A-Posteriori Identifiability of the Maxwell Slip Model of Hysteresis

Demosthenes D. Rizos \* Spilios D. Fassois ++

\* Swiss Federal Laboratory of Material Testing and Research (EMPA)
Ueberlandstrasse 129, CH-8600, Duebendorf, Switzerland
(Tel: +41-44-823-4989; e-mail: dimosthenis.rizos@empa.ch;
demosthenis.rizos@gmail.com).
++ Department of Mechanical Engineering and Aeronautics
University of Patras GR 265 00, Greece
(Tel/Fax: +30-2610-969-495; email: fassois@mech.upatras.gr).

Abstract: The a–posteriori identifiability of the Maxwell Slip model of hysteresis is addressed. The necessary and sufficient conditions that guarantee that the available data are informative enough are provided. An Output Error type estimator is subsequently postulated and its consistency is established. It is specifically shown that the estimates converge in probability to their actual counterparts under easily verifiable conditions on the Maxwell Slip model structure, the excitation, and mild assumptions on the additive measurement noise. The analytical results are verified through Monte Carlo simulations.

1. INTRODUCTION

This study focuses on the identification of systems with hysteresis based upon the Maxwell Slip model. This particular model is also referred to as Generalized Maxwell Resistive Capacitor model [Lee et al., 2000] or Prandtl–Ishlinskii model of Stop-Type [Visintin, 1991, p.83], and belongs to the general class of rheological models of hysteresis [Visintin, 1991, pp.40-41].

The general problem tackled may be posed via the following two questions:

- Is it worthwhile employing the Maxwell Slip model for the identification of actual systems with hysteresis?
- How can a Maxwell Slip model be identified based on a single pair of excitation - response signals?

The first question may be answered by considering the model advantages, that is its simplicity and physical interpretation, as well as its capability of reproducing rate - independent hysteresis [Oh et al., 2005] with nonlocal memory [Lampaert et al., 2002]. Due to these, the Maxwell Slip model has been employed in several applications, with friction (see Lampaert and Swevers [2001], Lampaert et al. [2002], Rizos and Fassois [2003], Al-Bender et al. [2003], Parltiz et al. [2004], Rizos and Fassois [2004, 2005a,b], Worden et al. [2007], Rizos and Fassois [2009], Moerlooze et al. [2010]) and piezoelectric actuator modelling (see Goldfarb and Celanovic [1997], Kuhnen [1999], Choi et al. [2002], Georgiou and Mrad [2006] for a representative collection) being probably the most frequent. The common conclusion from all these applications is that the Maxwell Slip model appears capable of representing the measured hysteresis quite adequately.

This study mainly aims at providing answers for the second question. Toward this end it is primarily assumed that the data generating process is perfectly described by a certain Maxwell Slip model (actual model). Hence, the question is what are the conditions under which a proper estimator manages to discriminate the actual model from the pool of models belonging to the Maxwell Slip model structure. This essentially is the concept of “a-posteriori identifiability” – an issue that has not been yet addressed for Maxwell Slip models.

An Output Error estimator (which belongs to the wide class of Prediction Error estimators) is also postulated in this study, and its properties are analyzed. First the conditions that guarantee that the expected global minimum of the estimation criterion is unique and corresponds to the actual model parameters are briefly reviewed. These are related to the concepts of a–priori identifiability and excitation richness.

Subsequently, the conditions for convergence of the sample global minimum of the estimation criterion to its expected value are derived. Following this, the main contribution of the current study which concerns estimator consistency is established.

2. THE MAXWELL SLIP MODEL

2.1 Mathematical Description

The Maxwell Slip model consists of M operators of the stop – type in parallel configuration (Fig. 1). Each operator has negligible inertia and is characterized by its own linear stiffness $k_i$ and maximum spring deformation $\Delta_i$ (threshold). The difference between the excitation (common to all operators) displacement $x(t)$ and the $i$-th operator’s current mass position $x_i(t)$ provides the corresponding current (unobservable) spring deformation $\delta_i$. The $i$-th operator remains sticking as long as $|\delta_i| < \Delta_i$, while slipping, that is $|\delta_i| = \Delta_i$, until the exerted displacement reaches a local extremum, that is the exerted velocity ($v = \frac{dx}{dt}$) goes through zero (changes sign).

