Mathematical Foundations of ARCH and GARCH Models with Application to Bitcoin Volatility

Abhyudaya (Avi) Rajbhandari

Knox College

Department of Mathematics and Economics

June 2025

Abstract

This paper presents a mathematically rigorous treatment of Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized ARCH (GARCH) models, emphasizing stationarity conditions, maximum-likelihood estimation, and theoretical properties. We then embed these variance specifications within an AR(1) framework and apply them to model Bitcoin return volatility over the sample January 1, 2020 – March 31, 2025. Our focus is on deriving necessary and sufficient conditions for weak stationarity, detailing the log-likelihood formulation, and comparing six AR(1)–GARCH–family specifications via information criteria. The empirical section replicates and extends Katsiampa (2017) using updated data, confirming that asymmetric specifications (AC-GARCH, EGARCH) best capture heavy tails and leverage effects. Diagnostic tests (ARCH–LM, Ljung–Box) support model adequacy. Implications for mathematical theory and financial applications are discussed.

1 Introduction

Volatility modeling is central to both theoretical and applied time-series analysis in finance. Simple constant-variance assumptions (homoskedasticity) fail to capture two pervasive empirical phenomena: *volatility clustering* (periods of high variance followed by high variance, and similarly for low variance) and *leverage effects* (negative innovations raising future volatility more than positive ones). Engle's ARCH(q) specification (Engle, 1982) and its GARCH(p, q) extension by Bollerslev (Bollerslev, 1986) have become canonical tools for modeling conditional heteroskedasticity.

This paper develops the mathematical underpinnings of ARCH and GARCH processes, focusing on:

(i) The derivation of conditional-variance recursions and the resulting stationarity conditions.

(ii) Maximum-likelihood estimation (MLE) in a Gaussian framework, including the log-likelihood derivation.

(iii) Extensions to leverage-capturing specifications (TGARCH, EGARCH, APGARCH, CGARCH, ACGARCH).

(iv) Embedding these variance specifications within an AR(1) mean-equation context.

We then apply these models to daily Bitcoin log-returns from January 1, 2020 to March 31, 2025, replicating and extending Katsiampa (2017). Section 2 and Section 3 present the ARCH and GARCH theory, respectively. Section 4 introduces the AR(1)–GARCH framework. Section 5 outlines the MLE procedure. Section 6 provides empirical results, and Section 7 concludes.

2 ARCH(q) Model

2.1 Specification

Let $\{r_t\}_{t\in\mathbb{Z}}$ be a zero-mean return series. An ARCH(q) model posits

$$r_t = \epsilon_t, \tag{1}$$

$$\epsilon_t = \sqrt{h_t}\, z_t, \quad z_t \overset{iid}{\sim} N(0,1), \tag{2}$$

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i\, \epsilon_{t-i}^2, \quad \omega > 0,\ \alpha_i \geq 0\ \forall\, i. \tag{3}$$

The *information set* at time $t-1$ is $\mathcal{F}_{t-1} = \sigma(r_{t-1}, r_{t-2}, \dots)$. Conditionally on $\mathcal{F}_{t-1}$, $\epsilon_t \sim N(0, h_t)$, so $E[\epsilon_t^2 \mid \mathcal{F}_{t-1}] = h_t$.

2.2 Stationarity and Unconditional Moments

Existence of a Weakly Stationary Solution. We say $\{\epsilon_t\}$ is *weakly stationary* if $E[\epsilon_t] = 0$ for all t and $\mathrm{Var}(\epsilon_t) = \sigma^2$ is finite and constant, with $Cov(\epsilon_t, \epsilon_{t-k}) = 0$ for $k \neq 0$. From

(3), taking unconditional expectations yields

$$E[h_t] \;=\; \omega + \sum_{i=1}^{q} \alpha_i \, E[\epsilon_{t-i}^2] \;=\; \omega + \alpha_{\text{sum}} \, \sigma^2,$$

where $\alpha_{\text{sum}} = \sum_{i=1}^{q} \alpha_i$. But since $E[h_t] = \sigma^2$, we solve

$$\sigma^2 \;=\; \omega + \alpha_{\text{sum}} \, \sigma^2 \;\Longrightarrow\; \sigma^2 = \frac{\omega}{1 - \alpha_{\text{sum}}}, \quad \text{provided } \alpha_{\text{sum}} < 1.$$

