A model for stocks dynamics
based on a non-Gaussian path integral

Giovanni Paolinelli   Gianni Arioli

Dipartimento di Matematica
Politecnico di Milano
piazza Leonardo da Vinci 32, 20133 Milano, Italy
giovanni.paolinelli@polimi.it, gianni.arioli@polimi.it

Abstract

We introduce a model for the dynamics of stock prices based on a non quadratic path integral. The model is a generalization of Ilinski's path integral model, more precisely we choose a different action, which can be tuned to different time scales. The result is a model with a very small number of parameters that provides very good fits of some stock prices and indices fluctuations.

Keywords: Quantum-finance, path integral, gauge theory, financial markets, fat tails, non-Gaussian dynamics.

1   Introduction

Stochastic calculus has been a fundamental tool in finance since the publication of the Ph.D. thesis of Louis Bachelier [19] in 1900. Much more recently, alternative approaches based on the techniques used in quantum mechanics and quantum field theory have been proposed. In particular, the path integral approach to quantum mechanics introduced by Feynman in [20, 21] has turned out to be particularly effective for financial applications, see also [4-12].

Ilinski in [1, 2] establishes two analogies between finance and quantum mechanics: arbitrage is compared with the Lagrangian action, and the invariance of the behaviour of the market with respect to the choice of the currency of reference is seen as a gauge invariance. The framework proposed by Ilinski has been considered and generalized in many papers, see e.g., [9, 14, 15, 16, 18].

In classical mechanics, the evolution of a dynamical system can be described as the trajectory that minimizes the Lagrangian action. In quantum mechanics one can describe the evolution of a system only in probabilistic

terms, but the least action principle is not completely lost: loosely speaking, the transition probability between different states can be computed with the aid of a path integral by considering all trajectories, assigning a probability depending on the action to each trajectory and averaging among all trajectories. Such probability is maximal for the least action trajectory, and Plank's constant h determines how fast the probability decreases when other trajectories are considered.

Classical finance is based on a condition which resembles the principle of least action: arbitrage, i.e., the possibility to make a positive amount of money without risk, is minimized by the market, the minimum being zero.

Ilinski proposes a new approach where arbitrage behaves as the action in quantum mechanics. The most probable trajectory for a stock price in the financial market is the one corresponding to zero arbitrage, but trajectories with some amount of arbitrage are possible, although their probability decreases exponentially with respect to the arbitrage.

The second analogy proposed by Ilinski is gauge invariance. The behaviour of the investors, and hence the trajectory of the prices in a market, does not depend on the choice of currency. Then, all quantities used in the theory have to be invariant under the action of a Lie group $\mathbb{R}^+$, which represent the change between currencies.

Ilinski describes the financial market as a principal fibre-bundle (PFB) with gauge group $\mathbb{R}^+$. This concept is borrowed from physics as well, and it is proposed as the best way to generate the gauge-invariant quantities, since principal fibre-bundles are specifically designed to describe environments where gauge-invariance is required.

The model described in [1] corresponds to a Geometric Brownian Motion (GBM), which generates a normal probability density function; this is not satisfactory, since stocks prices are known to display a leptokurtic distribution. Large price variations are underestimated by the GBM, causing significant errors in the computation of quantities depending on the stock price trajectories, such as the price of derivatives. More precisely, in [17] it is shown that, if the fat tails are ignored, then the rate of return of an investment with no risk of financial loss and the term premium (the compensation that investors require for bearing the risk that short-term Treasury yields do not evolve as they expected) are miscomputed.

Ilinski also introduces in [2] an improved model combining the minimization of the arbitrage with a perturbation generated by the impact of the orders.

This model has been numerically studied and improved in [9], where the numerical computations show a good agreement between the theoretical model and the real data concerning the short term dynamics.

The results of [9] clearly show that the perturbation plays a fundamental role in the good agreement between the model and the real market, since it is responsible of the change of the distribution of prices from mesokurtic to

2

leptokurtic. A similar approach has been developed in [14, 15, 16], where a model based on quantum field theory has been developed.

All these results imply that the minimization of the arbitrage does not suffice to describe the short term dynamics of the financial market.

The perturbation of the GBM is used also in classical finance, with stochastic volatility models; this approach has the main drawback of a high number of parameters. In this paper we propose to maintain the basic ideas introduced by Ilinski, but we choose a different action. In the following sections, the interpretation of the financial market as principal fibre bundle is assumed, and we refer to [1, 2] for an introduction of the topic.

2   The non-Gaussian path integral

Many financial models consist in a Gaussian noise and an additional perturbation, see e.g. [9, 14, 29]. Clearly, it is necessary to choose one or more parameters for each random noise. This means that the models presented in [9, 14, 29] are characterized by a high number of parameters. The main idea that we present in this paper is a model where the main dynamics is not Gaussian, so that a perturbation is not required, and only parameters for the main dynamics are introduced. Our goal is a model with a precision similar to the model in [9, 14, 29], but simpler and with a lower number of parameters. We begin by observing how the path integral encodes the Gaussian behaviour. Consider the stochastic differential equation

$$\dot{x}(t) = \eta(t) \tag{1}$$

where

$$x(t) = \log(S_t).$$

The noise η does not need to be Gaussian; let $\tilde{D}(\eta)$ be the probability distribution of η and define $L(\eta)$ by

$$\tilde{D}(\eta) = e^{-L(\eta)}. \tag{2}$$

To underline the analogy with quantum mechanics, we call the function $L(\eta)$ the Lagrangian. In [28, see 20.1.13] it is shown that the path integral

$$P(x_b, t_b | x_a, t_a) = \int D\tilde{\eta} \int Dx \exp[-\int_{t_a}^{t_b} L(\tilde{\eta}(t)) dt] \delta(\dot{x} - \tilde{\eta})$$

$$= \int Dx \exp[-\int_{t_a}^{t_b} L(\dot{x}) dt] \tag{3}$$

returns the probability distribution of $x_b = x(t_b)$ when $x(t_a) = x_a$.

If we compare the transition probability formula proposed by Ilinski

$$P(S(T)|S(0))_{GBM} \propto \int D \log(S) e^{-\frac{1}{2\sigma^2} \int_0^T (\partial_t \log(S_t) - r)^2 dt}, \tag{4}$$

3

with (3), we see that the Lagrangian associated to the GBM described by Ilinski is

$$L_{GBM} = \frac{1}{2\sigma^2}(\partial_t \log(S_t) - r)^2 \tag{5}$$

and then

$$\tilde{D}(\partial_t \log(S_t)) \simeq N(r_{GBM}, \sigma^2) \tag{6}$$

with

$$r_{GBM} = r - \sigma^2 T/2\,.$$

The previous analysis, together with formula (5), explains why the log-returns in (4) are normally distributed and why the Gaussian nature of the perturbation is encoded in the structure of the Lagrangian. The latter observation is the starting point of our model.

