A welcome to the jungle of continuous-time multivariate non-Gaussian models based on Lévy processes applied to finance

A welcome to the jungle of continuous-time multivariate non-Gaussian models based on Lévy processes applied to finance

Michele Leonardo Bianchi[a,*], Asmerilda Hitaj[b], Gian Luca Tassinari[c]

[a] *Financial Stability Directorate, Bank of Italy,*
micheleleonardo.bianchi@bancaditalia.it
[*] Corresponding author

[b] *Department of Economics, University of Insubria,*
asmerilda.hitaj@uninsubria.it

[c] *Department of Economics, Department of Management, University of Bologna,*
gianluca.tassinari2@unibo.it

This version: June 19, 2022

Abstract. In this paper we review the large and growing literature on continuous-time multivariate non-Gaussian models based on Lévy processes applied to finance and proposed in the literature in the last years. We explain the empirical motivation and the idea behind each approach. Then, we study the models focusing on the parsimony of the number of parameters, the properties of the dependence structure, and the computational tractability. For each parametric class we analyze the main features, we provide the characteristic function, the marginal moments up to order four, the covariances and the correlations. Furthermore, we survey the methods proposed in literature to calibrate these models on the time-series of log-returns, with a view toward practical applications and possible numerical issues. Finally, to empirically assess the differences between models, we conduct an analysis on a five-dimensional series of stock index log-returns.

Keywords: multivariate non-Gaussian processes, moments matching, two-step procedure, expectation-maximization maximum likelihood, generalized method of moments.

1 Introduction

Many problems of practical interest in finance, such as portfolio selection, multi-asset derivative pricing, or the estimation of systemic risk measures are multi-dimensional. The multivariate normal model is usually applied to solve these problems, mainly because both the theoretical and practical complexity of a model increases as soon as one moves from a Gaussian to a non-Gaussian framework. As observed by Roncalli (2020) in discussing default correlation models, only classical models (i.e. Gaussian and Student-t copulas) are really used by professionals. As far as we know, even if an extensive academic literature on these topics is available, multivariate non-Gaussian models are usually put into practice only by more sophisticated investment firms and hedge funds. It is well-known that the multivariate normal distribution has two main drawbacks: (1) its margins are normally distributed and, therefore, it is not capable of capturing heavy tails and asymmetries; (2) its dependence structure is symmetric and, mostly, it is not able of generating asymmetry of dependence during extreme market movements and the tails margins are asymptotically independent.

It is notorious that the historical distribution of asset returns is asymmetric and heavy-tailed. In particular, stock log-returns typically have a distribution characterized by a negative skewness and a positive excess of kurtosis. Furthermore, asset price dynamics exhibits spikes and jumps. These features are incompatible with the Gaussian distribution hypothesis for asset log-returns and the geometric Brownian motion assumption for asset price dynamics. A geometric Brownian motion presents continuous trajectories not able to generate large and sudden price changes as those often found in stock markets. The smile phenomenon observed in option prices provides further empirical evidence against normality: implied volatilities depend on both the strike and the option maturity. On the contrary, the geometric Brownian motion hypothesis implies a constant implied volatility. Further empirical evidence for the presence of jumps in the price process of the underlying is provided by the smile observed on short-term options.

Bedendo et al. (2010) showed that in calm market conditions the choice of the copula function used to model the dependence structure between the underlying assets does not significantly affect multi-asset option prices. Conversely, in volatile market scenarios both linear correlation and tail dependence strengthen and the specification of the dependence structure becomes much more relevant. For example, the dependence between the components of a portfolio and the components of a basket have an impact on portfolio risk measurement, in derivative pricing, or in the systemic risk estimation. Thus, multivariate models able to capture also non-linear dependence among margins are necessary.

In light of the above, in this paper we review different multivariate non-Gaussian models based on Lévy processes to capture the dynamics of log-returns processes. We note that all the necessary technical background on Lévy processes can be found in the following texts Applebaum (2009), Bertoin (1996) and Sato (1999). Lévy processes with jumps are capable of producing trajectories that are consistent with the discontinuous large and sudden movements observed on asset prices and infinitely divisible return distributions with heavy tails and skewness. Furthermore, exponential Lévy option pricing models with jumps can lead to different smile patterns and to high skews for short maturity options. Additionally, multivariate exponential Lévy models for asset prices are able to generate both linear and non-linear dependence. For all analyzed models we look

for consistent and computationally efficient estimation procedures.

A multivariate model should have the following properties: (a) the density function can be written in closed form or in quasi-closed form (i.e. evaluated through an efficient and well-known numerical method); (b) the characteristic function has a closed form allowing one to explore the properties of the model or the change of measure needed to price derivative contracts; (c) the distribution of a linear combination of the margins can be easily obtained: in finance, the knowledge of the distribution of portfolio returns (i.e. the weighted average of the individual asset returns) can be very useful to evaluate portfolio risk measures and to solve portfolio optimization problems; (d) the model is able to explain four stylized facts about financial time-series, that is heavy tails, negative skewness, asymmetric dependence, and volatility clustering (see Allen and Satchell (2014) and Bianchi et al. (2016)); (e) the model can be extended to price derivatives or, at least, there is an efficient method to draw random samples from the model; (f) the number of parameters with respect to the multivariate normal does not increase too much (possibly linearly) by the number of margins, that is the model should be flexible enough but it should not be overparameterized; (g) there is at least one robust estimation algorithm and, ideally, there exists a package written in some commonly used programming language allowing one to perform the estimation.

It should be noted that in this paper we review only continuous-time models based on Lévy processes that usually are not capable of capturing neither the volatility clustering effect nor the leverage effect. The volatility clustering effect is the tendency of large changes in asset prices (either positive or negative) to be followed by large changes, and small changes to be followed by small changes. The leverage effect is the empirically observed fact that negative shocks have a stronger impact on the variance than positive shocks of the same magnitude (i.e. bad news raises the future volatility more than good news). From a practical perspective, it is possible to add volatility clustering dynamics by preliminary filtering the log-return data through a GARCH model and subsequently by calibrating the multivariate models on the standardized log-returns, as done for example in Bianchi et al. (2016).

The empirical exercise conducted in this paper does not want to be a definitive or fully comprehensive analysis on the models discussed in this paper. It is clear that it can be improved under multiple perspectives, for example by considering a GARCH model to filter the data, by comparing the models under a risk management perspective or by studying a larger sample of stocks (see Bianchi and Tassinari (2020)). What we want to do here is a step back before beginning these more numerical intensive and challenging empirical studies. We would like to shed light on some theoretical, technical and practical aspects researchers and practitioners may face when they start implementing these models. This work is something between a review paper and an advanced tutorial, or to be even more precise, it is a welcome to the jungle of multivariate non-Gaussian models based on Lévy processes applied to finance.

A multivariate non-Gaussian Lévy model can be built by following different approaches. It is possible (1) to consider a linear combination of independent Lévy processes, as done for example by Kawai (2009), Kaishev (2013), Ballotta and Bonfiglioli (2016) and Ballotta et al. (2019); (2) to time-change a multivariate Lévy process with a univariate or a multivariate subordinator (e.g. Barndorff-Nielsen et al. (2001), Luciano and Semeraro (2010b), Hitaj et al. (2018), and Semeraro (2019)); (3) to define its Lévy

measure, as done by Rosinski (2007) and Bianchi et al. (2011); (4) to specify the univariate models separately from the dependence structure, i.e. the Lévy copula, to form a multivariate model (see Tankov (2016) for a review). As shown in Kallsen and Tankov (2006), similarly to copulas for probability measures, Lévy copulas represent a flexible approach for building multidimensional Lévy processes. More precisely, a Lévy copula gives a representation of the Lévy measure of a multivariate Lévy process, which allows to specify separately the Lévy measures of the marginal processes and their dependence structure. In this paper we consider only the first two approaches. Even if the third and fourth approaches are elegant from a theoretical perspective, they are not simple to use in practical applications. Furthermore, models based on Lévy copulas can be estimated when jumps are observable. This is the case, for instance, in insurance models where jumps represent claims whose timing and amounts are known (see Tankov (2016)). However, this is not the case in most financial applications, unless some simplifying restrictive assumptions are adopted.

In this work we discuss and empirically assess the fitting performance of several continuous-time multivariate heavy-tailed and semi heavy-tailed models applied to finance and proposed in the literature in the last years. It is clear that there is a strict connection between the multivariate non-Gaussian distributions reviewed in this paper and continuous-time processes. In all cases it is always possible to define a multivariate Lévy process whose increments are distributed as the multivariate random variables described in this work and representing the multivariate time-series of log-returns.

For each model we analyze the main properties and the more useful formulas, that is: the characteristic function, the probability density function (in the case it can be written in closed form), the marginal moments up to order four, covariances and correlations. After having described how to calibrate these models, analyzed the computational tractability and possible numerical issues, we empirically compare them on a real market dataset on the basis of some fitting error measures. Just for comparison purposes, we calibrate also the multivariate normal model.

It is interesting to observe that there are several goodness-of-fit tests for the assessment of univariate models. For example there is the well-known Kolmogorov-Smirnov test in which the distance between the theoretical and the empirical cumulative distribution function is considered. However, there are not leading approaches to assess the goodness-of-fit of multivariate models. Considering the existing literature on this topic, in order to compare the fitting performance of the multivariate models under analysis we also use the distance between historical and model correlations and the nonparametric distance between observed and simulated data defined by Li et al. (2009). We remind that in this paper we do not compare the models under a risk management perspective, that is we do not backtest them by considering the performance of portfolio allocation strategies or the ability to produce good risk measures. As observed by Bossu (2014) and Meissner (2014), correlation is almost as ubiquitous as volatility in quantitative finance and it is critical in many areas such as investments, trading, and risk management, as well as in financial crises and in financial regulation. For this reason it seems important to be able to fit observed historical correlations together with marginal moments. However, it should be noted that it is not obvious how the in-sample fit of marginal moments and correlations affects the out-of-sample performance in portfolio allocation or in risk modeling. This analysis goes beyond the aim of this paper and is left for future reasearch.

The paper is organized as follows. In Section 2 we review the normal mean-variance mixture models, where the mixing variable has a semi-heavy tailed distribution, and in Section 3 we consider extensions based on a multivariate mixing distribution referred to as one factor subordinated models. Both models belong to the class of time-changed Brownian motions: while the former has a quasi-closed formula for the joint density function, the latter does not but it may have a greater flexibility in fitting market data thanks to its richer dependence structure. A multivariate model based on tempered stable distribution and its major capability to fit the margins is discussed in Section 4. A further extension is proposed in Section 5 where a more flexible multifactorial model is considered. This model is able to capture dependence separately and independently both in positive and negative jumps and in their finite and infinite activity components. Multivariate models based on linear combinations of Lévy processes are presented in Section 6. Differences in estimation methods are described in Section 7. In Section 8 we describe the data analyzed in the empirical study, we discuss the main empirical results and we identify some computational issues. After having summarized the main results, Section 9 concludes. Finally, in the Appendix we provide the formulas of expected value, variance, skewness, excess of kurtosis and correlation for each model described in the paper. These formulas are generally useful in practical applications to finance.

2 Normal mean-variance mixture models

As observed by Arellano-Valle and Azzalini (2020), in the last few decades, a number of models have been proposed where a multivariate normal variable represents the basic constituent but with the superposition of another random component, either in the sense that the normal mean value or the variance-covariance matrix or both these components are subject to the effect of another random variable of continuous type. The multivariate non-normal distributions analyzed in this section are based on this idea and they are known as multivariate normal mean-variance (or variance-mean) mixtures. These distributions share much of the structure of the multivariate normal distribution, but they allow asymmetry, heavy tails and both linear and non-linear dependence.

In particular, a random vector Y has a multivariate normal mean-variance mixture distribution (NMV) if the following equality in law holds

$$Y = \mu + \theta S + \sqrt{S} Q Z, \tag{2.1}$$

where μ, $\theta \in \mathbb{R}^n$, Q is a square matrix of order n such that QQ' is positive definite, S is a positive random variable, $Z \sim N(0, I_n)$ and S is independent from Z.
Furthermore, if the mixing variable S is infinitely divisible then Y is infinitely divisible and its law uniquely determines a time-changed Brownian motion whose subordinator at time one has the law of S (see Barndorff-Nielsen et al. (2001) and Tassinari and Bianchi (2014)). The time-changed Brownian motion construction is well known from the theory of stochastic processes. In the theoretical work of Monroe (1978), it is shown under what conditions a stochastic process is equivalent to a time-changed Brownian motion. In particular, Monroe showed that any semimartingale can be written as a time-changed Brownian motion. The application to finance dates back to Clark (1973) and it is based on the following evidence: "The different evolution of prices series on different days is due

to the fact that information is available to traders at a varying rate. On days when no new information is available trading is slow, and the price process evolves slowly. On days when new information violates old expectation, trading is brisk, and the price process evolves much faster". In periods of high volatility, time runs faster than in periods of low volatility. The physical time in the Brownian motion is substituted by a non-negative non-decreasing Lévy process, that is a subordinator (stochastic or business time) providing tail effects. By subordination, it is possible to capture empirically observed anomalies that contradict the classical lognormality assumption for asset prices. Madan et al. (1998) introduced the asymmetric version of the variance-gamma process as a model for log-returns by time-changing a Brownian motion with drift with an independent Gamma subordinator and they provided a methodology to explicitly compute European option prices in this context based on the use of hypergeometric functions. Geman et al. (2001) argued that price processes being semimartingales due to the no-arbitrage condition, they are time-changed Brownian motions. They considered pure jump Lévy processes of finite variation with an infinite arrival rate of jumps as models for the logarithm of asset prices, expressed as the difference between two increasing random processes that account for the upward and downward moves of the market, and they represented them as time-changed Brownian motions. Specifically, the time changes considered in their work were represented by Poisson processes, gamma processes, general subordinators, and the inverse local time of Brownian motion at zero.
The time-changed Brownian motion construction can be extended to a multivariate framework. Indeed, it is possible to subordinate a multivariate Brownian motion by a univariate or a multivariate stochastic time (see also Bianchi et al. (2019)).

Therefore, in order to model asset returns it is possible to build an n-dimensional Lévy process whose increments follow an infinitely divisible NMV distribution by simply time-changing a multivariate Brownian motion with a common *one*-dimensional subordinator. While in Luciano and Schoutens (2006) and Tassinari and Corradi (2013) a model with independent Brownian motions was proposed, in Leoni and Schoutens (2008), Tassinari (2009), Wu et al. (2009), Tassinari and Corradi (2014), Tassinari and Bianchi (2014), Bianchi et al. (2016), and Bianchi and Tassinari (2020) correlated Brownian motions were considered. Furthermore, according to Frahm (2004), this family of distributions belongs to the class of elliptical variance-mean mixtures. Elliptical and generalized elliptical heavy-tailed distributions have been widely studied (see e.g. Kring et al. (2009), Dominicy et al. (2013), Bianchi et al. (2019)).

Let $Y = \{Y_t, t \geq 0\}$ be a multivariate process such that the following equality holds

$$Y_t = \mu t + \theta S_t + D_\sigma W_{S_t}, \tag{2.2}$$

where $S = \{S_t, t \geq 0\}$ is a *one*-dimensional subordinator, $W = \{W_t, t \geq 0\}$ is an n-dimensional Wiener process with $corr\,[W_{j,t}, W_{k,t}] = \rho_{jk}$ independent from S, and D_σ is a diagonal matrix with diagonal elements $\sigma_j > 0$ for $j = 1, ..., n$. For each discrete time step Δt the distribution of the increments of the process belongs to the NMV family

$$Y_{\Delta t} = \mu \Delta t + \theta S_{\Delta t} + \sqrt{S_{\Delta t}} D_\sigma A Z,$$

where $S_{\Delta t}$ denotes the distribution of the subordinator increments which is independent from Z, A is the lower Cholesky decomposition of a correlation matrix Ω, that is, $\Omega^{1/2} = A$

and $Q = D_\sigma A = \Sigma^{1/2}$. The characteristic function of Y_t defined in equation (2.2) is given by

$$\Psi_{Y_t}(u) = \exp\left(itu'\mu + tl_{S_1}(\varphi(u))\right), \tag{2.3}$$

where $l_{S_1}(.)$ is the Laplace exponent of the subordinator, and $\varphi(u)$ is the characteristic exponent of the multivariate Brownian motion, that is

$$\begin{aligned} \varphi(u) &= iu'\theta - \frac{1}{2}u'\Sigma u \\ &= \sum_{j=1}^{n} iu_j\theta_j - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} u_j u_k \sigma_j \sigma_k \rho_{jk}, \end{aligned} \tag{2.4}$$

where $u \in \mathbb{R}^n$ and the matrix Σ has elements $\Sigma_{jk} = \sigma_j\sigma_k\rho_{jk}$. Since Σ is a variance-covariance matrix, we can rewrite equation (2.4) using matrix notation, as follows

$$\varphi(u) = iu'\theta - \frac{1}{2}u'D_\sigma\Omega D_\sigma u,$$

where D_σ is a diagonal matrix with diagonal $\sigma \in \mathbb{R}^n_+$, and Ω is the correlation matrix of the Brownian motions with elements ρ_{jk}.