The discrete–time $i$-th spring deformation $\delta_i(t), [\forall i = 1, \ldots, M]$ is defined as:

$$\delta_i(t) \triangleq x(t) - x_i(t) \quad (1)$$

while its evolution is described as [Rizos and Fassois, 2004]:

---

* This work was supported by the VolkswagenStiftung (Grant no I/76938).
Fig. 1. The Maxwell Slip model structure.

The model response (total force), \( f(t, \theta, \delta_M(1)) \), is obtained by summing up the spring forces of all operators:

\[
f(t, \theta, \delta_M(1)) = \sum_{i=1}^{M} k_i \cdot \delta_i(t)
\]

with \( t = 1, 2, \ldots \) referring to normalized discrete time. \( \theta_M \triangleq [k_M', d_M']^T \in \Theta_M \) (parameter space) stands for the model parameter vector:

\[
k_M \triangleq [k_1 \ldots k_M]^T, \quad d_M \triangleq [\Delta_1 \ldots \Delta_M]^T
\]

while (2)-(3) subject to initial (unmeasured and unobserved) spring deformations \( \delta_M(1) \):

\[
\delta_M(1) \triangleq [\delta_1(1) \ldots \delta_M(1)]^T
\]

with \( \delta_i(1) \) being the \( i \)-th operator initial spring deformation.

Careful consideration of (2) reveals that each operator has three distinct operational regimes:

1. **Positive slip** \([\text{slip}(+)]\) \( \iff \delta_i(t) = \Delta_i \)
2. **Stick** \( \iff \delta_i(t) < \Delta_i \)
3. **Negative slip** \([\text{slip}(-)]\) \( \iff \delta_i(t) = -\Delta_i \)

2.2 Basic Model Structure Assumptions

The basic assumptions underpinning the Maxwell Slip model structure are as follows:

**Assumption A1.** Each spring stiffness \( k_i \) is a nonzero finite constant, that is \( k_i \neq 0 \) and \( |k_i| < \infty \), \( \forall i \in 1, \ldots, M \). □

**Assumption A2.** The thresholds, \( \Delta_i (\forall i = 1, \ldots, M) \), are distinct and strictly positive constants, that is (without loss of generality) \( 0 < \Delta_1 < \ldots < \Delta_M < \infty \). □

**Assumption A3.** The parameter space \( \Theta_M \subset \mathbb{R}^{2M} \) is compact (bounded and closed [Bernstein, 2005, p.356]). □

3. MAXWELL SLIP MODEL ESTIMATION

3.1 The Actual Model

Prior to any analysis, the data generating process (actual model) needs to be defined. Thus, it is presently assumed that the data generating process is provided by the following definition:

**Definition 1.** (Actual model). The observation sequence

\[
\{y(t, \theta^{o}_{MP}, \delta^{o}_{MP}(1)) : t = 1, \ldots, N\}
\]

is generated by a Maxwell Slip model with \( M^o \) operators and initial state vector \( \delta^{o}_{MP}(1) \). It is parameterized by \( \theta^{o}_{MP} \in \Theta_{MP} \) and is subject to assumptions A1 - A3. The actual model is excited by a displacement sequence \( \{x(t) : t = 1, \ldots, N\} \) and provides a response sequence \( f(t, \theta^{o}_{MP}, \delta^{o}_{MP}(1)) \) which is corrupted by measurement stochastic noise \(\epsilon(t)\). Thus, \( \forall t = 1, 2, \ldots, N, \)

\[
y(t, \theta^{o}_{MP}, \delta^{o}_{MP}(1)) = f(t, \theta^{o}_{MP}, \delta^{o}_{MP}(1)) + \epsilon(t)
\]

with \( E[\epsilon(t)] = 0 \) and \( E[\epsilon^2(t)] < \infty \), \( \forall t \}, E[\cdot] \) and \( N \) stand for the statistical expectation and the number of observations, respectively.