Remark. Formula (4) represents a transition probability of the price $S(t)$ expressed as an integral in $D \log S_t$, while (3) represents the transition probability of $x = \log(S)$ expressed as an integral in Dx. This difference is due to the fact that (3) is not written in an explicit gauge-invariant form, since the measure $D \log(S) = DS/S$ is gauge-invariant ($D \log(S) = DS/S$), while the measure DS is not.

3 The new model

3.1 Setting

In order to present the basic hypotheses underlying the new model, we begin with the definition of the path-space in the financial context. Given a time interval $[0, T]$, where $T \in \mathbb{R}^+$ represents our choice of time horizon, we define

$$X_{S_0, S_T} = \{S \in BV([0, T], \mathbb{R}^+) \,:\, S(0) = S_0,\, S(T) = S_T\} \tag{7}$$

as the set of all the paths connecting two prices $S_0, S_T \in \mathbb{R}^+$. The set of all admissible paths is

$$X_T = \bigcup_{S_0, S_T \in \mathbb{R}^+} X_{S_0, S_T}\,.$$

We consider the BV functions in order to take in account both continuous and non-continuous stock-price paths.

The basic hypotheses for our model are:

- **(HP1)** There exists a functional $\mathcal{A} : X_T \to \mathbb{R}$, invariant with respect to the action of the gauge group $\mathbb{R}^+$, and related to the amount of money that is possible to earn with some trading strategy during the time horizon $[0, T]$.

- **(HP2)** The market evolves fluctuating around the configuration of minimal value of $\mathcal{A}$.

- **(HP3)** The financial market is not fully efficient, small amounts of arbitrage are allowed.

3.2 Stochastic minimization

In order to state rigorously what we mean by "fluctuation around the configuration of minimal action", we introduce the concept of *stochastic minimization*. Assume that $\mathcal{A} : X_T \to \mathbb{R}^+$ is an action which satisfies **HP1**, let $P : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function, and let

$$p : X_T \to \mathbb{R}^+ , \quad p(S) = P(\mathcal{A}(S)). \tag{8}$$

Since P is decreasing, the classical path $\bar{S}$ which minimizes the action, corresponds to the path that maximizes p. By *stochastic minimization* we are assuming that all paths are possible, with probability assigned by p. All quantities related to the path of prices (probability of the final price or related quantities, such as the value of a derivative product) has to be computed on each path, and the result is the weighted average with respect to p on all paths. There are two main differences with respect to the computation of a propagator in quantum mechanics: first, in quantum mechanics the action is multiplied by the imaginary unit and the transition probability is the modulus square of the path integral; second, for the financial path integral we only require that P is a decreasing function, not necessarily an exponential. The analogy with quantum mechanics also concerns the methods adopted to compute the transition probability, as we show in Section 5. It is straightforward to observe that the system described by Ilinski consists in the *stochastic minimization* of the arbitrage, with respect to an exponential function P.

3.3 The new action

We denote by $\mathcal{A}_{arb}$ the amount of money that is possible to earn with arbitrage in the time horizon $[0, T]$. Ilinski shows in [1] that

$$\mathcal{A}_{arb}(S_t, T) = \beta_2 \int_0^T (\partial_t \log(S_t) - r_t)^2 dt = \beta_2 \int_0^T \Omega(t)^2 dt, \tag{9}$$

where r is the net non-risky interest rate and β_2 is a constant with the dimension of time. The function $\Omega(t)$ is the curvature of the principal fibre-bundle considered by Ilinski, therefore $\mathcal{A}_{arb}$ is gauge invariant, see [2, Chap. 5]. Note that the curvature has the dimension of time inverse, therefore β_2 needs to have the dimension of time; similarly, in quantum mechanics case the action is divided by h. Note that β_2 weighs the paths variation with respect to $\bar{S}$, that is it corresponds to the inverse of the variance squared.

Since the curvature is gauge invariant, it is a natural component of a gauge-invariant functional. A different gauge invariant functional is the $L^1_{[0,T]}$-norm of $\Omega(t)$

$$\mathcal{A}_{me}(S_t, T) = \beta_1 \int_0^T |\partial_t \log(S_t) - r_t| dt = \beta_1 \int_0^T |\Omega(t)| dt. \qquad (10)$$

In this case β_1 provides a scale for the variation of the paths with respect to $\bar{S}$. Adopting the same point of view of Ilinski, we consider $\mathcal{A}_{me}(S_t, T)$ as the maximal amount of money that is possible to earn in the time horizon $[0, T]$. The strategy to obtain that gain consists in borrowing cash and buying the stock when the stock outperform the cash, and vice versa. This strategy involves some risk, since it is not possible to know in advance what is the right case. It is straightforward to check that the functionals (9) and (10) satisfy the hypotheses **HP1**. Note that

$$\min \mathcal{A}_{me}(S_t, T) = \min \mathcal{A}_{arb}(S_t, T) = 0, \iff \partial_t \log(S_t) = r_t$$

and, if $T_2 > T_1$ and $S_1(t) = S_2(t)$ for all $t \in [0, T_1]$, then

$$\mathcal{A}_{me}(S_1(t), T_1) < \mathcal{A}_{me}(S_2(t), T_2) \quad \text{and} \quad \mathcal{A}_{arb}(S_1(t), T_1) < \mathcal{A}_{arb}(S_2(t), T_2).$$

It is also important to note that, if $\beta_1 = \beta_2$ and assuming that S_t does not minimize the actions, then

$$\mathcal{A}_{me}(S_t, T) > \mathcal{A}_{arb}(S_t, T). \qquad (11)$$

The previous inequality implies that, if we use the same scale to measure the revenues from $\mathcal{A}_{me}$ and $\mathcal{A}_{arb}$, then the first strategy is more rewarding than the second one.

This last fact may seem counter-intuitive. We recall that (11) is computed over the continuous-time version of the process S_t. In the real case the stock price does not vary continuously, since a price variation happens only when there is a new order with a different price. The time scale associated to the price variation between two successive orders is called the "tick-by-tick" variation of the price, and it is the time frame that must be selected to compute the previous quantities. Log-price variations in this context are always very small, implying (11). The analysis of the probability density functions of the one minute dynamics confirms this statement.

We now show that $\mathcal{A}_{me}$ is a reasonable choice for the action functional in the case of a short time horizon, whereas $\mathcal{A}_{arb}$ is more suitable in the case of a long time horizon.

3.3.1   The short time horizon

Assuming a closed market, we note that the amount of money that a trader gains corresponds to the loss of all the other traders. The fairest situation is the one without losses, which implies no gain as well.

