Supplementary Material for "Bubble Modeling and Tagging: A Stochastic Nonlinear Autoregression Approach"

Xuanling Yang[1], Dong Li[2], and Ting Zhang[3]

[1]*Beijing University of Chemical Technology*
[2]*Tsinghua University, and* [3]*University of Georgia*

The Supplementary Material contains part of simulation results and all technical proofs of theorems and propositions in the article. It also provides the explicit expressions of $\mathcal{I}$ and $\mathcal{J}$ in Theorem 3.

S1 Part of simulation results

In the simulation studies, the error ε_t follows

- $\mathcal{N}(0,1)$;
- the Laplace distribution with density

$$h(x) = \frac{1}{\sqrt{2}} \exp\big(-\sqrt{2}|x|\big), \quad x \in \mathbb{R};$$

- the standardized Student's t_5 (st_5) with density

$$h(x) = \frac{8}{3\sqrt{3}\pi}(1 + x^2/3)^{-3}, \quad x \in \mathbb{R}.$$

Table S.1 reports the bias, empirical standard deviation (ESD), and asymptotic standard deviation (ASD) of the QMLE $\widehat{\theta}_n$ for Cases I-III. Here, the ASD of θ_0 is simulated by extra time series of length 10,000, and 2,000 replications are used to reduce the estimated bias. From the table, we can see that the QMLE performs well irrespective of infinite variance or heavy-tailedness issues. The biases are small and all the ESDs are close to the corresponding ASDs.

To see the overall approximation of the QMLE $\widehat{\phi}_n$, Fig. S.1 displays the histogram of $\sqrt{n}\big(\widehat{\phi}_n - \phi_0\big)$ when the sample size $n = 400$. From the figure, we can see that $\sqrt{n}\big(\widehat{\phi}_n - \phi_0\big)$ is always asymptotically normal irrespective of infinite variance or heavy-tailedness of y_t.

Tables S.2 and S.3 report the complete results of the finite-sample performance of the two tagging methods described in Section 5.

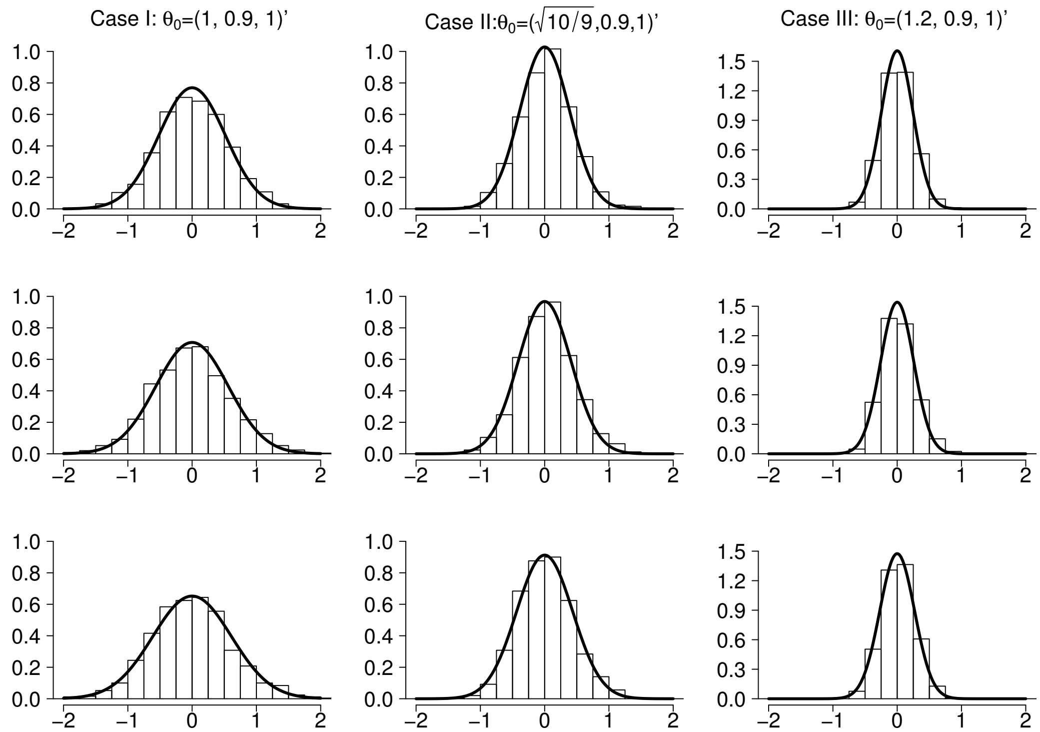

Figure S.1: The histogram of $\sqrt{n}(\widehat{\phi}_n - \phi_0)$ with the sample size $n = 400$. The left column panel corresponds to Case I, i.e., y_t is weakly stationary; the middle to Case II, and the right to Case III, i.e., y_t has an infinite variance, respectively. The upper row panel is when $\varepsilon_t \sim \mathcal{N}(0,1)$, the middle when $\varepsilon_t \sim$ the Laplace distribution, and the lower when $\varepsilon_t \sim \mathrm{st}_5$, respectively.

