MODELLING VOLATILITY AND CORRELATIONS WITH A HIDDEN MARKOV DECISION TREE.

PHILIPPE CHARLOT

GREQAM & Aix-Marseille University
Centre de la Vieille Charité
2, rue de la Charité
13236 Marseille cedex 02
France

VERY PRELIMINARY DRAFT – COMMENTS ARE MOST WELCOME!

ABSTRACT. The goal of the present paper is to present a new multivariate GARCH model with time-varying conditional correlation. Since the seminal work of Bollerslev (1990), conditional correlation models have become a attractive field in economics. Different specifications have been developed to study both empirical findings and practical use like asymmetry, change in regime but also estimation of large correlation matrix (see, e.g. Silvennoinen and Teräsvirta (2009) for a survey of recent advances). Among this field of research, our work focus on change in regime specification based on tree structure.

Indeed, tree-structured dynamic correlation models has been developed to analyse volatility and co-volatility asymmetries (see Dellaportas and Vrontos (2007)) or linking the dynamics of the individual volatilities with the dynamics of the correlations (see Audrino and Trojani (2006)). The common approach of these models is to partitioning the space of time series recursively using binary decisions. This can be interpreted as a deterministic decision tree. At the opposite, the approach that we adopt for this paper is developed around the idea of hierarchical architecture with a Markov temporal structure. Our model is based on an extension of Hidden Markov Model (HMM) introduced by Jordan, Ghahramani, and Saul (1997). It is a factorial and coupled HMM.

Hence, our model is based is a stochastic decision tree liking the dynamics of univariate volatility with the dynamics of the correlations. It can be view as a HMM which is both factorial and dependent coupled. The factorial decomposition provides a factorized state space. This state space decomposition is done using state dependent and time-varying transition probabilities given an input variable. The top level of the tree can be seen as a *master* process and the following levels as *slave* processes. The constraint of a level on the following is done via a coupling transition matrix which produce the ordered hierarchy of the structure. As the links between decision states are driven with Markovian dynamics and the switch from one level to the following is done via a coupling transition matrix, this architecture gives a fully probabilistic decision tree. Estimation is done in one step using maximum likelihood.

We also perform an empirical analysis of our model using real financial time series. Results show that our hidden tree-structured model can be an interesting alternative to deterministic decision tree.
Keywords: Multivariate GARCH; Dynamic correlations; Regime switching; Hidden Markov Decision Trees.
JEL Classification: C32, C51, G1, G0.

E-mail address: ph.charlot@gmail.com.
Date: this version january 2011 (first draft april 2010).

1. Introduction

The past few years, multivariate volatility models have been a field of very active research. Related to this topic, the study of multivariate GARCH with conditional correlations has attracted a strong interest since the seminal paper of Bollerslev (1990). Introducing dynamic in the conditional correlations has led this pioneering approach to creating a new class of model, the well-known Dynamic Conditional Correlation (DCC) models. Proposed by Engle and Sheppard (2001) and by a slightly different manner by Tse and Tsui (2002), the DCC becomes an attractive parametrization.

This new class now includes many extensions. Engle (2002) generalized the model by introducing full matrix instead of scalars to drive the dynamics of the process (Generalized DCC). Therefore, this unrestricted specification requires a large number of parameters and make estimation unfeasible. To solve the curse of dimensionality while relaxing constraints on the dynamics, various extensions have been proposed. Most of them are based on combinations of specific matrix, partitioned vectors or block matrix (see Billio and Caporin (2009), Hafner and Franses (2009) and Billio, Caporin, and Gobbo (2006)) but also clustering techniques (see Zhou and Chan (2008)). A DCC incorporate exogenous variables is proposed by Vargas (2008) to identify factors that could lead correlations. A semi-parametric approach is presented by Franses, Hafner, and van Dijk (2005), using a transformed Nadaraya-Watson estimator. Feng (2007) also presents a local estimator for the correlations based on k-nearest-neighbors (k-NN) methods.

Others extensions focus on the asymmetric effect in the conditional correlation. Asymmetry refers to the upward motion of volatility more after a negative than after a positive shock. An asymmetric extension has been proposed by Cappiello, Engle, and Sheppard (2006) to study worldwide linkages in the dynamics of the correlations of selected bonds and equity markets. Vargas (2006) extends the Block DCC of Billio, Caporin, and Gobbo (2003) to take account of the asymmetry between blocks of asset returns. Cajigas and Urga (2006) also propose an asymmetric generalized DCC where standardized residuals follow an asymmetric multivariate Laplace distribution.

Another direction of extensions refers to introduce regime switch in the conditional correlations to take into account empirical findings in case of turbulent periods. In this context, Silvennoinen and Teräsvirta (2005) assume the conditional correlation matrix varies smoothly between two constant correlations matrix. The link between these two extreme matrix is done via a conditional logistic function according to a exogenous or endogenous transition variable. This model has been improved to allows another transition around the first one (see Silvennoinen and TerÃď svirta (2009)). Differently from the previous STAR approach, Pelletier (2006) suggested to introduce Markov switch to the correlation. In this way, correlations are constants within regime but vary from one regime to another. Both previous approach assume correlations evolve between constant correlations matrix. Billio and Caporin (2005) explained a different strategy with an extension of the scalar DCC of Engle and Sheppard (2001) where both the parameters and the unconditional correlation are driven by a hidden Markov chain.

Alongside these developments, the technical issue of the curse of dimensionality, which underlies most of the developments in the field of multivariate GARCH, has been the subject of recent advances. This problem of the explosion in the number of parameters arises quickly: with k time series, the numbers of correlation to estimate is $k(k-1)/2$. Several strategies have been proposed. Palandri (2009) decomposes the correlations and partial correlations. This decomposition transforms a high dimensional optimization problem into a set of simple estimates. Engle, Shephard, and Sheppard (2008) also exploited the idea of decompose a complex estimation problem into a set of simpler models to construct a composite likelihood which is obtained from the summation of likelihood of subsets of assets. Engle and Kelly (2009) suggests a radically different strategy based on an updated approximation of the correlation. At each time, they assumes all pairwise correlations are equal. Engle (2008) also exposes an alternate solution based on approximation. This aptly named *MacGyver* method is based on the bivariate estimation of each pair and then calculate the median of theses estimators.

This brief summary of the literature is far from exhaustive and we refer to Bauwens, Laurent, and Rombouts (2006) for a general survey on multivariate GARCH. The paper of Silvennoinen and Teräsvirta (2009) gives an useful review of recent advances and Engle (2009) outlines a thorough presentation of conditional correlations models.

The model we present is an extension of multivariate GARCH with conditional correlations based on a tree-structure. Models based on tree structure have already been proposed. Barone-Adesi and Audrino (2006) apply a tree structure developed by Audrino and Buhlmann (2001) and Audrino and Trojani (2006) to a rolling window averaged conditional correlation estimator. This method allows thresholds in the conditional correlations. The estimator is a convex combination of realized correlations and estimates of averaged correlations. As the averaged conditional correlation is constructed with univariate GARCH models, this approach has the advantage to required few parameters. This study has been extended using non-parametric functions based on functional gradient descent to estimate univariate volatilities by Audrino (2006). Also based on the idea of Audrino and Buhlmann (2001), the model suggested by Trojani and Audrino (2005) adopts the two-step procedure exposed in Engle and Sheppard (2001). In first time, conditional variance are extracted with tree-structured GARCH. Conditional correlations are then computed from standardized residuals and have also their own tree-structured dynamic. An tree-structured extension of the DCC is suggested by Dellaportas and Vrontos (2007) to study volatility and co-volatility asymmetries. Making a link between conditional variances and correlations, this model allows to exhibit multivariate thresholds. Using a sequence of binary decisions rules, each terminal node is matched with a multivariate GARCH.

One common feature of theses approaches is they are all based on the idea of binary tree. Time series are recursively partitioned using binary decisions. The resulting tree is purely deterministic. At the opposite, in this paper, we propose en extension of the DCC based on a stochastic decision tree. Our model links the univariate volatilities with the correlations via a hidden stochastic decision tree. Hidden Markov Decision Tree (HMDT) model is an extension of Hidden Markov Model (HMM) introduced by Jordan, Ghahramani, and Saul (1997). Based on a Markov temporal structure, the architecture is at the opposite of the classical deterministic approach on a binary decision tree. Our proposed models can be estimated via maximum likelihood.