3.2 The Estimator

An Output Error type estimator is postulated. The output error for an arbitrary Maxwell Slip model with \( M^o \) operators is given as:

\[
e(t, \theta_{MP}, \delta_{MP}(1)) \triangleq y(t, \theta_{MP}, \delta_{MP}(1)) - f(t, \theta_{MP}, \delta_{MP}(1))
\]

and the estimator is defined as the minimizer of the quadratic cost functional:

\[
Q_N(\theta_{MP}, \delta_{MP}(1)) \triangleq \frac{1}{N} \sum_{t=1}^{N} e^2(t, \theta_{MP}, \delta_{MP}(1))
\]

with respect to the unknown parameters \( \theta_{MP} \) and \( \delta_{MP}(1) \). Thus:

\[
\begin{bmatrix}
\theta_{MP}^T \\
\delta_{MP}^T
\end{bmatrix}_{N, \delta_{MP}(1)} \triangleq \arg \min_{\theta_{MP}, \delta_{MP}(1)} Q_N(\theta_{MP}, \delta_{MP}(1))
\]

For more information regarding the numerical implementation of the estimator, the interested reader is referred to [Rizos and Fassois, 2004].

3.3 Estimator Basic Properties

It is at this point worthwhile to present some of the main properties of the above estimator. By applying the statistical expectation operator to (10) and considering Definition 1, one obtains:

\[
E_Q(\theta_{MP}, \delta_{MP}(1)) = \frac{1}{N} \sum_{t=1}^{N} E[\epsilon^2(t)] +
\]

\[
\frac{1}{N} \sum_{t=1}^{N} \left(f(t, \theta_{MP}, \delta_{MP}(1)) - f(t, \theta_{MP}, \delta_{MP}(1))\right)^2 < \infty
\]

As it is clear the global minimum of \( E_Q(\theta_{MP}, \delta_{MP}(1)) \), with respect to \( \theta_{MP} \) and \( \delta_{MP}(1) \) is achieved iff \( f(t, \theta_{MP}, \delta_{MP}(1)) = f(t, \theta_{MP}, \delta_{MP}(1)) \), \( \forall t \).

The above equation plays a key role for the subsequent a-posteriori identifiability analysis. The necessary and sufficient conditions for fulfilling the following equivalence \( \forall \theta_{MP} \in \Theta_{MP} \) need to be provided:

\[
f(t, \theta_{MP}, \delta_{MP}(1)) = f(t, \theta_{MP}, \delta_{MP}(1)) \iff
\]

\[
\begin{cases}
\theta_{MP} \equiv \theta_{MP} \\
\delta_{MP}(1) \equiv \delta_{MP}(1)
\end{cases}
\]

thus proving that the global minimum of \( E_Q(\theta_{MP}, \delta_{MP}(1)) \) is unique and can be only achieved by the actual model.

1. The superscript "o" designates the actual model.
4. A-POSTERIORI IDENTIFIABILITY

The a-posteriori identifiability, that is the conditions under which the estimator is capable of discriminating the actual system from all alternatives Maxwell Slip models with \(M^r\) operators based on provided excitation (displacement) – response (force) data, is established.

Towards this direction, it suffices to show that:

\[
[\theta_{M^r}; \delta_{M^r}]^T = \arg \min_{\theta_{M^r}, \delta_{M^r}} Q_N(\theta_{M^r}, \delta_{M^r}(1)), \forall N (13)
\]

so that minimization of the cost function yields the actual parameter vector. However, due to the stochastic nature of the estimates [see (8) – (9)], it is quite difficult to directly manipulate \(Q_N(\theta_{M^r}, \delta_{M^r}(1))\). Thus, an asymptotic \((N \to \infty)\) indirect approach is adopted (the interested reader is referred to Pötscher and Prucha [1997] for more details on this general procedure), which includes the following main steps:

**Convergence:** Find the conditions under which \(Q_N(\theta_{M^r}, \delta_{M^r}(1))\) converges in a proper statistical sense to \(E_Q N(\theta_{M^r}, \delta_{M^r}(1))\) (its expected value) uniformly \(^2\) over the parameter space as \(N \to \infty\).