Table S.1: Numerical simulation results.

n		$(\phi_0, p_0, \sigma_0^2)$								
		$(1, 0.9, 1)$			$(\sqrt{10/9}, 0.9, 1)$			$(1.2, 0.9, 1)$		
		$\widehat{\phi}_n$	$\widehat{p}_n$	$\widehat{\sigma}_n^2$	$\widehat{\phi}_n$	$\widehat{p}_n$	$\widehat{\sigma}_n^2$	$\widehat{\phi}_n$	$\widehat{p}_n$	$\widehat{\sigma}_n^2$
	$\varepsilon_t \sim \mathcal{N}(0,1)$									
200	Bias	-0.0012	-0.0042	0.0098	0.0026	-0.0077	-0.0059	0.0018	-0.0038	-0.0060
	ESD	0.0431	0.0467	0.1673	0.0307	0.0376	0.2047	0.0185	0.0284	0.1993
	ASD	0.0367	0.0438	0.1559	0.0274	0.0373	0.1652	0.0176	0.0297	0.1913
400	Bias	-0.0002	-0.0026	0.0044	0.0009	-0.0030	0.0014	0.0007	-0.0020	0.0056
	ESD	0.0265	0.0321	0.1112	0.0193	0.0266	0.1160	0.0126	0.0220	0.1378
	ASD	0.0259	0.0310	0.1102	0.0194	0.0263	0.1168	0.0124	0.0210	0.1353
800	Bias	0.0002	-0.0010	0.0009	0.0009	-0.0021	-0.0006	0.0005	-0.0004	-0.0009
	ESD	0.0192	0.0217	0.0774	0.0145	0.0193	0.0864	0.0087	0.0145	0.0979
	ASD	0.0183	0.0219	0.0779	0.0137	0.0186	0.0826	0.0088	0.0148	0.0957
	$\varepsilon_t \sim$ Laplace									
200	Bias	0.0020	-0.0068	0.0143	0.0036	-0.0093	-0.0067	0.0035	-0.0045	-0.0098
	ESD	0.0465	0.0492	0.2341	0.0323	0.0400	0.2350	0.0205	0.0313	0.2738
	ASD	0.0399	0.0463	0.2199	0.0292	0.0384	0.2299	0.0183	0.0301	0.2626
400	Bias	0.0015	-0.0056	-0.0070	0.0018	-0.0037	0.0034	0.0006	-0.0015	0.0004
	ESD	0.0289	0.0326	0.1545	0.0227	0.0286	0.1580	0.0135	0.0201	0.1950
	ASD	0.0282	0.0328	0.1555	0.0206	0.0272	0.1626	0.0129	0.0213	0.1857
800	Bias	-0.0001	-0.0007	0.0034	0.0015	-0.0024	0.0000	0.0007	-0.0014	-0.0040
	ESD	0.0214	0.0235	0.1116	0.0150	0.0201	0.1179	0.0093	0.0151	0.1314
	ASD	0.0199	0.0232	0.1099	0.0146	0.0192	0.1150	0.0092	0.0151	0.1313
	$\varepsilon_t \sim \mathrm{st}_5$									
200	Bias	0.0002	-0.0062	0.0046	0.0031	-0.0085	-0.0101	0.0029	-0.0042	-0.0074
	ESD	0.0541	0.0527	0.2512	0.0345	0.0416	0.2461	0.0209	0.0315	0.2948
	ASD	0.0433	0.0487	0.2712	0.0309	0.0395	0.2824	0.0191	0.0304	0.3224
400	Bias	0.0021	-0.0050	0.0016	0.0018	-0.0043	0.0063	0.0019	-0.0028	-0.0071
	ESD	0.0312	0.0354	0.1833	0.0234	0.0291	0.1880	0.0149	0.0215	0.1910
	ASD	0.0306	0.0345	0.1918	0.0219	0.0279	0.1997	0.0135	0.0215	0.2280
800	Bias	-0.0001	-0.0014	0.0083	0.0004	-0.0019	0.0019	0.0009	-0.0013	-0.0033
	ESD	0.0224	0.0239	0.1353	0.0150	0.0195	0.1312	0.0091	0.0151	0.1523
	ASD	0.0216	0.0244	0.1356	0.0155	0.0197	0.1412	0.0096	0.0152	0.1612

Table S.2: The values (in percentage) of P, P0, and P1 for RBT_1-RBT_4 and NBT when $n = 200$.

	$\phi_0 = 1, p_0 = 0.9$			$\phi_0 = \sqrt{10/9}, p_0 = 0.9$			$\phi_0 = 1.2, p_0 = 0.9$		
Method	P	P0	P1	P	P0	P1	P	P0	P1
				$\varepsilon_t \sim \mathcal{N}(0,1)$					
RBT_1	90.93	53.15	94.73	92.15	59.34	95.37	93.44	65.62	96.12
RBT_2	84.64	21.33	91.25	85.45	25.99	91.67	87.81	38.80	93.01
RBT_3	90.47	51.22	94.48	91.68	57.51	95.11	93.75	67.68	96.25
RBT_4	91.49	56.11	95.04	92.81	62.85	95.73	95.02	74.09	96.99
NBT	87.01	33.20	92.56	87.63	36.48	92.87	89.16	44.62	93.75
				$\varepsilon_t \sim$ Laplace					
RBT_1	90.56	51.39	94.51	91.66	56.92	95.12	92.89	62.91	95.85
RBT_2	84.50	20.82	91.16	85.24	24.77	91.58	88.02	39.95	93.16
RBT_3	90.24	50.10	94.33	91.40	55.91	94.98	93.58	66.95	96.19
RBT_4	91.42	55.71	94.98	92.55	61.56	95.62	94.88	73.72	96.95
NBT	86.53	30.93	92.28	87.22	34.55	92.67	89.11	44.52	93.76
				$\varepsilon_t \sim st_5$					
RBT_1	90.63	51.80	94.56	91.87	58.01	95.20	93.10	63.99	95.94
RBT_2	84.49	20.64	91.16	85.33	25.62	91.59	87.97	39.76	93.10
RBT_3	90.36	50.61	94.41	91.61	57.11	95.06	93.77	67.86	96.26
RBT_4	91.40	55.71	94.98	92.81	62.86	95.72	94.97	74.05	96.97
NBT	86.69	31.66	92.38	87.47	35.97	92.77	89.22	44.92	93.79