In Sections 2.1 and 2.2 we will describe two parametric models belonging to the NMV class: the multivariate generalized hyperbolic (MGH) and the multivariate normal tempered stable (MNTS) model. Before going into details, we would like to describe the main properties of the two models, to underline the main differences and to discuss some practical aspects.

It is important to note that unlike the MNTS case, if the increments of the process at a given time scale (e.g. daily) follow a MGH distribution, on a different time scale (e.g. yearly) the increments follow an infinitely divisible distribution different from the MGH (see Cont and Tankov (2003)), making the MGH distribution less convenient when one needs to work with data with different time scales. This problem can arise, for example, in options pricing models that make use of both daily returns and implied volatilities (see Tassinari and Bianchi (2014) and Bianchi and Tassinari (2020)).

As shown in McNeil et al. (2005) the portfolio constructed as a linear combination of generalized hyperbolic (GH) margins has a GH distribution. Kim et al. (2012) proved that the same is true in the normal tempered stable (NTS) case. Thus, these models can be easily applied to evaluate widely known portfolio risk measures and to solve asset allocation and portfolio optimization problems (see Bianchi and Tassinari (2020)).

Unlike the MGH random variable, it is not possible to obtain in closed form the probability density function of the MNTS random variable. However, the density of a MNTS random variable can be obtained by a numerical integration that combines the density of a multivariate normal distribution and the density of a univariate tempered stable mixing distribution, which can be evaluated by means of a fast Fourier transform (FFT) (see Stoyanov and Racheva-Iotova (2004) and Bianchi et al. (2017)).

In estimating these models, it is usually not possible to resort to direct maximization of the likelihood function as the number of parameters is large. To overcome this obstacle, in the estimation of the parameters of the multivariate normal mean-variance mixture distributions, the use of the expectation-maximization (EM) maximum likelihood estimation method is particularly convenient as it allows to find the parameters

of the multivariate Gaussian distribution and those of the mixing distribution separately (see Protassov (2004), Hu (2005), and McNeil et al. (2005) in the MGH case, and Bianchi et al. (2016) in the MNTS).

While the parameter estimation of the MGH model is well-known in the literature, the estimation of the MNTS model is more challenging. As proposed by Fallahgoul et al. (2016), the parameters μ and Σ of the model in equation (2) can be estimated using the sample mean vector and sample variance-covariance matrix for stock market returns and the parameters of the subordinator can be obtained by considering the average of the margin estimates. Alternatively, Kim et al. (2012) conducted an empirical analysis on the Dow Jones Industrial Average (DJIA) index and 29 of the 30 component stocks. They estimated the subordinator parameters on the DJIA index returns and the vector θ was estimated on the margins. Finally, both μ and Σ were estimated by considering the sample covariances together with the univariate estimates (see also Kurosaki and Kim (2021) and Kim (2022) for applications to portfolio allocation). However, we do not explore these two methods in our empirical study, since we will rely on the EM estimation approach as it will be described in Section 7.3.

Finally, from a practical perspective, the simulation of random draws from the multivariate distributions defined in equation (2.1) is simple: one only needs an efficient algorithm to simulate the univariate mixing distribution and a standard procedure to generate multivariate normal random samples.

2.1 The multivariate generalized hyperbolic distribution

The GH distribution, introduced by Barndorff-Nielsen (1977), has received a lot of attention in the financial-modeling literature (see Eberlein and Keller (1995), Prause (1999), and Eberlein et al. (2002)). Many well known distributions, like for example the student's t, the skew-t, the variance gamma (VG) and the normal inverse Gaussian (NIG), belong to the GH parametric family. In this section we review the multivariate extension of the GH distribution.

Let $S = \{S_t,\ t \geq 0\}$ be a generalized inverse Gaussian process (GIG), i.e., a Lévy process in which the law of S_1 is generalized inverse Gaussian with parameters ϵ, ψ, χ, where ψ and χ are both nonnegative and not simultaneously 0. Therefore, the number of parameters to be estimated in case of a MGH distribution is $(n^2 + 5n)/2 + 3$.

We denote the law of S_1 as $GIG\,(\epsilon, \chi, \psi)$. The density function of S_1 is

$$f(x; \epsilon, \psi, \chi) = \frac{1}{2K_\epsilon\left(\sqrt{\chi\psi}\right)} \left(\frac{\psi}{\chi}\right)^{\frac{\epsilon}{2}} x^{\epsilon-1} \exp\left(-\frac{1}{2}\left(\frac{\chi}{x} + \psi x\right)\right), x > 0,$$

and its characteristic function is

$$\Psi_{S_1}(u) = \left(1 - \frac{2iu}{\psi}\right)^{-\frac{\epsilon}{2}} \frac{K_\epsilon\left(\sqrt{\chi(\psi - 2iu)}\right)}{K_\epsilon\left(\sqrt{\chi\psi}\right)}. \tag{2.5}$$

If in equation (2.2) we select a subordinator $S = \{S_t, t \geq 0\}$ such that the characteristic function of S_1 is (2.5), then the process $Y = \{Y_t, t \geq 0\}$ is referred to as the MGH process with parameters (ϵ, χ, ψ, θ, μ, Σ).

Finally, using (2.3) we get the characteristic function of the MGH process with linear drift

$$\begin{aligned}\Psi_{Y_t}(u) = &\exp(iu'\mu t)\left(1-\frac{2}{\psi}\left(iu'\theta-\frac{1}{2}u'\Sigma u\right)\right)^{-\frac{\epsilon t}{2}} \\ &\left(\frac{K_\epsilon\left(\sqrt{\chi\left(\psi-2\left(iu'\theta-\frac{1}{2}u'\Sigma u\right)\right)}\right)}{K_\epsilon\left(\sqrt{\chi\psi}\right)}\right)^t.\end{aligned} \tag{2.6}$$

Setting $u_i = 0$, $\forall i \neq j$, into (2.6) we get the characteristic function of the log-return process of the j-th underlying asset

$$\begin{aligned}\Psi_{Y_{j,t}}(u_j) = &\exp(iu_j\mu_j t)\left(1-\frac{2}{\psi}\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)\right)^{-\frac{\epsilon t}{2}} \\ &\left(\frac{K_\epsilon\left(\sqrt{\chi\left(\psi-2\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)\right)}\right)}{K_\epsilon\left(\sqrt{\chi\psi}\right)}\right)^t.\end{aligned} \tag{2.7}$$

Setting $t = 1$ into (2.6) and into (2.7) we get the characteristic function of the MGH and GH distributions.

Comparing the characteristic function of the MGH distribution with the one of Y_t we can notice that the GH distribution is infinitely divisible but not closed under convolution. Thus, if Y_1 is a MGH random variable, Y_t is not. If $\epsilon = -1/2$, G_1 follows an inverse Gaussian distribution with parameters $\gamma = \sqrt{\chi}$ and $\eta = \sqrt{\psi}$. If $\chi = 0$, G_1 follows a gamma distribution $\alpha = \epsilon$ and $\beta = \psi/2$. In the first case we get the multivariate normal inverse Gaussian (MNIG) model and in the second one the multivariate variance gamma (MVG) considered in Tassinari and Bianchi (2014).

2.2 The multivariate normal tempered stable distribution

The tempered stable family was introduced by Boyarchenko and Levendorskii (2000, 2002) and studied in deep by Rosinski (2007). The multivariate model discussed in this section is based on the classical tempered stable (CTS) distribution as described by Kim et al. (2012), Bianchi et al. (2016) and Bianchi and Tassinari (2020) and extends the works of Prause (1999), Leoni and Schoutens (2008) and Wu et al. (2009) to the CTS case. We refer to this model as multivariate normal tempered stable (MNTS). We observe that the elliptical tempered stable (ETS) distribution defined in Fallahgoul et al. (2016) is a subclass of the multivariate symmetric normal tempered stable (MSNTS) distribution, and a symmetric MNTS is a subclass of the tempered infinitely divisible family introduced by Bianchi et al. (2011).

The process $S = \{S_t, t \geq 0\}$ is said to be a CTS subordinator with parameters α, $\lambda > 0$, $C > 0$, $0 < \alpha < 1$ if the characteristic function of S_t is given by

$$\phi_{S_t}(u) = E[\exp(iuS_t)] = \exp\left(tC\Gamma(-\alpha)((\lambda - iu)^\alpha - \lambda^\alpha)\right), \tag{2.8}$$

where $u \in \mathbb{R}$ and $\Gamma(\cdot)$ is the gamma function. From equation (2.8) it is possible to compute the Laplace exponent of the CTS subordinator

$$l_{S_t}(u) = \ln \phi_{S_t}(-iu) = tC\Gamma(-\alpha)((\lambda - u)^\alpha - \lambda^\alpha).$$

If in equation (2.2) we select a subordinator $S = \{S_t, t \geq 0\}$ with characteristic function (2.8), then the process $Y = \{Y_t, t \geq 0\}$ is referred to as the MNTS process with parameters (a, λ, C, θ, μ, Σ). If one selects $\alpha = a/2$ and $C = \lambda^{1-\alpha}/\Gamma(1-\alpha)$, equation (2.8) can be written as

$$\phi_{S_t}(u) = \exp\left(-t\frac{2\lambda^{1-\frac{a}{2}}}{a}\left((\lambda - iu)^{\frac{a}{2}} - \lambda^{\frac{a}{2}}\right)\right),$$

and the multivariate distribution $Y_{\Delta t}$ defined as

$$Y_{\Delta t} = \mu\Delta t + \theta(S_{\Delta t} - \Delta t) + \sqrt{S_{\Delta t}} D_\sigma A Z,$$

is the MNTS distribution analyzed by Kim et al. (2012).

Using (2.3) we get the characteristic function of the MNTS process with linear drift

$$\Psi_{Y_t}(u) = \exp\left(t\left(iu'\mu + C\Gamma\left(-\frac{a}{2}\right)\left(\left(\lambda - iu'\theta + \frac{1}{2}u'\Sigma u\right)^{\frac{a}{2}} - \lambda^{\frac{a}{2}}\right)\right)\right). \tag{2.9}$$

Setting $u_i = 0$, $\forall i \neq j$, into (2.9) we get the characteristic function of the j-th marginal distribution

$$\Psi_{Y_{j,t}}(u_j) = \exp\left(t\left(iu_j\mu_j + C\Gamma\left(-\frac{a}{2}\right)\left(\left(\lambda - iu_j\theta_j + \frac{1}{2}u_j^2\sigma_j^2\right)^{\frac{a}{2}} - \lambda^{\frac{a}{2}}\right)\right)\right). \tag{2.10}$$

In case of MNTS distribution the number of parameters to be estimated is $(n^2 + 5n)/2+3$. If $\omega = 1/2$, S_1 follows an inverse Gaussian distribution with parameters $\gamma = -\frac{C\Gamma(-\omega)}{\sqrt{2}}$ and $\eta = \sqrt{2\lambda}$, and Y_1 follows the MNIG distribution with parameters $\alpha = -\frac{C\Gamma(-\omega)}{\sqrt{2}}$ and $\beta = \sqrt{2\lambda}$, and Y_1 follows the MVG distribution described in Tassinari and Bianchi (2014).

3 Generalized NMV mixture models

A multivariate NMV distribution is based on a common *one*-dimensional mixing variable. This corresponds to a multidimensional return process with a unique stochastic time-change, which implies the uniqueness of the business time for all assets. As shown by Harris (1986) this feature seems to be inconsistent with empirical evidence. A more realistic assumption is that each return has its own change of time, that is each margin has its own mixing variable. The multidimensional mixing distribution is based on the empirical finding that different stocks are affected by both a common and an idiosyncratic component. Semeraro (2010) discussed a generalization of the VG process, Luciano and Semeraro (2010a) proposed a generalization of the definition of NMV distribution based on an n-dimensional mixing variable. The authors had four desired features in mind: (1) the existence of characteristic functions in closed form; (2) the ability to capture a

wide range of linear and non-linear dependence; (3) the possibility of fitting separately the marginal distributions and the correlation parameters; (4) the development of a consistent framework in which assets and portfolio processes belong to the same class. Recently, Rathgeber et al. (2019) conducted a large simulation study on these models in order to identify the best fitting method for multivariate models.

A random vector Y has a multivariate generalized normal mean-variance mixture distribution (GNMV) if the following equality in law holds

$$Y = \mu + MD_G\theta + QD_{\sqrt{G}}Z, \tag{3.1}$$

with μ, $\theta \in \mathbb{R}^n$, M and Q are square matrices of order n, QQ' is positive definite, G is an n dimensional positive random vector whose j-th component is G_j, D_G and $D_{\sqrt{G}}$ are diagonal matrices with diagonal elements G_j and $\sqrt{G_j}$, respectively, and $Z \sim N(0, I_n)$ is independent from G. If the mixing variable G is infinitely divisible then Y is infinitely divisible and its law uniquely determines a Lévy process.

Following Luciano and Semeraro (2010a) the characteristic function of the random variable Y can be written as

$$\Psi_Y(u) = \exp(iu'\mu) \exp\left(l_G\left(iD_\theta M'u - \frac{1}{2}D_{Q'u}Q'u\right)\right),$$

where $l_G(.)$ is the Laplace exponent of the multivariate mixing variable G, D_θ and $D_{Q'u}$ are diagonal matrices with diagonal elements the vectors θ and $Q'u$, respectively. If we set $M = I_n$ and $G_j = S$ for all j in (3.1) we obtain (2.1).

Barndorff-Nielsen et al. (2001) proved that a random vector Y has GNMV distribution if and only if it is the law at time one of a Lévy process obtained by subordination of a $\mathbb{R}^n_+$-parameter Brownian motion with a multidimensional subordinator whose distribution is given by G. Luciano and Semeraro (2010a) and Luciano and Semeraro (2010b) built multivariate Lévy processes with GH, compound Poisson, NIG and VG margins using the multivariate subordination technique. In particular, they proposed two different techniques to build n-dimensional Lévy processes through subordination leading to different multivariate models with the same marginal processes. We refer to these two class of processes as the α-models and the $\rho\alpha$-models. Recently, an extension of the α-model based on the VG distributional assumption and weak-subordination and allowing a wider range of dependence has been proposed and applied to finance (see Buchmann et al. (2017), Michaelsen and Szimayer (2018), Madan (2018) and Buchmann et al. (2019)).

3.1 The αGH distribution

Lo and Wang (2000) provided empirical evidence that business time as measured by trades presents a significant common component. Semeraro (2010), Luciano and Semeraro (2010a) and Luciano and Semeraro (2010b) proposed to build multivariate subordinators able to capture both a time-change common to all assets and an idiosyncratic one. In particular, they used the random additive effect distributions proposed by Barndorff-Nielsen et al. (2001) to get a multivariate stochastic clock $G = \{G_t, t \geq 0\}$ containing both a common and an asset specific time-change:

$$G_t = X_t + \alpha S_t, \tag{3.2}$$

where $X = \{X_t, t \geq 0\}$ is an n-dimensional subordinator with independent components, $S = \{S_t, t \geq 0\}$ is a *one*-dimensional subordinator independent by X, α is a $n \times 1$ vector with positive elements. Let $Y = \{Y_t, t \geq 0\}$ be a multivariate process such that the following equalities hold

$$\begin{aligned} Y_t &= \mu t + Y_t^I \\ &= \mu t + B_{G_t} \\ &= \mu t + D_{G_t}\theta + D_\sigma W_{G_t}, \end{aligned} \tag{3.3}$$

where

- $Y^I = \{Y_t^I, t \geq 0\}$ is constructed by subordinating an n-dimensional arithmetic Brownian motion $B = \{B_t, t \geq 0\}$, where $B_t = \theta t + D_\sigma W_t$, with independent components with the subordinator (3.2);
- $W = \{W_t, t \geq 0\}$ is an n-dimensional Wiener process with $corr\left[W_{j,t}, W_{k,t}\right] = 0$ for $j \neq k$;
- D_σ is a diagonal matrix with diagonal elements $\sigma_j \in \mathbb{R}_+$ for all j,
- $\theta \in \mathbb{R}^n$ is a vector of parameters.

For each discrete time step Δt the distribution of the increments of the process is given by

$$Y_{\Delta t} = \mu \Delta t + D_{G_{\Delta t}}\theta + D_\sigma D_{\sqrt{G_{\Delta t}}} Z,$$

where $G_{\Delta t}$ denotes the distribution of the subordinator increments which is independent of Z. The distribution of $Y_{\Delta t}$ belongs to the class of the GNMV distribution with $M = I_n$ and $Q = D_\sigma$.