**Identifiability:** Find the conditions under which the global minimizer of \(E_Q N(\theta_{M^r}, \delta_{M^r}(1))\) is unique and provides the actual model parameters [essentially proof of (12)].

### 4.1 A–Priori Identifiability

The first task to handle is to prove that there exists at least one excitation displacement sequence for which the actual model parameters constitute the unique global minimizers of \(E_Q N(\theta_{M^r}, \delta_{M^r}(1))\). This leads to the a–priori identifiability (or uniqueness of representation) question \(^3\) [Rizos and Fassois, 2004].

It has been recently shown that assumptions A1 - A2 are necessary and sufficient for achieving the Maxwell Slip model structure a–priori identifiability with respect to \(\theta_{M^r}\) (see Rizos [2008], Rizos and Fassois [2011]). Moreover it has been concluded that the model structure is not a–priori identifiable with respect to \(\delta_{M^r}(1)\) [Rizos and Fassois, 2004]. Thus any attempt to estimate \(\delta_{M^r}(1)\) fails, and therefore the transient response, due to the \(\theta_{M^r}\), should be properly removed. For this reason the first excitation condition, designated as C0, has been introduced (also see Rizos and Fassois [2004]).

**Definition 2.** (Condition C0). The excitation \(x(t)\) needs to be designed such that it gets increased [resp: decreased], in an arbitrary way, up to a critical max [resp: min] value \(x(t_{cr})\), in order for all operators to be forced to slip[+] [resp: slip[–]] at \(t = t_{cr}\).

If condition C0 is properly applied, the available for estimation data set starts from the time \(t_{cr} \geq 1\), and the estimation is implemented only with respect to \(\theta_{M^r}\) [Rizos and Fassois, 2004]. Hence the estimator is postulated as:

\[
\hat{\theta}_{M^r,N} \triangleq \arg \min_{\theta_{M^r}} Q_N(\theta_{M^r}) (15)
\]

while (12) is now written as:

\[
f_t(\hat{\theta}_{M^r,N}) = f_t(\theta_{M^r}), \forall t \geq t_{cr} \iff \theta_{M^r} = \theta_{M^r} (16)
\]

### 4.2 Excitation Richness

The a-priori identifiability is only necessary for achieving a-posteriori identifiability. In general, an a-priori identifiable actual model (both assumptions A1 – A2 are valid), should be properly excited in order to reveal its unique character. Thus the necessary and sufficient excitation conditions, which a displacement sequence of general form should fulfill for validating (16), need to be derived.

These necessary and sufficient excitation conditions are qualitatively described as follows (see Rizos [2008] and Rizos and Fassois [2011] for the corresponding proof and a detailed relevant study).

Given that the condition C0 is satisfied, two conditions should be additionally met by the exerted displacement:

**C1:** All operators of the actual model should come at least once into their stick regime after time \(t_{cr}\).

**C2:** \(\forall t = 1, \ldots , M^r - 1\), there should be three successive time instants, say \(t_i, t_i+1\) and \(t_i+2\), after time \(t_{cr}\), such that the \(i\)-th operator slips[+] [or slips[–]], while the \((i+1)\)-th sticks.

Based on the above, the following conclusion is reached: On condition that the transient response has been removed by employing condition C0, \(\theta_{M^r}\) (actual model parameter vector) is the unique global minimizer of \(E_Q N(\theta_{M^r})\) iff the actual model satisfies assumptions A1–A2 (is a–priori identifiable) and gets excited by a displacement sequence that fulfills conditions C1 – C2.

### 4.3 Consistency

Within the previous subsection the conditions under which \(\hat{\theta}_{M^r,N}\) is the unique global minimizer of the \(E_Q N(\theta_{M^r})\) were provided. The next step is to find the conditions for the following statistical uniform convergence (in this context see the general procedures in [Pötscher and Prucha, 1997, p.37]):

\[
\lim \left\{ \sup_{\theta_{M^r}} |Q_N(\theta_{M^r}) - E_Q N(\theta_{M^r})| \right\} = 0, \forall \theta_{M^r} \in \Theta_{M^r} (17)
\]

Such a convergence allows for [Ljung, 1997, p.237]:

\[
\lim \left\{ \hat{\theta}_{M^r,N} - \theta_{M^r} \right\} = 0 (18)
\]

which will prove the consistency of the postulated estimator. Note that lim is the probability limit operator [Mendel, 1995, p.96].