Table S.3: The values (in percentage) of P, P0, and P1 for RBT_1-RBT_4 and NBT when $n = 200$.

	$\phi_0 = 1, p_0 = 0.5$			$\phi_0 = \sqrt{10/9}, p_0 = 0.5$			$\phi_0 = 1.2, p_0 = 0.5$		
Method	P	P0	P1	P	P0	P1	P	P0	P1
				$\varepsilon_t \sim \mathcal{N}(0,1)$					
RBT_1	66.70	66.79	66.69	67.77	67.87	67.75	69.93	70.05	69.92
RBT_2	68.13	68.24	68.11	69.45	69.56	69.43	72.16	72.32	72.15
RBT_3	68.09	68.20	68.08	69.37	69.48	69.35	71.93	72.09	71.92
RBT_4	68.32	68.43	68.31	69.55	69.66	69.53	72.42	72.58	72.41
NBT	66.85	66.95	66.83	67.97	68.07	67.95	70.00	70.16	69.99
				$\varepsilon_t \sim$ Laplace					
RBT_1	67.84	68.01	67.98	68.75	68.85	68.73	70.60	70.72	70.63
RBT_2	70.07	70.26	70.21	71.28	71.41	71.26	73.55	73.70	73.58
RBT_3	69.95	70.15	70.09	71.25	71.38	71.24	73.53	73.67	73.55
RBT_4	70.18	70.38	70.32	71.39	71.52	71.37	73.81	73.97	73.83
NBT	68.13	68.32	68.28	69.00	69.12	68.99	70.75	70.88	70.77
				$\varepsilon_t \sim \text{st}_5$					
RBT_1	67.05	67.16	67.12	67.92	67.97	67.78	69.83	69.91	69.77
RBT_2	68.93	69.07	69.00	70.16	70.23	70.03	72.74	72.85	72.68
RBT_3	68.88	69.02	68.95	70.08	70.14	69.94	72.55	72.66	72.49
RBT_4	68.99	69.13	69.06	70.20	70.28	70.07	72.92	73.04	72.85
NBT	67.33	67.47	67.40	68.30	68.37	68.16	70.31	70.42	70.24

S2 Technical Proofs

S2.1 Proof of Theorem 1

When $\phi_0 = 0$, then $\{y_t\}$ reduces to an i.i.d. sequence $\{\varepsilon_t\}$, and in this case all results hold clearly. Without loss of generality, we assume that $\phi_0 \neq 0$ in what follows. It suffices to verify the conditions in Theorem 19.1.3 in Meyn and Tweedie (2009). It is clear that $\{y_t\}$ defined by (1.1), with initial value y_0, is an homogeneous Markov chain on $\mathbb{R}$ endowed with its Borel σ-field $\mathcal{B}(\mathbb{R})$. Denote by λ the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The transition probabilities of $\{y_t\}$ are given, for $y \in \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$, by

$$\mathbf{P}(y, B) = \mathbb{P}(y_t \in B | y_{t-1} = y) = p_0 \mathbb{P}(\varepsilon_1 + \phi_0 |y| \in B) + (1 - p_0)\mathbb{P}(\varepsilon_1 \in B).$$

First, since $\mathbf{P}(\cdot, B)$ is continuous, for any $B \in \mathcal{B}(\mathbb{R})$, the chain $\{y_t\}$ has the Feller property.

Second, note that the density of ε_1 is positive over $\mathbb{R}$, we have $\mathbf{P}(y, B) > 0$ whenever $\lambda(B) > 0$. Thus the chain $\{y_t\}$ is λ-irreducible. Further, it can also be shown that the k-step transition probabilities $\mathbf{P}^k(y, B) = \mathbb{P}(y_t \in B | y_{t-k} = y) = \int_{\mathbb{R}} \mathbf{P}^{k-1}(x, B)\mathbf{P}(y, dx) > 0$ by an inductive approach for any integer $k \geq 1$, whenever $\lambda(B) > 0$, which establishes the aperiodicity of the chain $\{y_t\}$.