The characteristic function of Y_t defined in equation (3.3) is given by

$$\begin{aligned} \Psi_{Y_t}(u) &= \exp\left(itu'\mu\right) \Psi_{Y_t^I}(u) \\ &= \exp\left(itu'\mu\right) \exp\left(t \sum_{j=1}^{n} l_{X_{j,1}}(\psi_j(u_j))\right) \exp\left(t l_{S_1}(\sum_{j=1}^{n} \alpha_j \psi_j(u_j))\right) \\ &= \exp\left(itu'\mu\right) \exp\left(t \sum_{j=1}^{n} l_{X_{j,1}}(iu_j\theta_j - \frac{1}{2}u_j^2\sigma_j^2)\right) \exp\left(t l_{S_1}(\sum_{j=1}^{n} \alpha_j (iu_j\theta_j - \frac{1}{2}u_j^2\sigma_j^2))\right), \end{aligned} \tag{3.4}$$

where $l_{X_{j,1}}(.)$ and $l_{S_1}(.)$ are the Laplace exponents of the subordinators $X_{j,t}$ and S_t, respectively. Choosing $X_{j,1}$, S_1 and α_j opportunely, Luciano and Semeraro (2010a) and Luciano and Semeraro (2010b) proposed different multivariate models with GH, compound Poisson, NIG and VG margins. The number of parameters to be estimated in these multivariate distributions increase linearly with the number of margins. In this paper, we review only the GH model, and we refer to it as αGH model. Following Luciano and Semeraro (2010a) we build a multivariate subordinator $G = \{G_t, t \geq 0\}$ on $\mathbb{R}_+^n$ with dependent GIG margins $G_j = \{G_{j,t}, t \geq 0\}$, $j = 1, ..., n$, by defining

$$G_{j,t} = X_{j,t} + \alpha_j S_t = R_{j,t} + P_{j,t} + \frac{1}{\psi_j} S_t, \tag{3.5}$$

where $R_{j,1}$, $P_{j,1}$ and S_1 are independent with $R_{j,1} \sim GIG(-\epsilon, \chi_j, \psi_j)$, $P_{j,1} \sim \Gamma\left(\epsilon - a, \frac{\psi_j}{2}\right)$, and $S_1 \sim \Gamma\left(a, \frac{1}{2}\right)$. If $\epsilon > 0$, $\psi_j > 0$, $0 < a < \epsilon$, and $\chi_j \geq 0$ for all j, then $G_{j,1} \sim GIG(\epsilon, \chi_j, \psi_j)$, that is all the margins of the multivariate subordinator at time one follow a generalized inverse Gaussian law.
Using (3.4) and setting $t = 1$ we get the characteristic function of the αGH distribution

$$\begin{aligned}\Psi_{Y_1}(u) &= \prod_{j=1}^{n}\left(1 - \frac{2}{\psi_j}\left(iu_j\theta_j - \frac{1}{2}\sigma_j^2 u_j^2\right)\right)^{a-\frac{\epsilon}{2}} \frac{K_\epsilon\left(\sqrt{\chi_j\left(\psi_j - 2\left(iu_j\theta_j - \frac{1}{2}u_j^2\sigma_j^2\right)\right)}\right)}{K_\epsilon\left(\sqrt{\chi_j\psi_j}\right)} \\ &\quad \exp(iu'\mu)\left(1 - \sum_{j=1}^{n}\frac{2}{\psi_j}\left(iu_j\theta_j - \frac{1}{2}\sigma_j^2 u_j^2\right)\right)^{-a}.\end{aligned} \tag{3.6}$$

Setting $u_i = 0$, $\forall i \neq j$, into (3.6) we get the characteristic function (2.7) of the GH law.

Setting $\epsilon = 1$ the marginal processes are hyperbolic and we get the αHYP. If $a \to 0$ the αGH process degenerates into the MGH model with independent univariate GH processes. If $\chi_j \to 0$ for all j, the αGH process degenerates into the αVG. If one sets $\chi_j = \delta_j^2$, $\psi_j = \alpha_j^2 - \beta_j^2$, $\theta_j = \beta_j$, $\mu_j = 0$, $\sigma_j = 1$ for all j, we get the αGH process of Luciano and Semeraro (2010a).

The estimation of this model in Luciano and Semeraro (2010a) was performed in two steps by fixing $\epsilon = 1$. The restrictions on the parameters of the random variables $R_{j,1}$, $P_{j,1}$ and S_1 in the right hand side of (3.5) ensure that the margins are still GH distributed. Guillaume (2013), following Luciano and Semeraro (2010b) and Semeraro (2010), but removing the restrictions on single variable parameters, proposed the generalized αVG model whose margins are no longer VG distributed but still result to be infinitely divisible. The same principle can be followed to generalize the αNIG and the αGH models.

3.2 The $\rho\alpha$GH distribution

The α-models are obtained time changing a multivariate Brownian motion with independent components with the subordinator (3.2). The only source of dependence among different assets is due to the timing of the jumps. The $\rho\alpha$-models extend the α-models allowing the dependence of both time and size of the jumps. The introduction of these models is motivated by the search of a greater accuracy in fitting both marginal distributions and sample correlations. According to Luciano et al. (2016) a model should provide a good fit of the marginal distributions and should be flexible enough to capture correlations and portfolio returns. It should be noted that portfolio returns reflect linear and non-linear dependence. Notwithstanding the greater flexibility of the $\rho\alpha$-models, assets and portfolio distributions belong to the same class.

Let $Y = \{Y_t, t \geq 0\}$ be a multivariate process such that the following equalities hold

$$\begin{aligned}Y_t &= \mu t + Y_t^I + Y_t^\rho \\ &= \mu t + B_{X_t} + B_{S_t}^\rho \\ &= \mu t + D_{X_t}\theta + D_\sigma W_{X_t} + D_{S_t}\theta^\alpha + D_{\sigma^\alpha} W_{S_t}^\rho,\end{aligned}$$

where

- $Y^I = \{Y_t^I, t \geq 0\}$ is constructed by subordinating an n-dimensional arithmetic Brownian motion $B = \{B_t, t \geq 0\}$ with an n-dimensional subordinator $X = \{X_t, t \geq 0\}$ with independent components $X_j = \{X_{j,t}, t \geq 0\}$;
- $Y^\rho = \{Y_t^\rho, t \geq 0\}$ is constructed by subordinating an n-dimensional Brownian motion $B^\rho = \{B_t^\rho, t \geq 0\}$ with the common *one*-dimensional subordinator $S = \{S_t, t \geq 0\}$;
- $W = \{W_t, t \geq 0\}$ and $W^\rho = \{W_t^\rho, t \geq 0\}$ are independent n-dimensional Wiener processes, with $corr\left[W_{j,t}, W_{k,t}\right] = 0$ and $corr\left[W_{j,t}^\rho, W_{k,t}^\rho\right] = \rho_{jk}$ for $j \neq k$;
- $X = \{X_t, t \geq 0\}$ and $S = \{S_t, t \geq 0\}$ are an n-dimensional and a *one*-dimensional independent subordinators, independent of $W = \{W_t, t \geq 0\}$ and $W^\rho = \{W_t^\rho, t \geq 0\}$;
- θ^α, σ^α and α are n-dimensional vectors with $\theta^\alpha = \theta \times \alpha$, $\sigma^\alpha = \sigma \times \sqrt{\alpha}$ (the symbol $\times$ stands for the component-wise product of two vectors), where $\alpha_j \in \mathbb{R}_+$, for all j;
- D_{X_t}, D_σ, D_{S_t} and D_{σ^α} are diagonal matrices with diagonal elements $X_{j,t}$, σ_j, S_t, and $\sigma_j\sqrt{\alpha_j}$ respectively, for all j.

The parameter α_j must be chosen so that

$$\begin{aligned} Y_{j,t} &= \mu_j t + Y_{j,t}^I + Y_{j,t}^\rho \\ &= \mu_j t + B_{j,X_{j,t}} + B_{j,S_t}^\rho \\ &= \mu_j t + \theta_j X_{j,t} + \sigma_j W_{j,X_{j,t}} + \theta_j \alpha_j S_t + \sigma_j \sqrt{\alpha_j} W_{j,S_t}^\rho \end{aligned} \tag{3.7}$$

can be written as a time-changed Brownian motion

$$Y_{j,t} = \mu_j t + \theta_j G_{j,t} + \sigma_j W_{G_{j,t}},$$

where $G_{j,t} = X_{j,t} + \alpha_j S_t$ for all j. For each discrete time step Δt the distribution of the increments of the process can be written as

$$Y_{\Delta t} = \mu \Delta t + D_{X_{\Delta t}}\theta + D_\sigma D_{\sqrt{X_{\Delta t}}} Z_{(1)} + \theta^\alpha S_{\Delta t} + \sqrt{S_{\Delta t}} D_\sigma^\alpha A Z_{(2)},$$

where $X_{\Delta t}$ and $S_{\Delta t}$ denote the distributions of the subordinators increments, $Z_{(i)}$ $(i = 1, 2)$ are independent $N(0, I_n)$ random vectors, A is the lower Cholesky decomposition of the correlation matrix of W^ρ.

As shown in Luciano and Semeraro (2010b), the characteristic function of Y_t defined in equation (3.7) is given by

$$\begin{aligned} \Psi_{Y_t}(u) &= \exp(itu'\mu)\, \Psi_{Y_t^I}(u)\, \Psi_{Y_t^\rho}(u) \\ &= \exp(itu'\mu) \exp\left(t \sum_{j=1}^n l_{X_{j,1}}(\varphi_j(u_j))\right) \exp\left(t l_{S_1}(\varphi^{\rho\alpha}(u))\right) \\ &= \exp(itu'\mu) \exp\left(t \sum_{j=1}^n l_{X_{j,1}}(iu_j\theta_j - \frac{1}{2}u_j^2\sigma_j^2)\right) \exp\left(t l_{S_1}(iu'\theta^\alpha - \frac{1}{2}u'\Sigma^{\rho\alpha}u)\right), \end{aligned} \tag{3.8}$$

where $l_{X_{j,1}}(.)$ and $l_{S_1}(.)$ are the Laplace exponents of the subordinators $X_{j,t}$ and S_t, respectively, and $\Sigma^{\rho\alpha} = Var(B_1^{\rho})$ is a positive definite matrix with elements $\Sigma_{jk}^{\rho\alpha} = \sigma_j\sigma_k\sqrt{\alpha_j}\sqrt{\alpha_k}\rho_{jk}$.

Choosing $X_{j,1}$, S_1 and α_j opportunely, Luciano and Semeraro (2010b) and Luciano et al. (2016) proposed different multivariate models with compound Poisson, VG, NIG, and GH margins. We review only the last model and we refer to it as $\rho\alpha$GH model.

Considering the GIG subordinator defined in (3.5) and using (3.8) for $t = 1$, we get the characteristic function of the $\rho\alpha$GH distribution

$$\begin{aligned}\Psi_{Y_1}(u) &= \prod_{j=1}^{n}\left(1-\frac{2}{\psi_j}\left(iu_j\theta_j-\frac{1}{2}\sigma_j^2u_j^2\right)\right)^{a-\frac{\epsilon}{2}}\frac{K_\epsilon\left(\sqrt{\chi_j\left(\psi_j-2\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)\right)}\right)}{K_\epsilon\left(\sqrt{\chi_j\psi_j}\right)}\\ &\quad \exp\left(iu'\mu\right)\left(1-2\left(iu'\theta^{\alpha}-\frac{1}{2}u'\Sigma^{\rho\alpha}u\right)\right)^{-a}.\end{aligned} \tag{3.9}$$

Setting $u_i = 0$, $\forall i \neq j$, into (3.9) we get the characteristic function (2.7) of the GH law.

Setting $\epsilon = 1$ the marginal processes are hyperbolic and we get the $\rho\alpha$HYP. If $a \to 0$ the $\rho\alpha$GH process degenerates into the MGH model with independent univariate GH processes. If $\rho_{jk} = 0$ for all $j \neq k$ then we obtain the αGH model. If $\chi_j \to 0$ for all j, the $\rho\alpha$GH process degenerates into the $\rho\alpha$VG process which includes both the MVG and the αVG. Observe that it is not possible to obtain the MGH model and, since by construction ϵ must be positive, it is not possible to obtain multivariate models with NIG marginal processes. If one sets $\epsilon = \lambda$, $\theta^{\alpha} = \mu^{\rho}$, $\chi_j = \delta_j^2$, $\psi_j = \gamma_j^2 - \beta_j^2$, $\theta_j = \beta_j$, $\mu_j = 0$, $\sigma_j = 1$ for all j, we get the $\rho\alpha$GH process of Luciano and Semeraro (2010b).

4 The multivariate mixed TS distribution

In almost all the models reviewed in this paper the number of parameters increases quadratically with the number of margins, except in the case of the multivariate distributions described in Section 3.1 and 6.2. In practice it is difficult to estimate highly parametrized models. This problem is well known if we deal with portfolio allocation, especially when the investment universe is large. Inspired by the simplicity of the multivariate αVG distribution considered in Luciano and Semeraro (2010b) and Semeraro (2010), the multivariate mixed tempered stable (MMixedTS) distribution has been proposed in Hitaj et al. (2018). According to the authors the reason of proposing the MMixedTS distribution was due to the fact that, nevertheless the simplicity of the αVG model, this one is not able to capture all the stylized facts of assets returns as shown in Hitaj and Mercuri (2013). In particular, the multivariate αVG distribution is not able to generate negatively correlated margins with skewness of the same sign and positively correlated margins with skewness of opposite sign.

The multivariate mixed tempered stable model is built on the basis of the standardized classical tempered stable (stdCTS) distribution. A process $Y = \{Y_t, t \geq 0\}$ with values in $\mathbb{R}^n$ is called MMixedTS if for each margin j the following equality holds

$$Y_{j,t} = \mu_j t + \beta_j V_{j,t} + \sqrt{V_{j,t}}X_{j,t},$$

where $V_j = \{V_{j,t}, t \geq 0\}$ is the j-th component of the multivariate subordinator $V = \{V_t, t \geq 0\}$, defined as

$$V_{j,t} = G_{j,t} + a_j \Lambda_t,$$

in which $G_{j,t}$ and Λ_t are nonnegative infinitely divisible random variables with $G_{j,t}$ and Λ_t mutually independent, $a_j \geq 0$ and

$$X_{j,t}|V_{j,t} \sim stdCTS\left(\alpha_j, \lambda_{+,j}\sqrt{V_{j,t}}, \lambda_{-,j}\sqrt{V_{j,t}}\right).$$

for j from 1 to n.

In particular, if for each j, $G_{j,t} \sim \Gamma(c_j t, m_j)$, $\Lambda_t \sim \Gamma(\bar{n}t, k)$, and $a_j = \frac{k}{m_j}$, then $V_{j,t} \sim \Gamma((c_j + \bar{n})\, t, m_j)$ that guarantees infinite divisibility, necessary for the definition of multivariate MixedTS-Γ. In this case the number of parameters to be estimated is $7n + 1$.

Using matrix notation the MMixedTS distribution can be written as

$$Y = \mu + D_\beta V + D_V^{\frac{1}{2}} X$$

where $\mu \in \mathbb{R}^n$, $D_\beta \in \mathbb{R}^{n \times n}$ with $D_\beta = diag\,(\beta_1, \ldots \beta_n)$, $V \in \mathbb{R}^n$ is a random vector with positive elements, D_V is a random matrix positive defined, such that $D_V = diag\,(V_1, \ldots V_n)$, and X is a stdCTS random vector.

The characteristic function of the MMixedTS process is

$$\begin{aligned}\Psi_{Y_t}(u) &= \exp\left(i\sum_{j=1}^{n} u_j \mu_j t + tl_{\Lambda_1}\left(\sum_{j=1}^{n}\left(ia_j u_j \beta_j + a_j \varphi_{stdCTS}\left(u_j; \lambda_{+,j}, \lambda_{-,j}, \alpha_j\right)\right)\right)\right) \\ &\quad \prod_{j=1}^{n} \exp\left(tl_{G_{j,1}}\left(iu_j\beta_j + \varphi_{stdCTS}\left(u_j; \lambda_{+,j}, \lambda_{-,j}, \alpha_j\right)\right)\right),\end{aligned}$$

where the $\varphi_{stdCTS}\left(u; \alpha, \lambda_+, \lambda_-\right)$ is the characteristic exponent of a stdCTS random variable defined as

$$\varphi_{stdCTS}\left(u;\ \lambda_+,\ \lambda_-,\ \alpha\right) = \frac{(\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha}{\alpha\left(\alpha - 1\right)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)} + \frac{iu\left(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}\right)}{\left(\alpha - 1\right)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)}.$$

Hitaj et al. (2018) showed that the MMixtedTS-Γ model is able to overcome the limits of the multivariate αVG distribution preserving its simplicity. A first application of the MMixedTS distribution to portfolio selection is considered in Hitaj et al. (2019).

5 Multifactorial subordinated models

As described in Luciano and Semeraro (2010a) a multivariate Brownian motion can be subordinated by considering a single factor G defined as an n-dimensional positive random vector. This model can be extended to a multifactorial model as proposed by Marfé (2012a). Further extensions of this model have been proposed by Marfé (2012b) and Boen and Guillaume (2019b). Marfé (2012a) introduced a flexible multidimensional pure jump model with generalized variance gamma (GVG) margins. In particular, the model

is able to describe very different patterns of nonlinear dependence once margins are fixed, and also when linear dependence is fixed. The model flexibility is provided by two separate dependence structures for the positive and the negative jump components of the margins. Additionally, for each jump component it is possible to introduce dependence separately in its finite activity part, that is in large and rare jumps, and in its infinite activity part, that is in small and frequent jumps. In this section we first describe the multivariate generalized gamma (MGG) process. Then, we review the multivariate generalized variance gamma (MGVG) process. Similarly to the VG process, which can be built either as a difference of two independent gamma processes or by time-changing a Brownian motion with an independent gamma process, the MGVG process can be obtained either as difference of two independent MGG processes, or through Brownian motion subordination using the MGG process as a subordinator. We analyze in details only this second approach.

