cardinality.
3. It makes unjustified assumptions about the cardinality-constrained solution space and artificially constrains asset weights based on these assumptions.
4. It can never theoretically guarantee finding the global optima, so one can only hope that the results don't lie too far away from the global solutions.

Cardinality	TE_{min}	TE_{max}
26	0.0078	0.0065
28	0.0025	0.0063
30	0.0025	0.0063
50	0.0025	0.0063
100	0.0022	0.0061
150	0.0011	0.0059
200	0.0004	0.0060
250	0.0003	0.0055
300	0.0002	0.0057
350	0.00005	0.0047
400	0.0004	0.0054
450	0.00011	0.0039
492	0.00003	0.0039

Table 1: The range of tracking errors investigated for portfolios of different cardinalities. For each cardinality, a set of 10 points lying on the Pareto frontier between TE_{min} and TE_{max} is generated.

Because solvers do not solve such problems to global optimality, many users are not aware of just how bad current solutions can be, so it's interesting to examine this more closely.

3.3 Results

3.3.1 Mapping the frontier using the Octeract solver

The Pareto frontiers for S&P 500 portfolios with fixed cardinalities ranging from $K = 26$ to $K = 492$ are illustrated in Figure 1. One can see that as cardinality is decreased, greater levels of return can be achieved for the same level of risk at high risk levels. These results are intuitive in that one

expects to have a higher chance of significantly beating the market if the portfolio becomes more concentrated (lower cardinality), such that investments are made on the higher performing assets. On the other hand, the high cardinality portfolios correspond to increased diversification which closely follow the benchmark, and therefore their returns in excess of the market performance are expected to be lower. These results, shown here with unprecedented clarity, are in line with the foundations of modern portfolio theory. The Octeract solver can be used to find the best results in this plot *automatically* - the only user intervention necessary is to declare their goal.

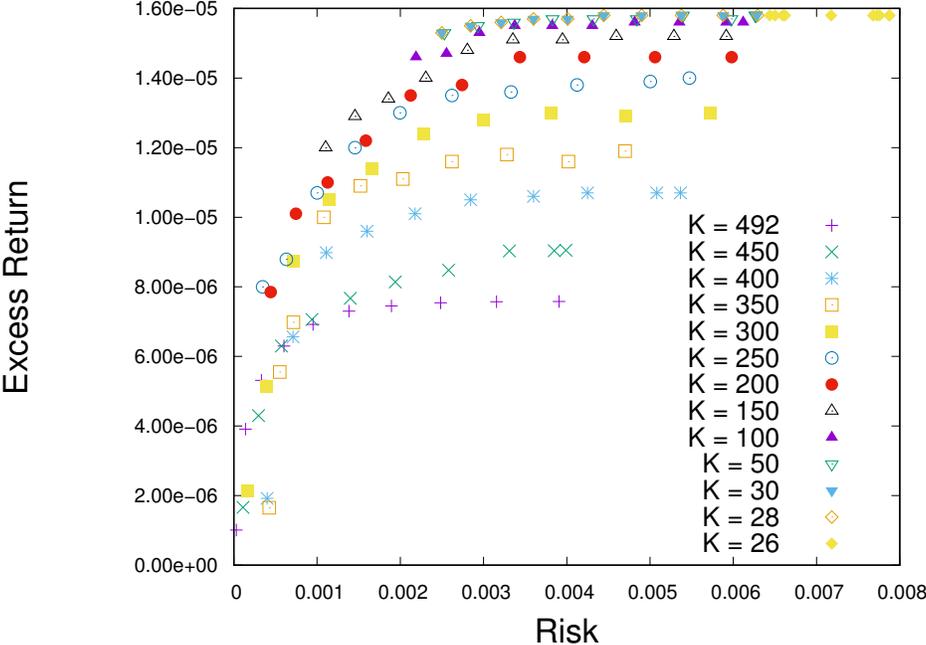


Figure 1: Pareto frontiers for S&P 500 portfolios with fixed cardinalities ranging from $K = 26$ to $K = 492$.