Third, let $V(x) = \log(1 + |x|)$, $x \in \mathbb{R}$. Then, by a simple calculation, it follows that

$$\mathbb{E}\{V(y_t)|y_{t-1}=y\}=(1-p_0)\mathbb{E}\{\log(1+|\varepsilon_1|)\}+p_0\mathbb{E}\{\log(1+|\phi_0|y|+\varepsilon_1|)\}.$$

Thus, we have that

$$\begin{aligned}
&\lim_{|y|\to\infty}\frac{\mathbb{E}\{V(y_t)|y_{t-1}=y\}}{V(y)}\\
&=\lim_{|y|\to\infty}\frac{(1-p_0)\mathbb{E}\{\log(1+|\varepsilon_1|)\}}{\log(1+|y|)}+p_0\lim_{|y|\to\infty}\frac{\mathbb{E}\{\log(1+|\phi_0|y|+\varepsilon_1|)\}}{\log(1+|y|)}\\
&=0+p_0\lim_{|y|\to\infty}\left(\frac{\log|y|}{\log(1+|y|)}+\frac{\mathbb{E}\left\{\log\left(1/|y|+\left|\phi_0+\varepsilon_1/|y|\right|\right)\right\}}{\log(1+|y|)}\right)\\
&=p_0.
\end{aligned}$$

Since $p_0\in[0,1)$, for fixed $\delta\in(0,1-p_0)$, i.e., $p_0<1-\delta<1$, there exists a constant $M>0$ such that

$$\mathbb{E}\{V(y_t)|y_{t-1}=y\}\leq(1-\delta)V(y),\quad \text{when } |y|>M.$$

To sum up the above arguments, by Theorem 19.1.3 in Meyn and Tweedie (2009), there exists a geometrically ergodic solution to model (1.1). The solution is unique since $\mathbb{E}(\log|s_t\phi_0|)=-\infty$. Thus, the results hold and then the proof is complete. □

S2.2 Proof of Theorem 2

Consider $\beta_n(\theta):=\{L_n(\theta)-L_n(\theta_0)\}/n$, $\theta\in\Theta$. By the strong law of large numbers for stationary and ergodic sequences and the inequality $\log x+$

$x^{-1} - 1 \geq 0$ for $x > 0$, a conditional argument yields that

$$
\begin{aligned}
\beta_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\Bigg\{&\log\frac{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2}{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2} \\
&+ \frac{(y_t - p\phi|y_{t-1}|)^2}{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2} - \frac{(y_t - p_0\phi_0|y_{t-1}|)^2}{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2}\Bigg\} \\
\xrightarrow{\text{a.s.}} \mathbb{E}\Bigg\{&\log\frac{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2}{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2} \\
&+ \frac{(y_t - p\phi|y_{t-1}|)^2}{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2} - \frac{(y_t - p_0\phi_0|y_{t-1}|)^2}{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2}\Bigg\} \\
= \mathbb{E}\Bigg\{&\log\frac{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2}{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2} \\
&+ \frac{p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2}{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2} - 1 + \frac{(p\phi - p_0\phi_0)^2 y_{t-1}^2}{p(1-p)\phi^2 y_{t-1}^2 + \sigma^2}\Bigg\} \geq \ 0,
\end{aligned}
$$

where the equality holds if and only if

$$
p(1-p)\phi^2 y_{t-1}^2 + \sigma^2 = p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2 \quad \text{and} \quad (p\phi - p_0\phi_0)^2 = 0 \quad \text{a.s.},
$$

equivalently, $\{p(1-p)\phi^2 - p_0(1-p_0)\phi_0^2\}\, y_{t-1}^2 = \sigma_0^2 - \sigma^2$ a.s. Then

$$
p(1-p)\phi^2 - p_0(1-p_0)\phi_0^2 = 0, \quad \sigma_0^2 - \sigma^2 = 0.
$$

Combining with $(p\phi - p_0\phi_0)^2 = 0$, we have $\phi = \phi_0$, $p = p_0$ and $\sigma^2 = \sigma_0^2$, i.e., $\theta = \theta_0$. The remainder of the proof can be completed by a standard compactness argument and it is thus omitted. □

S2.3 Proof of Theorem 3

Let $q_t(\theta) = p(1-p)\phi^2 y_{t-1}^2 + \sigma^2$ and $q_t := q_t(\theta_0)$. Then the first- and second-order partial derivatives of $q_t(\theta)$ with respect to θ are respectively as follows

$$\frac{\partial q_t(\theta)}{\partial \theta} = \begin{pmatrix} 2p(1-p)\phi y_{t-1}^2 \\ (1-2p)\phi^2 y_{t-1}^2 \\ 1 \end{pmatrix},$$

$$\frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} = \begin{pmatrix} 2p(1-p)y_{t-1}^2 & 2(1-2p)\phi y_{t-1}^2 & 0 \\ & -2\phi^2 y_{t-1}^2 & 0 \\ & & 0 \end{pmatrix}. \tag{S2.1}$$

Using the notation $q_t(\theta)$, we have

$$\ell_t(\theta) = \log\{q_t(\theta)\} + \frac{(y_t - p\phi|y_{t-1}|)^2}{q_t(\theta)}.$$

