Risk Dynamics in Trade Networks

Rafael M. Frongillo
Harvard University
raf@cs.berkeley.edu

Mark D. Reid
The Australian National University & NICTA
mark.reid@anu.edu.au

Abstract

We introduce a new framework to model interactions among agents which seek to trade to minimize their risk with respect to some future outcome. We quantify this risk using the concept of risk measures from finance, and introduce a class of trade dynamics which allow agents to trade contracts contingent upon the future outcome. We then show that these trade dynamics exactly correspond to a variant of randomized coordinate descent. By extending the analysis of these coordinate descent methods to account for our more organic setting, we are able to show convergence rates for very general trade dynamics, showing that the market or network converges to a unique steady state. Applying these results to prediction markets, we expand on recent results by adding convergence rates and general aggregation properties. Finally, we illustrate the generality of our framework by applying it to agent interactions on a scale-free network.

1 Introduction

The study of dynamic interactions between agents who each have a different stake in the world is of broad interest, especially in areas such as multiagent systems, decision theory, and economics. In this paper, we present a new way to model such dynamic interactions, based on the notion of risk measures from the finance literature.

The agents in our model will each hold a *position*, which states how much the agent stands to gain or lose for each possible outcome of the world. The overall outlook of an agent's position will be quantified by their *risk measure*, which simply captures the "riskiness" of their position. To minimize their risks, agents change their positions by trading *contingent contracts* amongst themselves; these contracts state that the owner is entitled to some amount of money which depends on this future outcome. Traders can be thought of as residing on a *network*, the edges or hyperedges of which dictate which agents can trade directly.

This simple setting gives rise to several natural questions, which we would like to understand: Given a set of agents with initial positions, can a stable equilibrium be found, where no agents can trade further for mutual benefit? If such an equilibrium exists, can the agents converge to it using a trading protocol, and if so what is the rate of convergence? How does the structure of the underlying network affect change these answers? What is the distribution of the agents' risks at equilibrium, and how does an agent's final risk depend on his position in the network? This paper addresses and provides answers to many of these questions.

Our model is heavily inspired by the work of Hu and Storkey (2014), who use risk-measure agents to draw connections between machine learning and prediction markets. Another motivation comes from Abernethy et al. (2014), who study a prediction market setting with risk-averse traders whose beliefs over the outcomes are members of an exponential family of distributions. Both papers analyze the steady-state equilibrium of the market, leaving open the question of how, and how fast, the market may arrive at that equilibrium. In fact, both papers specifically point to rates and conditions for convergence in their future work.

The contributions of this paper are threefold. First, we develop a natural framework to model the interactions of networked agents with outcome-contingent utilities, which is tractable enough to answer many of the questions posed above. Second, by showing that our trading dynamics can be recast as a randomized coordinate descent algorithm, we establish convergence rates for trading networks and/or agent models which are more general than the two prediction market papers above. Third, along the way to showing our rates, we adapt and generalize existing coordinate descent algorithms from the optimization literature, *e.g.* Nesterov (2012) and Richtárik and Takáč (2014), which may be of independent interest.

2 Setting

Let Ω be a finite set of possible outcomes. Following (Föllmer and Schied, 2004), a *position* is simply a function from outcomes to the reals, $X : \Omega \to \mathbb{R}$. Positions can be thought of as random variables which are intended to represent outcome-contingent monetary values. Denote by $\mathbb{1} : \Omega \to \mathbb{R}$ the constant position with $\mathbb{1} : \omega \mapsto 1$. The set of all positions under consideration will be denoted $\mathcal{X}$ and will be assumed to be closed under linear combination and contain at least all the outcome-independent positions $\{\alpha\mathbb{1} : \alpha \in \mathbb{R}\}$. We will denote by Δ the set of probability distributions over Ω, namely $\Delta = \{p \in [0,1]^{\Omega} \colon \langle p, \mathbb{1}\rangle = 1\}$, where $\langle p, x\rangle = \sum_{\omega\in\Omega} p(\omega)x(\omega)$ is the inner product. Note that $\langle p, X\rangle = \mathbb{E}_{\omega\sim p}\left[X(\omega)\right]$, the mean under p.