This happens when the value of the functional $\mathcal{A}_{me}$ is equal to 0. Market configurations with $\mathcal{A}_{me} > 0$ are associated with losses among other players and therefore they should be considered less probable.

In the short term dynamics, the minimization of the arbitrage is in contrast with **HP3**. The functional (9) is proportional to T and it is smaller than $\mathcal{A}_{me}$, i.e., it is negligible.

3.3.2   The long time horizon

Since the amount of arbitrage is directly proportional to the time horizon T, we can assume that in this case it is non negligible. Clearly, inequality (11) still holds, but in this case we have to consider that arbitrage does not involve risk. It is reasonable to assume that investors prefer the arbitrage strategy albeit it is less profitable, since the gain is sure. In this situation, it is essential to minimize the arbitrage in order to prevent the possibility to extract considerable amount of money without risk from the market in order to respect **HP3**.

3.3.3   Intermediate-time horizon

To consider an intermediate-time horizon, we assume that it is possible to modify continuously the action functional from $\mathcal{A}_{arb}$ to $\mathcal{A}_{me}$. Given the structure of the previous functionals, the most natural choice seems to be

$$\mathcal{A}_p(S_t, T) = \beta_p \int_0^T |\partial_t \log(S_t) - r_t|^p dt, \quad p \in [1, 2]. \tag{12}$$

This action may be interpreted as a hybrid strategy which involves both arbitrage and risk trading.

The previous formula is unconventional in the theory of path integrals, where the action is usually quadratic plus some perturbation. The previous definitions have to be intended in a formal sense; their interpretation and numerical evaluation is discussed in section 4.

3.4   The new stochastic dynamics

We now apply the concept of stochastic minimization to the action functionals previously defined, giving also the explicit formulation of the probabilities adopted.

The arbitrage strategy proposed by Ilinski can, in principle, be followed by every trader in the market. This means that it would be very simple to make a profit with an arbitrage transaction, i.e., to extract consistent amount of money from the market without risk. In order to forbid this eventuality, it is necessary that big arbitrage opportunities are very unlikely. This justifies the choice adopted by Ilinski which uses the exponential function.

The short time horizon case is different. The most probable configuration of the market does not admit capital losses, but different traders have different trading experience. We can assume that the more experienced trader will make a profit over the unskilled one; this means that high values of the functional $\mathcal{A}_p$ with $p \in [1, 2)$; should be more likely, given the same amounts of arbitrage. To account for this, we propose as probability density function

$$P_\gamma(\mathcal{A}_p) \propto e^{-\beta_p \mathcal{A}_p^\gamma} \quad \text{with} \quad \gamma \in (0, 1);$$

to maintain the analogy with the GBM case we define β by

$$\beta_p = \frac{1}{2\sigma^p}.$$

Note that the parameter σ is not the volatility, but it is related to it. Considering the action $\mathcal{A}_p$ with the probability defined above and **HP2**, the fluctuations of the market around the minimum action configuration turn out to be described by

$$P_{\gamma,p}(\mathcal{A}_p) \propto e^{-\frac{1}{2\sigma^p} \left(\int_0^T |\partial_t \log(S_t) - r_t|^p dt \right)^\gamma}, \tag{13}$$

and the GBM is recovered when $\gamma = 1$ and $p = 2$.

It is interesting to compare the pdf obtained with the GBM against different (γ, p) cases. The condition $\gamma < 1$, i.e. the intermediate time horizon, implies that small amounts of action $\mathcal{A}_p$ are less likely to happen with respect to the same amount of arbitrage in the GBM case. This is in agreement with our hypothesis, since by **HP3** small amount of arbitrage are allowed. Moreover, the arbitrage is not the quantity minimized in the short therm dynamics. On the other side, large values of the action $\mathcal{A}_p$ are more likely to happen with respect to the same amount of arbitrage in the GBM case; this is in agreement with the condition that requires the minimization of the arbitrage.

Recalling the path integral formalism introduced by Feynman in [20], we obtain that the transition probability is the gauge-invariant weighted mean with respect to $P_{\gamma,p}$ over the space of all paths connecting $S(0)$ with $S(T)$, i.e.,

$$P(S(T)|S(0)) = \sum_{S_t \in X_{S_0, S_T}}^{\sim} P_{\gamma,p}(\mathcal{A}_p(S_t)),$$

where the tilde notation indicates that the sum is computed adopting a gauge-invariant measure. The above formula means that the transition probability $P(S(T)|S(0))$ is composed of the contributions of all the possible paths connecting the initial and the final states. Each path contributes inversely with respect the amount of action associated to it; the magnitude of the contributions is controlled by the probability $P_{\gamma,p}$. An interesting mathematical problem is to define a way to sum all the paths, respecting

8

the gauge invariance; this will be discussed below.

The sum over the path space is an infinite dimensional integral

$$\int DS = \lim_{N \to \infty} \prod_{i=1}^{N-1} \int_0^\infty dS_i.$$

Note that the definition above is not gauge invariant. Given a gauge group G, a gauge-invariant function $f(x)$ and a measure $d\mu(x)$, we require that

$$\int f(x)d\mu(x) = \int f(gx)d\mu(gx) \quad \forall g \in G.$$

Such measure is known as the Haar measure of the group. For the group $\mathbb{R}^+$ the Haar measure exists and it is unique, up to a multiplicative constant; it is given by

$$dS/S = d\log(S).$$

Finally, the formula for the transition probability is

$$P(S(T)|S(0)) \propto \int D \log(S) e^{-\frac{1}{2\sigma^p} \left(\int_0^T |\partial_t \log(S_t) - r|^p dt \right)^\gamma}. \tag{14}$$

4   Definition and approximation of the path integrals

There exists a rigorous theory of path integrals only when the Lagrangian is quadratic. When a small perturbation is added, as it happens e.g. in quantum electrodynamics, it is possible to evaluate a path integral using perturbation techniques based on Feynman diagrams, but such techniques are not mathematically rigorous; moreover, they involve infinite quantities which require renormalization theory to be dealt with. In our model, when $p \neq 2$ there is no quadratic part to use as a starting point, therefore it is necessary to provide both a rigorous definition and a computation algorithm to formula (14).