3.3.2 Comparisons with the cardinality heuristic

In Figures 2, 3 and 4, we compare the solutions obtained using the heuristic trial-and-error method described in Section 3.1 to the globally optimal solutions obtained by the Octeract Solver for portfolios with fixed cardinalities $K = 250$, $K = 100$, and $K = 50$ respectively.

We first examine Figure 2, where we include the solutions that would be typically be found by portfolio software in the financial industry. These results are clustered in the bottom left corner of the graph corresponding to very low risk, but very low return portfolios.

If we are able to find global solutions, and the risk is increased *just slightly* to $\sim 5e-4$ (which corresponds to a 2% aggregate return fluctuation), we can achieve about a **1000% increase** in the expected excess return! This is a remarkable result which exemplifies that current software can struggle to find solutions that lie in the low risk region corresponding to much higher expected returns. On the other hand, with the Octeract solver, the user can immediately find these solutions by just entering the desired tracking error bounds.

For the high risk regions in all three cardinality-constrained portfolios, we see stark differences

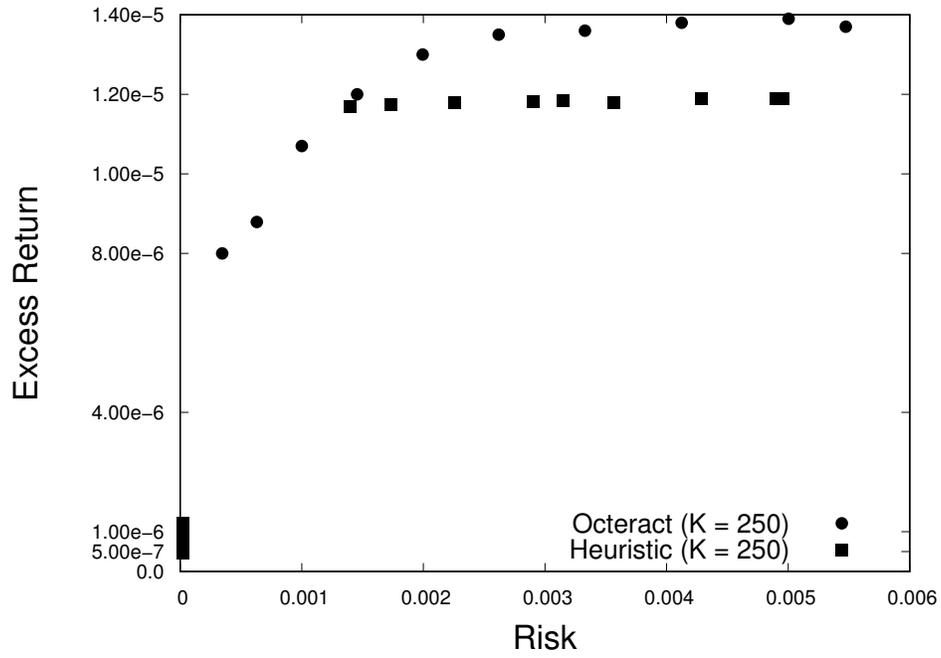


Figure 2: A comparison of the solutions obtained via a heuristic trial-and-error method used in industry to the set of globally optimal solutions for cardinality 250.

