# Information Differential Geometry of Incomplete Markets

Yuan Gao* and Kian Guan Lim† and Kah Hwa Ng‡

First Draft: July 4, 2001
This Draft: October 18, 2001

**Abstract**

We propose a new information geometrical framework for the study of pricing and portfolio optimization problems of incomplete markets. We show that the dual of the utility function is the entropy on the manifold of equivalent martingale measures. The optimal portfolio and arbitrage-free equilibrium pricing problems are transformed to a entropy optimization problem over the manifold. We then show the economic relations between the utility and the Riemannian structures of the manifold. We use stochastic volatility HJM model to illustrate the finite and infinite dimensional specifications of the incomplete market models.

**Keywords**: Incomplete market, information differential geometry, utility maximization, stochastic volatility HJM model

*JEL Classifications*:D52, G11, G13

*Mathematics Subject Classifications (2000)*:62P05, 91B24, 91B28

---

*Ph.D. candidate, Center for Financial Engineering, National University of Singapore, Kent Ridge Post Office, P.O. Box 1130, Singapore 911105. Email: cfegaoy@nus.edu.sg Fax: (065)8745430.
I would like to thank Prof. Yeneng Sun and Dr. Xiaoqing Liu for their helpful comments. I am also grateful to Prof. Robert Jarrow for his valuable comments. I am especially thankful to Prof. Hans Föllmer for his precious time and encouraging comments. Suggestions and comments are welcomed, all errors are ours.
†Professor, Singapore Management University

‡Associate Professor, Department of Finance and Accounting, National University of Singapore.

# 1 Introduction

This is the first paper of a series of our works on the geometry of incomplete markets.

The structure of the paper is as follows. We start by introducing the concept of entropy of the market which seems to have little relation to our pricing and portfolio optimization problems of an incomplete market. The relevance will be clear in the section presenting the utility maximization theorem. We then study the relations between the geometric structure and the utility structure of the market. We present some examples of specifications which reduce to the Föllmer-Schweizer minimal martingale measure ([7]). We also touch the infinite dimensional analysis for the general case.

## 1.1 The Main Problem

The problem that we are trying to tackle in this paper is the following. In an incomplete market, we want to know how to find an equivalent martingale measure of which the expectation can act as a pricing operator. The price should be arbitrage-free and correspond to some equilibrium such as maximum utility for a representative agent. We also want to know how to characterize such a arbitrage-free equilibrium equivalent martingale measure.

## 1.2 Brief Review of Results of the Incomplete Market

The idea of this paper is inspired by the works of Amari and Nagaoka (2000)[17], Frittelli (2000) [8], [9], Cvitanić, Schachermayer and Wang (2001)[5], Kramkov and Schachermayer (1999)[23].

The paper of Frittelli (2000) [9] stimulated our research in the geometry of equivalent martingale measures. Frittelli showed the existence and uniqueness of the minimal relative entropy measure. There are another two important and interesting results in the paper. One is the notion of utility martingale measures of which the Radon-Nikodym derivatives are proportional to the marginal utility of terminal wealth. Another is the relation between the relative entropy and the exponential utility. The relation provides a link between a mathematical quantity and an economical quantity. This enable us to ask the following question which drove our research.

> Is there any relationship between a given utility function and a general relative entropy function of the space of equivalent martingale measures of incomplete markets?

Another work of Frittelli (2000) [8] provides some insights to the above question. He used the concept of "generalized distance" between probability measures to show that the pricing measure for a given utility could be obtained by minimizing some distance in the space of probability measures. He gave examples, for exponential utility, the distance function is actually relative entropy, for Power HARA (hyperbolic absolute risk aversion) utility, the distance is the Kakutani-Hellinger distance, et al..

Either doing distance minimizing or general entropy maximizing, there is little use in finance without a financial economics reason. The reason is closely related to utility maximization and some rigorous treatments are in the works of Kramkov and Schachermayer (1999)[23], Cvitanić et al. (2001)[5], Schachermayer (1999) [27]. They showed that the utility maximization is equivalent to the conjugate function minimization in the space of Radon-Nikodym derivatives with respect to $P$. Their results do help us to provide a financial economical interpretation of our geometrical approach.

In a recent paper (which we only knew when we were finishing our paper) by Goll and Rüschendorf (2001) [12], they combine both the ideas of utility maximization and distance minimization in probability spaces. They show the insight between utility maximization and distance minimization of probability measures. They also remarked the homogenous properties in initial wealth of some most commonly used utility functions. Their definition of $f$-divergence is essentially the same as our definition of entropy. We realize similar results in a different setting and use it as an interpretation of the geometry of the incomplete market.

## 1.3 The Main Ideas of the Paper

We refer to figure 1 as an illustration of the geometry of an incomplete market.

The empirical measure $P$ is in the set of all measures of an incomplete market. The set $\mathcal{M}$ is the set of equivalent martingale measures.

The relation between utility and entropy is known in various forms in the works of Kramkov and Schachermayer(1999)[23], Cvitanić et al. (2001)[5], Schachermayer (1999) [27]. In the work of Goll and Rüschendorf (2001) [12], they study the distance between empirical measure and the martingale measure. They show that the measure $Q^*$ which minimize their "distance" from $Q$ to $P$ over $\mathcal{M}$ is the equilibrium pricing measure.

But they did not show how to find $Q^*$ in general. We study the problem of approximating $Q^*$ in a subset $\mathcal{S}$ of $\mathcal{M}$ which has a differential geometric structure.

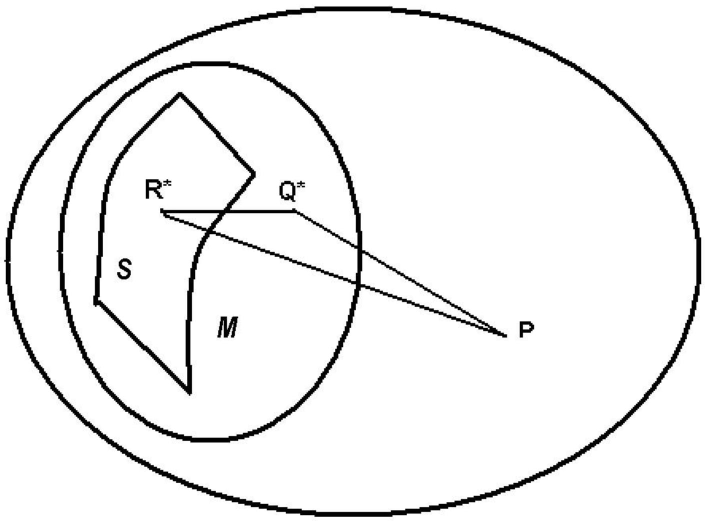


**Figure 1: Manifolds of EMM**

Illustration of the geometry of the incomplete markets.

In order to show that the dual relation still holds for the best approximation $R^*$, we prove it again under different conditions for some "nice" subset of the equivalent martingale measures.

Our purpose to study the finite dimensional subset is to seek for some analytical pricing formula for derivatives securities in an incomplete market.

To find $R^*$, it is important to understand the geometry of $\mathcal{S}$. So we focus on the distance between two martingale measures $R^*$ and $Q^*$. We also study the relation between the utility, risk aversion and the Riemannian structure of the manifold of equivalent martingale measures.

## 1.4 Our Approach and Focus

Our work may look similar to the works of Frittelli (2000) [8], Goll and Rüschendorf (2001) [12] at a first sight. But there are some fundamental modelling differences. Instead of seeking generality, we concentrate on some simpler subset of the space of all equivalent martingale measures $\mathcal{M}$. This subset has similar properties of the general abstract space.

First, instead of modelling a point in the space of $\mathcal{M}$ by the measure $Q$ itself, we model it using its Radon-Nikodym derivative with respect to the empirical measure, $dQ/dP$. So instead of discussing the "distance" between the $Q$ and $P$, we discussed the "entropy" of the point $dQ/dP$ and cross entropy between $dQ_1/dP$ and $dQ_2/dP$ which serves as a measure of distance between two martingale densities. The empirical measure $P$ is actually not in our space. By doing some parameterization on the market price of risk, the Radon-Nikodym density has the form of an exponential family which is well-studied in information geometry. Also the cross entropy between two Radon-Nikodym densities allows us to study the convergence to the maximum utility density and incomplete market equilibrium.

Second, our space is a subset of $\mathcal{M}$. In order to have a finite dimensional manifold, we have to restrict ourself to the densities with some form of incomplete market price of risk processes. Our optimization domain is much smaller and our manifold of equivalent martingale measures is not a convex density space. So the general results on convex probability spaces ( [12], [27]) can not be applied directly to our space. However, we can still get similar properties in our simpler subset by assuming smoothness regularities. So we reduce the domain of optimization from a general abstract space of $\mathcal{M}$ to a smaller manifold. In this simpler manifold, the incomplete market modelling becomes more concrete and a step nearer industrial applications.

Third, our set of admissible strategies are different from Kramkov and Schachermayer (1999)[23], Cvitanić et al. (2001)[5], Schachermayer (1999) [27]. We do not have the martingale representation property since our manifold is not the set of all equivalent measures (see Jacka (1992) [18] Theorem 3.4). We model the extra term as some profit to be taken away from the terminal wealth. By making use of the Utility and Entropy Maximization Theorems, we obtain the arbitrage-free equilibrium prices in the sense of Davis (1997 [6]) for homogenous differentiable entropy markets (definition 4.1).

Fourth, our manifold is "richer" in structure since we assume a differentiable manifold structure rather than the general weak-star topology. We realize that there is a natural Riemannian metric induced by the entropy. This structure in the manifold plays a crucial role in the optimization of entropy by affecting the gradient vectors. We also provide a geometrical characterization of the optimal Radon-Nikodym density and propose a gradient flow algorithm.

We focus on providing a link between an incomplete market pricing problem and an information geometrical problem. Our approach reveals the fact that differential geometric methods could be used to compute the optimal $dQ/dP$.

Brody and Hughston (2001) [1], [2] have used information geometry to study the positive interest rate term structures. Our study differs from Brody and Hughston (2001) [1], [2]. They study the discount bond densities using the Fisher-Rao metric.

We study the incomplete market Radon-Nikodym densities using the metric induced from the utility function. In particular, the metric induced by exponential utility is the Fisher-Rao metric.

We would also like to point out that the study of manifolds of the Radon-Nikodym densities is different from the one of stochastic processes on manifolds or manifolds of forward rate curves.

# 2 Model of the Incomplete Financial Market

## 2.1 Market Securities and Utility Function

We consider the finite time horizon $[0, T]$. Our model of the market consists of the numéraire bank account and zero coupon bond which matures at time $T$. We denote the price of the bond after dividing the numéraire by $S(t, T)$.

Our assumptions on the utility function are different from the ones in Kramkov and Schachermayer (1999)[23], Cvitanić et al. (2001)[5], Schachermayer (1999) [27]. We drop their assumption on asymptotic elasticity but we need a much smoother structure.

The utility function $U : (0, \infty) \to \mathbb{R}$ for non-negative wealth is assumed to be, continuously differentiable, strict concave ($U'' < 0$), strictly increasing ($U' > 0$). The inverse of $U'$ exist and smooth, it is denoted by $I$. Also

$$\lim_{x \to \infty} U'(x) = 0.$$

This will ensure the Fenchel-Legendre transformation ([22]) yields the important dual function

$$V(y) \equiv \sup_{x>0} \{U(x) - xy\}, y > 0,$$

which has the following properties.

The conjugate function $V(y)$ equals $U(I(y)) - yI(y)$, is continuously differentiable, strictly decreasing ($V'(y) < 0$), strictly convex ($V''(y) > 0$), and satisfies $V' = -I$.

We refer the proof to page 96 of [22].

## 2.2 Stochastic Volatility HJM Model

We use the two factor stochastic volatility Heath-Jarrow-Morton (HJM) model ([14]) as an example of incomplete market model to explain our approach. The multi-factor extension is just a matter of notation.

Given the probability triblet $(\Omega, \mathcal{F}, P)$, we consider the following forward rate dynamics (cf. [30])

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_1(t),$$
$$d\sigma(t,T) = \beta(t,T)dt + \nu(t,T)dW_2(t),$$

with a two dimensional Brownian motion $(W_1, W_2)^\top$.

Let the market prices of risk of $W_1$ and $W_2$ be $\phi$, $\xi$ satisfying the Novikov's condition.

Since $W_2$ is totally "unhedgeable", we call it *incomplete Brownian component* and $\xi$ *incomplete market prices of risk*.

The HJM no arbitrage condition "fixes" the *complete market prices of risk* $\phi_t$ as

$$\phi_t = \frac{\int_t^T \sigma(t,s)ds}{2} - \frac{\int_t^T \alpha(t,s)ds}{\int_t^T \sigma(t,s)ds}, \tag{2.1}$$

where

$$\int_t^T \sigma(t,s)ds = \int_t^T \sigma(0,s)ds + \int_t^T \int_0^t \beta(z,s)dzds + \int_t^T \int_0^t \nu(z,s)dW_2(z)ds.$$

By Girsanov transformation, we have the following form of the Radon-Nikodym density and the new Brownian motion under $Q_\xi$

$$\frac{dQ_\xi}{dP} = \exp\Big(\int_0^T \phi_s dW_1(s) + \int_0^T \xi_s dW_2(s) - \frac{1}{2}\int_0^T \phi_s^2 ds - \frac{1}{2}\int_0^T \xi_s^2 ds\Big),$$
$$\begin{pmatrix} \widetilde{W}_1(t) \\ \widetilde{W}_2(t) \end{pmatrix} = \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} - \begin{pmatrix} \int_0^t \phi_s ds \\ \int_0^t \xi_s ds \end{pmatrix}.$$

Each equivalent martingale measure is then related to each incomplete market price of risk process $(\xi_t)_{0\le t\le T}$.

### 2.3 Manifold of Equivalent Martingale Measures

Let $\mathcal{I}$ be the infinite dimensional space of "integrable" processes

$$\mathcal{I} \equiv \Big\{(\xi_t)_{0\le t\le T}\Big|E_P\Big[\exp(\frac{1}{2}\int_0^T \xi_s^2 ds)\Big] < \infty, \ \ \xi \text{ being adapted and predictable}\Big\}.$$

Let $\mathcal{M}$ be the space of equivalent martingale measures. We identify a martingale measure $Q_\xi$ with its density $dQ_\xi/dP$.