The paper is organized as follows. The model is defined section. We first briefly recall the basics of the DCC of Engle and Sheppard (2001) and then outline our model. Section the model is applied to data examples. In the first application, we apply our model to exchange rate data with in-sample and out-of-sample analysis. Final remarks concludes and proposes areas for further research in section.

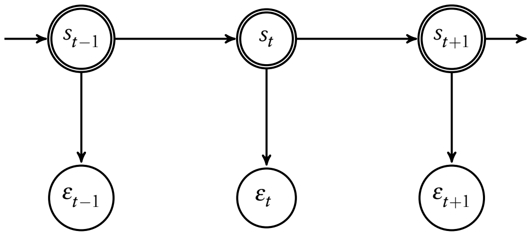
2. The Model

2.1. **What is a Hidden Markov Decision Tree?**

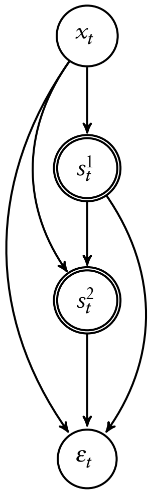
The approach that we adopt for this paper is developed around the idea of hierarchical architecture with a Markov temporal structure. This model can be explained in one of two ways.

2.1.1. *A Hierarchical Mixtures-of-Experts with Markov dependencies.*

This approach can be first interpreted as a hybrid model that blends the Hierarchical Mixtures-of-Experts (HME) of Jordan and Jacobs (1994) and Hidden Markov models (HMM). The HME is based on the *divide-and-conquer* algorithm. This methodology is to fit the data by dividing the space into nested subspaces and estimate a sub-model for each of these regions (see figure 1(a) for the representation of HME as a directed acyclic graph). The data are fitted by a model which is the overall combination of each submodels. This philosophy has given rise to numerous approaches as, without being exhaustive the CART (see Breiman, Friedman, Olshen, and Stone (1984)), MARS (see Friedman (1991)) or GUIDE (see Loh (2002)). HME is itself an generalization of the Mixture-of-Experts (ME) models of Jacobs, Jordan, Nowlan, and Hinton (1991). It assumes that the data can be fitted by a piecewise function. The approach is based on two elements. First, the models has *n experts* to fit the data in *n* regions. An expert can be a constant, AR or GARCH process. The second element is a *gating network*, that associate to each region with a expert given an input variable. The final output is is a convex combination of the expert outputs given the input. The ME architecture has been applied to time series by Carvalho and Tanner (2006). This architecture has only one level of experts associated with one gating network. The HME generalize this approach by allowing different level of experts matched by various gating functions. Huerta, Jiang, and Tanner (2003) use a two layers HME to study the US industrial production index with *trend-stationary process experts* and *difference-stationary process experts*. As explained in McLachlan and Peel (2000), the main difference between CART/MARS/GUIDE approaches and HME divide the covariate space with probabilistic scheme



(a) HIDDEN MARKOV MODEL.



(b) HIERARCHICAL MIXTURE OF EXPERTS.

Figure 1: The hidden Markov model and the hierarchical mixture of experts as graphical models.

whereas the first use hard boundaries. It also may be pointing out of the link between ME architecture and (S)TAR models. Indeed, in a STAR-GARCH model the transition function can be interpreted as the gating network and the two GARCH component as the experts. The same comparison is true for the HME and the multiple-regime STAR of van Dijk and Franses (1999).

Hidden Markov models has emerged as a standard tools to modeling time series with regime switching. This approach has been extensively used, whether in graphical models (see Rabiner (1989)) or in econometrics with its counterpart Markov-switching (see among others Hamilton (1994). HMM is based on the assumption that each observation is linked to a finite number of hidden states via a probability distribution (see figure 1(b)). The conditional probability distribution of the hidden variable generally assumes to be a first order Markov chain. In time series analysis, this has first been used by Hamilton (1989) to make a regime switching AR process and generalized by Krolzig (1997) to the VAR case. Example of applications to GARCH models can be found in Hamilton and Susmel (1994), Gray (1996), Dueker (1997), Klaassen (2002) and Haas, Mittnik, and Paollela (2004). Markov-switching based extensions of multivariate GARCH models with dynamic correlations have also been proposed. Billio and Caporin (2005) propose a DCC model with Markov switch in the unconditional correlation process. Application of this MS-DCC model to daily stock market indices from January 2000 to December 2003 highlights the advantage of introducing Markov jumps during turbulent periods. Pelletier (2006) also suggested a Markov-switching DCC. This regime switching for dynamic correlations (RSDC) model assumes that correlations are constant within regime but vary from one regime to another.

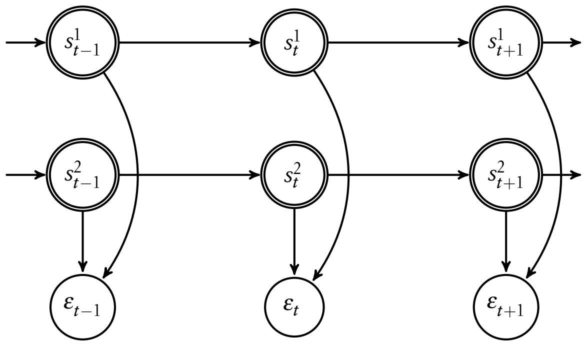
The HMDT can be seen as a hybrid model combining the two previous approach. The result gives a probabilistic decision tree where the decision at each time is conditional of the decision at the previous time.

2.1.2. *A factorial and coupled HMM.* In this paper, we adopt another look of the HMDT. Contrary to the previous approach where the HMDT is defined as a hierarchical mixture with Markovian dynamics, we can also consider this architecture as a pure HMM extension. Indeed, HMDT can be view as both a factorial HMM and a coupled HMM.

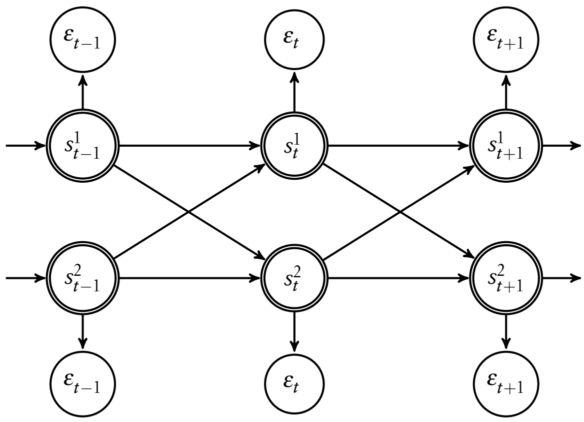
The factorial hidden Markov (FHMM) model has been proposed by Ghahramani and Jordan (1997). This extension of the basic HMM assumes the state variable is factored into various state variables. Indeed, each of these has its independent Markovian dynamic and the output is the combination of these several underlying processes (see figure 2(a)) . Because the model has various independent hidden Markov models in parallel, the resulting state space of the model is the Cartesian product of the parallel sub-processes. The FHMM has been widely used in speech recognition because of its ability to model time series generated from various independent sources (see among others Roweis (2000) for an application to microphone source separation). The distributed state representation of the FHMM also provide an interesting framework for handwriting recognition (see Williams, Toussaint, and Storkey (2007) for the use of a primitive model based on FHMM to represent natural handwriting data).

Coupled hidden Markov models (CHMM) are also an extension of the classical HMM. It consists of a model with several HMMs whose Markov chains interact together (see Brand (1997)). There are two ways to allow the states to have interdependencies. We can qualify as *independent coupling* when various processes are coupled at the output. From this point of view, FHMM can be considered as an independent coupled HMM. Another kind of making interactions between chains is to assume that processes are dependent. In *dependent coupling*, the current state is dependent of the other states of its own chain but also depends on the previous state of another chain and so on (see figure 2(b)). Dependent coupled hidden Markov models include many architecture, as models with different lag, nested loop and non-linear dependency relationship. Coupled HMM have been used in various fields like, without being exhaustive, facial event mining (see Ma, Zhou, Celenk, and Chelberg (2004)), video-realistic speech animation (see Xie and Liu (2007)), audio-visual speech recognition (see Nefian, Liang, Pi, Xiaoxiang, Mao, and Murphy (2002)).