A MGG process is the Lévy process $\hat{G} = \{\hat{G}_t, t \geq 0\}$ on $\mathbb{R}^n_+$, where each component $\hat{G}_j = \{\hat{G}_{j,t}, t \geq 0\}$, $j = 1, ..., n$, is defined as the linear combination of independent subordinators, that is

$$\hat{G}_{j,t} = G_{j,t} + q_j G_{c,t} + G^*_{j,N_{j,t}} + p_j G^*_{c,N_{c,t}} = G_{j,t} + q_j G_{c,t} + X_{j,t} + p_j X_{c,t}$$

with $G_{j,1} \sim \Gamma\left(\frac{1-k_j}{q_j} - c_1, \frac{1}{q_j}\right)$, $G_{c,1} \sim \Gamma(c_1, 1)$, $G^*_{j,1} \sim \Gamma\left(1, \frac{1}{p_j}\right)$, $G^*_{c,1} \sim \Gamma(1, 1)$, $N_{j,1} \sim Poiss\left(\frac{k_j}{p_j} - c_2\right)$, $N_{c,1} \sim Poiss(c_2)$, $X_{j,1} \sim CP\left(\frac{k_j}{p_j} - c_2, 1, \frac{1}{p_j}\right)$ and $X_{c,1} \sim CP(c_2, 1, 1)$ where $0 < c_1 < \min_j \frac{1-k_j}{q_j}$, $0 < c_2 < \min_j \frac{k_j}{p_j}$, and $CP(\lambda, \alpha, \beta)$ denotes the law at time one of a compound Poisson process with jump intensity λ and jump size $\Gamma(\alpha, \beta)$.

The construction in equation (5) allows to express each margin as linear combination of two common factors, $G_{c,t}$ and $X_{c,t}$, and two idiosyncratic factors, $G_{j,t}$ and $X_{j,t}$. Each marginal process can be decomposed into the sum of an infinite activity component, $G_{j,t} + q_j G_{c,t}$, and a finite activity part, $X_{j,t} + p_j X_{c,t}$. Furthermore, all the margins at time one follow a generalized gamma law and we denote this writing $\hat{G}_{j,1} \sim \hat{\Gamma}(1 - k_j, q_j, k_j, p_j)$. The joint characteristic function of $\hat{G} = \{\hat{G}_t, t \geq 0\}$ is given by

$$\begin{aligned}
\Psi_{\hat{G}_t}(u) &= \prod_{j=1}^{n} \Psi_{G_{j,t}}(u_j)\, \Psi_{X_{j,t}}(u_j)\, \Psi_{G_{c,t}}\left(\sum_{j=1}^{n} u_j q_j\right) \Psi_{X_{c,t}}\left(\sum_{j=1}^{n} u_j p_j\right) \\
&= \prod_{j=1}^{n} \exp\left(\frac{itu_j(k_j - c_2 p_j)}{1 - iu_j p_j}\right) (1 - iu_j q_j)^{t\left(\frac{k_j - 1}{q_j} + c_1\right)} \\
&\quad \exp\left(\frac{itc_2 \sum_{j=1}^{n} u_j p_j}{1 - i\sum_{j=1}^{n} u_j p_j}\right) \left(1 - i\sum_{j=1}^{n} u_j q_j\right)^{-tc_1}.
\end{aligned} \tag{5.1}$$

Setting $u_i = 0$, $\forall i \neq j$, into (5.1) we get the characteristic function of $\hat{G}_j = \{\hat{G}_{j,t}, t \geq 0\}$

$$\Psi_{\hat{G}_{j,t}}(u_j) = \exp\left(\frac{iu_j t k_j}{1 - iu_j p_j}\right) (1 - iu_j q_j)^{t\frac{k_j - 1}{q_j}}.$$

Marfé (2012a) defined the MGVG process as the Lévy process $Y = \{Y_t, t \geq 0\}$ on $\mathbb{R}^n$ obtained as the difference of two independent MGG processes. This construction

allows to model explicitly dependence separately and independently both in positive and negative jumps and in their finite and infinite activity components.

Hovewer, Marfé (2012a) provided an alternative way to build a MGVG process through subordination using the MGG process as a subordinator. As already observed, we analyse in details only this second approach. Let $Y = \{Y_t, t \geq 0\}$ be a multivariate process such that the following equalities in law hold

$$\begin{aligned} Y_t &= \mu t + Y_t^{1,I} + Y_t^{2,I} + Y_t^{1,\rho} + Y_t^{2,\rho} \\ &= \mu t + B_{G_t}^{1,I} + B_{X_t}^{2,I} + B_{G_{c,t}}^{1,\rho} + B_{X_{c,t}}^{2,\rho}, \end{aligned}$$

where

- $Y^{1,I} = \{Y_t^{1,I}, t \geq 0\}$ is constructed by subordinating an n-dimensional arithmetic Brownian motion $B^{1,I} = \{B_t^{1,I}, t \geq 0\}$ with an n-dimensional subordinator $G = \{G_t, t \geq 0\}$ with independent components $G_j = \{G_{j,t}, t \geq 0\}$, i.e $B_G^{1,I} = \{D_{G_t}\theta + D_\sigma W_{G_t}^{1,I}, t \geq 0\}$;

- $Y^{2,I} = \{Y_t^{2,I}, t \geq 0\}$ is constructed by subordinating an n-dimensional arithmetic Brownian motion $B^{2,I} = \{B_t^{2,I}, t \geq 0\}$ with an n-dimensional subordinator $X = \{X_t, t \geq 0\}$ with independent components $X_j = \{X_{j,t}, t \geq 0\}$, i.e $B_X^{2,I} = \{D_{X_t}\theta + D_\sigma W_{X_t}^{2,I}, t \geq 0\}$;

- $Y^{1,\rho} = \{Y_t^{1,\rho}, t \geq 0\}$ is constructed by subordinating an n-dimensional Brownian motion $B^{1,\rho} = \{B_t^{1,\rho}, t \geq 0\}$ with the common *one*-dimensional subordinator $G_c = \{G_{c,t}, t \geq 0\}$, i.e $B_{G_c}^{1,\rho} = \{D_{G_{c,t}}\theta^q + D_{\sigma^q} W_{G_{c,t}}^{1,\rho}, t \geq 0\}$;

- $Y^{2,\rho} = \{Y_t^{2,\rho}, t \geq 0\}$ is constructed by subordinating an n-dimensional Brownian motion $B^{2,\rho} = \{B_t^{2,\rho}, t \geq 0\}$ with the common *one*-dimensional subordinator $X_c = \{X_{c,t}, t \geq 0\}$, i.e $B_{X_c}^{2,\rho} = \{D_{X_{c,t}}\theta^p + D_{\sigma^p} W_{X_{c,t}}^{2,\rho}, t \geq 0\}$;

- $W^{l,I} = \{W_t^{l,I}, t \geq 0\}$ and $W^{l,\rho} = \{W_t^{l,\rho}, t \geq 0\}$ are independent n-dimensional Wiener processes, with $corr\left[W_{j,t}^{l,I}, W_{k,t}^{l,I}\right] = 0$ and $corr\left[W_{j,t}^{l,\rho}, W_{k,t}^{l,\rho}\right] = \rho_{jk}$ for $j \neq k$ and $l = 1, 2$;

- G, X, G_c and X_c are independent subordinators, independent of $W^{l,I}$ and $W^{l,\rho}$ for $l = 1, 2$;

- D_{G_t}, D_{X_t}, $D_{G_{c,t}}$, $D_{X_{c,t}}$, D_σ, D_{σ^q} and D_{σ^p} are diagonal matrices with diagonal elements $G_{j,t}$, $X_{j,t}$, $G_{c,t}$, $X_{c,t}$, σ_j, $\sigma_j\sqrt{q_j}$ and $\sigma_j\sqrt{p_j}$, respectively, with σ_j, q_j and $p_j \in \mathbb{R}_+$ for all j.

- μ, θ, θ^q and θ^p are vectors in $\mathbb{R}^n$ with $\theta^q = \theta \times q$, $\theta^p = \theta \times p$ where q and $p \in \mathbb{R}_{0^+}^n$.

From independence and following (3.8) we get the characteristic function of the MGVG

process

$$\Psi_{Y_t}(u) = \prod_{j=1}^{n}\left(1-q_j\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)\right)^{t\left(\frac{k_j-1}{q_j}+c_1\right)}\exp\left(\frac{t\left(k_j-c_2p_j\right)\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)}{1-p_j\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)}\right)$$
$$\left(1-\left(iu'\theta^q-\frac{1}{2}u'\Sigma^q u\right)\right)^{-tc_1}\exp\left(t\left(iu'\mu+\frac{c_2\left(iu'\theta^p-\frac{1}{2}u'\Sigma^p u\right)}{1-\left(iu'\theta^p-\frac{1}{2}u'\Sigma^p u\right)}\right)\right), \tag{5.2}$$

where $\Sigma^q = Var\left(B_1^{1,\rho}\right)$ and $\Sigma^p = Var\left(B_1^{2,\rho}\right)$ are positive definite matrices with elements $\Sigma^q_{jk} = \sigma_j\sigma_k\sqrt{q_j}\sqrt{q_k}\rho_{jk}$ and $\Sigma^p_{jk} = \sigma_j\sigma_k\sqrt{p_j}\sqrt{p_k}\rho_{jk}$. Setting $u_i = 0$, $\forall i \neq j$, into (5.2) we get the characteristic function of the j-th marginal GVG process

$$\Psi_{Y_{j,t}}(u) = \left(1-q_j\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)\right)^{t\frac{k_j-1}{q_j}}\exp\left(t\left(iu_j\mu_j+\frac{k_j\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)}{1-p_j\left(iu_j\theta_j-\frac{1}{2}u_j^2\sigma_j^2\right)}\right)\right).$$

Marfé (2012a) suggested to estimate the MGVG process in two steps. First, estimate the margins with maximum likelihood estimation, recovering the density function from the characteristic function by using the FFT algorithm. Then, given the estimates of margins, estimate the common parameters to calibrate the empirical correlations or the empirical co-skewnesses or both at the same time.

6 Linear combination of Lévy processes

As observed by Ballotta et al. (2019), Lévy processes offer agile distribution modeling for asset prices, but they also present significant estimation challenges, especially in a multivariate setting. A consistent and computationally efficient estimation procedure is needed for portfolio risk measurement and management. While in the previous Sections we reviewed models obtained through subordination, here we analyze a different approach. We discuss how to construct multivariate Lévy models using affine linear transformations of random vectors with independent Lévy components as proposed in Kawai (2009), Kaishev (2013) and further studied by Ballotta and Bonfiglioli (2016) and Ballotta et al. (2019). The main idea is to follow the *divide and conquer* approach, which allows one to solve the curse of dimensionality because each estimation procedure involves only a subsection of the overall parameter space. After a linear trasnformation, a multivariate estimation is splitted into many univariate estimations. These approaches are based on the independent component analysis (ICA) and the principal component analysis (PCA).

6.1 ICA based multivariate linear models

The idea behind this approach is to find a matrix $A \in \mathbb{R}^{n\times n}$ and a random vector $X = (X_1, \ldots, X_n)'$ with infinitely divisible, independent, and standardized components such that the law of the vector AX approximates the law of the standardized log-returns while the correlation matrix of AX approximates a given correlation matrix.

We define a new random vector Z with n entries as follows

$$Z = AX + b, \tag{6.1}$$

where $b \in \mathbb{R}^n$. Requiring that X is a square integrable random vector and assuming, without loss of generality, $E[XX'] = I_n$ with $E[X] = 0$, we have that the following equality holds

$$var[Z] = AA'.$$

If X is an infinitely divisible random vector, we have that Z inherits this property from which it is possible to determine its Lévy measure and the associated characteristic function. The corresponding Lévy process $Z = \{Z_t, t \geq 0\}$ is defined as follows:

$$Z_t = AX_t + bt. \tag{6.2}$$

Assuming that each component of the vector X is not normally distributed it is possible to separate the estimation of matrix A from the estimation of parameters of each component in X, through the ICA proposed in Comon (1994). In the ICA approach, the dependence structure of the components in the vector Z is described through the matrix A, called mixing matrix, that can be easily computed using the *FastICA* algorithm developed in Hyvärinen and Erkki (2000). Through this approach, Madan and Yen (2004) developed a multivariate VG model for asset returns and introduced a portfolio selection procedure based on the maximization of the expected CARA utility function. This approach has been further investigated in Hitaj et al. (2015) and Mercuri and Rroji (2018), where the components of the vector X are assumed to be independent and mixed tempered stable distributed.

In the following, we discuss two alternative approaches for constructing multivariate Lévy models through a scheme described in equation (6.1). The first method proposed in Kawai (2009) and based on the CTS distribution, and the second one proposed in Kaishev (2013) in which a multivariate Lévy process is built as a linear combination of independent gamma processes.

In both cases analyzed in Sections 6.1.1 and 6.1.2, we first standardize the margins and then we apply the *FastICA* algorithm to find the independent components of the vector X. For each margin j we have

$$\frac{Y_j - \mu_j}{\sigma_j} = Z_j = A_j X + b_j, \tag{6.3}$$

where μ_j and σ_j are the empirical mean and empirical standard deviation of Y_j, b_j is equal to zero and A_j is the j-th row of the matrix A. We assume that theoretical means and standard deviations are estimated without errors. By construction means, standard deviations and correlation matrix of the model in equation (6.3) correspond with the empirical ones. The equation (6.3) allows one to obtain the characteristic function of Y_j given the estimates of the standardized independent component X. The characteristic function of the linear combination in equation (6.2) can be written as follows

$$\Psi_{Z_t}(u) = \exp(iu'bt)\Psi_{X_t}(A'u).$$

Additionally, given the moments of the independent components, it is possible to compute the moments of the original margin Y_j, for j from 1 to n. From the homogeneity

property and the additivity property of independent random variables, it follows that the cumulant of order k of a linear combination of independent random variables is a linear combination of their cumulants of the same order with coefficients raised to the power k.

Thus, these models can be estimated by considering the moment matching approach as described in Section 7.1 or by applying a maximum likelihood estimation on each standardized univariate independent component, and by computing, then, the density of each margin Y_j as linear combination of the independent standardized components.

6.1.1 Multivariate linear classical tempered stable model

As proposed in Kawai (2009), assume that $X = \{X_t, t \geq 0\}$ is a Lévy process in $\mathbb{R}^n$ without a Gaussian component and

$$var\left[X_{1,t}\right] = \ldots = var\left[X_{n,t}\right] = t\xi^2 \tag{6.4}$$

holds, with $\xi > 0$. The j-th component of the stochastic process $Y = \{Y_t, t \geq 0\}$ is defined by the equality

$$Y_{j,t} = \sum_{l=1}^{n} c_{j,l} X_{l,t}.$$

For a fixed correlation matrix $\mho$, the transformation matrix K such that $KK' = \mho$ can be obtained for example through a singular value decomposition. The restriction (6.4) implies additional constraints on the marginal parameters of $X_{j,t}$ during the calibration procedure. In order to be in the Kawai's framework and to avoid this additional constraint in the calibration algorithm, we apply the *FastICA* algorithm to the standardized multivariate returns as in equation (6.3). Then on each independent component j we estimate a univariate standardized tempered stable model having the following characteristic function with parameters $(\alpha_j,\, \lambda_{j+},\, \lambda_{j-})$

$$\begin{aligned}\phi_{X_j}(u) = E[\exp(iuX_j)] = \exp\Big(& -iu\big(C\Gamma(1-\alpha_j)(\lambda_{j+}^{\alpha_j-1} - \lambda_{j-}^{\alpha_j-1})\big) \\ & + C\Gamma(-\alpha_j)((\lambda_{j+} - iu)^{\alpha_j} - \lambda_{j+}^{\alpha_j} + (\lambda_{j-} + iu)^{\alpha_j} - \lambda_{j-}^{\alpha_j})\Big)\end{aligned}$$

where

$$C = (\Gamma(2-\alpha_j)(\lambda_{j+}^{\alpha_j-2} + \lambda_{j-}^{\alpha_j-2}))^{-1},$$

and the cumulants are $c_1(X_j) = 0$, $c_2(X_j) = 1$,

$$c_3(X_j) = C\Gamma(3-\alpha_j)(\lambda_{j+}^{\alpha_j-3} - \lambda_{j-}^{\alpha_j-3}),$$

and

$$c_4(X_j) = C\Gamma(4-\alpha_j)(\lambda_{j+}^{\alpha_j-4} - \lambda_{-}^{\alpha_j-4}).$$

We refer to this multivariate Lévy process built as linear combination of independent CTS processes as multivariate linear classical tempered stable (MLCTS) model. While Kawai (2009) considered a process under the so-called mean-correcting martingale measure, we estimate the model under the historical measure where the mean of each margins corresponds to the empirical one.