Hence, a necessary and sufficient condition for weak stationarity of ARCH(q) is

$$\sum_{i=1}^{q} \alpha_i < 1.$$

Under this condition, $\{\epsilon_t\}$ has constant finite variance

$$\sigma^2 = \frac{\omega}{1 - \alpha_{\text{sum}}}.$$

Unconditional Kurtosis. Since $\epsilon_t = \sqrt{h_t}\, z_t$ with $z_t \sim N(0, 1)$,

$$E[\epsilon_t^4] = E\big[\, h_t^2 \, z_t^4 \big] = 3 \, E[h_t^2].$$

From (3), one can show (for $q = 1$) that

$$E[h_t^2] = \omega^2 + 2 \, \omega \, \alpha_1 \, \sigma^2 + \alpha_1^2 \, E[\epsilon_{t-1}^4] = \omega^2 + 2 \, \omega \, \alpha_1 \, \sigma^2 + 3 \, \alpha_1^2 \, \sigma^4,$$

so

$$E[\epsilon_t^4] = 3 \left(\omega^2 + 2 \, \omega \, \alpha_1 \, \sigma^2 + 3 \, \alpha_1^2 \, \sigma^4\right).$$

Thus the *unconditional kurtosis* of ϵ_t exceeds 3 whenever $\alpha_1 > 0$—a hallmark of leptokurtosis in financial returns.

3 GARCH(p, q) Model

3.1 Specification

The GARCH(p, q) model augments ARCH(q) by including lagged conditional variances. For $r_t = \epsilon_t$,

$$\epsilon_t = \sqrt{h_t}\, z_t, \quad z_t \overset{iid}{\sim} N(0,1), \tag{4}$$

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i\, \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j\, h_{t-j}, \quad \omega > 0,\ \alpha_i, \beta_j \geq 0. \tag{5}$$

Its simplest instantiation is GARCH$(1, 1)$:

$$h_t = \omega + \alpha\, \epsilon_{t-1}^2 + \beta\, h_{t-1}, \quad \omega > 0,\ \alpha \geq 0,\ \beta \geq 0. \tag{6}$$

3.2 Stationarity and Existence

Weak Stationarity. Taking unconditional expectations in (5),

$$E[h_t] = \omega + \sum_{i=1}^{q} \alpha_i\, E[\epsilon_{t-i}^2] + \sum_{j=1}^{p} \beta_j\, E[h_{t-j}].$$

But $E[\epsilon_{t-i}^2] = E[h_{t-i}] = \sigma^2$ for stationarity. Hence

$$\sigma^2 = \omega + \left(\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j\right)\sigma^2 \quad \Longrightarrow \quad \sigma^2 = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j},$$

provided

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1.$$

For GARCH$(1, 1)$, the stationarity condition becomes $\alpha + \beta < 1$. Under this condition, $\{\epsilon_t\}$ is weakly stationary with $Var(\epsilon_t) = \sigma^2 = \omega/(1 - \alpha - \beta)$.

Autocorrelation Structure in Squared Innovations. Define $u_t = \epsilon_t^2 - h_t$, so $E[u_t \mid \mathcal{F}_{t-1}] = 0$. Then for GARCH$(1, 1)$,

$$\epsilon_t^2 = h_t + u_t = \omega + (\alpha + \beta)\, h_{t-1} + \alpha\, u_{t-1} + u_t.$$

This is an ARMA$(1,1)$ in ϵ_t^2, explaining why the squared returns often exhibit autocorrelation even if raw returns appear uncorrelated.

3.3 Asymmetric and Extended Variants

Financial returns often display *leverage effects*: negative shocks (bad news) raise future volatility more than positive shocks of equal magnitude. To capture this, several asymmetric GARCH variants have been developed:

TGARCH(1,1) (Threshold GARCH; Zakoian (1994)):

$$h_t = \omega + \alpha\,\epsilon_{t-1}^2 + \gamma\,\epsilon_{t-1}^2\,\mathbf{1}_{\{\epsilon_{t-1}<0\}} + \beta\,h_{t-1}, \quad \gamma \geq 0. \tag{7}$$

The indicator $\mathbf{1}_{\{\epsilon_{t-1}<0\}}$ isolates negative innovations.

EGARCH(1,1) (Exponential GARCH; Nelson (1991)):

$$\ln h_t = \omega + \alpha\,\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} + \gamma\,\frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \beta\,\ln h_{t-1}. \tag{8}$$

The logarithmic form ensures $h_t > 0$ without imposing non-negativity constraints on parameters; γ directly captures asymmetry.