In order to proceed, some additional assumptions regarding the measurement stochastic noise \(\epsilon(t)\) should made.

**Assumption A4.** The measurement noise sequence \(\{\epsilon(t) : t = 1, \ldots , N\}\) is a zero mean, quasi-stationary signal [Ljung, 1997, p.27] such that:

\[
\epsilon(t) = \sum_{r=0}^{\infty} h_r(r) \cdot w(t-r) = H_r(\theta) \cdot w(t) (19)
\]

where \(w(t) : t = 1, \ldots , N\) is a sequence of independent random variables with \(Ew(t) = 0\) and \(Ew^2(t) = \sigma^2_w(t)\) and bounded.

---

\(^2\) This allows for the transfer of some properties of \(Q_N(\theta_{M^r}, \delta_{M^r}(1))\) to its limiting function [Ljung, 1997, p.237].

\(^3\) Prove that (12) is valid for at least one excitation displacement sequence.
forth moments. \( \{H_t(\mathcal{B}), t = 1, 2, \ldots \} \) is a uniformly stable family of filters [Ljung, 1997, p.21]. Note that \( \mathcal{B} \) stands for the backshift operator, that is \( \mathcal{B}^r \cdot x(t) \triangleq x(t - r) \).

The following proposition then deals with estimator consistency.

**Proposition 1.** (Estimator consistency). Let the actual model be provided by Definition 1 and be excited by a displacement sequence that satisfies conditions C0 – C2. Assume also that the measurement stochastic noise \( \varepsilon(t) \) satisfies assumption A4. Then the estimator provided by (15) is consistent, that is:

\[
\lim\left( \hat{\theta}_{M'}^N \right) = \theta_{M'}^0 \quad (20)
\]

Note that assumptions A1-A3 are implicitly included in Definition 1.

**Proof.** The proof is based on Lemma 3.1 of Pötscher and Prucha [Pötscher and Prucha, 1997, p.17], which claims that if the Uniform Law of large Numbers (ULLN) (17) exists (see Appendix A – Lemma 4) and the actual parameter vector is the identifiably unique minimizer of \( E_0(\hat{\theta}_{M'}) \) (see Appendix A – Lemma 2) then the postulated estimator is consistent.

4.4 Discussion

It is worthwhile discussing the assumptions and conditions employed within Proposition 1.

Due to the lack of identifiability for \( \delta_{M'}(1) \), the transient model response should be removed from the data set [Rizos and Fassios, 2004]. Condition C0 essentially serves this purpose.

Assumptions A1 – A2 and conditions C1 – C2 are necessary and sufficient for fulfilling (16), and therefore ensuring the estimator uniqueness.

Assumption A3 (parameter space compactness) is only sufficient for obtaining (17). If not valid, then the uniform convergence may be difficult to achieve [Pötscher and Prucha, 1997, p.26]. Nevertheless, this is a rather formal assumption, without reducing the practical significance of Proposition 1.

Although assumption A4 is rather mild, it may be possible to somewhat relax. Indeed, broader stochastic noise sequences may be considered under the condition that there is a proper Law of Large Numbers (LLN) for obtaining the required uniform convergence of the sample cost function to its expected value (see [Pötscher and Prucha, 1997, p.45] for an excellent collection of limiting theorems). However, the requirement for bounded second order statistical moments cannot be relaxed, since the prediction error estimators’ asymptotic properties rely on the convergence of the sample second moments to their population counterparts.

