A Generalized Central Limit Theorem with Applications to Econometrics and Finance

Shige Peng

Institute of Mathematics
Shandong University, Fudan University
Minerva Foundation Lecturer, Columbia University
peng@sdu.edu.cn

Presented in
Seminar of Econometrics
Princeton University
November 25, 2008, Princeton, NJ

In this talk we present our recent result (see the attached paper [Peng2008]) of central limit theorem under uncertainty of probability measures and distributions (or ambiguity), a new type of law of large number is also derived in this general result. Since the paper is written in a style of mathematics, we now give explanations for their applications to finance and possibly in econometrics. A less rigorous language in the sense of mathematics will be used in this part.

This result is a generalization of the classical central limit theorem (CLT in short). The classical CLT is in the framework of a given probability space $(\Omega, \mathcal{F}, P)$, i.e., without uncertainty of probability measures. Our new result can provide a new argument to explain a puzzle why so many practitioners, e.g., traders and risk officials in financial markets can widely use normal distribution without serious data analysis or even with data inconsistence. We will show that in many situations these practitioners are right to calculate according the model of normal distribution, provided that they know how to chose the corresponding parameters, even under a situation of strong uncertainty of distributions so that the historical data cannot at all support the normal distribution hypothesis. However there are still many other situations where we must use a new type of 'sublinear distribution' $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$ and the related new calculation in the place of the classical normal distribution.

Since the law of large number and central limit theorem are two fundamental results in statistics and econometrics, it is natural to ask if this new result can be applied to this domains or/and to mathematical finance and other areas where probability and distribution uncertainties cannot be neglected. Many efforts are still needed in our future researches.

In econometrics, finance and many other domains of economics we often face to calculate of a quantity $E[\varphi(X)]$ for many different purpose, where φ is a given function and X a random variable (examples can be given such as vNM utilities, portfolio selections, option pricing, risk measures (VaR), pricing of mortgage based assets, option pricing, and many other situations).

It is an interesting phenomenon that normal distributions are widely applied to the above types of calculation and analysis. In many practical situation in finance people just use the normal distribution model:

$$X \sim \mathcal{N}(\mu, \sigma^2), \quad \mu = E[X], \quad \sigma^2 = E[(X-\mu)^2].$$

Or its multidimensional version. The historical data is used to estimates μ and σ^2 and then the value $E[\varphi(X)]$ calculate is calculated by

$$E[\varphi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \varphi(x) \exp(\frac{-(x-\mu)^2}{2\sigma^2}) dx.$$

Many academic researchers give critiques to the practitioners of quantitative finance for their wide abuses of normal distributions. It is also true that in many situations their data can be far from being normal distributed (fat tails, thin tails is usually the critics for theoretical analysis). In fact there exists a big gap between practitioners and academic researchers at this point since, on the other side, people in quantitative finance usually think that academic people are far from be the 'real world' of practical finance.

Why normal distributions are so widely applied? A very convincing explanation is from the classical central limit theorem.

Theorem 1 ***(Classical CLT)*** *Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d sequence and let $\mu = E[X_1]$ and $\sigma^2 = E[(X_1-\mu)^2]$, then*

$$\lim_{n\to\infty} E[\varphi(\sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}})] = E[\varphi(X)], \quad X \sim \mathcal{N}(0, \sigma^2).$$

The power and beauty of of this theorem come from the fact that the above sum tends to $\mathcal{N}(0, \sigma^2)$ regardless the original distribution of X_i, provided that $X_i \sim X_1$, for all $i = 2, 3, \cdots$ and that X_1, $X_2, \cdots$ are mutually independent. Since in finance, it is common that a big risk position X is in fact a sum a large (even huge) number of small quantities of assets, thus it is natural to apply the above CLT to assume that X is normal or nearly normal distributed. Application of CLT in finance can be traced back to the well-known thesis [B1900] of Louis Bachelier, in which the above small quantity was considered as a small step of random walk. This pioneer work deeply influenced [Osborne1959], B. Mandelbrot [Mande1963], Fama [Fama1965], P. Samuelson [Samuelson1965], R. Merton [Merton1973], F. Black and M. Scholes's assumption of geometric Brownian motion for the underlying prices of stocks in their option pricing models [BS1973].

Banks treat many of their risk positions of their mortgage based assets based on the similar reasoning.

But a question is: can they really be sure that $\{X_i\}_{i=1}^{\infty}$ is i.i.d. or approximately i.i.d.? If we ask this question to a trader or a risk manager of a bank, often the answer is not so clear. We know also many discussions and empirical analysis on the volatility uncertainty of stocks.

From this start point, it's more realistic to be within a situation of the uncertainty of distribution functions (of X_i) which in turn induces the model uncertainty of our probability space $(\Omega, \mathcal{F}, P)$. We will see that this types of small model uncertainty of distributions does not automatically disappear after the cumulation of $\sum X_i/\sqrt{n}$.

In order to treat this type of model uncertainty, we better not assume that the above X_i come from a same distribution. It is then more realistic to assume that the distribution of X_i is belong to a subset of distributions $\{F_\theta(x) : \theta \in \Theta\}$.

More generally speaking, we can not be sure that we have one probability space, instead we assume to have a set of probabilities $Q \in \mathcal{Q}$ and we don't know which is the real one. A safe robust calculation of an expectation of a loss position X is to take it's upper expectation

$$\hat{\mathbb{E}}[X] = \sup_{Q \in \mathcal{Q}} E_Q[X],$$

Remark 2 *In general one probability measure Q corresponds one expectation $E_Q[\cdot]$. The inverse is also true: one expectation $E[\cdot]$ corresponds one probability measure.*

Example 3 *Consider an economic agent with vNM type of utility function $u(X) = E[U(X)]$ where U is a concave and increasing function. If his uncertainty subset of probabilities is $\mathcal{Q}$, then his robust utility is calculated through*

$$\hat{u}(X) = \inf_{Q \in \mathcal{Q}} E_Q[X] = -\hat{\mathbb{E}}[-U(X)].$$

Example 4 *If a trader of a Bank shorts a call option $\varphi(X) := \max\{X - k, 0\}$, with a price p, where X is the underlying asset and k is the corresponding strike price. Then the risk of his position $p - \varphi(X)$ is measured by*

$$\varphi(x) = \hat{\mathbb{E}}[\varphi(X) - p] = \hat{\mathbb{E}}[\varphi(X)] - p.$$

The above defined upper expectation $\hat{\mathbb{E}}[\cdot]$ is a very useful notion. It is easy to check that it satisfies the following properties:

a) monotonicity: $\hat{\mathbb{E}}[X] \geq \hat{\mathbb{E}}[Y]$, if $X \geq Y$;

b) constant preserving: $\hat{\mathbb{E}}[c] = c$;

c) sub-additivity: $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$

d) positive homogeneity: $\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X]$, for constants $\lambda \geq 0$.

Properties c) + d) is called sub-linearity, an operator $\hat{\mathbb{E}}$ satisfying a), b), c) and d) is called an upper (or sublinear) expectation. If the inequality in c) becomes equality, then $\hat{\mathbb{E}}$ becomes the classical (linear) expectation and thus there is no probability uncertainty.

It is worth to point out that, recently the above types of sublinear expectations has been received an important attention. Many paper propose to use $\hat{\rho}(X) = \hat{\mathbb{E}}[-X]$ to be the measure of risk of a financial position X. This type of risk measure is called **coherent risk measure**. The following representation theorem is an important and basic result:

Theorem 5 *(**[ADEH1999]**, see also [Del2002], [FS2004] and my recent lecture notes) Let $\hat{\mathbb{E}}$ be a given a sublinear expectation. Then there exists a subset $\{E_\theta : \theta \in \Theta\}$ of linear expectations such that*

$$\hat{\mathbb{E}}[X] = \sup_{\theta \in \Theta} E_\theta[X], \quad \textit{for each } X.$$

This result tell us that, inversely a sublinear expectation also corresponds an uncertainty subset of probabilities.

Our basic point view is that: in a world with uncertainty of probabilities, it is much better to work with

Definition 6 *A triple $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a sublinear expectation space, where $\hat{\mathbb{E}}$ sublinear expectation $\hat{\mathbb{E}}$ defined on a linear space $\mathcal{H}$ of random variables $X : \Omega \mapsto \mathbb{R}$ satisfying $X = (X_1, \cdots, X_n)$, $X_i \in \mathcal{H}$ implies $\varphi(X) \in \mathcal{H}$, where φ is any function defined on $\mathbb{R}^n$ satisfying the following locally Lipschitz condition: $\varphi \in C_{Lip}(\mathbb{R}^n)$, i.e.,*

$$|\varphi(x) - \varphi(y)| \leq C(1 + |x|^m + |y|^m)|x - y|, \ \forall x, y \in \mathbb{R}^n.$$

In this model a specific $\hat{\mathbb{E}}$ gives us a specific size of our model uncertainty. Under this framework we can see that, for each given n-dimensional random vector $X = (X_1, \cdots, X_n)$, $X_i \in \mathcal{H}$:

$$\hat{\mathbb{F}}_X[\varphi] := \hat{\mathbb{E}}[\varphi(X)] : \varphi \in C_{Lip}(\mathbb{R}^n) \mapsto \mathbb{R}.$$

$\hat{\mathbb{F}}_X[\cdot]$ forms again a sublinear expectation on $(\mathbb{R}^n, C_{Lip}(\mathbb{R}^n))$. $\hat{\mathbb{F}}_X[\cdot]$ is called the sublinear distribution of X under $\hat{\mathbb{E}}$. In fact since $\hat{\mathbb{F}}_X[\cdot]$ is a sublinear expectation, thus there exists a subset of linear distributions $\{F_\theta, \theta \in \Theta\}$ such that

$$\hat{\mathbb{F}}_X[\varphi] = \sup_{\theta \in \Theta} \int_{\mathbb{R}^n} \varphi(x) dF_\theta(x)$$

Namely, the sub-linearity of $\hat{\mathbb{F}}_X[\cdot]$ describes the distribution uncertainty of X.

Definition 7 *In a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$, two n-dimensional random vector X and Y is said to be identically distributed, denoted by $X \sim Y$, if their sublinear distributions coincide. Namely $\hat{\mathbb{E}}[\varphi(X)] = \hat{\mathbb{E}}[\varphi(Y)]$, for each function φ.*

It is clear that $X \sim Y$ implies that the distribution uncertainties of X and Y are the same. It does not necessarily imply that the distribution of X is the same as that of Y.

Definition 8 *An n-dimensional random vector Y is said to be independent of another m-dimensional random vector X if for each function $\varphi \in C_{Lip}(\mathbb{R}^m \times \mathbb{R}^n)$ we have*

$$\hat{\mathbb{E}}[\varphi(X,Y)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x,Y)]_{x=X}].$$

Y is independent of X means that the distribution uncertainty of Y does not change for any realization of $X = x$.

It is clear that the above i.i.d. notion is more realistic. We begin to give our new central limit theorem with generalizes Theorem 1. The result presented here is in fact a special situation of Theorem 5.1 of the attached paper in the sense that here we only discuss 1-dimensional case (corresponding 1-dimensional normal distribution) whereas in Theorem 5.1 of the attached paper consider multi-dimensional cases. Theorem 5.1 also cover a part of the law of large number.

Theorem 9 *(New CLT [Peng2007], [Peng2008], See also my lecture notes) Let $\{X_i\}_{i=1}^{\infty}$ be a i.i.d. in a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ in the following sense: $X_i \sim X_1$ and X_{i+1} is independent to $(X_1, \cdots, X_i)$ for each $i = 1, 2, \cdots$ We also assume that $\mu = \hat{\mathbb{E}}[X_1] = -\hat{\mathbb{E}}[-X_1]$, then we denote*

$$\overline{\sigma}^2 = \hat{\mathbb{E}}[X_1^2], \quad \underline{\sigma}^2 = -\hat{\mathbb{E}}[-X_1^2]$$

Then for each convex function φ we have

$$\hat{\mathbb{E}}[\varphi(\sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}})] \to \hat{\mathbb{E}}[\varphi(X)], \text{ with } X \sim \mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2]).$$

Where X's distribution $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$ is a generalized normal distribution, called G-normal distribution (or G-distribution in the attached paper). A special case is $\underline{\sigma}^2 = \overline{\sigma}^2$, in this case X is classically normal distributed: $X \sim \mathcal{N}(0, \overline{\sigma}^2)$. Before given the full definition of this generalized form $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$ of our normal distribution, we first present the following two special and typical situation which has significant implication to the problem we have discussed at the beginning.