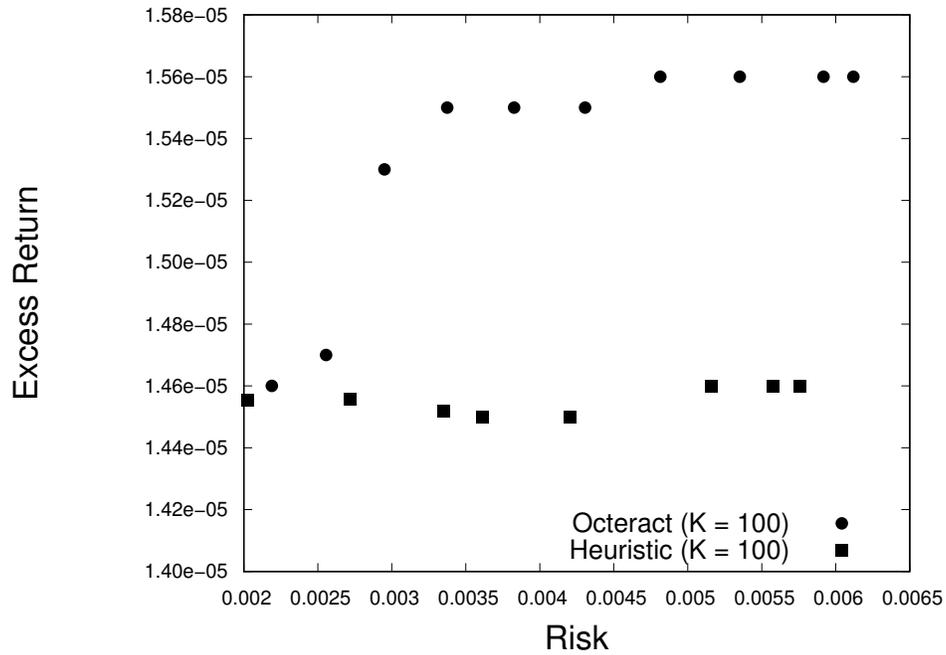


Figure 3: A comparison of the solutions obtained via a heuristic trial-and-error method used in industry to the set of globally optimal solutions for cardinality 100.

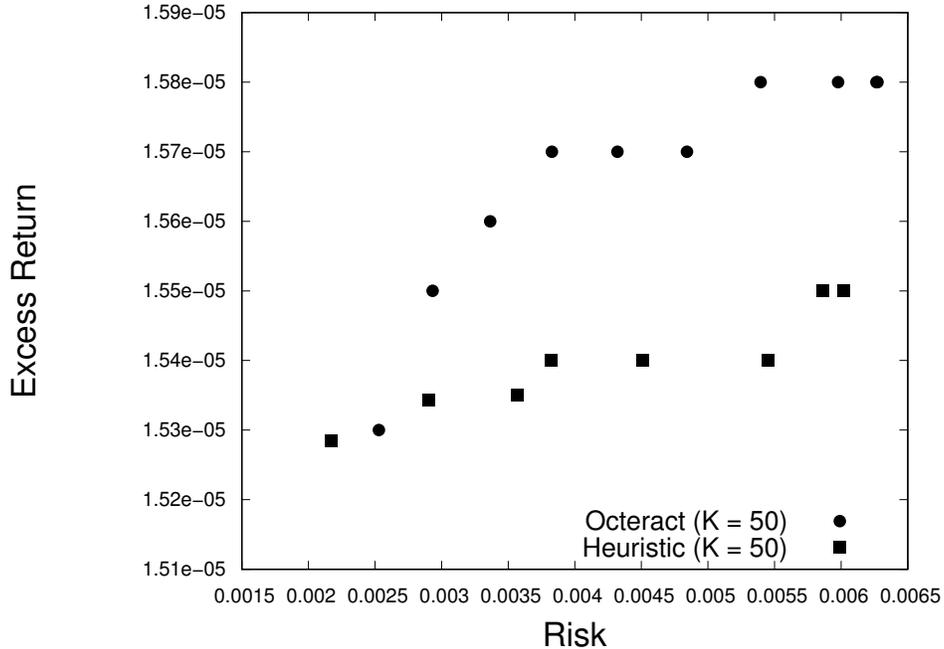


Figure 4: A comparison of the solutions obtained via a heuristic trial-and-error method used in industry to the set of globally optimal solutions for cardinality 50.

between the heuristic solutions and the set of Pareto-optimal solutions. Differences of up to 14% relative to the global solutions are found for $K = 250$, whereas for $K = 100$ and $K = 50$, differences of up to 6% and 2% are observed respectively. In order to explore the high risk region with the heuristic method, we had to adjust the scaling parameter λ by trial-and-error until solutions in the relevant region were found. This is *not* something that is done by portfolio optimization software for cardinality-constrained portfolios, and we stress that there exists no automated or robust method in this heuristic framework to find even non-global solutions within a given range of risk.