In this framework, HMDT can be view as a HMM which is both factorial and dependent coupled. The factorial decomposition provides a factorized state space. This state space decomposition is done using constant or state dependent time-varying transition probabilities given an input variable. As described in Brand (1997), the top level of the tree can be seen as a *master* process and the following levels as *slave* processes. The constraint of a level on the following is done via a coupling



(a) Factorial hidden Markov model.



(b) Coupled hidden Markov model.

Figure 2: Factorial and Coupled HMM as graphical models.

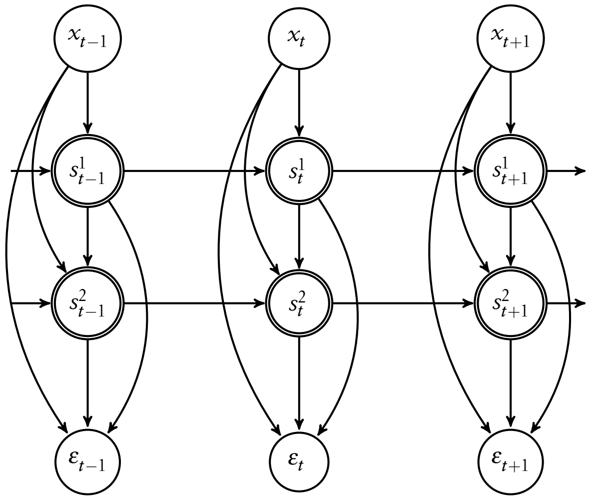


Figure 3: The hidden Markov decision tree model views as a directed acyclic graph.

transition matrix which produce the ordered hierarchy of the structure. As the links between decision states are driven with Markovian dynamics and the switch from one level to the following is done via a coupling transition matrix, this architecture gives a fully probabilistic decision tree.

2.2. **Hidden Markov decision tree for correlations.**

2.2.1. *Starting point.* The model we propose in this paper is an extension of the DCC of the scalar DCC (see Engle and Sheppard (2001)). Given y_t an K dimensional time series of length T the DCC class models assumes:

$$y_t|\mathscr{F}_{t-1} \overset{iid}{\sim} \mathscr{N}(0, H_t) \tag{2.1}$$

where $\mathscr{F}_{t-1}$ refers to the information set at time $t-1$. Conditional variance-covariance matrix of returns y_t is expressed as follows:

$$H_t = D_t R_t D_t \tag{2.2}$$

where R_t is a $K \times K$ constant correlation matrix. the matrix D_t is a $K \times K$ diagonal matrix containing univariate time varying standard deviations:

$$D_t = \text{diag}\{h^{1/2}_{i,t}\} \tag{2.3}$$

for $i = 1, ..., K$. Getting the matrix D_t is generally referenced as the so-called *degarching* filtration required to construct standardized residuals expressed as:

$$\varepsilon_t = D_t^{-1} r_t \tag{2.4}$$

The conditional correlations are simply the expectation of the standardized residuals:

$$\mathbb{E}_{t-1}[\varepsilon_t \varepsilon_t'] = D_t^{-1} H_t D_t^{-1} = R_t \tag{2.5}$$

In the scalar DCC of Engle and Sheppard (2001), correlations are computed using a combination of the conditional covariance of the standardized residuals $Q - t$ which follows a scalar BEKK process:

$$Q_t = (1-a-b)\overline{V} + a\varepsilon_{t-1}\varepsilon_{t-1}' + bQ_{t-1} \tag{2.6}$$

where $\overline{Q}$ is the unconditional covariance matrix of the standardized residuals and a and b are two positive scalars such that $a + b < 1$. The conditional correlation matrix is then obtained by:

$$R_t = \text{diag}\{Q_t\}^{-1/2} Q_t \text{diag}\{Q_t\}^{-1/2} \tag{2.7}$$

Our model has the DCC of Engle and Sheppard (2001) as starting point. We propose to improve this specification by introducing a hidden decision. To study the relationship between univariate volatility and correlation, this tree has two levels. The first established a splitting rule to discriminate low and high regime of variance while the second level discriminates low and high correlations regime.

Thus, this construction is a suitable method to investigate the link between high (low) volatility and high (low) correlations. Specifically, it allows us to study the probability of being in a type regime of correlations given the regime of the univariate conditional variance.

2.2.2. *Hidden tree structure.* As point out in section 2.1.2, a HMDT can be view as a factorial and coupled HMM. The factorial HMM decomposition is used to divide the space of the time series into low/high conditional variance for each series and for low/high sequence of correlations. Transition matrix used in our model can be both static or time-varying and can depend of endogenous or exogenous variables.

Formally, at the first level, each time series have its own conditional variance driven by a transition matrix with time varying probabilities. Univariate high and low conditional variance follow a two state Markov process of first order such. It means the state s_t is independent of state lagged state except s_{t-1}:

$$\mathbb{P}[s_{t+1} = i | s_t = j, ..., s_t = h] = \mathbb{P}[s_{t+1} = i | s_t = j] \tag{2.8}$$

To introduce time-varying transition probabilities, we can use the specification proposed by Diebold, Lee, and Weinbach (1994). In that way, transition probabilities are assumed to follow a logistic function of a endogenous or exogenous variable. Elements of the transition matrix $\mathrm{P}^{\mathrm{k}}_{\mathrm{vol}}$ for the conditional variance of the k^{th} with row i and column j is defined as:

$$p^k_{ij}(t) = \mathbb{P}[s_{t+1} = j | s_t = i; x_{t-1}, \gamma^k_n], \; i, j = 1, .., N \tag{2.9}$$

Introducing time-variation, transition probabilities are then expressed as:

$$p^k_{ij}(t) = \frac{\exp(x'_{t-1}\gamma^k_n)}{1 + \exp(x'_{t-1}\gamma^k_n)} \tag{2.10}$$

with x_{t-1} is the conditioning vector referencing as the input of figure 3 and γ its coefficient. As we partition the space of the univariate conditional variance in two subspaces, low and high variance, the transition matrix for the variance of the k^{th} time serie is of size 2×2 and can be expressed as follow:

$$\mathrm{P}^{\mathrm{k}}_{\mathrm{vol}} = \begin{bmatrix} p^k_{11}(t) & 1 - p^k_{22}(t) \\ 1 - p^k_{11}(t) & p^k_{22}(t) \end{bmatrix} \tag{2.11}$$

Since all univariate volatility processes have the same Markovian dynamic specification, individual transition matrix can be aggregate to constitute the first level of the decision tree. Individual HMMs are aggregated in a factorial HMM representation. The dynamic of the univariate volatility level can be resumed in a general transition $\mathrm{P}_{\mathrm{vol}}$ by the cross product of the transition matrix of univariate volatility models $\mathrm{P}^{\mathrm{k}}_{\mathrm{vol}}$:

$$\mathrm{P}_{\mathrm{vol}} = \bigotimes_{i=1}^{K} \mathrm{P}^{\mathrm{i}}_{\mathrm{vol}} \tag{2.12}$$

This factorial representation allows to represent all the dynamics of the K univariate volatilities containing 2 states with a single transition matrix of size $2^K \times 2^K$.

The same specification is used for the second level. This level discriminates between low and high correlations. Thus, the decision step is represented by a 2-by-2 transition matrix written as:

$$\mathrm{P}_{\mathrm{corr}} = \begin{bmatrix} p^c_{11}(t) & 1 - p^c_{22}(t) \\ 1 - p^c_{11}(t) & p^c_{22}(t) \end{bmatrix} \tag{2.13}$$

whose elements can be both static and time-varying. In that case, this is the same specification as defined in the equation (2.10). Given the transition matrix of the first and second level, the partition of the space is represented by transition matrix P expressed as:

$$\mathrm{P} = \mathrm{P}_{\mathrm{vol}} \otimes \mathrm{P}_{\mathrm{corr}} \tag{2.14}$$

of size $2^{K+1} \times 2^{K+1}$.

