Research Statement

My research focuses on the **design and analysis of decision-making models under uncertainty with limited information**. It is grounded in applications, in particular *portfolio and revenue management*, where stochasticity is often difficult to characterize using traditional probabilistic techniques because of fast-changing environments, for stock prices, or of limited data, for new products. My overarching goal is to incorporate uncertainty in an intuitive, tractable manner, which (i) is well-suited to the information available, i.e., does not require managers to make restrictive assumptions on the underlying probability distributions, (ii) leads to tractable formulations that can be solved efficiently for the large data sets encountered in practice, and (iii) yields theoretical insights into the structure of the optimal solution, to foster a better understanding of the computer-generated strategy. These features play an important role in practitioners adopting quantitative approaches. To achieve this research's goals, I use tools from *robust optimization* as well as *data-driven optimization*.

Methodology

<u>*Background*</u>

Robust optimization was originally developed to address parameter uncertainty in mathematical programming problems, and protects the system against the worst-case values of the parameters within a set. For instance, in portfolio management, the objective is to maximize the minimum wealth of the investor, where the minimum is computed over an uncertainty set representing possible values for the stock prices. *Data-driven optimization* incorporates historical data directly into the decision-making model, instead of using past observations to estimate underlying probabilities and then injecting these (possibly erroneous) values into the optimization problem. It is particularly useful when only a few data points are available, making it difficult to compute uncertainty sets accurately, or when the use of uncertainty sets yields overly conservative solutions because of the problem structure.

A key issue in robust optimization is the tractability of the resulting formulations. Because it is inherently a *max-min approach*, a naïve way to try solving the problem would be to minimize the worst-case uncertainty for a candidate allocation, then update the allocation to improve the objective for that worst-case uncertainty, and iterate between the maximization and minimization until an optimal solution is found. That is clearly not practical. Instead, the tractability of robust optimization hinges on the ability to reformulate the inner minimization problem as an equivalent maximization one, to create one large maximization problem that can be solved *in one step*. (The new problem must also have an appropriate structure such as convexity, which guarantees that the algorithm, when it converges, has truly found the optimal solution.) The tool used in convex optimization to transform minimization problems into maximization problems is called "strong duality". It is particularly easy to implement when the problem is linear. When the problem is nonlinear, however, strong duality might not hold; even if it does, the dual of a nonlinear problem is more complicated to derive than that of a linear problem and cannot necessarily be expressed in closed form. Such closed-form expressions are critical for the robust problem to be tractable. This makes it significantly more difficult to incorporate robust optimization ideas into nonlinear problems. Earlier work (Ben-Tal and Nemirovski [1]) has focused on the case where the problem is *convex* in the decision variables but *affine* in the uncertainty.

<u>*Contributions*</u>

My main methodological focus lies in the field of **robust nonlinear optimization**, which has received little attention in the literature because of the issues outlined above. I consider the following broad classes of problems, each of which requires its own solution tools: (i) problems that are nonlinear and non-convex in the uncertainty, (ii) problems that are nonlinear but convex in the uncertainty, and (iii) problems that are linear in the uncertainty and nonlinear in the decision variables. My goals – and ultimately my contributions – are primarily to reformulate such nonlinear problems as **tractable mathematical models** and to provide **theoretical insights** into the optimal robust solution, in particular comparing it with its nominal counterpart and highlighting the impact of the problem parameters on the optimal strategy. While I have been able to solve many robust problems arising in my research using off-the-shelf software, the problem structure has sometimes required the development of problem-specific algorithms or heuristics, again to establish tractability.

In case (i), I have investigated in detail a special type of problems that are non-convex in the uncertain parameters: *piecewise linear* problems stemming from robust linear optimization with recourse in presence of right-hand-side uncertainty. Such problems arise for instance in revenue management with uncertain demand. (In a max-min setting, the function to be minimized is then concave, which is the "opposite" of the desirable property of convexity.) In Thiele et. al. [16], we show how to efficiently solve the robust problem to optimality using a *cutting-plane algorithm* because that problem has too many constraints to be tractable "as is"; this is in sharp contrast with robust problems commonly encountered in the literature, which are similar in size to their nominal counterparts and are solved using off-the-shelf software as linear programming problems or second-order cone programming problems. The cutting-plane algorithm represents a distinctive feature of our contribution. For a special case commonly encountered in practice, called simple recourse, we also present another algorithm that does not involve cutting planes and instead solves the problem as a series of linear programming problems. We also provide insights into the worst-case value of the uncertainty.

