Supplementary material to: S. Clinet and Y. Potiron, "Disentangling sources of high frequency market microstructure noise"

We give the assumptions related to Proposition 1 and a detailed proof of the consistency of the BIC. First, defining

\[ \chi(\theta) := \mathbb{E} \left[ (\Delta \phi(Q_{t_1}, \theta) - \Delta \phi(Q_{t_1}, \theta_0))^2 \right], \]

we assume

[A] For any \( m \in \mathcal{M} \), \( \chi \) admits a unique minimum \( \tilde{\theta}^{(m)} \) on the interior of \( m \).

Note that [A] is automatically satisfied for linear models such as (2.1) as soon as the variance-covariance matrix of the vector of returns of information \( \Delta Q_{t_1} \) is positive definite.

We also define

\[ W_i(\theta) := \phi(Q_i, \theta) - \phi(Q_i, \theta_0), \]

and for any \( i, j, k, l \in \mathbb{N} \), and for any multi-indices \( q = (q_1, q_2) \), \( r = (r_1, r_2, r_3, r_4) \), where the subcomponents of \( q \) and \( r \) are \( d \) dimensional multi-indices, the following quantities conditioned on the price process

\[ \mathbb{E}[W_i(\theta) \mid X] = 0 \ a.s., \]

\[ \rho^q_j(\theta) := \mathbb{E} \left[ \frac{\partial^{q_1} W_i(\theta)}{\partial \theta^{q_1}} \frac{\partial^{q_2} W_{i+j}(\theta)}{\partial \theta^{q_2}} \mid X \right] = \mathbb{E} \left[ \frac{\partial^{q_1} W_i(\theta)}{\partial \theta^{q_1}} \frac{\partial^{q_2} W_{i+j}(\theta)}{\partial \theta^{q_2}} \mid X \right] \ a.s., \]

\[ \kappa^r_{j,k,l}(\theta) := \text{cum} \left( \frac{\partial^{r_1} W_i(\theta)}{\partial \theta^{r_1}}, \frac{\partial^{r_2} W_{i+j}(\theta)}{\partial \theta^{r_2}}, \frac{\partial^{r_3} W_{i+k}(\theta)}{\partial \theta^{r_3}}, \frac{\partial^{r_4} W_{i+l}(\theta)}{\partial \theta^{r_4}} \right) \mid X \]

\[ = \text{cum} \left( \frac{\partial^{r_1} W_i(\theta)}{\partial \theta^{r_1}}, \frac{\partial^{r_2} W_{i+j}(\theta)}{\partial \theta^{r_2}}, \frac{\partial^{r_3} W_{i+k}(\theta)}{\partial \theta^{r_3}}, \frac{\partial^{r_4} W_{i+l}(\theta)}{\partial \theta^{r_4}} \right) \ a.s., \]

where \( \rho^q_j(\theta) \) and \( \kappa^r_{j,k,l}(\theta) \) are assumed independent of \( n \). The following assumption is directly taken from [Clinet and Potiron, 2019]:

[B] The impact function \( \phi \) is supposed to be of class \( C^m \) in \( \theta \) with \( m > \bar{d}/2 + 2 \). Moreover, for any \( i = 0, \ldots, m \) and \( 0 \leq |q|, |r| \leq m \), we have

\[ \sup_{\theta \in \Theta} \sum_{j=0}^{+\infty} \rho^q_j(\theta) < \infty \ a.s., \]

\[ \sup_{\theta \in \Theta} \sum_{j,k,l=0}^{+\infty} \kappa^r_{j,k,l}(\theta) < \infty \ a.s., \]

\[ \mathbb{E} \left[ \sup_{\theta \in \Theta} \left| \frac{\partial^j \Delta \phi(Q_{t_1}, \theta)}{\partial \theta^j} \right| ^p \mid X \right] < \infty \ a.s., \text{for any } p \geq 1, \ 0 \leq j \leq 2, \]

\[ \frac{\partial \rho_0(\theta)}{\partial \theta} = 0 \iff \theta = \theta_0. \]

We are now ready to prove Proposition 1.
Step 1. We prove our claim when $m_0$ is a submodel of $m$, and so $d > d_0$ where $d$ is the number of parameters of $m$. By Theorem 3.1 from [Clinet and Potiron, 2019], and up to some reordering of the subcomponents of $\theta$, the estimator $\hat{\theta}(m)$ is consistent and asymptotically normal, toward the limit $v_0 = (\hat{\sigma}^2_0, \hat{\theta}_0^0, \cdots, \hat{\theta}_0^{d_0}, 0, \cdots, 0)$ where $\hat{\sigma}^2_0 = \int_0^T \sigma_s^2 ds + \sum_{0 \leq s \leq T} \Delta J_s^2$. We slightly reformulate the problem as follows: introducing $\tilde{w}(m) = ((\hat{\sigma}^2)(m), N^{1/2}(\hat{\theta}(m) - \theta_0))$, and $w_0 = (\hat{\sigma}^2_0, 0, \cdots, 0)$, we have by a Taylor expansion, for some $\overline{w} \in [\tilde{w}(m), w_0]$$$
abla \theta \left| \begin{array}{c} 2(l_{\text{exp}}(\tilde{w}(m)) - l_{\text{exp}}(v_0)) = 2(L_{\text{exp}}(\tilde{w}(m)) - L_{\text{exp}}(w_0)) \\
 = -N(\tilde{w}(m) - w_0)\nabla^2 L_{\text{exp}}(\overline{w})(\tilde{w}(m) - w_0) \rightarrow d \chi^2(d) \end{array} \right.$$ \text{by application of Theorem 3.1 and Lemma C.15 from [Clinet and Potiron, 2019], and with $L_{\text{exp}}^{(d)}$ being the restriction of $L_{\text{exp}}$ on $m$, where $L_{\text{exp}}$ is defined in (C.88), p.323 of [Clinet and Potiron, 2019], and $H_{\text{exp}}(m) = -N\hat{\theta}_2 L_{\text{exp}}(m) / \partial w^2$.}$