Given this space partition the first and the second level are then linked. This relation is fully probabilistic and attribute a weight to the decision related to the correlation given the decision of the univariate volatility. This relation is done using a *coupling* matrix. Coupling matrix are generally used in the context of coupled HMMs where a set of Markov chain influence each other. In our purpose, the matrix we use verifies the Markov property. Thus, the dynamic of the junction between the two levels is done with an abstract Markov chain. This linking relation does not emit directly observation but plays the part of attributing a weight between volatility and correlations. This coupling matrix is of size 2-by2 and is written as:

$$\mathrm{P}_{\text{coupl}} = \begin{bmatrix} c_{11}^{(}t) & 1-c_{22}^{(}t) \\ 1-c_{11}^{(}t) & c_{22}^{(}t) \end{bmatrix} \tag{2.15}$$

The elements of this matrix can also be time-varying given the input and are defined with the logistic specification of the equation (2.10).

Thereby, the dynamic of the probabilistic tree can be can be summarized as follow. All the ordered hierarchy is based on a set of HMMs which make the partition of the space. This partition is slightly different from that usually considered in binary tree models where the space is partitioned recursively using binary rules. In our model, the partition is static and is defined a priori. The decision process occurs according to a cascade of ordered synchronous HMMs. The coupling process has not the Markovian property but can be interpreted as vector of probabilities, whose sum is equal to one. By making the link between the HMMs, the coupling matrix captures interprocess influences. Conditional time-varying transition and coupled probabilities allow the structure to adapt for each observation.

In our model, we use a set of HMMs to the coupling vector is used to weight influence of the univariate volatility on the correlations. The first level discriminates each time serie between low and high volatility while the second level ordered between high and low correlations. The input variable can be both endogenous or exogenous. A consensual choice could be time calendar.

2.2.3. *Specification for univariate volatilities.* In our decision tree, the first level distinguishes between low and high volatility. Thus, this decision step parametrizes a univariate regime switching GARCH model. Among the literature of time series, many specifications have been proposed. Most of them have the drawback to require approximation schemes to avoid the path dependency problem. This is the case of the models suggested by Gray (1996) and Klaassen (2002). Beyond this numerical aspect, specification with approximations can entail difficulties in the interpretation of the processes corresponding to each regime. This motivation led us to use the Markov-switching GARCH proposed by Haas, Mittnik, and Paollela (2004). This model is expressed as follow:

$$\begin{pmatrix} h_{1,t} \\ \vdots \\ h_{N,t} \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_N \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} y_{t-1}^2 + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix} \odot \begin{pmatrix} h_{1,t-1} \\ \vdots \\ h_{N,t-1} \end{pmatrix} \tag{2.16}$$

where $\odot$ means the element-by-element multiplication. Stationarity condition implies $\alpha_n + \beta_n < 1$ for each $n = 1, ..., N$. Besides its computational advantages, this model has a clear cut interpretation. It assumes the conditional variance can switch between N separate GARCH models that evolve in parallel. Because we need two regimes, each time serie involves eight parameters.

2.2.4. *Specification for correlations.* The second level discriminates between low and high conditional correlations. Similarly with the first level, we use a specification which does not require approximation in order to facilitate the interpretation of the regimes. Haas and Mittnik (2008) have proposed an extension of the Markov-switching GARCH to the multivariate case. Since the conditional covariances of the standardized residuals has a BEKK dynamic, we can use this Markov-switching extension in the scalar case. It can be written as:

$$\begin{pmatrix} Q_{1,t} \\ \vdots \\ Q_{N,t} \end{pmatrix} = \begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_N \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \varepsilon_{t-1}\varepsilon_{t-1}' + \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} \odot \begin{pmatrix} Q_{1,t-1} \\ \vdots \\ Q_{N,t-1} \end{pmatrix} \tag{2.17}$$

Stationarity conditions imply the intercept Ω_n to be a positive definite matrix and $a_n + b_n < 1$ for $n = 1, ...N$. This specification requires $K(K+1)/2 + 4$ parameters.

2.3. **Estimation.** Estimation of the model is done with maximum likelihood. One of the advantages of the DCC of Engle (2002) is the possibility to estimate the model in a two steps approach. The idea is to estimate the parameters related to the univariate volatilities in a first step and in a second step whose related to the correlations. In our case, because of the relation between the degarching process and the correlation, we must estimate the model in one step. With the assumption of normality, the log-likelihood can be written:

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left(K\log(2\pi) + \log(|H_t|) + y_t' H_t^{-1} y_t\right) \tag{2.18}$$

To maximize the likelihood we need to make inference on the state of the various Markov chains of the model. The strategy developed in our paper is to convert the complex dynamic of the factorial and coupled HMM into a simple regular HMM so that we can use standards tools on filtering and smoothing probabilities. A conversion relationship between coupled HMM and standard HMM has been proposed by Brand (1997). Given $\mathrm{P}_{\mathrm{S|S}}$ and $\mathrm{P}_{\mathrm{S'|S'}}$ two transition matrix, $\mathrm{P}_{\mathrm{S|S'}}$ and $\mathrm{P}_{\mathrm{S'|S}}$ two coupling matrix, the relationship of Brand (1997) is given by:

$$(\mathrm{P}_{\mathrm{S|S}} \otimes \mathrm{P}_{\mathrm{S'|S'}}).R(\mathrm{P}_{\mathrm{S'|S}} \otimes \mathrm{P}_{\mathrm{S|S'}}) \tag{2.19}$$

where R is a row permute operator swaping fast and slow indices. In our case, because we use only one coupling vector in a downward direction, the regular HMM representation $\mathrm{P}_{\mathrm{reg}}$ of our model can be expressed as:

$$\mathrm{P}_{\mathrm{reg}} = (\mathrm{P}_{\mathrm{vol}} \otimes \mathrm{P}_{\mathrm{corr}}) \odot (\mathrm{P}_{\mathrm{coupl}} \otimes (\iota\iota')) \tag{2.20}$$

with ι a vector of ones of length 2^{K+1}.

The conversion into regular HMM allows to use standard tools about the inference of the Markov chains. Let ξ_{jt} the probability to be in regime j given the information set available in $t-1$ and η_{jt} the density under the regime j. The probability to be in each regime at time t given the observations set up to t, written $\hat{\xi}_{t|t}$, can be calculated using the following expression of the so-called Hamilton's filter:

$$\hat{\xi}_{t|t} = \frac{(\hat{\xi}_{t|t-1} \odot \eta_t)}{\iota'(\hat{\xi}_{t|t-1} \odot \eta_t)} \tag{2.21}$$

and:

$$\hat{\xi}_{t|t+1} = \mathrm{P}_{\mathrm{reg}} \times \hat{\xi}_{t|t} \tag{2.22}$$

where $\odot$ denotes element by-element multiplication. The conversion to regular HMM has the advantage to be easy to compute. Interpretation is also natural. Each state of the regular representation corresponds of a combination of possible cases, e.g. the first series in high or low volatility period, the second in high or low, etc. and the correlation in low and high period. The main drawback of the conversion is the complexity of the maximum likelihood estimation. Optimization methods based on gradient methods using numerical derivatives can have difficulties to escape from local maxima. That therefore justifies to test several initial conditions.

3. Application

For the first application, we propose to illustrate our proposed model on a daily dataset. We consider futures prices on 10 year treasury bonds and the S&P 500 from September 1994 to February 2003 (2201 observations). The two series comes from DataStream and are plotted in figure 4. Returns are calculated by taking 100 times the first difference of the logarithm of each series minus the sample mean. Descriptives statistics of the log-returns are given in table 1. All the results reported in the analysis were generated with Matlab on Linux.