Case (ii), where the problem is nonlinear but convex in the uncertainty, has a key application in portfolio management and is described in detail below under "Application 1: Portfolio Management" for greater clarity. Case (iii), where the nonlinear problem is linear in the uncertain parameters, arises for instance in pricing under uncertainty; see "Application 2: Revenue Management." Revenue management also offers a broad class of settings for which robust optimization yields overly conservative solutions, in particular when there is only one source of uncertainty (one random product); robust optimization then plans for the very worst case in the range, which is not a realistic behavior to expect from a decision-maker. This calls for a data-driven optimization approach, also described under "Application 2."

My results make important methodological contributions to the field of robust nonlinear optimization by showing how some robust nonlinear problems can be solved efficiently without linearization. They also offer a systematic approach to characterizing the optimal robust strategy and comparing it with its nominal counterpart.

Application 1: Portfolio management

<u>*Background*</u>

I now describe my research in **robust portfolio management**, which I conduct in support of the Financial Engineering thrust area in my department, a program for which I am the leading researcher at this time. My work builds on the well-established Log-Normal model of asset prices while addressing its limitations. This model, which has become a cornerstone of portfolio management, assumes that the logarithms of the stock returns obey Normal distributions; it then leads to elegant closed-form formulas, in particular in option pricing (Black and Scholes [3]). Empirical studies, however, suggest that the true distributions of the logarithms of the stock returns have *fatter tails* than what the Normal distribution suggests, i.e., decision-makers assuming a Log-Normal model for asset prices underestimate the risk of tail (extreme) events. Unfortunately, no other distribution has emerged as a consistently better fit for all stocks. See, e.g., Jansen and de Vries [4]. This, coupled with managers' risk aversion, makes robust optimization an *excellent candidate* to protect the investor's portfolio against adverse events in an intuitive, tractable manner.

Previous robust optimization approaches, however, have modeled uncertainty at the level of the stock returns, instead of their logarithms (also called *continuously compounded rates of return*), in spite of the latter being more accurate drivers of uncertainty. This distinction matters because different functional forms change the worst-case value of the uncertainty and ultimately the optimal solution. The earlier model, although simplistic with regard to stock price behavior, was appealing because of its more obvious tractability. When uncertainty is incorporated at the level of the stock returns, the max-min formulation is *bilinear* and strong duality can be applied easily to produce a maximization problem using a *simple* set of rules. When uncertainty is incorporated at the level of the logarithms of the stock returns, which matches real-life stock price behavior more closely, the max-min formulation is *linear* in the portfolio allocation but *nonlinear* in the uncertain parameters.

<u>*Contributions*</u>

The robust optimization approach I propose, which I call Log-robust portfolio management, models the logarithms of the stock returns – previously assumed to obey Normal distributions – as uncertain parameters belonging to an uncertainty set, and then maximizes the investor's worst-case wealth over that set. The problem includes constraints on the asset allocation, such as a budget constraint. While the *nonlinearities* that arise from this setup complicate matters substantially compared to the linear case, one of the key conclusions of my work is that the Log-robust portfolio management approach does remain tractable.

In fact, the robust optimization problem is ultimately **linear** when short sales are not allowed, as we show in Kawas and Thiele [5]. ("Short sales" refers to the practice of selling shares the manager does not yet own, in the hope the price will decrease before he needs to acquire the actual shares. The manager makes money from the difference between the selling and buying prices.) The uncertainty sets we use are polyhedral, with range forecasts for each uncertain parameter and an additional constraint limiting the number of independent uncertain parameters that can simultaneously take their worst-case value (see Bertsimas and Sim [2]); the results do not hold for the other type of uncertainty sets commonly encountered in the literature, namely, ellipsoids. When short sales are allowed, the problem is no longer convex in the uncertainty, but can be reformulated as one if stock prices are independent, in which case the optimal robust solution is found by solving a series of linear programming problems, and of convex programming problems with one variable; the case with correlated assets is addressed via heuristics (Kawas and Thiele [6]). The approach can also be extended to incorporate risk-return tradeoffs, i.e., minimizing downside risk while imposing a constraint on the future nominal wealth of the investor, in the spirit of Markowitz [8]; the resulting problem is no longer linear but is a convex programming problem for which we discuss efficient implementation strategies (Kawas and Thiele [7]).

In addition to maintaining tractability, the Log-robust optimization approach offers many **theoretical insights** into the optimal strategy. For instance, when assets are independent and short sales are not allowed, we show in Kawas and Thiele [5] that there exists an index *j* such that it is optimal to invest in the stocks with the *j* highest nominal returns, in amounts inversely proportional to the volatility of their continuously compounded rates of return. In particular, solving the robust optimization problem in this case provides the manager with the index *j*, which measures the degree of diversification of the portfolio; the allocation is then determined uniquely. Allowing short sales in that setting introduces a second index *k*, such that it is optimal to invest in stocks 1 to *j* and short sell stocks *k* to *n*, with *n* the total number of stocks (see [6]). The fact that the robust policy is quite intuitive hopefully encourages managers to implement it in real-life settings. In addition, many insights are available regarding the impact of the decision-maker's aversion to ambiguity, measured by the size of the uncertainty set, on portfolio diversification.