When viewed as a vector space in $\mathbb{R}^{\Omega}$, the set of positions

$\mathcal{X}$ introduced above is a subspace of dimension $k \leq |\Omega|$. Hence, it must have a basis of size k, and thus we must have some $\phi : \Omega \to \mathbb{R}^k$ with the property that for all $X \in \mathcal{X}$, there is some $r \in \mathbb{R}^k$ such that $X(\omega) = r \cdot \phi(\omega) = \sum_i r_i \phi(\omega)_i$ for all $\omega \in \Omega$.

We will make extensive use of this "compressed" form of $\mathcal{X}$, which we denote by $\mathcal{R} = \mathbb{R}^k$. Define the counterpart $X[r] \in \mathcal{X}$ of $r \in \mathcal{R}$ to be the position $X[r] : \omega \mapsto r \cdot \phi(\omega)$. The presence of outcome-independent positions then translates into the existence of some $r^\$ \in \mathcal{R}$ satisfying $X[r^\$] = \mathbb{1}$. Finally, we denote by $\Pi := \mathrm{conv}(\phi(\Omega))$ the convex hull of the basis function ϕ.

As intuition about ϕ and Π, it is helpful to draw analogy to the setting of prediction markets. As we will see in Section 4, the function ϕ can be thought of as encoding the payoffs of each of k outcome-contingent contracts, or *securities*, where contract i pays $\phi(\omega)_i$ for outcome ω. The space Π then becomes the set of possible beliefs $\{\langle p, \phi\rangle : p \in \Delta\}$ of the expected value of the securities.

Risk Measures

Following Hu and Storkey (2014), agents in our framework will each quantify their uncertainty in positions via a (convex monetary) *risk measure* $\rho : \mathcal{R} \to \mathbb{R}$ satisfying, for all $r, r' \in \mathcal{X}$:

- *Monotonicity*: $\forall \omega\ X[r](\omega) \leq X[r'](\omega) \Rightarrow \rho(r) \geq \rho(r')$.
- *Cash invariance*: $\rho(r + c \cdot r^\$) = \rho(r) - c$ for all $c \in \mathbb{R}$.
- *Convexity*: $\rho(\lambda r + (1-\lambda) r') \leq \lambda\rho(r) + (1-\lambda)\rho(r')$ for all $\lambda \in [0, 1]$.
- *Normalization*: $\rho(0) = 0$.

The reasonableness of these properties is usually argued as follows (see, *e.g.*, (Föllmer and Schied, 2004)). Monotonicity ensures that positions that result in strictly smaller payoffs regardless of the outcome are considered more risky. Cash invariance captures the idea that if a guaranteed payment of $\$c$ is added to the payment on each outcome then the risk will decrease by $\$c$. Convexity states that merging positions results in lower risk. Finally, normalization is for convenience, stating that a position with no payout should carry no risk.

In addition to these common assumptions, we will make two regularity assumptions:

- *Expressiveness*: ρ is everywhere differentiable, and $\mathrm{closure}\{\nabla\rho(r) : r \in \mathcal{R}\} = \Pi$.
- *Strict risk aversion*: the convexity inequality above is strict unless $r - r' = \lambda r^\$$ for some $\lambda \in \mathbb{R}$.

Expressiveness is related to the dual formulation given below; roughly, it says that the agent must take into account every possible distribution over outcomes when calculating the risk. Strict risk aversion says that an agent should strictly prefer a mixture of positions, unless of course the difference is outcome-independent.

A key result concerning convex risk measures is the following representation theorem (cf. Föllmer and Schied (2004, Theorem 4.15), Abernethy, Chen, and Vaughan (2013, Theorem 3.2)).

Theorem 1 (Convex Risk Representation)**.** *A functional $\rho : \mathcal{R} \to \mathbb{R}$ is a convex risk measure if and only if there is a closed convex function $\alpha : \Pi \to \mathbb{R} \cup \{\infty\}$ such that*

$$\rho(r) = \sup_{\pi \in \mathrm{relint}(\Pi)} \langle \pi, -r\rangle - \alpha(\pi). \tag{1}$$

Here $\mathrm{relint}(\Pi)$ denotes the relative interior of Π, the interior relative to the affine hull of Π. Notice that if f^* denotes the convex conjugate $f^*(y) := \sup_x \langle y, x\rangle - f(x)$, then this theorem states that $\rho(r) = \alpha^*(-r)$. This result suggests that the function α can be interpreted as a *penalty function*, assigning a measure of "unlikeliness" $\alpha(\pi)$ to each expected value π of the securities defined above. Equivalently, $\alpha(\langle p, \phi\rangle)$ measures the unlikeliness of distribution p over the outcomes. We can then see that the risk is the greatest expected loss under each distribution, taking into account the penalties assigned by α.