The model proposed by Ilinski has a Gaussian Lagrangian, therefore it is defined as

$$P(S(T)|S(0))_{GBM} = \int D \log(S) e^{-\frac{1}{2\sigma^2} \int_0^T (\partial_t \log(S_t) - r_t)^2 dt}$$

$$\approx \lim_{N \to \infty} \prod_{i=1}^{N-1} \int_0^\infty \frac{dS_i}{S_i} \exp \left[-\frac{T/N}{2\sigma^2} \sum_{i=0}^{N-1} \left(\frac{\log(S_{i+1}) - \log(S_i)}{T/N} - r \right)^2 \right]. \tag{15}$$

The Gaussian nature of (15) guarantees two important properties: the finite dimensional approximation does not depend on N and the probability distribution derived from the calculation is again log-normal.
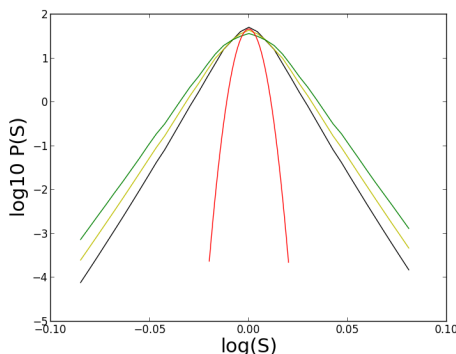
9

Figure 1: Output of the algorithm when $\gamma = 0.4$, $p = 1.4$ and $\sigma = 0.014$, and $N = 10$ (blue), $N = 11$ (yellow), $N = 12$ (green). The red curve represents a Gaussian pdf with $\sigma = 0.00128$.

Consider now (14). The most natural definition would be the analogous of (15), but our computations suggest that the sequence does not approach a finite limit, see Figure 1. In order to propose a well posed definition, we consider first the case $\gamma = 1$, that is

$$P(S(T)|S(0)) \propto \int D\log(S)e^{-\frac{1}{2\sigma^p}\int_0^T |\partial_t \log(S_t) - r_t|^p dt} \, ,$$

and we compare our problem with a standard problem in probability theory.

Consider a collection of i.i.d. stable Lévy random variables $\{X_n\}$ with characteristic function

$$\phi(q) = e^{-\gamma|q|^\alpha} \quad \text{with} \quad \alpha \le 2 \, .$$

The previous distribution is stable, see e.g. [4, Sec 4.2]; this amounts to say that the sum of N rescaled random variables $\{X_n\}$

$$\tilde{S}_N = \sum_{n=0}^{N} X_n/N^{1/\alpha}$$

is also a Lévy random variable with the parameters independent from N, provided that we also rescale the probability $P(\tilde{S}) \mapsto P(\tilde{S})N^{1/\alpha}$.

Clearly, when $\alpha = 2$ we recover the Gaussian case, [4, Sec 3.3], and the rescaled variables are also Gaussian.

A similar phenomenon happens with the path integral definition. We set

$$P(S(T)|S(0)) = \int D\log(S)e^{-\frac{1}{2\sigma^2}\int_0^T |\partial_t \log(S_t) - r_t|^p dt}$$

$$\approx \lim_{N\to\infty} \prod_{i=1}^{N-1} \int_0^\infty \frac{dS_i}{S_i} \exp\left[-\frac{(T/N)^{p/2}}{2\sigma^p} \sum_{i=0}^{N-1} \left| \frac{\log(S_{i+1}) - \log(S_i)}{T/N} - r \right|^p \right],$$

$$(16)$$

10

and we check that the numerical approximation appears to be convergent, see Figures 2 and 3 (left). We recover the familiar quadratic definition in the case $p = 2$.

In order to support the choice of the exponent $p/2$, for some choices of (p, σ) we estimated the path integrals with different values of N, and we observed that the resulting pdf's converge when N is increased. Figures 2 and 3 show the results for some of these choices. The relative errors of the computations with $\gamma = 1$ is less than 3% .
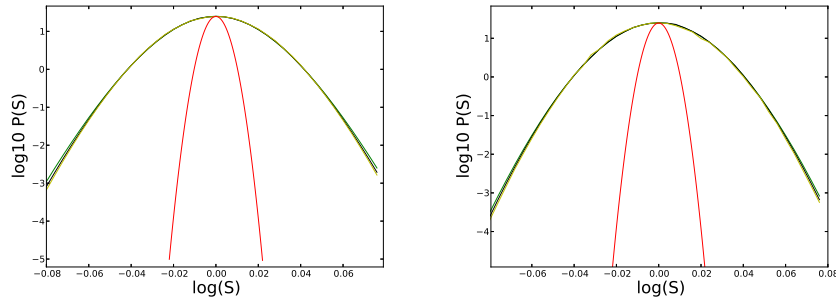
Figure 2: Left: results with $\gamma = 1$, $p = 1.2$, $\sigma = 0.0027$, $\Delta \log S = 0.16$. Right: results with $\gamma = 1$, $p = 1.5$ and $\sigma = 0.0037$, $\Delta \log S = 0.16$. For both pictures we have N=9 (green), N=12 (blue) and N=15 (yellow). The red curve represents a Gaussian distribution with σ=0.00128.

We point out that the distribution computed with the path integral is different from the one associated to the action.

$$\int D \log(S) e^{-\frac{1}{2\sigma^p} \int_0^T |\partial_t \log(S_t) - r_t|^p dt} \neq e^{-\frac{1}{2\sigma^p T} |\log(S_T) - \log(S_0)|^p} ,$$

see Figure 4.
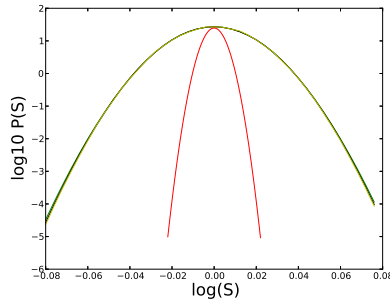
Figure 3: Results with $\gamma = 1$, $p = 1.7$ and $\sigma = 0.004$, $\Delta \log S = 0.16$ with N=9 (green), N=12 (black) and N=15 (yellow). The red curve represents a Gaussian distribution with σ=0.00128.

11
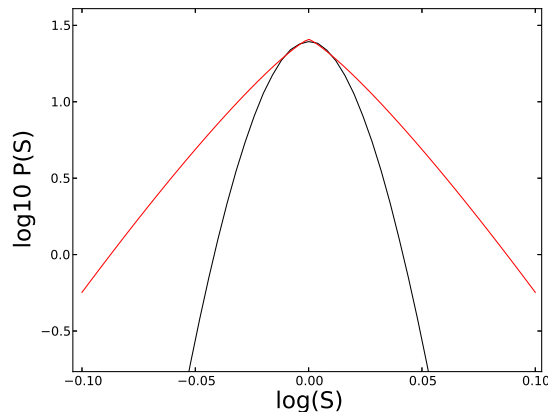
Figure 4: Comparison between the graph in Figure 2 right (black), and the pdf associated to the action (red)

The case with $\gamma \neq 1$ needs some further generalization of the definition of the path integral. With the same approach, we found that a well posed definition is

$$P(S(T)|S(0)) = \int D\log(S)e^{-\frac{1}{2\sigma^2}\left(\int_0^T |\partial_t \log(S_t)-r_t|^p dt\right)^\gamma}$$

$$\approx \lim_{N\to\infty} \prod_{i=1}^{N-1} \int_0^\infty \frac{dS_i}{S_i} \exp\left[-\frac{1}{2\sigma^p}\left(\sum_{i=0}^{N-1} \left|\frac{\log(S_{i+1})-\log(S_i)}{T/N} - r\right|^p (T/N)^{f(p,\gamma)}\right)^\gamma\right],$$

(17)

with

$$f(p,\gamma) = p - (p/2)^\gamma/\gamma.$$

Figures 5 and 6 show the results of the estimates of the integral for some values of (γ, p, σ). The relative errors are negligible when $N \leq 12$. When $N = 13, 14, 15$ it is respectively 5%, 8% and 10%.