My work on robust portfolio management offers an avenue for new research that has broad appeal because of its algorithmic tractability and of the numerous theoretical insights it offers.

Application 2: Revenue management

Background

The robust nonlinear problems I focus on in revenue management are: (i) nonlinear pricing problems with capacitated resources, which are linear in the demand uncertainty, and as such are well-suited for a *robust optimization approach*, and (ii) piecewise linear problems that arise in inventory control and have a limited number of sources of randomness (e.g., one product with stochastic demand), for which robust optimization yields conservative solutions and *data-driven optimization* is better suited.

Pricing problems under capacity offer a specific set of challenges in robust optimization because the uncertainty affects constraint (capacity) and objective (revenue) in two opposite manners: (i) the worst-case revenue is achieved when the demand is *less* than its nominal value, so the capacity constraint is under-utilized, but (ii) protecting the constraint against uncertainty means guaranteeing feasibility for any demand in the uncertainty set, including the best case of *high demand*. The primary domain of application is the service industry, especially airlines and hotels. Because these two cases of higher and lower demand than expected cannot happen at the same time, the decision-maker maximizing his worst-case revenue but unwilling to be overly conservative will naturally consider the possibility of overbooking in setting prices. (This does not necessarily mean the manager will let overbooking happen in practice; the reservation system can for instance stop sales when capacity is achieved.)

Contributions

A novelty of my approach in robust pricing management is that, instead of defining the uncertainty set by limiting the number of independent parameters that can take their worst-case value, as is common in the literature, I consider an *overbooking budget*, which captures the amount by which the decision-maker is willing to use the constraint beyond its capacity. In Thiele [14], I establish the existence of a *reference price* for the product, which plays a critical role in understanding how uncertainty affects the optimal strategy. In particular, I derive conditions describing when it is optimal for the manager to *increase* prices from their nominal levels when uncertainty is present. I extend this work to the multi-product case in Thiele [15].

My focus in **data-driven revenue management** has been on designing algorithms for exogenous, *non-stationary* demand processes, which have received little attention in the literature but represent an important class of real-life processes. When demand obeys the same distribution from period to period, it is of course optimal to gather as much information as possible and keep all historical data. When demand exhibits an upward or downward trend, however, the oldest data points become obsolete and should be discarded. But if demand is cyclical, data points corresponding to the same phase in the cycle should be kept, including those that were observed many time periods ago. My then-doctoral student and I have investigated two types of algorithms, to match each of these problem settings. In the first one, we adjust the size of the sample size adaptively for non-stationary processes with a trend (Metan and Thiele [9]). In the second one, we cluster data points according to their estimated phase in the cycle; this latter algorithm outperforms the traditional Holt-Winters method for seasonal demand (Metan and Thiele [10], [11]).

While data-driven optimization does not require estimating underlying probability distributions, it also has disadvantages. In particular, fitting data to a distribution allows **unobserved extreme events** to be incorporated in the decision-making process; relying on past data *alone* will under-protect the system against adverse outcomes that have yet to occur. To address this issue, we have considered a hybrid approach adding a correction term to the optimal data-driven order quantity in the newsvendor problem. This correction term uses a range for unobserved demand values and goes to zero as the number of data points increases (Metan and Thiele [12]).

In the same line of thought, we have studied the impact of incorporating both scenarios and range forecasts, where scenarios reflect high, medium or low demand, and demand in each case is given a separate range, in inventory management (Metan and Thiele [13]). This work provides a closed-form expression of the critical probability of the baseline (medium) scenario, at which the manager switches between predefined optimal strategies. (The other two scenarios have equal probability. This does not require the demand to be symmetric.) The proposed tri-scenario framework, when the probability of the baseline scenario exceeds the threshold value, achieves a significant cost decrease compared to the traditional robust approach applied to the full demand support. This is because the optimal order no longer depends on the whole demand support, but only on the demand range in the middle scenario. The baseline probability is expressed in terms of the cost parameters. Incidentally, this result could be used to motivate a common modeling technique in robust optimization, which is to use range forecasts that are a little smaller than the true support of the random variables.

Recognition

I have received various kinds of recognition for my research, from academia, industry and funding agencies. I am the single Principal Investigator on **two National Science Foundation grants**, one funding my work on robust portfolio management ("Robust Portfolio Management With Uncertainty Compounded Rates of Returns", Grant CMMI-0757983, 2008-2011, $199,964) and the other funding my work on robust revenue management ("Robustness and Performance in Data-Driven Revenue Management", Grant CMMI-0540143, 2005-2009, $262,000).