With two series, the dynamic of the univariate volatility level can be summarized in a 4-by-4 transition matrix (see equation (2.12)). The second level has a 2-by-2 transition matrix. Thus, given equation (2.14), the space is partitioned into eight sub-spaces, each corresponding of a special case. Let $h_{t,Y}$ and $h_{t,P}$ the univariate volatilities of the bonds and the S&P 500. The superscript h denotes a high regime of volatility and l a low regime. A high regime of correlations is written by R_t^h and a low by R_t^l. The enumeration of the various cases corresponds to:

(1) $\{h_{t,Y}^h, h_{t,P}^h\}, \{R_t^h\}$

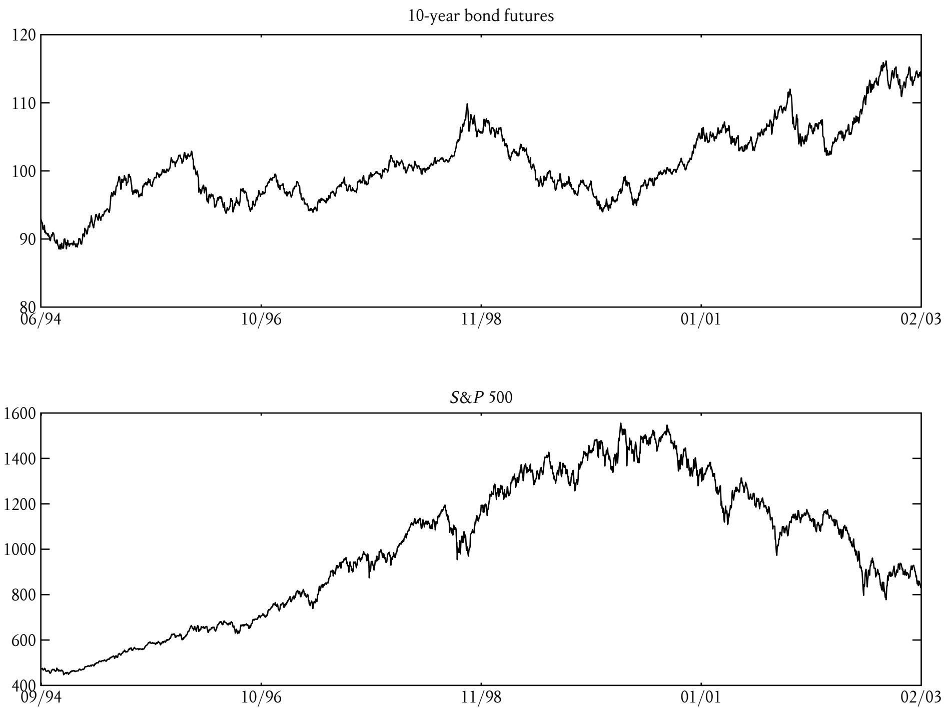


Figure 4: Futures prices of 10-year bond and S&P500 1994 to July 2003.

	Bond $	S&P500
Minimum	-2.2154	-7.7621
Maximum	1.4175	5.7549
Mean	0.0094	0.0258
Variance	0.1430	1.5191
Skewness	-0.5310	-0.1236
Kurtosis	5.4602	6.5324
Jarque-Bera	658.2178 (1.0000e−03)	1.1494$e+$03 (1.0000e−03)
Box-Ljung Q(25)	30.9139 (0.1919)	29.8141 (0.2313)
Box-Ljung Q(50)	65.4385 (0.0703)	64.6047 (0.0802)

Table 1: Descriptive statistics for the 10 year treasury bonds and the S&P 500 futures prices.

(2) $\{h^h_{t,Y}, h^h_{t,P}\}, \{R^l_t\}$
(3) $\{h^h_{t,Y}, h^l_{t,P}\}, \{R^h_t\}$
(4) $\{h^h_{t,Y}, h^l_{t,P}\}, \{R^l_t\}$
(5) $\{h^l_{t,Y}, h^h_{t,P}\}, \{R^h_t\}$
(6) $\{h^l_{t,Y}, h^h_{t,P}\}, \{R^l_t\}$
(7) $\{h^l_{t,Y}, h^l_{t,P}\}, \{R^h_t\}$
(8) $\{h^l_{t,Y}, h^l_{t,P}\}, \{R^l_t\}$

With two series, the decision can be summarized as in figure 5. The coupling matrix is then of size 2×2. This coupling matrix links each state $\{h^i_{t,Y}, h^j_{t,P}\}$, $i, j = h, l$ with a state $\{R^i_t\}$, $i = h, l$. The relation introduced by this vector can be view as a weighted function.

With two series, the number of parameters related to the first level is 8 and 12 for the correlations. With 2 parameters for the coupling matrix, the total number of parameters to estimate is 30. As explained in the previous section, all the decisions of the first level is resumed in a transition matrix which is the cross product of the two transition matrix related to each serie.

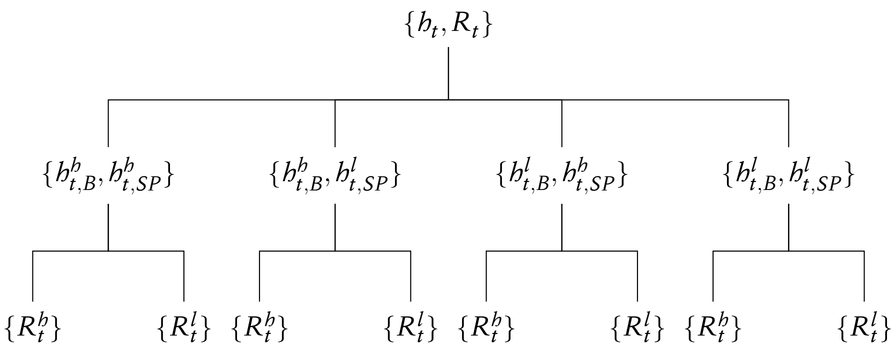


Figure 5: Structure of the decision process for a bivariate dataset.

3.1. **Results of the model.** We first apply our model with constant transition probabilities. This configuration has no input. Estimated correlations are plotted on figure 6(a). Our results confirm those of Colacito and Engle (2006), who had previously studied this dataset. The correlation are positive in the first part of the sample and becomes negative at this end.

The figure 6(b) shows the smoothed probabilities of the regimes selected by the decision process. Contrary to deterministic decision tree, interpretation is far from obvious. This is due to the fact that deterministic tree attribute a model for each partition of the space. Because of the threshold classification, each partition is represented. In our case, we discriminate among various regime. That why the smoothed probabilities of the figure 6(b) represents the combination of regime that fit the data.

The model clearly identifies three combinations of regimes. The identification of the nature of regimes can be done using the table 2 of estimated parameters. We first note that the 10-year bond futures serie does not seem to have a change of regime. The following interpretation of theses combinations are, by chronological order:

(1) $\{h^2_{t,Y}, h^1_{t,P}, R^1_t\}$: normal volatility for bond, low volatility for S&P500 and positive correlations regime.
(2) $\{h^2_{t,Y}, h^2_{t,P}, R^1_t\}$: normal volatility for bond, high volatility for S&P500 and positive correlations regime.
(3) $\{h^2_{t,Y}, h^1_{t,P}, R^2_t\}$: normal volatility for bond, low volatility for S&P500 and negative correlations regime.

The results of the estimated transition probability matrices clearly shows that the first regime are persistent for the bond. Our model does not identify low and high volatility period. This is not the case of the index where the model identifies two regimes. This is also the same for the correlations and the first part of the sample is clearly a regime of positive correlations while and the second part has negative correlations.

The estimated results of the coupling transition matrix gives us the relationship between the first and the second level. This 2-by-2 transition matrix with constant probabilities is written as:

$$\mathrm{P}_{\text{coupl}} = \begin{bmatrix} c_{11} & 1-c_{22} \\ 1-c_{11} & c_{22} \end{bmatrix} \tag{3.1}$$

where the first state link the first state link the volatility level with the positive correlation regime and the second rely the first level with the negative correlation regime. Estimated parameters are:

$$\widehat{\mathrm{P}}_{\text{coupl}} = \begin{bmatrix} 0.2522 & 0.1371 \\ 0.7478 & 0.8629 \end{bmatrix} \tag{3.2}$$

First level				
	10-year bond futures		S&P500	
Parameters	Estimate	standard errors	Estimate	standard errors
ω_1	0.0027	4.07e-6	0.0082	5.50e-6
α_1	0.0434	1.66e-4	0.0694	1.92e-4
β_1	0.9394	5.45e-4	0.9263	1.91e-4
ω_2	0.0342	0.0101	0.0157	6.48e-5
α_2	0.1652	0.3214	0.0911	0.0013
β_2	0.8201	0.3342	0.8992	0.001
p_{11}	0.9981	1.50e-4	0.9998	5.08
p_{22}	0.0010	7.69e-8	0.6522	0.0154
Coupling matrix				
Parameters	Estimate	standard errors		
p_{11}	0.2522	0.0584		
p_{22}	0.8629	0.0049		
Second level				
Parameters	Estimate	standard errors		
a_1	0.1532	0.1372		
b_1	0.7349	0.2126		
const	$\begin{bmatrix} 0.1281 & 0.2852 \\ & 0.6880 \end{bmatrix}$	$\begin{bmatrix} 8.44e-4 & 0.0013 \\ & 0.0092 \end{bmatrix}$		
a_2	0.0138	2.76e-5		
b_2	0.9567	2.61e-4		
const	$\begin{bmatrix} 0.0037 & -0.0087 \\ & 0.0164 \end{bmatrix}$	$\begin{bmatrix} 5.32e-6 & 1.53e-5 \\ & 0.0037 \end{bmatrix}$		
p_{11}	0.8843	0.0154		
p_{22}	0.2502	0.0305		

Table 2: Model estimates.