6.1.2 Multivariate linear gamma model

Kaishev (2013) proposed a new class of processes defined as linear combination of independent gamma processes, called LG processes. In this paper we consider and estimate a special case of this processes, where the univariate standardized independent components are define as sum of independent gamma processes, that is:

$$X_j = \sum_{k=1}^{d} d_k G_{k,t},$$

where $G_{k,t} \sim \Gamma(a_k t, \lambda)$. We refer to it as multivariate linear gamma (MLG) model. We assume $d = 2$ and on each independent component j we estimate a univariate linear gamma model having the following characteristic function with parameters (λ, a_{j+}, a_{j-})

$$\begin{aligned}\phi_{X_j}(u) = E[\exp(iuX_j)] = \exp\Big(&a_{j+}\log(\lambda) - a_{j+}\log(\lambda - id_{j+}u) \\ &+ a_{j-}\log(\lambda) - a_{j-}\log(\lambda + id_{j+}u)\Big),\end{aligned}$$

where

$$\begin{aligned}d_{j+} &= \lambda\sqrt{\frac{a_{j-}}{a_{j+}(a_{j+}+a_{j-})}}, \\ d_{j-} &= \lambda\sqrt{\frac{a_{j+}}{a_{j-}(a_{j+}+a_{j-})}},\end{aligned}$$

and the cumulants are $c_1(X_j) = 0$, $c_2(X_j) = 1$,

$$c_3(X_j) = 2\lambda^{-3}(a_{j+}d_{j+}^3 - a_{j-}d_{j-}^3)$$

and

$$c_4(X_j) = 6\lambda^{-4}(a_{j+}d_{j+}^4 + a_{j-}d_{j-}^4).$$

6.2 PCA based multivariate linear models

A further approach to build multivariate models based on linear combination of independent Lévy processes has been recently proposed by Ballotta and Bonfiglioli (2016) and Ballotta et al. (2019). This approach can be viewed as a further extension of the methods described in Section 6, even if Ballotta et al. (2019) proposed an estimation approach based on the principal component analysis (PCA).

For the case of an n-dimensional model, the authors suggested a 2-step estimation procedure in which a common factor Υ has to be estimated first and then n univariate estimations should be conducted, one per each idiosyncratic component. The model is defined as linear combination of two independent Lévy processes, the first representing a common risk component, the second representing the idiosyncratic risks. The first component is the first principal component defined through the PCA. Even if it is possible to extend the model to the first k principal components to capture k different common risk factors, we consider only the first one in the empirical application.

Let $X_j = \{X_{j,t}, t >= 0\}$ and $\Upsilon = \{\Upsilon_t, t >= 0\}$ be two independent Lévy processes belonging to the same parametric family (e.g. CTS Lévy processes with possible different parameters), then $Y_j = \{Y_{j,t}, t >= 0\}$ can be defined as follows

$$Y_j = X_j + f_j \Upsilon,$$

where f_j is the j-th component of the vector $f \in \mathbb{R}^n$. Given the characteristic functions of X_j and Υ, it is simple to obtain the characteristic function of Y_j as well as its cumulants. According to equation (3) in Ballotta et al. (2019), it can be shown that the correlation is given by the following formula

$$corr(Y_j, Y_k) = \frac{f_j f_k var[\Upsilon]}{\sqrt{var[Y_j]}\sqrt{var[Y_k]}}.$$

The estimation of this model can be conducted by maximum likelihood estimation (MLE) through the FFT: it is fast to implement, and its complexity does not increase with the number of components of the multivariate model. First, the parameters of Υ are estimated, then n independent MLE are performed to estimate the parameters of X_j and the vector f. By construction, this second step can be parallelized in a straightforward way. While Ballotta et al. (2019) conducted the empirical analysis by considering NIG and Merton jump-diffusion processes, we assume that the risk components are CTS distributed. We refer to this multivariate Lévy process built as linear combination of independent CTS processes as multivariate linear classical tempered stable (MLCTS) model.

7 Estimation methods

From a theoretical standpoint, a good estimator should satisfy the following properties: (1) the expected value of the estimator should be equal to the true value of the parameter (*unbiasedness*); (2) as the dimension of the sample increases the estimator should converge in probability to the true value of the parameter (*consistency*); (3) among the unbiased estimators the selected one should be that with the smallest variance (*efficiency*). The knowledge of the sample distribution of an estimator allows to perform hypothesis testing on model parameters. In this section we discuss different estimation methods used in the literature. For each method, we explain the underlying theoretical requirements that ensure proper statistical properties of estimators. Recall that the aim is to estimate the parameters of the log-returns process $Y = \{Y_t, t \geq 0\}$.

7.1 Moments matching (or brute force)

The knowledge of the characteristic function of a multivariate parametric model allows to derive the theoretical moments of the margins and of the joint distribution. Since theoretical moments are expressed as a function of the unknown parameters, it is possible to estimate model parameters by minimizing the distance between empirical and theoretical moments. This simple approach can be applied to estimate all multivariate models discussed in this contribution. We refer to this sort of moments matching estimation method as the *brute force* approach. More in details we minimize the Euclidean norm of

the difference between the first four empirical and theoretical marginal moments and the Frobenius norm of the difference between empirical and theoretical correlation matrices, that is

$$\min_{\Theta} \left(\sum_{i=1}^{4} w_i \| m_i^* - m_i(\Theta) \| + w_\rho \| \rho_i^* - \rho(\Theta) \|_F \right), \tag{7.1}$$

where m_i^* and $m_i(\Theta)$ are the empirical and the theoretical marginal moments of order i, ρ^* and $\rho(\Theta)$ are the empirical and theoretical correlation matrices, w_i and w_ρ are weights.

To take into account the characteristics of each model, including the number of parameters, and to avoid numerical errors in the optimization algorithm, we use different weights w_i and w_ρ (see Section 8 for the choice of the weights). Since this type of approach strictly depends on the starting point in the optimization algorithm, we randomly draw 100 different starting points and select as result the point of minimum distance among the 100 solutions. A careful selection of both the upper and lower parameters bounds is needed to have a satisfactory performance of the optimization algorithm. Since the theoretical moments have a closed form formula, the algorithm is fast for all models. A similar approach will be used in the two-step approach described in Section 7.4 to minimize the distance between empirical and theoretical correlation matrices to find the common parameters governing the dependence structure.

It should be noted that for some models to ensure a proper correlation matrix we apply the hypersphere decomposition as described in Rebonato and Jäckel (1999), that is the correlation matrix of dimension n is decomposed as the product of a lower triangular matrix B and its transpose B'. This lower triangular matrix is function of $n(n-1)/2$ angles and BB' is by construction a correlation matrix.

7.2 Maximum likelihood estimation

Let us consider a multivariate random vector, that is a random variable Y that assumes values on $\mathbb{R}^n$ with an assigned probability law. Given a set of T observations $\{Y^k = Y_{t_k} - Y_{t_{k-1}}\}_{k=1,\dots,T}$, the log-likelihood function can be written as

$$LL(\Theta; Y^1, \dots, Y^T) = \sum_{k=1}^{T} \log f_Y(Y^k; \Theta), \tag{7.2}$$

where Θ is the set of parameters. The idea behind the MLE is to choose the vector Θ that maximizes the likelihood, or equivalently, the logarithm of the likelihood of the observed sample, that is

$$\max_{\Theta} LL(\Theta; Y^1, \dots, Y^T). \tag{7.3}$$

Under mild conditions, the method ensures the consistency property while the efficiency is attained only asymptotically. Moreover, estimators converge in law to the Gaussian distribution with rate $\frac{1}{\sqrt{T}}$. The procedure can be used if it exists a closed form formula of the joint density function. However, as soon as the dimension increases, the optimization problem in equation (7.3) becomes infeasible.

We will consider the MLE algorithm only in the univariate case to estimate the parameter of the margins (e.g. in the two-step procedure described in Section 7.4 or for estimating the linear models described in Section 6). In the GH case there is a closed form

formula for the density function, therefore the likelihood function is simple to compute. In all other cases we will compute the density function by means of the FFT as discussed in details in Bianchi et al. (2019).

7.3 Expectation maximization MLE method

In the subclass of multivariate infinitely divisible distributions that can be written as a mixture, the maximum likelihood approach can be performed using the expectation maximization (EM) algorithm proposed by Dempster et al. (1977). In the class of normal mean-variance mixtures it is necessary to be able to evaluate the posterior distribution of the mixing random variable. For univariate distributions explicit expressions for estimators of parameters have been given in Dimitris (2002) for the univariate NIG and in Loregian et al. (2012) for the univariate VG. Liu and Rubin (1994), Hu (2005) and McNeil et al. (2005) study the EM-based maximum likelihood algorithm for estimating the parameters of the MGH distribution. Bianchi et al. (2016) proposed a simple expectation-maximization maximum likelihood estimation procedure for the MNTS model where the density function of the mixing random variable is computed by means of a FFT procedure.

The density function of a normal mean-variance mixtures distribution can be written as

$$f_Y(y;\Theta) = \int_0^\infty f_{Y|S}(y|s;\mu,\theta,\Sigma)h(s;\Theta_h)ds, \tag{7.4}$$

where $Y|S \sim N(\mu + \theta S, S\Sigma)$ (see Hu (2005)), h is the density function of the mixing random variable with parameter set Θ_h (e.g. the set of parameters of the GIG distribution in the MGH case), and Θ is the set of all model parameters. In the MNTS case the density function h is computed by means of a FFT procedure, that is the characteristic function is inverted to calculate the density function h and the density f_Y in equation (7.4) has to be found by numerical integration.

Given a set of T observations $\{Y^k = Y_{t_k} - Y_{t_{k-1}}\}_{k=1,\dots,T}$, the log-likelihood can be written as

$$LL(\Theta; Y^1, \dots, Y^T) = \sum_{k=1}^{T} \log f_Y(Y^k;\Theta). \tag{7.5}$$

We consider the following likelihood function instead of the likelihood in equation (7.5)

$$\begin{aligned} LL(\Theta; Y^1, \dots, Y^T, S^1, \dots, S^T) &= \sum_{k=1}^{T} \log f_{Y,S}(Y^k, S^k;\Theta) \\ &= \sum_{k=1}^{T} \log f_{Y|S}(Y^k|S^k;\mu,\theta,\Sigma) + \sum_{k=1}^{T} \log h_S(S^k;\Theta_h) \\ &= L_1(\mu,\theta,\Sigma;Y|S) + L_2(\Theta_h;S), \end{aligned} \tag{7.6}$$

where $\{S^k = S_{t_k} - S_{t_{k-1}}\}_{k=1,\dots,T}$ the latent mixing variables. In order to find a MLE based on (7.6), we consider the following iterative algorithm.

1. Set $i = 1$ and select a starting value for $\Theta^{(1)}$, that is $\mu^{(1)} \in \mathbb{R}^n$ is the sample mean, $\theta^{(1)} \in \mathbb{R}^n$ is the zero vector, $V \in \mathbb{R}^n \times \mathbb{R}^n$ is the sample covariance matrix.

2. By considering that

$$f_{S|Y^k}(s; Y^k, \Theta) = \frac{f_{Y|S}(Y^k|s; \mu, \theta, \Sigma) h(s; \Theta_h)}{f_Y(Y^k; \Theta)}, \tag{7.7}$$

compute the following weights

$$\begin{aligned} \delta_k^{(\cdot)} &= E(S^{k^{-1}}|Y^k, \Theta^{(\cdot)}), \\ \eta_k^{(\cdot)} &= E(S^k|Y^k, \Theta^{(\cdot)}), \\ \rho_k^{(i)} &= (Y^k - \mu^{(i)})' \left(\Sigma^{(i)}\right)^{-1} (Y^k - \mu^{(i)}), \end{aligned} \tag{7.8}$$

The expectations in equation (7.8) are evaluated by numerical integration. In equation (7.7), $f_{Y|S}$ can be written in closed form since $Y|S^k \sim N(\mu + \theta S^k, S^k\Sigma)$ (see Hu (2005)). While in the MNTS case the density h is computed by means of a FFT procedure, and the denominator is evaluated by numerical integration, in the MGH case both functions have a closed form formula.

3. Evaluate the average values

$$\bar{\delta}^{(i)} = \sum_{k=1}^{T} \delta_k^{(i)}, \qquad \bar{\eta}^{(i)} = \sum_{k=1}^{T} \eta_k^{(i)}.$$

4. Get the estimates

$$\theta^{(i+1)} = \frac{N^{-1} \sum_{k=1}^{T} \delta_k^{(i)} (\bar{Y} - Y^k)}{\bar{\delta}^{(i)} \bar{\eta}^{(i)} - 1},$$

$$\begin{aligned} \mu^{(i+1)} &= \frac{N^{-1} \sum_{k=1}^{T} \delta_k^{(i)} Y^k - \theta^{(i+1)}}{\bar{\delta}^{(i)}}, \\ \Psi &= \frac{1}{N} \sum_{k=1}^{T} \delta_k^{(i)} (Y^k - \mu^{(i+1)})(Y^k - \mu^{(i+1)})' - \bar{\eta}^{(i)} \theta^{(i+1)} \theta^{(i+1)'}, \\ \Sigma^{(i+1)} &= \frac{|V|^{1/n} \Psi}{|\Psi|^{1/n}}. \end{aligned}$$

5. Set $\Theta^{(i')} = \{\Theta_h^{(i)}, \theta^{(i+1)}, \mu^{(i+1)}, \Sigma^{(i+1)}\}$ and calculate the new weight $\bar{\eta}^{(i')}$ as done in Steps 2 and 3.

6. To complete the calculation of $\Theta^{(i+1)}$, find Θ_h that maximize the likelihood function in equation (7.5), that is

$$LL(\Theta^{(i+1)}; Y^1, \ldots, Y^T) = \sum_{k=1}^{T} \log f_Y(Y^k; \Theta^{(i+1)}),$$

where $\Theta^{(i+1)} = \{a, \lambda, C, \theta^{(i+1)}, \mu^{(i+1)}, \Sigma^{(i+1)}\}$.

7. If $i < 1,000$ and $LL(i) - LL(i-1) > 1e-5$, increment iteration count i and go to step 2, otherwise, stop the algorithm.

As discussed in McNeil et al. (2005), the algorithm converges to the MLE because it produces improved parameter estimates at each step, in the sense that the value of the original likelihood is continually increased. In the applications contained in this paper, the algorithm converges long before reaching 1,000 iterations.

7.4 Two-step procedure

In literature several authors have considered a two-step procedure for the estimation of non-Gaussian multivariate models. The idea is to split parameters into two groups: the parameters of the first group can be estimated on the margins, usually by MLE, while the parameters of the second group, used to capture the dependence structure, can be estimated by minimizing some distance between the theoretical and empirical higher co-moments. While marginal parameters are estimated using the MLE method discussed in Section 7.2, the Frobenius norm is usually applied to minimize the distance between theoretical and empirical correlation matrices (see Section 7.1). A two-step procedure was used in Marfé (2012a), Hitaj and Mercuri (2013), Luciano et al. (2016) and Boen and Guillaume (2019a).

Here we discuss the conditions required for applying this method and statistical properties of estimators. If it is not possible to identify the set of parameters that completely characterize the margins this method should not be applied. For example this is the case of NMV as the parameters governing the dependence structure affect all margins. A situation when the two-step procedure can be applied refers to the case where the dependence structure is introduced through a multivariate mixing random variable as for example in Luciano and Semeraro (2010b). Notice that it is necessary to impose constraints on the parameters of the subordinator leading to marginal distributions which become functions only of parameters of the first group (see Guillaume (2013) for a discussion about these constraints in case of multivariate αVG and αNIG).

7.5 Estimation methods based on the characteristic function

Most of the multivariate models reviewed in this paper do not have a closed form formula for the density function. This makes the likelihood-based estimation inconvenient. However, these models can be represented through some transformations of the probability function such as the characteristic function or the Laplace transform. We discuss briefly the generalized method of moments (GMM) proposed in Hansen (1982) that can be used for the estimation of all the multivariate models described in this paper. The aim of this procedure is to estimate the vector of parameters $\Theta \in \mathbb{R}^p$ from a model based on the following vector of q unconditional moment conditions:

$$E\left[g\left(\Theta, Y\right)\right] = 0 \tag{7.9}$$

where $g\left(\cdot\right) : \Xi \times \mathbb{R}^n \to \mathbb{R}^q$, $\Xi \subset \mathbb{R}^p$ is a compact space. For a given sample $Y^1, Y^2, \ldots, Y^T$, we replace the expectation in (7.9) with the sample mean and obtain:

$$\bar{g}_T\left(\Theta\right) = \frac{1}{T}\sum_{k=1}^{T} g\left(\Theta, Y^k\right).$$

The GMM estimator depends on the choice of a positive definite weighting matrix $F_T \in \mathbb{R}^{q\times q}$ and is the solution of the following minimization problem

$$\hat{\Theta}\left(F_T\right) = \min_{\Theta\in\Xi}\ \bar{g}_T\left(\Theta\right)' F_T \bar{g}_T\left(\Theta\right). \tag{7.10}$$

To find an estimator, we need at least as many moment conditions as the number of parameters. In particular we have the classical method of moments (MM) for $q = p$ and the GMM for $q > p$. Under mild conditions, for any positive definite weighting matrix F_T, the GMM produces consistent estimators. Moreover the asymptotic distribution of $\hat{\Theta}$ is

$$\sqrt{T}\left(\hat{\Theta} - \Theta\right) \sim N\left(0, H\right)$$

where H is the asymptotic variance-covariance matrix defined as

$$H = \left(J'FJ\right)^{-1} J'FRFJ\left(J'FJ\right)^{-1}$$

with $J \in \mathbb{R}^{q\times p}$

$$J = E\left[\frac{\partial g\left(\Theta, Y\right)}{\partial \Theta'}\right]$$

and $R \in \mathbb{R}^{q\times q}$

$$R = E\left[g\left(\Theta, Y\right) g\left(\Theta, Y\right)'\right],$$

and $F_T \xrightarrow{P} F$ as $T \to \infty$. An appropriate choice of matrix F improves the efficiency within the class of GMM type estimators. The most efficient estimator is obtained if

$$F_T \xrightarrow{P} R^{-1} \tag{7.11}$$

and, in that case, the variance-covariance matrix H becomes

$$H = \left(J'R^{-1}J\right)^{-1}. \tag{7.12}$$

Several algorithms have been proposed in literature in order to obtain an estimator with variance-covariance matrix that approaches matrix H in (7.12).