In addition, I received an **IBM Faculty Award** in Service Sciences, Management and Engineering ($34,000) in 2007, based on a project proposal on robust and adaptive lending. I am also the co-Principal Investigator on a project in data-driven adaptive forecasting and inventory control sponsored by the

CELDi – Center for Engineering Logistics and Distribution – **academia-industry consortium** ($99,981, including $44,516 for Lehigh University, the rest going to Dr. Kevin Taaffe at Clemson University).

Conclusions

My research has taken exciting directions over the past five years, with my main contributions centering on robust nonlinear optimization problems, including theoretical insights and computational tractability. I am particularly enthusiastic about my growing involvement in the financial engineering program at Lehigh. My interactions with local companies, especially those that contribute projects to the Master of Science students in Analytical Finance (which I supervise), have helped me gain a better understanding of the problems faced by practitioners and give my research additional relevance. In future work, I would like to incorporate *discrete* events revealed over time, such as a customer defaulting on his debt or a bond rating downgrade. I suspect this will lead me to integrate robust and stochastic techniques in one unified framework. I plan to keep developing close collaborations with industry leaders to explore robust optimization models that are well-suited to the information at hand, and that require novel techniques because of the underlying structure of the uncertainty. I believe this area offers many avenues for novel methodologies and will ultimately help managers make better decisions.

References

[1] Ben-Tal, A., and A. Nemirovski. "Robust convex optimization." *Mathematics of Operations Research* 23(4): 769-805, 1998.

[2] Bertsimas, D., and M. Sim. "The price of robustness." *Operations Research* 52(1): 35-53, 2004.

[3] Black, F., and M. Scholes. "The pricing of options and corporate liabilities." *Journal of Political Economy* 81: 637-659, 1973.

[4] Jansen, D., and C. de Vries. "On the frequency of large stock returns: putting booms and busts in perspective." *The Review of Economics and Statistics* 73(1): 18-24, 1991.

[5] Kawas, B., and A. Thiele. "A Log-robust optimization approach to portfolio management." *OR Spectrum* to appear, 2009.

[6] Kawas, B., and A. Thiele. "Short sales in Log-robust portfolio management." Technical Report, Lehigh University, submitted to *Computers & Operations Research*, Bethlehem, PA, 2008.

[7] Kawas, B., and A. Thiele. "Risk-return tradeoffs in Log-robust portfolio management." Technical Report, Lehigh University, Bethlehem, PA, submitted to *European Journal of Operational Research*, 2009.

[8] Markowitz, H. *Portfolio selection: efficient diversification of investments.* John Wiley & Sons, New York, 1959.

[9] Metan, G., and A. Thiele. "An adaptive algorithm for the optimal sample size in the non-stationary data-driven newsvendor problem." Proceedings of the 10th INFORMS Computing Society conference, E. Baker, A. Joseph, A. Mehrotra and M. Trick, editors, *Extending the Horizons: Advances in Computing, Optimization, and Decision Technologies*, pp.77-96, Springer, New York, 2007.

[10] Metan, G., and A. Thiele. "A dynamic and data-driven approach to the newsvendor problem under seasonal demand." in P. Pardalos, K. Furman and W. Chaovalitwongse, editors, *Optimization and Logistics Challenges in the Enterprise*, pp.277-304, Springer, New York, 2009.

[11] Metan, G., A. Thiele, "Integrated forecasting and inventory control for seasonal demand," Proceedings of the 11th INFORMS Computing Society conference, J. Chinneck, B. Kristjansson and M. Saltzman, editors, *Operations Research and Cyber-Infrastructure*, pp.427-441, Springer, New York, 2009.

[12] Metan, G., and A. Thiele. "Protecting the data-driven newsvendor against rare events: a correction-term approach." Accepted to *Algorithmic Operations Research* subject to minor revision, 2009.

[13] Metan, G., and A. Thiele. "Scenario design in robust inventory management." Technical Report, Lehigh University, Bethlehem, PA, under revision for second round at *European Journal of Operational Research*, 2008.

[14] Thiele, A. "Single-product pricing via robust optimization." Technical report, Lehigh University, Bethlehem, PA, under revision for second round at *Operations Research*, 2006.

[15] Thiele, A. "Multi-product pricing via robust optimization." *Journal of Revenue and Pricing Management* 8(1): 67-80, 2009.

[16] Thiele, A., T. Terry, and M. Epelman. "Robust linear optimization with recourse." Technical Report, Lehigh University, Bethlehem, PA, submitted to *Naval Research Logistics*, 2009.