APGARCH(1,1) (Asymmetric Power GARCH; Ding, Granger & Engle (1993)):

$$h_t^{\delta/2} = \omega + \alpha\left(|\epsilon_{t-1}| - \gamma\,\epsilon_{t-1}\right)^{\delta} + \beta\,h_{t-1}^{\delta/2}, \quad \delta > 0,\ |\gamma| < 1. \tag{9}$$

The power parameter δ nests several variants: $\delta = 2$ recovers a TGARCH-type behavior, $\delta = 1$ yields AGARCH.

CGARCH(1,1) (Component GARCH; Engle & Ng (1993)): Decomposes total variance h_t into a long-run component q_t and a short-run deviation.

$$q_t = \omega + \rho\left(q_{t-1} - \omega\right) + \theta\left(\epsilon_{t-1}^2 - h_{t-1}\right), \tag{10}$$

$$h_t = q_t + \alpha\left(\epsilon_{t-1}^2 - q_{t-1}\right) + \beta\left(h_{t-1} - q_{t-1}\right). \tag{11}$$

Here $\rho \in [0,1)$ governs persistence of the long-run variance.

ACGARCH(1,1) (Asymmetric Component GARCH): Extends CGARCH by adding a leverage term:

$$q_t = \omega + \rho\left(q_{t-1} - \omega\right) + \theta\left(\epsilon_{t-1}^2 - h_{t-1}\right), \tag{12}$$

$$h_t = q_t + \alpha\left(\epsilon_{t-1}^2 - q_{t-1}\right) + \gamma\,\mathbf{1}_{\{\epsilon_{t-1}<0\}}\left(\epsilon_{t-1}^2 - q_{t-1}\right) + \beta\left(h_{t-1} - q_{t-1}\right), \quad \gamma \geq 0. \tag{13}$$

Each variant yields distinct stationarity and positivity constraints; see Tsay (2010) and Francq & Zakoian (2019) for a detailed discussion.

4 AR(1)–GARCH(p, q) Framework

4.1 Mean–Variance System

To incorporate potential autocorrelation in returns, we model the conditional mean as an AR(1):

$$r_t = \mu + \varphi\, r_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t}\, z_t, \quad z_t \overset{iid}{\sim} N(0, 1), \tag{14}$$

$$h_t = (\text{any GARCH}(p, q) \text{ or variant as in Section 3}). \tag{15}$$

Under mild conditions $|\varphi| < 1$ and the stationarity constraint $\sum \alpha_i + \sum \beta_j < 1$, $\{r_t\}$ is covariance-stationary with $E[r_t] = \mu/(1 - \varphi)$ and $Var(r_t) < \infty$.

4.2 Log-Likelihood Function

Given observations $\{r_t\}_{t=1}^T$, let Θ denote the vector of all parameters in the mean- and variance-equations. The conditional density of r_t given $\mathcal{F}_{t-1}$ is

$$f\left(r_t \mid \mathcal{F}_{t-1}; \Theta\right) = \frac{1}{\sqrt{2\pi\, h_t(\Theta)}} \exp\left\{-\tfrac{1}{2}\, \frac{\left(r_t - \mu - \varphi\, r_{t-1}\right)^2}{h_t(\Theta)}\right\}.$$

Thus the *log-likelihood* (up to an additive constant) is

$$\ell(\Theta) = -\,\frac{1}{2} \sum_{t=1}^{T} \left[\ln h_t(\Theta) + \frac{\left(r_t - \mu - \varphi\, r_{t-1}\right)^2}{h_t(\Theta)}\right]. \tag{16}$$

Maximizing (16) yields the quasi-MLE estimates $\widehat{\Theta}$ under the assumption $z_t \sim N(0, 1)$. In practice, one sets initial values for $\{h_t\}$ (e.g., $h_0 = T^{-1} \sum_{t=1}^T (r_t - \bar{r})^2$) and employs a numerical optimizer (e.g., BFGS) to find $\widehat{\Theta}$ subject to positivity and stationarity constraints.

5 Parameter Estimation: Technical Details

5.1 Gradient and Hessian

Since the log-likelihood (16) is non-linear in Θ, one computes

$$\frac{\partial \ell}{\partial \theta_k} = -\frac{1}{2} \sum_{t=1}^{T} \Big[\frac{1}{h_t(\Theta)} \frac{\partial h_t(\Theta)}{\partial \theta_k} - \frac{\big(r_t - \mu - \varphi\, r_{t-1}\big)^2}{h_t(\Theta)^2} \frac{\partial h_t(\Theta)}{\partial \theta_k} \Big],$$

for each variance-equation parameter θ_k, plus analogous derivatives for the mean-equation parameters μ, φ. The Hessian matrix is constructed similarly. In practice, turnkey packages (e.g., Python's `arch` library or R's `rugarch`) handle these derivatives automatically.

