An Iterative Approximation of the Sublinear Expectation of an Arbitrary Function of $G$-normal Distribution and the Solution to the Corresponding $G$-heat Equation

Yifan Li, Reg Kulperger

May 1, 2018

Abstract

It has been a well-known problem in the $G$-framework that it is hard to compute the sublinear expectation of the $G$-normal distribution $\hat{E}[\varphi(X)]$ when $\varphi$ is neither convex nor concave, if not involving any PDE techniques to solve the corresponding $G$-heat equation. Recently, we have established an efficient iterative method able to compute the sublinear expectation of arbitrary functions of the $G$-normal distribution, which directly applies the Nonlinear Central Limit Theorem in the $G$-framework to a sequence of variance-uncertain random variables following the Semi-$G$-normal Distribution, a newly defined concept with a nice Integral Representation, behaving like a ladder in both theory and intuition, helping us climb from the ground of classical normal distribution to approach the peak of $G$-normal distribution through the iteratively maximizing steps. The series of iteration functions actually produce the whole solution surface of the $G$-heat equation on a given time grid.
Contents

1 Introduction 3

2 Preliminaries 3

3 Main Theoretical Results 6
   3.1 The Semi-G-normal Distribution in one dimension 6
   3.2 The Iterative Approximation of the G-normal Distribution in one dimension 9
   3.3 Extension to the d-dimensional Situation 15

4 Implementation 16
   4.1 The 1-dimensional Situation 16
   4.2 The d-dimensional Situation 20

5 Assessment of the Iterative Method 22

6 Concluding Remarks and Further Development 22

References 23
1 Introduction

(This is a short version and will be expanded to a longer version for general audience.)

There is a long existing thinking gap between classical normal and $G$-normal distribution.

For instance, the result

$$\hat{E}[\varphi(N(0, [\sigma^2, \bar{\sigma}^2]))] \geq \sup_{v \in [\sigma, \bar{\sigma}]} E[\varphi(N(0, \sigma^2))]$$

indicates that the uncertainty set of $G$-normal distribution is much larger that a class of linear
normal distributions with $\sigma \in [\sigma, \bar{\sigma}]$. Especially, the strict inequality

$$\hat{E}[(N(0, [\sigma^2, \bar{\sigma}^2]))] > 0 = E[(N(0, \sigma^2))] > -\hat{E}[-(N(0, [\sigma^2, \bar{\sigma}^2]))]$$

tells us that the “skewness” of $G$-normal distribution becomes uncertain so its “symmetry” is un-
certain, which is quite counter-intuitive for a “normal” distribution. Is it possible to understand the
$G$-normal distribution from our familiar classical normal distribution? In other words, is it possible
to use the linear expectation of linear normal distribution (or the linear heat equation) to approach
the sublinear expectation of $G$-normal distribution (or the nonlinear $G$-heat equation)? This paper
gives positive answers to both of these questions by providing a ladder from the ground of $N(0, \sigma^2)$
to approach the peak of $N(0, [\sigma^2, \bar{\sigma}^2])$: the Semi-$G$-normal distribution $\hat{N}(0, [\sigma^2, \bar{\sigma}^2])$, which is a
classical normal distribution scaled by a sublinear “constant” (the maximal distribution).

2 Preliminaries

The $G$-expectation or sublinear expectation framework (also called the $G$-framework), motivated by
the problems with ambiguity or uncertainty, is a generalization of the linear probability framework.
Similar to the Choquet expectation, the sublinear expectation can be represented as the supreme
of a class of linear expectations. Intuitively, the linear expectation mainly considers the “average”,
while the sublinear expectation focuses on the “bound” to create a interval to cover the uncertainty
which is hard to be described by a certain distribution. Please turn to Peng (2007, 2008) for more
details.

Definition 2.1. A sublinear expectation space is defined as a triple $(\Omega, \mathcal{H}, \hat{E})$. $\Omega$ is a given set (also
known as a sample space). $\mathcal{H}$ is a linear space of real valued functions defined on $\Omega$ satisfying $c \in \mathcal{H}$
for each constant $c$ and $[X] \in \mathcal{H}$ if $X \in \mathcal{H}$, which can be regarded as the space of random variables.
$\hat{E}$ is a sublinear expectation which is a functional $\hat{E} : \mathcal{H} \rightarrow \mathbb{R}^d$ satisfying:

1. Monotonicity: $\hat{E}[X] \geq \hat{E}[Y]$ if $X \geq Y$;
2. Constant preserving: $\hat{E}[c] = c$ for $c \in \mathbb{R}^d$;
3. Sub-additivity: For each $X, Y \in \mathcal{H}$, $\hat{E}[X + Y] \leq \hat{E}[X] + \hat{E}[Y]$;
4. Positive homogeneity: $\hat{E}[\lambda X] = \lambda \hat{E}[X]$ for $\lambda \geq 0$.

If only monotonicity and constant preserving are satisfied, $\hat{E}$ is called a nonlinear expectation and
$(\Omega, \mathcal{H}, \hat{E})$ is called a nonlinear expectation space.

In the following context, we often use capital letters like $X := (X_1, X_2, \ldots, X_d), d \in \mathbb{N}_+$ to denote
the random variables (or vectors) in $\mathcal{H}$. Meanwhile, if $X \in \mathcal{H}$, we also have $\varphi(X) \in \mathcal{H}$ for every $\varphi$
in $C_{l.Lip}(\mathbb{R}^d)$ which is the linear space of functions satisfying the locally Lipchitz property:

$$|\varphi(x) - \varphi(y)| \leq C_\varphi |1 + |x|^k + |y|^k||x - y|$$

3
for \( x, y \in \mathbb{R}^d \), some \( k \in \mathbb{N} \) and \( C_\varphi > 0 \) depending on \( \varphi \). If not specified, we will always stay in the sublinear expectation space \((\Omega, \mathcal{H}, \hat{E})\) and the function space \( C_{l,Lip}(\mathbb{R}^d) \) (or \( C_{l,Lip} \) in short, which can be replaced by other spaces). Our computation in this space is usually different from the linear expectation \( E \) mainly because of the sub-additivity and positive homogeneity of \( \hat{E} \). Here are some useful tools to understand and deal with \( \hat{E} \).

**Proposition 2.2.** If \( X \) has the mean-uncertainty: \( \underline{\mu} < \bar{\mu} \) with \( \underline{\mu} := -\hat{E}[-X] \) and \( \bar{\mu} := \hat{E}[X] \), for \( \lambda \in \mathbb{R} \), we will have

\[
\hat{E}[\lambda X] = \begin{cases} 
\lambda \hat{E}[X] & \lambda \geq 0 \\
-\lambda \hat{E}[-X] & \lambda < 0
\end{cases}
\]

where \( \lambda^+ := \max\{\lambda, 0\} \) and \( \lambda^- := \max\{-\lambda, 0\} \).

**Proposition 2.3.** If \( X \) has the mean-certainty: \( \underline{\mu} = \bar{\mu} =: \mu \) namely \( -\hat{E}[-X] = \hat{E}[X] = \mu \), for \( \lambda \in \mathbb{R} \), we will have

\[
\hat{E}[\lambda X] = \lambda \hat{E}[X](= \lambda \mu)
\]

and furthermore,

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

**Proof.** The first one is directly from the proposition 2.2. The second one comes from

\[
\hat{E}[Y + \lambda X] = \hat{E}[Y] + \lambda \hat{E}[X] = \hat{E}[Y] + \lambda \hat{E}[X]
\]

and

\[
\hat{E}[Y + \lambda X] = \hat{E}[Y - \lambda(-X)] \geq \hat{E}[Y] - \hat{E}[\lambda(-X)] = \hat{E}[Y] + \lambda \hat{E}[X].
\]

\( \Box \)

**Proposition 2.4.** For \( p, q > 0 \), \( \frac{1}{p} + \frac{1}{q} = 1 \), we have

\[
\hat{E}[|XY|] \leq (\hat{E}[|X|^p])^{1/p}(\hat{E}[|Y|^q])^{1/q}.
\]

**Proposition 2.5.** For \( p \geq 1 \), we have

\[
(\hat{E}[|X + Y|^p])^{1/p} \leq (\hat{E}[|X|^p])^{1/p} + (\hat{E}[|Y|^p])^{1/p}.
\]

We will use \( \hat{E} \) to redefine distributions and independence.

**Definition 2.6** (Distributions). \( \mathbb{F}_X \) is called the *distribution* of \( X \), which is a functional: \( \mathbb{F}_X[\varphi] := \hat{E}[\varphi(X)] : \varphi \in C_{l,Lip}(\mathbb{R}^d) \to \mathbb{R} \). \( X \) and \( Y \) are *identically distributed*, denoted by \( X \overset{d}{=} Y \), iff for any \( \varphi \in C_{l,Lip} \), \( \hat{E}[\varphi(X)] = \hat{E}[\varphi(Y)] \), namely, \( \mathbb{F}_X[\varphi] = \mathbb{F}_Y[\varphi] \). A sequence \( \{X_n\}_{n=1}^\infty \) converges in distribution to \( X \), denoted as \( X_n \overset{d}{\to} X \), iff for any \( \varphi \in C_{l,Lip} \), \( \lim_{n \to \infty} \hat{E}[\varphi(X_n)] = \hat{E}[\varphi(X)] \).

**Definition 2.7** (Independence). \( Y \) is *independent from \( X \)*, denoted by \( X \overset{\rightarrow}{\longrightarrow} Y \), iff for any \( \varphi \in C_{l,Lip} \)

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

Intuitively, any realization of \( X \) will have no effects on the distribution of \( Y \).