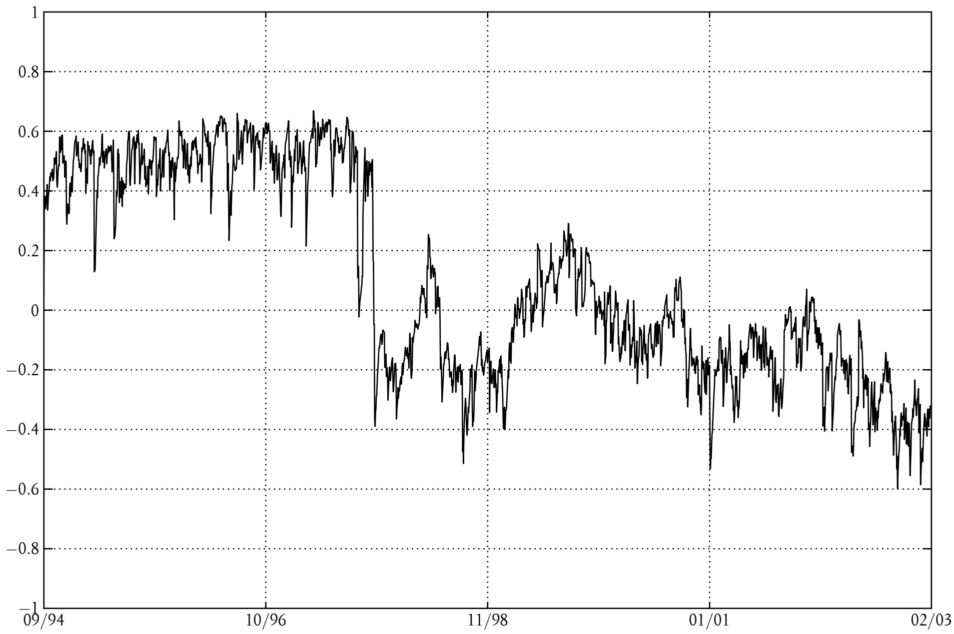
This result means that in general, volatility is associated with negative correlations.

3.2. **Deepening of the relationship between first and second level.** In the later application, we use a very simple transition matrix to link the first and the second level. This specification has the advantage of being attractive from a computational point of view but seems to be quite restrictive. Indeed the linking relation is common to all case of the first stage. In this subsection, we extend the later specification by introducing a specific relation for all the possible case at the first level. In that case, there are four two-by-two coupling matrix matrix P^i, $i = 1, \ldots, 4$. Each of these links a pair $\{h_{t,B}, h_{t,SP}\}$ with a case $\{R_t\}$. The decision of this extended case can be summarized as in figure 7.

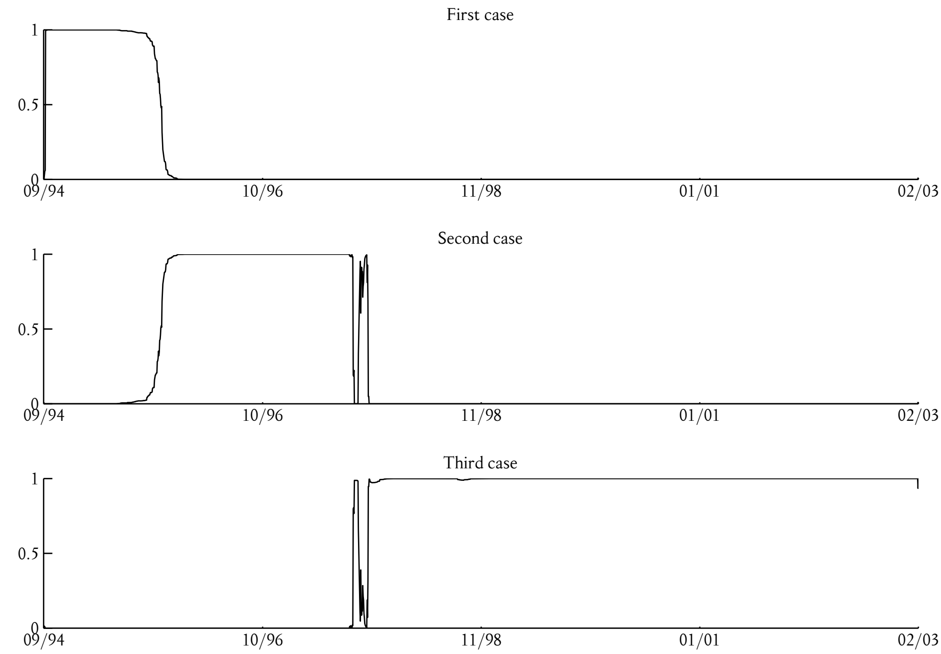
While the simple model needs only two parameters for the coupling matrix, the extended case needs eight parameters. Estimation is done using maximum likelihood using a regular HMM representation. Estimated correlations are plotted in figure 8(a) and the smoothed probabilities of the relevant regimes are plotted in figure 8(b); estimated parameters are in table 3.

Comparing to the previous specification, the introduction of specific coupling matrix increases the explanatory power of the model. Estimation of the model exhibits six cases:

(1) $\{h^1_{t,Y}, h^1_{t,P}, R^1_t\}$: low volatility for bond, low volatility for S&P500 and negative correlations regime.
(2) $\{h^1_{t,Y}, h^1_{t,P}, R^2_t\}$: low volatility for bond, low volatility for S&P500 and positive correlations regime.



(a) ESTIMATED CORRELATIONS.



(b) SMOOTHED PROBABILITIES.

Figure 6: Estimated correlations and smoothed probabilities.

(3) $\{h^1_{t,Y}, h^2_{t,P}, R^2_t\}$: low volatility for bond, high volatility for S&P500 and positive correlations regime.
(4) $\{h^2_{t,Y}, h^1_{t,P}, R^1_t\}$: high volatility for bond, low volatility for S&P500 and negative correlations regime.
(5) $\{h^2_{t,Y}, h^1_{t,P}, R^2_t\}$: high volatility for bond, low volatility for S&P500 and positive correlations regime.
(6) $\{h^2_{t,Y}, h^2_{t,P}, R^2_t\}$: high volatility for bond, high volatility for S&P500 and positive correlations.

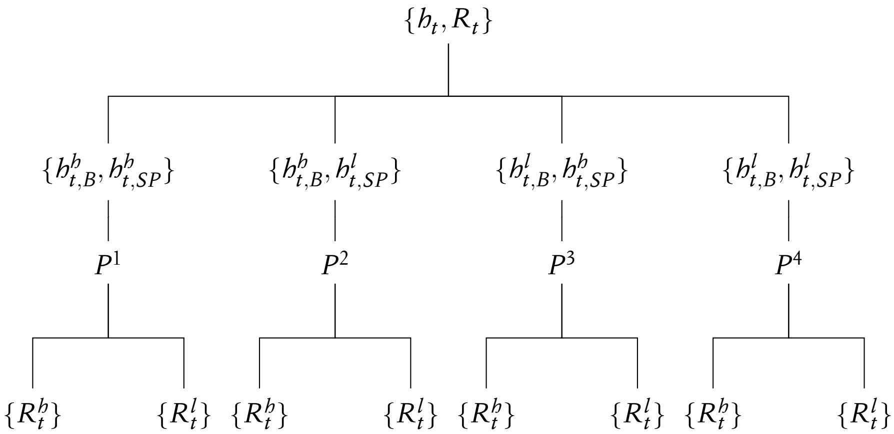
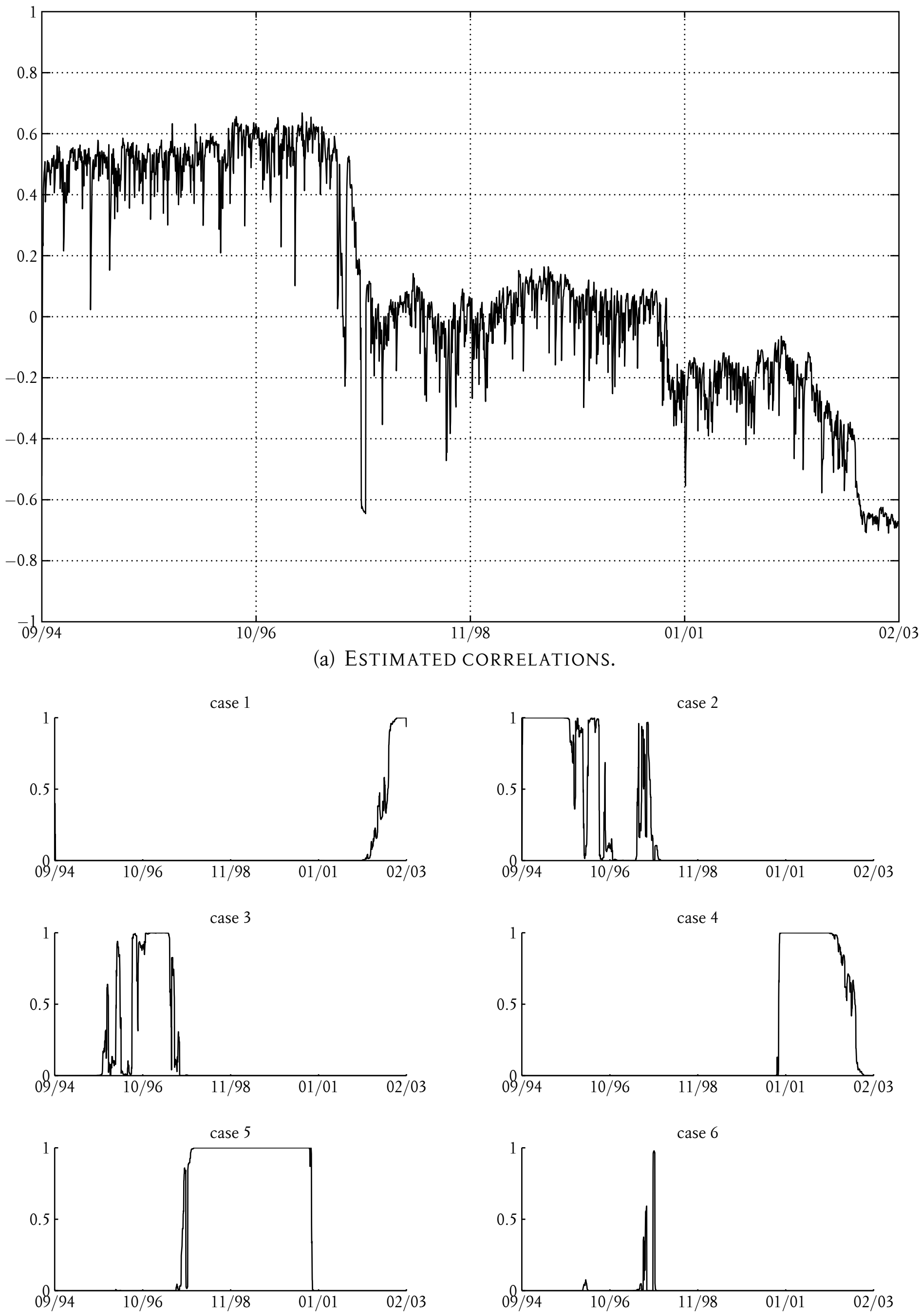


Figure 7: Structure of the decision process for a bivariate dataset of the extended version.

While the previous specification only shows us that in general volatility is associated with negative correlations, our extended model gives a more precise information about the relationship between volatility and correlations. Given the estimated parameters of the four coupling matrix, we can deduce that there is a strong probability to have positive correlation when bond and S&P500 are in low volatility regime and when bond has low volatility and S&P500 high has high volatility. On the other hand, negative correlations seems to be associated with a high volatility of the bond. This particularity is more signification when the S&P500 is in low volatility regime.

4. Conclusion

The model we have presented in this paper has a natural and easy interpretation. Based on the Hidden Markov Decision Tree introduced by Jordan, Ghahramani, and Saul (1997), we develop a stochastic decision tree liking the dynamics of univariate volatility with the dynamics of the correlations. Unlike conventional approaches based on deterministic decision tree, our model allows a probabilistic point of view of the relationship between univariate volatility and correlations. We have restricted ourselves the complexity of the tree-structure. More complex architecture remains to be done. It also remains to develop estimation tools more suitable when the dimension of the problem increases.



(a) ESTIMATED CORRELATIONS.

(b) SMOOTHED PROBABILITIES.

Figure 8: Estimated correlations and smoothed probabilities.

REFERENCES

AUDRINO, F. (2006): "The impact of general non-parametric volatility functions in multivariate GARCH models," *Computational Statistics & Data Analysis*, 50(11), 3032–3052.

AUDRINO, F. T., AND P. BUHLMANN (2001): "Tree-structured GARCH models," *Journal of the Royal Statistical Society, Series B*, 63(4), 727–744.

First level				
	10-year bond futures		S&P500	
Parameters	Estimate	standard errors	Estimate	standard errors
ω_1	0.0017	1.03e-06	0.0115	0.00015
α_1	0.0406	0.00029	0.0867	0.00129
β_1	0.9488	7.55e-05	0.8391	0.00933
ω_2	0.0137	6.35e-05	0.0831	0.00013
α_2	0.0706	0.00050	0.3582	1.11e-05
β_2	0.9195	0.00078	0.6417	1.06e-05
p_{11}	0.9987	0.00227	0.9999	1.02e-28
p_{22}	0.0010	6.272e-07	0.0001	1.70e-07
Coupling matrix				
Parameters	Estimate	standard errors		
p^1_{11}	0.7558	0,002485		
p^1_{22}	0.09319	1,28e-05		
p^2_{11}	0.7661	0,00066		
p^2_{22}	0.11846	1,66e-05		
p^3_{11}	0.2874	0,00045		
p^3_{22}	0.0267	8,71e-07		
p^4_{11}	0.5686	0,00026		
p^4_{22}	1.00e-05	6,43e-08		
Second level				
Parameters	Estimate	standard errors		
a_1	8.8334e-04	8.473e-06		
b_1	0.9716	0.00086		
const	$\begin{bmatrix} 0.0008 & -0.0031 \\ & 0.0127 \end{bmatrix}$	$\begin{bmatrix} 1.12e-05 & 0.00014 \\ & 0,00211 \end{bmatrix}$		
a_2	0.0344	0.00044		
b_2	0.4325	0.04093		
const	$\begin{bmatrix} 0.0834 & 0.1107 \\ & 0.1836 \end{bmatrix}$	$\begin{bmatrix} 0,00111 & 0,01074 \\ & 0,11735 \end{bmatrix}$		
p_{11}	0.8223	0.00144		
p_{22}	0.9714	6.76e-06		

Table 3: Model estimates.

AUDRINO, F. T., AND A. TROJANI (2006): "Estimating and Predicting Multivariate Volatility Thresholds in Global Stock Markets," *Journal of Applied Econometrics*, 21(3), 345–369.

BARONE-ADESI, G., AND F. AUDRINO (2006): "Average conditional correlation and tree structures for multivariate GARCH models," *Journal of Forecasting*, 25(8), 579–600.

BAUWENS, L., S. LAURENT, AND J. ROMBOUTS (2006): "Multivariate GARCH Models : a Survey," *Journal of Applied Econometrics*, 21(1), 79–109.

BILLIO, M., AND M. CAPORIN (2005): "Multivariate Markov switching dynamic conditional correlation GARCH representations for contagion analysis.," *Statistical Methods and Applications*, 14(2), 145–161.

——— (2009): "A generalized Dynamic Conditional Correlation model for portfolio risk evaluation," *Mathematics and Computers in Simulation*, 79(8).

BILLIO, M., M. CAPORIN, AND M. GOBBO (2003): "Block Dynamic Conditional Correlation Multivariate GARCH models," Working Paper 03.03, GRETA.

——— (2006): "Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation," *Applied Financial Economics Letters*, 2(2), 123–130.

BOLLERSLEV, T. (1990): "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH model," *Review of Economics and Statistics*, 72, 498–505.