Risk-Based Agents

We are interested in the interaction between two or more agents who express their preferences for positions via risk measures. Burgert and Rüschendorf (2006) formalise this problem by considering N agents with risk measures ρ_i for $i = 1, \ldots, N$ and asking how to split a position $r \in \mathcal{R}$ in to per-agent positions r_i satisfying $\sum_i r_i = r$ so as to minimise the total risk $\sum_i \rho_i(r_i)$. They note that the value of the total risk is given by the *infimal convolution* $\wedge_i \rho_i$ of the individual agent risks — that is,

$$(\wedge_i \rho_i)(r) := \inf\left\{\sum_i \rho_i(r_i) : \sum_i r_i = r\right\}. \tag{2}$$

A key property of the infimal convolution, which will underly much of our analysis, is that its convex conjugate is the sum of the conjugates of its constituent functions. See *e.g.* Rockafellar (1997) for a proof.

$$(\wedge_i \rho_i)^* = \sum_i \rho_i^*\ . \tag{3}$$

Hu and Storkey (2014) identify a special, market making agent with risk ρ_0 that aims to keep its risk constant rather than minimising it. The risk minimising agents trade with the market maker by paying the market maker $\rho_0(-r)$ dollars in exchange for receiving position r, thus keeping the market maker's risk constant. We will revisit these special constant-risk interactions in Section 4. For now, we will consider quite general trading dynamics.

3 Trade Dynamics

We now describe how agents may interact with one another, by introducing certain dynamics of trading among agents. Recall that we have N agents, and each agent i is endowed with a risk measure ρ_i. We further endow agent i with an initial position $r_i^0 \in \mathcal{R}$, and let $r^0 = \sum_i r_i^0$. We will start time at $t = 0$ and denote the position of trader i at time t by r_i^t.

A crucial concept throughout the paper is that of surplus. Given a subset of the agents willing to trade among themselves, we can quantify the total net drop in risk that group can achieve.

Definition 1. *Given $r_S = \{r_i\}_{i \in S}$ for some subset of agents S, the S-*surplus *of r is the function $\Phi : \mathcal{R}^{|S|} \to \mathbb{R}$ defined by $\Phi_S(r_S) = \sum_{i \in S} \rho_i(r_i) - (\wedge_i \rho_i)(\sum_{i \in S} r_i)$. In particular, $\Phi(r) := \Phi_{[N]}(r)$ is the* surplus function.

We now define trade functions, which are efficient in the sense that all of this surplus is divided, perhaps unevenly, among the agents present. A trade dynamic will then be simply a distribution over trade functions.

Definition 2. *Given some subset of nodes $S \subseteq [N]$, we say a function $f : \mathcal{R}^N \to \mathcal{R}^N$ is a* trade function on S *if*

1. $\sum_{i \in S} f(r)_i = \sum_{i \in S} r_i$,
2. *the S-surplus is allocated, meaning $\Phi_S(f(r)_S) = 0$,*
3. *for all $j \notin S$ we have $f(r)_j = r_j$.*

The following result shows that trade functions have remarkable structure: once the subset S is specified, the trade function is completely determined, up to cash transfers. In other words, the surplus is removed from the position vectors, and then it is redistributed as cash to the traders, and the choice of trade function is merely in how this redistribution takes place. The fact that the derivatives match has strong intuition from prediction markets: agents must agree on the price. Note that all proofs may be found in the full version of the paper Frongillo and Reid (2014).

Theorem 2. *The trade functions on any $S \subseteq [N]$ are unique up to zero-sum cash transfers. Moreover, if f is a trade function on S, then $\nabla \rho_i(f(r)_i) = \pi^*_S$ for all i, where $\pi^*_S = \min_{\pi \in \Pi} \sum_{i \in S} \alpha_i(\pi) - \langle \pi, \sum_{i \in S} r_i \rangle$.*

Our notion of trade dynamics, defined below, is quite intuitive — predefined groups of agents S_i gather at random to negotiate a trade which minimizes their total risk, subject to the constraint that trading may only be among those gathered.