**Definition 2.8** (i.i.d.). \( \{X_i\}_{i=1}^\infty \) is *i.i.d.* iff \( X_{i+1} \overset{d}{=} X_i \) and \( (X_1, X_2, \ldots, X_i) \overset{\rightarrow}{\longrightarrow} X_{i+1} \) for each \( i \in \mathbb{N}_+ \).
Let $\bar{X}$ be an independent copy of $X$, which means $\bar{X} \overset{d}{=} X$ and $X \rightarrow \bar{X}$.

**Definition 2.9** (Maximal Distribution). $X$ follows Maximal Distribution iff, for any independent copy $\bar{X}$,

$$X + \bar{X} \overset{d}{=} 2X.$$  

This is the sublinear version of a constant. A more specific definition is that $X$ follows the maximal distribution $M(\Gamma)$ if there exists a bounded, closed and convex set $\Gamma \subset \mathbb{R}^d$ such that for any $\varphi \in C_{l,Lip}(\mathbb{R}^d)$,

$$F_X[\varphi] = \hat{E}[\varphi(X)] = \max_{v \in \Gamma} \varphi(v).$$

For $d = 1$, we have $X \sim M[\mu, \overline{\mu}]$ with mean-uncertainty: $\mu := -\hat{E}[-X]$ and $\overline{\mu} := \hat{E}[X]$.

**Definition 2.10** (G-normal Distribution). $X$ follows G-normal Distribution iff, for any independent copy $\bar{X}$,

$$X + \bar{X} \overset{d}{=} \sqrt{2}X.$$  

For $d = 1$, we have $X \sim N(0, [\sigma^2, \overline{\sigma}^2])$ with variance-uncertainty: $\sigma^2 := -\hat{E}[-X^2]$ and $\overline{\sigma}^2 := \hat{E}[X^2]$.

Let $\mathbb{S}_d$ denote the set of all real-valued $d \times d$ symmetric matrices.

**Theorem 2.11** (G-normal distribution Characterized by the G-heat Equation). $X$ follows the $d$-dimensional G-normal distribution, iff $u(t, x) := \hat{E}[\varphi(x + \sqrt{1-t}X)]$ is the solution to the G-heat Equation defined on $[0, \infty] \times \mathbb{R}^d$:

$$u_t + G(D^2u) = 0, \quad u|_{t=1} = \varphi$$

where $G(A) := \frac{1}{2}\hat{E}[(AX, X)] : \mathbb{S}_d \rightarrow \mathbb{R}$, which is a sublinear function characterizing the distribution of $X$. For $d = 1$, we have $G(a) = \frac{1}{2}(\sigma^2a^+ - \overline{\sigma}^2a^-)$ and when $\sigma^2 > 0$, this is also called the Black-Scholes-Barenblatt equation with volatility uncertainty.

**Remark 2.11.1**. We can use the function $G(A) := \frac{1}{2}\hat{E}[(AX, X)]$ to characterize the definition of G-normal distribution. In fact, $G(A)$ can be further expressed as

$$G(A) = \frac{1}{2} \sup_{V \in \mathcal{V}} \text{tr}[AV]$$

where $\mathcal{V} = \{BB^T : B \in \mathbb{S}_d\}$ is a collection of non-negative definite symmetric matrices which can be treated as the uncertainty set of the covariance matrices.

**Definition 2.12** (G-normal distribution with Characterization). Let $X$ be any $d$-dimensional G-normal distributed random vector. To be specific, we say $X \sim N(0, \mathcal{V})$ with the set of covariance matrices $\mathcal{V}$ if its distribution is characterized by the G-heat equation with

$$G(A) := \frac{1}{2}\hat{E}[(AX, X)] = \frac{1}{2} \sup_{V \in \mathcal{V}} \text{tr}[AV].$$

In other words, $\mathcal{V}$ is the set corresponding to the $G$ function characterizing the distribution of $X$. In order to show the covariance-uncertainty of $N(0, \mathcal{V})$, we can expand the details of $\mathcal{V}$ as

$$\mathcal{V} := \{V = (\rho_{ij}\sigma_i\sigma_j)_{d \times d} : \sigma_i^2 \in [\sigma^2, \overline{\sigma}^2], \rho_{ij} = \rho_{ji} = \begin{cases} 1 & i = j, i, j = 1, 2, \ldots, d, \\ \in [\underline{\rho}_i, \overline{\rho}_i] & i \neq j, \end{cases} \}.\]$$

When $d = 1$, we say $X \sim N(0, [\sigma^2, \overline{\sigma}^2])$ with the variance interval $[\sigma^2, \overline{\sigma}^2]$ if its distribution is characterized by the G-heat equation with

$$G(a) := \frac{1}{2}\hat{E}[aX^2] = \frac{1}{2}(\sigma^2a^+ - \overline{\sigma}^2a^-).$$
Theorem 2.13 (The Nonlinear Central Limit Theorem by Peng). Consider a sequence of i.i.d. 
$\{X_i\}_{i=1}^{\infty}$ with mean-certainty $\hat{E}[X_1] = -\hat{E}[-X_1] = 0$. Let $X$ be a $G$-normal distributed random variable characterized by the function $G(A) := \frac{1}{2}\hat{E}[(AX_1, X_1)]$. Then for any $\varphi \in C_{l,\text{Lip}},$

$$
\lim_{n \to \infty} \hat{E}[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i)] = \hat{E}[\varphi(X)].
$$

For $d = 1$, let $\sigma^2 := -\hat{E}[-X_1^2]$ and $\sigma^2 := \hat{E}[X_1^2]$. Then we have $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ converges in distribution to $X \sim N(0, (\sigma^2, \sigma^2))$.

We will call Theorem 2.13 the nonlinear CLT in short. We also have the convergence rate of nonlinear CLT from Song (2017).

Theorem 2.14 (The Convergence Rate of Nonlinear CLT). Under the setting of Theorem 2.13 when $d = 1$, for bounded and Lipschitz continuous $\varphi$ (i.e. for any $x, y \in \mathbb{R}$, $|\varphi(x)| \leq M$ and $|\varphi(x) - \varphi(y)| \leq C_\varphi |x - y|), there exist $\alpha \in (0, 1)$ depending on $\sigma$ and $\bar{\sigma}$, and $C_{\alpha,G} > 0$ depending on $\alpha$, $\sigma$ and $\bar{\sigma}$ such that

$$
\sup_{C_\varphi \leq 1} |\hat{E}[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i)] - \hat{E}[\varphi(X)]| \leq C_{\alpha,G} \frac{\hat{E}[|X_1|^{2+\alpha}]}{n^{\frac{\alpha}{2}}} = O\left(\frac{1}{n^{\frac{\alpha}{2}}}\right).
$$

3 Main Theoretical Results

The iterative algorithm is based on a new concept, the Semi-$G$-normal distribution $\hat{N}(0, [\sigma^2, \sigma^2])$ with a smaller uncertainty set than the $G$-normal distribution. This new distribution has a nice integral representation and a general connection with the $G$-normal distribution, that is the CLT (central limit theorem) in the $G$-framework. For reader’s convenience, we will first give the results in one dimension and they can be easily extended to multi-dimensional cases.

3.1 The Semi-$G$-normal Distribution in one dimension

Definition 3.1 (The Semi-$G$-normal Distribution). $W$ follows the Semi-$G$-normal distribution, denoted as, $W \sim \hat{N}(0, [\sigma^2, \sigma^2])$, if there exist $Y \sim N(0, [1, 1])$ and $Z \sim M[\sigma, \bar{\sigma}], \bar{\sigma} \geq \sigma \geq 0$ with independent relation $Z \sim Y$, such that

$$
W = Z \cdot Y
$$

where “$\cdot$” is the number multiplication (which can be omitted) and the direction of independence here cannot be reversed.

Remark 3.1.1. $Y \sim N(0, [1, 1])$ can be regarded as the classical standard normal distribution $N(0, 1)$ since the corresponding $G$-heat equation will be reduced to classical heat equation when $\sigma$ and $\bar{\sigma}$ coincide.

Remark 3.1.2 (the mean and variance of $W$). It is not hard to show that it has a certain zero mean:

$$
\hat{E}[W] = \hat{E}[ZY] = \hat{E}[\hat{E}[zY]_{z=Z}] = \hat{E}[\hat{E}[zY]_{z=Z}] = 0
$$

and $-\hat{E}[-W] = 0$. For the variance, we have

$$
\hat{E}[W^2] = \hat{E}[Z^2Y^2] = \hat{E}[\hat{E}[z^2Y^2]_{z=Z}] = \hat{E}[\hat{E}[z^2Y^2]_{z=Z}] = \hat{E}[z^2 \cdot 1]_{z=Z} = \max_{z \in [\sigma, \bar{\sigma}]} z^2 = \sigma^2
$$
and similarly,

$$-\hat{\mathbb{E}}[W^2] = -\hat{\mathbb{E}}[Z^2] = -\hat{\mathbb{E}}[E[-z^2 Y^2]_{z=Z}] = -\max_{z \in [\sigma, \bar{\sigma}]} (-z^2) = \min_{z \in [\sigma, \bar{\sigma}]} z^2 = \sigma^2.$$

**Theorem 3.2** (The Integral Representation of the Semi-$G$-normal Distribution). Let $W \sim \hat{N}(0, [\sigma^2, \bar{\sigma}^2])$ then for any $\varphi \in C_l.lip(\mathbb{R})$, we have

$$\mathbb{E}[\varphi(W)] = \max_{z \in [\sigma, \bar{\sigma}]} E[\varphi(N(0, z^2))] = \max_{z \in [\sigma, \bar{\sigma}]} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \varphi(zy) dy.$$