Figures 6 and 5 show the results of our tests concerning the well posedness of (17). The probability density functions are studied over a range of more than 4 orders of magnitude, that is a wider range than what we use in the comparison with real data. In order to estimate the convergence, for a given integer M we compute the approximations with $N = M$ and $N = M+1$; then we compute the average of the differences of the probabilities of the highest and lowest price in the two cases. The results for different choices of M are shown in Figure 7 along with the relative errors. Clearly, we only have numerical evidence that our definition of non Gaussian path integral is well posed, more precisely that the limits in (16) and (17) exist, are finite, and they are consistent with the standard Gaussian case.
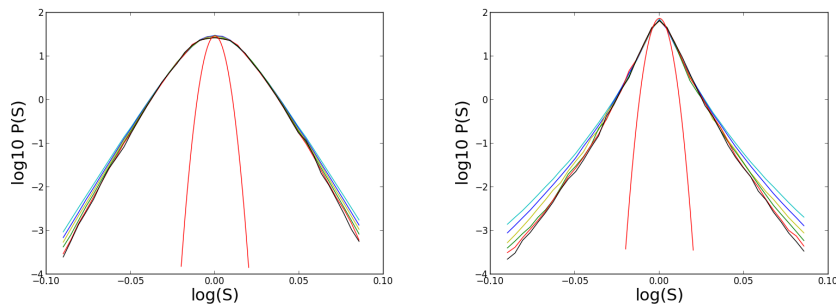
Figure 5: Left: results with $\gamma = 0.5$, $p = 1.5$ and $\sigma = 0.015$. Right: results with $\gamma = 0.2$, $p = 1.3$ and $\sigma = 0.014$. The dimensions are: N=10 (azure), N=11 (blue), N=12 (yellow), N=13 (green), N=14 (red) and N=15 (black).
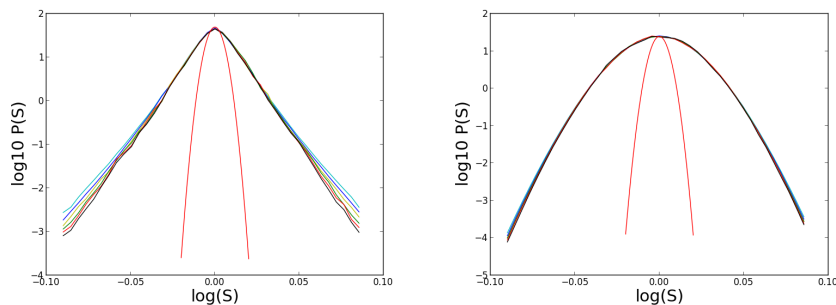

Figure 6: Left: results with $\gamma = 0.29$, $p = 1.275$ and $\sigma = 0.013$. Right: results with $\gamma = 0.8$, $p = 1.7$ and $\sigma = 0.0085$. The dimensions are: N=10 (azure), N=11 (blue), N=12 (yellow), N=13 (green), N=14 (red) and N=15 (black).

5 Numerical Methods

To compute the conditional probability $P(S(T)|S(0))$ we need to estimate multi-dimensional integrals.

The crude Monte Carlo (CMC) method is the simplest to implement, but it is very slow. Markov-Chain Monte Carlo (MCMC) methods are faster, but they require a separate computation to normalize the pdf, thus making this approach less efficient for our purposes.

We choose instead the Quasi Monte Carlo (QMC) method. The main difference between the QMC and CMC consists in the disposition of the generated points. The QMC algorithm can generate a pattern of points which are distributed more evenly than the plain CMC. The better distribution of the points in the QMC yields a convergence rate of the integral of order $\log(N)^D/N$, where N is the number of points and D the dimension of the in-
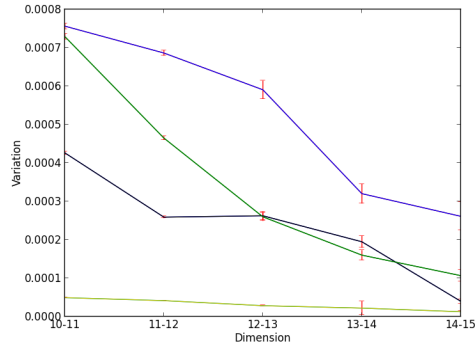
Figure 7: The black and green lines refer to the PDFs in Figure 5. The blue and yellow lines refer to the PDFs in Figure 6.

tegral, while the CMC has order $N^{-1/2}$. Clearly, the QMC method a better choice when D is small, while CMC performs better in higher dimensions. Due to this differences all the simulations shown in section 6 are performed with a QMC methods; whereas the simulation shown in section 4, which are characterized by higher dimensions, are carried on with the usual CMC methods.

There are many algorithms to generate a set of QMC points. We choose Sobol, because it ensures the same homogeneity of the points in all the directions of the domain.

The QMC algorithm is deterministic method are characterized by a fixed disposition; as opposed with the quasi-random points generated by the CMC algorithm, therefore errors cannot be estimated by repeating the algorithm. Instead, to obtain different estimates of the integral we shift the points by a random value and then repeat the computation. Details about the Sobol algorithm and the random shift method to estimate the error can be found in [30].

We estimated all the probability density functions with 40 final prices. On average the computation required a few minutes in the cases treated with the QMC, using on a laptop with a current dual core processor. We remark that the algorithm is fully parallelizable; in particular the computation times on a GPU may be much smaller.

6 Results

We present here some comparisons between real data and our model. The real data consist in 3 months price-sheets of the *AMAZON*, *GENERAL ELECTRIC* and *APPLE* stocks. To obtain the approximate probability density function associated to the stock, we follow the method introduced

14

in the reference [19], i.e. we consider a set of historical data as a set of instances of a stochastic variable. We first compute X_i by

$$X_i = \log(P(t_i)/P(t_{i-1})), \qquad t_i - t_{i-1} = \tau,$$

then we build a histogram of the values X_i with N bins. The histograms shown in the figures below display the number of counts ΔC in each histogram bin, divided by the bin width $\Delta S/N$. The result is then normalized.