Using the Girsanov transformation (**??**), we have the following coordinate over $\mathcal{M}$. Namely, the bijective mapping

$$\xi \in \mathcal{I} \longmapsto \frac{dQ_\xi}{dP} \in \mathcal{M}.$$

Since $\mathcal{I}$ is infinite dimensional, so is $\mathcal{M}$. If we only consider a finite dimensional subset of $\mathcal{I}$ which the incomplete market price of risk process $\xi$ is specified up to a real parameter vector $\theta$, then we have obtain the following subset of $\mathcal{M}$

$$\mathcal{S} \equiv \Big\{\frac{dQ_\theta}{dP} \equiv \frac{dQ_{\xi(\theta)}}{dP} \in \mathcal{M}\Big|\theta \in \mathbb{R}^n\Big\}. \tag{2.2}$$

In the following way

$$\theta \in \mathbb{R}^n \longmapsto \frac{dQ_\theta}{dP} \in \mathcal{S},$$

we obtain a parameterized family of densities. This is also a *Manifold of Equivalent Martingale Measures* from an information differential geometric point of view.

The specifications of the manifold $\mathcal{S}$ are discussed in detail in our companion paper [10].

# 3 Entropy of the Incomplete Market

### 3.1 Maximum Entropy over $\mathcal{M}$

**Definition 3.1.** *For a real number $y > 0$, the* entropy *of a point $dQ_\xi/dP \in \mathcal{M}$ is defined as*

$$H_y(Q_\xi) \equiv -E_P\Big[V(y\frac{dQ_\xi}{dP})\Big].$$

The following theorem is due to Goll and Rüschendorf ([12]) using a result of Liese and Vajda ([24], Proposition 8.2). Note that the convexity of $\mathcal{M}$ is crucial.

**Theorem 3.2.** *(Goll and Rüschendorf, 2001) Fix any $y > 0$. If the maximum entropy point among $\mathcal{M}$ exists, i.e. there exist a unique point $Q^* \in \mathcal{M}$ such that*

$$H_y(Q^*) = \max_{Q_\xi \in \mathcal{M}} H_y(Q_\xi);$$

*then we have*

*(i) the maximum entropy point is unique,*

*(ii) for this maximal $dQ^*/dP$, we have*

$$E_P\Big[V'(y\frac{dQ^*}{dP})(y\frac{dQ^*}{dP} - y\frac{dQ}{dP})\Big] \leq 0, \forall\ Q \in \mathcal{M}.$$

The measure $Q^*$ is the "right" pricing measure for the incomplete market as shown by them.

We want to find a method to approximate $Q^*$ starting with some tractable subset of $\mathcal{M}$ which may not be convex. For the development of the geometry, the subset will be the manifold of equivalent martingale measures under some specification.

To see that a manifold of equivalent martingale measures is not convex in general, we have to show

$$Q_\xi \in \mathcal{S}, Q_\rho \in \mathcal{S} \ \not\Rightarrow\ \lambda Q_\xi + (1-\lambda)Q_\rho \in \mathcal{S}.$$

It follows from the fact that if

$$\frac{dQ_\eta}{dP} = \lambda\frac{dQ_\xi}{dP} + (1-\lambda)\frac{dQ_\rho}{dP},$$

then the incomplete market price of risk process of the convex combination is

$$\eta_t = \frac{\lambda\xi_t(\frac{dQ_\xi}{dP})_t + (1-\lambda)\rho_t(\frac{dQ_\rho}{dP})_t}{\lambda(\frac{dQ_\xi}{dP})_t + (1-\lambda)(\frac{dQ_\rho}{dP})_t}$$

using stochastic differential equation of the density.

## 3.2 Assumptions

We will be interested in the market of which the preference is sufficiently smooth and so is its dual. The fixed number $y$ below is related to the initial wealth of a portfolio.

**Assumption 1.** *(continuity and smoothness) In $\mathcal{M}$, the entropy $H_y(Q_\xi)$ is continuous as a function of $y$ and $\xi$ with respect to the corresponding topology. It is differentiable with respect to $y$ and Fréchet-differentiable with respect to $\xi$. We can interchange differentiation with the expectation under $P$.*

Let $\mathcal{E}$ be a fixed subset of $\mathcal{M}$ which can be $\mathcal{M}$ itself and can be of finite or infinite dimensional.

**Assumption 2.** *(existence) The maximum entropy point among the subset $\mathcal{E}$ of $\mathcal{M}$ exist, i.e. there exist a point $Q_0 \in \mathcal{E}$ such that*

$$H_y(Q_0) = \max_{Q_\xi \in \mathcal{E}} H_y(Q_\xi).$$

*Furthermore, for a fixed $Q_n \in \mathcal{M}$, the function of $\epsilon \in [0,1]$ and $Q \in \mathcal{E}$,*

$$f(\epsilon, Q) = E_P\Big[V\big(y\frac{dQ_n}{dP} + \epsilon(y\frac{dQ}{dP} - y\frac{dQ_n}{dP})\big)\Big],$$

*attains a minimum point over $[0,1] \otimes \mathcal{E}$.*

## 3.3 Examples of Existence

We know that $H_y(Q_\xi)$ is bounded above for all $Q_\xi \in \mathcal{M}$ due to Jensen's inequality, that is

$$-H_y(Q_\xi) = E_P\Big[V(y\frac{dQ_\xi}{dP})\Big] \geq V\Big(E_P[y\frac{dQ_\xi}{dP}]\Big) = V(y).$$

**Definition 3.3.** *A functional $f$ on a Banach space is* coercive *if $f(x) \to \infty$ as $\|\xi\| \to \infty$.*

**Proposition 3.4.** *With assumption 1, if $-H_y(Q_\theta)$ is coercive as functional of $\theta$ in $\mathbb{R}^n$, then the maximum entropy point exist over the manifold of equivalent martingale measures $\mathcal{S}$ defined in (2.2).*

*Proof.* We know that $f = -H_y(Q_\theta)$ is bounded below and so $\inf f$ exist.

We quote Theorem 1.6 of [13]. Let $f$ be a weakly semi-continuous functional bounded below on the reflexive Banach space. If $f$ is coercive, then $\inf f$ is attained at a point.

We know that $\mathbb{R}^n$ is Hilbert space and by our smoothness assumption, the minimum of $f$ is attained. Hence the maximum entropy point exist. $\square$

By the same argument, we have

**Proposition 3.5.** *Let $\mathcal{E} \subseteq \mathcal{M}$ be the subset of which any incomplete market price of risk process $\xi$ is a deterministic Lebesque square integrable function on $[0, T]$. If assumption 1 is satisfied and $-H_y(Q_\xi)$ is coercive as functional of $\xi$ in $L^2[0, T]$, then the maximum entropy point over $\mathcal{E}$ exist.*

Similarly, we could check whether the function $f(\epsilon, Q_\xi)$ in subsection 3.2 is coercive for the existence.

**Example 3.6.** *(Log Utility) Let $V(z) = -logz$ and let $\mathcal{E}$ be the subset of equivalent martingale measures of which the incomplete market price of risk is of the form $\xi_s(\theta) = \theta W_2(s)$. That is, the density is*

$$\frac{dQ_\theta}{dP} = e^{G(\theta)}$$

*where*

$$G(\theta) = F(W_1, W_2) + \theta \int_0^T W_2(s) dW_2(s) - \frac{1}{2}\theta^2 \langle \int_0^T W_2(s) dW_2(s) \rangle$$

*for some functional $F$ on the Brownian motion $W_1$ and $W_2$. In this case,*

$$f(\theta) \equiv E_P\Big[V(\frac{dQ_\theta}{dP})\Big]$$

*is coercive since*

$$f(\theta) = -E_P\Big[F(W_1, W_2) + \theta \int_0^T W_2(s) dW_2(s) - \frac{1}{2}\theta^2 \langle \int_0^T W_2(s) dW_2(s) \rangle\Big]$$

*goes to infinity as $\theta$ goes to infinity.*

The coercive condition is relatively easy to check. There are of course other assumptions that can do the same job such as the Palais-Smale conditions. We refer to our another companion paper [11] for more discussions.

## 3.4 Entropy Maximization Theorem

**Theorem 3.7.** *(Entropy Maximization Theorem) Fix any $y > 0$ and let $\mathcal{M}, \mathcal{E}, H_y$ satisfy the assumption 1, 2, then either*

*(i) there exist a sequence of measures $Q_0 \in \mathcal{E}, Q_1, \ldots, Q_n, \ldots$ in $\mathcal{M}$ not in $\mathcal{E}$ such that the sequence $H_y(Q_n)$ is strictly increasing, i.e. $E_P[V(y\frac{dQ_n}{dP})] > E_P[V(y\frac{dQ_{n+1}}{dP})]$;*

*or*

*(ii) there exist a set* $\mathcal{E}_N \equiv \mathcal{E} \cup \Big\{Q_1, \ldots, Q_N\Big\}, N \geq 0$ *such that* $Q_N$ *is the unique maximum entropy point over* $\mathcal{E}_N$ *and*

$$E_P\Big[V'(y\frac{dQ_N}{dP})(y\frac{dQ_N}{dP} - y\frac{dQ}{dP})\Big] \leq 0, \forall\, Q \in \mathcal{E}_N.$$

## 3.5 Proof of the Entropy Maximization Theorem

We need the following elementary fact.

**Lemma 3.8.** *Let* $f(\epsilon)$ *be a twice differentiable strictly convex function on* $[0,1]$*, i.e.* $f''(\epsilon) > 0$*. If* $f(1) \geq f(0)$*, then either*

*(i)* $f'(0) \geq 0$ *and* $f(0)$ *is the unique minimum of* $f$ *over* $[0,1]$*,*

*or*

*(ii)* $\exists\, \epsilon^* \in (0,1)$ *such that* $f(\epsilon^*)$ *attains the minimum of* $f$ *over* $[0,1]$ *and* $f(0) > f(\epsilon^*)$*.*

*Proof.* (Proof of the Entropy Maximization Theorem)

We will minimize $-H_y$ in the proof. Consider the following program.

Step 0: Solve

$$\min_{Q \in \mathcal{E}} E_P\Big[V(y\frac{dQ}{dP})\Big] \tag{3.1}$$

and obtain the minimal point at $Q_0 \in \mathcal{E}$ and let $\mathcal{E}_0 = \mathcal{E}$.

Step 1: Consider the minimization over the "net" formed by drawing the convex combinations from $Q_0$ to all $Q \in \mathcal{E}_0$, i.e.,

$$\min_{\epsilon \in [0,1], Q \in \mathcal{E}_0} f_1(\epsilon, Q) \tag{3.2}$$

where

$$f_1(\epsilon, Q) = E_P\Big[V\Big(y\frac{dQ_0}{dP} + \epsilon(y\frac{dQ}{dP} - y\frac{dQ_0}{dP})\Big)\Big].$$

Check whether

$$E_P\Big[V'(y\frac{dQ_0}{dP})(y\frac{dQ}{dP} - y\frac{dQ_0}{dP})\Big] \geq 0, \forall\, Q \in \mathcal{E}_0.$$

If so, we claim that $Q_0$ is unique minimum of the problem (3.2) and (3.1) (**Claim 1**) and we are done. Otherwise, we solve the problem (3.2) and obtain

$$Q_1 = Q_0 + \widehat{\epsilon}_1(\widehat{Q}_1 - Q_0), \quad \widehat{\epsilon}_1 \in (0,1), \widehat{Q}_1 \in \mathcal{E}_0,$$

and we claim that $E_P[y\frac{dQ_1}{dP}] < E_P[y\frac{dQ_0}{dP}]$ (**Claim 2**).

We proceed by induction.

Step $n+1$ after Step $n$: we let $\mathcal{E}_n = \mathcal{E} \cup \Big\{Q_1, \dots, Q_n\Big\}$. Before solving

$$\min_{\epsilon\in[0,1],Q\in\mathcal{E}_n} f_{n+1}(\epsilon, Q), \tag{3.3}$$

where

$$f_{n+1}(\epsilon, Q) = E_P\Big[V\Big(y\frac{dQ_n}{dP} + \epsilon(y\frac{dQ}{dP} - y\frac{dQ_n}{dP})\Big)\Big],$$

we check whether

$$E_P\Big[V'(y\frac{dQ_n}{dP})(y\frac{dQ}{dP} - y\frac{dQ_n}{dP})\Big] \geq 0, \forall\ Q \in \mathcal{E}_n.$$

If so, we know that $Q_n$ is unique minimum of problem (3.3) by Claim 1. Let $N = n$ and we are done.

Otherwise, we solve the problem and obtain

$$Q_{n+1} = Q_n + \widehat{\epsilon}_{n+1}(\widehat{Q}_{n+1} - Q_n), \quad \widehat{\epsilon}_{n+1} \in (0,1), \widehat{Q}_{n+1} \in \mathcal{E}_n,$$

and we have $E_P[y\frac{dQ_{n+1}}{dP}] < E_P[y\frac{dQ_n}{dP}]$ by Claim 2. Note that

$$\min_{\epsilon,Q\in\mathcal{E}_N} f_{n+1} = \min\Big\{\min_{\epsilon,Q\in\mathcal{E}} f_{n+1}, \min_{\epsilon} f_{n+1}(\epsilon, Q_1), \dots, \min_{\epsilon} f_{n+1}(\epsilon, Q_n)\Big\}$$

and from assumption 2, $\min_{\epsilon,Q\in\mathcal{E}_N} f_{n+1}$ exist.

Thus we have shown that this program will produce either (i) or (ii) of the theorem. We only have to check Claim 1 and Claim 2.