In this paper we use the R package `gmm` developed in Chaussé (2010), where the optimal matrix F is estimated using the heteroskedastic auto-correlation consistent (HAC) approach proposed in Newey and West (1987). Then, as g we select the distance between the empirical and theoretical characteristic function. For a given grid $u_j \in \mathbb{R}^n$, with j from 1 to q, g is defined as

$$g\left(\Theta, Y, u_j\right) = e^{i\langle u_j, Y\rangle} - \Psi_{Y,\Theta}\left(u_j\right),$$

where $\langle\ ,\ \rangle$ is the scalar product. The moment conditions require

$$E\left[g\left(\Theta, Y, u_j\right)\right] = 0,$$

where the j-th element of the vector function $\bar{g}_T\left(\Theta\right)$ is

$$\bar{g}_{T,j}\left(\Theta\right) = \frac{1}{T}\sum_{k=1}^{T}\left(e^{i\left\langle u_j, Y^k\right\rangle} - \Psi_{Y^k,\Theta}\left(u_j\right)\right). \tag{7.13}$$

It is evident that the choice of the grid is crucial (see Section 8 for the choice of the grid). Increasing the grid size u_j implies numerical instability and R^{-1} may not be defined.

8 Empirical analysis

In this paper we have tried to highlight the differences between various multivariate distributions applied to finance. Also, different estimation methods have been discussed. In order to select a *good* multivariate model it is necessary to understand the most important features we want to be able to replicate, the computational burden of the choice and the characteristics of the algorithm selected in the estimation. In this section we compare the different multivariate non-normal models with the multivariate normal one to which we refer to as *MNormal*. This model is completely characterized through the mean vector μ and covariance matrix Σ.

In Table 1 the number of parameters as a function of the number of margins is reported for each multivariate model. The models with a linear dependence on the number of margins are the MMixedTS, the αGH and the PCA based linear model with k equal to 1. For the other models the dependence is quadratic.

Model	number of parameters	simulation algorithms
MNormal	$\frac{n^2+3n}{2}$	*rnorm*
MGH	$\frac{n^2+5n}{2}+3$	*rnorm, rgig*
MNTS	$\frac{n^2+5n}{2}+3$	*rnorm, FFT*
αGH	$5n+2$	*rnorm, rgig*
$\rho\alpha$GH	$\frac{n^2+9n}{2}+2$	*rnorm, rgig*
MMixedTS	$7n+1$	*rgamma, FFT*
MGVG	$\frac{n^2+11n}{2}+2$	*rnorm, rgamma, rpois*
ICA MLG	n^2+5n	*rgamma or FFT*
ICA MLCTS	n^2+5n	*FFT*
PCA MLCTS	$5n+(n+5)k$	*FFT*

Table 1: Number of parameters as a function of the number of margins n and possible simulation algorithms implemented in R. The number of principal components is equal to k. We refer to the inverse transform algorithm based on the fast Fourier transform as *FFT*.

In this section we report the estimation results of the models we discussed. We calibrate each model with three different algorithms (only the MMixedTS is estimated with two different approaches). We estimate the models under the so-called historical probability measure, that is by using observed market returns without relying on additional data, like for example option prices (see Bianchi and Tassinari (2020)).

The analysis is performed on Datastream daily dividend-adjusted closing prices from July 1, 2003 through June 29, 2018 for five stock indexes: the Deutsche Aktienindex 30 (ticker DAX), the Cotation Assistée en Continu 40 (ticker CAC), the Financial Times Stock Exchange Milano Indice di Borsa (ticker FTSEMIB), Índice Bursátil Español (ticker IBEX), Amsterdam Exchange Index (ticker AEX) representing five major European indexes. Table 2 reports the summary of statistics, for each European indexes, calculated on daily log-returns. The time period in this study includes the high volatility period after the Lehman Brothers filing for Chapter 11 bankruptcy protection (September 15, 2008), the eurozone sovereign debt crisis, during which, in November 2011, the spread between the 10-year Italian BTP and the German Bund with the same maturity exceeded 500 basis points, and the recent Italian political turmoil at the end of May 2018.

	min	max	mean	std	skewness	ex.kurtosis
DAX	-0.074	0.1080	0.03%	1.31%	-0.0510	6.2406
CAC	-0.0947	0.1059	0.01%	1.34%	-0.0458	6.9923
FTSEMIB	-0.1333	0.1088	-0.0%	1.50%	-0.2547	6.1572
IBEX	-0.1319	0.1348	0.01%	1.41%	-0.1459	8.3714
AEX	-0.0959	0.1003	0.02%	1.26%	-0.1998	8.8435

Table 2: Summary statistics of daily log-returns.

We estimate the models using the methods discussed in Section 7. The code is implemented in R language. Three estimation error measures based on the distance between theoretical and empirical distribution function are considered.

The first error measure is the Kolmogorov-Smirnov distance (KS)

$$KS = \sup_x \left| \hat{F}(x) - F(x) \right|, \tag{8.1}$$

applied to all margins. If the margins do not have closed form formula for the density function, the evaluation of equation (8.1) is conducted by means of the FTT as discussed in details in Bianchi et al. (2019). Given the number of observations, the KS distance provides a satisfactory result if it is less than 0.03 (i.e. the KS test has a p-value grater than 0.05). The estimation error, in terms of margins (calculated using KS distance), is reported in Table 3, columns labeled KS_1, KS_2, KS_3, KS_5, for all the models under analysis.

The second measure is represented by the distances between empirical and theoretical moments considered in Section 7.1. The estimation error in terms of moments is reported in Table 3, columns labeled, *mean*, *sd*, *skewness*, *ex-kurtosis*. The distance between historical and model correlations is reported in the column labeled *rho*.

The third error measure is given by the norm of the vector $\bar{g}_{T,j}$ defined in equation (7.13), which gives the average distance between the empirical and the theoretical characteristic function given a grid $u_j \in \mathbb{R}^n$, with j from 1 to 50. The average distance between the empirical and the theoretical characteristic function is reported in Table 3, column labeled $\|\bar{g}\|^2$.

As fourth error measure, we evaluate a distance between multivariate densities, using the nonparametric approach as proposed in Li et al. (2009) that considers all components together and implemented in the *npdeneqtest* function of the R package *np* (see also Xia and Grabchak (2022)). This distance is based on kernel density estimators between observed and simulated data. For each model, we compare the 5-dimensional market data to the simulated samples based on parameters estimated for the ten models. The number of observations of the simulated sample is the same of the observed one (i.e. 3,843 trading days). The estimates of the considered multivariate distance, to which we refer to as *md*, are reported in the last column of Table 3.

It should be noted that for all the ten models analyzed in this paper there is a stochastic representation allowing us to simulate random variates from these non-normal models by drawing random samples from the multivariate normal and univariate random variables (see Table 1). While the simulation of the Gamma, Poisson, generalized inverse Gaussian random variables are obtained through the *rgamma* and *rpois* of the *stats* package of R and the *rgig* function of the *ghyp* package of R, the simulation of CTS random

draws is not a simple task and an ad-hoc implementation is needed. We consider the inverse transform algorithm and the cumulative distribution function is evaluated through the fast Fourier transform (see Bianchi et al. (2017) and Bianchi et al. (2019) for a detailed description).

A visual assessment of the empirically findings explored in this paper are provided in Figure 1 where we show the simulated log-returns drawn from the ten parametric models fitted on market data and compare them to observed market log-returns. For each model we take the parameters of the best performing estimation method. The number of simulations is equal to the number of observed market data. The scatterplots shown in Figure 1 seem coherent with the results reported in Table 3.

Analyzing the results in Table 3 and by looking at Figure 1, as expected, even if the estimation method based on the moments (i.e. the brute force approach) is very simple to implement, it does not always provide a satisfactory estimation error. However, it can be a good starting point for the other estimation algorithms considered in this study. Due to the large number of parameters, in some cases it is not easy to understand which can be a good starting point for the optimization procedure. The brute force approach gives the possibility to explore a multivariate model without having to implement complex algorithms or without having to wait too long for the algorithm convergence. However, in some cases the estimation error is large if compared to more robust methods (e.g. the EM or the GMM algorithm). Even if the GMM algorithm can be applied to all models having a characteristic function in closed form, the computing time of this algorithm as well as possible numerical issues may be an obstacle for large scale practical applications. The GMM is a very general estimation approach and for this reason it may be more difficult to use in comparison with ad-hoc estimation approaches implemented for specific cases. However, with a proper selection of the grid and of the parameter boundaries, the GMM provide satisfactories results.

As starting points of the moments-matching approach, we draw 100 random starting points in the parameters space. The w_i and w_ρ depend on the selected parametric model and the choice is done after exploring the dataset and the algorithm itself. We try to select the weights in a way that all moments have a similar importance in the optimization algorithm. For the MNTS, the MGH, the αGH and the MLCTS model based on PCA we consider only mean, standard deviation and correlation in equation (7.1); for the $\rho\alpha$GH we consider also the skewness; for the MGVG and the MMixedTS we consider also both skewness and excess kurtosis; for the multivariate linear models based on ICA we consider all marginal moments up to order four. This is the reason why in Table 3 in both the MNTS and the MGH case the error in fitting the skewness and the excess kurtosis is large and in the other cases the error in fitting moments of order higher than two is not so big.

Both the EM and the two-step approaches are reasonably fast: the maximization of the likelihood function is conducted only on univariate models. The EM approach applied to the MGH and MNTS models works properly even for large scale practical applications (see Bianchi and Tassinari (2020) and Bianchi et al. (2020)). Our optimization algorithms in R do not rely on parallel computing techniques and the code implements the L-BFGS-B method. However, while it is not so simple to parallelize the optimization algorithm, it should be noted that the n independent MLE steps of the models leveraging on univariate MLE can be run in parallel without great effort (e.g. it is possible to write an efficient

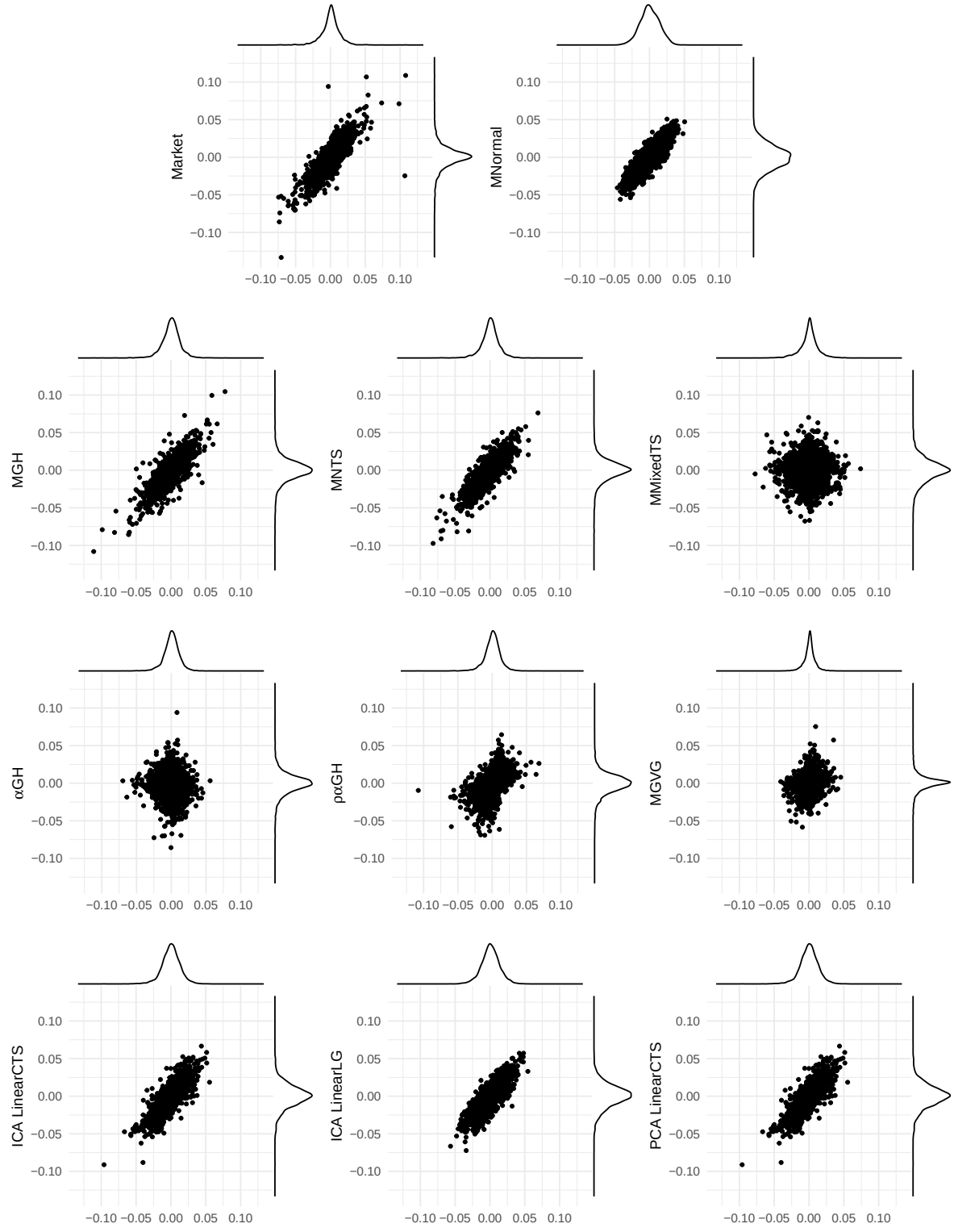

Figure 1: Bivariate scatterplots of market and simulated data. The DAX (x-axis) and the FTSEMIB (y-axis) indexes are considered. Depicted market data are daily log-returns from July 1, 2003 through June 29, 2018. The densities of the margins are reported on the secondary axes.

R code with the packages *foreach* and *doParallel*). This is the case for the two-step approaches and the linear models (i.e. αGH, $\rho\alpha$GH, MGVG, MLG and MLCTS) which are based on the *divide et impera* concept: the estimation procedure is simplified, the dimensionality problem is solved and the models provide a consistent and parallelizable parameters estimation.

In the GMM estimation the multivariate grid of dimension n is selected as follows. For the first dimension we consider a vector of q equally spaced points in the interval between minimum and maximum observed returns. Then, after having fixed a seed, to obtain the vector representing the second dimension, we randomly permute the vector obtained for the first dimension. The same approach is considered for all other dimensions up to n. The value of q depends on the model and it ranges from 15 to 50. As starting point we consider the estimates obtained through the moments-matching approach. First, we obtain a preliminary estimate by considering as weighting matrix the identity matrix, then we conduct a second estimation with the weighting matrix given in equation (7.11). The selection of the starting point and of the grid largely affects the final result of the optimization procedure. The estimates obtained through the GMM approach are usually not far from the starting point. The GMM approach seems to work better for models with a simpler dependence structure and a smaller number of parameters (i.e. the MMixedTS model). A proper selection of q is needed to avoid possible numerical issues of the R package *gmm*. Even if the GMM approach is reasonably fast, it is not always simple to obtain satisfactory results in terms of margins fitting and convergence properties. This may be caused by the large number of parameters involved in the optimization problem. In order to speed up the GMM algorithm and to avoid loops, the characteristic function should take as input a matrix u, instead of a vector u, and implement the code leveraging on matrix operations. This can be done for all parametric models analyzed in this work.

From an estimation error standpoint, some models have a very good performance in fitting the margins, but they show a bad correlation fitting (e.g. the MMixedTS and the MGVG). For the MMixedTS the large error in fitting the empirical correlation matrix seems to be due to the number of parameters, too small to explain the behavior of both margins and correlations. The MMixedTS is the best performer in fitting the margins. The MGH and the MNTS models have a satisfactory performance, even if the correlation fitting is not as good as for other competitors. At least for the data analyzed in this paper, the $\rho\alpha$GH seems to show the better mix between estimation errors and computational tractability, even if the two-step procedure is not elegant from a pure statistical perspective. The ICA based linear models are simple to estimate, mainly because the multivariate estimation problem is converted to a set of univariate problems. However, the performance is not always good enough and some numerical issues in the FFT inversion of the characteristic function may affect the evaluation of the estimation errors. These issues are caused by the fact that the model parameters are estimated on the independent components and the margins are obtained by multiplying these components by small numbers, that is by the elements of the matrix A and of the vector σ (see equation (6.3)). The PCA based linear model with CTS components is more efficient from a computational standpoint and it is simpler to implement, at least if one considers the first principal component only. In our view, this last model is very promising, even if it has a less flexible dependence structure in comparison with the $\rho\alpha$GH. The estimation procedure has a computational complexity equivalent to the estimation of a non-normal

univariate model.