The error bars for the real data are estimated as $\sigma_{bins}\Delta S/N$; where ΔS is the width of the histogram x-bars. In order to compare our computations with the results obtained in [2, 3, 14], we plot the probability density functions in logarithmic scale. The error associated to the numerical computation has been estimated and it is omitted since it is negligible with respect to the error over the data. The numerical computations are performed with the previous techniques approximating the path integrals with a 10 dimensional integral.

6.1 Short-term dynamics

We present here the results for three stocks, *AMAZON*, *GENERAL ELECTRIC* and *APPLE*.

The data used are the price-sheet from September 20, 2017 until December 20, 2017, with sample frequency 1, 5 and 30 minutes. In this section $r = 0$, unless otherwise stated.

6.1.1 1 minute

In this section we present the results for the one minute dynamics. The Figures 8, 9 and 10 show in blue the probability density function derived from the data, while the black line is the probability density function computed with the model presented above. The dataset consists of 30000 prices.

6.1.2 5 minutes

The five minutes dynamics is presented in the Figures 11, 12 and 13. The dataset consists of 6000 prices.

The values of the parameters γ and p is the same for the three stocks. We also note that the shared parameters have a higher values with respect to the one minute dynamics.

6.1.3 30 minutes

Finally, we present the final results for the half-hour time frame. Figures 14, 15 and 16 show the results of our computations, together with the probability density functions computed from the real data. In this case the dataset
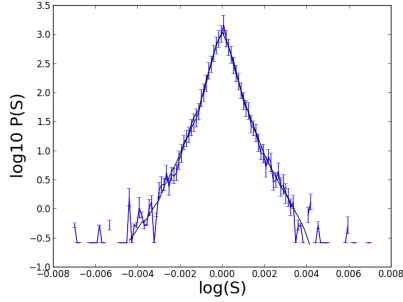
Figure 8: AMAZON stock, $\sigma = 0.035$, $p = 1.15$, $\gamma = 0.15$, $\Delta \log(S) = 0.011$, $N = 120$ and $T = 1m$.
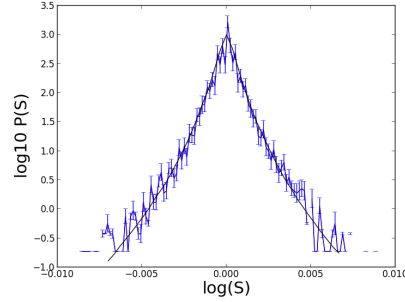

Figure 9: GENERAL ELECTRIC Stocks $\sigma = 0.0326$, $p = 1.15$, $\gamma = 0.15$, $\Delta \log(S) = 0.014$, $N = 120$ and $T = 1m$.
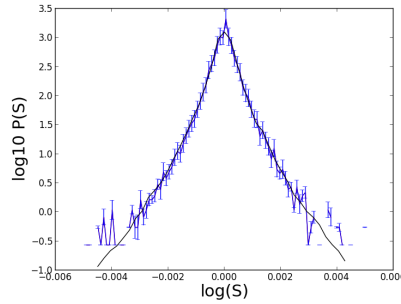

Figure 10: APPLE Stocks $\sigma = 0.0321$, $p = 1.15$, $\gamma = 0.15$, $\Delta \log(S) = 0.009$, $N = 120$ and $T = 1m$.

contains 1000 prices; in order to maintain the number of prices per bin close to the other cases: N is between 11 and 17.

6.2   The analysis of the short-term dynamics

In the previous section we have analysed the dynamics of 3 similar stocks for market capitalization, in 3 different time horizons. The results of the computations shown above are consistent with the theory. The values of γ and p converge to the values associated to the GBM, when the time horizon is increased. It is also interesting to note that the values of the previous parameters are shared among all the stocks analysed for each time horizon treated. This last fact seems to suggest that the three stocks share a common dynamics.
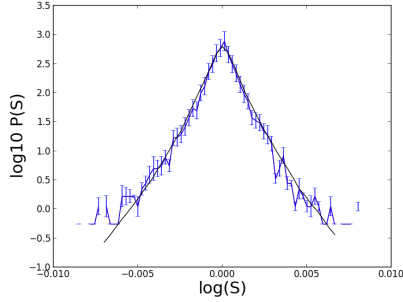
Figure 11: AMAZON stock, $\sigma = 0.0119$, $p = 1.2$, $\gamma = 0.2$, $\Delta \log(S) = 0.016$, $N = 75$ and $T = 5m$.
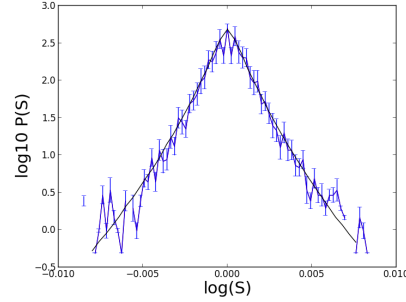

Figure 12: GENERAL ELEC-TRIC Stocks $\sigma = 0.0111$, $p = 1.2$, $\gamma = 0.2$, $\Delta \log(S) = 0.014$, $N = 77$ and $T = 5m$.
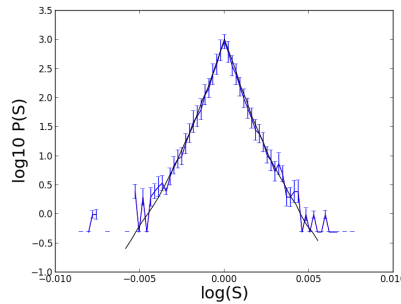

Figure 13: APPLE Stocks $\sigma = 0.0107$, $p = 1.2$, $\gamma = 0.2$, $\Delta \log(S) = 0.011$, $N = 75$ and $T = 5m$.

6.3 Long-term dynamics

We present here the results with time intervals of one day and one week. In order to have enough data to approximate properly the probability density function, we need a consistent extension of the time horizon, which now is 30 years. We select data from the February 01, 1988 until February 01, 2018.

Considering such time intervals, it is better to study the dynamics of indices, since stock prices may be affected by capital increases, splits and other events which would require more care in the selection of the data. We select the *Dow Jones* and the *S&P500*.