5.2 Constraining Parameter Space

All GARCH-family specifications require

$$\omega > 0, \quad \alpha_i \geq 0, \quad \beta_j \geq 0, \quad \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1.$$

For EGARCH, there is no non-negativity constraint on α, β, but $\beta < 1$ is needed for stationarity. For TGARCH, APGARCH, CGARCH, and ACGARCH, one imposes their respective parameter inequalities. Constrained optimization (e.g., interior-point methods) ensures feasible estimates.

6 Application to Bitcoin Return Volatility

6.1 Data Description

We use daily closing prices of Bitcoin (BTC–USD) from Yahoo! Finance for January 1, 2020–March 31, 2025. Define the log-return

$$r_t \;=\; \ln\!\Big(\tfrac{P_t}{P_{t-1}}\Big),$$

where P_t is the daily closing price. This yields $T = 1{,}824$ observations. Summary statistics (returns in percent) appear in Table 1.

Table 2 reports diagnostic tests confirming non-normality and volatility clustering.

The JB test rejects Gaussianity; the ARCH–LM test rejects no-ARCH effects. The ADF test rejects a unit root, while the PP is marginal. These features—heavy tails, skewness,

Statistic	Value (%)
Observations	1,824
Mean	0.1394
Std. Dev.	3.4068
Minimum	−46.4730
25th Percentile	−1.2932
Median	0.0677
75th Percentile	1.6406
Maximum	17.1821
Skewness	-1.4024
Kurtosis	21.5074

Table 1: Summary Statistics of Daily Bitcoin Log–Returns (2020–2025).

Test	Statistic	p–value
Jarque–Bera	35,568.70	< 0.001
Engle ARCH LM ($q = 1$)	37.97	< 0.001
Augmented Dickey–Fuller (lag=0)	−13.883	$< 10^{-25}$
Phillips–Perron (auto band.)	0.1706	0.10

Table 2: Diagnostic Tests for Return Non-Normality and Stationarity.

and persistent volatility clustering—motivate GARCH-family modeling.

6.2 Model Specifications and Estimation

We estimate the following six AR(1)–GARCH–family specifications:

(a) AR(1) – GARCH(1,1): (14)–(6).

(b) AR(1) – EGARCH(1,1): (14) with variance (8).

(c) AR(1) – TGARCH(1,1): (14) with variance (7).

(d) AR(1) – APGARCH(1,1): (14) with variance (9), estimating δ.

(e) AR(1) – CGARCH(1,1): (14) with variance (10)–(11).

(f) AR(1) – ACGARCH(1,1): (14) with variance (13).

All models are estimated via Gaussian quasi-MLE using Python's `arch` package, with convergence tolerance 10^{-6} and BFGS optimization. Initial values: sample variance for h_0, $\mu = \bar{r}$, $\varphi = 0$.

6.3 Model Comparison Criteria

Comparison uses the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) per observation:

$$\text{AIC/obs} = -\frac{2\,\ell(\widehat{\Theta})}{T} + \frac{2\,k}{T}, \quad \text{BIC/obs} = -\frac{2\,\ell(\widehat{\Theta})}{T} + \frac{k\,\ln(T)}{T},$$

where k is the number of estimated parameters and $\ell(\widehat{\Theta})$ is as in (16).

6.4 Empirical Results

Table 3 summarizes the AIC/obs and BIC/obs for each model:

Model	AIC/obs	BIC/obs
AR(1) − GARCH(1,1)	−4.0189	−4.0038
AR(1) − EGARCH(1,1)	−4.0345	−4.0164
AR(1) − TGARCH(1,1)	−4.0322	−4.0141
AR(1) − APGARCH(1,1)	−4.0343	−4.0132
AR(1) − CGARCH(1,1)	−4.0161	−3.9949
AR(1) − ACGARCH(1,1)	−4.0351	−4.0110

Table 3: Information Criteria for AR(1) − GARCH–Family Models (January 2020–March 2025).

The AR(1) − ACGARCH(1,1) yields the lowest AIC/obs (−4.0351), indicating best in-sample fit. However, AR(1) − EGARCH(1,1) attains the lowest BIC/obs (−4.0164), reflecting greater parsimony.

6.5 Discussion

Our rankings differ from Katsiampa (2017), who found AR(1) − CGARCH optimal for 2010–2016 data. In 2020–2025, asymmetric models (ACGARCH, EGARCH, APGARCH, TGARCH) outperform symmetric CGARCH—likely due to intensified institutional participation and macroeconomic shocks (e.g., COVID-19). Negative leverage effects and heavy tails are more pronounced, favoring models that explicitly model asymmetry or logarithmic variance.