Corollary 10 *$X \sim \mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$, then If φ is a convex function then*

$$\hat{\mathbb{E}}[\varphi(X)] = \frac{1}{\sqrt{2\pi\overline{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(x) \exp(\frac{-x^2}{2\overline{\sigma}^2}) dx$$

But φ is a concave function, then

$$\hat{\mathbb{E}}[\varphi(X)] = \frac{1}{\sqrt{2\pi\underline{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(x) \exp(\frac{-x^2}{2\underline{\sigma}^2}) dx.$$

Remark 11 *The above two situations can explain why in practical situation, specially in finance and also in many theoretical studies in economics, people can widely use normal distribution: since in many situation the functions φ under consideration are convex or concave.*

Example 12 *An economic agent with vNM utility function U, then, as the above example, his robust utility can be approximately calculated by $-\hat{\mathbb{E}}[-U(X)]$. Since $-U$ is a convex function, thus we have*

$$-\hat{\mathbb{E}}[-U(X)] = \frac{1}{\sqrt{2\pi\overline{\sigma}^2}} \int_{-\infty}^{\infty} U(x) \exp(\frac{-x^2}{2\overline{\sigma}^2}) dx.$$

Example 13 *Similarly when we calculate the risk measure of the short position $p-(X-k)^+$ by $\hat{\mathbb{E}}[(X-k)^+ - p]$, since $(x-k)^+$ is convex, then we have*

$$\hat{\mathbb{E}}[(X-k)^+ - p] = \frac{1}{\sqrt{2\pi\overline{\sigma}^2}} \int_{-\infty}^{\infty} (x-k)^+ \exp(\frac{-x^2}{2\overline{\sigma}^2}) dx - p.$$

Remark 14 *Here it is important to mention that, it is possible that a recent history data $\{x_i\}_{i=1}^N$ of X is not at all approximated by a Gaussian. An extreme example is when $\underline{\sigma} = 0$ and $\overline{\sigma} = 1$. In this case $\delta_{\{0\}}$ is inside the distribution uncertainty subset of X. Thus it is very possible that our recent data for X is $x_i \equiv 0$, $i = 1, 2, \cdots$ but if, when φ is convex, we still use $X \sim \mathcal{N}(0,1)$ to calculate $\hat{\mathbb{E}}[\varphi(X)]$. But if φ is concave, then $\hat{\mathbb{E}}[\varphi(X)] = \varphi(0)$.*

We can still given example that with, a with unstable 'dirty history data' of X, the calculation of $\hat{\mathbb{E}}[\varphi(X)]$ is still by using Gaussian $\mathcal{N}(0, \underline{\sigma}^2)$ if φ is concave and $\mathcal{N}(0, \overline{\sigma}^2)$.

We now explain the general situation how to calculate $\hat{\mathbb{E}}[\varphi(X)]$ for $X \sim \mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$. Before doing this we give the definition of the sublinear distribution $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$:

Definition 15 *A random variable X on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is said to be $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$-distributed, with*

$$\overline{\sigma}^2 = \mathbb{E}[X^2], \quad \underline{\sigma}^2 = -\mathbb{E}[-X^2],$$

if for each a,b ≥ 0, we have $aX + b\bar{X} \sim \sqrt{a^2+b^2}X$, where $\bar{X}$ is an independent copy of X (namely $\bar{X}$ is independent of X and identically distributed with respect to X).

Remark 16 *This type of characterization, or definition, of normal distribution was introduced by P. Lévy and was systematically applied by Mandelbrot and Fama to modeling the normal distribution behavior of a stock prices. It is interesting to see that we just change our ground to our uncertainty formulation to arrive of new definition of normal distribution $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$.*

The mean of X is zero, this can be easily check by the definition: $\hat{\mathbb{E}}[X]=-\hat{\mathbb{E}}[-X]=0$. We also denote $\mu+X\sim\mathcal{N}(\mu,[\underline{\sigma}^2,\overline{\sigma}^2])$.

The calculation of $\hat{\mathbb{E}}[\varphi(X)]$ is via:

$$u(t,x):=\hat{\mathbb{E}}[\varphi(x+\sqrt{t}X)],\quad t\in[0,1],\ x\in\mathbb{R}.$$

We can check from the above definition that u is the solution of the following

$$\partial_t u=G(\partial_{xx}u),\quad u(0,x)=\varphi(x).$$

where

$$G(a)=\frac{\overline{\sigma}^2}{2}a^+-\frac{\underline{\sigma}^2}{2}a^-,\quad a^+=\max\{a,0\},\ a^-=(-a)^+.$$

Once we solve this parabolic partial differential equation, then it is clear that $\hat{\mathbb{E}}[\varphi(X)]=u(1,0)$.

Remark 17 *When φ is convex, it is easy to check that the solution u is convex in x, and thus u solves the heat equation $\partial u=\frac{\overline{\sigma}^2}{2}\partial_{xx}u$. We the can calculate $\hat{\mathbb{E}}[\varphi(X)]=u(1,0)$ as $X\sim\mathcal{N}(0,\overline{\sigma}^2)$. The case where φ is concave is similar.*

In the continuous time setting, Bachelier has defined Brownian motion. Similarly we can also define the corresponding G-Brownian motion under the uncertainty of probabilities for continuous time finance (see [Peng2006a] and [Peng2006b]).

References

[ADEH1999] Ph. Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath, Coherent Measures of Risk, Mathematical Finance 9, 203-228.

[B1900] Bachelier, L., Theorie de la Speculation, Paris, 1900.

[BS1973] Black, Fischer and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." Journal of Political Economy, vol. 81 (May-June), pp. 637-54.

[Del2002] Delbaen, F. (2002), Coherent Risk Measures (Lectures given at the Cattedra Galileiana at the Scuola Normale di Pisa, March 2000), Published by the Scuola Normale di Pisa.

[Fama1965] Fama, E. F. 1965. The behaviour of stock market prices. Journal of Business 38:34-105.

[FS2004] Föllmer & Schied (2004) Statistic Finance, Walter de Gruyter.

[Merton1973] Robert Merton, 1973. "Theory of Rational Option Pricing." The Bell Journal of Economics and Management Science, vol. 4 (Spring), pp. 141-83.

[Osborne1959] Osborne, M. F. M. 1959. Brownian motion in the stock market. Operations Research 7:145-73.

[Mande1963] Mandelbrot, B. 1963. The variation of certain speculative prices. Journal of Business 36: 394-419.

[Peng2007] Peng, S. (2007) Law of large numbers and central limit theorem under nonlinear expectations, in arXiv:math.PR/0702358v1 13 Feb 2007

[Peng2008] A New Central Limit Theorem under Sublinear Expectations, in arXiv:0803.2656v1 [math.PR] 18 Mar 2008.

[Peng2006a] Peng, S. (2006) G–Expectation, G–Brownian Motion and Related Stochastic Calculus of Ito's type, in arXiv:math.PR/0601035v2 3Jan 2006, in Proceedings of 2005 Abel Symbosium, Springer.

[Peng2006b] Peng, S. (2006) Multi-Dimensional G-Brownian Motion and Related Stochastic Calculus under G-Expectation, Stochastic Processes and their Applications Volume 118, Issue 12, December 2008, Pages 2223-2253

[Samuelson1965] Samuelson, Paul A. (1965). Rational theory of warrant pricing, Industrial Management Review, 6, 13-31

A New Central Limit Theorem under Sublinear Expectations

Shige PENG*
Institute of Mathematics
Shandong University
250100, Jinan, China
peng@sdu.edu.cn

version: March 18, 2008

Abstract

We describe a new framework of a sublinear expectation space and the related notions and results of distributions, independence. A new notion of G-distributions is introduced which generalizes our G-normal-distribution in the sense that mean-uncertainty can be also described. W present our new result of central limit theorem under sublinear expectation. This theorem can be also regarded as a generalization of the law of large number in the case of mean-uncertainty.

1 Introduction

The law of large numbers (LLN) and central limit theorem (CLT) are long and widely been known as two fundamental results in the theory of probability and statistics. A striking consequence of CLT is that accumulated independent and identically distributed random variables tends to a normal distributed random variable whatever is the original distribution. It is a very useful tool in finance since many typical financial positions are accumulations of a large number of small and independent risk positions. But CLT only holds in cases of model certainty. In this paper we are interested in CLT with mean and variance-uncertainty.

Recently problems of model uncertainties in statistics, measures of risk and superhedging in finance motivated us to introduce, in [13] and [14] (see also [11], [12] and references herein), a new notion of sublinear expectation, called "G-expectation", and the related "G-normal distribution" (see Def. 4.5) from which we were able to define G-Brownian motion as well as the corresponding

*The author thanks the partial support from The National Basic Research Program of China (973 Program) grant No. 2007CB814900 (Financial Risk).

stochastic calculus. The notion of G-normal distribution plays the same important rule in the theory of sublinear expectation as that of normal distribution in the classic probability theory. It is then natural and interesting to ask if we have the corresponding LLN and CLT under a sublinear expectation and, in particular, if the corresponding limit distribution of the CLT is a G-normal distribution. This paper gives an affirmative answer. We will prove that the accumulated risk positions converge 'in law' to the corresponding G-normal distribution, which is a distribution under sublinear expectation. In a special case where the mean and variance uncertainty becomes zero, the G-normal distribution becomes the classical normal distribution. Technically we introduce a new method to prove a CLT under a sublinear expectation space. This proof of our CLT is short since we borrow a deep interior estimate of fully nonlinear PDE in [5]. The assumptions of our CLT can be still improved.

This paper is organized as follows: in Section 2 we describe the framework of a sublinear expectation space. The basic notions and results of distributions, independence and the related product space of sublinear will be presented in Section 3. In Section 4 we introduce a new notion of G-distributions which generalizes our G-normal-distribution in the sense that mean-uncertainty can be also described. Finally, in Section 5, we present our main result of CLT under sublinear expectation. For reader's convenience we present some basic results of viscosity solutions in the Appendix.

This paper is a new and generalized version of [15] in which only variance uncertainty was considered for random variables instead random random vectors. Our new CLT theorem can be applied to the case where both mean-uncertainty and variance-uncertainty cannot be negligible. This theorem can be also regarded as a new generalization of LLN. We refer to [9] and [10] for the developments of LLN with non-additive probability measures.

2 Basic settings

For a given positive integer n we will denote by $\langle x, y\rangle$ the scalar product of x, $y \in \mathbb{R}^n$ and by $|x| = (x,x)^{1/2}$ the Euclidean norm of x. We denote by $\mathbb{S}(n)$ the collection of $n \times n$ symmetric matrices and by $\mathbb{S}_+(n)$ the non negative elements in $\mathbb{S}(n)$. We observe that $\mathbb{S}(n)$ is an Euclidean space with the scalar product $\langle P, Q\rangle = tr[PQ]$.

In this paper we will consider the following type of spaces of sublinear expectations: Let Ω be a given set and let $\mathcal{H}$ be a linear space of real functions defined on Ω such that if $X_1, \cdots, X_n \in \mathcal{H}$ then $\varphi(X_1, \cdots, X_n) \in \mathcal{H}$ for each $\varphi \in C_{l.Lip}(\mathbb{R}^n)$ where $C_{l.Lip}(\mathbb{R}^n)$ denotes the linear space of (local Lipschitz) functions φ satisfying

$$|\varphi(x) - \varphi(y)| \leq C(1 + |x|^m + |y|^m)|x - y|, \quad \forall x, y \in \mathbb{R}^n,$$
$$\text{for some } C > 0,\ m \in \mathbb{N} \text{ depending on } \varphi.$$

$\mathcal{H}$ is considered as a space of "random variables". In this case we denote $X = (X_1, \cdots, X_n) \in \mathcal{H}^n$.

Remark 2.1 *In particular, if $X, Y \in \mathcal{H}$, then $|X|$, $X^m \in \mathcal{H}$ are in $\mathcal{H}$. More generally $\varphi(X)\psi(Y)$ is still in $\mathcal{H}$ if $\varphi, \psi \in C_{l.Lip}(\mathbb{R})$.*

Here we use $C_{l.Lip}(\mathbb{R}^n)$ in our framework only for some convenience of technicalities. In fact our essential requirement is that $\mathcal{H}$ contains all constants and, moreover, $X \in \mathcal{H}$ implies $|X| \in \mathcal{H}$. In general $C_{l.Lip}(\mathbb{R}^n)$ can be replaced by the following spaces of functions defined on $\mathbb{R}^n$.

- $\mathbb{L}^\infty(\mathbb{R}^n)$: the space bounded Borel-measurable functions;
- $C_b(\mathbb{R}^n)$: the space of bounded and continuous functions;
- $C_b^k(\mathbb{R}^n)$: the space of bounded and k-time continuously differentiable functions with bounded derivatives of all orders less than or equal to k;
- $C_{unif}(\mathbb{R}^n)$: the space of bounded and uniformly continuous functions;
- $C_{b.Lip}(\mathbb{R}^n)$: the space of bounded and Lipschitz continuous functions.