We recall that there exists a tradeoff between model fitting, prediction accuracy and parsimony. In practice, the return generating process in unknown and selecting a model on the basis of fitting accuracy may increase the propensity to choose an unnecessarily complicated models. This is due to the fact we do not know when we are using observations advantageously to characterize the model and when we are fitting observation noise. This last will contaminate the model and probably degrading its predictive capabilities. In general simplicity or complexity does not equal accuracy, but finding the middle ground that provides the most accurate predictions is problematic. In this paper we deal with estimation and fitting accuracy of the multivariate non-Gaussian distributions but we do not deal with the predictive accuracy of the models.

9 Conclusions

In this paper we provide a guide for the use of continuous-time multivariate non-Gaussian models based on Lévy processes with a view toward applications to finance. After a detailed analysis of the theoretical structure behind a sample of multivariate models proposed in the financial literature, we observed their performance in terms of fitting on a five-dimensional series of log-returns. The contribution of the paper is not only to present models with a unifying notation but also to give some inputs for the practical implementation of their estimation algorithms. For each model we provide the necessary formulas and methods needed to find a preliminary estimate that can be used as starting point of more complex and robust algorithms. Additionally, we propose different estimation methods which can be used in practical applications.

The parametric models reviewed is this paper have a different level of complexity from both a theoretical and practical standpoint. We show that it is not always true that a greater level of complexity provides a better estimation performance, at least for the data considered in this study. In most cases the multivariate estimation problem can be decomposed in different steps with computational complexity similar to a univariate estimation problem. When this decomposition is not possible, we show how to perfom a satisfactory parameters estimation.

Empirical motivation, statistical properties of estimators and computational tractability are important features that should be taken into accout when selecting a model to be used in practice. This work should be read as a tutorial allowing one to have a practical view on different multivariate non-Gaussian models based on Lévy processes proposed in the quantitative finance literature in the last years. The work does not pretend to be an integrative survey providing a complete picture of the existing literature and solving the still open problems.

		KS_1	KS_2	KS_3	KS_4	KS_5	mean	sd	skewness	ex-kurtosis	rho	$\|var(\bar{g})^{-1/2}\bar{g}\|^2$	$\|\bar{g}\|^2$	md
Mnormal	MLE	0.079	0.077	0.079	0.074	0.080	0.00e+0	0.00e+00	6.98e-1	36.632	0.000		1.07e-16	146.05
MGH	moments	0.056	0.096	0.073	0.077	0.081	1.20e-5	7.73e-5	5.60e+1	364.456	0.001		1.73e-11	191.96
	EM	0.027	0.014	0.022	0.018	0.021	1.96e-6	2.53e-3	4.97e+0	193.856	0.014		5.97e-12	10.81
	GMM	0.034	0.058	0.040	0.043	0.039	4.70e-3	7.08e-4	5.40e+1	359.648	0.005	0.049	1.24e-6	127.15
MNTS	moments	0.061	0.045	0.067	0.060	0.054	5.72e-4	4.67e-4	3.03e+0	33.685	0.030		1.55e-9	189.77
	EM	0.028	0.015	0.022	0.019	0.022	2.45e-4	4.87e-3	6.39e-1	36.708	0.015		3.43e-9	47.62
	GMM	0.027	0.023	0.058	0.044	0.026	1.52e-4	7.23e-3	3.39e+0	33.088	0.038	0.060	2.36e-9	311.00
MMixedTS	moments	0.035	0.017	0.032	0.011	0.039	1.97e-5	1.96e-4	1.98e-2	4.613	0.781		5.69e-10	466.39
	GMM	0.034	0.016	0.031	0.010	0.038	1.41e-4	2.10e-4	1.98e-2	4.604	0.781	0.989	2.23e-9	460.52
AlphaGH	moments	0.096	0.096	0.102	0.100	0.100	1.37e-4	7.03e-4	8.32e+0	13.481	0.021		2.07e-9	506.31
	MLE + correlation	0.034	0.038	0.024	0.028	0.043	1.67e-3	6.79e-3	1.60e+0	37.617	0.384		2.12e-7	448.28
	GMM	0.079	0.074	0.068	0.060	0.086	1.89e-3	3.68e-3	8.48e+0	13.233	0.028	0.074	1.98e-7	479.16
RhoAlphaGH	moments	0.154	0.110	0.121	0.103	0.207	1.37e-3	3.10e-2	1.42e-3	39.408	0.103		1.91e-7	312.73
	MLE + correlation	0.034	0.038	0.024	0.028	0.043	1.67e-3	6.79e-3	1.60e+0	37.617	0.000		2.12e-7	218.58
	GMM	0.088	0.061	0.063	0.047	0.104	5.30e-4	6.22e-3	2.80e-1	39.873	0.065	0.104	2.12e-8	259.14
MGVG	moments	0.080	0.100	0.016	0.043	0.062	1.02e-4	2.01e-4	1.92e-2	25.148	0.528		4.11e-7	441.04
	MLE + correlation	0.014	0.016	0.013	0.012	0.018	1.01e-4	2.65e-3	4.17e-1	54.392	1.035		5.83e-6	244.67
	GMM	0.148	0.117	0.086	0.087	0.137	2.40e-3	1.30e-2	3.67e+0	24.110	0.590	0.424	1.75e-7	310.12
ICA LinearLG	FastICA + moments	0.090	0.059	0.059	0.068	0.068	2.40e-5	2.40e-5	5.88e-4	7.732	0.000		1.03e-9	90.92
	FastICA + MLE	0.091	0.064	0.066	0.063	0.061	2.40e-5	2.40e-5	3.71e-1	7.620	0.000		5.68e-10	85.56
	FastICA + GMM	0.067	0.043	0.057	0.059	0.055	0.00e+0	1.95e-6	4.90e-1	6.811	0.000	0.994	1.03e-9	115.46
ICA LinearCTS	FastICA + moments	0.091	0.066	0.061	0.067	0.079	2.40e-5	2.40e-5	2.40e-5	7.741	0.000		8.24e-10	113.98
	FastICA + MLE	0.102	0.077	0.068	0.077	0.090	2.40e-5	2.40e-5	1.46e-1	5.942	0.000		9.57e-10	70.35
	FastICA + GMM	0.091	0.066	0.061	0.067	0.079	2.40e-5	2.40e-5	2.40e-5	7.741	0.000	0.998	8.24e-10	112.15
PCA LinearCTS	PCA + moments	0.034	0.035	0.046	0.036	0.044	5.74e-5	3.09e-2	5.79e-1	56.988	0.066		1.10e-9	556.17
	PCA + MLE	0.046	0.038	0.041	0.039	0.047	1.86e-3	2.02e-2	2.03e+0	38.635	0.383		2.36e-7	407.39
	PCA + GMM	0.032	0.033	0.045	0.034	0.041	4.17e-4	2.92e-2	5.77e-1	60.167	0.068	0.844	1.22e-8	559.58

Table 3: For each model and estimation approach we report the univariate Kolmogorov-Smirnov distance and the errors in fitting empirical moments in the period from July, 1 2003 to June, 29 2018. We estimate the model by considering the following algoritms: we minimize the distance between empirical and theoretical moments (*moments*), the expectation-maximization (*EM*), the generalized method of moments (*GMM*), the two-step procedure in which we first estimate the univariate margins by maximum likelihood estimation (MLE) and then we minimize the distance between empirical and theoretical correlations (*MLE + correlations*), the two-step procedure in which we first apply the FastICA or the PCA algoritm and then we minimize the errors between empirical and theoretical moments (*FastICA or PCA + moments*), the two-step procedure in which we first apply the FastICA or the PCA algoritm and then we estimate the standardized independent components by MLE or GMM (*FastICA or PCA + MLE or GMM*). While $\|var(\bar{g})^{-1/2}\bar{g}\|^2$ represents the GMM objective function at the optimal point, $\|\bar{g}\|^2$ represents the average distance between the empirical and the theoretical characteristic function with $q = 50$. Finally, *md* is the nonparametric distance between observed and simulated data proposed by Li et al. (2009).

References

D. Allen and S. Satchell. The four horsemen: heavy-tails, negative skew, volatility clustering, asymmetric dependence. *Discussion Paper, University of Sydney*, n. 4, 2014.

D. Applebaum. Lévy processes and stochastic calculus. *Cambridge University Press*. Cambridge, 2009.

R. B. Arellano-Valle and A. Azzalini. A formulation for continuous mixtures of multivariate normal distributions.. *Preprint*, arXiv:2003.13076.

L. Ballotta and E. Bonfiglioli. Multivariate asset models using Lévy processes and applications. *The European Journal of Finance*, 22(13):1320–1350, 2016.

L. Ballotta, G. Fusai, A. Loregian, and M.F. Perez. Estimation of multivariate asset models with jumps. *Journal of Financial and Quantitative Analysis*, 54(5):2053–2083, 2019.

O.E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 353(1674):401–419, 1977.

O.E. Barndorff-Nielsen, J. Pedersen, and K.I. Sato. Multivariate subordination, self-decomposability and stability. *Advances in Applied Probability*, 33(1):160–187, 2001.

M. Bedendo, F. Campolongo, E. Joossens, and F. Saita. Pricing multiasset equity options: How relevant is the dependence function? *Journal of Banking and Finance*, 34(4):788–801, 2010.

J. Bertoin. Lévy processes. *Cambridge University Press*. Cambridge, 1996.

M.L. Bianchi and G.L. Tassinari. Forward-looking portfolio selection with multivariate non-Gaussian models. *Quantitative Finance*, 2020.

M.L. Bianchi, S.T. Rachev, Y.S. Kim, and F.J. Fabozzi. Tempered infinitely divisible distributions and processes. *Theory of Probability and Its Applications (TVP)*, 55(1): 59–86, 2011.

M.L. Bianchi, G.L. Tassinari, and F.J. Fabozzi. Riding with the four horsemen and the multivariate normal tempered stable model. *International Journal of Theoretical and Applied Finance*, 19(4), 2016.

M.L. Bianchi, S.T. Rachev, and F.J. Fabozzi. Tempered stable Ornstein-Uhlenbeck processes: A practical view. *Communications in Statistics - Simulation and Computation*, 423-445(1):59–86, 2017.

M.L. Bianchi, S.V. Stoyanov, G.L. Tassinari, F.J. Fabozzi, and S.M. Focardi. *Handbook of heavy-tailed distributions in asset management and risk management*. World Scientific, 2019.

M.L. Bianchi, G. De Luca, and G. Rivieccio. CoVaR and backtesting: a comparison between a copula approach and parametric models. *Book of short papers - SIS 2020*, 2020.

L. Boen and F. Guillaume. Building multivariate Sato models with linear dependence. *Quantitative Finance*, 19(4):619–645, 2019a.

L. Boen and F. Guillaume. Towards a Δ-Gamma Sato multivariate model. *Review of Derivatives Research*, pages 1–39, 2019b.

S. Bossu. *Advanced Equity Derivatives: Volatility and Correlation*. Wiley, 2014.

S.I. Boyarchenko and S.Z. Levendorskii. Option pricing for truncated Lévy processes. *International Journal of Theoretical and Applied Finance*, 3(3), 2000.

S.I. Boyarchenko and S.Z. Levendorskii. *Non-Gaussian Merton-Black-Scholes theory*. River Edge, New Jersey: World Scientific Publishing, 2002.

B. Buchmann, B. Kaehler, R. Maller and A. Szimayer. Multivariate subordination using generalised Gamma convolutions with applications to Variance Gamma processes and option pricing. *Stochastic Processes and their Applications*. 127(7):2208–2242, 2017.

B. Buchmann, K.W. Lu, and D.B. Madan. Calibration for weak variance-alpha-gamma processes. *Methodology and Computing in Applied Probability*, 21(4):1151–1164, 2019.

P. Chaussé. Computing generalized method of moments and generalized empirical likelihood with R. *Journal of Statistical Software*, 34(11):1–35, 2010.

P.K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.

P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3): 287–314, 1994.

R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman and Hall/CRC, 2003.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

K. Dimitris. An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution. *Statistics and Probability Letters*, 57(1):43–52, 2002.

Y. Dominicy, H. Ogata, and D. Veredas. Inference for vast dimensional elliptical distributions. *Computational Statistics*, 28(4):1853–1880, 2013.

E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1(3):288–299, 1995.

E. Eberlein, , and K. Prause. The generalized hyperbolic model: Financial derivatives and risk measures. In H. Geman, D. Madan, S.R. Pliska, and T. Vorst, editors, *Mathematical Finance, Bachelier Congress*, pages 245–267. Springer Berlin Heidelberg, 2002.

H.A. Fallahgoul, Y.S. Kim, and F.J. Fabozzi. Elliptical tempered stable distribution. *Quantitative Finance*, 16(7):1069–1087, 2016.

G. Frahm. *Generalized elliptical distributions: theory and applications.* PhD thesis, Universität zu Köln, 2004.

H. Geman, D. Madan and M. Yor. Time changes for Lévy processes. *Mathematical Finance*. 11 (1): 79–96, 2001.

F. Guillaume. The αVG model for multivariate asset pricing: calibration and extension. *Review of Derivatives Research*, 16(1):25–52, 2013.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, pages 1029–1054, 1982.

L. Harris. Cross-security tests of the mixture distribution. *Journal of Financial and Quantitative Analysis*, 21(1):39–46, 1986.

A. Hitaj and L. Mercuri. Portfolio allocation using multivariate variance gamma models. *Financial markets and portfolio management*, 27(1):65–99, 2013.

A. Hitaj and L. Mercuri. Hedge fund portfolio allocation with higher moments and MVG models. In: Batten J.A., MacKay P., Wagner N. (eds), *Advances in Financial Risk Management.* Palgrave Macmillan, London, 331–346, 2013.

A. Hitaj, L. Mercuri, and E. Rroji. Portfolio selection with independent component analysis. *Finance Research Letters*, 15:146–159, 2015.

A. Hitaj, F. Hubalek, L. Mercuri, and E. Rroji. On properties of the MixedTS distribution and its multivariate extension. *International Statistical Review*, 86(3):512–540, 2018.

A. Hitaj, L. Mercuri, and E. Rroji. Sensitivity analysis of the mixed tempered stable parameters with implications in portfolio optimization. *Computational Management Science*, 16(1–2):71–95, 2019.

W. Hu. *Calibration Of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization and portfolio credit risk.* PhD thesis, Florida State University, 2005.

A. Hyvärinen and O. Erkki. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000.

V.K. Kaishev. Lévy processes induced by Dirichlet (b-)splines: Modeling multivariate asset price dynamics. *Mathematical Finance*, 23(2):217–247, 2013.

J. Kallsen and P. Tankov. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis*, 97(7):1551–1572, 2006.

R. Kawai. A multivariate Lévy process model with linear correlation. *Quantitative Finance*, 9(5):597–606, 2009.

Y.S. Kim, R. Giacometti, S.T. Rachev, F.J. Fabozzi, and D. Mignacca. Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model. *Annals of Operations Research*, 201(1):325–343, 2012.

Y.S. Kim. Portfolio Optimization and Marginal Contribution to Risk on Multivariate Normal Tempered Stable Model. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04613-7, 2022.

T. Kurosaki and Y. S. Kim. Cryptocurrency portfolio optimization with multivariate normal tempered stable process and Foster-Hart risk. *Finance Research Letters*. https://doi.org/10.1016/j.frl.2021.102143, 2021

S. Kring, S.T. Rachev, M. Höchstötter, F.J. Fabozzi, and M.L. Bianchi. Multi-tail generalized elliptical distributions for asset returns. *Econometrics Journal*, 12(2):272–291, 2009.

P. Leoni and W. Schoutens. Multivariate smiling. *Wilmott magazine, March*, 2008.

Li, Q. and Maasoumi, E. and Racine, S.R. A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics*. 148 (2): 186–200, 2009

C. Liu and D.B. Rubin. The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. *Biometrika*, 81(4):633–648, 1994.

A.W. Lo and J. Wang. Trading volume: Definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies*, 13(2):257–300, 2000.

A. Loregian, L. Mercuri, and E. Rroji. Approximation of the variance gamma model with a finite mixture of normals. *Statistics and Probability Letters*, 82(2):217–224, 2012.

E. Luciano and W. Schoutens. A multivariate jump-driven financial asset model. *Quantitative Finance*, 6(5):385–402, 2006.

E. Luciano and P. Semeraro. A generalized normal mean-variance mixture for return processes in finance. *International Journal of Theoretical and Applied Finance*, 13(3): 415–440, 2010a.

E. Luciano and P. Semeraro. Multivariate time changes for Lévy asset models: Characterization and calibration. *Journal of Computational and Applied Mathematics*, 233 (8):1937–1953, 2010b.

E. Luciano, M. Marena, and P. Semeraro. Dependence calibration and portfolio fit with factor-based subordinators. *Quantitative Finance*, 16(7):1037–1052, 2016.

D. Madan, P. Carr and E. Chang. The variance gamma process and option pricing. *Review of Finance*. 2 (1):79–105, 1998.

D.B. Madan. Instantaneous portfolio theory. *Quantitative Finance*, 18(8):1345–1364, 2018.