Proof of Claim 1 and 2:

Fix any $Q \neq Q_0, \in \mathcal{E}_0$, the function $f_1(\epsilon, Q)$ is strictly convex in $\epsilon$ since

$$\frac{\partial^2}{\partial\epsilon^2} f_1(\epsilon, Q) = E_P\Big[V''\Big(y\frac{dQ_0}{dP} + \epsilon(y\frac{dQ}{dP} - y\frac{dQ_0}{dP})\Big)(y\frac{dQ}{dP} - y\frac{dQ_0}{dP})^2\Big] > 0.$$

Note that

$$\frac{\partial}{\partial\epsilon} f_1(0, Q) = E_P\Big[V'(y\frac{dQ_0}{dP})(y\frac{dQ}{dP} - y\frac{dQ_0}{dP})\Big],$$

and $f_1(1, Q) = E_P[V(y\frac{dQ}{dP})] \geq f_1(0, Q) = E_P[V(y\frac{dQ_0}{dP})]$. So we can apply Lemma 3.8.

If $\frac{\partial}{\partial\epsilon} f_1(0, Q) \geq 0$, $\forall Q \in \mathcal{E}_0$, we know that $f_1(0, Q)$ is the unique minimum on $[0,1] \otimes \{Q\}$. Since $Q$ is an arbitrary element of $\mathcal{E}_0$, we have the result of claim 1.

On the other hand, if there exist a $Q \in \mathcal{E}_0$ such that $\frac{\partial}{\partial\epsilon} f_1(0,Q) < 0$, then there exist $\epsilon^*$ such that

$$E_P[V(y\frac{dQ_0}{dP})] = f_1(0,Q) > f_1(\epsilon^*,Q) \geq f_1(\widehat{\epsilon}_1,\widehat{Q}_1) = E_P[V(y\frac{dQ_1}{dP})].$$

We have proved Claim 2. $\square$

**Remark 3.9.** *If the sequence $E_P[V(y\frac{dQ_n}{dP})]$ does converge to $E_P[V(y\frac{dQ^*}{dP})]$, then by uniqueness of $Q^*$, we have that in some suitable topology*

$$Q_n \longrightarrow Q^*,\ Q^* = \sum_{k=1}^{\infty} p_k \widetilde{Q}_k,$$

*for some $\widetilde{Q}_k \in \mathcal{E}$ and $\sum_{k=1}^{\infty} p_k = 1$. This is because $Q_n$ can be written as sum of convex combinations of elements in $\mathcal{E}$.*

# 4 Entropy and Maximum Utility Equilibrium

In the Entropy Maximization Theorem, if we have $\mathcal{E}_N$ that satisfies (4.5), we could not proceed our approximation of $Q^*$. However, we will show that such a subset will have nice dual relation with the equilibrium of the incomplete market just like the optimal measure $Q^*$ over $\mathcal{M}$.

## 4.1 Homogenous Entropy

In the above section the entropy of the market depends on $y$, we want to study the entropy which is independent of $y$ and a quantity of the whole incomplete market. This leads to the following

**Definition 4.1.** *The entropy is* homogenous in initial wealth *if*

$$\arg\max_{Q\in\mathcal{E}} H_y(Q) = \arg\max_{Q\in\mathcal{E}} H_z(Q),\ \forall\ y,z>0, \forall\mathcal{E}\subseteq\mathcal{M}.$$

*In this case, we can define the* homogenous entropy *to be $H(Q) = c_1 H_{c_2}(Q) - c_3$ by suitably choosing the constants $c_1, c_2 > 0, c_3 \in \mathbb{R}$ for easy computation.*

**Assumption 3.** *The entropy of incomplete market derived from the utility function is homogenous in initial wealth.*

## 4.2 Examples of Homogenous Entropy

It is interesting that the entropies of most commonly used utilities are homogenous. This is also remarked in [12].

**Example 4.2.** *(Exponential Utility) Let $U(x) = -e^{-x}, x > 0$. Simple computation shows*

$$\begin{aligned} & I(z) = -\log z, V(z) = -z + z\log z, \\ & -H_y(Q) = E_P[V(y\frac{dQ}{dP})] = [-y + y\log y] + yE_P\Big[\frac{dQ}{dP}\log\big(\frac{dQ}{dP}\big)\Big]. \end{aligned}$$

*So we have that $\forall\ y > 0$,*

$$\begin{aligned} & \max_Q -E_P[V(y\frac{dQ}{dP})] \\ \Leftrightarrow\ & \max_Q -E_P\Big[\frac{dQ}{dP}\log\big(\frac{dQ}{dP}\big)\Big] \\ \Leftrightarrow\ & \min_Q E_P\Big[\frac{dQ}{dP}\log\big(\frac{dQ}{dP}\big)\Big]. \end{aligned}$$

*Hence the entropy is homogeneous and we define the homogenous entropy to be*

$$H(Q) = -E_P[\frac{dQ}{dP}\log(\frac{dQ}{dP})] = H_1(Q) - 1. \tag{4.1}$$

*The entropy is the Shannon's entropy or relative entropy (up to a sign difference) or Kullback-Leiber information in information geometry ( [25], [3]).*

*This generalizes the idea of Frittelli (2000) [9], namely, minimal relative entropy measure maximizes the exponential utility.*

**Example 4.3.** *(Log Utility) Let $U(x) = \log x, x > 0$. Similar computation yields*

$$\begin{aligned} & V(z) = -\log z - 1, \\ & -H_y(Q) = E_P[V(y\frac{dQ}{dP})] = -\log y - 1 - E_P[\log(\frac{dQ}{dP})], \\ & H(Q) = -E_P[-\log(\frac{dQ}{dP})] = H_1(Q) - 1. \end{aligned}$$

*Clearly, the entropy is homogenous. The homogeneous entropy is the $\alpha$-divergence with $\alpha = -1$ from information geometry (page 57 of [17]).*

**Example 4.4.** *(Power Hyperbolic Absolute Risk Aversion (HARA) Utility) Let $U(x) = -qx^p, p < 1, p \neq 0, \frac{1}{p} + \frac{1}{q} = 1$ or $q = \frac{p}{p-1}$. After some algebra, we have*

$$V(y\frac{dQ}{dP}) = y^q p(-\frac{1}{qp})^q(\frac{dQ}{dP})^q.$$

*It is clearly homogenous and we can define homogenous entropy to be*

$$H(Q) = -E_P[sign(p)(\frac{dQ}{dP})^q].$$

*When* $p = -1$ *or* $q = 1/2$*, the above homogenous entropy is the Hellinger distance or* $\alpha$*-divergence with* $\alpha = 0$ *between* $Q$ *and* $P$*.*

**Example 4.5.** *(Constant Relative Risk Aversion (CRRA) Utility) The utility is* $U(x) = \frac{x^p}{p}, 0 < p < 1$*. Similar to HARA utility, we have*

$$E_P[V(y\frac{dQ}{dP})] = y^{\frac{p}{p-1}}(\frac{1}{p} - 1)E_P[(\frac{dQ}{dP})^{\frac{p}{p-1}}].$$

*It is homogenous again and we have*

$$H(Q) = -E_P[(\frac{dQ}{dP})^{\frac{p}{p-1}}].$$

## 4.3 Approximation of the Equilibrium Wealth

A portfolio is a triplet $(x, h_t, Y_t)$, where the constant $x > 0$ is the initial wealth and the strategy $h_t$ is a predictable $S$-integrable process specifying the number of bonds to hold in the portfolio.

The $\mathcal{F}_t$ measurable random variable $Y_t$ is used to model the possible *cost* or injection of cash flow if $Y_t > 0$. It is *over-hedging profit* to be taken away from trading wealth or consumption if $Y_t < 0$ with positive wealth. We will see that it is an error term due to approximation of the maximum entropy measure $Q^*$.

The wealth process up to time $T$ is then

$$X_T^x = x + \int_0^T h(s)dS(s,T) + Y_T, \tag{4.2}$$

where the integration is with respect to the martingale discount bond price process $S$.

Let $\mathcal{E}_N$ be the subset of equivalent martingale measures as in section 3.

An *semi-equilibrium strategy* should end up with a non-negative wealth and risk neutral expected profit for all measures in $\mathcal{E}_N$.

We define up to time $T$ the set of trading wealths from $\mathcal{E}_N$ equilibrium strategies with a fixed initial wealth $x > 0$

$$\mathcal{X}_x(0) = \Big\{X > -x \ \text{ P a.s.} \ \Big| \ X = \int_0^T hdS + Y, E_Q[Y] \le 0, \forall\, Q \in \mathcal{E}_N\Big\}.$$

The strategy has bounded lost of trading wealth.

Note that every $\int_0^T hdS$ is a martingale under a measure $Q$, where $Q \in \mathcal{E}_N$. This is because the HJM no-arbitrage construction makes the discount zero coupon bond a martingale. Hence

$$\forall\ Q \in \mathcal{E}_N, \forall X \in \mathcal{X}_x(0), E_Q[X] = E_Q[Y] \leq 0.$$

We study the *semi-equilibrium* by maximizing the expected utility of terminal wealth over the set of $\mathcal{E}_N$ equilibrium strategies.

**Definition 4.6.** *The* indirect utility $u$ *and its dual* $v$ *are defined as*

$$u(x) \equiv \max_{X \in \mathcal{X}_x(0)} E_P[U(x+X)], x > 0, \tag{4.3}$$

$$v(y) \equiv \min_{Q \in \mathcal{E}_N} E_P\big[V(y\frac{dQ}{dP})\big], y > 0. \tag{4.4}$$

## 4.4 Utility Maximization Theorem

**Theorem 4.7.** *(Utility Maximization Theorem) Under the assumption 1, 2, 3, let* $\mathcal{E}_N$ *be the subset in (ii) of the Entropy Minimization Theorem. That is,* $\mathcal{E}_N$ *has a unique maximum entropy point* $Q_N$ *such that*

$$E_P\Big[V'(y\frac{dQ_N}{dP})(y\frac{dQ_N}{dP} - y\frac{dQ}{dP})\Big] \leq 0, \forall\ Q \in \mathcal{E}_N. \tag{4.5}$$

*Then we have*

*(i)* $u(x)$ *is well defined, i.e., the maximum is attained at some* $\widehat{X} \in \mathcal{X}_x(0)$,

*(ii) the maximal wealth* $\widehat{X}$ *for the value function* $u(x)$ *is given by* $\widehat{X} = I(y\frac{dQ_N}{dP}) - x$, *where* $y = u'(x)$.

**Remark 4.8.** *(Error Term* $Y$*) Notice that our definition of* $\mathcal{X}_x(0)$ *has an error term* $Y$ *compared to the definitions in [23], [5], [27] and [12]. Their dual optimization domain is set* $\mathcal{M}$. *Their optimal strategy with zero initial wealth can be represented as a stochastic integral* $\int_0^T hdS$. *Our dual domain* $\mathcal{E}_N$ *is much small than theirs. The smaller the set* $\mathcal{E}_N$ *is the bigger the set* $\mathcal{X}_x(0)$ *is. We are further away from a martingale representation of the optimal wealth in* $\mathcal{X}_x(0)$. *In the language of Jacka (1992, [18], Theorem 3.4), the maximum strategy* $I(y\frac{dQ_N}{dP})$ *is* $Q_N$*-attainable if and only if*

$$E_{Q_N}\Big[I(y\frac{dQ_N}{dP})\Big] - E_Q\Big[I(y\frac{dQ_N}{dP})\Big] = E_P\Big[I(y\frac{dQ_N}{dP})(\frac{dQ_N}{dP} - \frac{dQ}{dP})\Big] \geq 0$$

*for all $Q$ in the collection of martingale measures equivalent to $Q_N$ which is just $\mathcal{M}$. This is the same as saying*

$$E_P\Big[V'(y\frac{dQ_N}{dP})(\frac{dQ_N}{dP}-\frac{dQ}{dP})\Big] \leq 0,\ \forall\ Q \in \mathcal{M}$$

*if and only if $Y = 0\ \ P$ a.s. for the optimal wealth $\widehat{X}$.*

The theorem gives rise to an interesting term $u'(x)$ that looks like a "marginal certainty equivalent" of the terminal wealth $x + \widehat{X}$.

**Corollary 4.9.** *Under the assumptions of theorem 4.7, we have at equilibrium points $\widehat{X}, x, y, Q_N$,*

$$E_P[U'(\widehat{X}+x)] = u'(x).$$

*Proof.* From the theorem, we have at equilibrium

$$\widehat{X}+x = I(y\frac{dQ_N}{dP}), \Rightarrow\ U'(\widehat{X}+x) = y\frac{dQ_N}{dP},$$
$$\Rightarrow E_P[U'(\widehat{X}+x)] = yE_P\Big[\frac{dQ_N}{dP}\Big] = y = u'(x).$$

□

## 4.5 Proof of Utility Maximization Theorem

**Lemma 4.10.** *The value function function $v(y)$ is continuously differentiable with*

$$v'(y) = -E_{Q_N}\Big[I(y\frac{dQ_N}{dP})\Big] = -E_P\Big[\frac{dQ_N}{dP}I(y\frac{dQ_N}{dP})\Big].$$

*Proof.* We know there exist an unique $Q_N \in \mathcal{E}_N$ such that $v(y) = E_P[V(y\frac{dQ_N}{dP})]$ and $Q_N$ is independent of $y$ by homogeneous assumption 3. By smoothness assumption,

$$\frac{\partial}{\partial y}E_P[V(y\frac{dQ_N}{dP})] = E_P[\frac{\partial}{\partial y}V(y\frac{dQ_N}{dP})] = E_P[V'(y\frac{dQ_N}{dP})\frac{dQ_N}{dP}]$$
$$= v'(y) = -E_P\Big[\frac{dQ_N}{dP}I(y\frac{dQ_N}{dP})\Big],$$

where the last equality follows from $V' = -I$. □

**Lemma 4.11.** *Given $y > 0$ defined by $-v'(y) = x$, there exist $\widehat{X} \in \mathcal{X}_x(0)$, such that*

$$x + \widehat{X} = I(y\frac{dQ_N}{dP}) \geq 0 \ P \ a.s. \quad and \ \ x = E_{Q_N}\Big[I(y\frac{dQ_N}{dP})\Big].$$

*Proof.* From Lemma 4.10, we have $-v'(y) = E_{Q_N}[I(y\frac{dQ_N}{dP})]$. So we have

$$x = E_{Q_N}[I(y\frac{dQ_N}{dP})], \tag{4.6}$$

since $-v'(y) = x$.