Definition 2.2 *A **sublinear expectation** $\hat{\mathbb{E}}$ on $\mathcal{H}$ is a functional $\hat{\mathbb{E}} : \mathcal{H} \mapsto \mathbb{R}$ satisfying the following properties: for all $X, Y \in \mathcal{H}$, we have*

(a) Monotonicity:	*If $X \geq Y$ then $\hat{\mathbb{E}}[X] \geq \hat{\mathbb{E}}[Y]$.*
(b) Constant preserving:	*$\hat{\mathbb{E}}[c] = c$.*
(c) Sub-additivity:	*$\hat{\mathbb{E}}[X] - \hat{\mathbb{E}}[Y] \leq \hat{\mathbb{E}}[X - Y]$.*
(d) Positive homogeneity:	*$\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X], \ \forall \lambda \geq 0$.*

*(In many situation **(c)** is also called property of self–domination). The triple $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a **sublinear expectation space** (compare with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$). If only **(c)** and **(d)** are satisfied $\hat{\mathbb{E}}$ is called a sublinear functional.*

Remark 2.3 *Just as in the framework of a probability space, a sublinear expectation space can be a completed Banach space under its natural norm $\|\cdot\| = \hat{\mathbb{E}}[|\cdot|]$ (see [11]-[16]) and by using its natural capacity $\hat{c}(\cdot)$ induced via $\hat{\mathbb{E}}[|\cdot|]$ (see [4] and [3]). But the results obtained in this paper do not need the assumption of the space-completion.*

Lemma 2.4 *Let $\mathbb{E}$ be a sublinear functional defined on $(\Omega, \mathcal{H})$, i.e., **(c)** and **(d)** hold for $\mathbb{E}$. Then there exists a family $\mathcal{Q}$ of linear functional on $(\Omega, \mathcal{H})$ such that*

$$\mathbb{E}[X] := \sup_{E \in \mathcal{Q}} E[X], \quad \forall E \in \mathcal{Q}.$$

*and such that, for each $X \in \mathcal{H}$, there exists a $E \in \mathcal{Q}$ such that $\mathbb{E}[X] := E[X]$. If we assume moreover that **(a)** holds (resp. **(a)**, **(b)** hold) for $\mathbb{E}$, then **(a)** also holds (resp. **(a)**, **(b)** hold) for each $E \in \mathcal{Q}$.*

Proof. Let $\mathcal{Q}$ be the family of all linear functional dominated by $\mathbb{E}$, i.e., $E[X] \leq \mathbb{E}[X]$, for all $X \in \mathcal{H}$, $E \in \mathcal{Q}$. We first prove that $\mathcal{Q}$ is non empty. For a given $X \in \mathcal{H}$, we denote $L = \{aX : a \in \mathbb{R}\}$ which is a subspace of $\mathcal{H}$. We define $I : L \to \mathbb{R}$ by $I[aX] = a\mathbb{E}[X]$, $\forall a \in \mathbb{R}$, then $I[\cdot]$ forms a linear functional on L and $I \leq \mathbb{E}$ on L. Since $\mathbb{E}[\cdot]$ is sub-additive and positively homogeneous, by Hahn-Banach theorem (see e.g. [19]pp102) there exists a linear functional E on $\mathcal{H}$ such that $E = I$ on L and $E \leq \mathbb{E}$ on $\mathcal{H}$. Thus E is a linear functional dominated by $\mathbb{E}$ such that $\mathbb{E}[X] := E[X]$. We now define

$$\mathbb{E}_{\mathcal{Q}}[X] \triangleq \sup_{E \in \mathcal{Q}} E[X].$$

It is clear that $\mathbb{E}_{\mathcal{Q}}=\mathbb{E}$.

If **(a)** holds for $\mathbb{E}$, then for each non negative element $X \in \mathcal{H}$, for each $E \in \mathcal{Q}$, $E[X] = -E[-X] \geq -\mathbb{E}[-X] \geq 0$, thus **(a)** also holds for E. If moreover **(b)** holds for $\mathbb{E}$, then for each $c \in \mathbb{R}$, $-E[c] = E[-c] \leq \mathbb{E}[-c] = -c$ and $E[c] \leq \mathbb{E}[c] = c$, we get $E[c] = c$. The proof is complete. ■

Example 2.5 *For some $\varphi \in C_{l.Lip}(\mathbb{R})$, $\xi \in \mathcal{H}$, let $\varphi(\xi)$ be a gain value favorable to a banker of a game. The banker can choose among a set of distribution $\{F(\theta, A)\}_{A \in \mathcal{B}(\mathbb{R}), \theta \in \Theta}$ of a random variable ξ. In this situation the robust expectation of the risk for a gamblers opposite to the banker is:*

$$\hat{\mathbb{E}}[\varphi(\xi)] := \sup_{\theta \in \Theta} \int_{\mathbb{R}} \varphi(x) F(\theta, dx).$$

3 Distributions, independence and product spaces

We now consider the notion of the distributions of random variables under sublinear expectations. Let $X = (X_1, \cdots, X_n)$ be a given n-dimensional random vector on a sublinear expectation space $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}})$. We define a functional on $C_{l.Lip}(\mathbb{R}^n)$ by

$$\hat{\mathbb{F}}_X[\varphi] := \hat{\mathbb{E}}[\varphi(X)] : \varphi \in C_{l.Lip}(\mathbb{R}^n) \mapsto (-\infty, \infty). \tag{1}$$

The triple $(\mathbb{R}^n, C_{l.Lip}(\mathbb{R}^n), \hat{\mathbb{F}}_X[\cdot])$ forms a sublinear expectation space. $\hat{\mathbb{F}}_X$ is called the distribution of X.

Definition 3.1 *Let X_1 and X_2 be two n–dimensional random vectors defined respectively in sublinear expectation spaces $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}}_1)$ and $(\Omega_2, \mathcal{H}_2, \hat{\mathbb{E}}_2)$. They are called identically distributed, denoted by $X_1 \overset{d}{=} X_2$, if*

$$\hat{\mathbb{E}}_1[\varphi(X_1)] = \hat{\mathbb{E}}_2[\varphi(X_2)], \quad \forall \varphi \in C_{l.Lip}(\mathbb{R}^n).$$

It is clear that $X_1 \overset{d}{=} X_2$ if and only if their distributions coincide.

Remark 3.2 *If the distribution $\hat{\mathbb{F}}_X$ of $X \in \mathcal{H}$ is not a linear expectation, then X is said to have distributional uncertainty. The distribution of X has the following four typical parameters:*

$$\bar{\mu} := \hat{\mathbb{E}}[X], \quad \underline{\mu} := -\hat{\mathbb{E}}[-X], \qquad \bar{\sigma}^2 := \hat{\mathbb{E}}[X^2], \quad \underline{\sigma}^2 := -\hat{\mathbb{E}}[-X^2].$$

The subsets $[\underline{\mu}, \bar{\mu}]$ and $[\underline{\sigma}^2, \bar{\sigma}^2]$ characterize the mean-uncertainty and the variance-uncertainty of X. The problem of zero-mean uncertainty have been studied in [P3], [P4]. In this paper the mean uncertainty will be in our consideration.

The following simple property is very useful in our sublinear analysis.

Proposition 3.3 *Let $X, Y \in \mathcal{H}$ be such that $\hat{\mathbb{E}}[Y] = -\hat{\mathbb{E}}[-Y]$, i.e. Y has no mean uncertainty. Then we have*

$$\hat{\mathbb{E}}[X + Y] = \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y].$$

In particular, if $\hat{\mathbb{E}}[Y] = \hat{\mathbb{E}}[-Y] = 0$, then $\hat{\mathbb{E}}[X + Y] = \hat{\mathbb{E}}[X]$.

Proof. It is simply because we have $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$ and

$$\hat{\mathbb{E}}[X + Y] \geq \hat{\mathbb{E}}[X] - \hat{\mathbb{E}}[-Y] = \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y].$$

■

The following notion of independence plays a key role:

Definition 3.4 *In a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ a random vector $Y = (Y_1, \cdots, Y_n)$, $Y_i \in \mathcal{H}$ is said to be independent to another random vector $X = (X_1, \cdots, X_m)$, $X_i \in \mathcal{H}$ under $\hat{\mathbb{E}}[\cdot]$ if for each test function $\varphi \in C_{l.Lip}(\mathbb{R}^m \times \mathbb{R}^n)$ we have*

$$\hat{\mathbb{E}}[\varphi(X, Y)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x, Y)]_{x=X}].$$

Remark 3.5 *In the case of linear expectation, this notion of independence is just the classical one. It is important to note that under sublinear expectations the condition "Y is independent to X" does not implies automatically that "X is independent to Y".*

Example 3.6 *We consider a case where $X, Y \in \mathcal{H}$ are identically distributed and $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = 0$ but $\bar{\sigma}^2 = \hat{\mathbb{E}}[X^2] > \underline{\sigma}^2 = -\hat{\mathbb{E}}[-X^2]$. We also assume that $\hat{\mathbb{E}}[|X|] = \hat{\mathbb{E}}[X^+ + X^-] > 0$, thus $\hat{\mathbb{E}}[X^+] = \frac{1}{2}\hat{\mathbb{E}}[|X| + X] = \frac{1}{2}\hat{\mathbb{E}}[|X|] > 0$. In the case where Y is independent to X, we have*

$$\hat{\mathbb{E}}[XY^2] = \hat{\mathbb{E}}[X^+\bar{\sigma}^2 - X^-\underline{\sigma}^2] = (\bar{\sigma}^2 - \underline{\sigma}^2)\hat{\mathbb{E}}[X^+] > 0.$$

But if X is independent to Y we have

$$\hat{\mathbb{E}}[XY^2] = 0.$$

The independence property of two random vectors X, Y involves only the joint distribution of (X, Y). The following construction tells us how to construct random vectors with given sublinear distributions and with joint independence.

Definition 3.7 *Let $(\Omega_i, \mathcal{H}_i, \hat{\mathbb{E}}_i)$, $i = 1, 2$ be two sublinear expectation spaces. We denote by*

$$\mathcal{H}_1 \times \mathcal{H}_2 := \{Z(\omega_1, \omega_2) = \varphi(X(\omega_1), Y(\omega_2)) : (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2,$$
$$(X, Y) \in (\mathcal{H}_1)^m \times (\mathcal{H}_2)^n,\ \varphi \in C_{l.Lip}(\mathbb{R}^m \times \mathbb{R}^n),\ m, n = 1, 2, \cdots\},$$

and, for each random variable of the above form $Z(\omega_1, \omega_2) = \varphi(X(\omega_1), Y(\omega_2))$,

$$(\hat{\mathbb{E}}_1 \times \hat{\mathbb{E}}_2)[Z] := \hat{\mathbb{E}}_1[\bar{\varphi}(X)], \quad \textit{where } \bar{\varphi}(x) := \hat{\mathbb{E}}_2[\varphi(x, Y)],\ x \in \mathbb{R}^m.$$

It is easy to check that the triple $(\Omega_1 \times \Omega_2, \mathcal{H}_1 \times \mathcal{H}_2, \hat{\mathbb{E}}_1 \times \hat{\mathbb{E}}_2)$ forms a sunlinear expectation space. We call it the product space of sublinear expectation of $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}}_1)$ and $(\Omega_2, \mathcal{H}_2, \hat{\mathbb{E}}_2)$. In this way we can define the product space of sublinear expectation

$$(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathcal{H}_i, \prod_{i=1}^n \hat{\mathbb{E}}_i)$$

of any given sublinear expectation spaces $(\Omega_i, \mathcal{H}_i, \hat{\mathbb{E}}_i)$, $i = 1, 2, \cdots, n$. In particular, when $(\Omega_i, \mathcal{H}_i, \hat{\mathbb{E}}_i) = (\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}}_1)$ we have the product space of the form $(\Omega_1^{\otimes n}, \mathcal{H}_1^{\otimes n}, \hat{\mathbb{E}}_1^{\otimes n})$.

The following property is easy to check.

Proposition 3.8 *Let X_i be n_i-dimensional random vectors in sublinear expectation spaces $(\Omega_i, \mathcal{H}_i, \hat{\mathbb{E}}_i)$, for $i = 1, \cdots, n$, respectively. We denote*

$$Y_i(\omega_1, \cdots, \omega_n) := X_i(\omega_i), \quad i = 1, \cdots, n.$$

Then Y_i, $i = 1, \cdots, n$ are random variables in the product space of sublinear expectation $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathcal{H}_i, \prod_{i=1}^n \hat{\mathbb{E}}_i)$. Moreover we have $Y_i \overset{d}{=} X_i$ and Y_{i+1} is independent to $(Y_1, \cdots, Y_i)$, for each i.

Moreover, if $(\Omega_i, \mathcal{H}_i, \hat{\mathbb{E}}_i) = (\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}}_1)$ and $X_i = X_1$, for all i, then we also have $Y_i \overset{d}{=} Y_1$. In this case Y_i ia called independent copies of Y_1 for $i = 2, \cdots, n$.

The situation "Y is independent to X" often appears when Y occurs after X, thus a very robust sublinear expectation should take the information of X into account. We consider the following example: Let $Y = \psi(\xi, \theta)$, $\psi \in C_b(\mathbb{R}^2)$, where ξ and X are two bounded random variables in a classical probability space $(\Omega, \mathcal{F}, P)$ and θ is a completely unknown parameter valued in a given interval $[a, b]$. We assume that ξ is independent of X under P in the classical sense.

On the space $(\Omega, \mathcal{H})$ with $\mathcal{H} := \{\varphi(X,Y) : \varphi \in C_{l.Lip}(\mathbb{R}^2)\}$, we can define the following three robust sublinear expectations:

$$\mathbb{E}_1[\varphi(X,Y)] = \sup_{\theta\in[a,b]} E_P[\varphi(X,\psi(\xi,\theta)], \ \ \mathbb{E}_2[\varphi(X,Y)] = E_P[\sup_{\theta\in[a,b]} \varphi(X,\psi(\xi,\theta)],$$
$$\mathbb{E}_3[\varphi(X,Y)] = E_P[\{\sup_{\theta\in[a,b]} E_P[\varphi(x,\psi(\xi,\theta)]\}_{x=X}].$$

But it is seen that only under the sublinear expectation $\mathbb{E}_3$ that Y is independent to X.