D.B. Madan and J. Yen. Asset allocation for CARA utility with multivariate Lévy returns. *Handbook of Financial Engineering*, 2004.

R. Marfé. A generalized variance gamma process for financial applications. *Quantitative Finance*, 12(1):75–87, 2012a.

R. Marfé. A multivariate pure-jump model with multi-factorial dependence structure. *International Journal of Theoretical and Applied Finance*, 15(04), 2012b.

A. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: Concepts, techniques, and tools*. Princeton University Press, 2005.

G. Meissner. *Correlation risk modeling and management: An applied guide including the Basel III correlation framework-with interactive models in Excel/VBA*. Wiley, 2014.

L. Mercuri and E. Rroji. Risk parity for mixed tempered dtable distributed sources of risk. *Annals of Operations Research*, 260(1):375–393, 2018.

M. Michaelsen and A. Szimayer. Marginal consistent dependence modelling using weak subordination for Brownian motions. *Quantitative Finance*, 18(11):1909–1925, 2018.

I. Monroe. Processes that can be embedded in Brownian motion. *The Annals of Probability*, 6(1):42–56, 1978.

W.K. Newey and K.D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

K. Prause. *The generalized hyperbolic model: Estimation, financial derivatives, and risk measures*. PhD thesis, Albert-Ludwigs-Universität zu Freiburg im Breisgau, 1999.

R.S. Protassov. EM-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with fixed λ. *Statistics and Computing*, 14(1): 67–77, 2004.

A.W. Rathgeber, J. Stadler, and S. Stöckl. Financial modelling applying multivariate Lévy processes: New insights into estimation and simulation. *Physica A: Statistical Mechanics and its Applications*, 532:121386, 2019.

R. Rebonato and P. Jäckel. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. *Journal of Risk*, 2(2):17–27, 1999.

T. Roncalli. *Handbook of financial risk management*. Chapman and Hall/CRC, 2020.

J. Rosinski. Tempering stable processes. *Stochastic processes and their applications*, 117: 677–707, 2007.

K. Sato. Lévy processes and infinitely divisible distributions. *Cambridge University Press*. Cambridge, 1999.

P. Semeraro. A multivariate Variance Gamma model for financial application. *International Journal of Theoretical and Applied Finance*, 11(8), 2008.

P. Semeraro. A note on the multivariate generalized asymmetric Laplace motion. *Communications in Statistics–Theory and Methods*, 49:2339–2355, 2020.

S. Stoyanov and B. Racheva-Iotova. Numerical methods for stable modeling in financial risk management. In S.T. Rachev, editor, *Handbook of computational and numerical methods in finance*, pages 299–329. Birkhäuser, 2004.

P. Tankov. Lévy copulas: Review of recent results. In M. Podolskij, R. Stelzer, S. Thorbjørnsen, and A.E.D. Veraart, editors, *The fascination of probability, statistics and their applications: In honour of Ole E. Barndorff-Nielsen*, pages 127–151. Springer International Publishing, Cham, 2016.

G.L. Tassinari. *Pricing equity and debt tranches of collateralized funds of hedge fund obligations*. PhD thesis, University of Bergamo, 2009.

G.L. Tassinari and M.L. Bianchi. Calibrating the smile with multivariate time-changed Brownian motion and the Esscher transform. *International Journal of Theoretical and Applied Finance*, 17(4), 2014.

G.L. Tassinari and C. Corradi. Pricing equity and debt tranches of collateralized funds of hedge fund obligations: An approach based on stochastic time change and Esscher-transformed martingale measure. *Quantitative Finance*, 13(12), 2013.

G.L. Tassinari and C. Corradi. Valuation of collateralized funds of hedge fund obligations: A basket option pricing approach. In M Corazza and C. Pizzi, editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 277–288. Springer, Berlin Heidelberg New York, 2014.

Y.C. Wu, S.L. Liao, and S.D. Shyu. Closed-form valuations of basket options using a multivariate normal inverse Gaussian model. *Insurance: Mathematics and Economics*, 44(1):95–102, 2009.

Y. Xia and M. Grabchak. Estimation and simulation for multivariate tempered stable distributions. *Journal of Statistical Computation and Simulation*, 92(3):451–475, 2022.

Moments

In this Appendix we provide the formulas of expected value, variance, skewness, excess kurtosis and correlation of the models described in the paper. Recall that the cumulant of order j of a random variable X with cumulant generating function $\varphi_X(u)$ can be computed as

$$c_j(X) = \frac{\partial^j}{\partial u^j}\varphi_X(u)|_{u=0},$$

and the following equalities hold

$$\begin{aligned} E[X] &= c_1(X), \\ var[X] &= c_2(X), \\ skew[X] &= \frac{c_3(X)}{c_2(X)^{3/2}}, \\ kurt[X] &= 3 + \frac{c_4(X)}{c_2(X)^2}. \end{aligned}$$

MNTS model

$$E\left[Y_{j,1}\right] = \mu_j + E\left[S_1\right]\theta_j,$$

$$var\left[Y_{j,1}\right] = var\left[S_1\right]\left(\theta_j^2 + \frac{\sigma_j^2\lambda}{1-\omega}\right),$$

$$skew\left[Y_{j,1}\right] = skew\left[S_1\right]\left(\theta_j^3 + \frac{3\theta_j\sigma_j^2\lambda}{2-\omega}\right)\left(\theta_j^2 + \frac{\sigma_j^2\lambda}{1-\omega}\right)^{-\frac{3}{2}},$$

$$kurt\left[Y_{j,1}\right] = 3 + \left(kurt\left[S_1\right] - 3\right)\left[\theta_j^4 + \frac{3\sigma_j^2\lambda}{3-\omega}\left(2\theta_j^2 + \frac{\sigma_j^2\lambda}{2-\omega}\right)\right]\left(\theta_j^2 + \frac{\sigma_j^2\lambda}{1-\omega}\right)^{-2},$$

$$corr\left[Y_{j,1}; Y_{k,1}\right] = \frac{\theta_j\theta_k + \frac{\sigma_{jk}\lambda}{1-\omega}}{\sqrt{\left(\theta_j^2 + \frac{\sigma_j^2\lambda}{1-\omega}\right)\left(\theta_k^2 + \frac{\sigma_k^2\lambda}{1-\omega}\right)}},$$

where

$$\begin{aligned} E\left[S_1\right] &= -\omega C\Gamma(-\omega)\lambda^{\omega-1}, \\ var\left[S_1\right] &= \omega(\omega-1)C\Gamma(-\omega)\lambda^{\omega-2}, \\ skew\left[S_1\right] &= (2-\omega)\left[\omega(\omega-1)C\Gamma(-\omega)\lambda^{\omega}\right]^{-\frac{1}{2}}, \\ kurt\left[S_1\right] &= 3 + (\omega-2)(\omega-3)\left[\omega(\omega-1)C\Gamma(-\omega)\lambda^{\omega}\right]^{-1}. \end{aligned}$$

MGH model

$$E\left[Y_{j,1}\right] = \mu_j + E\left[G_1\right]\theta_j,$$
$$var\left[Y_{j,1}\right] = E\left[G_1\right]\sigma_j^2 + var\left[G_1\right]\theta_j^2,$$
$$skew\left[Y_{j,1}\right] = \frac{c_3\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^{3/2}},$$
$$kurt\left[Y_{j,1}\right] = 3 + \frac{c_4\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^2},$$
$$corr\left[Y_{j,1}; Y_{k,1}\right] = \frac{\sigma_{jk} + \theta_j\theta_k\Delta\left(\frac{\chi}{\psi}\right)^{\frac{1}{2}}}{\sqrt{\left[\sigma_j^2 + \theta_j^2\Delta\left(\frac{\chi}{\psi}\right)^{\frac{1}{2}}\right]\left[\sigma_k^2 + \theta_k^2\Delta\left(\frac{\chi}{\psi}\right)^{\frac{1}{2}}\right]}},$$

where

$$c_3\left[Y_{j,1}\right] = 3var\left[G_1\right]\theta_j\sigma_j^2 + c_3\left[G_1\right]\theta_j^3,$$
$$c_4\left[Y_{j,1}\right] = 3var\left[G_1\right]\sigma_j^4 + 6c_3\left[G_1\right]\theta_j^2\sigma_j^2 + c_4\left[G_1\right]\theta_j^4,$$
$$cov\left[Y_{j,1}; Y_{k,1}\right] = E\left[G_1\right]\sigma_{jk} + var\left[G_1\right]\theta_j\theta_k,$$
$$\Delta = \left(\frac{K_{\epsilon+2}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)} - \frac{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)}\right),$$

and with

$$E\left[G_1\right] = \left(\frac{\chi}{\psi}\right)^{\frac{1}{2}}\frac{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)},$$
$$var\left[G_1\right] = \left(\frac{\chi}{\psi}\right)\left[\frac{K_{\epsilon+2}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)} - \left(\frac{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)}\right)^2\right],$$
$$c_3\left[G_1\right] = \left(\frac{\chi}{\psi}\right)^{\frac{3}{2}}\left[\frac{K_{\epsilon+3}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)} - \frac{3K_{\epsilon+2}\left(\sqrt{\chi\psi}\right)K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}^2\left(\sqrt{\chi\psi}\right)} + 2\left(\frac{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)}\right)^3\right],$$
$$c_4\left[G_1\right] = \left(\frac{\chi}{\psi}\right)^2\left[\frac{K_{\epsilon+4}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)} - \frac{4K_{\epsilon+3}\left(\sqrt{\chi\psi}\right)K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}^2\left(\sqrt{\chi\psi}\right)} - 3\left(\frac{K_{\epsilon+2}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)}\right)^2\right] +$$
$$+ 6\left(\frac{\chi}{\psi}\right)^2\left[\frac{2K_{\epsilon+2}\left(\sqrt{\chi\psi}\right)K_{\epsilon+1}^2\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}^3\left(\sqrt{\chi\psi}\right)} - \left(\frac{K_{\epsilon+1}\left(\sqrt{\chi\psi}\right)}{K_{\epsilon}\left(\sqrt{\chi\psi}\right)}\right)^4\right].$$

αGH model

The margins are GH distributed with parameters $(\mu_j, \theta_j, \sigma_j, \chi_j, \psi_j, \epsilon)$, for each $j = 1, .., n$. While marginal moments are as in MGH case, the correlations are given by

$$corr\left[Y_{j,1}; Y_{k,1}\right] = \frac{4a\left(\psi_j\psi_k\right)^{-1}\theta_j\theta_k}{\sqrt{\left[E\left[G_{1,j}\right]\sigma_j^2 + var\left[G_{1,j}\right]\theta_j^2\right]\left[E\left[G_{1,k}\right]\sigma_k^2 + var\left[G_{1,k}\right]\theta_k^2\right]}},$$

where

$$E\left[G_{j,1}\right] = \left(\frac{\chi_j}{\psi_j}\right)^{\frac{1}{2}} \frac{K_{\epsilon+1}\left(\sqrt{\chi_j\psi_j}\right)}{K_\epsilon\left(\sqrt{\chi_j\psi_j}\right)}, \tag{9.1}$$

and

$$var\left[G_{j,1}\right] = \left(\frac{\chi_j}{\psi_j}\right)\left[\frac{K_{\epsilon+2}\left(\sqrt{\chi_j\psi_j}\right)}{K_\epsilon\left(\sqrt{\chi_j\psi_j}\right)} - \left(\frac{K_{\epsilon+1}\left(\sqrt{\chi_j\psi_j}\right)}{K_\epsilon\left(\sqrt{\chi_j\psi_j}\right)}\right)^2\right]. \tag{9.2}$$

$\rho\alpha$GH model

The margins are GH distributed with parameters $(\mu_j, \theta_j, \sigma_j, \chi_j, \psi_j, \epsilon)$, for each $j = 1, .., n$. While marginal moments are as in MGH case, the correlations are given by

$$corr\left[Y_{j,1}; Y_{k,1}\right] = \frac{2a\left(\psi_j\psi_k\right)^{-1}\left(\sigma_j\sigma_k\rho_{jk}\sqrt{\psi_j}\sqrt{\psi_k} + 2\theta_j\theta_k\right)}{\sqrt{\left[E\left[G_{1,j}\right]\sigma_j^2 + var\left[G_{1,j}\right]\theta_j^2\right]\left[E\left[G_{1,k}\right]\sigma_k^2 + var\left[G_{1,k}\right]\theta_k^2\right]}}$$

where $E\left[G_{j,1}\right]$ and $var\left[G_{j,1}\right]$ are as in equations (9.1) and (9.2).

MMixedTS model

$$E\left[Y_{j,1}\right] = \mu_j + \beta_j\frac{c_j + \bar{n}}{m_j},$$

$$var\left[Y_{j,1}\right] = \left(1 + \frac{\beta_j^2}{m_j}\right)\frac{(c_j + \bar{n})}{m_j},$$

$$skew\left[Y_{j,1}\right] = \frac{c_3\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^{3/2}},$$

$$kurt\left[Y_{j,1}\right] = 3 + \frac{c_4\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^2},$$

$$corr\left[Y_{j,1}, Y_{k,1}\right] = \frac{\frac{\beta_j\beta_k}{m_jm_k}\bar{n}}{\sqrt{\left(1 + \frac{\beta_j^2}{m_j}\right)\frac{(c_j+\bar{n})}{m_j}}\sqrt{\left(1 + \frac{\beta_k^2}{m_k}\right)\frac{(c_k+\bar{n})}{m_k}}}.$$

where

$$c_3\left[Y_{j,1}\right] = \left[(2-\alpha_j)\frac{\lambda_{+,j}^{\alpha_j-3} - \lambda_{-,j}^{\alpha_j-3}}{\lambda_{+,j}^{\alpha_j-2} + \lambda_{-,j}^{\alpha_j-2}} + \left(3 + 2\frac{\beta_j^2}{m_j}\right)\frac{\beta_j}{m_j}\right]\frac{(l_j + \bar{n})}{m_j},$$

$$c_4\left[Y_{j,1}\right]=\beta_j^4\left(3+\frac{6}{l_j+\bar{n}}\right)\frac{\left(l_j+\bar{n}\right)^2}{m_j^4}+6\beta_j^2\frac{l_j+\bar{n}}{m_j^3}\left(l_j+\bar{n}+2\right)+$$
$$+4\beta_j\left(2-\alpha_j\right)\left(\frac{\lambda_{+,j}^{\alpha_j-3}-\lambda_{-,j}^{\alpha_j-3}}{\lambda_{+,j}^{\alpha_j-2}+\lambda_{-,j}^{\alpha_j-2}}\right)\frac{l_j+\bar{n}}{m_j^2}+\left(3-\alpha_j\right)\left(2-\alpha_j\right)\left(\frac{\lambda_{+,j}^{\alpha_j-4}+\lambda_{-,j}^{\alpha_j-4}}{\lambda_{+,j}^{\alpha_j-2}+\lambda_{-,j}^{\alpha_j-2}}\right)\frac{l_j+\bar{n}}{m_j}.$$

MGVG model

$$E\left[Y_{j,1}\right]=\mu_j+E\left[\hat{G}_{j,1}\right]\theta_j=\mu_j+\theta_j,$$
$$var\left[Y_{j,1}\right]=E\left[\hat{G}_{j,1}\right]\sigma_j^2+var\left[\hat{G}_{j,1}\right]\theta_j^2=\sigma_j^2+\left(1-k_j\right)q_j\theta_j^2+2k_jp_j\theta_j^2,$$
$$skew\left[Y_{j,1}\right]=\frac{c_3\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^{3/2}},$$
$$kurt\left[Y_{j,1}\right]=3+\frac{c_4\left[Y_{j,1}\right]}{var\left[Y_{j,1}\right]^2},$$
$$corr\left[Y_{j,1};Y_{k,1}\right]=\frac{\sigma_j\sigma_k\rho_{jk}\left(c_1\sqrt{q_jq_k}+c_2\sqrt{p_jp_k}\right)+\theta_j\theta_k\left(c_1q_jq_k+2c_2p_jp_k\right)}{\sqrt{var\left[Y_{j,1}\right]var\left[Y_{k,1}\right]}}$$

where

$$c_3\left[Y_{j,1}\right]=3var\left[\hat{G}_{j,1}\right]\theta_j\sigma_j^2+c_3\left[\hat{G}_{j,1}\right]\theta_j^3$$
$$c_4\left[Y_{j,1}\right]=3var\left[\hat{G}_{j,1}\right]\sigma_j^4+6c_3\left[\hat{G}_{j,1}\right]\theta_j^2\sigma_j^2+c_4\left[\hat{G}_{j,1}\right]\theta_j^4$$
$$cov\left[Y_{j,1};Y_{k,1}\right]=E\left[\hat{G}_{j,1}\right]\sigma_{jk}+var\left[\hat{G}_{j,1}\right]\theta_j\theta_k,$$

and with

$$E\left[\hat{G}_{j,1}\right]=1,$$
$$var\left[\hat{G}_{j,1}\right]=\left(1-k_j\right)q_j+2k_jp_j,$$
$$c_3\left[\hat{G}_{j,1}\right]=2(1-k_j)q_j^2+6k_jp_j^2,$$
$$c_4\left[\hat{G}_{j,1}\right]=6\left(1-k_j\right)q_j^3+24k_jp_j^3.$$