Now define $\widehat{X} = I(y\frac{dQ_N}{dP}) - x$, all we need to show is that $\widehat{X} \in \mathcal{X}_x(0)$.

First, since $I = (U')^{-1} \geq 0$, so $\widehat{X} + x \geq 0$ P a.s..

Second, from (4.6), $E_{Q_N}[\widehat{X}] = 0$. By Martingale Representation Theorem, we have

$$\widehat{X} = E_{Q_N}[\widehat{X}] + \int_0^T h_t dS(t,T) + Y_T = \int_0^T hdS + Y_T,$$

for some processes $h$ and $Q_N$ martingale $Y_t$.

For $\forall\ Q \in \mathcal{E}_N$, to show that $E_Q[Y_T] \leq 0$ , all we need to show is $E_Q[\widehat{X}] \leq 0$ since $E_Q[\int_0^T hdS] = 0$. By our assumption on $Q_N$,

$$\begin{aligned} E_Q[\widehat{X}] &= E_Q[I(y\frac{dQ_N}{dP})] - x \\ &= E_Q[I(y\frac{dQ_N}{dP})] - E_{Q_N}[I(y\frac{dQ_N}{dP})] \\ &= E_P\Big[V'(y\frac{dQ_N}{dP})(\frac{dQ_N}{dP} - \frac{dQ}{dP})\Big] \leq 0. \end{aligned}$$

□

With the help of the above lemmas which are different from the ones of Cvitanić et al.(2001)[5], the main proof follows from theirs. For completeness, it is included below.

*Proof.* (Proof of the Utility Maximization Theorem)

From the definition of $V$, we have $V(y) \geq U(x) - xy, x, y > 0$. So $\forall\ X \in \mathcal{X}_x(0), \forall\ Q \in \mathcal{E}_N$, we have

$$\begin{aligned} U(x+X) &\leq V(y\frac{dQ}{dP}) + (x+X)y\frac{dQ}{dP}, \\ E_P[U(x+X)] &\leq E_P[V(y\frac{dQ}{dP})] + E_P[\frac{dQ}{dP}(x+X)y] \\ &\leq E_P[V(y\frac{dQ}{dP})] + E_Q[(x+X)y] \\ &\leq E_P[V(y\frac{dQ}{dP})] + xy, \\ \Rightarrow u(x) &\leq \inf_{s>0}\{v(s) + xs\} \end{aligned}$$

where the last equality follows from the fact that $E_Q[X] \leq 0$ for $\forall\ Q \in \mathcal{E}_N$ for admissible trading gains.

So we see that the maximization problem on the left hand side is dominated by the minimization problem on the right hand side.

Since $V(z) = U(I(z)) - zI(z)$, we have $V(y\frac{dQ}{dP}) = U(I(y\frac{dQ}{dP})) - y\frac{dQ}{dP}I(y\frac{dQ}{dP})$. So right hand side of the above inequality becomes

$$E_P[V(y\frac{dQ}{dP})] + xy = E_P\Big[U(I(y\frac{dQ}{dP}))\Big] + y(x - E_Q\Big[I(y\frac{dQ}{dP})\Big]).$$

Comparing this with the left hand side of the above inequality, we find that if

$$x + X = I(y\frac{dQ}{dP}), x = E_Q\Big[I(y\frac{dQ}{dP})\Big], \tag{4.7}$$

then the equality holds.

If we could find $y, Q, X$ such that the equalities (4.7) hold, we then obtain the maximal point $\widehat{X}$ with some $y$ and the theorem will be proved.

From lemma 4.11, we have found them for $y$ with $-v'(y) = x$.

We want to find the expression of $y$ as some function of $x$. If $-v'(y) = x$, the following inequality $u(x) \leq \inf_{s>0}\{v(s)+xs\}$ becomes equality, i.e. $u(x) = \inf_{s>0}\{v(s)+xs\}$. From ( [22]), this shows that $v$ is the dual of $u$ just like $V$ is the dual of $U$. So we have $v' = -(u')^{-1}$ (just like $V' = -I$). So we have

$$-v'(y) = x \ \Rightarrow\ (u')^{-1}(y) = x \ \Rightarrow\ y = u'(x).$$

It completes the proof. $\square$

## 4.6 Arbitrage-free Equilibrium Pricing of Incomplete Market

The Utility Maximization Theorem tells us that if the equivalent martingale measures of the incomplete market are assumed to be in $\mathcal{E}_N$ with some assumptions on the incomplete market price of risk processes, then we could still achieve arbitrage-free equilibrium "approximately" by accepting some error term $Y$ in our equilibrium wealth.

Given a random discounted contingent claim $\widetilde{X} \geq 0$ at time $T$, we assume $P(\widetilde{X} > 0) > 0$.

A price of $\widetilde{X}$ is *arbitrage free* if it fall in the following interval,

$$\mathcal{I}_{NA} = [\inf_{Q\in\mathcal{E}_N} E_Q[\widetilde{X}], \sup_{Q\in\mathcal{E}_N} E_Q[\widetilde{X}]].$$

This interval is smaller than the one in Frittelli (2000 [8]), Karatzas and Kou (1996) [21].

We denote the semi-equilibrium wealth at terminal time $T$ with initial wealth $x > 0$ as $X^x$, i.e.

$$E_P[U(X^x)] = \max_{X\in\mathcal{X}_x(0)} E_P[U(x + X)].$$

We have replaced the wealth $X_T^{x,\pi}$ in [6] and [4] by the "approximately" maximal expected utility wealth $X^x$ in the following definition. Note that our $X^x$ has an extra error term $Y$.

The *semi-equilibrium price* (or marginal rate of substitution price) in the sense of Davis (1997, [6]) is

$$\widehat{p} = \frac{E_P[U'(X^x)\widetilde{X}]}{u'(x)}.$$

**Corollary 4.12.** *(Arbitrage-free Semi-equilibrium Pricing) Under the assumptions of Utility Maximization Theorem, the semi-equilibrium price of $\widetilde{X}$ is arbitrage-free and is given by*

$$E_{Q_N}[\widetilde{X}],$$

*where $Q_N$ is unique maximum entropy point of $\mathcal{E}_N$.*

*Proof.* We have for any initial wealth $x > 0$, $X^x = I(u'(x)\frac{dQ_N}{dP})$ and

$$E_P[U'(X^x)\widetilde{X}] = E_P[u'(x)\frac{dQ_N}{dP}\widetilde{X}] = u'(x)E_{Q_N}[\widetilde{X}].$$

So $\widehat{p} = E_{Q_N}[\widetilde{X}]$. Since $Q \in \mathcal{E}_N$, $\widehat{p} \in \mathcal{I}_{NA}$. □

# 5 Distance between Two Equivalent Martingale Measures

From the above section, we see that the problem of semi-equilibrium pricing in a incomplete market is equivalent to maximizing the entropy in $\mathcal{E}_N$. The reason we want to study semi-equilibrium and some subset $\mathcal{E}_N$ first is that some stronger geometrical structure might exist in some subset $\mathcal{E}_N$.

The key geometrical concept is the *distance* between two equivalent martingale measures which has a financial meaning.

**Convention**: From now on, for simplicity of notation, let's denote the original dual of the utility by $\widetilde{V}$ and define the *homogeneous dual* function by

$$V(\frac{dQ}{dP}) \equiv c_1 \widetilde{V}(c_2 \frac{dQ}{dP}) + c_3,\ \ c_1, c_2 > 0, \tag{5.1}$$

so that the homogenous entropy has the similar form of

$$H(Q) \equiv -E_P\Big[V(\frac{dQ}{dP})\Big].$$

Note that $V$ is also strictly convex.

## 5.1 Cross Entropy between two Martingale Measures

The concept of cross entropy is crucial for the information geometry of the incomplete market.

The cross entropy ([25], [26]) is a measure of dissimilarity between two Radon-Nikodym densities of the incomplete market.

**Definition 5.1.** *The* cross entropy *between two densities in $\mathcal{M}$ with respect to a entropy $H$ is defined as*

$$D(Q, \widetilde{Q}) \equiv H(\widetilde{Q}) - H(Q) + \frac{d}{d\epsilon} H\big(\widetilde{Q} + \epsilon(Q - \widetilde{Q})\big)\Big|_{\epsilon=0}.$$

The following form of cross entropy is frequently used.

**Proposition 5.2.** *The cross entropy satisfies*

$$\begin{aligned} D(Q, \widetilde{Q}) &= H(\widetilde{Q}) - H(Q) + E_P\Big[V'(\frac{d\widetilde{Q}}{dP})(\frac{d\widetilde{Q}}{dP} - \frac{dQ}{dP})\Big] \\ &= E_P\Big[V(\frac{dQ}{dP}) - V(\frac{d\widetilde{Q}}{dP}) - V'(\frac{d\widetilde{Q}}{dP})(\frac{dQ}{dP} - \frac{d\widetilde{Q}}{dP})\Big]. \end{aligned}$$

It's proved by direct calculation.

**Proposition 5.3.** *(Rao and Nayak, 1985) The cross entropy satisfies*

$$D(Q,\widetilde{Q}) = 0 \text{ if } Q = \widetilde{Q},$$
$$D(Q,\widetilde{Q}) > 0 \text{ if } Q \neq \widetilde{Q}.$$

*Proof.* Since $V$ is strictly convex, $H$ is strictly concave. We have for $Q \neq \widetilde{Q}$

$$\lim_{\epsilon\downarrow 0} \frac{H\big(\widetilde{Q}+\epsilon(Q-\widetilde{Q})\big)-H(\widetilde{Q})}{\epsilon} = \lim_{\epsilon\downarrow 0} \frac{H\big((1-\epsilon)\widetilde{Q}+\epsilon Q\big)-H(\widetilde{Q})}{\epsilon}$$
$$> \lim_{\epsilon\downarrow 0} \frac{(1-\epsilon)H(\widetilde{Q})+\epsilon H(Q)-H(\widetilde{Q})}{\epsilon}$$
$$= -[H(\widetilde{Q})-H(Q)].$$

Hence $D(Q,\widetilde{Q}) > 0$. The case where $Q = \widetilde{Q}$ is obvious. □

## 5.2 Price of Information

Let $Q$ and $\widetilde{Q}$ be two equivalent martingale measures in $\mathcal{M}$. If the point $Q$ is the only martingale measure, then the market would be complete. So the incomplete market can be thought of a collection of infinitely many complete markets.

By results of Kramkov and Schachermayer (1999) [23], Schachermayer (1999) [27], the terminal wealth correspond to the maximum expected utility is given by

$$X = I(c_2 \frac{dQ}{dP})$$

where $c_2 > 0$ is related to the initial wealth if the market is indeed complete at the point $Q$ as in the (5.1). Similarly, at $\widetilde{Q}$, we have the maximum expected utility wealth

$$\widetilde{X} = I(c_2 \frac{d\widetilde{Q}}{dP}).$$

Note that when we want to compare two points, the market can no longer be thought of as complete. Hence, although both $X$ and $\widetilde{X}$ use the same constant $c_2$, it does not mean that their initial wealth are the same.

When we switch from one martingale measure $Q$ to another martingale measure $\widetilde{Q}$, the information contained in the measure are changed. We will quantify how much the extra information is worth using difference in the ability to achieve maximum utility.

Since more wealth will lead to higher utility. We have to normalize the expected utility by the mean of the wealth as we are comparing the ability to achieve maximum utility from the same basis.

The *relative price of information* of switching from $\widetilde{Q}$ to $Q$ is defined as

$$p_r(Q, \widetilde{Q}) \equiv \left| \Big[ E_P[U(X)] - c_2 E_Q[X] \Big] - \Big[ E_P[U(\widetilde{X})] - c_2 E_Q[\widetilde{X}] \Big] \right|.$$

Note that we have to use the same measure to compute the mean of the wealth. This is because we want to compare the relative ability just like we use the same measure $P$ to compute the expected utility.

If we use different measure to compute the mean of the wealth, we get the *absolute price of information* defined as

$$p_a(Q, \widetilde{Q}) \equiv \left| \Big[ E_P[U(X)] - c_2 E_Q[X] \Big] - \Big[ E_P[U(\widetilde{X})] - c_2 E_{\widetilde{Q}}[\widetilde{X}] \Big] \right|.$$

This is to compare the absolute ability in achieving expected utility normalized by its own price of the wealth.

## 5.3 Relation between Entropy Difference, Cross Entropy and Price of Information

Since $\widetilde{V}$ is the original dual, we have

$$\widetilde{V}(y) = U(I(y)) - zI(y),\ I = (U')^{-1},\ \widetilde{V}'(y) = -I(y).$$

The new dual $V$ is related to $U$ by the following lemma.