In the previous equation, $\rightarrow d \chi^2(d)$ stands for the convergence in law toward a chi-squared distribution with $d$ degrees of freedom. We have a similar result for $m_0$, and thus $L_{\text{exp}}(\tilde{w}(m)) - L_{\text{exp}}(\tilde{w}(m)) = O_F(1)$. This, in turn, implies that $\text{BIC}(m) - \text{BIC}(m_0) \sim (d-d_0)\log(N) \rightarrow^F + \infty$. 

Step 2. We prove our claim when $m_0$ is not a submodel of $m$. We recall that, by definition of the likelihood process, we have

$$\hat{\theta}(m) \in \arg\min_{\theta \in m} \Delta \tilde{Z}(\theta)^T \Delta \tilde{Z}(\theta),$$

and

$$(\hat{\sigma}^2)(m) = T^{-1} \Delta \tilde{Z}(\hat{\theta}(m))^T \Delta \tilde{Z}(\hat{\theta}(m)),$$

with $\tilde{Z}_t(\theta) = Z_t - \phi(Q_t, \theta)$. By direct calculation similar to that of Section C.4 from [Clinet and Potiron, 2019], we have the uniform convergence for $\theta \in m$

$$N^{-1} \Delta \tilde{Z}(\theta)^T \Delta \tilde{Z}(\theta) \rightarrow^F \chi(\theta) = \mathbb{E} \left[ (\Delta \phi(Q_t, \theta) - \Delta \phi(Q_t, \theta_0))^2 \right].$$

As a direct consequence, we obtain that $\hat{\theta}(m) \rightarrow^F \hat{\theta}(m)$ where we recall that $\hat{\theta}(m)$ is the unique minimum of $\chi$ on the interior of $m$ by Assumption [A]. Similarly, we easily obtain that

$$(\hat{\sigma}^2)(m) = \Delta_N^{-1} \chi(\hat{\theta}(m)) + o_F(\Delta_N^{-1}),$$

where $\Delta_N = T/N$, and where $\chi(\hat{\theta}(m)) > 0$ by the identifiability assumption (2.14) from [Clinet and Potiron, 2019] along with the fact that $\hat{\theta}(m) \neq \theta_0$. Moreover, we have

$$\frac{\partial^2 l_{\text{exp}}(\tilde{w}(m))}{\partial v^2} = \begin{pmatrix} \frac{-T}{\hat{\sigma}^4(m)} & 0 \\
0 & \frac{1}{\hat{\sigma}^2(m)} \frac{\partial^2 (\Delta \tilde{Z}(\tilde{w}(m))^T \Delta \tilde{Z}(\hat{\theta}(m)))}{\partial \theta^2} \end{pmatrix}.$$
and therefore by a Taylor expansion at $\tilde{\nu}^{(m)}$, we get for some $\nu \in [v_0, \tilde{\nu}^{(m)}]$

$$2(l^{(m)}_{\exp}(\tilde{\nu}^{(m)}) - l^{(m)}_{\exp}(v_0)) = (\tilde{\nu}^{(d)} - v_0)^T \frac{\partial^2 l^{(m)}_{\exp}(\nu)}{\partial \nu^2} (\tilde{\nu}^{(m)} - v_0)$$

$$= \frac{-T((\tilde{\sigma}^2)^{(m)} - \sigma_0^2)}{2(\tilde{\sigma}^4)^{(m)}\Delta_N} - \frac{(\hat{\theta}^{(m)} - \theta_0)^T}{2(\tilde{\sigma}^2)^{(m)}\Delta_N} \frac{\partial^2 \chi(\tilde{\theta}^{(m)})}{\partial \theta^2} (\tilde{\theta}^{(m)} - \theta_0)$$

Now, since $\hat{\theta}^{(m)}$ is the unique minimum of $\chi$ on the interior of $m$, we deduce that $\frac{\partial^2 \chi(\tilde{\theta}^{(m)})}{\partial \theta^2}$ is a positive matrix and thus $-T\Delta_N^{-1} \frac{(\tilde{\theta}^{(m)} - \theta_0)^T}{2\chi(\tilde{\theta}^{(m)})} \frac{\partial^2 \chi(\tilde{\theta}^{(m)})}{\partial \theta^2} (\tilde{\theta}^{(m)} - \theta_0) \leq 0$. Therefore, we have $2(l^{(m)}_{\exp}(\tilde{\nu}^{(m)}) - l^{(m)}_{\exp}(v_0)) \leq -\frac{T}{2}\Delta_N^{-1} + o_{\mathbb{P}}(\Delta_N^{-1})$. Thus,

$$\text{BIC}(m) - \text{BIC}(m_0) \geq \frac{N}{2} + \frac{(d - d_0)\log(N)}{o_{\mathbb{P}}(N)} + o_{\mathbb{P}}(N),$$

which proves our claim.

References