6.3.1   1 day dynamics

The one day dynamics is presented in Figures 17 and 18. The dimension of our dataset is approximatively of 7000 prices, where the price selected is price at the closing time.
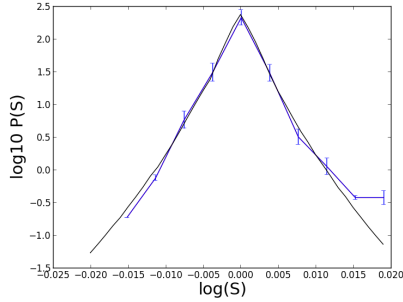
Figure 14: AMAZON stock, $\sigma = 0.00945$, $p = 1.23$, $\gamma = 0.23$, $\Delta \log(S) = 0.034$, $N = 11$, $r = -0.0001$ and $T = 30m$.
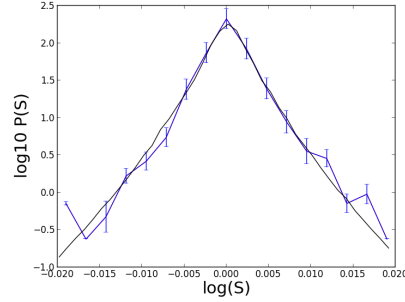

Figure 15: GENERAL ELEC-TRIC Stocks $\sigma = 0.0099$, $p = 1.23$, $\gamma = 0.23$, $\Delta \log(S) = 0.04$, $N = 17$, $r = 0.0002$ and $T = 30m$.
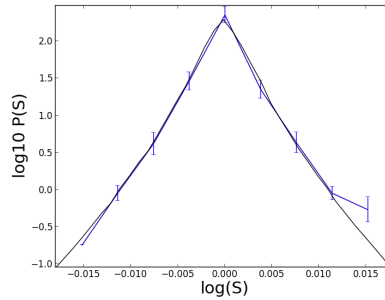

Figure 16: APPLE Stocks $\sigma = 0.0094$, $p = 1.23$, $\gamma = 0.23$, $\Delta \log(S) = 0.034$, $N = 11$, $r = -0.00015$ and $T = 30m$.

6.3.2 The week dynamics

The results for the weekly dynamics is presented in the Figures 19 and 20, the data-set contains almost 1100 prices, selected at the Friday closing time. The histograms of these computations are characterized by a lower number of prices per bin. Since we are considering prices with one week sampling frequency, the variance of each bin is very high, therefore we need a larger N to maintain low data errors

6.4 Final analysis

The results with a time horizon of one minute show a better agreement with the data than [2, 3, 14]. Our results are similar to those in [9]; but we have 3 parameters instead of 5.

It is also interesting to note that the same model is capable to provide a good agreement with 5 different time horizons in both cases of indexes and
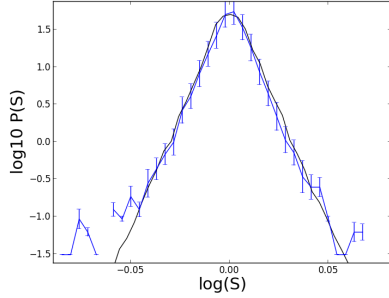
18

Figure 17: Dow Jones Index, $\sigma = 0.062$, $p = 1.35$ and $\gamma = 0.35$, $\Delta \log(S) = 0.12$, $N = 70$, $r = 0.0005$ and $T = 1d$.
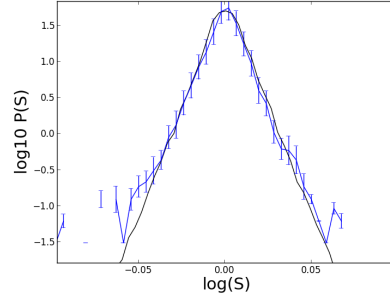

Figure 18: S&P500 Index, $\sigma = 0.062$, $p = 1.35$ and $\gamma = 0.35$, $\Delta \log(S) = 0.12$, $N = 70$, $r = 0.0005$ and $T = 1d$.

stocks.

The parameter values used in the previous computations are consistent with the financial interpretation. We note that the values of (γ, p) both grow with the enlargement of the time horizon; this relation is justified by the theoretical background of the theory presented in section 3. In all the cases, different stocks and indexes share the same value of the parameters γ and p. The difference associated to the liquidity is expressed by a different value of σ. This fact seems to suggest that, given a time frame, there is a common dynamics for all the cases.

7 The Chapman-Kolmogorov equation

The Geometric Brownian Motion is a Markov process. This means that the probability density function obtained from the path integral formulation satisfies the gauge-invariant semigroup equation

$$P(S(T)|S(0))_{GBM} = \int D \log(S_K) P(S(T)|S(K))_{GBM} P(S(K)|S(0))_{GBM} , \tag{18}$$

where $K \in (0, T)$.

The previous equation is also known as the Chapman-Kolmogorov equation and it is a general property of Markovian processes. A straightforward calculation shows that (18) is satisfied if and only if $\gamma = 1$. In [28, Sec. 20.1.18], it is stated that the semigroup property is reasonably satisfied for time frames higher than 15 minutes only outside the deep tails of rare events. Also, the model proposed for the one minute dynamics, that is the Boltzmann distribution, satisfies the semigroup equation, but it does not achieve a good agreement in the deep tails region, see [28, Figure 20.10]. The model in (14) behaves in the same manner; Figure 21 compares the
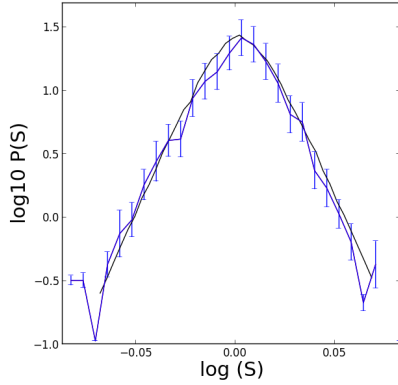
Figure 19: Down Jones Index, $\sigma = 0.023$, $p = 1.42$ and $\gamma = 0.42$, $\Delta \log(S) = 0.14$, $T = 7d$, $r = 0.002$ and $N = 50$.
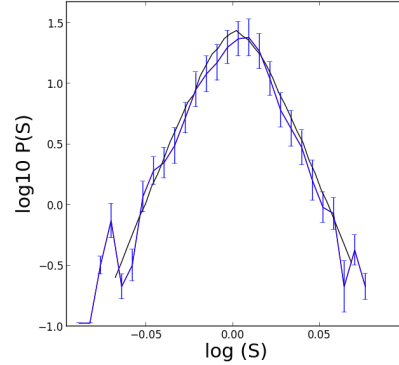
Figure 20: S&P Index $\sigma = 0.023$, $p = 1.42$ and $\gamma = 0.42$, $\Delta \log(S) = 0.14$, $T = 7d$, $r = 0.002$ and $N = 50$.

results of the model with $\gamma = 1$ with the *GE* data shown in Figure 12. The agreement is satisfactory only in the central region.
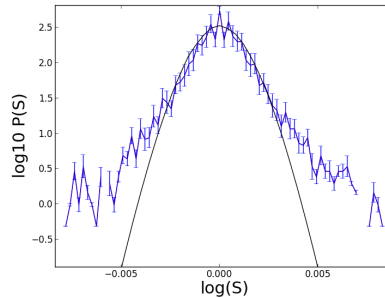

Figure 21: Output of the algorithm with $\gamma = 1$, $p = 1.2$ and $\sigma = 0.00037$.