Remark 3.9 *It is possible that the above parameter θ is in fact a function of X and ξ: $\theta = \Theta(X,\xi)$ where Θ is a completely unknown function valued in $[a,b]$, thus $Y = \psi(\xi,\Theta(X,\xi))$ is dependent to X in the classical sense. But since Θ is a completely unknown function a robust expectation is $\mathbb{E}_3$.*

Definition 3.10 *A sequence of d-dimensional random vectors $\{\eta_i\}_{i=1}^{\infty}$ in $\mathcal{H}$ is said to converge in distribution under $\hat{\mathbb{E}}$ if for each $\varphi \in C_b(\mathbb{R}^n)$ the sequence $\left\{\hat{\mathbb{E}}[\varphi(\eta_i)]\right\}_{i=1}^{\infty}$ converges.*

4 G-distributed random variables

Given a pair of d-dimensional random vectors (X,Y) in a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$, we can define a function

$$G(p,A) := \hat{\mathbb{E}}[\frac{1}{2}\langle AX, X\rangle + \langle p, Y\rangle], \quad (p,A) \in \mathbb{S}(d) \times \mathbb{R}^d \tag{2}$$

It is easy to check that $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ is a sublinear function monotonic in $A \in \mathbb{S}(d)$ in the following sense: For each $p, \bar{p} \in \mathbb{R}^d$ and $A, \bar{A} \in \mathbb{S}(d)$

$$\left\{ \begin{array}{rl} G(p+\bar{p}, A+\bar{A}) & \leq G(p,A) + G(\bar{p},\bar{A}), \\ G(\lambda p, \lambda A) & = \lambda G(p,A), \quad \forall \lambda \geq 0, \\ G(p,A) & \geq G(p,\bar{A}), \ if \quad A \geq \bar{A}. \end{array} \right. \tag{3}$$

G is also a continuous function.

The following property is classic. One can also check it by using Lemma 2.4.

Proposition 4.1 *Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a sublinear function monotonic in $A \in \mathbb{S}(d)$ in the sense of (3) and continuous in $(0,0)$. Then there exists a bounded subset $\Theta \in \mathbb{R}^d \times \mathbb{R}^{d\times d}$ such that*

$$G(p,A) = \sup_{(q,Q)\in\Theta} [\frac{1}{2} tr[AQQ^T] + \langle p, q\rangle], \quad \forall (p,A) \in \mathbb{R}^d \times \mathbb{S}(d).$$

The classical normal distribution can be characterized through the notion of stable distributions introduced by P. Lévy [6] and [7]. The distribution of a d-dimensional random vector X in a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called stable if for each $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^d$, there exists $\mathbf{c} \in \mathbb{R}^d$ and $d \in \mathbb{R}$ such that

$$\langle \mathbf{a}, X \rangle + \langle \mathbf{b}, \bar{X} \rangle \overset{d}{=} \langle \mathbf{c}, X \rangle + d,$$

where $\bar{X}$ is an independent copy of X.

The following G-normal distribution plays the same role as normal distributions in the classical probability theory:

Proposition 4.2 *Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given sublinear function monotonic in $A \in \mathbb{S}(d)$ the sense of (3) and continuous in $(0,0)$. Then there exists a pair of d-dimensional random vectors (X, Y) in some sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ satisfying (2) and the following condition:*

$$(aX + b\bar{X}, a^2Y + b^2\bar{Y}) \overset{d}{=} (\sqrt{a^2+b^2}X, (a^2+b^2)Y\ ), \quad \forall a, b \geq 0, \tag{4}$$

where $(\bar{X}, \bar{Y})$ is an independent copy of (X, Y). The distribution of (X, Y) is uniquely determine by G.

Example 4.3 *For the sublinear function $\bar{G} : \mathbb{R}^d \mapsto \mathbb{R}$ defined by $\bar{G}(p) := G(p, 0)$, $p \in \mathbb{R}^d$, we can concretely construct a d-dimensional random vector Y in some sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ satisfying*

$$\bar{G}(p) := \hat{\mathbb{E}}[\langle p, Y \rangle], \quad p \in \mathbb{R}^d \tag{5}$$

and the following condition:

$$a^2Y + b^2\bar{Y} \overset{d}{=} (a^2+b^2)Y, \quad \forall a, b \in \mathbb{R}, \tag{6}$$

where Y is an independent copy of Y. In fact we can take $\Omega = \mathbb{R}^d$, $\mathcal{H} = C_{l.Lip}(\mathbb{R}^d)$ and $Y(\omega) = \omega$. To define the corresponding sublinear expectation $\hat{\mathbb{E}}$, we apply Proposition 4.1 to find a subset $\bar{\Theta} \in \mathbb{R}^d$ such that

$$\bar{G}(p) = \sup_{q \in \bar{\Theta}} \langle p, q \rangle, \qquad p \in \mathbb{R}^d. \tag{7}$$

Then for each $\xi \in \mathcal{H}$ of the form $\xi(\omega) = \varphi(\omega)$, $\varphi \in C_{l.Lip}(\mathbb{R}^d)$. $\omega \in \mathbb{R}^d$ we set

$$\hat{\mathbb{E}}[\xi] = \sup_{\omega \in \bar{\Theta}} \varphi(\omega). \tag{8}$$

It is easy to check that the distribution of Y satisfies (5) and (6). It is the so-called worst case distribution with respect to the subset of mean uncertainty $\bar{\Theta}$. We denote this distribution by $\mathcal{U}(\bar{\Theta})$.

Example 4.4 *For the sublinear and monotone function $\hat{G} : \mathbb{S}(d) \mapsto \mathbb{R}$ defined by $\hat{G}(A) := G(0, A)$, $A \in \mathbb{S}(d)$ the d-dimensional random vector X in Proposition 4.2 satisfies*

$$\hat{G}(A) := \frac{1}{2}\hat{\mathbb{E}}[\langle AX, X \rangle], \quad p \in \mathbb{R}^d \tag{9}$$

and the following condition:

$$aX + b\bar{X} \overset{d}{=} \sqrt{a^2 + b^2}X, \quad \forall a, b \in \mathbb{R}, \tag{10}$$

where $\bar{X}$ is an independent copy of X. In particular, for each components X_i of X and $\bar{X}_i$ of $\bar{X}$, we have $\sqrt{2}\hat{\mathbb{E}}[X_i] = \hat{\mathbb{E}}[X_i + \bar{X}_i] = 2\hat{\mathbb{E}}[X_i]$ and $\sqrt{2}\hat{\mathbb{E}}[-X_i] = \hat{\mathbb{E}}[-X_i - \bar{X}_i] = 2\hat{\mathbb{E}}[-X_i]$ it follows that X has no mean uncertainty:

$$\hat{\mathbb{E}}[X_i] = \hat{\mathbb{E}}[-X_i] = 0, \quad i = 1, \cdots, d.$$

On the other hand, by Proposition 4.1 we can find a bounded subset $\hat{\Theta} \in \mathbb{S}_+(d)$ such that

$$\frac{1}{2}\hat{\mathbb{E}}[\langle AX, X\rangle] = \hat{G}(A) = \frac{1}{2} \sup_{Q \in \bar{\Theta}} tr[AQ], \quad A \in \mathbb{S}(d). \tag{11}$$

If $\hat{\Theta}$ is a singleton $\hat{\Theta} = \{Q\}$, then X is a classical zero-mean normal distributed with covariance Q. In general $\hat{\Theta}$ characterizes the covariance uncertainty of X.

Definition 4.5 *(**G-distribution**) The pair of d-dimensional random vectors (X, Y) in the above proposition is called G-distributed. X is said to be $\hat{G}$-normal distributed. We denote the distribution of X by $X \overset{d}{=} \mathcal{N}(0, \hat{\Theta})$.*

Proposition 4.8 and Corollary 4.9 show that a G-distribution is a uniquely defined sublinear distribution on $(\mathbb{R}^{2d}, C_{l.Lip}(\mathbb{R}^{2d}))$. We will show that a pair of G-distributed random vectors is characterized, or generated, by the following parabolic PDE defined on $[0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$:

$$\partial_t u - G(D_y u, D_x^2 u) = 0, \tag{12}$$

with Cauchy condition $u|_{t=0} = \varphi$, where $D_y = (\partial_{y_i})_{i=1}^d$, $D_x^2 = (\partial^2_{x_i, x_j})_{i,j=1}^d$. (12) is called the G-heat equation.

Remark 4.6 *We will use the notion of viscosity solutions to the generating heat equation (12). This notion was introduced by Crandall and Lions. For the existence and uniqueness of solutions and related very rich references we refer to Crandall, Ishii and Lions [2] (see Appendix for the uniqueness). We note that, in the situation where $\underline{\sigma}^2 > 0$, the viscosity solution (12) becomes a classical $C^{1+\frac{\alpha}{2}, 2+\alpha}$-solution (see [5] and the recent works of [1] and [18]). Readers can understand (12) in the classical meaning.*

Definition 4.7 *A real-valued continuous function $u \in C([0, T] \times \mathbb{R}^d)$ is called a viscosity subsolution (respectively, supersolution) of (12) if, for each function $\psi \in C_b^3((0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d)$ and for each minimum (respectively, maximum) point $(t, x, y) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ of $\psi - u$, we have*

$$\partial_t \psi - G(D_y \psi, D_x^2 \psi) \leq 0 \quad (\textit{respectively}, \ \geq 0).$$

u is called a viscosity solution of (12) if it is both super and subsolution.

Proposition 4.8 *Let (X, Y) be G-distributed. For each $\varphi \in C_{l.Lip}(\mathbb{R}^d \times \mathbb{R}^d)$ we define a function*

$$u(t, x, y) := \hat{\mathbb{E}}[\varphi((x + \sqrt{t}X, y + tY)], \ (t, x) \in [0, \infty) \times \mathbb{R}.$$

Then we have

$$u(t + s, x, y) = \hat{\mathbb{E}}[u(t, x + \sqrt{s}X, y + sY)], \quad s \geq 0. \tag{13}$$

We also have the estimates: For each $T > 0$ there exist constants $C, k > 0$ such that, for all $t, s \in [0, T]$ and $x, y \in \mathbb{R}$,

$$|u(t, x, y) - u(t, \bar{x}, \bar{y})| \leq C(1 + |x|^k + |y|^k + |\bar{x}|^k + |\bar{y}|^k)|x - y| \tag{14}$$

and

$$|u(t, x, y) - u(t + s, x + y)| \leq C(1 + |x|^k + |y|^k)(s + |s|^{1/2}). \tag{15}$$

Moreover, u is the unique viscosity solution, continuous in the sense of (14) and (15), of the generating PDE (12).

Proof. Since

$$\begin{aligned}
u(t, x, y) - u(t, \bar{x}, \bar{y}) &= \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X, y + tY)] - \hat{\mathbb{E}}[\varphi(\bar{x} + \sqrt{t}X, \bar{y} + tY)] \\
&\leq \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X, y + tY) - \varphi(\bar{x} + \sqrt{t}X, \bar{y} + tY)] \\
&\leq \hat{\mathbb{E}}[C_1(1 + |X|^k + |Y|^k + |x|^k + |y|^k + |\bar{x}|^k + |\bar{y}|^k)] \\
&\quad \times (|x - \bar{x}| + |y - \bar{y}|) \\
&\leq C(1 + |x|^k + |y|^k + |\bar{x}|^k + |\bar{y}|^k)(|x - \bar{x}| + |y - \bar{y}|).
\end{aligned}$$

We then have (14). Let $(\bar{X}, \bar{Y})$ be an independent copy of (X, Y). Since (X, Y) is G-distributed, then

$$\begin{aligned}
u(t + s, x, y) &= \hat{\mathbb{E}}[\varphi(x + \sqrt{t + s}X, y + (t + s)Y)] \\
&= \hat{\mathbb{E}}[\varphi(x + \sqrt{s}X + \sqrt{t}\bar{X}, y + sY + t\bar{Y})] \\
&= \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x + \sqrt{s}\widetilde{x} + \sqrt{t}\bar{X}, y + s\widetilde{y} + t\bar{Y})]_{(\widetilde{x}, \widetilde{y}) = (X, Y)}] \\
&= \hat{\mathbb{E}}[u(t, x + \sqrt{s}X, y + sY)].
\end{aligned}$$

We thus obtain (13). From this and (14) it follows that

$$\begin{aligned}
&u(t + s, x, y) - u(t, x, y) = \hat{\mathbb{E}}[u(t, x + \sqrt{s}X, y + sY) - u(t, x)] \\
&\leq \hat{\mathbb{E}}[C_1(1 + |x|^k + |y|^k + |X|^k + |Y|^k)(\sqrt{s}|X| + s|Y|)].
\end{aligned}$$

Thus we obtain (15). Now, for a fixed $(t, x, y) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$, let $\psi \in C_b^{1,3}([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d)$ be such that $\psi \geq u$ and $\psi(t, x, y) = u(t, x, y)$. By (13) and Taylor's expansion it follows that, for $\delta \in (0,t)$,

$$\begin{aligned}
0 &\le \hat{\mathbb{E}}[\psi(t-\delta, x+\sqrt{\delta}X, y+\delta Y) - \psi(t,x,y)] \\
&\le \bar{C}(\delta^{3/2}+\delta^2) - \partial_t\psi(t,x,y)\delta \\
&+ \hat{\mathbb{E}}[\langle D_x\psi(t,x,y), X\rangle \sqrt{\delta} + \langle D_y\psi(t,x,y), Y\rangle \delta + \frac{1}{2}\left\langle D_x^2\psi(t,x,y)X, X\right\rangle \delta] \\
&= -\partial_t\psi(t,x,y)\delta + \hat{\mathbb{E}}[\langle D_y\psi(t,x,y), Y\rangle + \frac{1}{2}\left\langle D_x^2\psi(t,x,y)X, X\right\rangle]\delta + \bar{C}(\delta^{3/2}+\delta^2) \\
&= -\partial_t\psi(t,x,y)\delta + \delta G(D_y\psi, D_x^2\psi)(t,x,y) + \bar{C}(\delta^{3/2}+\delta^2).
\end{aligned}$$

From which it is easy to check that

$$[\partial_t\psi - G(D_y\psi, D_x^2\psi)](t,x,y) \le 0.$$

Thus u is a viscosity supersolution of (12). Similarly we can prove that u is a viscosity subsolution of (12). ■

Corollary 4.9 *If both (X,Y) and $(\bar{X},\bar{Y})$ are G-distributed with the same G, i.e.,*

$$G(p,A) := \hat{\mathbb{E}}[\frac{1}{2}\langle AX, X\rangle + \langle p, Y\rangle] = \hat{\mathbb{E}}[\frac{1}{2}\left\langle A\bar{X}, \bar{X}\right\rangle + \left\langle p, \bar{Y}\right\rangle], \ \ \forall (p,A) \in \mathbb{S}(d)\times\mathbb{R}^d.$$

then $(X,Y) \overset{d}{=} (\bar{X},\bar{Y})$. In particular, $X \overset{d}{=} -X$.