Definition 3. *Given m subsets $\mathcal{S} = \{S_i\}_{i=1}^m$ and m trade functions f_i on S_i, and a distribution $p \in \Delta_m$ with full support, a* trade dynamic *is the randomized algorithm which selects f_i with probability p_i and takes $r^{t+1} = f_i(r^t)$. A* fixed point *r of the trade dynamic is a point with $f_i(r) = r$ for all $i \in [m]$.*

We now give a few natural instantiations of trade dynamics which we will use throughout the paper. Let G be a directed graph with a node for each agent. An *edge dynamic* has a trade function $f_{(i,j)}$ on $\{i, j\}$ for each edge (i, j) in G, where if $r' = f_{(i,j)}(r)$ we have $\rho_j(r'_j) = \rho_j(r_j)$ and $\rho_i(r'_i) = \rho_i(r_i) - \Phi_{\{i,j\}}(r_{\{i,j\}})$. In other words, the agents minimize their collective risks, but agent i takes all of the surplus. Similarly, a *node dynamic* has a trade function f_i for each node $i \in [N]$, on $S_i = \{j : (i, j) \in E(G)\} \cup \{i\}$, the out-neighborhood of i, and $r' = f_i(r)$ satisfies $\rho_j(r'_j) = \rho_j(r_j)$ for $j \in S_i \setminus \{i\}$ while $\rho_i(r'_i) = \rho_i(r_i) - \Phi_{S_i}(r_{S_i})$.

A third dynamic we will consider uses a notion of fairness; call a trade function f on S *fair* if it satisfies $\rho_i(f(r)_i) = \rho_i(r_i) - \frac{1}{|S|}\Phi_S(r_S)$ for all $i \in S$. Then a *fair trade dynamic* is simply a mixture of fair trade functions. Returning to the graph theme, we may define fair versions of the node and edge dynamics above, in the natural way.

For all these types of trade dynamics, we will see that the only crucial property is that of *connectedness*, which ensures that trades can eventually travel from one agent to any other. Given this property, we show a quite general equilibrium result.

Definition 4. *A trade dynamic with subsets $\mathcal{S}$ is* connected *if the hypergraph with nodes $[N]$ and hyperedges $\mathcal{S}$ is a connected hypergraph.*

Theorem 3. *Let $\pi^* = \min_{\pi \in \Pi} \sum_i \alpha_i(\pi) - \langle \pi, r^0 \rangle$. There exists $r^* \in \mathcal{R}^N$ such that for all connected trade dynamics D, the unique fixed point of D is r^*, up to zero-sum cash transfers. Moreover, $\Phi(r^*) = 0$ and $\nabla \rho_i(r^*_i) = \pi^*$ for all i.*

The result of Theorem 3 is somewhat surprising — not only is there a unique equilibrium (up to cash transfers) for all connected dynamics, but all connected dynamics have the *same* equilibrium! If one restricts to connected graphical networks, this means that the equilibrium does not depend on the network structure. The power of our framework is that the equilibrium analysis holds regardless of the way agents interact, as long as information is allowed to spread to all agents eventually. In fact, one could even consider an arbitrary process choosing subsets S^t of agents to trade at each time t; if the set $\mathcal{S}$ of subsets which are visited infinitely often yields a connected hypergraph, then the proof Theorem 3 still applies.

Now that the existence of an equilibrium has been established, we turn to the question of convergence. The proof of our rates is in the full version; the technique is to show that our trade dynamics are performing a type of randomized coordinate descent algorithm, where the coordinate subspaces correspond to subsets S of agents, and then use standard techniques to analyze it. Let us briefly see why coordinate descent is a useful analogy for our dynamics. Recall that we have m subsets of agents S_i, and that each trade function f_i only modifies the positions of agents in S_i. Thinking of $(r_1, \ldots, r_N)$ as a large Nk vector (recall $\mathcal{R} = \mathbb{R}^k$), the trade function f_i is thus modifying only $|S_i|$ blocks of k entries. Moreover, f_i is *minimizing* the sum of the risks of agents in S_i. Hence, ignoring for now the constraint that the sum of the positions remain constant, f_i is performing a block coordinate descent step of the surplus function Φ on this block of coordinates.