**Proof.** This is quite straightforward because:

$$\mathbb{E}[\varphi(W)] = \mathbb{E}[\varphi(Y)] = \mathbb{E}[\mathbb{E}[\varphi(Y_Z)_{Z=Z}]] = \mathbb{E}[\mathbb{G}(Z)]$$

where $G(z) := \hat{\mathbb{E}}[\varphi(Y_Z)] = E[\varphi(N(0, z^2))]$ can be proved to be in $C_l.lip$ based on $\varphi \in C_l.lip$. Specifically, we have

$$|G(x) - G(y)| = |E[\varphi(x \cdot Z) - \varphi(y \cdot Z)]| 
\leq E[C_\varphi(1 + |xZ|^k + |yZ|^k)] Z \cdot |x - y| 
= C_\varphi : E[|Z| + |Z^{k+1}|x^k + |Z^{k+1}|y^k]|x - y| 
\leq C_G(1 + |x|^k + |y|^k)|x - y|$$

where $C_G = C_\varphi \max\{E[|Z|], E[|Z^{k+1}|]\}$. Therefore,

$$\mathbb{E}[\varphi(W)] = \mathbb{E}[\mathbb{G}(Z)] = \max_{z \in [\sigma, \bar{\sigma}]} G(z) 
= \max_{z \in [\sigma, \bar{\sigma}]} E[\varphi(N(0, z^2))] 
= \max_{z \in [\sigma, \bar{\sigma}]} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \varphi(zy) dy. \quad \blacksquare$$

**Remark 3.2.1** (Why is it called a “semi” one?). The comparison theorem of parabolic PDEs ($G$-heat vs Heat Equations) tells us that

$$\hat{\mathbb{E}}[\varphi(N(0, [\sigma^2, \bar{\sigma}^2]))] \geq \max_{v \in [\sigma, \bar{\sigma}]} E[\varphi(N(0, v^2))] = \hat{\mathbb{E}}[\varphi(\hat{N}(0, [\sigma^2, \bar{\sigma}^2]))]$$

whose inequality is mostly strict (like $\varphi(x) = x^3$) and becomes equal when $\varphi$ is convex or concave. From the representation theorem of $\hat{\mathbb{E}}$: $\hat{\mathbb{E}}[\varphi(X)] = \sup_{P \in \mathbb{Q}} E_P[\varphi(X)]$, we have the intuition that $Q_{\text{Semi-G-normal}} \subset Q_{\text{G-normal}}$ where $Q_{\text{Semi-G-normal}}$ consists of measures corresponding to classical normal distributions with $\sigma \in [\sigma, \bar{\sigma}]$.

**Remark 3.2.2.** Intuitively, the $G$-normal distribution is more “uncertain” than the semi-$G$-normal distribution. Explicitly speaking, we already have the following representation for the $G$-normal distributed $X \sim N(0, [\sigma^2, \bar{\sigma}^2])$ (see Denis et al. (2011)):

$$\hat{\mathbb{E}}[\varphi(X)] = \sup_{\theta \in \Theta^\Omega} E_P[\varphi(\int_0^1 \theta_s dB_s)]$$

where $\Theta = [\sigma, \bar{\sigma}]$, $\Theta^\Omega := \{ \theta : \theta_t \in \Theta, \text{ for } t \in [0, 1] \}$, the set of all processes valuing in $[\sigma, \bar{\sigma}]$ in the time range $[0, 1]$ and $B$ is the classical Brownian motion in $(\Omega, \mathcal{F}, P)$. Meanwhile, from the integral representation (Theorem 3.2), we can show that for $W \sim \hat{N}(0, [\sigma^2, \bar{\sigma}^2])$,

$$\hat{\mathbb{E}}[\varphi(W)] = \sup_{\theta \in \Theta^\Omega} E_P[\varphi(\int_0^1 \hat{\theta} dB_s)]$$
where \( \bar{\theta} = \int_{0}^{1} \theta_s ds (\in [\underline{\tau}, \bar{\tau}]) \), the average of the process \( \theta \) over the time interval [0, 1]. To summarize, the \( G \)-normal distribution has the uncertainty set consisting of all processes valuing in \( \Theta \), while the semi-\( G \)-normal distribution only has the set made up of all constant processes valuing in \( \Theta \).

**Corollary 3.2.1** (the connection with the \( G \)-normal distribution). When \( \varphi \) is convex or concave and \( \varphi \in C^2(\mathbb{R}) \), for \( X \sim N(0, [\underline{\sigma}^2, \bar{\sigma}^2]) \) and \( W \sim N(0, [\underline{\varphi}^2, \bar{\varphi}^2]) \), we have

\[
\hat{E}[\varphi(X)] = \hat{E}[\varphi(W)].
\]

**Proof.** The well-known integral representation of \( G \)-normal distribution under convexity (or concavity) directly comes from the solution of the classical heat equation because \( u(t, x) \) will be convex (or concave, respectively) to \( x \) then \( u_{xx} \geq 0 \) (or \( \leq 0 \), respectively), giving us

\[
\hat{E}[\varphi(X)] = \begin{cases} 
E[\varphi(N(0, \bar{\sigma}^2))] & \text{\( \varphi \) is convex} \\
E[\varphi(N(0, \underline{\sigma}^2))] & \text{\( \varphi \) is concave}.
\end{cases}
\]

For the semi-\( G \)-normal distribution, by using its representation with \( G(z) := E[\varphi(zY)](z \in [\underline{\varphi}, \bar{\varphi}] \) and \( Y \sim N(0, 1) \), we only need to prove that

\[
\hat{E}[\varphi(W)] = \max_{z \in [\underline{\varphi}, \bar{\varphi}]} G(z) = \begin{cases} 
G(\bar{\sigma}) & \text{\( \varphi \) is convex} \\
G(\underline{\varphi}) & \text{\( \varphi \) is concave}.
\end{cases}
\]

First of all, \( \varphi \) has the Taylor expansion

\[
\varphi(x) = \varphi(0) + \varphi^{(1)}(0)x + \varphi^{(2)}(\xi_x)\frac{x^2}{2}
\]

where \( \xi_x \in (0, x) \).

1. When \( \varphi \) is convex, \( \varphi^{(2)}(\xi_x) \geq 0 \). The Taylor expansion tells us that:

\[
G(z) = E[\varphi(zY)]
= E[\varphi(0) + \varphi^{(1)}(0)zY + \varphi^{(2)}(\xi_{zY})\frac{z^2}{2}Y^2]
= \varphi(0) + \frac{1}{2}E[\varphi^{(2)}(\xi_{zY})(zY)^2]
\]

where \( \xi_{zY} \in (0, zY) \) is a random variable depending on \( Y \). Let \( M := zY \sim N(0, z^2) \), then

\[
K(z) := E[\varphi^{(2)}(\xi_{zY})(zY)^2] = E[\varphi^{(2)}(\xi_M)M^2] = \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{m^2}{2z^2}\right)\varphi^{(2)}(\xi_m)m^2 dm.
\]

In order to consider the monotone property of \( K(z) \), work on its derivative:

\[
K'(z) = \frac{d}{dz} \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{m^2}{2z^2}\right)\varphi^{(2)}(\xi_m)m^2 dm
= \int \frac{1}{\sqrt{2\pi}} \left[ \frac{d}{dz} \exp\left(-\frac{m^2}{2z^2}\right) \right] \varphi^{(2)}(\xi_m)m^2 dm
= \int \frac{1}{\sqrt{2\pi}} \left[ \frac{m^2}{z^3} \exp\left(-\frac{m^2}{2z^2}\right) \right] \varphi^{(2)}(\xi_m)m^2 dm \geq 0.
\]

\[
\geq 0 \text{ for } z \in [\underline{\varphi}, \bar{\varphi}]
\]
This tells us \( K(z) \) is increasing with respect to \( z \in [\sigma, \overline{\sigma}] \), then \( K(z) \) reaches its maximum at \( z = \sigma \). Hence,
\[
\hat{E}[\varphi(W)] = \max_{z \in [\sigma, \overline{\sigma}]} G(z) = \max_{z \in [\sigma, \overline{\sigma}]} (\varphi(0) + \frac{K(z)}{2}) = G(\sigma).
\]

2. When \( \varphi \) is concave, then \(-\varphi\) is convex. Replace the \( \varphi \) above with \(-\varphi\) and repeat the same procedure, we have
\[
-G(z) = -E[\varphi(zY)] \\
= E[(-\varphi)(zY)] \\
= (-\varphi)(0) + \frac{z^2}{2} E[(-\varphi)^{(2)}(\xi_{zY})Y^2] \\
\]
and
\[
K'(z) = \int \frac{1}{\sqrt{2\pi}} \left[ \frac{m^2}{z^3} \exp \left( -\frac{m^2}{2z^2} \right) \right] (\varphi)^{(2)}(\xi_m) m^2 dm \geq 0 \\
\]
Hence, \(-G(z)\) is increasing with respect to \( z \), that is, \( G(z) \) is decreasing according to \( z \). Therefore,
\[
\hat{E}[\varphi(X)] = \max_{z \in [\sigma, \overline{\sigma}]} G(z) = G(\sigma). \quad \Box
\]

The initial motivation of the semi-\( G \)-normal distribution is that we want to create a tool or ladder to help us better understand and handle the \( G \)-normal distribution, especially based on what we already know about the classical normal distribution, which turns out to be feasible from the nice properties of semi-\( G \)-normal distribution and thanks to the constructed theory in \( G \)-framework (like the nonlinear CLT by Peng). The following result is one of the exciting results from the semi-\( G \)-normal distribution to better understand and compute the expectation of the \( G \)-normal distribution.