It is worth noting that the relative differences between the heuristic and global solutions increase significantly with higher cardinalities, as this is of importance in real-life financial strategies. In particular, portfolio managers generally aim to construct portfolios both with good diversification (as high a cardinality as possible to ensure their exposure to risk is kept relatively low), and an asset composition that generates maximum return for that diversification.

3.3.3 Insights

Our results suggest that, although largely unaware of the extent of the limitations, portfolio managers face three main issues:

1. For low risk strategies, their software cannot find the globally optimal solutions which provide a much greater increase in return (e.g. 1000% for $K = 250$ in the case study) for just a small increase in exposure to risk.
2. For higher risk strategies, they have a very large probability of missing out on the globally optimal strategies which generate the maximum expected return, assuming they even find

heuristic solutions in the desired risk range.

3. For cardinality-constrained problems, solutions up to 100% better are never located.

This is purely a technological limitation; after all, if a portfolio manager or quantitative analyst wishes to construct a cardinality-constrained portfolio with an amount of risk between two bounds, their software is *fundamentally unable to do this*.

Instead, they are forced to blindly go through processes of trial and error to traverse the Risk-Return space, in hopes of finding solutions close to the Pareto frontier that satisfy their risk constraints. No industrial software can confirm how close these solutions are to global optimality, and as clearly demonstrated in Figures 2, 3 and 4, they are often significantly far away for realistic portfolio compositions. This discrepancy demonstrates that modern decision-making methodologies are neither robust nor guaranteed, a fact that *severely impacts their predictive power*, even if the input data is of very good quality.

4 Conclusions

In this case study, we explored a fundamental problem in portfolio management: whether the inability of the vast majority of active funds to beat the index is linked to technological limitations in optimization technology.

Our computational results clearly demonstrate that the quality of the input data can be of secondary import, because the models are solved incorrectly to begin with. If portfolios are constructed using incorrect criteria (e.g., up to 1000% worse behaviour than what managers' own models can predict), these portfolios are bound to underperform, regardless of the quality of the input data. The only way to make sound decisions and to capitalise on good quality data is to be able to solve the models properly in the first place.

We illustrated this by computing the Pareto frontiers for a range of cardinality-constrained S&P 500 portfolios using real input data based on historical benchmark analysis. The frontiers were generated by solving a series of extended non-convex Markowitz optimization problems using the Oocteract solver, which is the first commercially available solver that can solve models of this type to guaranteed global optimality. The globally optimal solutions are compared to solutions obtained using a heuristic technique currently employed in the financial industry for portfolios with fixed cardinalities of 250, 100, and 50.

For cardinality 250, we show that solutions that would typically be found using current portfolio optimization software correspond to very low risk but very low return portfolios. The heuristic technique is unable to locate the just slightly riskier solutions which provide a massive increase in expected return. For example, an increase of 1000% is observed in the expected return for a small increase in risk of order $5e-4$. In the high risk regions for all three cardinalities, significant differences are found between the heuristic and global solutions, with divergences of up to 14% being observed. For active funds who profit on marginal percentages, 14% is a massive difference in predictive power, let alone 1000%. These solutions are utterly beyond reach, as any heuristic framework is fundamentally unable to reliably find even non-global solutions for a given range of risk.

Our results suggest that portfolio managers who have requirements to limit the number of unique assets in their portfolios are extremely likely to be missing out on the best possible low-risk and high-risk construction strategies. With the financial climate changing as larger amounts

of money are being shifted towards index funds, active fund managers are under great pressure to justify their performance fees when there is evidence showing that their current strategies are underperforming the benchmarks. Our results indicate that this underperformance is intrinsically connected to their inability to quantitatively test new portfolio optimization models/strategies and find the best possible solutions due to hard limitations in currently available software.

References

- [1] H. Markowitz. PORTFOLIO SELECTION. *The Journal of Finance*, 7:77–91, 1952.
- [2] C. Newlands and M. Marriage. 99% of actively managed US equity funds underperform, 2016.