A simple calculation yields the first-order partial derivatives of $\ell_t(\theta)$ with respect to θ

$$\frac{\partial \ell_t(\theta)}{\partial \theta} = \frac{1}{q_t(\theta)}\frac{\partial q_t(\theta)}{\partial \theta} - \frac{2|y_{t-1}|(y_t - p\phi|y_{t-1}|)}{q_t(\theta)}\vartheta - \frac{(y_t - p\phi|y_{t-1}|)^2}{[q_t(\theta)]^2}\frac{\partial q_t(\theta)}{\partial \theta}, \tag{S2.2}$$

where $\vartheta = (p, \phi, 0)'$, and the second-order partial derivatives

$$\frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta'} = \left\{\frac{1}{q_t(\theta)} - \frac{(y_t - p\phi|y_{t-1}|)^2}{[q_t(\theta)]^2}\right\}\frac{\partial^2 q_t(\theta)}{\partial\theta\partial\theta'}$$
$$+ \left\{\frac{2(y_t - p\phi|y_{t-1}|)^2}{[q_t(\theta)]^3} - \frac{1}{[q_t(\theta)]^2}\right\}\frac{\partial q_t(\theta)}{\partial\theta}\frac{\partial q_t(\theta)}{\partial\theta'}$$
$$+ \frac{2|y_{t-1}|(y_t - p\phi|y_{t-1}|)}{[q_t(\theta)]^2}\left\{\frac{\partial q_t(\theta)}{\partial\theta}\vartheta' + \vartheta\frac{\partial q_t(\theta)}{\partial\theta'}\right\}$$
$$+ \frac{2y_{t-1}^2}{q_t(\theta)}\vartheta\vartheta' - \frac{2|y_{t-1}|(y_t - p\phi|y_{t-1}|)}{q_t(\theta)}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

By the Taylor expansion, by the definition of $\widehat{\theta}_n$, we have

$$0 = \frac{1}{\sqrt{n}}\frac{\partial L_n(\widehat{\theta}_n)}{\partial\theta} = \frac{1}{\sqrt{n}}\frac{\partial L_n(\theta_0)}{\partial\theta} + \frac{1}{n}\frac{\partial^2 L_n(\theta^*)}{\partial\theta\partial\theta'}\sqrt{n}\big(\widehat{\theta}_n - \theta_0\big),$$

where $\theta^* \in \Theta$ and satisfies $\|\theta^* - \theta_0\| \le \|\widehat{\theta}_n - \theta_0\|$. Note that the continuity of $\partial^2\ell_t(\theta)/\partial\theta\partial\theta'$ in θ and the strong law of large numbers for stationary and ergodic sequences, it is not hard to get

$$\frac{1}{n}\frac{\partial^2 L_n(\theta^*)}{\partial\theta\partial\theta'} = \frac{1}{n}\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta'} + o_p(1) = \mathcal{J} + o_p(1).$$

Further, let $\mathcal{F}_t = \sigma(y_j : j \le t)$ be the σ-algebra generated by the random variables $\{y_j : j \le t\}$. By the expressions in (S2.1) and (S2.2), and the following facts

$$\begin{aligned} \mathbb{E}\{y_t - p_0\phi_0|y_{t-1}||\mathcal{F}_{t-1}\} &= \mathbb{E}\{(s_t - p_0)\phi_0|y_{t-1}| + \varepsilon_t|\mathcal{F}_{t-1}\} = 0, \\ \mathbb{E}\{(y_t - p_0\phi_0|y_{t-1}|)^2|\mathcal{F}_{t-1}\} &= p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \sigma_0^2 = q_t, \end{aligned} \tag{S2.3}$$

we have that

$$\mathbb{E}\left\{\frac{\partial \ell_t(\theta_0)}{\partial \theta}\middle|\mathcal{F}_{t-1}\right\}=0,$$

i.e., $\{\partial \ell_t(\theta_0)/\partial\theta\}$ is a martingale difference sequence with respect to $\{\mathcal{F}_t\}$. Thus, by the martingale central limit theorem in Brown (1971), it follows that

$$\frac{1}{\sqrt{n}}\frac{\partial L_n(\theta_0)}{\partial \theta}=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_t(\theta_0)}{\partial \theta}\xrightarrow{d}\mathcal{N}(0,\mathcal{I}),$$

where

$$\mathcal{I}=\mathbb{E}\left\{\frac{\partial \ell_t(\theta_0)}{\partial \theta}\frac{\partial \ell_t(\theta_0)}{\partial \theta'}\right\}. \tag{S2.4}$$

Finally, we have

$$\sqrt{n}\big(\widehat{\theta}_n-\theta_0\big)=-[\mathcal{J}+o_p(1)]^{-1}\frac{1}{\sqrt{n}}\frac{\partial L_n(\theta_0)}{\partial \theta}\xrightarrow{d}\mathcal{N}(0,\mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}). \tag{S2.5}$$

For the explicit expressions of $\mathcal{I}$ and $\mathcal{J}$, please see Section S3 in this Supplementary Material. The proof is complete. □

S2.4 Proof of Theorem 4

According to the definition of $\widehat{\eta}_t$, by Theorem 2 and the strong law of large numbers for a stationary and ergodic sequence, we first have the following facts, as $n \to \infty$,