Proof. For each $\varphi \in C_{l.Lip}(\mathbb{R}^d\times\mathbb{R}^d)$ we set

$$\begin{aligned}
u(t,x,y) &:= \hat{\mathbb{E}}[\varphi(x+\sqrt{t}X, y+tY)], \\
\bar{u}(t,x,y) &:= \hat{\mathbb{E}}[\varphi(x+\sqrt{t}\bar{X}, y+t\bar{Y})], \ (t,x) \in [0,\infty)\times\mathbb{R}.
\end{aligned}$$

By the above Proposition, both u and $\bar{u}$ are viscosity solutions of the G-heat equation (12) with Cauchy condition $u|_{t=0} = \bar{u}|_{t=0} = \varphi$. It follows from the uniqueness of the viscosity solution that $u \equiv \bar{u}$. In particular

$$\hat{\mathbb{E}}[\varphi(X,Y)] = \hat{\mathbb{E}}[\varphi(\bar{X},\bar{Y})].$$

Thus $(X,Y) \overset{d}{=} (\bar{X},\bar{Y})$. ■

Corollary 4.10 *Let (X,Y) be G-distributed. For each $\psi \in C_{l.Lip}(\mathbb{R}^d)$ we define a function*

$$v(t,x) := \hat{\mathbb{E}}[\psi((x+\sqrt{t}X+tY)], \ (t,x) \in [0,\infty)\times\mathbb{R}^d.$$

Then v is the unique viscosity solution of the following parabolic PDE

$$\partial_t v - G(D_x v, D_x^2 v) = 0, \qquad v|_{t=0} = \psi. \tag{16}$$

Moreover we have $v(t,x+y) \equiv u(t,x,y)$, where u is the solution of the PDE (12) with initial condition $u(t,x,y)|_{t=0} = \psi(x+y)$.

4.1 Proof of Proposition 4.2

We now proceed to prove Proposition 4.2. Let $u = u^\varphi$ be the unique viscosity solution of the G-heat equation (12) with $u^\varphi|_{t=0} = \varphi$, Then we take $\widetilde{\Omega} = \mathbb{R}^{2d}$, $\widetilde{\mathcal{H}} = C_{l.Lip}(\mathbb{R}^{2d})$, $\widetilde{\omega} = (x, y) \in \mathbb{R}^{2d}$. The corresponding sublinear expectation $\widetilde{\mathbb{E}}[\cdot]$ is defined by, for each $\xi \in \mathcal{H}$ of the form $\xi(\omega) = (\varphi(x, y))_{(x,y)\in\mathbb{R}^{2d}} \in C_{l.Lip}(\mathbb{R}^{2d})$, $\widetilde{\mathbb{E}}[\xi] = u^\varphi(1, 0)$. The monotonicity and sub-linearity of u^φ with respect to φ are known in the theory of viscosity solution. For reader's convenience we provide a new and simple proof in the Appendix (see Corollary 6.4 and Corollary 6.5). The positive homogeneity of $\widetilde{\mathbb{E}}[\cdot]$ is easy to be checked.

We now consider a pairs of d-dimensional random vectors $(\widetilde{X}, \widetilde{Y})(\omega) = (x, y)$. We have

$$\hat{\mathbb{E}}[\varphi(\widetilde{X}, \widetilde{Y})] = u^\varphi(1, 0), \quad \forall \varphi \in C_{l.Lip}(\mathbb{R}^{2d}).$$

In particular, just set $\varphi_0(x, y) = \frac{1}{2}\langle Ax, x\rangle + \langle p, y\rangle$, we can check that

$$u^{\varphi_0}(t, x, y) := G(p, A)t + \frac{1}{2}\langle Ax, x\rangle + \langle p, y\rangle .$$

We thus have

$$\widetilde{\mathbb{E}}[\frac{1}{2}\left\langle A\widetilde{X}, \widetilde{X}\right\rangle + \left\langle p, \widetilde{Y}\right\rangle] = u^{\varphi_0}(t, 0)|_{t=1} = G(p, A), \quad (p, A) \in \mathbb{R}^d \times \mathbb{S}(n).$$

To prove that the distribution of $(\widetilde{X}, \widetilde{Y})$ satisfies condition (4), we follow Definition 3.7 to construct a product space of sublinear expectation

$$(\Omega, \mathcal{H}, \hat{\mathbb{E}}) = (\widetilde{\Omega} \times \widetilde{\Omega}, \widetilde{\mathcal{H}} \times \widetilde{\mathcal{H}}, \widetilde{\mathbb{E}} \times \widetilde{\mathbb{E}})$$

and introduce two pair of random vectors

$$(X, Y)(\omega_1, \omega_2) = \omega_1, \ (\bar{X}, \bar{Y})(\omega_1, \omega_2) = \omega_2, \quad (\omega_1, \omega_2) \in \widetilde{\Omega} \times \widetilde{\Omega}.$$

By Proposition 3.8 both $(X, Y) \overset{d}{=} (\widetilde{X}, \widetilde{Y})$ and $(\bar{X}, \bar{Y})$ is an independent copy of (X, Y). For each $\varphi \in C_{l.Lip}(\mathbb{R}^{2d})$ and for each fixed $\lambda > 0$, $(\bar{x}, \bar{y}) \in \mathbb{R}^{2d}$, since the function v defined by $v(t, x, y) := u^\varphi(\lambda t, \bar{x} + \sqrt{\lambda}x, \bar{y} + \lambda y)$ solves exactly the same equation (12) but with Cauchy condition

$$v|_{t=0} = \varphi(\bar{x} + \sqrt{\lambda} \times \cdot, \bar{y} + \lambda \times \cdot).$$

Thus

$$\begin{aligned}\hat{\mathbb{E}}[\varphi(\bar{x} + \sqrt{\lambda}X, \bar{y} + \lambda Y)] &= v(t, \bar{x}, \bar{y})|_{t=1} \\ &= u^{\varphi(\sqrt{\lambda}\times\cdot, \lambda\times\cdot)}(t, \bar{x}, \bar{y})|_{t=1} = u^\varphi(\lambda, \bar{x}, \bar{y}).\end{aligned}$$

By the definition of $\hat{\mathbb{E}}$, for each $t > 0$ and $s > 0$,

$$\begin{aligned}\hat{\mathbb{E}}[\varphi(\sqrt{t}X + \sqrt{s}\bar{X}, tY + s\bar{Y})] &= \widetilde{\mathbb{E}}[\widetilde{\mathbb{E}}[\varphi(\sqrt{t}x + \sqrt{s}\bar{X}, ty + s\bar{Y})]_{(x,y)=(X,Y)}] \\ &= u^{u^\varphi(s,\cdot,\cdot)}(t, 0, 0) = u^\varphi(t + s, 0, 0) \\ &= \hat{\mathbb{E}}[\varphi(\sqrt{t+s}X, (t + s)Y)].\end{aligned}$$

Namely $(\sqrt{t}X+\sqrt{s}\bar{X}, tY+s\bar{Y}) \overset{d}{=} (\sqrt{t+s}X, (t+s)Y)$. Thus the distribution of (X,Y) satisfies condition (4).

It remains to check that the functional $\widetilde{\mathbb{E}}[\cdot] : C_{l.Lip}(\mathbb{R}^{2d}) \mapsto \mathbb{R}$ forms a sublinear expectation, i.e., (a)-(d) of Definition 2.2 are satisfied. Indeed, (a) is simply the consequence of comparison theorem, or the maximum principle of viscosity solution (see [CIL], the prove of this comparison theorem as well as the sub-additivity (c) are given in the Appendix of [P6]). It is also easy to check that, when $\varphi \equiv c$, the unique solution of (12) is also $u \equiv c$; hence (b) holds true. (d) also holds since $u^{\lambda\varphi} = \lambda u^{\varphi}$, $\lambda \geq 0$. The proof is complete.

5 Central Limit Theorem

Theorem 5.1 *(Central Limit Theorem) Let a sequence $\{(X_i, Y_i)\}_{i=1}^{\infty}$ of $\mathbb{R}^d \times \mathbb{R}^d$-valued random variables in $(\mathcal{H}, \hat{\mathbb{E}})$. We assumed that $(X_{i+1}, Y_{i+1}) \overset{d}{=} (X_i, Y_i)$ and (X_{i+1}, Y_{i+1}) is independent to $\{(X_1, Y_1), \cdots, (X_i, Y_i)\}$ for each $i = 1, 2, \cdots$. We assume furthermore that,*

$$\hat{\mathbb{E}}[X_1] = \hat{\mathbb{E}}[-X_1] = 0,$$

Then the sequence $\{\bar{S}_n\}_{n=1}^{\infty}$ defined by

$$\bar{S}_n := \sum_{i=1}^{n} (\frac{X_i}{\sqrt{n}} + \frac{Y_i}{n})$$

converges in law to $\xi + \zeta$:

$$\lim_{n\to\infty} \hat{\mathbb{E}}[\varphi(\bar{S}_n)] = \widetilde{\mathbb{E}}[\varphi(\xi+\zeta)], \tag{17}$$

for all functions $\varphi \in C(\mathbb{R}^d)$ satisfying a polynomial growth condition, where (ξ, ζ) is a pair of G-distributed random vectors and where the sublinear function $G : \mathbb{S}(d) \times \mathbb{R}^d \mapsto \mathbb{R}$ is defined by

$$G(p, A) := \hat{\mathbb{E}}[\langle p, Y_1 \rangle + \frac{1}{2} \langle AX_1, X_1 \rangle], \quad A \in \mathbb{S}(d), \quad p \in \mathbb{R}^d.$$

Corollary 5.2 *The sum $\sum_{i=1}^{n} \frac{X_i}{\sqrt{n}}$ converges in law to $\mathcal{N}(0, \hat{\Theta})$, where the subset $\hat{\Theta} \subset \mathbb{S}_+(d)$ is defined in (11) for $\hat{G}(A) = G(0, A)$, $A \in \mathbb{S}(d)$. The sum $\sum_{i=1}^{n} \frac{Y_i}{n}$ converges in law to $\mathcal{U}(\bar{\Theta})$, where the subset $\bar{\Theta} \subset \mathbb{R}^d$ is defined in (7) for $\bar{G}(p) = G(p, 0)$, $p \in \mathbb{R}^d$. If we take in particular $\varphi(y) = d_{\bar{\Theta}}(y) = \inf\{|x-y| : x \in \bar{\Theta}\}$, then by (8) we have the following generalized law of large number:*

$$\lim_{n\to\infty} \hat{\mathbb{E}}[d_{\bar{\Theta}}(\sum_{i=1}^{n} \frac{Y_i}{n})] = \sup_{\theta \in \bar{\Theta}} d_{\bar{\Theta}}(\theta) = 0. \tag{18}$$

Remark 5.3 *If Y_i has no mean-uncertainty, or in other words, $\bar{\Theta}$ is a singleton: $\bar{\Theta} = \{\bar{\theta}\}$ then (18) becomes*

$$\lim_{n\to\infty} \hat{\mathbb{E}}[|\sum_{i=1}^{n} \frac{Y_i}{n} - \bar{\theta}|] = 0.$$

To our knowledge, the law of large numbers with non-additive probability measures have been investigated with a quite different framework and approach from ours (see [9], [10]).

To prove this theorem we first give

Lemma 5.4 *We assume the same condition as Theorem 5.1. We assume furthermore that there exists $\beta > 0$ such that, for each A, $\bar{A} \in \mathbb{S}(d)$ with $A \geq \bar{A}$, we have*

$$\hat{\mathbb{E}}[\langle AX_1, X_1\rangle] - \hat{\mathbb{E}}[\langle \bar{A}X_1, X_1\rangle] \geq \beta tr[A - \bar{A}]. \tag{19}$$

Then (17) holds.