### 3.2 The Iterative Approximation of the \( G \)-normal Distribution in one dimension

**Lemma 3.3** (General connection between the Semi-\( G \)-normal and the \( G \)-normal distribution). Consider a sequence of i.i.d. \( \{W_i\}_{i=1} \sim \mathcal{N}(0, [\sigma^2, \overline{\sigma}^2]) \) and \( X \sim \mathcal{N}(0, [\sigma^2, \overline{\sigma}^2]) \), then we have
\[
\lim_{n \to \infty} \hat{E}[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i)] = \hat{E}[\varphi(X)], \forall \varphi \in C_{l.Lip}(\mathbb{R}).
\]
In other words, \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i \) converges in distribution to the \( G \)-normal distributed \( X \).

**Lemma 3.3** is a direct result of the CLT (Theorem 2.13) in the \( G \)-framework.

**Theorem 3.4** (The Iterative Approximation of the \( G \)-normal Distribution). Consider a \( G \)-normal distributed random variable \( X \sim \mathcal{N}(0, [\sigma^2, \overline{\sigma}^2]) \). For any \( \varphi \in C_{l.Lip}(\mathbb{R}) \) and integer \( n \geq 1 \), consider the series of iteration functions \( \{\varphi_{i,n}\}_{i=1}^{n} \) with initial function \( \varphi_{0,n}(x) := \varphi(x) \) and iterative relation:
\[
\varphi_{i+1,n}(x) := \max_{v \in [\sigma, \overline{\sigma}]} E[\varphi_{i,n}(N(x, v^2/n))], i = 0, 1, \ldots, n-1.
\]
The final iteration function for a given \( n \) is \( \varphi_{n,n} \). As \( n \to \infty \), we have \( \varphi_{n,n}(0) \to \hat{E}[\varphi(X)] \).
Proof. Set the initial function $\varphi_{0,n}(x) := \varphi(x)$ and the iteration

$$\varphi_{i,n}(x) := \max_{v \in [\sigma, \bar{\sigma}]} E[\varphi_{i-1,n}(N(x, \frac{v^2}{n}))].$$

In order to use the integral representation of the Semi-G-normal distribution (Theorem 3.2) in the next stage, we want each iteration function to be in the function space $C_{l,Lip}$. For convenience, we omit the subscript $n$ for a while and let $\varphi_i := \varphi_{i,n}$. We also know that the optimal $v$ will be some value in $[\sigma, \bar{\sigma}]$ depending on $x$, i.e.

$$\varphi_i(x) = E[\varphi_{i-1}(N(x, \frac{\sigma^2}{n})), v_x \in [\sigma, \bar{\sigma}]].$$

By induction, we only need to show that given $\varphi_{i-1} \in C_{l,Lip}$, we also have $\varphi_i \in C_{l,Lip}$, for $i = 1, 2, \ldots, n$. Suppose for $\varphi_{i-1}$ we have a constant $C_{i-1}$ and an positive integer $k$ such that

$$|\varphi_{i-1}(x) - \varphi_{i-1}(y)| \leq C_{i-1}(1 + |x|^k + |y|^k)|x - y|.$$

Let $Z \sim N(0, 1)$ and consider

$$|\varphi_i(x) - \varphi_i(y)| = |E[\varphi_{i-1}(N(x, \frac{\sigma^2}{n}))) - E[\varphi_{i-1}(N(y, \frac{\sigma^2}{n}))]|$$

$$= |E[\varphi_{i-1}(N(x, \frac{\sigma^2}{n}) + \frac{v_x}{\sqrt{n}}Z) - E[\varphi_{i-1}(N(y, \frac{\sigma^2}{n}) + \frac{v_y}{\sqrt{n}}Z)]|$$

$$\leq E|\varphi_{i-1}(N(x, \frac{\sigma^2}{n}) + \frac{v_x}{\sqrt{n}}Z) - \varphi_{i-1}(N(y, \frac{\sigma^2}{n}) + \frac{v_y}{\sqrt{n}}Z)|$$

$$\leq E[C_{i-1}(1 + |x + \frac{v_x}{\sqrt{n}}Z|^k + |y + \frac{v_y}{\sqrt{n}}Z|^k)|(x + Z) - (y + Z)|]$$

$$= C_{i-1}|x - y|(1 + E|x + \frac{v_x}{\sqrt{n}}Z|^k + E|y + \frac{v_y}{\sqrt{n}}Z|^k).$$

For given $\omega \in \Omega$, let $z := Z(\omega)$, we have

$$|x + \frac{v_x}{\sqrt{n}}z|^k = |(x + \frac{v_x}{\sqrt{n}}z)^k| \leq \sum_{j=1}^{k} \binom{k}{j} |x|^j |\frac{v_x}{\sqrt{n}}z|^{k-j}$$

$$\leq \sum_{j=1}^{k} \binom{k}{j} \max\{|x|^k, |\frac{v_x}{\sqrt{n}}z|^k\} \leq 2^k(|x|^k + |\frac{v_x}{\sqrt{n}}|^k|z|^k) \leq 2^k(|x|^k + |\frac{\sigma}{\sqrt{n}}|^k|z|^k)$$

where the $2^k$ can be improved to max$\{2^{k-1}, 1\}$ but they are both constants so $2^k$ is good enough for us. Then $|x + \frac{v_x}{\sqrt{n}}Z(\omega)|^k \leq 2^k(|x|^k + |\frac{\sigma}{\sqrt{n}}|^k|Z(\omega)|^k)$; taking expectation on both sides, we have

$$E|x + \frac{v_x}{\sqrt{n}}Z|^k \leq 2^k(|x|^k + |\frac{\sigma}{\sqrt{n}}|^k|E|Z|^k).$$

Therefore,

$$|\varphi_i(x) - \varphi_i(y)| \leq C_{i-1}|x - y|(1 + E|x + \frac{v_x}{\sqrt{n}}Z|^k + E|y + \frac{v_y}{\sqrt{n}}Z|^k)$$

$$\leq C_{i-1}|x - y|(1 + 2^{k+1}|\frac{\sigma}{\sqrt{n}}|^k |E|Z|^k + 2^k|x|^k + 2^k|y|^k)$$

$$\leq C_i(1 + |x|^k + |y|^k)|x - y|.$$
where \( C_i := C_{i-1} \cdot \max\{1 + 2^{k+1} |\frac{\eta}{n}|^k E|Z|^k, 2^k\} \).

Considering a sequence of nonlinear i.i.d. \( \{W_i\}_{i=1}^{n} \sim \tilde{N}(0, [\sigma^2, \sigma^2]) \) and \( W_{i,n} := \frac{1}{\sqrt{n}} W_i \), since \( \varphi_i \in C_{l.Lip} \), we can apply the integral representation of the Semi-G-normal distribution at each step. Then we have

\[
\hat{E}[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i)] = \hat{E}[\varphi_{0,n}(\sum_{i=1}^{n} W_{i,n})]
\]

\[
= \hat{E}[\hat{E}[\varphi_{0,n}(\sum_{i=1}^{n-1} w_{i,n} + W_{n,n})|w_{i,n}=W_{i,n}, i=1,\ldots,n-1]]
\]

\[
= \hat{E}[\left[ \max_{v_n \in [0,\sigma]} E[\varphi_{0,n}(\sum_{i=1}^{n-1} w_{i,n} + N(0, v_n^2/n))]|w_{i,n}=W_{i,n}, i=1,\ldots,n-1]\right]]
\]

\[
= \hat{E}[\left[ \max_{v_n \in [0,\sigma]} E[\varphi_{0,n}(N(\sum_{i=1}^{n-1} w_{i,n}, v_n^2/n))]|w_{i,n}=W_{i,n}, i=1,\ldots,n-1]\right]]
\]

\[
= \hat{E}[\varphi_{1,n}(\sum_{i=1}^{n-1} W_{i,n})]
\]

\[
= \hat{E}[\left[ \max_{v_{n-1} \in [0,\sigma]} E[\varphi_{1,n}(N(\sum_{i=1}^{n-2} w_{i,n}, v_{n-1}^2/n))]|w_{i,n}=W_{i,n}, i=1,\ldots,n-2\right]]
\]

\[
= \ldots
\]

\[
= \hat{E}[\left[ \max_{v_2 \in [0,\sigma]} E[\varphi_{n-2,n}(N(w_{1,n}, v_2^2/n))]|w_{1,n}=W_{1,n}\right]]
\]

\[
= \hat{E}[\varphi_{n-1,n}(W_{1,n})]
\]

\[
= \max_{v_1 \in [0,\sigma]} E[\varphi_{n-1,n}(N(0, v_1^2/n))] = \varphi_{n,n}(0)
\]

According to Lemma 3.3 we have

\[
\hat{E}[\varphi(X)] = \lim_{n \rightarrow \infty} \hat{E}[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i)] = \lim_{n \rightarrow \infty} \varphi_{n,n}(0). \quad \square
\]

Remark 3.4.1. From the proof, we note that the iteration function can also be expressed as the sublinear expectation of the semi-G-normal distribution (letting \( W_0 := 0 \)):

\[
\varphi_{i,n}(x) = \hat{E}[\varphi(x + \sum_{j=0}^{i} \frac{W_{n-j}}{\sqrt{n}})] = \hat{E}[\varphi(x + \sum_{j=0}^{i} \frac{W_{j}}{\sqrt{n}})]
\]

for \( i = 0, 1, \ldots, n \).