$$
\begin{aligned}
\bar{\eta} = {} & \phi_0 \frac{1}{n} \sum_{t=1}^{n} (s_t - p_0)|y_{t-1}| I(|y_{t-1}| \le a) + \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t I(|y_{t-1}| \le a) \\
& + (p_0\phi_0 - \widehat{p}_n\widehat{\phi}_n) \frac{1}{n} \sum_{t=1}^{n} |y_{t-1}| I(|y_{t-1}| \le a) \\
\overset{a.s.}{\to} {} & \phi_0 \mathbb{E}(s_t - p_0) \mathbb{E}\{|y_{t-1}| I(|y_{t-1}| \le a)\} + \mathbb{E}(\varepsilon_t) \mathbb{P}(|y_{t-1}| \le a) \\
& + 0 \cdot \mathbb{E}\{|y_{t-1}| I(|y_{t-1}| \le a)\} = 0
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{1}{n} \sum_{t=1}^{n} (\widehat{\eta}_t - \bar{\eta})^2 = \frac{1}{n} \sum_{t=1}^{n} \widehat{\eta}_t^2 - \bar{\eta}^2 \\
& = \phi_0^2 \frac{1}{n} \sum_{t=1}^{n} (s_t - p_0)^2 y_{t-1}^2 I(|y_{t-1}| \le a) + \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2 I(|y_{t-1}| \le a) \\
& \quad + (p_0\phi_0 - \widehat{p}_n\widehat{\phi}_n)^2 \frac{1}{n} \sum_{t=1}^{n} y_{t-1}^2 I(|y_{t-1}| \le a) \\
& \quad + 2\phi_0 \frac{1}{n} \sum_{t=1}^{n} (s_t - p_0) \varepsilon_t |y_{t-1}| I(|y_{t-1}| \le a) \qquad \text{(S2.6)} \\
& \quad + 2\phi_0 (p_0\phi_0 - \widehat{p}_n\widehat{\phi}_n) \frac{1}{n} \sum_{t=1}^{n} (s_t - p_0) y_{t-1}^2 I(|y_{t-1}| \le a) \\
& \quad + 2(p_0\phi_0 - \widehat{p}_n\widehat{\phi}_n) \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t |y_{t-1}| I(|y_{t-1}| \le a) \\
& \overset{a.s.}{\to} \; p_0(1 - p_0)\phi_0^2 \mathbb{E}\{y_{t-1}^2 I(|y_{t-1}| \le a)\} + \sigma_0^2 \mathbb{P}(|y_{t-1}| \le a),
\end{aligned}
$$

where σ_η^2 is defined in (4.3). Further, using the preceding expression of $\bar{\eta}$, we have that $\bar{\eta} = O_p(1/\sqrt{n})$ by the martingale central limit theorem in

Brown (1971) and Theorems 2-3. Similarly, we can get

$$\begin{aligned}
&\frac{1}{n}\sum_{t=k+1}^{n}(\widehat{\eta}_t-\bar{\eta})(\widehat{\eta}_{t-k}-\bar{\eta})-\frac{1}{n}\sum_{t=k+1}^{n}\widehat{\eta}_t\widehat{\eta}_{t-k}\\
&\qquad=\frac{n-k}{n}\bar{\eta}^2-\bar{\eta}\frac{1}{n}\sum_{t=k+1}^{n}\widehat{\eta}_t-\bar{\eta}\frac{1}{n}\sum_{t=k+1}^{n}\widehat{\eta}_{t-k}=O_p(1/n)
\end{aligned}$$

for each fixed $k\geq 0$. Using above facts, we have that

$$\sqrt{n}\widehat{\rho}_{nk}=(1+o_p(1))\left\{\frac{1}{\sigma_\eta^2\sqrt{n}}\sum_{t=k+1}^{n}\widehat{\eta}_t\widehat{\eta}_{t-k}\right\}+o_p(1). \tag{S2.7}$$

Next, it suffices to consider the joint limiting distribution of

$$(\sigma_\eta^2\sqrt{n})^{-1}\sum_{t=k+1}^{n}\widehat{\eta}_t\widehat{\eta}_{t-k},\quad k=1,...,M.$$

To this end, let $\eta_t(\theta)=(y_t-p\phi|y_{t-1}|)I(|y_{t-1}|\leq a)$, then $\widehat{\eta}_t=\eta_t(\widehat{\theta}_n)$ and $\eta_t=\eta_t(\theta_0)$ in (4.2). Denote

$$\rho_{nk}(\theta)=\frac{1}{n\sigma_\eta^2}\sum_{t=k+1}^{n}\eta_t(\theta)\,\eta_{t-k}(\theta),\quad \theta\in\Theta,\quad k\geq 1.$$

Note that $\partial\eta_t(\theta)/\partial\theta=-\vartheta|y_{t-1}|I(|y_{t-1}|\leq a)$, where $\vartheta=(p,\phi,0)'$, and

$$\begin{aligned}
\frac{\partial\rho_{nk}(\theta_0)}{\partial\theta'}&=\frac{1}{n\sigma_\eta^2}\sum_{t=k+1}^{n}\left\{\frac{\partial\eta_t(\theta_0)}{\partial\theta'}\eta_{t-k}+\eta_t\frac{\partial\eta_{t-k}(\theta_0)}{\partial\theta'}\right\}\\
&=-\vartheta_0'\frac{1}{n\sigma_\eta^2}\sum_{t=k+1}^{n}\{\eta_{t-k}|y_{t-1}|I(|y_{t-1}|\leq a)+\eta_t|y_{t-k-1}|I(|y_{t-k-1}|\leq a)\}\\
&=\frac{u_k}{\sigma_\eta^2}\vartheta_0'+o_p(1)
\end{aligned}$$

with $u_k=-\mathbb{E}\{\eta_{t-k}|y_{t-1}|I(|y_{t-1}|\leq a)$, by the law of large numbers and $\mathbb{E}(\eta_t)=0$. Then, by the Taylor expansion, the law of large numbers, and