**Lemma 5.4.** *Let $V(y) = c_1 \widetilde{V}(c_2 y) + c_3,\ c_1, c_2 > 0$, we have*

$$V'(y) = -c_1 c_2 I(c_2 y), \quad V''(y) = -\frac{c_1 c_2^2}{U''(I(c_2 y))}.$$

*Proof.* The first equality is obvious. The second equality follows from the fact that

$$\frac{dI(y)}{dy} = \frac{1}{U''(x)}$$

where $x = I(y)$. It is because $y = U'(x)$, we have

$$\frac{dI(y)}{dy} = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{\frac{d}{dx}U'(x)} = \frac{1}{U''(x)}.$$

□

**Proposition 5.5.** *The absolute price of information is proportional to the difference of entropy and the relative price of information is proportional to the cross entropy. That is*

$$p_a(Q, \widetilde{Q}) = \frac{1}{c_1}\left|H(\widetilde{Q}) - H(Q)\right|, \quad p_r(Q, \widetilde{Q}) = \frac{1}{c_1}D(Q, \widetilde{Q}).$$

*Proof.* Straightforward computation yields

$$\begin{aligned}
&H(\widetilde{Q}) - H(Q) = E_P\Big[V(\frac{dQ}{dP}) - V(\frac{d\widetilde{Q}}{dP})\Big] \\
&= c_1 E_P\Big[\widetilde{V}(c_2\frac{dQ}{dP}) - \widetilde{V}(c_2\frac{d\widetilde{Q}}{dP})\Big] \\
&= c_1 E_P\Big[U(I(c_2\frac{dQ}{dP})) - c_2\frac{dQ}{dP}I(c_2\frac{dQ}{dP}) - \Big(U(I(c_2\frac{d\widetilde{Q}}{dP})) - c_2\frac{d\widetilde{Q}}{dP}I(c_2\frac{d\widetilde{Q}}{dP})\Big)\Big] \\
&= c_1\Big[\Big(E_P[U(X)] - c_2E_Q[X]\Big) - \Big(E_P[U(\widetilde{X})] - c_2E_{\widetilde{Q}}[\widetilde{X}]\Big)\Big].
\end{aligned}$$

Hence the first equality is proved. Similarly for the cross entropy, we have

$$\begin{aligned}
&D(Q, \widetilde{Q}) = E_P\Big[V(\frac{dQ}{dP}) - V(\frac{d\widetilde{Q}}{dP}) + V'(\frac{d\widetilde{Q}}{dP})(\frac{d\widetilde{Q}}{dP} - \frac{dQ}{dP})\Big] \\
&= c_1 E_P\Big[U(I(c_2\frac{dQ}{dP})) - c_2\frac{dQ}{dP}I(c_2\frac{dQ}{dP}) - \Big(U(I(c_2\frac{d\widetilde{Q}}{dP})) - c_2\frac{d\widetilde{Q}}{dP}I(c_2\frac{d\widetilde{Q}}{dP})\Big) \\
&\qquad - c_2 I(c_2\frac{d\widetilde{Q}}{dP})(\frac{d\widetilde{Q}}{dP} - \frac{dQ}{dP})\Big] \\
&= c_1 E_P\Big[U(I(c_2\frac{dQ}{dP})) - c_2\frac{dQ}{dP}I(c_2\frac{dQ}{dP}) - \Big(U(I(c_2\frac{d\widetilde{Q}}{dP})) - c_2\frac{dQ}{dP}I(c_2\frac{d\widetilde{Q}}{dP})\Big)\Big] \\
&= c_1\Big[\Big(E_P[U(X)] - E_Q[X]\Big) - \Big(E_P[U(\widetilde{X})] - E_Q[\widetilde{X}]\Big)\Big].
\end{aligned}$$

Hence the second equality is proved. □

We have seen that the third term in the definition of the cross entropy switch the expectation of $\widetilde{X}$ from $\widetilde{Q}$ to $Q$.

**Remark 5.6.** *We have also shown that the expression inside the absolute value of $p_r(Q, \widetilde{Q})$ is nonnegative since the cross entropy is nonnegative. It does not mean that the ability of $Q$ is always higher than $\widetilde{Q}$ since it is only a relative difference using the same measures. On the contrary, the absolute price of information does reflect the absolute ability due to the Utility Maximization Theorem.*

The relative price of information reflects the difference and is a good measure of "distance" between information. However, two different information could have zero absolute price of information. Hence we use the relative price of information as a financial interpretation of the cross entropy and the distance between two equivalent martingale measures.

# 6 Finite Dimensional Information Differential Geometric Structure of the Incomplete Market

We are in the position to discuss the differential geometry of the subset or the finite dimensional manifold of equivalent martingale measures $\mathcal{S}$ defined in subsection 2.3.

## 6.1 Geometry and Infinitesimal Distance between two Equivalent Martingale Measures

At this point, one might ask the following question.

> Why not view the entropy $H(Q_\theta) = -E_P[V(\frac{dQ_\theta}{dP})]$ as a real function of $\theta$ and skip the "complicated" differential geometry?

Qualitatively, the entropy is a functional of the density $\frac{dQ_\theta}{dP}$. The financial informational differences between two densities are more complicated than the difference of the two parameters. By viewing the entropy only as a function of $\theta$, one will miss the true relations between the densities.

Quantitatively, the measure of infinitesimal distance between two densities is affected by the two views. This is related to the concept of Riemannian metric in information differential geometry ([17]).

We first explore the implication of viewing $H$ as a function of $\theta$. This means the gradient vector of the entropy function is given as

$$\nabla H = \frac{\partial H}{\partial \theta_i}\frac{\partial}{\partial \theta_i}, \tag{6.1}$$

where we assume Einstein summation convention (indices appearing twice means summing over). The terms $\frac{\partial}{\partial \theta_i}$ mean the basis vectors.

Implicitly this formula assume the Euclidean geometry in $\mathcal{S}$ with coordinate $\theta$. In this geometry, the infinitesimal distance between two points $\frac{dQ_\theta}{dP}$ and $\frac{dQ_{\theta+d\theta}}{dP}$ is given by

$$ds^2 = \sum_i d\theta_i^2.$$

But each point in $\mathcal{S}$ is a state-price density, the above measure of difference between two densities does not capture enough financial information. As a result the gradient flow using the above formula will not lead to the maximal entropy point efficiently.

## 6.2 Riemannian Metric on the Manifold of Equivalent Martingale Measures

What is the alternative measure of infinitesimal distance?

From the previous section, we see that the cross entropy contains rich financial information about the incomplete market. So we will use the cross entropy to define a infinitesimal distance.

In geometry terms, we want to define a Riemannian metric. From Rao (1987,[25]), Burbea and Rao (1982,[3]), we have the following information metric induced from the cross entropy.

**Proposition 6.1.** *(Rao, 1987) The formal expansion of the cross entropy between two infinitesimally near points are*

$$D(Q_\theta, Q_{\theta+d\theta}) = \frac{1}{2} g_{ij} d\theta_i d\theta_j + o(\sum_i d\theta_i^2),$$

*where*

$$g_{ij} = E_P\Big[V''(\frac{dQ_\theta}{dP})\frac{\partial}{\partial\theta_i}\Big(\frac{dQ_\theta}{dP}\Big)\frac{\partial}{\partial\theta_j}\Big(\frac{dQ_\theta}{dP}\Big)\Big]. \tag{6.2}$$

The proof is the same as [25] after replacing the Lebesgue measure by $dP$ and their density by $\frac{dQ_\theta}{dP}$. For easy reference, we reproduce here.

*Proof.* Write $f(x) = D(Q_\theta, Q_x)$ where $x \in \mathbb{R}^n$. We have from Proposition 5.3 that $f(\theta) = 0$ and

$$\frac{\partial}{\partial x_i} f(x)\Big|_{x=\theta} = 0, \ \forall \ i$$

since the cross entropy obtains the minimal value zero at $x = \theta$. Apply Taylor expansion to $f(x)$ around $\theta$ and note that $\partial\theta_i = \partial x_i$, we have

$$f(\theta + dx) = \frac{1}{2}\Big(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big) dx_i dx_j + o(\sum_i dx_i^2),$$

where

$$\begin{aligned}
&\Big(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big)\Big|_{x=\theta} = \frac{\partial^2}{\partial x_i \partial x_j}\Big(E_P\Big[V(\frac{dQ_\theta}{dP}) - V(\frac{dQ_x}{dP}) + V'(\frac{dQ_x}{dP})(\frac{dQ_x}{dP} - \frac{dQ_\theta}{dP})\Big]\Big) \\
&= \frac{\partial}{\partial x_i}\Big(E_P\Big[-V'(\frac{dQ_x}{dP})\frac{\partial}{\partial x_j}(\frac{dQ_x}{dP}) + V''(\frac{dQ_x}{dP})\Big(\frac{\partial}{\partial x_j}(\frac{dQ_x}{dP})\Big)(\frac{dQ_x}{dP} - \frac{dQ_\theta}{dP}) \\
&\qquad + V'(\frac{dQ_x}{dP})\frac{\partial}{\partial x_j}(\frac{dQ_x}{dP})\Big]\Big) \\
&= \frac{\partial}{\partial x_i}\Big(E_P\Big[V''(\frac{dQ_x}{dP})\Big(\frac{\partial}{\partial x_j}(\frac{dQ_x}{dP})\Big)(\frac{dQ_x}{dP} - \frac{dQ_\theta}{dP})\Big]\Big) \\
&= E_P\Big[V''(\frac{dQ_x}{dP})\Big(\frac{\partial}{\partial x_j}(\frac{dQ_x}{dP})\Big)\Big(\frac{\partial}{\partial x_i}(\frac{dQ_x}{dP})\Big)\Big]\Big|_{x=\theta}
\end{aligned}$$

□

Since $V$ is strictly convex, the metric is indeed positive definite. That is

$$\begin{aligned}
&g_{ij} d\theta_i d\theta_j \\
=&E_P\Big[V''(\frac{dQ_\theta}{dP})\frac{\partial}{\partial\theta_i}\Big(\frac{dQ_\theta}{dP}\Big)d\theta_i \frac{\partial}{\partial\theta_j}\Big(\frac{dQ_\theta}{dP}\Big)d\theta_j\Big] \\
=&E_P\Big[V''(\frac{dQ_\theta}{dP})\Big(d(\frac{dQ_\theta}{dP})\Big)^2\Big] > 0,
\end{aligned}$$

for non-zero $\frac{dQ_\theta}{dP}$.

The Riemannian metric defined in 6.2 is called *incomplete market Riemannian metric.*

We use the cross entropy to measure the difference between two near densities in order to capture the financial information.

So the infinitesimal distance between two equivalent martingale measures is proportional to the infinitesimal relative price of information or the cross entropy given by

$$ds^2 = 2D(Q_\theta, Q_{\theta+d\theta}) = g_{ij} d\theta_i d\theta_j.$$

With this metric the gradient of the entropy function on the Riemannian manifold $\mathcal{S}$ is given by

$$\nabla H = g^{ij}\frac{\partial H}{\partial\theta_j}\frac{\partial}{\partial\theta_i}, \tag{6.3}$$

where $g^{ij}$ is the matrix inverse of the metric $g_{ij}$.

Since the incomplete market metric contains financial information that is closely related to the utility and arbitrage-free equilibrium pricing, the above gradient is more suitable for incomplete market modelling.

**Example 6.2.** *(Fisher metric from exponential utility) We consider the exponential utility with the suitably chosen* $V(z) = z \log z$ *and* $V''(z) = \frac{1}{z}$*. The incomplete market metric is*

$$
\begin{aligned}
g_{ij} &= E_P\Big[\frac{1}{\frac{dQ_\theta}{dP}} \frac{\partial}{\partial \theta_i}(\frac{dQ_\theta}{dP}) \frac{\partial}{\partial \theta_j}(\frac{dQ_\theta}{dP})\Big] \\
&= E_P\Big[\frac{dQ_\theta}{dP} \frac{\partial}{\partial \theta_i}(\log \frac{dQ_\theta}{dP}) \frac{\partial}{\partial \theta_j}(\log \frac{dQ_\theta}{dP})\Big].
\end{aligned}
$$

*This is also called the Fisher Information metric which plays a central role in information geometry of statistics and econometrics.*

## 6.3 Incomplete Market Riemannian Metric and Absolute Risk Aversion

**Proposition 6.3.** *Let*

$$
X_\theta = I(c_2 \frac{dQ_\theta}{dP}), \quad \frac{dQ_\theta}{dP} = e^{G(\theta)}.
$$

*That is, let* $X_\theta$ *be the maximum expected utility wealth as if the market is complete at the equivalent martingale measure* $Q_\theta \in \mathcal{S}$ *and* $G(\theta)$ *be the log density. Then incomplete market Riemannian metric is related to the Pratt-Arrow absolute risk aversion* $A(X_\theta)$ *and risk tolerance* $T(X_\theta)$ *by*

$$
g_{ij} = c_1 c_2 E_{Q_\theta}\Big[\frac{1}{A(X_\theta)}(\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})\Big] = c_1 c_2 E_{Q_\theta}\Big[T(X_\theta)(\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})\Big],
$$

*where*

$$
A(X_\theta) \equiv -\frac{U''(X_\theta)}{U'(X_\theta)}, \quad T(X_\theta) \equiv -\frac{U'(X_\theta)}{U''(X_\theta)} = \frac{1}{A(X_\theta)}.
$$

*Proof.* The proof is straight forward computation.

From Lemma 5.4, 6.1 and the fact that $z = U'(I(z))$,

$$\begin{aligned} g_{ij} &= E_P\Big[V''(\frac{dQ_\theta}{dP})\frac{\partial}{\partial\theta_i}\Big(\frac{dQ_\theta}{dP}\Big)\frac{\partial}{\partial\theta_j}\Big(\frac{dQ_\theta}{dP}\Big)\Big] \\ &= E_P\Big[(-\frac{c_1c_2^2}{U''(I(c_2\frac{dQ_\theta}{dP}))})(\frac{dQ_\theta}{dP})(\frac{\partial G}{\partial\theta_i})(\frac{dQ_\theta}{dP})(\frac{\partial G}{\partial\theta_j})\Big] \\ &= E_P\Big[(\frac{dQ_\theta}{dP})(-\frac{c_1c_2U'(I(c_2\frac{dQ_\theta}{dP}))}{U''(I(c_2\frac{dQ_\theta}{dP}))})(\frac{\partial G}{\partial\theta_i})(\frac{\partial G}{\partial\theta_j})\Big] \\ &= c_1c_2E_{Q_\theta}\Big[\Big(-\frac{U'(X_\theta)}{U''(X_\theta)}\Big)(\frac{\partial G}{\partial\theta_i})(\frac{\partial G}{\partial\theta_j})\Big]. \end{aligned}$$

□

**Remark 6.4.** *We also have another interpretation of the infinitesimal distance between two equivalent martingale measures as*

$$ds^2 = g_{ij}d\theta_i d\theta_j = c_1c_2E_{Q_\theta}\Big[(dG)^2T(X_\theta)\Big],$$

*where* $dG = \frac{\partial G}{\partial\theta_i}d\theta_i$*. We compare two equivalent martingale measures* $Q_\theta$ *and* $Q_\eta$*. If the product of the risk neutral expected risk tolerance of the the maximum expected utility wealth at* $Q_\theta$ *and the change of log density is bigger than the one of* $Q_\eta$*, then the infinitesimal distance around* $Q_\theta$ *is also bigger than the one of* $Q_\eta$*.*

## 6.4 Linear Risk Tolerance Incomplete Market Model

The linear risk tolerance utility covers exponential utility, log utility and HARA utility. We have

$$T(x) = \alpha + \beta x \Leftrightarrow \begin{cases} \text{exponential utility} & \text{if } \beta = 0, \\ \text{log utility} & \text{if } \beta = 1, \\ \text{HARA utility} & \text{if } \beta \neq 0, \neq 1. \end{cases}$$

This is because integrating the above expression yields

$$U(x) = \begin{cases} -\alpha e^{-\frac{x}{\alpha}} & \text{if } \beta = 0, \\ \log(\alpha + x) & \text{if } \beta = 1, \\ \frac{1}{\beta}\Big(\frac{1}{1-\frac{1}{\beta}}\Big)\Big(\alpha+\beta x\Big)^{1-\frac{1}{\beta}} & \text{if } \beta \neq 0, \neq 1. \end{cases}$$

Note that $\beta$ could be negative. If $T(x)$ is a constant, then the utility is exponential and the incomplete market Riemannian metric is the Fisher information metric.