Furthermore, the following result shows that \( \{\varphi_{i,n}\}_{i=0}^{n} \) produces the whole solution surface of the corresponding G-heat equation on a given time grid. The proof involves the convergence rate of the nonlinear CLT (Theorem 2.14) from Song (2017), which considers the functions \( \varphi \) being bound and Lipschitz. If this is not the only function space, we can adapt the above result to the space \( C_{l.Lip} \) (locally Lipschitz) to get the Corollary 3.4.1.
Corollary 3.4.1. Consider the G-heat equation defined on \([0, \infty) \times \mathbb{R}\):

\[
    u_t + G(u_{xx}) = 0, \quad u|_{t=1} = \varphi
\]

where \(G(a) := \frac{1}{2}E[aX^2] - \frac{1}{2}(\sigma^2 a^+ - \sigma^2 a^-)\) and \(\varphi \in C_{t,Lip}(\mathbb{R})\). Then for the iterations \(\{\varphi_{i,n}\}_{i=0}^{n}\) in Theorem 3.6, for each \(p \in (0, 1]\), we have

\[
|u(1 - p, x) - \varphi_{[np], n}(x)| = |(\frac{1}{n}) \sum_{i=0}^{[np]} W_i| \leq C_{\varphi}(1 + |x|^k)O\left(\frac{1}{(np)^{\alpha/2}}\right),
\]

where \(\lfloor np \rceil\) is the floor (or integer) part of \(np\). When \(p = 0\), we have \(u(1, x) = \varphi(x) = \varphi_{0,n}(x)\).

Remark 3.4.2. If \(\varphi\) is required to be bounded and Lipschitz, we may remove the local factor \((1 + |x|^k)\) from the error part to get a uniform error approximation regardless of \(x\).

Proof. For each \(p \in (0, 1]\), consider the error when approximating \(u(1 - p, x)\), which can be approximated by

\[
    u(1 - \frac{[np]}{n}, x) = \mathbb{E}[\varphi(x + \sum_{j=0}^{[np]} \frac{X_j}{\sqrt{n}})] \approx \mathbb{E}[\varphi(x + \sum_{i=0}^{[np]} \frac{W_i}{\sqrt{n}})] = \varphi_{[np], n}(x).
\]

Specifically speaking, we intend to work on the error

\[
    |u(1 - p, x) - \varphi_{[np], n}(x)| \leq |u(1 - p, x) - u(1 - \frac{[np]}{n}, x)| + |u(1 - \frac{[np]}{n}, x) - \varphi_{[np], n}(x)|.
\]

Before diving into these two parts, we can prepare the converging property of \(p_n := \frac{[np]}{n}\). Actually, the inequality

\[
p_n = \frac{[np]}{n} \leq \frac{np}{n} \leq \frac{[np] + 1}{n} = p_n + \frac{1}{n},
\]

tells us that

\[
    |p_n - p| < \frac{1}{n}.
\]

The \((1)\) part involves the continuity of \(u\) on the time dimension specified by doing the Taylor expansion:

\[
(1) \leq |u_t(1 - p, x)|||p_n - p| + O(|p_n - p|^2). \quad O\left(\frac{1}{n^2}\right)
\]

We are looking for bound of \(|u_t|\). Fortunately, according to the properties of the solution to the G-heat equation (see Pao (2012) and Wang (1992)), there exist constant \(C > 0\) and \(\beta > 0\), such that

\[
|u_t(t, x) - u_t(s, x)| \leq C|t - s|^{\frac{\beta}{2}}.
\]

By letting \(s = 0\) and \(c_x := |u_t(0, x)|\), we have

\[
|u_t(t, x)| \leq |u_t(t, x) - u_t(0, x)| + |u_t(0, x)| \leq C|t|^{\frac{\beta}{2}} + c_x.
\]
Therefore, for a fixed \( x \), we have
\[
(1) \leq (C|1 - p| + c_x)|p_n - p| + O(\frac{1}{n^2}) = c_xO\left(\frac{1}{n}\right).
\]
The (2) part can be rewritten as follows:
\[
(2) = |\hat{E}[\varphi(x + \frac{1}{\sqrt{n}}\sum_{j=0}^{np} X_j)] - \hat{E}[\varphi(x + \frac{1}{\sqrt{n}}\sum_{j=0}^{np} W_j)]|
\]
\[
= |\hat{E}[\varphi(x + \frac{1}{\sqrt{np}}\sum_{j=0}^{np} X_j)] - \hat{E}[\varphi(x + \frac{1}{\sqrt{np}}\sum_{j=0}^{np} W_j)]|
\]
By letting \( X_{np} := \frac{1}{\sqrt{np}}\sum_{j=0}^{np} X_j \) and \( W_{np} := \frac{1}{\sqrt{np}}\sum_{j=0}^{np} W_j \), we have
\[
(2) = |\hat{E}[\varphi(x + \sqrt{p_n}X_{np})] - \hat{E}[\varphi(x + \sqrt{p_n}W_{np})]|
\]
\[
\leq |\hat{E}[\varphi(x + \sqrt{p_n}X_{np})] - \hat{E}[\varphi(x + \sqrt{p}W_{np})]|
\]
\[
+ |\hat{E}[\varphi(x + \sqrt{p_n}W_{np})] - \hat{E}[\varphi(x + \sqrt{p}W_{np})]|
\]
\[
= (2)_1 + (2)_2 + (2)_3
\]
where \((2)_1 + (2)_2\) involves the continuity of \( \varphi \) and the shrinking speed of \(|p_n - p|\) and \((2)_3\) is exactly fitted into the convergence rate of nonlinear CLT. In \((2)_1\) or \((2)_2\), we do not need to worry about the random variables since \( X_{np} \overset{d}{=} X \overset{d}{=} N(0, [\sigma^2, \sigma^2]) \) and \( W_{np} \overset{d}{=} X \) as \( n \to \infty \) from nonlinear CLT. Hence, let us work on a general expression by replacing \( X_{np} \) and \( W_{np} \) by \( Z_n \) satisfying \( Z_n \overset{d}{=} X \). We know \( \varphi \in C_{l, Lip} \) satisfying, for \( x, y \in \mathbb{R} \),
\[
|\varphi(x) - \varphi(y)| \leq C_{\varphi}(1 + |x|^k + |y|^k)|x - y|
\]
with \( C_{\varphi} > 0 \) and \( k \in \mathbb{N} \). When \( k = 0 \), \( \varphi \) is uniformly Lipschitz, then
\[
(2)_1 \text{or}(2)_2 = |\hat{E}[\varphi(x + \sqrt{p_n}Z_n)] - \hat{E}[\varphi(x + \sqrt{p}Z_n)]|
\]
\[
\leq |\hat{E}[[\varphi(x + \sqrt{p_n}Z_n) - \varphi(x + \sqrt{p}Z_n)]|
\]
\[
\leq \hat{E}[C_{\varphi}\sqrt{p_n - \sqrt{p}}||Z_n||] = C_{\varphi}\sqrt{p_n - \sqrt{p}}\hat{E}[||Z_n||]
\]
For the expectation part in the last line, since \( Z_n \overset{d}{=} X \), we can make \( n \) large enough so that
\[
\hat{E}[||Z_n||] \leq \hat{E}[||X||] + 1 =: K_0.
\]
For the \( p_n \) part, the Taylor expansion of \( \sqrt{x} \) at \( x = p \) tells us that
\[
|\sqrt{p_n} - \sqrt{p}| \leq \frac{1}{2\sqrt{p}}|p_n - p| + O(|p_n - p|^2) = O\left(\frac{1}{n\sqrt{p}}\right).
\]
Hence, when \( k = 0 \),
\[
(2)_1 \text{or}(2)_2 \leq C_1O\left(\frac{1}{n\sqrt{p}}\right)
\]
with $C_1 = C_\varphi K_0$. When $k \geq 1$, we have

$$(2)_1 \text{or } (2)_2 \leq \hat{E}[|\varphi(x + \sqrt{p_n}Z_n) - \varphi(x + \sqrt{p}Z_n)|]\leq \hat{E}[C_\varphi(1 + |x + \sqrt{p_n}Z_n|^k + |x + \sqrt{p}Z_n|^k)\sqrt{p_n} - \sqrt{p}|Z_n|]
\leq C_\varphi\hat{E}\left[(1 + 2^{k+1}|x|^k + 2^k|Z_n|^k(\sqrt{p_n}^k + |\sqrt{p}|^k))|Z_n|\right]\sqrt{p_n} - \sqrt{p}.$$

For the expectation part in the last line, by letting $n$ be large enough so that

$$(\hat{E}[|Z_n|^2])^{\frac{1}{2}} \leq (\hat{E}[|X|^2])^{\frac{1}{2}} + 1 = (E[|N(0, \sigma^2)|^2])^{\frac{1}{2}} + 1 = \sigma + 1 =: K_1,$$

and

$$(\hat{E}[|Z_n|^{2k}])^{\frac{1}{2}} \leq (\hat{E}[|X|^{2k}])^{\frac{1}{2}} + 1 = (E[|N(0, \sigma^2)|^{2k})^{\frac{1}{2}} + 1 = \sigma^k((2k-1)!!)^{\frac{1}{2}} + 1 =: K_2,$$

we have

$$\hat{E}\left[(1 + 2^{k+1}|x|^k + 2^{k+1}|Z_n|^k)|Z_n|\right] \leq (\hat{E}[(1 + 2^{k+1}|x|^k + 2^{k+1}|Z_n|^k)]^{\frac{1}{2}} (\hat{E}[|Z_n|^2])^{\frac{1}{2}} \leq \left(1 + 2^{k+1}|x|^k + 2^{k+1}(\hat{E}[|Z_n|^{2k}])^{\frac{1}{2}}\right)(\hat{E}[|Z_n|^2])^{\frac{1}{2}} \leq C_2(1 + |x|^k)$$

where $C_2 = K_1 \cdot \max\{1 + 2^{k+1}K_2, 2^{2k+1}\}$. Again, for the $p_n$ part, we have $|p_n - p| = O\left(\frac{1}{n\sqrt{p}}\right)$. Therefore, when $k \geq 1$, for a fixed $x$,

$$(2)_1 \text{or } (2)_2 \leq C_3(1 + |x|^k)O\left(\frac{1}{n\sqrt{p}}\right)$$

where $C_3 = C_\varphi C_2$. In a word, for $k \in \mathbb{N}$,

$$(2)_1 \text{or } (2)_2 \leq C_4(1 + |x|^k)O\left(\frac{1}{n\sqrt{p}}\right)$$

where $C_4 = \max\{C_1, C_3\}$.