Theorem 4. *For any connected trade dynamic, we have $\mathbb{E}[\Phi(r^t)] = O(1/t)$.*

Amazingly, Theorem 4 holds for *all* connected trade dynamics, as they each minimize the surplus in whichever S_i is chosen, and that is enough for the coordinate descent bounds to apply. In fact, it is more than enough: as the proof technique only tracks the optimality gap, the rates extend to less efficient trade dynamics, as long as they achieve a drop in surplus within a constant factor of an optimal dynamic. This suggests that our convergence results are robust with respect to the model of rationality one employs; if agents have bounded rationality and cannot compute positions which would exactly minimize their risk, but instead approximate it within a constant factor of the gradient update, the rate remains $O(1/t)$.

4 Application to Prediction Markets

Our analysis was motivated in part by work that considered the equilibria of prediction markets with specific models of trader behavior: traders as risk minimizers (Hu and Storkey, 2014); and traders with exponential utilities and beliefs from exponential families (Abernethy et al., 2014). In both cases, the focus was on understanding the properties of the market at convergence, and questions concerning whether and how convergence happened were left as future work. We now explain how this earlier work can be seen as a special case of our analysis with an appropriate choice of network structure and dynamics. In doing so we also generalize several earlier results.

Following Abernethy, Chen, and Vaughan (2013), a cost function-based prediction market consists of a collection of k outcome-dependent *securities* $\{\phi(\cdot)_i\}_{i=1}^k$ that pay $\phi(\omega)_i$ dollars should outcome $\omega \in \Omega$ occur. A *market maker* begins with an initial position $r^0 \in \mathcal{R}$, the *liability vector*, and a *cost function* $C : \mathcal{R} \to \mathbb{R}$. A trader who wishes to purchase a bundle of securities $r \in \mathcal{R}$ is charged $\text{price}(r) := C(r^t + r) - C(r^t)$ by the market maker which then updates its liability to $r^{t+1} = r^t + r$. The desirable properties for cost functions are quite different from those of risk measures (e.g. information incorporating, arbitrage-free), yet as observed by Hu and Storkey (2014), the duality-based representation of cost functions is essentially the same as the one for risk measures (compare Theorem 1 and (Abernethy, Chen, and Vaughan, 2013, Theorem 5)). In essence then, cost functions *are* risk measures, though because liability vectors measure losses and position vectors measure gains, we simply have $\rho_C(r) = C(-r)$.

In the prediction market of Hu and Storkey (2014), agents have risk measures ρ_i and positions r_i. A trade of r between such and agent a market maker with cost function C and position r^t makes the agent's new risk $\rho_i(r_i + r - \text{price}(r) \cdot r^\$)$ since the market maker charges $\text{price}(r)$ dollars for r. Similarly, one can check that the market maker's risk remains constant for all trades of this form.

An agent minimizing its risk implements the *trading function* (Definition 2) $f : (-r^t, r_i) \mapsto (-r^t - r, r_i + r)$ since $\min_r \rho_i(r_i + r - (C(r^t + r) - C(r^t)) \cdot r^\$) = \min_r \rho_i(r_i + r) + \rho_C(-r^t - r)$ by cash invariance of ρ_i, guaranteeing the surplus between the agent and market maker is zero. Thus, one could think of agents in a risk-based prediction market as residing on a star graph, with the market maker in the center. By Theorem 3, any trade dynamic which includes every agent with positive probability will converge, and Theorem 4 gives an $O(1/t)$ rate of convergence.

An important special case is where agents all share the same base risk measure ρ, but to different degrees b_i which intuitively correspond to a level of *risk affinity*. Specifically, let $\rho_i(r) = b_i \rho(r/b_i)$, where a higher b_i corresponds to a more risk-seeking agent.[1] As we now show, the market equilibrium gives agent i a share of the initial sum of positions r^0 proportional to his risk affinity, and the final "consensus" price of the market is simply that of a scaled version of r^0.

[1]Note however that agents are still risk-averse; only in the limit as $b \to \infty$ do the traders become risk-neutral.