The incomplete market Riemannian metric of the linear risk tolerance models are

$$g_{ij} = c_1 c_2 E_{Q_\theta}\Big[(\alpha + \beta X_\theta)(\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})\Big].$$

By Example 6.2, we know that the Fisher metric (denoted by $g_{ij}^F$) corresponds to risk tolerance $T(X_\theta) = \alpha = \frac{1}{c_1 c_2}$ so that

$$g_{ij}^F = c_1 c_2 \alpha E_{Q_\theta}\Big[(\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})\Big] = E_{Q_\theta}\Big[(\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})\Big].$$

Using the equality $E[XY] = E[X]E[Y] + COV(X,Y)$, we have the following relation between the general incomplete market Riemannian metric $g_{ij}$ and the Fisher metric $g_{ij}^F$ as

$$g_{ij} = c_1 c_2 \Big(g_{ij}^F E_{Q_\theta}[T(X_\theta)] + COV_{Q_\theta}[T(X_\theta), (\frac{\partial G}{\partial \theta_i})(\frac{\partial G}{\partial \theta_j})]\Big).$$

## 6.5 Gradient Flow of the Entropy and Dynamics of the Efficient Market

In order to obtain maximum utility, the market will adjust the geometric parameters of the incomplete MPR to reach the maximum entropy density. An efficient market should move the parameter from a arbitrary initial state to the optimal state according to the most efficient way.

Geometrically, the movement in the manifold $\mathcal{S}$ is a curve from a real interval to the parameter space of $\mathcal{S}$: $\lambda \in [a,b] \to c(\lambda) \in \mathbb{R}^n$. The gradient flow of $H$ is a curve whose tangent vectors are gradients of $H$. We know by Cauchy-Schwarz inequality that the function $H$ increase most along the direction of its gradient vector. So gradient flow is the most efficient way to the optimal point. For the gradient flow $c(\lambda)$, we have

$$\frac{dc_i}{d\lambda}\frac{\partial}{\partial \theta_i} = \nabla H = g^{ij}\frac{\partial H}{\partial \theta_j}\frac{\partial}{\partial \theta_i},$$

and

$$\frac{dc_i}{d\lambda} = g^{ij}\frac{\partial H}{\partial \theta_j}(Q_{c(\lambda)}). \tag{6.4}$$

The differential equation (6.4) describes the movement towards equilibrium in an efficient market.

We could make use of the above equation as an algorithm to solve the entropy maximization problem if analytic solution is difficult to get. The algorithm is as follows.

1. Pick an initial point $\theta = (\theta_i)$.
2. Compute $\frac{\partial H}{\partial \theta_i}$ at the current state $\theta$. If the absolute values are zero or less than $\epsilon$ (fixed precision), stop.
3. Otherwise, compute $g_{ij}$ and invert it to $g^{ij}$. Also compute $\delta_i = g^{ij} \frac{\partial H}{\partial \theta_j}$. Move to the next point of $\theta_i + \eta\delta_i$, where $\eta$ is the step size.
4. Repeat the second item at this new point until it is optimal.

We refer to [11] for more delicate algorithms.

## 6.6 Critical Points of the Entropy

**Definition 6.5.** *The* critical points *of the entropy $H$ is the points such that*

$$\frac{\partial H}{\partial \theta_i} = 0, \text{ for all } i.$$

The Hessian at critical point ([20]) is the tensor

$$\nabla dH = \Big(\frac{\partial^2 H}{\partial \theta_i \partial \theta_j}\Big) d\theta_i \otimes d\theta_j.$$

**Proposition 6.6.** *If the unique maximal entropy point $Q_\theta$ lays in the interior of the manifold $\mathcal{S}$, then it is a critical point of $H$ with the Hessian matrix of $H$*

$$\frac{\partial^2 H}{\partial \theta_i \partial \theta_j}$$

*being negative definite.*

*Proof.* The proof is standard in differential geometry.

At $\zeta_\theta$, the gradient vector vanishes,

$$g^{ij} \frac{\partial H}{\partial \theta_j}(Q_\theta) = 0 \text{ for all } i.$$

The above is a matrix equation and since $g^{ij}$ is a non-singular matrix, so the solution is the vector

$$\frac{\partial H}{\partial \theta_j}(Q_\theta) = 0 \text{ for all } j.$$

Hence it is a critical point.

Take any geodesic $c$ with tangent vector

$$\dot{c}(0) = \frac{dc_i}{dt}\frac{\partial}{\partial\theta_i}$$

(cf page 139 of [20]), we have

$$\nabla dH(\dot{c}(0), \dot{c}(0)) = \frac{d^2}{dt^2}H(Q_{c(t)})|_{t=0}.$$

Since

$$\frac{d^2}{dt^2}H(Q_{c(t)})|_{t=0} \leq 0,$$

for a local maximum, we have the result as

$$\nabla dH(\dot{c}(0), \dot{c}(0)) = \frac{dc_i}{dt}\Big(\frac{\partial^2 H}{\partial\theta_i\partial\theta_j}\Big)\frac{dc_j}{dt} \leq 0.$$

□

The above proposition gives the same optimal conditions if $H$ is viewed as a real function of $\theta$.

So the optimal point or the critical point does not depends on the Riemannian metric. But to find the critical point, we need to know the geometry.

## 6.7 Pricing Formula as the Limit of some Pricing Sequence

We could use some numerical optimization algorithm on the Riemannian manifold $\mathcal{S}$ to produce a sequence of parameters $\theta_n$ approaching the critical point or the maximum entropy point $\theta^*$.

Suppose further that for each parameter $\theta_n$ of the incomplete market price of risk process $\xi_{\theta_n}$, the pricing formula of a derivative contingent claim $\widetilde{X}$ is

$$C_{\widetilde{X}}(\theta_n) = E_{Q_{\theta_n}}[\widetilde{X}e^{-\int_0^T r_u du}]$$

for some continuous functional as if $Q_{\theta_n}$ is the correct pricing measure. Then the correct price in the sense of arbitrage-free semi-equilibrium is the limit of the sequence $C_{\widetilde{X}}(\theta_n)$. Write this symbolically,

$$C_{\widetilde{X}}(\theta^*) = \lim_{\|\nabla H(Q_{\theta_n})\| \to 0} C_{\widetilde{X}}(\theta_n).$$

Hence we have a new numerical method of computing the price in an incomplete market.

# 7 Examples of Finite Dimensional Manifold of Equivalent Martingale Measures

We show how to construct a finite dimensional manifold of equivalent martingale measures using the stochastic volatility HJM model.

## 7.1 Prior Specifications of Incomplete Market Prices of Risk

Inspired from Bayesian prior analysis, we will specify the prior form of the incomplete market price of risk processes for our models.

The market price of risk process $\xi$ is said to be *exponential* if it satisfies the following. For $\forall\ t \in [0,T]$, the stochastic integral can be written as

$$\int_0^t \xi_s dW_2(s) = \sum_{i=1}^n \theta_i J_i(t, W_1, W_2), \tag{7.1}$$

where $\theta_i$ are real parameters independent of $t$; $J_i(t, W_1, W_2)$ is a stochastic functional of $W_1(s), W_2(s), 0 \le s \le t$ and $t$. It is independent of $\theta_i$ a.s.. The independence means, for almost all $\omega \in \Omega$,

$$\frac{\partial}{\partial \theta_j} J_i(t, W_1, W_2)(\omega) = 0 \text{ for all } i, j. \tag{7.2}$$

The parameter $\theta = (\theta_i)$ are for the planning interval $[0,T]$. If the planning horizon is changed to $[t,T]$, we could have another parameter. So the parameter is actually function of the planning interval $\theta(t,T)$.

If we specify the incomplete market price of risk process to be deterministic and exponential, the above stochastic integrals in the Radon-Nikodym derivative become Gaussian.

**Example 7.1.** *(constant) Let* $\xi_s = \theta, \theta \in \mathbb{R}, \forall s \in [0,T]$*, then* $\int_0^t \xi_s dW_2(s) = \theta W_2(t)$*. So the functional* $J(t, W_1, W_2) = W_2(t)$*.*

**Example 7.2.** *(polynomial) Let* $\xi_s = \sum_{i=0}^n \theta_i s^i, \theta_i \in \mathbb{R}, \forall s \in [0,T]$*, then* $\int_0^t \xi_s dW_2(s) = \sum_{i=0}^n \theta_i \int_0^t s^i dW_2(s)$*. The functional* $J(t, W_1, W_2) = \int_0^t s^i dW_2(s)$*.*

**Example 7.3.** *(piece-wise constant) Let* $t_0 = 0 < t_1 < \ldots < t_{n+1} = T$ *be a finite partition of the interval* $[0,T]$*, a piece-wise constant incomplete market price of risk is of the form* $\xi_s = \sum_{i=0}^n \theta_i I_{[t_i, t_{i+1}]}(s)$*. We have the stochastic integral and functional as*

$$\int_0^t \xi_s dW_2(s) = \sum_{i=0}^n \theta_i \big[W_2(t \wedge t_{i+1}) - W_2(t \wedge t_i)\big].$$

Next, we will give an example of the Non-Gaussian stochastic integral model.

**Example 7.4.** *We could assume the dynamics of $\xi_s$ is given by*

$$d\xi_t = \theta_1 dt + \theta_2 dW_1(t) + \theta_3 dW_2(t),$$

*where $\theta_i$ are all real constants. Equivalently, we have*

$$\xi_s = \xi_0 + \theta_1 s + \theta_2 W_1(s) + \theta_3 W_2(s).$$

*Putting $\theta_0 = \xi_0$ as another parameter, the functional becomes*

$$\int_0^t \xi_s dW_2(s) = \theta_0 W_2(t) + \theta_1 \int_0^t s dW_2(s) + \theta_2 \int_0^t W_1(s) dW_2(s) + \theta_3 \int_0^t W_2(s) dW_2(s).$$

**Remark 7.5.** *(Financial Economic interpretation of the prior specification) In general, the complete market price of risk process is an adapted process related to the bonds and short rates. However, for the incomplete Brownian components, traders or representative agents of the market can not use any securities to hedge them. So when they are planning the trading at time zero for the finite horizon $[0, T]$, the practical thing to do is to assume certain "simple" form of the incomplete market price of risk process and solve for it according to their risk preferences or utility functions. One typical assumption could be the constant incomplete market price of risk over the planning horizon.*

## 7.2 Exponential Family of the Radon-Nikodym densities

Recall that an *n-dimensional exponential family* with respect to a $\sigma$-finite measure (or probability measure) $dP$ are of the form

$$p(\mathbf{x}, \theta) = \exp\Big[E(\mathbf{x}) + \sum_{i=1}^{m} \theta_i L_i(\mathbf{x}) - \psi(\theta)\Big], \tag{7.3}$$

where $\mathbf{x}$ is a measurable (or random) vector with respect to $dP$, $E, L_1, \ldots, L_m$ are functionals of $\mathbf{x}$, $\psi$ is a function of $\theta$, $m \geq n$ and the parameter space $\Theta = \{(\theta_i)_{i=1,\ldots,m}\}$ is a n-dimensional subspace of $\mathbb{R}^m$.

**Proposition 7.6.** *If the incomplete market price of risk $\xi_s$ is exponential, then the family of Radon-Nikodym densities $\frac{dQ_\xi}{dP}$ of the stochastic volatility HJM model is a n-dimensional exponential family.*

*Proof.* Direct verification shows

$$\begin{aligned}\frac{dQ_\xi}{dP} = \exp\Big[ &\int_0^T \phi_s dW_1(s) - \frac{1}{2}\int_0^T \phi_s^2 ds + \sum_{i=1}^{n} \theta_i J_i(T, W_1, W_2) \\ &- \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \theta_i\theta_j \langle J_i(T, W_1, W_2), J_j(T, W_1, W_2)\rangle_t\Big],\end{aligned}$$

where $\phi$ is specified by the drift and volatility using the HJM no-arbitrage condition (2.1). Comparing the above expression with the definition of exponential family concludes the proof. □

# 8 Minimal Martingale Measures and Examples of Specifications of Stochastic Volatility HJM Model

In this subsection, we give examples on the specifications of the complete and incomplete market price of risk $\phi_s$ and $\xi_s$ with the utility function to compute the maximum entropy density. More general specifications will lead to more complex analytical calculations.

**Remark 8.1.** *In the HJM model, the specification of the complete market price of risk $\phi_s$ does not matter too much. This is because for complete market deterministic volatility HJM models, the interest rate derivatives pricing does not depend on the empirical drift and the complete market price of risk. But for stochastic volatility HJM model, the specification of $\phi_s$ (or equivalently, the drift of the HJM forward rates) plays an important role for the determination and characterization of the incomplete counterpart $\xi_s$.*

**Remark 8.2.** *In order to have positive HJM interest rate specification, one could follow Jin and Glasserman (2001) [19]. They specified $\phi_s$ using Flesaker-Hughston formulation or general pricing kernels from Markov processes. We have to add the incomplete Brownian component $W_2$ to drive the pricing kernel processes. This is because the volatility can be recovered from pricing kernel processes.*

## 8.1 Constant Incomplete Market Price of Risk and Independent complete Market Price of Risk

We assume that $\xi_s$ for the interval $[0, T]$ to be a constant $\theta$. It is exponential.