Theorems 2-3, it follows that

$$\begin{aligned}\sqrt{n}\big(\rho_{nk}(\widehat{\theta}_n) - \rho_{nk}(\theta_0)\big) &= \frac{\partial \rho_{nk}(\theta_0)}{\partial \theta'}\sqrt{n}(\widehat{\theta}_n - \theta_0) + o_p(1) \\ &= \frac{u_k}{\sigma_\eta^2}\vartheta_0'\sqrt{n}(\widehat{\theta}_n - \theta_0) + o_p(1).\end{aligned}$$

Let $\boldsymbol{\rho}_n(\theta) = (\rho_{n1}(\theta), ..., \rho_{nM}(\theta))'$. It follows that

$$\sqrt{n}\boldsymbol{\rho}_n(\widehat{\theta}_n) = \sqrt{n}\boldsymbol{\rho}_n(\theta_0) + \frac{1}{\sigma_\eta^2}(u_1, \ldots, u_M)'\vartheta_0'\sqrt{n}(\widehat{\theta}_n - \theta_0) + o_p(1).$$

By (S2.5), we have

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = -\mathcal{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_t(\theta_0)}{\partial \theta} + o_p(1).$$

The martingale central limit theorem in Brown (1971) gives that

$$(\sqrt{n}\boldsymbol{\rho}_n(\theta_0), \sqrt{n}(\widehat{\theta}_n - \theta_0))' \xrightarrow{d} \mathcal{N}(0, \mathbf{G}).$$

Thus, $\sqrt{n}\boldsymbol{\rho}_n(\widehat{\theta}_n) \xrightarrow{d} \mathcal{N}(0, \mathbf{UGU}')$ by a matrix linear transformation.

Finally, note that, by (S2.6)-(S2.7),

$$\sqrt{n}\big(\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n(\widehat{\theta}_n)\big) = o_p(1)\sqrt{n}\boldsymbol{\rho}_n(\widehat{\theta}_n) + o_p(1) = o_p(1)O_p(1) + o_p(1) = o_p(1).$$

Thus, $\sqrt{n}\widehat{\boldsymbol{\rho}}_n \xrightarrow{d} \mathcal{N}(0, \mathbf{UGU}')$. The proof is complete. □

S2.5 Proof of Proposition 1

For any time point t, a k-th cumulative bubble collapses if $s_t = 0$, $s_{t-l} = 1$ for $1 \leq l \leq k$ and $s_{t-k-1} = 0$. Let $\{z_s^\diamond\}$ be a new auxiliary process that satisfies the recursion

$$z_s^\diamond = \begin{cases} \varepsilon_s, & \text{if } s \le t-k-1, \\ \phi_0|z_{s-1}^\diamond| + \varepsilon_s, & \text{if } s > t-k-1, \end{cases}$$

then $y_{t-1} = z_{t-1}^\diamond$ if a k-th cumulative bubble is formed at time $t-1$ to be collapsed at time t. Note that the process $\{z_s^\diamond\}$ is constructed using the innovation sequence $\{\varepsilon_t\}$, which is independent of the sequence $\{s_t\}$, we can show that the joint probability

$$\begin{aligned} &\mathbb{P}(r_t \le c_r, s_t = 0, s_{t-1} = \cdots = s_{t-k} = 1, s_{t-k-1} = 0) \\ &= \mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r, s_t = 0, s_{t-1} = \cdots = s_{t-k} = 1, s_{t-k-1} = 0) \\ &= p_0^k(1-p_0)^2\mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r). \end{aligned}$$

On the other hand, the marginal probability that a k-th cumulative bubble collapses at time t equals to $\mathbb{P}(s_t = 0, s_{t-1} = \cdots = s_{t-k} = 1, s_{t-k-1} = 0) = p_0^k(1-p_0)^2$, and thus it suffices to show that

$$\mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r) = \mathbb{P}(z_k \ge -c_r).$$

For this, note that the two vectors $(\varepsilon_{t-k}, \ldots, \varepsilon_t)$ and $(\varepsilon_0, \ldots, \varepsilon_k)$ share the same distribution, and thus by definition the two vectors $(z_{t-1}^\diamond, \ldots, z_{t-k-1}^\diamond)$ and $(z_k, \ldots, z_0)$ have the same joint distribution. By independence of ε_t and $z_{t-1}^\diamond$ we can then conclude that

$$\mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r) = \mathbb{P}(\varepsilon_k - \phi_0|z_{k-1}| \le c_r) = \mathbb{P}(\phi_0|z_{k-1}| - \varepsilon_k \ge -c_r).$$

If the innovation sequence $\{\varepsilon_t\}$ has a symmetric distribution, then $z_k = \phi_0|z_{k-1}| + \varepsilon_k$ has the same distribution as $\phi_0|z_{k-1}| - \varepsilon_k$, and the result follows. □

S2.6 Proof of Proposition 2

For any time point t, it constitutes a k-th cumulative bubble if $s_{t-l} = 1$ for $0 \le l \le k-1$ and $s_{t-k} = 0$. Let $\{z_s^\circ\}$ be a new auxiliary process that satisfies the recursion

$$z_s^\circ = \begin{cases} \varepsilon_s, & \text{if } s \le t-k, \\ \phi_0|z_{s-1}^\circ| + \varepsilon_s, & \text{if } s > t-k, \end{cases}$$

then by the independence of $\{\varepsilon_t\}$ and $\{s_t\}$ we can show that the joint probability

$$\begin{aligned} &\mathbb{P}(r_t \le c_r, s_t = 0, s_{t-1} = \cdots = s_{t-k} = 1, s_{t-k-1} = 0) \\ &\quad = \mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r, s_t = s_{t-k-1} = 0, s_{t-1} = \cdots = s_{t-k} = 1) \\ &\quad = p_0^k(1-p_0)^2\mathbb{P}(\varepsilon_t - \phi_0|z_{t-1}^\diamond| \le c_r). \end{aligned}$$