In (2)_3, we can directly apply the convergence rate of nonlinear CLT (Theorem 2.14):

$$(2)_3 = |\hat{E}[\varphi(x + \sqrt{p}\frac{1}{\sqrt{[np]}}\sum_{j=0}^{[np]} X_j)] - \hat{E}[\varphi(x + \sqrt{p}\frac{1}{\sqrt{[np]}}\sum_{j=0}^{[np]} W_j)]|
\leq |\hat{E}[\varphi(\frac{1}{\sqrt{[np]}}\sum_{j=0}^{[np]} X_j)] - \hat{E}[\varphi(\frac{1}{\sqrt{[np]}}\sum_{j=0}^{[np]} W_j)]| = C_\varphi O(\frac{1}{(np)^{\alpha/2}})$$

where $\varphi(a) := \varphi(x + \sqrt{pa})$, satisfying

$$|\varphi(a) - \varphi(b)| \leq C_\varphi(1 + |x + \sqrt{pa}|^k + |x + \sqrt{pb}|^k)\sqrt{p}|a - b|
\leq C_\varphi(1 + 2^k(|x|^k + |\sqrt{pa}|^k) + 2^k(|x|^k + |\sqrt{pb}|^k)|a - b|
\leq C_\varphi(1 + 2^{k+1}|x|^k + 2^k(|a|^k + |b|^k))|a - b|
\leq C_\varphi(1 + |a|^k + |b|^k)|a - b|$$

14
in which \( C_\tilde{\phi} = C_\phi \max\{1 + 2^{k+1}|x|^k, 2^k \} \leq C_\varphi (1 + 2^k + 2^{k+1}|x|^k) \). Hence,

\[
(2)_3 \leq C_5 (1 + |x|^k) O(\frac{1}{(n\alpha)^{\alpha/2}})
\]

where \( C_5 = C_\varphi 2^{k+1} \). To summarize, with \( \alpha \in (0, 1) \), for a given \( p \) and fixed \( x \), the error is

\[
|u(1 - p, x) - \varphi_{[np],n}(x)| \leq (1) + (2) \\
\leq c_x O(\frac{1}{n^p}) + 2C_4 (1 + |x|^{k}) O(\frac{1}{n^{\sqrt{p}}}) + C_5 (1 + |x|^{k}) O(\frac{1}{(n\alpha)^{\alpha/2}}) \\
= M(1 + |x|^{k}) O(\frac{1}{(n\alpha)^{\alpha/2}}).
\]

Since \( p \in (0, 1) \), we only need to consider what happens when \( p \) approaches to 0, in order to get a bound similar with a uniform bound, by letting \( q_n := \frac{1}{\sqrt{n}} \), we have

\[
\sup_{p>q_n} |u(1 - p, x) - \varphi_{[np],n}(x)| \leq M(1 + |x|^k) O(\frac{1}{(\sqrt{n})^{\alpha/2}}).
\]

**Remark 3.4.3.** This *iterative algorithm* actually allows us to not only approximate the \( \tilde{E}[\varphi(X)](\approx \varphi_{n,n}(0)) \) but also solve the \( G \)-heat equation \( (u(1 - p, x) \approx \varphi_{[np],n}(x)) \) for any \( \varphi \in C_{l, Lp}(\mathbb{R}) \) without involving any PDE techniques. (In the literature, for instance, the explicit solutions in [Hu (2012)] come from special PDE techniques suitable for functions like \( \varphi(x) = x^{2m+1}, m \in \mathbb{N}_+ \) which can be extended to functions satisfying \( \varphi(\lambda x) = \lambda^\alpha \varphi(x) \) with \( \alpha > 0 \).)

## 3.3 Extension to the \( d \)-dimensional Situation

The definition of semi-\( G \)-normal distribution can be naturally extended to multi-dimensional situation. Intuitively speaking, the multivariate semi-\( G \)-normal distribution can be treated as an analogue of the linear multivariate normal distribution: \( N(0, \mathbf{V}) = \mathbf{V}^{\frac{1}{2}} N(0, \mathbf{I}_d) \), where \( \mathbf{I}_d \) is a \( d \times d \) identity matrix.

Actually, the multivariate semi-\( G \)-normal distribution preserve many similar properties to classical one. For instance, under a special sequential independence setting, we can construct the multivariate semi-\( G \)-normal distribution from several univariate semi-\( G \)-normal distributed random variables. Furthermore, we can use the multivariate semi-\( G \)-normal distribution (constructed from univariate ones) to approach the multivariate \( G \)-normal distribution from the nonlinear CLT. More exploration can be found in the master’s thesis [Li (2018)].

**Definition 3.5** (the Semi-\( G \)-normal distribution in \( d \) dimension). Let a bounded, closed and convex subset \( \mathcal{V} \subset \mathcal{S}_d \) be the uncertainty set of covariance matrices, i.e.

\[
\mathcal{V} := \{ \mathbf{V} = (\rho_{ij}\sigma_i\sigma_j)_{d \times d} : \sigma_i^2 \in [\sigma_i^2, \sigma_i^2], \rho_{ij} = \rho_{ji} = \begin{cases} 1 & i = j \\
\in [\rho_{ij}, \rho_{ij}] & i \neq j, i, j = 1, 2, \ldots, d \end{cases} \}
\]

and \( \mathcal{V}^{\frac{1}{2}} := \{ \mathbf{V}^{\frac{1}{2}} : \mathbf{V} \in \mathcal{V} \} \) where \( \mathbf{V}^{\frac{1}{2}} \) is the symmetric square root of \( \mathbf{V} \). We say a \( d \)-dimensional random vector \( W : \Omega \rightarrow \mathbb{R}^d \) will follow the Semi-\( G \)-normal distribution, denoted as, \( W \sim \tilde{N}(0, \mathcal{V}) \), if there exist a \( d \)-dimensional \( G \)-normal distributed random vector

\[
Y \sim N(0, \mathbf{I}_d) : \Omega \rightarrow \mathbb{R}^d,
\]
and a $d \times d$-dimensional maximal distributed random matrix

$$Z \sim M(V_{1/2}^{1/2}) : \Omega \to \mathbb{R}^{d \times d},$$

as well as $Y$ is independent from $Z$, such that

$$W = Z \cdot Y$$

where "·" is the matrix multiplication (which can be omitted) and the direction of independence here cannot be reversed.

**Remark 3.5.1.** $Y$ can be regarded as the classical multivariate normal distribution with identity covariance matrix.

**Corollary 3.5.1** (the Integral Representation of the Semi-$G$-normal distribution in $d$ dimension). Consider a $d$-dimensional random vector $W \sim \hat{N}(0, V)$, where $V$ is the uncertainty set of covariance matrices. Then for any $\varphi \in C_{l,Lip}(\mathbb{R}^d)$, we have

$$\hat{E}[\varphi(W)] = \max_{V \in V} E[\varphi(N(0, V))] = \max_{V_{1/2} \in V_{1/2}} \int_{\mathbb{R}^d} \frac{1}{(2\pi)^{d/2}} \exp(-\frac{1}{2}y'y)\varphi(V_{1/2}y) \, dy.$$ 

Similarly, we can obtain the iterative approximation in multi-dimensional case.

**Theorem 3.6** (The Iterative Approximation of the $G$-normal Distribution in $d$ dimension). Consider a $G$-normal distributed random variable $X \sim N(0, V)$. For any $\varphi \in C_{l,Lip}(\mathbb{R}^d)$ and integer $n \geq 1$, consider the series of iteration functions $\{\varphi_{i,n}\}_{i=1}^{n}$ with initial function $\varphi_{0,n}(x) := \varphi(x)$ and iterative relation:

$$\varphi_{i+1,n}(x) := \max_{V \in V} E[\varphi_{i,n}(N(x, V/n))], i = 0, 1, \ldots, n - 1.$$ 

The final iteration function for a given $n$ is $\varphi_{n,n}$. As $n \to \infty$, we have $\varphi_{n,n}(0) \to \hat{E}[\varphi(X)]$.

### 4 Implementation

In this section, we will show the implementation of the iterative algorithm which provides a feasible way to approach $\hat{E}[\varphi(N(0, [\sigma^2, \sigma^2]))] \text{ by using } E[\varphi(N(x, v^2))]$.

#### 4.1 The 1-dimensional Situation

Consider a $G$-normal distributed random variable $X \sim N(0, [\sigma^2, \sigma^2])$ with $[\sigma, \bar{\sigma}] = [0.5, 1]$. For any $\varphi \in C_{l,Lip}(\mathbb{R})$, in order to compute $\hat{E}[\varphi(X)]$, setting a fixed large $n$ as the total number of iterations, we implement the following procedure:

1. Start from

$$\varphi_{0,n}(x) := \varphi(x);$$

2. Since we are iterating the functions on the infinite domain $\mathbb{R}$, in practice, we need to set up a finite grid to do interpolation at each step. Choose a large constant $K$ to decide the range of the numerical domain of $x$ then set up the spatial grid:

$$-K = x_0 < x_1 < x_2 < \ldots < x_L = K;$$
3. At the iteration step $i(=1, 2, \ldots, n-1)$, for each $x = x_j$, $j = 0, 1, \ldots, L$, evaluate

$$\varphi_{i+1,n}(x_j) := \max_{v \in [\varphi, \varphi]} E[\varphi_{i,n}(N(x_j, \frac{v^2}{n}))]$$

where the linear expectation can be computed from integration or MC (Monte Carlo) method (by generating a linearly i.i.d. standard normal sample: $Z_1, Z_2, \ldots, Z_M \sim N(0,1)$):

$$E[\varphi_{i,n}(N(x, \frac{v^2}{n}))] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} \exp\left(-\frac{(\frac{hv}{v})^2}{2}\right) \varphi_{i,n}(x+\frac{v}{\sqrt{n}}Z_k) dm$$

Then take maximum of $E[\varphi_{i,n}(N(x, \frac{v^2}{n}))]$ over $v \in [\varphi, \bar{\varphi}]$ to get $\varphi_{i+1,n}(x)$ by doing appropriate optimization: here we use a quasi-Newton method called the “L-BFGS-B” by Byrd et al. (1995) which “allows box constraints, that is each variable can be given a lower and/or upper bound”.