5. NUMERICAL RESULTS

In this section a numerical verification of the a-posteriori identifiability of the Maxwell Slip model is provided. Towards this end 900 excitation - response realizations are generated by a Maxwell Slip model characterized by \( M' = 3 \) operators and parameters indicated in the left part of Table 1. For all realizations the excitation (displacement) is \( x(t) = \cos(2t) \) (sampling frequency \( f_s = 250 \text{Hz} \)) while the responses \( f(t, \hat{\theta}_3^0) \) are corrupted by random signals. These noise signals are generated by passing a sequence \( w(t) \) of independent zero mean random variables through a time–varying ARMA (TARMA) model [Poulomenos and Fassios, 2004]. The corresponding standard deviations of the model response and \( w(t) \) signals are, 0.90 and 0.036, respectively. A representative realization of the excitation (displacement) and the resulting response (force) with and without additive noise are depicted in Fig. 2.

Two different types of Monte Carlo estimation results are obtained. In the first case, each estimation is based upon 3000 sample (12 s) long excitation–noise corrupted response signals, while in the second case it is based upon 6000 sample (24 s) long signals. This is done in order to demonstrate that the estimates get improved (converge with increasing amounts of data).

The Monte Carlo results (900 estimation procedures for \( N = 3000 \) and \( N = 6000 \)) are provided in Table 1. As it is evident, the sample means of all estimated parameters are in excellent agreement with their actual counterparts, with the \( N = 6000 \) case estimates being slightly closer (apart from \( k_2 \)). It is important to observe that the corresponding standard deviations of the \( N = 6000 \) estimates are almost half of their \( N = 3000 \) case counterparts. This indicates that the estimates get improved with increasing signal record length, reflecting estimator consistency.

6. CONCLUDING REMARKS

The a–posteriori identifiability of the Maxwell Slip model has been proven under minimal and easily verifiable conditions on the Maxwell Slip model structure and excitation and mild assumptions on the additive measurement noise. The main conclusion is that a Maxwell Slip model may be consistently identified from a single experiment. Further research is currently devoted to estimator asymptotic normality and the development of model structure determination procedures.
7. ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of this study via a VolkswagenStiftung Grant (Grant no I/76938) to the University of Patras.

Appendix A. PROOF OF ESTIMATOR CONSISTENCY

Lemma 1. (Lipschitz Condition). Consider two Maxwell Slip models with $\mathcal{M}^o$ operators, parameterized by $\theta_{\mathcal{M}^o} = [k_1 \ldots k_{\mathcal{M}^o} \Delta_1 \ldots \Delta_{\mathcal{M}^o}]^T \in \Theta_{\mathcal{M}^o}$ and $\theta_{\mathcal{M}^o} = [k_1 \ldots k_{\mathcal{M}^o} \Delta_1 \ldots \Delta_{\mathcal{M}^o}]^T \in \Theta_{\mathcal{M}^o}$, with the parameter space $\Theta_{\mathcal{M}^o}$ being subject to assumption A3. Both models are excited by the same $x(t)$, such that condition $C0$ is validated at $t = t_C$, for both. Then the following Lipschitz condition is valid for the model responses:

$$|f(t, \theta_{\mathcal{M}^o}) - f(t, \theta_{\mathcal{M}^o})| \leq C \cdot d_\theta(\theta_{\mathcal{M}^o}, \theta_{\mathcal{M}^o}), \quad \forall t \geq t_C \quad (A1)$$

with $C < \infty$ being an arbitrary positive constant and $d_\theta(\cdot)$ the $L_1$-norm distance:

$$d_\theta(\theta_{\mathcal{M}^o}, \theta_{\mathcal{M}^o}) = \sum_{i=1}^{\mathcal{M}} \left( |k_i - \bar{k}_i| + |\Delta_i - \bar{\Delta}_i| \right) \quad (A2)$$

\[
\text{Proof.} \quad \text{See Rizos [2008] or Rizos and Fassos [2011].} \\
\]

Lemma 2. (Identifiability uniqueness of $\theta_{\mathcal{M}^o}^o$). Under condition $C0$, $\theta_{\mathcal{M}^o}^o$ is the identifiably unique [Pötscher and Prucha, 1997, p.16] minimizer of $EQ_N(\theta_{\mathcal{M}^o})$ iff the actual system is subject to assumptions A1-A2 and its excitation conditions $C1$-$C2$.