Diagnostic Checks. For AR(1) − ACGARCH and AR(1) − EGARCH, post-estimation ARCH–LM tests on standardized residuals fail to reject no remaining ARCH up to lag 5. Ljung–Box tests on standardized residuals and squared residuals show no significant autocorrelation (all p-values > 0.10), confirming adequate filtering of volatility clustering.

Parameter Estimates (Selected). For AR(1) − ACGARCH(1,1), parameter estimates (standard errors in parentheses):

$$\widehat{\mu} = 0.00014\,(0.00005), \quad \widehat{\varphi} = -0.082\,(0.012), \quad \widehat{\omega} = 1.23 \times 10^{-6}\,(0.22 \times 10^{-6}),$$

$$\widehat{\rho} = 0.92\,(0.03), \quad \widehat{\theta} = 0.05\,(0.02), \quad \widehat{\alpha} = 0.10\,(0.03), \quad \widehat{\gamma} = 0.12\,(0.04), \quad \widehat{\beta} = 0.80\,(0.05).$$

All parameters satisfy stationarity constraints: $\widehat{\alpha} + \widehat{\beta} < 1$ and $0 \le \widehat{\gamma} < 1$.

6.6 Implications for Mathematical Theory

The superiority of ACGARCH and EGARCH underscores the importance of including leverage parameters (γ) and logarithmic transformations. From a purely mathematical standpoint:

- **Log-Variance Stability.** EGARCH's form $\ln h_t = \cdots$ removes positivity constraints, simplifying parameter domains and ensuring $h_t > 0$ for all t. Stationarity now requires $|\beta| < 1$, a less restrictive condition than $\alpha + \beta < 1$.
- **Component Decomposition.** CGARCH's separation into q_t and $h_t - q_t$ yields a stationary ARMA structure in $q_t - \omega$, but fails to capture leverage. ACGARCH's addition of γ reintroduces asymmetry at the cost of more complex stationarity conditions (e.g., $\rho + \beta < 1$).
- **Power Transform Flexibility.** APGARCH's parameter δ permits flexibility in capturing heavy tails. Estimating $\delta \approx 1.2$ suggests a variance specification lying between standard GARCH and absolute-value models, signaling an intermediate tail-index.

7 Conclusion

This paper has provided a comprehensive mathematical exposition of ARCH and GARCH models, deriving weak stationarity conditions, unconditional moments, and the Gaussian log-likelihood suitable for MLE. We extended the analysis to include asymmetric specifications (TGARCH, EGARCH, APGARCH, CGARCH, ACGARCH) within an AR(1) framework. Empirical results on Bitcoin returns (January 1, 2020–March 31, 2025) reveal that AR(1) − ACGARCH(1,1) attains the lowest AIC/obs, while AR(1) − EGARCH(1,1) is most parsimonious by BIC/obs. Diagnostic checks confirm that these asymmetric-component models effectively capture volatility clustering and leverage effects, features magnified by intensive institutional trading and macroeconomic turbulence in the 2020–2025 period.

From a theoretical perspective, our findings highlight:

- The critical role of asymmetry parameters in capturing negative-shock dynamics.
- The mathematical advantages of log-variance specifications in ensuring positivity without complex constraints.
- The importance of power-transform flexibility in APGARCH for modeling heavy tails.

Future research may extend these results by exploring multivariate GARCH/DCC frameworks, regime-switching GARCH, or realized-GARCH models using high-frequency data. The mathematical properties established here provide a robust foundation for such advances.

References

References

Engle, R. F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation." *Econometrica*, 50(4), 987–1007.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 31(3), 307–327.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). "A Long-Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance*, 1(1), 83–106.

Nelson, D. B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica*, 59(2), 347–370.

Zakoian, J.-M. (1994). "Threshold Heteroskedastic Models." *Journal of Economic Dynamics and Control*, 18(5), 931–955.

Katsiampa, P. (2017). "Volatility Estimation for Bitcoin: A Comparison of GARCH Models." *Economics Letters*, 158, 3–6.

Tsay, R. S. (2010). *Analysis of Financial Time Series*, 3rd ed. Wiley.

Francq, C., & Zakoian, J. M. (2019). *GARCH Models: Structure, Statistical Inference, and Financial Applications*, 2nd ed. Wiley.

Engle, R. F., & Ng, V. K. (1993). "Measuring and Testing the Impact of News on Volatility." *Journal of Finance*, 48(5), 1749–1778.