Then use $\varphi_{i,n}(x_j)$, $j = 0, 1, \ldots, L$, to fit the function $\varphi_{i,n}$ by choosing proper splines to do both interpolation and extrapolation based on the type of $\varphi$. For instance, for $\varphi(x) = x^3$, we can use the “fmm” method by Forsythe et al. (1977): “an exact cubic spline is fitted through the four points at each end of the data, and this is used to determine the end conditions.”

Finally, we have

$$\hat{E}[\varphi(X)] \approx \max_{v \in [\varphi, \bar{\varphi}]} E[\varphi_{n-1,n}(N(0, v^2))] = \varphi_{i,n}(0).$$

**Comment 4.0.1.** In the step 3, we notice that it is necessary to use these values $\varphi_{i,n}(x + \frac{v}{\sqrt{n}}Z_k)$, $k = 1, 2, \ldots, M$, with $M$ points in the neighbourhood of one points $x$, to estimate the $\varphi_{i,n}(x)$. If we use a larger grid (with $M^3$ points) for $\varphi_{i,n}$ to preserve the precision of $\varphi_{i,n}$ on a smaller grid (with $M^{i-1}$ points), we will be stuck into the so-called “nested situation” unless we are dealing with functions with bounded domain. As we increase $n$, namely, the number of iterations, even if we only want to compute one point of the last iterative function with certain precision, in the previous iterations, we will still need to prepare a series of grids which is enlarged exponentially with respect to the iteration step. Therefore, one crucial step here is the interpolation (and extrapolation) to avoid the nested dilemma, which can help us get a function with continuous domain from a fixed discrete grid. The splines should be chosen based on the type of $\varphi$ (polynomial, periodic and so on). Meanwhile, we need to appropriately choose the constant $K$ to determine the numerical domain, to make the spline model able to capture the pattern of the function so as to achieve the best extrapolation and interpolation performance.

**Comment 4.0.2.** When evaluating the linear expectation $E[\varphi_i(N(x, \frac{v^2}{n}))]$ (omitting $n$ for a while), to preserve the precision regardless of $x$, we can apply the MC with control variables. One common problem of Monte Carlo method is that given a normal sample, the error of the MC estimation will be enlarged when $x$ goes farther away from zero. Specifically, let $h := \frac{1}{n}$ and $Z \sim N(0,1)$, then

$$E[\varphi_i(N(x, \frac{v^2}{n}))] \approx \frac{1}{M} \sum_{m=1}^{M} \varphi_i(x + \sqrt{hv}Z_i) = \frac{1}{M} \sum_{m=1}^{M} \left( \varphi_i(x) + \frac{\varphi_i^{(1)}(x)}{1!} \sqrt{hv}Z_i + \frac{\varphi_i^{(2)}(x)}{2} hv^2 Z_i^2 + \frac{\varphi_i^{(3)}(x)}{6} hv^3 Z_i^3 \right)$$

Hence, the error can be expressed as,

$$\left| \frac{\varphi_i^{(1)}(x)}{1!} \sqrt{hv} \left( \frac{1}{M} \sum_{m=1}^{M} Z_i - 0 \right) + \frac{\varphi_i^{(2)}(x)}{2} hv^2 \left( \frac{1}{M} \sum_{m=1}^{M} Z_i^2 - 1 \right) \right|$$
Suppose $\epsilon_1 := \frac{1}{M} \sum_{m=1}^{M} Z_i - 0$ and $\epsilon_2 := \frac{1}{M} \sum_{m=1}^{M} Z_i^2 - 1$, for a given sample $\{Z_i\}_{i=1}^{M}$, $\epsilon_1$ and $\epsilon_2$ are fixed number, and even if we regenerate the $\{Z_i\}_{i=1}^{M}$ each time, the random variable $\epsilon_1$ and $\epsilon_2$ should also not have so much variation because we know

$$
\epsilon_1 \sim N(0, \frac{1}{M}), \quad \epsilon_2 \sim \frac{1}{M} \chi^2(M)
$$

then $\text{Var} [\epsilon_1] = \frac{1}{M}$ and $\text{Var} [\epsilon_2] = \frac{2}{M}$. However, for the case $\varphi_1(x) = x^3$, when $x$ goes farther away from zero, the values of $|\varphi_1^{(1)}(x)| = 3|x|^2$ and $|\varphi_1^{(2)}(x)| = 6|x|$ will become larger. Therefore, without loss of generality, for a fixed normal sample, the error of the MC will increase as $x$ goes away from zero. This problem can not be overcome by simply increase the $M$ (since the $|\varphi_1^{(1)}(x)|$ and $|\varphi_1^{(2)}(x)|$ will always expand the small error anyway and these enlarged errors will be cumulated as time goes further backward). Fortunately, we may try a type of variance reduction method for Monte Carlo method, called the control variables. We can use this method to preserve the consistent precision of MC method outside the neighbourhood of $x$. The idea is that, after getting the $\varphi_1^{(1)}(x), i = 1, 2$, approximate the $E[\varphi(x + \sqrt{hv}Z)]$ by

$$
E[\varphi(x + \sqrt{hv}Z) - \varphi(x) - \varphi(x) - \varphi^{(1)}(x)\sqrt{hv}Z - \varphi^{(2)}(x)hv^2Z^2] + \varphi(x) + \frac{\varphi^{(2)}(x)hv^2}{2}
$$

$$
\approx \frac{1}{M} \sum_{i=1}^{M} \left[ \varphi(x + \sqrt{hv}Z_i) - \varphi(x) - \varphi^{(1)}(x)\sqrt{hv}Z_i - \varphi^{(2)}(x)hv^2Z_i \right] + \varphi(x) + \frac{\varphi^{(2)}(x)hv^2}{2}
$$

$$
(= \frac{1}{M} \sum_{i=1}^{M} \left[ \varphi^{(3)}(\xi_x, Z_i) \right] + \varphi(x) + \frac{\varphi^{(2)}(x)hv^2}{2}).
$$

In this way, because of the boundedness of $\varphi^{(3)}(x) = 6$ for $\varphi(x) = x^3$, the error will not be enlarged by $\varphi^{(3)}(\xi_x, Z_i)$ when $x$ moves farther away zero.

**Remark 4.0.1.** Let $X \sim N(0, [\sigma^2, \bar{\sigma}^2])$. In the context of the 1-dimensional $G$-heat equation:

$$
u_t + G(\partial^2_x u) = 0, \quad u|_{t=1} = \varphi
$$

where $G(a) := \frac{1}{2} \hat{E}[aX^2] = \frac{1}{2}(\sigma^2 a^+ - \bar{\sigma}^2 a^-)$. We know there is a natural connection between the $G$-heat equation and $G$-normal distribution:

$$
u(t, x) = \hat{E}[\varphi(x + \sqrt{1-t}X)].
$$

Then we can make a collection of the iterative functions with properties:

1. $\varphi_{n,n}(0) \approx \hat{E}[\varphi(X)] = u(0,0)$;
2. $\varphi_{n,n}(x) \approx u(0,x)$;
3. In general, $\varphi_{k,n}(x) \approx u(1-k/n, x)$, for $k = 0, 1, \ldots, n$.

This will actually give us a surface of the solution $u(t, x)$ with $t = 1 - \frac{k}{n}, k = 0, 1, \ldots, n$.

In the following $2 \times 2$-layout figures, for the first row ($n = 50, K = 5$), Panel-(1,1) is the approximated solution paths of the $G$-heat equation with $\varphi(x) = x^3 = \varphi_{0,n}(x)$ (the black solid curve) and the curves whose right branches are moving up are the $\varphi_{i,n}$’s as the iteration proceeds. The Panel-(1,2) shows the approximated paths with $\varphi(x) = (1 - |x|)\mathbb{1}_{(1,1)}(x)$ (which belongs to a class of functions that, previously, are hard to deal with if not applying special PDE methods). The Panel-(2,1) is the pathwise comparison plot ($n = 100, K = 50, \varphi(x) = x^3$) with the explicit solution provided in [Hu (2012)] where we can only notice the error (with absolute value $\leq 0.004$) around $x = 0$ and we can see the stable accuracy beyond the spatial grid $[-K, K]$ (the black dashed lines). The Panel-(2,2) is the approximation of $\hat{E}[X^3]$ as $n$ increases ($K = 5$) (where the horizontal line labels the true value; here we use integration to compute $E$, if using MC, we will assign a variance for each point).
For curiosity, we can play with this algorithm \((n = 50, K = 5)\) by changing the terminal function to other \(\varphi \in Cl.Lip\): Panel-(1, 1) \((\varphi(x) = \sin x)\), Panel-(2, 1) \((\varphi(x) = \cos x)\), Panel-(1, 2) \((\varphi(x) = \sin x + \cos x)\), and Panel-(1, 1) \((\varphi(x) = 1/(1 + \exp(-x^2)))\).

(I am still working on the implementation when involving the viscosity solution like the case with \(\sigma = 0\).)
4.2 The $d$-dimensional Situation

For the multi-dimensional $G$-heat equation (with covariance uncertainty), the MC setting may relieve the iterative algorithm from the curse of dimensionality.