\[
\text{Proof.} \quad \text{This is straightforward as the actual model parameter vector, } \theta_{\mathcal{M}^o}^o, \text{ is the unique global minimizer of } EQ_N(\theta_{\mathcal{M}^o}) \text{ iff the actual system is subject to assumptions A1-A2 and excitation conditions } C1-C2 \text{ [see (16)]. Thus in this case the unique global minimizer of } EQ_N(\theta_{\mathcal{M}^o}) \text{ is independent of } N. \\
\]

Now consider a proper density function $d(\cdot)$ and $\lambda > 0$, as well as the set $\Theta_{\mathcal{M}^o} \triangleq \{ \theta_{\mathcal{M}^o} \in \Theta_{\mathcal{M}^o} : d(\theta_{\mathcal{M}^o}, \theta_{\mathcal{M}^o}^o) \geq \lambda \}$. Since $\theta_{\mathcal{M}^o}^o$ is the $N$-independent unique global minimizer of $EQ_N(\theta_{\mathcal{M}^o})$, then $\forall \lambda > 0$ and $\forall N$ sufficiently large:

$$\inf_{\Theta_{\mathcal{M}^o}} EQ_N(\theta_{\mathcal{M}^o}) - EQ_N(\theta_{\mathcal{M}^o}^o) > 0 \implies \liminf_{N \to \infty} \left[ \inf_{\Theta_{\mathcal{M}^o}} EQ_N(\theta_{\mathcal{M}^o}) - EQ_N(\theta_{\mathcal{M}^o}^o) \right] > 0 \quad (A3)$$

Therefore, $\theta_{\mathcal{M}^o}^o$ is the “identifiably unique” [Pötscher and Prucha, 1997, p.16] minimizer of $EQ_N(\theta_{\mathcal{M}^o})$.

\[
\text{Lemma 3. (ULLN). Let a sequence of random variables } \{ \varepsilon(t) : t = 1, \ldots, N \} \text{ be subject to assumption A4. Then:} \\
\]

$$\begin{align*}
\text{plim} \left( \frac{1}{N} \sum_{i=1}^{N} \varepsilon(t) \right) &= 0 \quad (A4) \\
\text{plim} \left( \frac{1}{N} \sum_{i=1}^{N} E[\varepsilon(t)] \right) &= O(1) \quad (A5) \\
\text{plim} \left( \frac{1}{N} \sum_{i=1}^{N} (\varepsilon^2(t) - E(\varepsilon^2(t))) \right) &= 0 \quad (A6)
\end{align*}$$

with $O$ designating the Big O notation.

Proof. Equation (A.4) may be proven via considering (19).

$$\frac{1}{N} \sum_{i=1}^{N} \varepsilon(t) = \frac{1}{N} \sum_{i=1}^{N} \varepsilon(t) \triangleq w(t) \sum_{i=1}^{N} h_i(r-t) \quad (A7)$$

Since $H_i(\theta)$ is a uniformly stable family of filters [Ljung, 1997, p.20] and $w(t)$ is a sequence of independent random variables with $E(w(t)) = 0$ and $\sigma^2_n(w(t)) < \infty$, then:

$$E(\varepsilon(t)) = 0, \quad \sigma^2_n(w(t)) = \left( \sum_{i=1}^{N} h_i(r-t) \right)^2 \cdot \sigma^2_n(w(t)) < \infty \quad (A8)$$

Then due to the 1st Kolmogorov theorem [Rao, 1974, p.114]:

$$\frac{1}{N} \sum_{i=1}^{N} w(t) \to 0 \quad w.p.1 \text{ as } N \to \infty \quad (A9)$$

The proof of (A.6) is provided by Ljung [Ljung, 1997, pp.34-35].

\[
\text{Lemma 4. (ULLN). Under assumptions A3 and A4 the cost function (14) converges i.i.d. to its expected counterpart uniformly over the parameter space, that is } \forall \theta_{\mathcal{M}^o} \in \Theta_{\mathcal{M}^o}: \\
\]

$$\begin{align*}
\text{plim} \left( \sup_{\Theta_{\mathcal{M}^o}} \left\{ EQ_N(\theta_{\mathcal{M}^o}) - EQ_N(\theta_{\mathcal{M}^o}^o) \right\} \right) &= 0 \quad (A12)
\end{align*}$$

\[
\text{Proof.} \quad \text{The required Uniform LLN for the current estimator is derived by employing corollary 3.1 proposed by Newey [1991], which includes three requirements for its verification:} \\
\]