Proof. We first prove (17) for $\varphi \in C_{b.Lip}(\mathbb{R}^d)$. For a small but fixed $h > 0$, let V be the unique viscosity solution of

$$\partial_t V + G(DV, D^2V) = 0, \ (t,x) \in [0, 1+h] \times \mathbb{R}^d, \ V|_{t=1+h} = \varphi. \tag{20}$$

Since (ξ, ζ) is G-distributed we have

$$V(h, 0) = \widetilde{\mathbb{E}}[\varphi(\xi + \zeta)], \quad V(1+h, x) = \varphi(x) \tag{21}$$

Since (20) is a uniformly parabolic PDE and G is a convex function thus, by the interior regularity of V (see Krylov [5], Example 6.1.8 and Theorem 6.2.3), we have

$$\|V\|_{C^{1+\alpha/2, 2+\alpha}([0,1]\times\mathbb{R}^d)} < \infty, \text{ for some } \alpha \in (0,1).$$

We set $\delta = \frac{1}{n}$ and $S_0 = 0$. Then

$$\begin{aligned} V(1, \bar{S}_n) - V(0,0) &= \sum_{i=0}^{n-1} \{V((i+1)\delta, \bar{S}_{i+1}) - V(i\delta, \bar{S}_i)\} \\ &= \sum_{i=0}^{n-1} \{[V((i+1)\delta, \bar{S}_{i+1}) - V(i\delta, \bar{S}_{i+1})] + [V(i\delta, \bar{S}_{i+1}) - V(i\delta, \bar{S}_i)]\} \\ &= \sum_{i=0}^{n-1} \left\{I_\delta^i + J_\delta^i\right\} \end{aligned}$$

with, by Taylor's expansion,

$$J_\delta^i = \partial_t V(i\delta, \bar{S}_i)\delta + \frac{1}{2}\left\langle D^2V(i\delta, \bar{S}_i)X_{i+1}, X_{i+1}\right\rangle \delta + \left\langle DV(i\delta, \bar{S}_i), X_{i+1}\sqrt{\delta} + Y_{i+1}\delta\right\rangle$$

$$I_\delta^i = \int_0^1 [\partial_t V((i+\beta)\delta, \bar{S}_{i+1}) - \partial_t V(i\delta, \bar{S}_{i+1})] d\beta\delta + [\partial_t V(i\delta, \bar{S}_{i+1}) - \partial_t V(i\delta, \bar{S}_i)]\delta$$
$$+ \frac{1}{2}\left\langle D^2 V(i\delta, \bar{S}_i) X_{i+1}, Y_{i+1}\right\rangle \delta^{3/2} + \frac{1}{2}\left\langle D^2 V(i\delta, \bar{S}_i) Y_{i+1}, Y_{i+1}\right\rangle \delta$$
$$+ \int_0^1 \int_0^1 \left\langle \Theta_{\beta\gamma}^i (X_{i+1}\sqrt{\delta} + Y_{i+1}\delta), X_{i+1}\sqrt{\delta} + Y_{i+1}\delta \right\rangle \gamma d\beta d\gamma$$

with

$$\Theta_{\beta\gamma}^i = D^2 V(i\delta, \bar{S}_i + \gamma\beta(X_{i+1}\sqrt{\delta} + Y_{i+1}\delta) - D^2 V(i\delta, \bar{S}_i).$$

Thus

$$\hat{\mathbb{E}}[\sum_{i=0}^{n-1} J_\delta^i] - \hat{\mathbb{E}}[-\sum_{i=0}^{n-1} I_\delta^i] \leq \hat{\mathbb{E}}[V(1, \bar{S}_n)] - V(0,0) \leq \hat{\mathbb{E}}[\sum_{i=0}^{n-1} J_\delta^i] + \hat{\mathbb{E}}[\sum_{i=0}^{n-1} I_\delta^i]. \quad (22)$$

We now prove that $\hat{\mathbb{E}}[\sum_{i=0}^{n-1} J_\delta^i] = 0$. For the 3rd term of J_δ^i we have:

$$\hat{\mathbb{E}}[\left\langle DV(i\delta, \bar{S}_i), X_{i+1}\sqrt{\delta}\right\rangle] = \hat{\mathbb{E}}[-\left\langle DV(i\delta, \bar{S}_i), X_{i+1}\sqrt{\delta}\right\rangle] = 0.$$

For the second term, we have, from the definition of the function G,

$$\hat{\mathbb{E}}[J_\delta^i] = \hat{\mathbb{E}}[\partial_t V(i\delta, \bar{S}_i) + G(DV(i\delta, \bar{S}_i), D^2 V(i\delta, \bar{S}_i))]\delta.$$

We then combine the above two equalities with $\partial_t V + G(DV, D^2 V) = 0$ as well as the independence of (X_{i+1}, Y_{i+1}) to $\{(X_1, Y_1), \cdots, (X_i, Y_i)\}$, it follows that

$$\hat{\mathbb{E}}[\sum_{i=0}^{n-1} J_\delta^i] = \hat{\mathbb{E}}[\sum_{i=0}^{n-2} J_\delta^i] = \cdots = 0.$$

Thus (22) can be rewritten as

$$-\hat{\mathbb{E}}[-\sum_{i=0}^{n-1} I_\delta^i] \leq \hat{\mathbb{E}}[V(1, \bar{S}_n)] - V(0,0) \leq \hat{\mathbb{E}}[\sum_{i=0}^{n-1} I_\delta^i].$$

But since both $\partial_t V$ and $D^2 V$ are uniformly α-hölder continuous in x and $\frac{\alpha}{2}$-hölder continuous in t on $[0,1] \times \mathbb{R}$, we then have

$$|I_\delta^i| \leq C\delta^{1+\alpha/2}[1 + |X_{i+1}|^{2+\alpha} + |Y_1|^{2+\alpha}].$$

It follows that

$$\hat{\mathbb{E}}[|I_\delta^i|] \leq C\delta^{1+\alpha/2}(1 + \hat{\mathbb{E}}[|X_1|^{2+\alpha} + |Y_1|^{2+\alpha}]).$$

Thus

$$-C(\frac{1}{n})^{\alpha/2}(1 + \hat{\mathbb{E}}[|X_1|^{2+\alpha} + |Y_1|^{2+\alpha}]) \leq \hat{\mathbb{E}}[V(1, \bar{S}_n)] - V(0,0)$$
$$\leq C(\frac{1}{n})^{\alpha/2}(1 + \hat{\mathbb{E}}[|X_1|^{2+\alpha} + |Y_1|^{2+\alpha}]).$$

As $n \to \infty$ we have

$$\lim_{n\to\infty} \hat{\mathbb{E}}[V(1, \bar{S}_n)] = V(0,0). \tag{23}$$

On the other hand, for each $t, t' \in [0, 1+h]$ and $x \in \mathbb{R}^d$, we have

$$|V(t,x) - V(t',x)| \leq C(\sqrt{|t-t'|.} + |t-t'|).$$

,

Thus $|V(0,0) - V(h,0)| \leq C(\sqrt{h} + h)$ and, by (23),

$$|\hat{\mathbb{E}}[V(1,\bar{S}_n)] - \hat{\mathbb{E}}[\varphi(\bar{S}_n)]| = |\hat{\mathbb{E}}[V(1,\bar{S}_n)] - \hat{\mathbb{E}}[V(1+h,\bar{S}_n)]| \leq C(\sqrt{h} + h).$$

It follows form (21) and (23) that

$$\limsup_{n\to\infty} |\hat{\mathbb{E}}[\varphi(\bar{S}_n)] - \widetilde{\mathbb{E}}[\varphi(\xi+\zeta)]| \leq 2C(\sqrt{h} + h).$$

Since h can be arbitrarily small we thus have

$$\lim_{n\to\infty} \hat{\mathbb{E}}[\varphi(\bar{S}_n)] = \widetilde{\mathbb{E}}[\varphi(\xi)].$$

■

We now give

Proof of Theorem 5.1. For the case when the uniform Elliptic condition 19 does not hold, we first follow an idea of Song [17] to introduce a perturbation to prove the above convergence for $\varphi \in C_{b.Lip}(\mathbb{R}^d)$. According to Definition 3.7 and Proposition 3.8 we can construct a sublinear expectation space $(\bar{\Omega}, \bar{\mathcal{H}}, \bar{\mathbb{E}})$ and a sequence of three random vectors $\{(\bar{X}_i, \bar{Y}_i, \bar{\eta}_i)\}_{i=1}^\infty$ such that, for each $n = 1, 2, \cdots$, $\{(\bar{X}_i, \bar{Y}_i)\}_{i=1}^n \overset{d}{=} \{(X_i, Y_i)\}_{i=1}^n$ and $(\bar{X}_{n+1}, \bar{Y}_{n+1}, \bar{\eta}_{n+1})$ is independent to $\{(\bar{X}_i, \bar{Y}_i, \bar{\eta}_i)\}_{i=1}^n$ and, moreover,

$$\bar{\mathbb{E}}[\psi(X_i, Y_i, \eta_i)] = \frac{1}{\sqrt{2\pi d}} \int_{\mathbb{R}^d} \hat{\mathbb{E}}[\psi(X_i, Y_i, x)] e^{-\frac{|x|^2}{2}} dx, \quad \forall \psi \in C_{l.Lip}(\mathbb{R}^{3\times d}).$$

We then use the following perturbation $X_i^\varepsilon = X_i + \varepsilon \eta_i$ for a fixed $\varepsilon > 0$. It is seen that the sequence $\{(\bar{X}_i^\varepsilon, \bar{Y}_i)\}_{i=1}^\infty$ satisfies all conditions in the above CLT, in particular

$$G_\varepsilon(p, A) := \bar{\mathbb{E}}[\frac{1}{2}\left\langle A\bar{X}_1^\varepsilon, \bar{X}_1^\varepsilon \right\rangle + \left\langle p, \bar{Y}_1 \right\rangle] = G(p, A) + \frac{\varepsilon^2}{2}\mathrm{tr}[A].$$

Thus it is strictly elliptic. We then can apply Lemma 5.4 to

$$\bar{S}_n^\varepsilon := \sum_{i=1}^n \left(\frac{X_i^\varepsilon}{\sqrt{n}} + \frac{Y_i}{n}\right) = \bar{S}_n + \varepsilon J_n, \quad J_n = \sum_{i=1}^n \frac{\eta_i}{\sqrt{n}}$$

and obtain

$$\lim_{n\to\infty} \hat{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon)] = \widetilde{\mathbb{E}}[\varphi(\xi + \zeta + \varepsilon\eta)].$$

where (ξ, ζ) is G-distributes and

$$\widetilde{\mathbb{E}}[\psi(\xi+\zeta,\eta)] = \frac{1}{\sqrt{2\pi d}}\int_{\mathbb{R}^d} \widetilde{\mathbb{E}}[\psi(\xi+\zeta,x)]e^{\frac{-|x|^2}{2}}dx, \quad \psi \in C_{l.Lip}(\mathbb{R}^{2d}).$$

Thus $(\xi+\varepsilon\eta,\zeta)$ is G_ε-distributed. But we have

$$\begin{aligned}|\hat{\mathbb{E}}[\varphi(\bar{S}_n)] - \hat{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon)]| &= |\hat{\mathbb{E}}[\varphi(\bar{S}_n)] - \hat{\mathbb{E}}[\varphi(\bar{S}_n+\varepsilon J_n)]| \\ &\leq \varepsilon C\hat{\mathbb{E}}[|J_n|] \leq C\varepsilon\end{aligned}$$

and similarly $|\widetilde{\mathbb{E}}[\varphi(\xi)] - \widetilde{\mathbb{E}}[\varphi(\xi+\varepsilon\eta)]| \leq C\varepsilon$. Sine ε can be arbitrarily small, it follows that

$$\lim_{n\to\infty}\hat{\mathbb{E}}[\varphi(\bar{S}_n)] = \widetilde{\mathbb{E}}[\varphi(\xi+\zeta)], \quad \forall\varphi \in C_{b.Lip}(\mathbb{R}^d).$$

On the other hand, it is easy to check that $\sup_n \hat{\mathbb{E}}[|\bar{S}_n|]+\widetilde{\mathbb{E}}[|\xi+\zeta|] < \infty$. We then can apply the following lemma to prove that the above convergence holds for the case where φ in $C(\mathbb{R}^d)$ with a polynomial growth condition. The proof is complete.

Lemma 5.5 *Let $(\hat{\Omega},\hat{\mathcal{H}},\hat{\mathbb{E}})$ and $(\widetilde{\Omega},\widetilde{\mathcal{H}},\widetilde{\mathbb{E}})$ be two sublinear expectation space and let $\zeta \in \hat{\mathcal{H}}$ and $\zeta_n \in \widetilde{\mathcal{H}}$, $n = 1,2,\cdots$, be given. We assume that, for a given $p \geq 1$ we have $\sup_n \hat{\mathbb{E}}[|Y_n|^p] + \widetilde{\mathbb{E}}[|Y|^p] \leq C$. If the convergence $\lim_{n\to\infty}\hat{\mathbb{E}}[\varphi(Y_n)] = \widetilde{\mathbb{E}}[\varphi(Y)]$ holds for each $\varphi \in C_{b.Lip}(\mathbb{R}^d)$, then it also holds for all function $\varphi \in C(\mathbb{R}^d)$ with growth condition $|\varphi(x)| \leq C(1+|x|^{p-1})$.*

Proof. We first prove that the above convergence holds for $\varphi \in C_b(\mathbb{R}^d)$ with a compact support. In this case, for each $\varepsilon > 0$, we can find a $\bar{\varphi} \in C_{b.Lip}(\mathbb{R}^d)$ such that $\sup_{x\in\mathbb{R}^d}|\varphi(x)-\bar{\varphi}(x)| \leq \frac{\varepsilon}{2}$. We have

$$\begin{aligned}&|\hat{\mathbb{E}}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \leq |\hat{\mathbb{E}}[\varphi(Y_n)] - \hat{\mathbb{E}}[\bar{\varphi}(Y_n)]| + |\widetilde{\mathbb{E}}[\varphi(Y)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y)]| \\ &+ |\hat{\mathbb{E}}[\bar{\varphi}(Y_n)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y_n)]| \leq \varepsilon + |\hat{\mathbb{E}}[\bar{\varphi}(Y_n)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y)]|.\end{aligned}$$

Thus $\limsup_{n\to\infty}|\hat{\mathbb{E}}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \leq \varepsilon$. The convergence must hold since ε can be arbitrarily small.