The numerical computation shown in Section 6 are based on a model which does not satisfies equation (18); indeed this model shows a good agreement also in the deep tails region.

We observe that the computation with $\gamma \simeq 1$, see Figures 2 (left),3, 5 (left),6 (right) have good overlap through all the finite dimensional approximations. The computations with $\gamma < 0.4$, see Figures 5 (right) and 6 (left), have a good overlap only in the higher dimensional approximations; and the differences between the extreme prices are higher with respect the cases with γ closer to 1. This suggests that the parameter γ plays an active role in the speed of convergence for the finite dimensional approximations.

8 Conclusion

In this paper we generalize the model proposed by Ilinski by using a Lagrangian which is not a function of the arbitrage. This generalization allows us to define a new model, where the main dynamics is non-Gaussian. The main advantage of this model consists in the low number of parameters with respect to other models with similar precision. More precisely, our model is characterized by 3 parameters, whereas the models described in [14],[9] and the Heston model have 5 parameters.

The application of a non-Gaussian dynamics to financial problems has already been proposed in [22, 23, 24, 28]. Our model is also characterized by the fact that it does not satisfy the semigroup property, i.e. the stochastic process described by the path integral, is not Markovian. Path integrals have already been used in order to solve stochastic differential equation with Non-Markovian noise, see e.g. [25, 26, 27]. Our work adds to the vast literature in this field and proposes a new conjecture for the computation of a particular form of path integrals. Our result are in agreement with the financial data analysed and with the theoretical background proposed.

Some interesting questions remain open. We analyse the prices of 3 stocks and 2 indexes with different time horizons, but we do not study the auto-aggregation phenomena because our dataset is too small for this kind of analysis. The study of a connection between the auto-aggregation constant α and the parameters (γ, p) is in our opinion worth investigating. Another interesting point may be the fact that the optimal value of γ seems to be $p - 1$.
Finally, we think that it would be important to study the connection between the parameter $\gamma \neq 1$ associated to the presence of memory and the relative implications with the convergence of the finite dimensional approximations.

References

[1] K. Ilinski, *Physics of Finance,* `http://lanl.arxiv.org/abs/hep-th/9710148`

[2] K. Ilinski, Physics of Finance: Gauge Modelling in Non-equilibrium Pricing, Wiley 2001

[3] R. N. Mantegna and H. E. Stanley, *Scaling Behaviour in the Dynamics of Economic Index*, Letters to Nature 06/07/1995, 376, 46-49

[4] R. N. Mantegna and H. E. Stanley, An Introduction to Econophysics , Cambridge University Press

[5] B.E. Baaquie, Quantum Finance Path Integrals and Hamiltonians for Options and Interest rates, Cambridge University Press, 2004.

[6] B.E. Baaquie, Interest Rates and Coupon Bonds in Quantum Finance, Cambridge University Press, 2009.

[7] G. Montagna, O. Nicrosini, N. Moreni, *A path integral way to option pricing*, Physica A 310 (2002) 450-466

[8] G. Montagna, M. Morelli, O. Nicrosini, P. Amato, M. Farina *Pricing derivatives by path integral and neural networks*, Physica A 324, 189-195, (2003)

[9] G. Paolinelli, G.Arioli, *A path integral based model for stocks and orders dynamics*, Physica A 510, 387-399, (2018)

[10] G Bormetti, G Montagna, N Moreni, O Nicrosini, *Pricing Exotic Options in a Path Integral Approach*, Quantitative Finance 6, 1, 55-66 (2006)

[11] E. Bennati, M. Rosa-Clot and S. Taddei, *A Path Integral Approach to Derivative Security Pricing: I. Formalism and Analytical Results*, Int. Journ. Theor. Appl. Finance 2 (1999) 381

[12] V. Linetsky, *The Path Integral Approach to Financial Modeling and Options Pricing*, Computational Economics 11 (1997) 129

[13] J.P.A. Devreese, D. Lemmens, J. Tempere, *Path integral approach to Asian options in the Black-Scholes model*, Physica A 389, 780-788 (2010)

[14] B. Dupoyet, H.R. Fiebig, D.P. Musgrove, *Gauge invariant lattice quantum field theory: Implications for statistical properties of high frequency financial markets*, Physica A 389 (2010) 107-116.

[15] B. Dupoyet, H.R. Fiebig, D.P. Musgrove, *Replicating financial market dynamics with a simple self-organized critical lattice model*, Physica A 390 (2011) 3120-3135.

[16] B. Dupoyet, H.R. Fiebig, D.P. Musgrove, *Arbitrage-free self-organizing markets with GARCH properties: Generating them in the lab with a lattice model*, Physica A 391 (2012) 4350-4363.

[17] P. Bidarkota, B. Dupoyet, *The impact of fat tails on equilibrium rates of return and term premia* , Journal of Economic Dynamics and Control 31 (2007) 887-905

[18] Samuel E. Vázquez, Simone Farinelli, *Gauge invariance, geometry and arbitrage*, The Journal of Investment Strategies 1 (2012)

[19] L.Bachelier, *Théorie de la Spéculation*, Annales Scientifiques de l'École Normale Supérieure 3 (17) (1900) 21-86.

[20] R. P. Feynman, *Space-time approach to non-relativistic quantum mechanics*, Rev. Mod. Phys. 20, 367-387 (1948)

[21] R. P. Feynman, Albert R. Hibbs, Quantum Mechanics and Path Integrals, Dover, 2010.

[22] H. Kleinert, *Option pricing for non-Gaussian price fluctuations*, Physica A 338, 151-159, (2004)

[23] H. Kleinert, *Option pricing from path integral for non-Gaussian Fluctuations. Natural martingale and application to truncated Lévy distributions*, Physica A 312 (2002) 217 - 242

[24] H. Kleinert, *Stochastic calculus for assets with non-Gaussian price Fluctuations* Physica A 311 (2002) 536-562

[25] A. J. McKane, H.C. Luckock, and A. J. Bray *Path integrals and non-Markov processes. I. General formalism* , Physical Review A Vol 41 Number 2, (1990)

[26] A. J. McKane, H.C. Luckock, and A. J. Bray *Path integrals and non-Markov processes. II. Escape rates and stationary distribution in the weak-noise limit*, Physical Review A Vol 41 Number 2, (1990)

[27] A. J. McKane, H.C. Luckock, and A. J. Bray *Path integrals and non-Markov processes. III. Calculation of the escape-rate pre-factor in the weak-noise limit* , Physical Review A Vol 41 Number 4, (1990)

[28] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, 5th Edition (2009)*, World Scientific Publication.

[29] Adrian A. Dragulescu and Victor M. Yakovenko ,*Probability distribution of returns in the Heston model with stochastic volatility*, Quantitative Finance Volume 2, (2002) - Issue 6.

[30] Paul Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer