Gulf Journal of Mathematics
Vol 20, Issue 2 (2025) 106-117

ISSN: 2309-4966
https://doi.org/10.56947/gjom.v20i2.3157

CONTRIBUTION TO STATISTICAL INFERENCE IN ARMA-GARCH RANDOM FIELD MODELING

WAFA MILI[1]* and SOUMIA KHARFOUCHI[2]

Abstract. In spatial series analysis, accurately modeling regional volatility and spatial dependence structures continue to pose a significant challenge. This article presents a novel quasi-maximum likelihood approach for estimating the coefficients of a two-dimensionally indexed ARMA-GARCH model. Our method is characterized by minimal assumptions when establishing the consistency of the proposed estimators that enhance the flexibility and applicability of the ARMA-GARCH models across diverse data sets. We conducted a Monte Carlo study that validates the theoretical findings of our research, demonstrating its suitability for spatial data scenarios. Our results contribute to advancements in image and signal processing applications, highlighting the broader implications of our work.

Keywords. Spatial Data, 2D ARMA-GARCH model, heteroscedasticity, Volatility, GQML, consistency
2020 Mathematics Subject Classification. Primary 46L55; Secondary 44B20

1. Introduction

In the realm of spatial series analysis, modeling the regional volatility and spatial dependence structures have been a longstanding challenge. The univariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model has garnered substantial interest due to its effectiveness in capturing volatility clustering and providing insights into financial time series data. GARCH models are particularly adept at modeling time-varying volatility, making them essential tools for understanding and forecasting fluctuations in financial markets.
Over time, various extensions and modifications of the GARCH framework have emerged, reflecting the growing recognition of these models in diverse applications such as portfolio selection, risk management, and option pricing. A notable example is the recent study by [7], which investigates the time-dependent volatility of cryptocurrencies using eleven different GARCH models. Their findings emphasize the superior forecasting capability of the integrated GARCH (IGARCH)

Date: Received: Jun 18, 2025; Accepted: Jul 27, 2025.
* Corresponding author

model, characterized by extreme persistence and infinite memory, integral features within the standard GARCH framework. Additionally, their study reveals that Value at Risk (VaR) serves as an effective measure of risk exposure in the cryptocurrency market, with asymmetric GARCH models further enhancing the precision of VaR predictions.

Despite these advancements, many studies overlook the nonlinearity of spatial dependency when extending GARCH models to spatial contexts. This gap highlights the need for careful consideration and advanced statistical methodologies when integrating autoregressive moving average (ARMA) structures with spatial GARCH processes. Addressing this challenge is crucial for effectively modeling spatial volatility and understanding complex dependencies in high-dimensional data sets [2], [5] and [3].

The present article introduces a novel quasi-maximum likelihood methodology for the estimation of coefficients in the two-dimensionally indexed ARMA-GARCH model. This approach constitutes a significant advancement in statistical inference for spatio-temporal processes, addressing the complex dependencies inherent in such models using a robust estimation framework suitable for high-dimensional data structures. In contrast to existing methods, our approach is characterized by minimal assumptions, ensuring flexibility and applicability across diverse data sets. The computational performance of our proposed method is particularly noteworthy, making it suitable for large-scale and high-dimensional data. The proposed methodology represents a significant expansion upon prior findings in [11], addressing scenarios where the GARCH process is latent and unobserved, functioning as the innovation process driving the dynamics of the observed ARMA model. This framework gains prominence due to the inherent limitations by assuming that the observed series solely represents noise. Numerous studies have explored this challenge; for instance, the authors in [10] use the two-dimensional ARMA-GARCH model as a comprehensive statistical framework for clutter modeling and anomaly detection. Their work establishes that this model constitutes a generalization of the causal Gauss Markov Random Field (GMRF), commonly applied within the context of cluster modeling despite its drawback of a constant conditional variance across space. In contrast, the implementation of GARCH models in cluster modeling demonstrates notable advantages.

Furthermore, researchers in [9] demonstrated that sub-band decomposition of synthetic aperture radar (SAR) images exhibits notably non-normal distributional statistics, best captured by using the two-dimensional ARMA-GARCH regression model and its extensions. In the domain of medical ultrasound (US) images, [8] introduced an innovative speckle suppression technique based on a 2D generalized Gaussian ARMA-GARCH model. A comprehensive overview of these models is available in [11] and [1].

The cornerstone of our methodology lies in proposing asymptotically consistent estimators for the parameter vector, ensuring convergence to the true parameter under increasing sample information, even when faced with uncertainties about the underlying distribution. This is achieved by seeking parameter estimates that minimize the Kullback-Leibler divergence, thereby quantifying and reducing the discrepancy between the true data-generating distribution and a specified parametric family that may be subject to model misspecification. Building on the work of [6], we extend their method to the complex domain of 2D ARMA-GARCH models, providing convergence results under conditions more relaxed than those found in the existing literature.

Throughout this article, we present a theoretical foundation for our methodology by addressing the challenges associated with estimating the parameters of 2D ARMA-GARCH models. Our approach opens avenues for improved results and forecasting in image and signal processing applications.

The remainder of this paper is organized as follows: it commences with a concise introduction, Section 2 presents the model definition and underlying assumptions that form the foundation of our approach. In Section 3, we introduce the Generalized Quasi-Maximum Likelihood Estimator (GQMLE) for our model, demonstrating the consistency of the proposed estimator. Section 4 details the results of our Monte Carlo simulation, which validates our theoretical findings and assesses the performance of our method under various conditions. Finally, Section 4 concludes the paper.

2. Model definition and assumptions

In this section, we examine a random field $(Y_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^2}$ specified on the same underlying probability space $(\Omega, \Im, P)$ and generated by a 2D ARMA random field model with errors generated by a 2D GARCH process. $(Y_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^2}$ is given by

$$\begin{cases} Y_{\mathbf{t}} = \sum_{\mathbf{k}\in S]\mathbf{0},(P_1,P_2)]} a_{\mathbf{k}} Y_{\mathbf{t}-\mathbf{k}} + \sum_{\mathbf{l}\in S]\mathbf{0},(Q_1,Q_2)]} b_{\mathbf{l}}\varepsilon_{\mathbf{t}-\mathbf{l}} + \varepsilon_{\mathbf{t}}, \\ \varepsilon_{\mathbf{t}} = \sqrt{h_{\mathbf{t}}}\eta_{\mathbf{t}}, \\ h_{\mathbf{t}} = w + \sum_{\mathbf{k}\in S]\mathbf{0},(r_1,r_2)]} \alpha_{\mathbf{k}} {\varepsilon^2}_{\mathbf{t}-\mathbf{k}} + \sum_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]} \beta_{\mathbf{l}} h_{\mathbf{t}-\mathbf{l}}, \end{cases} \tag{2.1}$$

where $\eta_{\mathbf{t}}$ is a sequence of independent and identically distributed (i.i.d.) random variables defined on the same probability space $(\Omega, \Im, P)$ defined such that its mean is zero and variance is unity. For any positive integer $\mathbf{t} = (t_1, t_2)$, $\mathbf{0} = (0, 0)$, $\mathbf{k} = (k_1, k_2)$, $\mathbf{l} = (l_1, l_2)$. let $S = \{s \in \mathbb{R}^d\}$ represents the spatial set of sites on which a family of random variables is defined. The $(a_{\mathbf{k}})$, $(b_{\mathbf{l}})$ (AR and MA coefficients), $w > 0$, $(\alpha_{\mathbf{k}})_{\mathbf{k}\in S]\mathbf{0},(r_1,r_2)]} \geq 0$, $(\beta_{\mathbf{l}})_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]} \geq 0$ are constant coefficients. The orders $P = (P_1, P_2)$, $Q = (Q_1, Q_2)$, $r = (r_1, r_2)$ and $s = (s_1, s_2)$ are assumed known, and consider $]0, (r_1, r_2)] = \{(m_1, m_2) \in \mathbb{Z}^2, 0 < m_1 \leq r_1, 0 < m_2 \leq r_2\}$ by arranging its terms by the lexico-graphic order. The vector parameter is

$$\begin{aligned}\theta &= \left((a_{\mathbf{k}})_{\mathbf{k}\in S]0,(P_1,P_2)]}, (b_{\mathbf{l}})_{\mathbf{l}\in S]\mathbf{0},(Q_1,Q_2)]}, w, (\alpha_{\mathbf{k}})_{\mathbf{k}\in S]\mathbf{0},(r_1,r_2)]}, (\beta_{\mathbf{l}})_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]}\right)' \\ &= \left(\underline{\vartheta'}, \underline{\varphi'}\right)',\end{aligned}$$

Strict stationarity of model (2.1) is a key consideration for reliable inference and forecasting in the remainder of this work. For this, one needs to consider both the ARMA part and the GARCH part separately. Indeed, the ARMA-GARCH process is a combination of an ARMA process for the mean and a GARCH process for the variance.

For $\mathbf{i}=(i_1,i_2)$, $\mathbf{z}=(z_1,z_2)\in\mathbb{C}^2$ and $\mathbf{z}^{\mathbf{i}}=z_1^{i_1}z_2^{i_2}$, let

$$\Psi_\theta(\mathbf{z})=1-\sum_{\mathbf{i}\in]\mathbf{0},(P_1,P_2)]}a_{\mathbf{i}}\mathbf{z}^{\mathbf{i}},$$

and

$$\Phi_\theta(\mathbf{z})=1+\sum_{\mathbf{j}\in S]\mathbf{0},(Q_1,Q_2)]}b_{\mathbf{j}}\mathbf{z}^{\mathbf{j}}.$$

Derived from these (AR) and (MA) polynomials, standard assumptions for strict stationarity of model (2.1) can be formulated as follows:

A1 : All roots of $\Psi(\mathbf{z})\Phi(\mathbf{z})$ are outside the unit circle (this condition ensures the strict stationarity of the ARMA part of (2.1), see [13] for more details).

A2 : the top Lyapunov exponent $\rho(\Gamma)<1$, where $\Gamma=E\left\{A_1(\mathbf{t})^{\frac{\gamma}{2}}+A_2(\mathbf{t})^{\frac{\gamma}{2}}\right\}$ (Refer to [6] for the formulations of $A_1(\mathbf{t})$ and $A_2(\mathbf{t})$ and further details about the strict stationarity of the GARCH part of (2.1).

3. GQMLE for model 2D ARMA-GARCH

Suppose that the observations $(Y_{\mathbf{t}})_{\mathbf{t}\in S[1,(N_1,N_2)]}$ correspond to a sample path of length N_1N_2 derived from a model solution characterized by both strict stationarity and non-anticipativity (2.1) The parameter space is $\theta\in\Theta\subset\mathbb{R}^K\times\,]0,+\infty[\,\times[0,\infty[^L$ and $K=(P_1+1)(P_2+1)+(Q_1+1)(Q_2+1)-2$ and $L=(r_1+1)(r_2+1)+(s_1+1)(s_2+1)-2$.

The unknown true parameter can be represented by

$$\begin{aligned}\theta_0&=\left((a_{0\mathbf{k}})_{\mathbf{k}\in S]\mathbf{0},(P_1,P_2)]},(b_{0\mathbf{l}})_{\mathbf{l}\in S]\mathbf{0},(Q_1,Q_2)]},w_0,(\alpha_{0\mathbf{k}})_{\mathbf{k}\in S]\mathbf{0},(r_1,r_2)]},(\beta_{0\mathbf{l}})_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]}\right)'\\&=\left(\underline{\vartheta}'_0,\underline{\varphi}'_0\right)'.\end{aligned}$$

Without loss of generality, consider $(Q_1,Q_2)\preceq(r_1,r_2)$, the initial values

$$Y_{\mathbf{t}},\mathbf{t}\in S\left[\mathbf{0},(1-r_1+Q_1-P_1,1-r_2+Q_2-P_2)\right],$$
$$\widetilde{\varepsilon}_{\mathbf{t}},\mathbf{t}\in S\left[(-r_1+Q_1,-r_2+Q_2),(1-r_1,1-r_2)\right]$$

and

$$\tilde{\nu}^2_{\mathbf{t}},\mathbf{t}\in S\left[(1-s_1,1-s_2),\mathbf{0}\right]$$

allow us to compute $\widetilde{\varepsilon}_{\mathbf{t}}(\underline{\vartheta})$, for
$\mathbf{t}\in\mathbf{S}\left[(-r_1+Q_1+1,-r_2+Q_2+1),\mathbf{N}\right]$, and $\tilde{\nu}^2_{\mathbf{t}}(\theta)$, $\mathbf{t}\in\mathbf{S}\left[\mathbf{1},\mathbf{N}\right]$, from

$$\left\{\begin{array}{l}\widetilde{\varepsilon}_{\mathbf{t}}=\widetilde{\varepsilon}^2_{\mathbf{t}}(\underline{\vartheta})=Y_{\mathbf{t}}-\sum_{\mathbf{t}\in S]\mathbf{0},(P_1,P_2)]}a_{\mathbf{k}}Y_{\mathbf{t-k}}-\sum_{\mathbf{l}\in S]\mathbf{0},(Q_1,Q_2)]}b_{\mathbf{l}}\widetilde{\varepsilon}_{\mathbf{t-l}}\\ \tilde{\nu}^2_{\mathbf{t}}=\tilde{\nu}^2_{\mathbf{t}}(\theta)=w+\sum_{\mathbf{k}\in S]\mathbf{0},(r_1,r_2)]}\alpha_{\mathbf{k}}\tilde{\varepsilon}^2_{\mathbf{t-k}}+\sum_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]}\beta_{\mathbf{l}}\tilde{\nu}^2_{\mathbf{t-l}}\end{array}\right.,\qquad(3.1)$$

An appropriate choice of initial values is made as follows:

$$\begin{aligned}&Y_{\mathbf{k}}=Y_1\text{ for all }\mathbf{k}\in S\left[\mathbf{0},(1-r_1+Q_1-P_1,1-r_2+Q_2-P_2)\right]\\&\widetilde{\varepsilon}^2_{\mathbf{k}}=\widetilde{\varepsilon}^2_1\text{ for all }\mathbf{k}\in S\left[(-r_1+Q_1,-r_2+Q_2),(1-r_1,1-r_2)\right],\\&\breve{\nu}^2_{\mathbf{t}}=\varepsilon^2_1\text{ for all }\mathbf{k}\in S\left[(1-s_1,1-s_2),\mathbf{0}\right],\end{aligned}\qquad(3.2)$$

with $\varepsilon_{\mathbf{t}} = \widetilde{\varepsilon}_{\mathbf{t}}\left(\underline{\vartheta}\right)$, the Gaussian QMLE is given by

$$\begin{aligned} L_{\mathbf{N}}\left(\theta\right) &= L_{\mathbf{N}}\left(\theta, \left(\widetilde{\varepsilon}_{\mathbf{t}}\left(\underline{\vartheta}\right)\right)_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]}\right) \\ &= \sum_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]} \frac{1}{\sqrt{2\pi\tilde{\nu}_{\mathbf{t}}^{2}\left(\theta\right)}} \exp\left(-\frac{\widetilde{\varepsilon}_{\mathbf{t}}^{2}\left(\underline{\vartheta}\right)}{2\tilde{\nu}_{\mathbf{t}}^{2}\left(\theta\right)}\right), \end{aligned}$$

Any measurable solution $\hat{\theta}_{\mathbf{N}}$ of

$$\hat{\theta}_{\mathbf{N}} = \arg\max_{\theta\in\Theta} L_{\mathbf{N}}\left(\theta\right) = \arg\min_{\theta\in\Theta} \tilde{I}_{\mathbf{N}}\left(\theta\right) \tag{3.3}$$

is termed a quasi-maximum likelihood estimator (QMLE) of θ where

$$\tilde{I}_{\mathbf{N}}\left(\theta\right) = \frac{1}{N_1 N_2} \sum_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]} \tilde{l}_{\mathbf{t}} \text{ and } \tilde{l}_{\mathbf{t}} = \tilde{l}_{\mathbf{t}}\left(\theta\right) = \frac{\widetilde{\varepsilon}_{\mathbf{t}}^{2}\left(\vartheta\right)}{\tilde{\nu}_{\mathbf{t}}^{2}\left(\theta\right)} + \log\tilde{\nu}_{\mathbf{t}}^{2}\left(\theta\right).$$

3.1. **Consistency.** To show the strong consistency, let us consider the following supplementary hypotheses:

A3 : $\theta_0 \in \Theta$ and Θ is compact.

A4 : $\ominus_{\underline{\vartheta}_0}\left(\mathbf{z}\right)$ and $\oplus_{\underline{\vartheta}_0}\left(\mathbf{z}\right)$ have no roots in common, $a_{0(P_1,P_2)} \neq 0$ or $b_{0(Q_1,Q_2)} \neq 0$.

A5 : $\forall\underline{\varphi} \in \left]0,\infty\right[\times \left[0,\infty\right[^{L}$, $\sum_{\mathbf{l}\in S]\mathbf{0},(s_1,s_2)]} \beta_{\mathbf{l}} < 1$.

A6 : if $\mathbf{r} > \mathbf{0}$, $\mathcal{A}_0\left(\mathbf{z}\right) = \sum\limits_{\mathbf{i}\in]\mathbf{0},(r_1,r_2)]} \alpha_{0\mathbf{i}}\mathbf{z}^{\mathbf{i}}$, $\mathcal{B}_0\left(\mathbf{z}\right) = 1 + \sum\limits_{\mathbf{j}\in S]\mathbf{0},(s_1,s_2)]} \beta_{0\mathbf{j}}\mathbf{z}^{\mathbf{j}}$ have no common roots, $\mathcal{A}_0\left(\mathbf{z}\right) \neq 0$, and $\alpha_{0(s_1,s_2)} + \beta_{0(s_1,s_2)} \neq 0$

Under assumptions **A1**$-$**A6**, $(Y_{\mathbf{t}})$ is taken to be the unique strictly stationary solution satisfying the non-anticipativity condition, to the underlying model equations (2.1). While Assumption **A2** (Lyapunov exponent $\rho(\Gamma) < 1$) characterizes strict stationarity, Assumption **A5** ($\sum\beta_{\mathbf{i}} < 1$) is required for our proof technique to to guarantee uniform geometric decay of initial condition effects. We can now state our first result.

Theorem 3.1. *(Consistency of the QML) Let $(\hat{\theta}_{\mathbf{N}})$ denote a sequence of QML estimators satisfying the condition specified in equation (3.3), subject to the initial conditions outlined in equation (3.2). Provided that the assumptions **A1**$-$**A6** are satisfied, almost surely $\hat{\theta}_{\mathbf{N}} \to \theta_0$, as $\mathbf{N} \to \infty$.*

Proof. We will follow the four steps below,

- **(a):** $\lim\limits_{\mathbf{N}\to\infty} \sup\limits_{\underline{\theta}\in\Theta} \left| I_{\mathbf{N}}\left(\underline{\theta}\right) - \tilde{I}_{\mathbf{N}}\left(\underline{\theta}\right)\right| = 0$, a.s.
- **(b):** $\exists\mathbf{t} \in \mathbb{Z}^2$ such that $\varepsilon_{\mathbf{t}}\left(\underline{\vartheta}\right) = \varepsilon_{\mathbf{t}}\left(\underline{\vartheta}_0\right)$ and $\nu_{\mathbf{t}}^{2}\left(\underline{\theta}\right) = \nu_{\mathbf{t}}^{2}\left(\underline{\theta}_0\right)$ $\;P_{\underline{\theta}_0}$ p.s $)\Longrightarrow$ $\underline{\theta} = \underline{\theta}_0$
- **(c):** If $\underline{\theta} \neq \underline{\theta}_0$, then $E_{\underline{\theta}_0}\left|l_{\mathbf{t}}\left(\underline{\theta}_0\right)\right| < E_{\underline{\theta}_0}\left|l_{\mathbf{t}}\left(\underline{\theta}\right)\right|$
- **(d):** For every parameter value $\underline{\theta}$ distinct from $\underline{\theta}_0$, A neighborhood $V\left(\underline{\theta}\right)$ containing $\underline{\theta}$ can be found such that,

$$\liminf_{\mathbf{N}\to\infty} \inf_{\underline{\theta}^*\in V(\underline{\theta})} \breve{I}_{\mathbf{N}}\left(\underline{\theta}^*\right) > E_{\underline{\theta}_0} l_1\left(\underline{\theta}_0\right) \; a.s.$$

To prove (a), we need to rewrite equation (3.1) in matrix form. For $\mathbf{t} = (t_1, t_2)$, we have

$$\underline{Y}_{(t_1,t_2)} = \underline{\mathbf{e}}_{(t_1,t_2)} + D_{f1}\underline{Y}_{(t_1,t_2-1)} + D_2\underline{Y}_{(t_1-1,t_2)}, \, (3.5)$$

$$\underline{\nu}^2_{(t_1,t_2)} = \underline{\mathbf{c}}_{(t_1,t_2)} + B_1\underline{\nu}^2_{(t_1,t_2-1)} + B_2\underline{\nu}^2_{(t_1-1,t_2)}, \tag{3.4}$$

where D_1, D_2 and B_1, B_2 are suitable $(P_1+1)\,P_2 \times (P_1+1)\,P_2$ and $(s_1+1)\,s_2 \times (s_1+1)\,s_2$ matrices, respectively, and

$$\underline{Y}_{(t_1,t_2)} = \begin{pmatrix} Y_{(t_1,t_2)} \\ Y_{(t_1,t_2-1)} \\ \vdots \\ Y_{(t_1,t_2-s_2+1)} \\ \vdots \\ Y_{(t_1-s_1,t_2)} \\ \vdots \\ Y_{(t_1-s_1,t_2-s_2+1)} \end{pmatrix}, \quad \underline{e}_{(t_1,t_2)} = \begin{pmatrix} \sum_{\mathbf{l}\in S]\mathbf{0},(Q_1,Q_2)]} b_{\mathbf{l}}\varepsilon_{\mathbf{t}-\mathbf{l}} + \varepsilon_{\mathbf{t}} \\ 0 \\ \vdots \\ 0 \end{pmatrix};$$

$$\underline{\nu}^2_{(t_1,t_2)} = \begin{pmatrix} \nu^2_{(t_1,t_2)} \\ \nu^2_{(t_1,t_2-1)} \\ \vdots \\ \nu^2_{(t_1,t_2-s_2+1)} \\ \vdots \\ \nu^2_{(t_1-s_1,t_2)} \\ \vdots \\ \nu^2_{(t_1-s_1,t_2-s_2+1)} \end{pmatrix}, \quad \underline{c}_{(t_1,t_2)} = \begin{pmatrix} \omega + \sum_{(k_1,k_2)\in S]\mathbf{0},\mathbf{r}]} \alpha_{(k_1,k_2)}\varepsilon^2_{(t_1-k_1,t_2-k_2)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Iterating (3.4) we obtain

$$\begin{aligned} \underline{Y}_{(t_1,t_2)} &= \sum_{n=0}^{t_1+t_2-1} \sum_{l=0}^{n} \Gamma^{l,k-1}\underline{e}_{(t_{1-l},t_{2-n+1})} + \Gamma^{t_1,t_2}\underline{Y}_{(0,0)} \\ &= \sum_{n\geq 0}\sum_{l=0}^{n*} \Gamma^{l,n-1}\underline{e}_{(t_{1-l},t_{2-n+1})}, \end{aligned} \tag{3.5}$$

$$\begin{aligned} \underline{\nu}^2_{(t_1,t_2)} &= \sum_{n=0}^{t_1+t_2-1} \sum_{l=0}^{n} \Gamma_*^{l,k-1}\underline{c}_{(t_{1-l},t_{2-n+1})} + \Gamma_*^{t_1,t_2}\underline{\nu}^2_{(0,0)} \\ &= \sum_{n\geq 0}\sum_{l=0}^{n} \Gamma_*^{l,n-1}\underline{c}_{(t_{1-l},t_{2-n+1})}, \end{aligned} \tag{3.6}$$

where $\Gamma^{n,m}(\mathbf{t})$ is defined as follows:

$\mathbf{T_1}$): $\Gamma^{0.1} = D_1$, $\Gamma^{1.0} = D_2$ et $\Gamma^{n,m} = D_1\Gamma^{n,m-1} + D_2\Gamma^{n-1,m}$
$\mathbf{T_2}$): $\Gamma_*^{0.1} = B_1$, $\Gamma_*^{1.0} = B_2$ et $\Gamma_*^{n,m} = B_1\Gamma_*^{n,m-1} + B_2\Gamma_*^{n-1,m}$

$\mathbf{T_3}$): $\Gamma^{0,0} = I_{(P_1+1)P_2}$, $\Gamma_*^{0,0} = I_{(s_1+1)s_2}$, $\Gamma^{-n,m} = \Gamma^{n,-m} = O_{(P_1+1)P_2}$ and $\Gamma_*^{-n,m} = \Gamma_*^{n,-m} = O_{(s_1+1)s_2}$.

Let $\underline{\tilde{\nu}}^2_{(t_1,t_2)}$ be defined as the vector derived by substituting $\nu^2_{(t_1,t_2)}$ by $\tilde{\nu}^2_{(t_1,t_2)}$ in $\underline{\nu}^2_{(t_1,t_2)}$, and let $\underline{\tilde{c}}_{(t_1,t_2)}$ constitute the vector obtained through the replacement of $\varepsilon^2_{\mathbf{i}}, \mathbf{i} \in S\left[(1-r_1,-r_2), \mathbf{0}\right]$, by the initial values (3.2). Assumptions **A1**, **A2** and **A3** ensure the almost sure convergence of $\sum_{n\geq 0}\sum_{l=0}^{n*}\Gamma^{l,n-1}\underline{c}_{(t_{1-l},t_{2-n+1})}$ and $\sum_{k\geq 0}\sum_{l=0}^{k}\Gamma^{l,k-1}\underline{c}_{(t_1-l,t_2-k+1)}$. Hence, $\left(\nu^2_{(t_1,t_2)} - \tilde{\nu}^2_{(t_1,t_2)}\right)$ and $\left(\varepsilon_{(t_1,t_2)} - \tilde{\varepsilon}_{(t_1,t_2)}\right)$ are the rest of convergent series, with for any $\mathbf{k} > \mathbf{0}$ and $\mathbf{i} \in S\left]\mathbf{0},(s_1,s_2)\right]$

$$\sup_{\theta\in\Theta}\left|\varepsilon_{(t_1,t_2)} - \tilde{\varepsilon}_{(t_1,t_2)}\right| \leq K\rho \quad \text{a.s}$$

where $\rho = \rho\left(B_1^{\frac{\gamma}{2}} + B_2^{\frac{\gamma}{2}}\right)$ for some $\gamma \in \left]0,2\right]$. Note here that assumption **A5** insures that $\rho < 1$.
We thus have, almost surely,

$$\left\|\underline{c}_n - \underline{\tilde{c}}_n\right\| \leq \sum_{\mathbf{k}\in S[\mathbf{1},(r_1,r_2)]}|\alpha_{\mathbf{k}}|\left|\tilde{\varepsilon}^2_{\mathbf{n-k}} - \varepsilon^2_{\mathbf{n-k}}\right| \leq K\rho^n\left(\sum_{\mathbf{k}\in S[1,(r_1,r_2)]}|\varepsilon_{n-\mathbf{k}}| + 1\right), \tag{3.7}$$

By (3.5)-(3.6)-(3.7), it follows that

$$\begin{aligned}
\left\|\underline{\nu}^2_{(t_1,t_2)} - \underline{\tilde{\nu}}^2_{(t_1,t_2)}\right\| &= \Bigg\|\sum_{\mathbf{n}=0}^{t_1+t_2-1}\sum_{l=0}^{\mathbf{n}}\Gamma^{\mathbf{l,n-1}}\left(\underline{c}_{(t_{1-l},t_{2-n+1})} - \underline{\tilde{c}}_{(t_{1-l},t_{2-n+1})}\right) + \\
&\quad \Gamma^{t_1,t_2}\left(\underline{\nu}^2_{(0,0)} - \underline{\tilde{\nu}}^2_{(0,0)}\right)\Bigg\| \\
&\leq K\sum_{\mathbf{n}\in S[1,t_1+t_2]}^{t_1+t_2}\rho^{t_1+t_2-n}\rho^n\left(\sum_{\mathbf{k}\in S[1,(r_1,r_2)]}|\varepsilon_{n-\mathbf{k}}| + 1\right) + K\rho^{t_1+t_2} \\
&\leq K\rho^{t_1+t_2}\sum_{n\in S[-r_1-r_2,t_1+t_2]}|\varepsilon_n| + 1
\end{aligned}$$

We thus have

$$\begin{aligned}
\sup_{\underline{\theta}\in\Theta}\left|I_N\left(\underline{\theta}\right) - \hat{I}_N\left(\underline{\theta}\right)\right| &\leq \frac{1}{N}\sum_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]}\left\{\sup_{\underline{\theta}\in\Theta}\left|\frac{\underline{\tilde{\nu}}^2_{(\mathbf{t_1},\mathbf{t_2})} - \underline{\nu}^2_{(\mathbf{t_1},\mathbf{t_2})}}{\underline{\tilde{\nu}}^2_{(\mathbf{t_1},\mathbf{t_2})}\underline{\nu}^2_{(\mathbf{t_1},\mathbf{t_2})}}\right|\varepsilon^2_{\mathbf{t}} + \right. \\
&\quad \left.\left|\log\left(1 + \frac{\underline{\nu}^2_{(\mathbf{t_1},\mathbf{t_2})} - \underline{\tilde{\nu}}^2_{(\mathbf{t_1},\mathbf{t_2})}}{\underline{\tilde{\nu}}^2_{(\mathbf{t_1},\mathbf{t_2})}}\right)\right| + \frac{\left|\varepsilon^2_{\mathbf{t}} - \tilde{\varepsilon}^2_{\mathbf{t}}\right|}{\underline{\tilde{\nu}}^2_{(\mathbf{t_1},\mathbf{t_2})}}\right\} \\
&\leq \left\{\sup_{\underline{\theta}\in\Theta}\max\left(\frac{1}{w^2},\frac{1}{w}\right)\right\}\frac{K}{N}\sum_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]}\rho^{t_1+t_2}\left(\varepsilon^2_{\mathbf{t}} + 1\right)\times \\
&\quad \left[\sum_{n\in S[-r_1-r_2,t_1+t_2]}\left(|\varepsilon_n| + 1\right)\right]
\end{aligned}$$

We substitute $(\varepsilon_{\mathbf{t}}^2+1)\times\left[\sum_{n\in S[-r_1-r_2,t_1+t_2]}(|\varepsilon_n|+1)\right]$ by $\xi_{\mathbf{t}}$, i.e.,

$$\left(\varepsilon_{\mathbf{t}}^2+1\right)\sum_{n\in S[-r_1-r_2,t_1+t_2]}(|\varepsilon_n|+1)=\xi_{\mathbf{t}},$$

then

$$\sup_{\underline{\theta}\in\Theta}\left|I_N(\underline{\theta})-\hat{I}_N(\underline{\theta})\right|\le\left\{\sup_{\underline{\theta}\in\Theta}\max\left(\frac{1}{w^2},\frac{1}{w}\right)\right\}\frac{K}{N}\sum_{\mathbf{t}\in S[\mathbf{1},\mathbf{N}]}\rho^{t_1+t_2}\xi_{\mathbf{t}}$$

It is therefore sufficient to demonstrate that for every positive real number m that $E\left(\rho^{t_1+t_2}\xi_{\mathbf{t}}\right)^m$ is finite

$$\begin{aligned}E\left(\rho^{t_1+t_2}\xi_{\mathbf{t}}\right)^{\frac{m}{2}}&\le\rho^{\frac{(t_1+t_2)m}{2}}\sum_{n\in S[-r_1-r_2,t_1+t_2]}\left(\varepsilon_{\mathbf{t}}^2|\varepsilon_n|+\varepsilon_{\mathbf{t}}^2+|\varepsilon_n|+1\right)^{\frac{m}{2}}\\&\le\rho^{\frac{(t_1+t_2)m}{2}}\sum_{n\in S[-r_1-r_2,t_1+t_2]}\left(\left\{E\left(\varepsilon_{\mathbf{t}}^{2m}\right)E|\varepsilon_n|^m\right\}^{\frac{1}{2}}+E|\varepsilon_{\mathbf{t}}|^m+E|\varepsilon_n|^{\frac{m}{2}}+1\right)\\&=O\left((t_1+t_2)\,\rho^{\frac{(t_1+t_2)m}{2}}\right).\end{aligned}$$

Thus, the proof of (a) has been established.

Part (b): Identifiability Condition

We show that if $\varepsilon_{\mathbf{t}}(\theta)=\varepsilon_{\mathbf{t}}(\theta_0)$ and $\nu_{\mathbf{t}}^2(\theta)=\nu_{\mathbf{t}}^2(\theta_0)$ almost surely, then $\theta=\theta_0$.

(1) *ARMA Component:* The equality $\varepsilon_{\mathbf{t}}(\theta)=\varepsilon_{\mathbf{t}}(\theta_0)$ implies:

$$\sum_{\mathbf{k}\in S[0,(P_1,P_2)]}a_{\mathbf{k}}Y_{\mathbf{t-k}}+\sum_{\mathbf{l}\in S[0,(Q_1,Q_2)]}b_{\mathbf{l}}\varepsilon_{\mathbf{t-l}}=\sum_{\mathbf{k}}a_{0\mathbf{k}}Y_{\mathbf{t-k}}+\sum_{\mathbf{l}}b_{0\mathbf{l}}\varepsilon_{\mathbf{t-l}}.$$

Under Assumption **A1** (roots of $\Psi_\theta(\mathbf{z})\Phi_\theta(\mathbf{z})$ outside the unit circle) and **A4** (no common roots), the spatial AR and MA polynomials are uniquely determined. Thus, $\underline{\vartheta}=\underline{\vartheta}_0$.

(2) *GARCH Component:* From $\nu_{\mathbf{t}}^2(\theta)=\nu_{\mathbf{t}}^2(\theta_0)$ and stationarity (**A2**):

$$w+\sum_{\mathbf{k}\in S[0,(r_1,r_2)]}\alpha_{\mathbf{k}}\varepsilon_{\mathbf{t-k}}^2+\sum_{\mathbf{l}\in S[0,(s_1,s_2)]}\beta_{\mathbf{l}}\nu_{\mathbf{t-l}}^2=w_0+\sum_{\mathbf{k}}\alpha_{0\mathbf{k}}\varepsilon_{\mathbf{t-k}}^2+\sum_{\mathbf{l}}\beta_{0\mathbf{l}}\nu_{\mathbf{t-l}}^2.$$

By **A6**, the polynomials $\mathcal{A}_\theta(\mathbf{z})$ and $\mathcal{B}_\theta(\mathbf{z})$ have no common roots, and non-negative coefficients ensure $\underline{\varphi}=\underline{\varphi}_0$.

Part (c): Kullback-Leibler Divergence Minimization

For $\theta\neq\theta_0$, we prove $E_{\theta_0}[l_{\mathbf{t}}(\theta_0)]<E_{\theta_0}[l_{\mathbf{t}}(\theta)]$.

$$\begin{aligned}E_{\theta_0}[l_{\mathbf{t}}(\theta)]&=E_{\theta_0}\left[\log\nu_{\mathbf{t}}^2(\theta)+\frac{\varepsilon_{\mathbf{t}}^2(\theta_0)}{\nu_{\mathbf{t}}^2(\theta)}\right]\\&=E_{\theta_0}\left[\log\nu_{\mathbf{t}}^2(\theta)+\frac{\nu_{\mathbf{t}}^2(\theta_0)}{\nu_{\mathbf{t}}^2(\theta)}\right]+E_{\theta_0}\left[\frac{(\varepsilon_{\mathbf{t}}(\theta_0)-\varepsilon_{\mathbf{t}}(\theta))^2}{\nu_{\mathbf{t}}^2(\theta)}\right].\end{aligned}$$

The first term is minimized when $\nu_{\mathbf{t}}^2(\theta)=\nu_{\mathbf{t}}^2(\theta_0)$ (by Jensen's inequality), and the second term vanishes only if $\varepsilon_{\mathbf{t}}(\theta)=\varepsilon_{\mathbf{t}}(\theta_0)$.

Part (d): Neighborhood Behavior
For any $\theta \neq \theta_0$, one can construct a neighborhood $V(\theta)$ for which:

$$\liminf_{N\to\infty} \inf_{\theta^*\in V(\theta)} \tilde{I}_N(\theta^*) > E_{\theta_0}[l_1(\theta_0)] \quad \text{a.s.}$$

(1) Uniform convergence of $\tilde{I}_N(\theta)$ to $E_{\theta_0}[l_{\mathbf{t}}(\theta)]$ follows from the ergodicity of the 2D ARMA-GARCH process (**A1**–**A2**).
(2) The strict convexity of $E_{\theta_0}[l_{\mathbf{t}}(\theta)]$ (from part (c)) ensures a local minimum at θ_0.

□

4. Monte Carlo simulation

To assess the finite-sample properties of the quasi-maximum likelihood (QML) estimator presented in Section 2, a rigorous evaluation framework is required , we conducted a series of simulation experiments. This Monte Carlo study allows us to evaluate the statistical properties of the QML estimator under controlled conditions by varying model parameters and sample sizes. Simulations were conducted for two first-order 2D ARMA-GARCH processes with parameters situated within the domain of the parameter space characterizing strictly stationary processes. The first model considered i.i.d. variables that follow a distribution characterized by a normal density function, whereas the second model is assumed to follow a centered Student t-distribution with five degrees of freedom.

The dataset was generated based on the first-order GARCH process described below:

$$\begin{cases} Y(i,j) &= aY(i,j-1) + bY(i-1,j) + \varepsilon(i,j), \\ \varepsilon(i,j) &= \sqrt{h(i,j)}\eta(i,j), \\ h(i,j) &= \omega + \alpha\varepsilon^2(i,j-1) + \beta h(i-1,j-1) \end{cases} \tag{4.1}$$

Parameters ω, α, β, a, and b are chosen in a manner that guarantees the existence and uniqueness of strictly stationary solutions of (4.1). Two options for $\{\eta(i,j);(i,j) \in \mathbb{Z}^2\}$ are chosen. To simplify the calculations, we assume that $\omega = 1$ is known. First, $\eta(i,j) \sim NID(0,1)$, $i = 1,\ldots,n$ and $j = 1,\ldots,m$ (model 1). Next, we assume that the innovations follow a centered Student t-distribution with five degrees of freedom in the second 2D ARMA-GARCH model (Model 2). The process initialization values are computed based on equation (3.2), specifically

$$\begin{aligned} Y(0,0) = Y(1,0) = Y(0,1) = Y(1,1), \\ \varepsilon(0,0) = \varepsilon(1,0) = \varepsilon(0,1) = \varepsilon(1,1), \\ h(0,0) = h(1,0) = h(0,1) = h(1,1). \end{aligned}$$

Let QML denotes the proposed quasi-maximum likelihood estimator. Data were generated for grid sizes of $(500,500)$, $(1000,1000)$, and $(1500,1500)$, resulting in 500 replications. Tables 1 and 2 display the averages in the 'mean columns' and the 'mean square error columns', which provide the means and The standard deviations of the sampling distributions of the average estimators across the 500 replications for both models 1 and 2, respectively. The average estimates for QMLE consistently approximated the true values across all grid sizes and parameter values. Additionally, as anticipated, increasing the sample size (N, M) significantly enhances bias reduction, supporting the strong consistency of QML. Finally, one can see from Table 2 that the QML maintains considerable robustness against deviations from the assumed innovation distribution in Model 2.