Now let φ be an arbitrary $C_b(\mathbb{R}^n)$-function. For each $N > 0$ we can find $\varphi_1, \varphi_2 \in C_b(\mathbb{R}^d)$ such that $\varphi = \varphi_1 + \varphi_2$ where φ_1 has a compact support and $\varphi_2(x) = 0$ for $|x| \leq N$, and $|\varphi_2(x)| \leq |\varphi(x)|$ for all x. It is clear that

$$|\varphi_2(x)| \leq \frac{\bar{C}(1+|x|^p)}{N}, \quad \forall x, \quad \text{where } \bar{C} = \sup_{x\in\mathbb{R}^d}|\varphi(x)|.$$

Thus

$$\begin{aligned}|\hat{\mathbb{E}}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| &= |\hat{\mathbb{E}}[\varphi_1(Y_n)+\varphi_2(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)+\varphi_2(Y)]| \\ &\leq |\hat{\mathbb{E}}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \hat{\mathbb{E}}[|\varphi_2(Y_n)|] + \widetilde{\mathbb{E}}[|\varphi_2(Y)|] \\ &\leq |\hat{\mathbb{E}}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \frac{\bar{C}}{N}(\hat{\mathbb{E}}[|Y_n| + \widetilde{\mathbb{E}}[|Y|]) \\ &\leq |\hat{\mathbb{E}}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \frac{\bar{C}C}{N}.\end{aligned}$$

We thus have $\limsup_{n\to\infty} |\hat{\mathbb{E}}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \leq \frac{\bar{C}C}{N}$. Since N can be arbitrarily large thus $\hat{\mathbb{E}}[\varphi(Y_n)]$ must converge to $\widetilde{\mathbb{E}}[\varphi(Y)]$. ■

6 Appendix: some basic results of viscosity solutions

We will use the following well-known result in viscosity solution theory (see Theorem 8.3 of Crandall Ishii and Lions [2]).

Theorem 6.1 *Let $u_i \in USC((0,T) \times Q_i)$ for $i = 1, \cdots, k$ where Q_i is a locally compact subset of $\mathbb{R}^{N_i}$. Let φ be defined on an open neighborhood of $(0,T) \times Q_1 \times \cdots \times Q_k$ and such that $(t, x_1, \cdots, x_k)$ is once continuously differentiable in t and twice continuously differentiable in $(x_1, \cdots, x_k) \in Q_1 \times \cdots \times Q_k$. Suppose that $\hat{t} \in (0,T)$, $\hat{x}_i \in Q_i$ for $i = 1, \cdots, k$ and*

$$\begin{aligned} w(t, x_1, \cdots, x_k) &:= u_1(t, x_1) + \cdots + u_k(t, x_k) - \varphi(t, x_1, \cdots, x_k) \\ &\leq w(\hat{t}, \hat{x}_1, \cdots, \hat{x}_k) \end{aligned}$$

for $t \in (0,T)$ and $x_i \in Q_i$. Assume, moreover that there is an $r > 0$ such that for every $M > 0$ there is a C such that for $i = 1, \cdots, k$,

$$\begin{aligned} &b_i \leq C \text{ whenever } (b_i, q_i, X_i) \in \mathcal{P}^{2,+} u_i(t, x_i), \\ &|x_i - \hat{x}_i| + |t - \hat{t}| \leq r \text{ and } |u_i(t, x_i)| + |q_i| + \|X_i\| \leq M. \end{aligned} \tag{24}$$

Then for each $\varepsilon > 0$, there are $X_i \in \mathbb{S}(N_i)$ such that
(i) $(b_i, D_{x_i}\varphi(\hat{t}, \hat{x}_1, \cdots, \hat{x}_k), X_i) \in \overline{\mathcal{P}}^{2,+} u_i(\hat{t}, \hat{x}_i)$, $i = 1, \cdots, k$;
(ii)

$$-(\frac{1}{\varepsilon} + \|A\|) \leq \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_k \end{bmatrix} \leq A + \varepsilon A^2$$

(iii) $b_1 + \cdots + b_k = \partial_t \varphi(\hat{t}, \hat{x}_1, \cdots, \hat{x}_k)$
where $A = D^2\varphi(\hat{x}) \in \mathbb{S}(N_1 + \cdots + N_k)$.

Observe that the above conditions (24) will be guaranteed by having u_i be subsolutions of parabolic equations given in the following two theorems, which is an improved version of the one in the Appendix of [16].

Theorem 6.2 *(Domination Theorem) $u_i \in USC([0,T] \times \mathbb{R}^N)$ be subsolutions of*

$$\partial_t u - G_i(t, x, u, Du, D^2u) = 0, \qquad i = 1, \cdots, k, \tag{25}$$

on $(0,T) \times \mathbb{R}^N$ such that, for given constants $\beta_i > 0$, $i = 1, \cdots, k$, $\left(\sum_{i=1}^k u_i(t,x)\right)^+ \to 0$, uniformly as $|x| \to \infty$. We assume that
(i) The functions

$$G_i : [0,T] \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{S}(N) \mapsto \mathbb{R}, \quad i = 1, \cdots, k,$$

are continuous in the following sense: for each $t \in [0,T)$, v_1, $v_2 \in \mathbb{R}$, x, y, p, $q \in \mathbb{R}^N$ and $Y \in \mathbb{S}(N)$,

$$\begin{aligned}&[G_i(t,x,v,p,X) - G_i(t,y,v,p,X)]^- \\ &\leq \bar{\omega}(1+(T-t)^{-1}+|x|+|y|+|v|)\omega(|x-y|+|p|\cdot|x-y|)\end{aligned}$$

where ω, $\bar{\omega} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ are given continuous functions with $\omega(0) = 0$.
(ii) Given constants $\beta_i > 0$, $i = 1, \cdots, k$, the following domination condition holds for G_i:

$$\sum_{i=1}^{k} \beta_i G_i(t,x,v_i,p_i,X_i) \leq 0, \tag{26}$$

for each $(t,x) \in (0,T) \times \mathbb{R}^N$ and $(v_i,p_i,X_i) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{S}(N)$ such that $\sum_{i=1}^k \beta_i v_i \geq 0$, $\sum_{i=1}^k \beta_i p_i = 0$, $\sum_{i=1}^k \beta_i X_i \leq 0$.

Then a similar domination also holds for the solutions: If the sum of initial values $\sum_{i=1}^k \beta_i u_i(0,\cdot)$ is a non-positive and continuous function on $\mathbb{R}^N$, then $\sum_{i=1}^k \beta_i u_i(t,\cdot) \leq 0$, for all $t > 0$.

Proof. We first observe that for $\bar{\delta} > 0$ and for each $1 \leq i \leq k$, the functions defined by $\tilde{u}_i := u_i - \bar{\delta}/(T-t)$ is a subsolution of:

$$\partial_t \tilde{u}_i - \tilde{G}_i(t,x,\tilde{u}_i + \bar{\delta}/(T-t), D\tilde{u}_i, D^2\tilde{u}_i) \leq -\frac{\bar{\delta}}{(T-t)^2}$$

where $\tilde{G}_i(t,x,v,p,X) := G_i(t,x,v+\bar{\delta}/(T-t),p,X)$. It is easy to check that the functions $\tilde{G}_i$ satisfy the same conditions as G_i. Since $\sum_{i=1}^k \beta_i u_i \leq 0$ follows from $\sum_{i=2}^k \beta_i \tilde{u}_i \leq 0$ in the limit $\bar{\delta} \downarrow 0$, it suffices to prove the theorem under the additional assumptions

$$\begin{array}{c}\partial_t u_i - G_i(t,x,u_i,Du_i,D^2u_i) \leq -c, \quad c := \bar{\delta}/T^2, \\ \text{and } \lim_{t\to T} u_i(t,x) = -\infty, \quad \text{uniformly in } [0,T) \times \mathbb{R}^N.\end{array} \tag{27}$$

To prove the theorem, we assume to the contrary that

$$\sup_{(s,x)\in[0,T)\times\mathbb{R}^N} \sum_{i=1}^{k} \beta_i u_i(t,x) = m_0 > 0$$

We will apply Theorem 6.1 for $x = (x_1, \cdots, x_k) \in \mathbb{R}^{k\times N}$ and

$$w(t,x) := \sum_{i=1}^{k} \beta_i u_i(t,x_i), \quad \varphi(x) = \varphi_\alpha(x) := \frac{\alpha}{2} \sum_{i=1}^{k-1} |x_{i+1} - x_i|^2.$$

For each large $\alpha > 0$ the maximum of $w - \varphi_\alpha$ achieved at some (t^α, x^α) inside a compact subset of $[0,T) \times \mathbb{R}^{k\times N}$. Indeed, since

$$M_\alpha = \sum_{i=1}^{k} \beta_i u_i(t^\alpha, x_i^\alpha) - \varphi_\alpha(x^\alpha) \geq m_0,$$

thus t^α must be inside an interval $[0, T_0]$, $T_0 < T$ and x^α must be inside a compact set $\{x \in \times\mathbb{R}^{k\times N} : \sup_{t\in[0,T_0]} w(t,x) \geq \frac{m_0}{2}\}$. We can check that (see [2] Lemma 3.1)

$$\begin{cases} \text{(i)} \ \lim_{\alpha\to\infty} \varphi_\alpha(x^\alpha) = 0. \\ \text{(ii)} \ \lim_{\alpha\to\infty} M_\alpha = \lim_{\alpha\to\infty} \beta_1 u_1(t^\alpha, x_1^\alpha) + \cdots + \beta_k u_k(t^\alpha, x_k^\alpha)) \\ \qquad = \sup_{(t,x)\in[0,T)\times\mathbb{R}^N} [\beta_1 u_1(t,x) + \cdots + \beta_k u_k(t,x)] \\ \qquad\quad = [\beta_1 u_1(\hat{t},\hat{x}) + \cdots + \beta_k u_k(\hat{t},\hat{x})] = m_0. \end{cases} \tag{28}$$

where $(\hat{t}, \hat{x})$ is a limit point of (t^α, x_1^α). Since $u_i \in$ USC, for sufficiently large α, we have

$$\beta_1 u_1(t^\alpha, x_1^\alpha) + \cdots + \beta_k u_k(t^\alpha, x_k^\alpha) \geq \frac{m_0}{2}.$$

If $\hat{t} = 0$, we have $\limsup_{\alpha\to\infty} \sum_{i=1}^k \beta_i u_i(t^\alpha, x_i^\alpha) = \sum_{i=1}^k \beta_i u_i(0, \hat{x}) \leq 0$. We know that $\hat{t} > 0$ and thus t^α must be strictly positive for large α. It follows from Theorem 6.1 that, for each $\varepsilon > 0$ there exists $b_i^\alpha \in \mathbb{R}$, $X_i \in \mathbb{S}(N)$ such that

$$(b_i^\alpha, D_{x_i}\varphi(x^\alpha), X_i) \in \bar{J}_{Q_i}^{2;+}(u_i)(t^\alpha, x_i^\alpha), \quad \sum_{i=1}^k \beta_i b_i^\alpha = 0, \text{ for } i = 1, \cdots, k, \tag{29}$$

and such that

$$-(\frac{1}{\varepsilon} + \|A\|)I \leq \begin{pmatrix} \beta_1 X_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \beta_{k-1}X_{k-1} & 0 \\ 0 & \ldots & 0 & \beta_k X_k \end{pmatrix} \leq A + \varepsilon A^2, \tag{30}$$

where $A = D^2\varphi_\alpha(x^\alpha) \in \mathbb{S}(kN)$ is explicitly given by

$$A = \alpha J_{kN}, \text{ where } J_{kN} = \begin{pmatrix} I_N & -I_N \ 0 \ldots & 0 & -I_N \\ \vdots & \ddots & \vdots & \vdots \\ 0 & 0 \ldots - I_N & I_N & -I_N \\ -I_N & 0 \ldots 0 & -I_N & I_N \end{pmatrix}.$$

The second inequality of (30) implies $\sum_{i=1}^k \beta_i X_i \leq 0$. Setting

$$p_1^\alpha = D_{x_1}\varphi_\alpha(x^\alpha) = \beta_1^{-1}\alpha(2x_1^\alpha - x_3^\alpha - x_2^\alpha),$$
$$\vdots$$
$$p_k^\alpha = D_{x_k}\varphi_\alpha(x^\alpha) = \beta_k^{-1}\alpha(2x_k^\alpha - x_{k-1}^\alpha - x_1^\alpha).$$

Thus $\sum_{i=1}^k \beta_i p_i^\alpha = 0$. This with (29) and (27) it follows that

$$b_i^\alpha - G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \leq -c, \quad i = 1, \cdots, k.$$

By (28)-(i) we also have $\lim_{\alpha\to\infty} |p_i^\alpha| \cdot |x_i^\alpha - x_1^\alpha| \to 0$. This, together with the the domination condition (26) of G_i, implies

$$\begin{aligned}
-kc = -\sum_{i=1}^k \beta_i b_i^\alpha - kc &\geq -\sum_{i=1}^k \beta_i G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \\
&\geq -\sum_{i=1}^k \beta_i G_i(t^\alpha, x_1^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \\
&\quad -\sum_{i=1}^k \beta_i [G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) - G_i(t^\alpha, x_1^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i)]^- \\
&\geq -\sum_{i=1}^k \beta_i \bar{\omega}(1 + (T-T_0)^{-1} + |x_1^\alpha| + |x_i^\alpha| + |u_i(t^\alpha, x_i^\alpha)|)\omega(|x_i^\alpha - x_1^\alpha| + |p_i^\alpha| \cdot |x_i^\alpha - x_1^\alpha|)
\end{aligned}$$