From equation (2.1), we know that $\phi_s$ depends on $W_1, W_2$ in general. But if we assume $\phi_s$ depends on $W_1$ only, then $W_2$ affects the drift of forward rate $\alpha(t, T)$ only through the volatility $\sigma(t, T)$ while $W_1$ drives the drift only through $\phi_s$.

We have

$$\frac{dQ_\theta}{dP} = \exp\Big(F(W_1) + \theta W_2(T) - \frac{1}{2}\theta^2 T\Big),$$

where $F(W_1)$ is a functional of $W_2(t), 0 \le t \le T$. We also have

$$\frac{\partial}{\partial\theta}(\frac{dQ_\theta}{dP}) = \frac{dQ_\theta}{dP}(W_2(T) - \theta T),$$
$$\frac{\partial^2}{\partial\theta^2}(\frac{dQ_\theta}{dP}) = \frac{dQ_\theta}{dP}\big[(W_2(T) - \theta T)^2 - T\big],$$
$$\frac{\partial H}{\partial\theta} = -E_P[V'(\frac{dQ_\theta}{dP})\frac{\partial}{\partial\theta}(\frac{dQ_\theta}{dP})],$$

$$\frac{\partial^2 H}{\partial\theta^2} = -E_P\Big[V''(\frac{dQ_\theta}{dP})\Big(\frac{\partial}{\partial\theta}(\frac{dQ_\theta}{dP})\Big)^2\Big] - E_P[V'(\frac{dQ_\theta}{dP})\frac{\partial^2}{\partial\theta^2}(\frac{dQ_\theta}{dP})]. \tag{8.1}$$

**Proposition 8.3.** *Let $\phi_1$ be independent of $W_2$ and $\phi_2 = \theta$ as the exponential parameter. Then the maximum entropy measure of Exponential, Log, HARA Utility is given by the minimal martingale measure, $Q_\theta$ where $\theta = 0$.*

*Proof.* The proof is computational. Note the fact that by Girsanov Theorem, under the measure $Q_\theta$, $\widetilde{W}_2(T) = W_2(T) - \theta T$ is a $Q_\theta$-Brownian motion.

(i) Exponential Utility $U(x) = -e^{-x}$.

The chosen new dual $V$ is $V(z) = z\log z$ and $V'(z) = 1 + \log z$. Solve

$$\frac{\partial H}{\partial\theta} = 0$$

for $\theta$. We have

$$\begin{aligned}
&- E_P\Big[\Big(1 + F(W_1) + \theta W_2(T) - \frac{1}{2}\theta^2 T\Big)\frac{dQ_\theta}{dP}\Big(W_2(T) - \theta T\Big)\Big] \\
&= -E_{Q_\theta}\Big[(1 + F(\widetilde{W}_1) + \theta\widetilde{W}_2(T) + \frac{1}{2}\theta^2 T)\widetilde{W}_2(T)\Big] = 0 \\
&\Rightarrow \quad \theta E_{Q_\theta}[\widetilde{W}_2^2(T)] = 0 \quad (\because E_{Q_\theta}[\widetilde{W}_2(T)] = 0) \\
&\Rightarrow \quad \theta T = 0 \\
&\Rightarrow \quad \theta = 0.
\end{aligned}$$

So we only have one critical point. We need to check using (8.1) that it is indeed a local maximum. At $\theta = 0$, the second term of $\partial^2 H/\partial\theta^2$ is

$$\begin{aligned}
&- E_P[V'(\frac{dQ_\theta}{dP})\frac{\partial^2}{\partial\theta^2}(\frac{dQ_\theta}{dP})] \\
&= -E_{Q_\theta}[(1 + F(\widetilde{W}_1) + \theta\widetilde{W}_2(T) + \frac{1}{2}\theta^2 T)(\widetilde{W}_2^2(T) - T)] \\
&= -E_{Q_\theta}[(1 + F(\widetilde{W}_1))(\widetilde{W}_2^2(T) - T)] \\
&= 0.
\end{aligned}$$

The first term of $\partial^2 H/\partial\theta^2$ is always less or equal to 0, so $\partial^2 H/\partial\theta^2 \leq 0$ and $\theta = 0$ is the maximum point.

(ii) Log Utility $U(x) = \log x$.

The homogeneous entropy is based on $V(z) = -\log z$. We have $V'(z) = -\frac{1}{z}$ and

$$\begin{aligned}
\frac{\partial H}{\partial\theta} &= E_P[\frac{1}{\frac{dQ_\theta}{dP}}\frac{dQ_\theta}{dP}(W_2(T) - \theta T)] \\
&= E_P[W_2(T) - \theta T] = 0 \\
&\Rightarrow \quad -\theta T = 0 \\
&\Rightarrow \quad \theta = 0.
\end{aligned}$$

Also

$$\frac{\partial^2 H}{\partial\theta^2} = E_P[-T] < 0.$$

(iii) HARA Utility $U(x) = -qx^p, 0 < p < 1, , p \neq 0, q = p/(p-1) < 0$.

The dual $V$ is chosen to be $V(z) = z^q$ and $V'(z) = qz^{q-1}$. Recall that if $X$ is normal $\mathcal{N}(0,T)$, then the moment generating function has the following properties

$$\begin{aligned}
&M(s) = E[e^{sX}] = e^{\frac{1}{2}Ts^2}, \\
&\frac{d}{ds}M(s) = E[e^{sX}X] = e^{\frac{1}{2}Ts^2}Ts, \\
&\frac{d^2}{ds^2}M(s) = E[e^{sX}X^2] = e^{\frac{1}{2}Ts^2}(T^2s^2 + T).
\end{aligned}$$

Similar to (i) and (ii),

$$\begin{aligned}
\frac{\partial H}{\partial\theta} &= -E_P\Big[q(\frac{dQ_\theta}{dP})^{q-1}\frac{dQ_\theta}{dP}(W_2(T) - \theta T)\Big] \\
&= -qE_{Q_\theta}\Big[e^{(q-1)[F(\widetilde{W}_1)+\theta\widetilde{W}_2(T)+\frac{1}{2}\theta^2 T]}\widetilde{W}_2(T)\Big] = 0 \\
&\Rightarrow \quad E_{Q_\theta}[e^{(q-1)\theta\widetilde{W}_2(T)}\widetilde{W}_2(T)] = 0 \\
&\Rightarrow \quad e^{\frac{1}{2}T(q-1)^2\theta^2}T(q-1)\theta = 0 \\
&\Rightarrow \quad \theta = 0.
\end{aligned}$$

And at $\theta = 0$,

$$\begin{aligned}
\frac{\partial^2 H}{\partial\theta^2} &= -E_P\Big[q^2(\frac{dQ_\theta}{dP})^q(W_2(T) - \theta T)^2 - qT(\frac{dQ_\theta}{dP})^q\Big] \\
&= -qE_{Q_\theta}\Big[(\frac{dQ_\theta}{dP})^{q-1}(qW_2^2(T) - T)\Big] \\
&= -qE_{Q_\theta}[e^{(q-1)F(\widetilde{W}_1)}]E_{Q_\theta}[qW_2^2(T) - T] < 0.
\end{aligned}$$

The first two terms in the above product is positive since $q < 0$ and the third term is $qT - T < 0$. Hence $\partial^2 H/\partial\theta^2 < 0$ and it completes the proof. $\square$

Under the above assumptions of the proposition, the minimal martingale measure ([7]) is the arbitrage-free equilibrium measure.

In particular, if one assumes that $\xi_s$ is a constant and the above three common utility functions, then one can just ignore the incomplete market price of risk (also called market price of volatility risk) and use the complete market setting for pricing. The result is automatically at the maximum expected utility.

With the above simplified assumptions, we do not need to use any geometry and the optimal point can be found by viewing $H$ as functions of $\theta$.

## 8.2 Log Utility, Linear and Stochastic Incomplete Market Price of Risk

The minimal martingale measure minimizes the entropy of log utility as shown in [28]. We show that our examples lead to the same conclusion as a parsimonious test of our models. These examples show that the maximum entropy point of a subset could be the maximum entropy point of the set $\mathcal{M}$.

Let us assume slightly more general that the complete market price of risk $\phi_s$ depends on both $W_1$ and $W_2$. Let the incomplete market price of risk be linear in time $\xi_s = \theta + \varphi s$. We have

$$\frac{dQ_{(\theta,\varphi)}}{dP} = \exp\Big(F(W_1, W_2) + \theta W_2(T) + \varphi \int_0^T s dW_2(s) - \frac{1}{2}\theta^2 T - \frac{1}{2}\varphi\theta T^2 - \frac{1}{6}\varphi^2 T^3\Big).$$

**Proposition 8.4.** *Under the above setup and log utility $U(x) = \log(x)$, the maximum entropy point is given by the coordinate $\theta = 0, \varphi = 0$, i.e. the minimal martingale measure.*

*Proof.* We solve the following system

$$\begin{aligned}\frac{\partial H}{\partial\theta} &= E_P\Big[W_2(T) - \theta T - \frac{1}{2}\varphi T^2\Big] = 0\\ \frac{\partial H}{\partial\varphi} &= E_P\Big[\int_0^T s dW_2(s) - \frac{1}{2}\theta T^2 - \frac{1}{3}\varphi T^3\Big] = 0.\end{aligned}$$

Since the above Gaussian integrals have zero means, the only solution is $\theta = 0, \varphi = 0$. The Hessian at this point is

$$\begin{pmatrix} -T & -\frac{1}{2}T^2 \\ -\frac{1}{2}T^2 & -\frac{1}{3}T^3 \end{pmatrix}$$

and is negative definite. □

**Proposition 8.5.** *Let the incomplete market price of risk be $\xi_s = \theta W_2(s)$. Under log utility $U(x) = \log x$, the maximum entropy point is $\theta = 0$.*

*Proof.* In this case,

$$\frac{dQ_\theta}{dP} = \exp\Big(F(W_1, W_2) + \frac{1}{2}\theta(W_2^2(T) - T) - \frac{1}{2}\theta^2 \int_0^T W_2^2(s)ds\Big).$$

So setting

$$\frac{\partial H}{\partial \theta} = E_P\Big[\frac{1}{2}(W_2^2(T) - T) - \theta \int_0^T W_2^2(s)ds\Big] = 0.$$

Since $E_P[W_2^2(T)] = T$ and

$$\frac{\partial^2 H}{\partial \theta^2} = E_P[\int_0^T W_2^2(s)ds] > 0,$$

we get the result. □

## 8.3 Examples of $\mathcal{E}_N$ in (ii) of the Entropy Maximization Theorem

We show examples of the subset $\mathcal{E}_N$ that has a unique maximum entropy point $Q_N$ such that

$$E_P\Big[V'(y\frac{dQ_N}{dP})(y\frac{dQ_N}{dP} - y\frac{dQ}{dP})\Big] \leq 0, \forall\, Q \in \mathcal{E}_N.$$

Under the setups of Proposition 8.3, it is easy to verify the following.

(1) Exponential Utility.

The dual is $V(z) = z\log z$ and $V'(z) = 1 + \log z$. At $\theta = 0, \forall\, \varphi \in \mathbb{R}^n$, we have

$$E_P\Big[\Big(1 + F(W_1)\Big)\Big(e^{F(W_1)} - e^{F(W_1)+\varphi W_2(T)-\frac{1}{2}\varphi^2 T}\Big)\Big]$$
$$= E_P\Big[\Big(1 + F(W_1)\Big)e^{F(W_1)}\Big] - E_P\Big[\Big(1 + F(W_1)\Big)e^{F(W_1)}\Big]E_P\Big[e^{\varphi W_2(T)-\frac{1}{2}\varphi^2 T}\Big] = 0,$$

since $E_P\Big[e^{\varphi W_2(T)-\frac{1}{2}\varphi^2 T}\Big] = 1$.

(2) Log Utility.

The dual $V(z) = -\log z$. At the optimal point $\theta = 0$, for $\forall\ \varphi \in \mathbb{R}^n$, we have

$$E_P\Big[V'(\frac{dQ_0}{dP})(\frac{dQ_0}{dP} - \frac{dQ_\varphi}{dP})\Big] = E_P\Big[-\frac{1}{e^{G(0)}}\Big(e^{G(0)} - e^{G(\varphi)}\Big)\Big]$$
$$= E_P\Big[-1 + e^{G(\varphi)-G(0)}\Big] = E_P\Big[-1 + e^{\varphi W_2(T)-\frac{1}{2}\varphi^2 T}\Big]$$
$$= -1 + 1 = 0.$$

(3) HARA Utility.

We have $V'(z) = qz^{q-1}$ and at $\theta = 0$

$$E_P\Big[qe^{(q-1)F(W_1)}\Big(e^{F(W_1)} - e^{F(W_1)+\varphi W_2(T)-\frac{1}{2}\varphi^2 T}\Big)\Big] = 0.$$

The next example are more general than the two cases of Proposition 8.4 and 8.5.

(4) Log Utility with arbitrary integrable $\xi_s$.