BRAND, M. (1997): "Coupled hidden Markov models for modeling interacting processes," Technical report 405, MIT Media Lab Perceptual Computing.

BREIMAN, L., J. FRIEDMAN, R. OLSHEN, AND C. STONE (1984): *Classification and Regression Trees.* Wadsworth and Brooks, Monterey, CA.

CAJIGAS, M., AND G. URGA (2006): "Dynamic Conditional Correlation Models with Asymmetric Multivariate Laplace Innovations," Discussion paper, Centre for Econometric Analysis, Cass Business School.

CAPPIELLO, L., R. F. ENGLE, AND K. SHEPPARD (2006): "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4(4), 537–572.

CARVALHO, A. X., AND M. A. TANNER (2006): "Modeling nonlinearities with mixtures-of-experts of time series models," *International journal of mathematics and mathematical sciences*, 9, 1–22.

COLACITO, R., AND R. ENGLE (2006): "Testing and Valuing Dynamic Correlations for Asset Allocation," *Journal of Business and Economic Statistics*, 24, 238–253.

DELLAPORTAS, P., AND I. D. VRONTOS (2007): "Modelling volatility asymmetries: a Bayesian analysis of a class of tree structured multivariate GARCH models," *Econometrics Journal*, 10(3), 503–520.

DIEBOLD, F., J. LEE, AND G. WEINBACH (1994): "Regime Switching with Time-Varying Transition Probabilities," in *Non-Stationary Time Series Analysis and Co-Integration*, ed. by C. Hargreaves. The Clarendon Press & Oxford University Press, advanced texts in econometrics edn.

DUEKER, M. J. (1997): "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility," *Journal of Business & Economic Statistics*, 15(1), 26–34.

ENGLE, R. (2002): "Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics*, 20, 339–350.

——— (2008): "High dimensional dynamic correlations," in *The Methodology and Practice of Econometrics: Papers in Honour of David F Hendry*, ed. by J. L. Castle, and N. Shephard. Oxford University Press.

——— (2009): *Anticipating correlations: a new paradigm for risk management*, The econometrics institute lectures. Princeton University Press.

ENGLE, R., AND B. KELLY (2009): "Dynamic Equicorrelation," Working Paper FIN-08-038, New York University.

ENGLE, R., AND K. SHEPPARD (2001): "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," Working Paper 2001–15, UCSD.

ENGLE, R. F., N. SHEPHARD, AND K. SHEPPARD (2008): "Fitting vast dimensional time-varying covariance models," Discussion paper.

FENG, Y. (2007): "A Local Dynamic Conditional Correlation Model," Working Paper 1592, MPRA.

FRANSES, P., C. HAFNER, AND D. VAN DIJK (2005): "Semi–Parametric Modelling of Correlation Dynamics," *Advances in Econometrics*, 20, 53–103.

FRIEDMAN, J. (1991): "Multivariate Adaptive Regression Splines," *Annals of Statistics*, 19(1), 181–214.

GHAHRAMANI, Z., AND M. I. JORDAN (1997): "Factorial hidden Markov models," *Machine Learning*, 29, 245–273.

GRAY, S. (1996): "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27–62.

HAAS, M., AND S. MITTNIK (2008): "Multivariate Regime-Switching GARCH with an Application to International Stock Markets," CFS Working Paper Series 2008-08, Center for Financial Studies.

HAAS, M., S. MITTNIK, AND M. PAOLLELA (2004): "A New Approach to Markov-Switching GARCH Models," *Journal of Financial Econometrics*, 2(4), 493–530.

HAFNER, C., AND P. H. FRANSES (2009): "A Generalized Dynamic Conditional Correlation Model: Simulation and Application to Many Assets," *Econometric Reviews*, 28(6), 612–631.

HAMILTON, J. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

——— (1994): *Time Series Analysis.* Princeton University Press.

HAMILTON, J., AND R. SUSMEL (1994): "Autoregressive Conditional Heteroskedasticity and Change in Regime," *Journal of Econometrics*, 64, 307–333.

HUERTA, G., W. JIANG, AND M. A. TANNER (2003): "Time series modeling via hierarchical mixtures," *Statistica Sinica*, 13(4), 1097–1118.

JACOBS, R. A., M. I. JORDAN, S. J. NOWLAN, AND G. E. HINTON (1991): "Adaptive mixtures of local experts," *Neural Computation*, 3, 79–87.

JORDAN, M., F. GHAHRAMANI, AND L. SAUL (1997): "Hidden Markov decision trees," in *Advances in Neural Information Processing Systems (NIPS) 9*, ed. by M. C. Mozer, M. I. Jordan, and T. Petsche. Cambridge MA: MIT Press.

JORDAN, M. I., AND R. A. JACOBS (1994): "Hierarchical Mixtures of Experts and the EM Algorithm," *Neural Computation*, 6(2), 181–214.

KLAASSEN, F. (2002): "Improving GARCH Volatility Forecasts with Regime-Switching GARCH," *Empirical Economics*, 27(2), 363–394.

KROLZIG, H. (1997): *Markov–switching vector autoregressions*vol. 454. Lecture notes in economics and mathematical systems edn.

LOH, W. (2002): "Regression trees with unbiased variable selection and interaction detection," *Statistica Sinica*, (12), 361–386.

MA, L., Q. ZHOU, M. CELENK, AND D. M. CHELBERG (2004): "Facial event mining using coupled hidden markov models," in *ICIP*, pp. 1405–1408.

MCLACHLAN, G., AND D. PEEL (2000): *Finite Mixture Models*, Wiley Series in Probability and Mathematical Statistics.

NEFIAN, A. V., L. LIANG, X. PI, L. XIAOXIANG, C. MAO, AND K. MURPHY (2002): "A coupled HMM for audio-visual speech recognition," in *International Conference on Acoustics, Speech and Signal Processing (CASSP'02)*, pp. 2013–2016.

PALANDRI, A. (2009): "Sequential conditional correlations: Inference and evaluation," *Journal of Econometrics*, 153(2), 122–132.

PELLETIER, D. (2006): "Regime switching for dynamic correlations," *Journal of Econometrics*, 131(1–2), 445–473.

RABINER, L. R. (1989): "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, 77(2), 257–286.

ROWEIS, S. T. (2000): "One Microphone Source Separation," in *Proceedings of NIPS*, pp. 793–799.

SILVENNOINEN, A., AND T. TERÄSVIRTA (2005): "Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations," Working Paper Series in Economics and Finance 0577, Stockholm School of Economics.

——— (2009): "Multivariate GARCH models," in *Handbook of Financial Time Series*, ed. by T. Andersen, R. Davis, J. Kreiss, and T. Mikosch. Springer, New York.

SILVENNOINEN, A., AND T. TERÃĎSVIRTA (2009): "Modeling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model," *Journal of Financial Econometrics*, 7(4), 373–411.

TROJANI, F., AND F. AUDRINO (2005): "A general multivariate threshold GARCH model with dynamic conditional correlations," University of St. Gallen Department of Economics working paper series 2005 2005-04, Department of Economics, University of St. Gallen.

TSE, Y., AND A. TSUI (2002): "A Multivariate GARCH Model with Time-Varying Correlations," *Journal of Business and Economic Statistics*, 20(3), 351–362.

VAN DIJK, D., AND P. FRANSES (1999): "Modelling multiple regimes in the business cycle," *Macroeconomic Dynamics*, 3, 311–340.

VARGAS, G. A. (2006): "An Asymmetric Block Dynamic Conditional Correlation Multivariate GARCH Model," *The Philippine Statistician*, 55(1–2), 83–102.

——— (2008): "What Drives the Dynamic Conditional Correlation of Foreign Exchange and Equity Returns?," MPRA Paper 7174, University Library of Munich, Germany.

WILLIAMS, B., M. TOUSSAINT, AND A. J. STORKEY (2007): "Modelling motion primitives and their timing in biologically executed movements.," in *NIPS*, ed. by J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis. MIT Press.

XIE, L., AND Z.-Q. LIU (2007): "A coupled HMM approach to video-realistic speech animation," *Pattern Recognition*, 40(8), 2325–2340.

ZHOU, T., AND L. CHAN (2008): "Clustered Dynamic Conditional Correlation Multivariate GARCH Model," in *DaWaK '08: Proceedings of the 10th international conference on Data Warehousing and Knowledge Discovery.*