The right side tends to zero as $\alpha \to \infty$, which induces a contradiction. The proof is complete. ■

Theorem 6.3 *(Domination Theorem) Let polynomial growth functions $u_i \in USC([0,T] \times \mathbb{R}^N)$ be subsolutions of*

$$\partial_t u - G_i(u, Du, D^2 u) = 0, \qquad i = 1, \cdots, k, \tag{31}$$

on $(0,T) \times \mathbb{R}^N$. We assume that $G_i : \mathbb{R} \times \mathbb{R}^N \times \mathbb{S}(\mathbb{N}) \mapsto \mathbb{R}$, $i = 1, \cdots, k$, are given continuous functions satisfying the following conditions: There exists a positive constant C, such that
(i) for all $\lambda \geq 0$

$$\lambda G_i(v, p, Y) \leq C G_i(\lambda v, \lambda p, \lambda Y);$$

(ii) Lipschitz condition:

$$|G_i(v_1, p, X) - G_i(v_2, q, Y)| \leq C(|v_1 - v_2| + |p - q| + \|X - Y\|),$$
$$\forall v_1, v_2 \in \mathbb{R},\ \forall p, q \in \mathbb{R}^N \ \text{ and } X, Y \in \mathbb{S}(N),$$

(iii) domination condition for G_i: for fixed constants $\beta_i > 0$, $i = 1, \cdots, k$,

$$\sum_{i=1}^k \beta_i G_i(v_i, p_i, X_i) \leq 0, \quad \text{for all } v_i \in \mathbb{R},\ p_i \in \mathbb{R}^N,\ X_i \in \mathbb{S}(N),$$
$$\text{such that } \sum_{i=1}^k \beta_i v_i \geq 0,\ \sum_{i=1}^k \beta_i p_i = 0,\ \sum_{i=1}^k \beta_i X_i \leq 0.$$

Then the following domination holds: If $\sum_{i=1}^k \beta_i u_i(0, \cdot)$ is a non-positive continuous function, then we have

$$\sum_{i=1}^k \beta_i u_i(t, x) \leq 0, \quad \forall (t, x) \in (0, T) \times \mathbb{R}^N.$$

Proof. We set $\xi(x) := (1+|x|^2)^{l/2}$ and

$$\tilde{u}_i(t,x) := u_i(t,x)e^{\lambda t}\xi^{-1}(x), \ i = 1, \cdots, k,$$

where l is chosen large enough so that $\sum_{i=1}^k |\tilde{u}_i(t,x)| \to 0$ uniformly. From condition (i) it is easy to check that for each $i = 1, \cdots$, $\tilde{u}_i$ is a subsolution of

$$\partial_t \tilde{u}_i - \tilde{G}_i(x, \tilde{u}_i, D\tilde{u}_i, D^2\tilde{u}_i) = 0, \tag{32}$$

where

$$\begin{aligned}&\tilde{G}_i(x,v,p,X) := -\lambda v \\ &\ + e^{\lambda t} C G_i(v, p + v\eta(x), X + p \otimes \eta(x) + \eta(x) \otimes p + v\kappa(x)).\end{aligned}$$

Here

$$\begin{aligned}\eta(x) &:= \xi^{-1}(x) D\xi(x) = k(1+|x|^2)^{-1}x, \\ \kappa(x) &:= \xi^{-1}(x) D^2\xi(x) = k(1+|x|^2)^{-1}I - k(k-2)(1+|x|^2)^{-2}x \otimes x.\end{aligned}$$

Since η and κ are uniformly bounded, one can choose a fixed but large enough $\lambda > 0$ such that $\tilde{G}_i(x,v,p,X)$ satisfies all conditions of G_i, $i = 1, \cdots, k$ in Theorem 6.2. The proof is complete by directly applying this theorem. ■

We have the following Corollaries which are basic in this paper:

Corollary 6.4 *(Comparison Theorem) Let $F_1, F_2 : \mathbb{R}^N \times \mathbb{S}(N) \mapsto \mathbb{R}$ be given functions satisfying similar conditions (i) and (ii) of Theorem 6.3. We also assume that, for each $p, q \in \mathbb{R}^N$ and X, $Y \in \mathbb{S}(N)$ such that $X \geq Y$, we have*

$$F_1(p, X) \geq F_2(p, Y).$$

Let $v_i \in$LSC$((0,T) \times \mathbb{R}^N)$, $i = 1, 2$, be respectively a viscosity supersolution of $\partial_t v - F_i(Dv, D^2v) = 0$ such that $v_1(0,\cdot) - v_2(0,\cdot)$ is a non-negative continuous function. Then we have $v_1(t,x) - v_2(t,x) \geq 0$ for all $(t,x) \in [0,\infty) \times \mathbb{R}^N$.

Proof. We set $\beta_1 = \beta_2 = 1$, $G_1(p,X) := -F_1(-p,-X)$ and $G_2 = F_2(p,X)$. It is observed that $u_1 := -v_1 \in$USC$((0,T) \times \mathbb{R}^N)$ is a viscosity subsolution of $\partial_t u - G_1(Du, D^2u) = 0$. For each p_1, $p_2 \in \mathbb{R}^N$ and X_1, $X_2 \in \mathbb{S}(N)$ such that $p_1 + p_2 = 0$ and $X_1 + X_2 \leq 0$, we also have

$$G_1(p_1, X_1) + G_2(p_2, X_2) = F_2(p_2, X_2) - F_1(p_2, -X_1) \leq 0$$

We thus can apply Theorem 6.3 to get $-u_1 + u_2 \leq 0$. The proof is complete. ■

Corollary 6.5 *(Domination Theorem) Let $F_i : \mathbb{R}^N \times \mathbb{S}(N) \mapsto \mathbb{R}$, $i = 0, 1$, given functions satisfying similar conditions (i) and (ii) of Theorem 6.3. Let $v_i \in$LSC$((0,T) \times \mathbb{R}^N)$ be viscosity supersolutions of $\partial_t v - F_i(Dv, D^2v) = 0$*

respectively for $i = 0, 1$ and let $v_2 \in USC((0,T)\times\mathbb{R}^N)$ be a viscosity subsolution of $\partial_t v - F_1(Dv, D^2v) = 0$. We assume that:

$$F_1(p,X) - F_1(q,Y) \le F_0(p-q, Z),$$
$$\forall p, q \in \mathbb{R}^N,\ X, Y, Z\ \in \mathbb{S}(N) \text{ such that } X - Y \le Z.$$

Then the following domination holds: If $v_0(0,\cdot)+v_1(0,\cdot)-v_2(0,\cdot)$ is a continuous and non-negative function then $v_0(t,\cdot)+v_1(t,\cdot)-v_2(t,\cdot) \ge 0$ for all $t > 0$.

Proof. We denote

$$G_i(p,X) := -F_i(-p,-X),\quad i = 0,1, \text{ and }\quad G_2(p,X) := F_1(p,X).$$

Observe that $u_i = -v_i \in$USC$((0,T)\times\mathbb{R}^N)$, $i = 0,1$, are viscosity subsolutions of $\partial_t u - G_i(Du, D^2u) = 0$, $i = 0,1$. We thus have, for each $X_0 + X_1 + X_2 \le 0$, $p_0 + p_1 + p_2 = 0$,

$$\begin{aligned} &G_0(p_0,X_0) + G_1(p_1,X_1) + G_2(p_2,X_2) \\ &= -F_0(-p_0,-X_0) - F_1(-p_1,-X_1) + F_1(p_2,X_2) \le 0. \end{aligned}$$

Theorem 6.3 can be applied, for the case $\beta_i = 1$, to get $\sum u_i \le 0$, or $v_0+v_1-v_2 \ge 0$. ■

Another co-product of Theorem 6.3 is:

Corollary 6.6 *(Concavity) Let $F : \mathbb{R}^N \times \mathbb{S}(N) \mapsto \mathbb{R}$ be a given function satisfying similar conditions (i) and (ii) of Theorem 6.3. We assume that F is monotone in X, i.e. $F(p,X) \ge F(p,Y)$ if $X \ge Y$, and that F is concave: for each fixed $\alpha \in (0,1)$,*

$$\alpha F(p,X)+(1-\alpha)F(q,Y) \le F(\alpha p+(1-\alpha)q, \alpha X+(1-\alpha)Y),\ \ \forall p,q \in \mathbb{R}^N,\ X, Y \in \mathbb{S}(N).$$

Let $v_i \in USC((0,T)\times\mathbb{R}^N)$, $i = 0,1$, be respectively viscosity subsolutions of $\partial_t v - F(Dv, D^2v) = 0$ and let $v \in LSC((0,T)\times\mathbb{R}^N)$ be viscosity supersolution of $\partial_t v - F(Dv, D^2v) = 0$ such that $\alpha v_1(0,\cdot)+(1-\alpha)v_2(0,\cdot)-v(0,\cdot)$ is a non-positive continuous function. Then for all $t \ge 0$ $\alpha v_1(t,\cdot)+(1-\alpha)v_2(t,\cdot)-v(t,\cdot) \ge 0$.

Proof. We set $\beta_1 = \alpha$, $\beta_2 = (1-\alpha)$, $\beta_3 = 1$ and denote

$$G_1(p,X) = G_2(p,X) := F(p,X),\ G_3(p,X) = -F(-p,-X).$$

Observe that $u_i = v_i \in$USC$((0,T)\times\mathbb{R}^N)$, $i = 1,2$, are viscosity subsolutions of $\partial_t u - G_i(Du, D^2u) = 0$, $u_3 = -v \in$ USC is a viscosity subsolution of $\partial_t u - G_3(Du, D^2u) = 0$. Since F is concave, thus for each $p_i \in \mathbb{R}^N$ and $X_i \in \mathbb{S}(N)$ such that $\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \le 0$, $\beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 = 0$, we have

$$\begin{aligned} \beta_1 G_1(p_1,X_1) + \beta_2 G_2(p_2,X_2) + \beta_3 G_3(p_3,X_3) &\le F(\beta_1 p_1 + \beta_2 p_2, \beta_1 X_1 + \beta_2 X_2) - F(-p_3,-X_3) \\ &\le F(-p_3, \beta_1 X_1 + \beta_2 X_2) - F(-p_3,-X_3) \\ &\le 0. \end{aligned}$$

Theorem 6.3 can be applied to prove that $\alpha v_1(t,\cdot)+(1-\alpha)v_2(t,\cdot) \le v(t,\cdot)$. The proof is complete. ■

References

[1] X. Cabre and L.A. Caffarelli, Fully nonlinear elliptic partial di erential equations, American Math. Society (1997).

[2] Crandall, M., Ishii, H., and Lions, P.-L. (1992) User'S Guide To Viscosity Solutions Of Second Order Partial Differential Equations, *Bulletin Of The American Mathematical Society,* **27**(1), 1-67.

[3] L. Denis, M. Hu and S. Peng, Function spaces and capacity related to a Sublinear Expectation: application to G-Brownian Motion Pathes, see arXiv:0802.1240v1 [math.PR] 9 Feb 2008.

[4] L. Denis and C. Martini, *A Theoretical Framework for the Pricing of Contingent Claims in the Presence of Model Uncertainty*, The Annals of Applied Probability, vol. 16, No. 2, pp 827-852, 2006.

[5] Krylov, N.V. Nonlinear Parabolic and Elliptic Equations of the Second Order, Reidel Publishing Company, 1987 (Original Russian version by Nauka, Moscow, 1985).

[6] P. Lévy, Calcul dés Probabilités, Gautier-Villars, 1925.

[7] P. Lévy, Processus Stochastic et Mouvement Brownian, Jacques Gabay, 2ème édition, Gautier-Villars, 1965.

[8] Maxwell, James Clerk, On the dynamic theory of gases, Brit. Assoc. Rep., 1859 (pt.2) p.9. in 1859;

[9] Marinacci, M. Limit laws for non-additive probabilities and their frequentist interpretation, Journal of Economic Theory 84, 145-195 1999.

[10] Maccheroni, F. and Marinacci, M. A strong law of large numbers for capacities, Preprint, 2004.

[11] Peng, S. (2004) Filtration Consistent Nonlinear Expectations and Evaluations of Contingent Claims, *Acta Mathematicae Applicatae Sinica,* English Series **20**(2), 1–24.

[12] Peng, S. (2005) Nonlinear expectations and nonlinear Markov chains, *Chin. Ann. Math.* **26B**(2), 159–184.

[13] Peng, S. (2006) G–Expectation, G–Brownian Motion and Related Stochastic Calculus of Itô's type, preprint (pdf-file available in arXiv:math.PR/0601035v1 3Jan 2006), to appear in *Proceedings of the 2005 Abel Symposium.*

[14] Peng, S. (2006) Multi-Dimensional G-Brownian Motion and Related Stochastic Calculus under G-Expectation, in arXiv:math.PR/0601699v1 28Jan 2006.

[15] Peng, S. (2007) Law of large numbers and central limit theorem under nonlinear expectations, in arXiv:math.PR/0702358v1 13 Feb 2007.

[16] Peng, S. (2007) G–Brownian Motion and Dynamic Risk Measure under Volatility Uncertainty, in arXiv:0711.2834v1 [math.PR] 19 Nov 2007.

[17] Song, Y. (2007) A general central limit theorem under Peng's G-normal distribution, Preprint.

[18] L. Wang, (1992) On the regularity of fully nonlinear parabolic equations: II, Comm. Pure Appl. Math. 45, 141-178.~~A general central limit theorem under~~

[19] K. Yosida, Functional Analysis, Sixth Edition, Springer-Verlag.