Table 1: Sample distribution of the estimators $\widehat{\alpha}^{qml}$, $\widehat{\beta}^{qml}$, $\widehat{a}^{qml}$ and $\widehat{b}^{qml}$ for Model 1

$true\ parameters$	$\alpha = 0.1$	$\beta = 0.8$	$a = 0.25$	$b = 0.075$
$(N, M) = (500, 500)$	$\widehat{\alpha}^{qml}$	$\widehat{\beta}^{qml}$	$\widehat{a}^{qml}$	$\widehat{b}^{qml}$
$mean$	0.079	0.821	0.2499	0.0749
$rmse$	0.0015	0.0016	0.0018	0.0021
$(N, M) = (1000, 1000)$				
$mean$	0.0788	0.820	0.2501	0.0751
$rmse$	0.001	9.6×10^{-4}	0.0011	0.0013
$(N, M) = (1500, 1500)$				
$mean$	0.0851	0.803	0.25	0.07505
$rmse$	4.8×10^{-5}	5.1×10^{-5}	7.2×10^{-5}	9.2×10^{-4}

Table 2: Sample distribution of the estimators $\widehat{\alpha}^{qml}$, $\widehat{\beta}^{qml}$, $\widehat{a}^{qml}$ and $\widehat{b}^{qml}$ for Model 2

$true\ parameters$	$\alpha = 0.1$	$\beta = 0.8$	$a = 0.25$	$b = 0.075$
$(N, M) = (500, 500)$	$\widehat{\alpha}^{qml}$	$\widehat{\beta}^{qml}$	$\widehat{a}^{qml}$	$\widehat{b}^{qml}$
$mean$	0.1426	0.8463	0.2497	0.0743
$rmse$	0.0049	0.0034	0.0024	0.0042
$(N, M) = (1000, 1000)$				
$mean$	0.1421	0.8459	0.2498	0.0751
$rmse$	0.0025	0.0018	0.0011	0.002
$(N, M) = (1500, 1500)$				
$mean$	0.1418	0.8459	0.25	0.075
$rmse$	0.0017	0.0012	7.6×10^{-4}	0.0018

5. Conclusion

This study presents an advancement in the estimation of two-dimensionally indexed ARMA-GARCH models through a quasi-maximum likelihood approach characterized by minimal assumptions. The implications of our research provide valuable tools for various applications in image and signal processing. Notably, we highlight the potential application of the proposed methodology in denoising ultrasound images, building on the work in [11] that utilized a generalized method of moments (GMM) approach.

In [11], the authors demonstrated the effectiveness of GMM in reducing noise while preserving important features within ultrasound images. However, we propose to enhance this algorithm by replacing the GMM estimation with our newly

developed QML method. This transition aims to improve the robustness and accuracy of noise suppression in ultrasound imaging.
Future work will focus on validating our approach across diverse datasets and exploring its potential in real-world scenarios such as anomaly detection and image denoising.

References

1. Amirmazlaghani, M., Amindavar, H. (2007) Image denoising using two- dimensional GARCH model. Systems, Signals and image processing. 397-400. https://doi.org/10.1109/iwssip.2007.4381125
2. Boulemnadjel, A., Hachouf, F. and Kharfouchi, S. (2016). GMM estimation of 2D-RCA models with applications to texture image classification IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society 25 (2):528-39. https://doi.org/10.1109/tip.2015.2494740
3. Etchison, T., Pantula, S. G., Brownie, C. (1994). Partial autocorrelation function for spatial processes. Statist. Probab. Lett. 21:9-19. https://doi.org/10.1016/0167-7152(94)90052-3
4. Francq, C., Zakoian, J.M. (2011): GARCH models: structure, statistical inference and financial applications (John Wiley and Sons, Hoboken. https://doi.org/10.1111/j.1751-5823.2011.00149-27.x
5. Kharfouchi, S. (2014): Inference for 2D-GARCH models, Stat. Probab. Lett., 92, pp. 99-108. https://doi.org/10.1016/j.spl.2014.05.007
6. Kharfouchi, S., W Mili.: (2023). Estimating 2-D GARCH models by quasi maximum of likelihood. Communications in Statistics-Theory and Methods 52 (17), 6275-6286. https://doi.org/10.1080/03610926.2022.2027452
7. Kumar, J., Jilowa, A. K., Deokar, M. (2024):Volatility modeling of cryptocurrency and identifying common GARCH model. Communications in Statistics: Case Studies, Data Analysis and Applications, 1-18. https://doi.org/10.1080/23737484.2024.2351998
8. Mazlaghani, M.A., Amindavar, H. (2012): Wavelet domain Bayesian processor for speckle removal in medical ultrasound images, IET Image Process. 6, (5), pp. 580-588. https://doi.org/10.1049/iet-ipr.2009.0110
9. Mazlaghani, M.A., Amindavar, H., Damjoo, A.M. (2009): Speckle suppression in SAR images using the 2D GARCH model, IEEE Trans. Image Process. 18, (2), pp. 250-259. https://doi.org/10.1109/tip.2008.2009857
10. Noiboar and I. Cohen (2007). Anomaly detection based on wavelet domain GARCH random field modeling. IEEE, Trans. Geos. Romote Sensing, 5(5). 1361-1373. https://doi.org/10.1109/tgrs.2007.893741
11. Raslain, S., Hachouf, F., Kharfouchi, S.: (2018). Using a generalised method of moment approach and 2D-generalised autoregressive conditional heteroscedasticity modelling for denoising ultrasound images1962IET Image Processing 12 (11), 2011-2022. https://doi.org/10.1049/iet-ipr.2018.5528
12. Romberg, J.K., Choi, H., Baraniuk, R.G. (1999): Bayesian wavelet-domain imagemodeling using hidden markov trees, In Image Processing (ICIP) IEEE, 1, pages 158-162. https://doi.org/10.1109/icip.1999.821586
13. Yao, Q. and Brockwell, P. (2006) Gaussian Maximum likelihood estimation for ARMA models II: Spatial processes. Bernoulli. 12(3), 403-429. https://doi.org/10.3150/bj/1151525128

[1] Departement of Mathematics, Freres Mentouri University Constantine, Constantine, Algeria.

Email address: wafa.mili@umc.edu.dz

[2] Biostim Laboratory, Faculty of Medicine, Constantine 3 University, Constantine, Algeria.

Email address: s kharfouchi@yahoo.fr