At $(0, 0)$ and $\forall\ \xi_s \in \mathcal{H}$, by the martingale property of the exponential, we have

$$E_P\Big[-1 + e^{G(\xi)-G(0)}\Big]$$
$$= E_P\Big[-1 + e^{\int_0^T \xi_s dW_2(s) - \frac{1}{2}\langle\int_0^T \xi_s dW_2(s)\rangle}\Big] = 0.$$

## 8.4 Example of General Specification

We assume $\phi_s$ depends on both $W_1$, $W_2$ and $\xi_s = \theta + \varphi s$. We assume Exponential Utility $U(x) = -e^x$. We can verify that in general,

$$\frac{\partial H}{\partial \theta}(0, 0) = -E_P\Big[(1 + F(W_1, W_2))e^{F(W_1,W_2)}W_2(T)\Big] \neq 0,$$
$$\frac{\partial H}{\partial \varphi}(0, 0) = -E_P\Big[(1 + F(W_1, W_2))e^{F(W_1,W_2)}(\int_0^T sdW_2(s))\Big] \neq 0.$$

The incomplete market metric or the Fisher metric is then

$$g_{\theta\theta} = E_P\Big[\frac{dQ_{(\theta,\varphi)}}{dP}\big(W_2(T) - \theta T - \frac{1}{2}\varphi T^2\big)^2\Big]$$
$$g_{\varphi\varphi} = E_P\Big[\frac{dQ_{(\theta,\varphi)}}{dP}(\int_0^T sdW_2(s) - \frac{1}{2}\theta T^2 - \frac{1}{3}\varphi T^3\big)^2\Big]$$
$$g_{\theta\varphi} = g_{\varphi\theta} = E_P\Big[\frac{dQ_{(\theta,\varphi)}}{dP}\big(W_2(T) - \theta T - \frac{1}{2}\varphi T^2\big)(\int_0^T sdW_2(s) - \frac{1}{2}\theta T^2 - \frac{1}{3}\varphi T^3\big)\Big].$$

For other utility functions and specifications, the incomplete market metric and the equation for critical point are not so easy to solve. Numerical computation might be needed.

# 9 Infinite Dimensional Analysis of the Entropy Functional

We will discuss the maximum entropy point $Q_{\xi^*}$ over $\mathcal{M}$, where the optimal incomplete market price of risk process $\xi^*_s$ is a stochastic process on the time interval $[0,T]$.

Let the integrable process $\eta_s \in \mathcal{H}$ be an arbitrary perturbation process. That is, we consider the entropy functional $H(Q_{\xi^*+\epsilon\eta})$ perturbed by a process $\eta_s$ and a very small positive number $\epsilon$.

## 9.1 Application of Malliavin Calculus to the Entropy Functional

We fix any incomplete market price of risk process $\xi$ first. Let us first view $H(Q_\xi)$ as a Brownian functional and define

$$\Phi(W_2) \equiv V\Big(\exp\Big[\int_0^T \phi_s dW_1(s) - \frac{1}{2}\int_0^T \phi_s^2 ds + \int_0^T \xi_s dW_2(s) - \frac{1}{2}\int_0^T (\xi_s)^2 ds\Big]\Big)$$

as functional of the Brownian motion $W_2$ only.

Together with the assumption 1, we assume further that the derivative in the Clark Formula $\partial\Phi(W_2; ds)$ (see page 365, [22]) of the functional $\Phi$ exist for the process $\xi_s$. That is

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\Phi(W_2+\epsilon\Lambda) - \Phi(W_2)\Big] = \int_0^T \Lambda_s \partial\Phi(W_2; ds),$$

where $\Lambda_s = \int_0^s \eta_u du$ and $\eta_s$ is an arbitrary perturbation process.

From (E.10) of Karatzas and Shreve (page 365, [22]), we have the following

$$E_P\Big[\int_0^T \Lambda_s \partial\Phi(W_2; ds)\Big] = E_P\Big[\Phi(W_2)\big(\int_0^T \eta_s dW_2(s)\big)\Big]. \tag{9.1}$$

The left hand side of the above expression is computed as

$$\begin{aligned}
&\frac{d}{d\epsilon}E_P\Big[\Phi(W_2+\epsilon\Lambda)\Big]\Big|_{\epsilon=0} \\
&= \frac{d}{d\epsilon}E_P\Big[V\Big(\exp\Big\{\int_0^T \phi_s dW_1(s) - \frac{1}{2}\int_0^T \phi_s^2 ds + \int_0^T \xi_s (dW_2(s) + \epsilon\eta_s ds) \\
&\qquad - \frac{1}{2}\int_0^T (\xi_s)^2 ds\Big\}\Big)\Big]\Big|_{\epsilon=0} \\
&= E_P\Big[V'\big(\frac{dQ_\xi}{dP}\big)\frac{dQ_\xi}{dP}\Big(\int_0^T \xi_s\eta_s ds\Big)\Big].
\end{aligned}$$

The right hand side of (9.1) is $E_P\Big[V(\frac{dQ_\xi}{dP})\Big(\int_0^T \eta_s dW_2(s)\Big)\Big]$.

Hence we have proved the following proposition.

**Proposition 9.1.** *Under the above assumptions and setup, we have for a incomplete market price of risk process* $\xi_s$

$$E_P\Big[V'(\frac{dQ_\xi}{dP})\frac{dQ_\xi}{dP}\Big(\int_0^T \xi_s\eta_s ds\Big)\Big] = E_P\Big[V(\frac{dQ_\xi}{dP})\Big(\int_0^T \eta_s dW_2(s)\Big)\Big], \tag{9.2}$$

*for arbitrary perturbation process* $\eta_s$.

## 9.2 Necessary Condition of the Optimal Incomplete Market Price of Risk Process

Similar to the calculus of variation, if $\xi_s^*$ is optimal of the functional $H(Q_\xi)$ for $\xi \in \mathcal{M}$, then for any perturbation $\eta_s$, we should have

$$\frac{\partial}{\partial \epsilon} H(Q_{\xi^*+\epsilon\eta})\Big|_{\epsilon=0} = 0.$$

So we compute under the assumption 1,

$$\begin{aligned}
&\frac{\partial}{\partial \epsilon} E_P\Big[V\Big(\exp\Big\{\int_0^T \phi_s dW_1(s) - \frac{1}{2}\int_0^T \phi_s^2 ds + \int_0^T \xi_s^* dW_2(s) \\
&\qquad + \epsilon\int_0^T \eta_s dW_2(s) - \frac{1}{2}\int_0^T (\xi_s^*)^2 ds - \epsilon\int_0^T \xi_s^*\eta_s ds - \frac{1}{2}\epsilon^2\int_0^T \eta_s^2 ds\Big\}\Big)\Big]\Big|_{\epsilon=0} \\
&= E_P\Big[V'(\frac{dQ_{\xi^*}}{dP})\frac{dQ_{\xi^*}}{dP}\Big(\int_0^T \eta_s dW_2(s) - \int_0^T \xi_s^*\eta_s ds\Big)\Big] = 0.
\end{aligned}$$

Use equation (9.2) with $\xi = \xi^*$, we have proved the following theorem.

**Theorem 9.2.** *(Characterization of the optimal incomplete market price of risk process) Under the above assumptions, we have for the maximum entropy point* $Q_{\xi^*}$ *with the optimal incomplete market price of risk process* $\xi_s^*$ *and arbitrary perturbation process* $\eta_s$ *that*

$$E_P\Big[\Big(V'(\frac{dQ_{\xi^*}}{dP})\frac{dQ_{\xi^*}}{dP} - V(\frac{dQ_{\xi^*}}{dP})\Big)\Big(\int_0^T \eta_s dW_2(s)\Big)\Big] = 0$$

## 9.3 Application of the Characterization Theorem

We consider the exponential utility which $V(z) = z \log z$ and $V'(z) = 1 + \log z$. Using characterization theorem, we have

$$
\begin{aligned}
& E_P\Big[\Big(V'(\frac{dQ_{\xi^*}}{dP})\frac{dQ_{\xi^*}}{dP} - V(\frac{dQ_{\xi^*}}{dP})\Big)\Big(\int_0^T \eta_s dW_2(s)\Big)\Big] \\
&= E_P\Big[\Big(\frac{dQ_{\xi^*}}{dP} + \frac{dQ_{\xi^*}}{dP}\log(\frac{dQ_{\xi^*}}{dP}) - \frac{dQ_{\xi^*}}{dP}\log(\frac{dQ_{\xi^*}}{dP})\Big)\Big(\int_0^T \eta_s dW_2(s)\Big)\Big] \\
&= E_P\Big[\Big(\frac{dQ_{\xi^*}}{dP}\Big)\Big(\int_0^T \eta_s dW_2(s)\Big)\Big] \\
&= E_P\Big[\Big(1 + \int_0^T (\frac{dQ_{\xi^*}}{dP})_s \xi_s^* dW_2(s) + \int_0^T (\frac{dQ_{\xi^*}}{dP})_s \phi_s dW_1(s)\Big)\Big(\int_0^T \eta_s dW_2(s)\Big)\Big] = 0 \\
&\Rightarrow \quad E_P\Big[\int_0^T (\frac{dQ_{\xi^*}}{dP})_s \xi_s^* \eta_s ds\Big] = 0,
\end{aligned}
$$

where we uses Itô isometry again. Putting $\eta = \xi^*$, we have

$$
E_P\Big[\int_0^T (\frac{dQ_{\xi^*}}{dP})_s (\xi_s^*)^2 ds\Big] = 0 \quad \Rightarrow \quad \xi_s^* = 0 \quad P \text{ a.s.}
$$

since $(\frac{dQ_{\xi^*}}{dP})_s > 0$.

Hence we have proved the following proposition.

**Proposition 9.3.** *If the maximum entropy measure for exponential utility exist and the the above assumptions are satisfied, the minimal measure is the maximum entropy measure.*

# 10 Conclusions

From a utility maximization point of view, it is surprising to see that the incomplete financial market moves towards equilibrium by maximizing its entropy just like a physical thermal system. Another surprise is that the state price densities of the incomplete market have a Riemannian geometrical structure just like the physical space-time described by the general theory of relativity.

We have shown that there exist an information geometrical structure in the incomplete market. The geometry is shown to be closely related to utility maximization equilibrium.

We hope to introduce a new mathematical tool, information geometry, for solving incomplete market problems in the future.

We have just discovered a new research direction and there are still a lot more to be discovered.

# References

[1] Dorje C. Brody and Lane P. Hughston. Interest rates and information geometry. *Preprint, Imperical College London and King's College*, 2000.

[2] Dorje C. Brody and Lane P. Hughston. Application of information geometry to interest rates theory. *Disordered and Complex Systems*, pages 281–287, 2001. P. Sollich, A.C.C. Coolen, L.P. Hughston and R.F. Streater (eds).

[3] Jacob Burbea and C. Radhakrishna Rao. Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *Journal of Multivariate Analysis*, 12:575–596, 1982.

[4] Jakša Cvitanić. Option trading under constrains. *Financial Mathematics*, page 164, 1996. W. J. Runggaldier (eds), Lectures given at the 3rd Session of the Centro Internazionale Matematico Estivo (C.I.M.E), Springer.

[5] Jakša Cvitanić, Walter Schachermayer, and Hui Wang. Utility maximization in incomplete markets with random endowment. *Finance and Stochastics*, 5:259–272, 2001.

[6] Mark H.A. Davis. Option pricing in incomplete markets. *Mathematics of Derivatives Securities*, pages 216–226, 1997. M. Dempster and S. Pliska (eds), Cambridge: Cambridge University Press.

[7] Hans Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*, pages 389–411, 1991. In M. Davis and R. Elliott (Eds.), Applied Stochastic Analysis, Volume 5 of Stochastics Monographs, Gordon and Breach, London/New York.

[8] Marco Frittelli. Introduction to a theory of value coherent with the no-arbitrage principle. *Finance and Stochastics*, 4:275–297, 2000.

[9] Marco Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39–52, 2000.

[10] Yuan Gao, Kian Guan Lim, and Ng Kah Hwa. Finite and infinite dimensional specification of the incomplete market model. 2001.

[11] Yuan Gao, Kian Guan Lim, and Ng Kah Hwa. Numerical optimization on the manifold of equivalent martingale measures for a incomplete market. 2001.

[12] Thomas Goll and Ludger Rüschendorf. Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5(4):557–581, 2001.

[13] Maria Do Rosário Grossinho and Stepan Agop Tersian. *An introduction to minimax theorems and their applications to differential equations*, page 11. Kluwer Academic Publishers, Boston, 2001.

[14] David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term strucutre of interest rates: A new methodology for contingent claims. *Econometrica*, 60(1):77–105, 1992.

[15] Shun ichi Amari. Information geometry of the EM and em algorithms for neural networks. *Neural Networks*, 8(9):1379–1408, 1995.

[16] Shun ichi Amari, Koji Kurata, and Hiroshi Nagaoka. Information geometry of boltzmann machines. *IEEE Transactions on Neural Networks*, 3(2):260–271, 1992.

[17] Shun ichi Amari and Hiroshi Nagaoka. *Methods of Information Geometry*. Oxford University Press, American Mathematical Society, first edition, 2000. Translated by Daishi Harada.

[18] S. D. Jacka. A martingale representation result and an application to incomplete financial markets. *Mathematical Finance*, 2(4):239–250, 1992.

[19] Yan Jin and Paul Glasserman. Equilibrium positive interest rates: A unified view. *Review of Financial Studies*, 14(1):187–214, 2001.

[20] Jürgen Jost. *Riemannian Geometry and Geometric Analysis*. Springer, Berlin, second edition, 1998.

[21] I. Karatzas and S. G. Kou. On the pricing of contingent claims under constraints. *The Annals of Applied Probability*, 6(2):321–369, 1996.

[22] Ioannis Karatzas and Steven E. Shreve. *Methods of mathematical finance*, pages 24,306. Springer-Verlag, New York, first edition, 1998.

[23] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility fuctions and optimal investment in incomplete markets. *The Annals of Applied Probability*, 9(3):904–950, 1999.

[24] F. Liese and I. Vajda. *Convex Statistical Distances*. Teubner, Leipzig, 1987.

[25] C. R. Rao. Differential metrics in probability spaces. *Differential Geometry in Statistical Inference IMS Lecture Notes and Monograph Series 10, IMS: Hayward, CA*, pages 217–240, 1987. S. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen and C. R. Rao,.

[26] C. Radhakrishna Rao and Tapan K. Nayak. Cross entropy, dissimilarity measures, and characterizations of quadratic entropy. *IEEE Transactions on Information Theory*, 31(5):589–593, 1985.

[27] Walter Schachermayer. Optimal investment in incomplete markets when wealth may become nagative. *Annals of Applied Probability*. forth coming.

[28] Martin Schweizer. A minimal property of the minimal martingale measure. *Statistics & Probability Letters*, 42:27–31, 1999.

[29] Lukasz Stettner. Option pricing in discrete-time incomplete market models. *Mathematical Finance*, 10(2):305–321, 2000.

[30] Mitch Warachka. Stochastic volatility in the Heath, Jarrow and Morton model. 2000.