Let $x := (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. Consider a $d$-dimensional $G$-normal distributed random vector $X \sim N(0, \mathcal{V})$, where $\mathcal{V}$ is the uncertainty set of covariance matrices. For any $\varphi \in C_{l,Lip}(\mathbb{R}^d)$, in order to compute $\mathbb{E}[\varphi(X)]$, similarly, set a large $n$ as the total number of iterations, then we need to do the procedure:

1. Start from
   $$\varphi_{0,n}(x) := \varphi(x);$$

2. We need to set up a grid in the domain $\mathbb{R}^d$, to avoid the curse of dimensionality in the sense that the rectangular grid points become sparser in higher dimension and much more with exponential rate ($L$ points in each dimension mean $L^d$ points in total). Therefore, here we use the Monte Carlo grid points sampling from a $d$-dimensional multivariate normal distribution:
   $$\{x_j\}_{j=0}^L \sim N(0, \sigma^2 I_d).$$

3. For $i = 1, 2, \ldots, n$, and for each $x = x_j$, $j = 0, 1, \ldots, L$, let
   $$\varphi_{i+1,n}(x) := \max_{V \in \mathcal{V}} E[\varphi_{i,n}(N(x, \frac{V}{\sqrt{n}}))];$$

   where, again, to deal with the curse of dimensionality, the expectation can be computed by Monte Carlo method which maintains its convergence rate regardless of the dimension by generating a linearly i.i.d. sample from standard multivariate normal $N(0, I_d)$: $Z_1, Z_2, \ldots, Z_M$ (and we can also apply the control variable method to protect the precision from the variation of $x$):
   $$E[\varphi_{i,n}(N(x, \frac{V}{\sqrt{n}}))] \approx \frac{1}{M} \sum_{k=1}^M \varphi_{i,n}(x + \frac{V}{\sqrt{n}} Z_k).$$

Take maximum of $E[\varphi_{i-1,n}(N(x, \frac{V}{\sqrt{n}}))]$ over $V \in \mathcal{V}$ by appropriate optimization (here we still use the “L-BFGS-B” method). Then use $\varphi_{i+1,n}(x_j)$, $j = 0, 1, \ldots, L$, to fit the function $\varphi_{i+1,n}$ by applying a proper spline model to do interpolation and extrapolation (here we work on the setting of Generalized Additive Model after testing and design the structure of splines based on the properties of $\varphi$).

Finally, we have

$$\hat{\mathbb{E}}[\varphi(X)] \approx \max_{V \in \mathcal{V}} E[\varphi_{n-1,n}(N(0, \frac{V}{\sqrt{n}}))].$$

Remark 4.0.2. Consider a 2-dimensional $G$-normal distributed random variable $X \sim N(0, \mathcal{V})$ where

$$\mathcal{V} := \left\{ V = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} : \sigma_i \in [0.5, 1], i = 1, 2; \rho \in [-0.5, 0.5] \right\}. $$

In the context of 2-dimensional $G$-heat equation:

$$u_t + G(D^2 u) = 0, \quad u|_{t=1} = \varphi$$
where $G(A) := \frac{1}{2} \mathbb{E}[(AX, X)] : S_d \to \mathbb{R}$, we also have the natural connection between the $G$-heat equation and $G$-normal distribution,

$$u(t, x) = \mathbb{E}[\varphi(x + \sqrt{1-t}X)], \ (t, x) \in [0, \infty] \times \mathbb{R}^2.$$

Then these results have been verified:

1. we have $\varphi_{n,n}(0) \approx \mathbb{E}[\varphi(X)] = u(0, 0)$ which repeats the above remark;
2. by replacing $\varphi(X)$ with a shifted version $\varphi(x + X)$, we directly get $\varphi_{n,n}(x) \approx u(0, x)$;
3. In general, we have $\varphi_{k,n}(x) \approx u(1 - \frac{k}{n}, x)$, for $k = 0, 1, \ldots, n$. In other words, the function in each iteration step has its connection with the solution of $G$-heat equation.

This is the solution surface with $\varphi(x) = x_1^3 + x_2^3$, the number of iteration steps $n = 10$, and

$$\mathcal{V} := \left\{ V = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \rho \sigma_1 \sigma_2 \end{pmatrix} : \sigma_i \in [0.5, 1], i = 1, 2; \rho \in [-0.5, 0.5] \right\}.$$
In the above figure, the black surface is $u(1, x) = \varphi(x)$ and the green one is $u(0, x)$.

(I am still working on more details for this and also the Generalized Additive Model with control variable method.)

5 Assessment of the Iterative Method

The strengths of this iterative method are well-worth mentioning:

1. It will work for any $\varphi$ applicable in the nonlinear central limit theorem, including some irregular $\varphi$ which makes $\hat{E}[\varphi(X)]$ difficult to compute even by using classical numerical PDE methods. This property has been verified in the proof of Theorem 3.6 and has been checked numerically for one dimension.

2. It will numerically solve the corresponding $G$-heat equation by not just giving one point or one path but actually directly providing the whole surface of $u(t, x)$, because the function in each iteration step has its connection with the $u(t, x)$ at one grid point $t$. This is given by Theorem 3.6.

3. It will give us a great visualization and intuition about how the solution surface of $G$-heat equation is evolved from the terminal function $\varphi(x)$ by looking at the procedure of iteration, and how the sublinear expectation of $G$-normal distribution is approached by iteratively maximizing the linear expectation of classical normal distribution. The inherent bridge or ladder here is the Semi-$G$-normal distribution which helps us climb from the stage of classical normal distribution to reach the stage of $G$-normal distribution. It partially fills in the long-existing thinking gap between them, since we have the inequality: $\hat{E}[\varphi(X)] \geq \sup_{v \in [\sigma, \sigma]} E[\varphi(N(0, v^2))]$ (for various $\varphi$, especially when $\varphi$ is neither convex nor concave, the strict greater relation is much more frequent than equality), which, in some sense, strictly separated the $G$-normal distribution with the classical one and made us feel a little risky to connect them for a long time;

4. It can be naturally extended to higher dimension both in theory and algorithm, since we already have the established multi-dimensional distributions in $G$-framework, as well as the algorithm for computation of linear expectation of classical multivariate normal distribution. Then it can solve the corresponding multi-dimensional $G$-heat equation attached with covariance uncertainty;

However, for numerical practice, we still need to reflect on how to properly choose the splines to achieve better interpolation and extrapolation performance.

6 Concluding Remarks and Further Development

The semi-$G$-normal distribution proves its own value by giving us this iterative approximation. This is a typical distribution connecting the maximal distribution from nonlinear framework and the classical normal distribution from linear framework.

In the master’s thesis Li (2018), we further explore the semi-$G$-normal distribution by considering its independence, its multivariate version (with the construction from univariate objects) and the semi-$G$-Brownian motion (with its connection to the $G$-Brownian motion. For the statistical side, we also provide a pseudo approach to simulate the semi-$G$-normal distribution and the estimation method for the variance interval of $\hat{N}(0, [\bar{\sigma}^2, \bar{\sigma}^2])$. The semi-$G$-normal distribution $\hat{N}(0, [\bar{\sigma}^2, \bar{\sigma}^2])$
behaves like the transition from the linear normal distribution $N(0, \sigma^2)$ to the sublinear $G$-normal distribution $N(0, [\sigma^2, \sigma^2])$.

The set of measures of the semi-$G$-normal distribution, consisting of a class of linear measures of $N(0, \sigma^2)$ with $\sigma \in [\sigma, \sigma]$, is smaller than that of the $G$-normal distribution. Hence, on the one hand, it has less unusual properties than the $G$-normal distribution and more similar properties to the linear normal distributions (e.g. its sublinear “skewness” is equal to zero); on the other hand, it has strong connections with $G$-normal distribution (e.g. the sublinear expectation with convexity and the nonlinear CLT). As a crucial concept in the thesis [Li (2018)], the semi-$G$-normal distribution is aimed for constructing a bridge between the linear and the sublinear world so that in the future, more tools on the land of linear expectation framework can be transformed to the island of the sublinear expectation framework to help the larger community in both the academy and industry to understand the intuition, learn the theory or algorithms and apply them to the real world problems with uncertainty.

References

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16(5), 1190–1208.

Denis, L., M. Hu, and S. Peng (2011). Function spaces and capacity related to a sublinear expectation: application to $G$-Brownian motion paths. *Potential Analysis* 34(2), 139–161.

Forsythe, G. E., C. B. Moler, and M. A. Malcolm (1977). *Computer methods for mathematical computations*. Prentice-Hall.

Hu, M. (2012). Explicit solutions of $G$-heat equation with a class of initial conditions by $G$-Brownian motion. *Nonlinear Analysis* 75(18), 6588–6595.

Li, Y. (2018). Statistical exploration in the $G$-expectation framework: the pseudo simulation and estimation of variance uncertainty. Master’s thesis, The University of Western Ontario, London, Canada. to be defended in July.

Pao, C.-V. (2012). *Nonlinear parabolic and elliptic equations*. Springer Science & Business Media.

Peng, S. (2007). $G$-expectation, $G$-Brownian motion and related stochastic calculus of itô type. In *Stochastic analysis and applications*, pp. 541–567. Springer.

Peng, S. (2008). Multi-dimensional $G$-Brownian motion and related stochastic calculus under $G$-expectation. *Stochastic Processes and their Applications* 118(12), 2223–2253.

Song, Y. (2017). *Normal Approximation by Stein’s Method under Sublinear Expectations*. ArXiv:1711.05384.

Wang, L. (1992). On the regularity theory of fully nonlinear parabolic equations: I. *Communications on pure and applied mathematics* 45(1), 27–76.