Theorem 5. *Let ρ be a given risk measure, and for each agent i choose an initial position $r_i^0 \in \mathcal{R}$ and risk defined by $\rho_i(r_i) = b_i \rho(r_i/b_i)$ for some $b_i > 0$. Let $r^0 = \sum_i r_i^0$, and define $r \in \mathcal{R}^N$ by $r_i = b_i r^0 / \sum_j b_j$. Then r is the unique point up to zero-sum cash transfers such that $\Phi(r) = 0$. Moreover, r satisfies for all i,*

$$\nabla \rho_i(r_i) = \nabla \rho \left(r^0 / \textstyle\sum_j b_j \right) . \tag{4}$$

This result generalizes those in §5 of Abernethy et al. (2014), where traders are assumed to maximize an expected utility of the form $U_b(w) = -b \exp(-w/b)$ under beliefs drawn from an exponential family with sufficient statistic given by the securities ϕ. The above result shows that exactly the same weighted distribution of positions at equilibrium occurs for *any* family of risk-based agents, not just those derived from exponential utility via certainty equivalents (Ben-Tal and Teboulle, 2007). In addition, this generalization shows that the agents need not have exponential family beliefs: their positions r_i act as general natural parameters, and $1/b_i$ acts as a general measure of risk aversion. Finally, this connection also means our analysis applies to their setting, addressing their future work on dynamics and convergence.

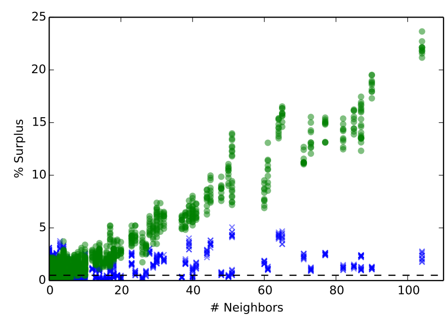

Figure 1: Percentage of captured surplus per trader vs. number of trading neighbors for fair edge dynamic (green circles) and fair node dynamic (blue crosses). The dashed black line shows the fair distribution for 200 agents (0.5%).

5 Conclusions

We have developed a framework to analyze arbitrary networks of risk-based agents, giving a very general analysis of convergence and rates, and addressing open issues in both Hu and Storkey (2014) and Abernethy et al. (2014). We view this as a foundation, which opens more questions than it answers. For example, can we improve the asymptotic rates of convergence? One potential technique would be to show that trading never leaves a bounded region, and carefully applying bounds for strongly convex functions (modulo the $r^\$$ direction), which could give a rate as fast as $O(1/2^t)$. An even broader set of questions has to do with the distribution of risk — how does the network topology effect the outcome on the agent level? As our experiments show (Fig. 1), even local properties of the network may have a strong effect on the final distribution of risks, and understanding this relationship is a very interesting future direction.

Acknowledgments

We would like to thank Matus Telgarsky for his generous help, as well as the lively discussions with, and helpful comments of, Sébastien Lahaie, Miro Dudík, Jenn Wortman Vaughan, Yiling Chen, David Parkes, and Nageeb Ali.

References

Abernethy, J.; Kutty, S.; Lahaie, S.; and Sami, R. 2014. Information aggregation in exponential family markets. In *Proceedings of the fifteenth ACM conference on Economics and computation*, 395–412. ACM.

Abernethy, J.; Chen, Y.; and Vaughan, J. W. 2013. Efficient market making via convex optimization, and a connection to online learning. *ACM Transactions on Economics and Computation* 1(2):12.

Ben-Tal, A., and Teboulle, M. 2007. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance* 17(3):449–476.

Burgert, C., and Rüschendorf, L. 2006. On the optimal risk allocation problem. *Statistics & decisions* 24(1/2006):153–171.

Föllmer, H., and Schied, A. 2004. Stochastic finance, volume 27 of de gruyter studies in mathematics.

Frongillo, R. M., and Reid, M. D. 2014. Risk dynamics in trade networks. *arXiv:1410.0413 [cs, math]*. arXiv: 1410.0413.

Hu, J., and Storkey, A. 2014. Multi-period trading prediction markets with connections to machine learning. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*.

Nesterov, Y. 2012. Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM Journal on Optimization* 22(2):341–362.

Richtárik, P., and Takáč, M. 2014. Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function. *Mathematical Programming* 144(1-2):1–38.

Rockafellar, R. 1997. *Convex analysis*. Princeton University Press.