On the other hand, the marginal probability that time t is a k-th cumulative bubble equals to $\mathbb{P}(s_t = \cdots = s_{t-k+1} = 1, s_{t-k} = 0) = p_0^k(1-p_0)$, and thus it suffices to show that $\{z_s^\circ\}_{t-k<s\le t}$ and $\{z_{s'}\}_{1\le s'\le k}$ share the same distribution. For this, note that the two vectors $(\varepsilon_{t-k}, \ldots, \varepsilon_t)$ and $(\varepsilon_0, \ldots, \varepsilon_k)$ share the same distribution, and they drive z_s°, $t-k < s \le t$, and $z_{s'}$, $1 \le s' \le k$, based on the same recursion, the result then follows. □

S3 Explicit expressions of $\mathcal{I}$ and $\mathcal{J}$ in Theorem 3

As for the explicit expressions of $\mathcal{I}$ and $\mathcal{J}$, by (S2.1) and (S2.3) and the following facts

$$\mathbb{E}\{(y_t - p_0\phi_0|y_{t-1}|)^3|\mathcal{F}_{t-1}\} = p_0(1-p_0)(1-2p_0)\phi_0^3|y_{t-1}|^3 + \kappa_3,$$

$$\mathbb{E}\{(y_t - p_0\phi_0|y_{t-1}|)^4|\mathcal{F}_{t-1}\} = p_0(1-p_0)(1-3p_0+3p_0^2)\phi_0^4 y_{t-1}^4 + 6\sigma_0^2 p_0(1-p_0)\phi_0^2 y_{t-1}^2 + \kappa_4,$$

a tedious algebraic calculation can yield

$$\begin{aligned}\mathcal{J} &= \mathbb{E}\left\{\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta'}\right\} \\ &= \mathbb{E}\left\{\frac{1}{[p_0(1-p_0)\phi_0^2 y_t^2 + \sigma_0^2]^2}\mathbf{A}_t\right\} + \mathbb{E}\left\{\frac{2y_t^2}{p_0(1-p_0)\phi_0^2 y_t^2 + \sigma_0^2}\right\}\mathbf{D},\end{aligned}$$

$$\begin{aligned}\mathcal{I} &= \mathbb{E}\left\{\frac{\partial \ell_t(\theta)}{\partial\theta}\frac{\partial \ell_t(\theta)}{\partial\theta'}\right\} \\ &= \mathbb{E}\left\{\frac{p_0(1-p_0)(1-2p_0)^2\phi_0^4 y_t^4 + 4\sigma_0^2 p_0(1-p_0)\phi_0^2 y_t^2 + (\kappa_4 - \sigma_0^4)}{[p_0(1-p_0)\phi_0^2 y_t^2 + \sigma_0^2]^4}\mathbf{A}_t\right\} \\ &\quad + \mathbb{E}\left\{\frac{2p_0(1-p_0)(1-2p_0)\phi_0^3 y_t^4 + 2|y_t|\kappa_3}{[p_0(1-p_0)\phi_0^2 y_t^2 + \sigma_0^2]^3}\mathbf{B}_t\right\} \\ &\quad + \mathbb{E}\left\{\frac{4y_t^2}{p_0(1-p_0)\phi_0^2 y_t^2 + \sigma_0^2}\right\}\mathbf{D}\end{aligned}$$

with $\kappa_3 = \mathbb{E}(\varepsilon_t^3)$, $\kappa_4 = \mathbb{E}(\varepsilon_t^4)$, and

$$\mathbf{A}_t = \begin{pmatrix} 4p_0^2(1-p_0)^2\phi_0^2 y_t^4 & 2p_0(1-p_0)(1-2p_0)\phi_0^3 y_t^4 & 2p_0(1-p_0)\phi_0 y_t^2 \\ & (1-2p_0)^2\phi_0^4 y_t^4 & (1-2p_0)\phi_0^2 y_t^2 \\ & & 1 \end{pmatrix},$$

$$\mathbf{B}_t = \begin{pmatrix} 4p_0^2(1-p_0)\phi_0 y_t^2 & p_0(3-4p_0)\phi_0^2 y_t^2 & p_0 \\ & 2(1-2p_0)\phi_0^3 y_t^2 & \phi_0 \\ & & 0 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} p_0^2 & p_0\phi_0 & 0 \\ & \phi_0^2 & 0 \\ & & 0 \end{pmatrix}.$$

Here, the elements in the lower triangles can be completed by symmetry.

References

Brown, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist. 42*(1), 59–66.

Meyn, S. and R. L. Tweedie (2009). *Markov Chains and Stochastic Stability* (2nd ed.). Cambridge University Press.

College of Mathematics and Physics, Beijing University of Chemical Technology, Beijing, China

E-mail: yangxl@buct.edu.cn

Department of Statistics and Data Science, Tsinghua University, Beijing, China

E-mail: malidong@tsinghua.edu.cn

Department of Statistics, University of Georgia, Athens, Georgia, USA.

E-mail: tingzhang@uga.edu