1. A compact parameter space (see assumption A3).
2. The following pointwise LLN, $\forall \theta_{\mathcal{M}^o} \in \Theta_{\mathcal{M}^o}$: \quad \text{plim} \left( \sup_{\Theta_{\mathcal{M}^o}} \left\{ EQ_N(\theta_{\mathcal{M}^o}) - EQ_N(\theta_{\mathcal{M}^o}^o) \right\} \right) = 0 \quad (A13)
3. The following Lipschitz condition, $\forall \theta_{\mathcal{M}^o}, \theta_{\mathcal{M}^o} \in \Theta_{\mathcal{M}^o}$ regarding the summands, that are: $q(t, \theta_{\mathcal{M}^o}) \triangleq \varepsilon^2(t, \theta_{\mathcal{M}^o})$, of (14):

$$|q(t, \theta_{\mathcal{M}^o}) - q(t, \theta_{\mathcal{M}^o})| \leq b(t) \cdot h(d_\theta(\theta_{\mathcal{M}^o}, \theta_{\mathcal{M}^o})) \quad (A14)$$

with the function $h : [0, \infty) \to [0, \infty)$ being continuous at zero as well as $h(0) = 0$. Also the function $b(t)$ is such that $\frac{1}{T} \sum_{t=1}^{T} b(t) = O(1)$

The second requirement may be achieved by considering also (7), (8), (14) and (19). Thus, following some algebra, the leftmost part of (A.13) becomes:

$$\begin{align*}
\text{plim} \left( \frac{1}{N} \sum_{i=1}^{N} (\varepsilon^2(t) - E(\varepsilon^2(t))) \right) + \text{plim} \left( \frac{2}{N} \sum_{i=1}^{N} (\varepsilon(t) - E(\varepsilon(t))) \right)
\end{align*}$$

with: $\varepsilon(t) \triangleq f(t, \theta_{\mathcal{M}^o}) - f(t, \theta_{\mathcal{M}^o}^o) \cdot \varepsilon(t).$
The leftmost part of (A.15) is zero due to Lemma 3 [see (A.6)]. As far as the rightmost part is concerned, first consider that due to assumption A3, $f(t, \theta_{\Delta}) = \int f(t, \theta_{\Delta}) < \infty$, $\forall t$ and $\forall \theta_{\Delta} \in \Theta_{\Delta}$. Thus $\mathbb{E} 0(t) = \infty$, $\forall t$. Since $\mathbb{E} (t)$ is also obtained by passing the sequence of independent random variables $w(t)$ through a uniformly stable filter with coefficients $h(t) = f(t, \theta_{\Delta})$ [in a similar manner as $\mathbb{E} (t)$ is obtained – see (19)], then the LLN described by (A.4) of Lemma 3 is also valid for the $\mathbb{E} (t)$ sequence. This proves that the rightmost part is also zero and therefore the pointwise LLN of (A.13) exists.

The third requirement may be met by employing Lemma 1 and (A.5) of Lemma 3 as follows:

Since $q(t, \theta_{\Delta}) \equiv \mathbb{E} (t, \theta_{\Delta})$, then after some algebra, using also (7) – (8):

$$q(t, \theta_{\Delta}) - q(t, \bar{\theta}_{\Delta}) \leq b(t) \cdot C \cdot h(d_{\Delta}(\theta_{\Delta}, \bar{\theta}_{\Delta}))$$

(A.16)

$$h(d_{\Delta}(\theta_{\Delta}, \bar{\theta}_{\Delta})) = \sum_{i=1}^{M_{\Delta}} \left( |k_i| - |\Delta_i - \bar{\Delta}_i| \right)$$

(A.17)

Due to the assumption A2 and (A.5) of Lemma 3:

$$\frac{1}{N} \sum_{i=1}^{N} E \{ b(t) \} = O(1)$$

(A.19)

Thus the required Lipschitz condition is also valid. \[ \square \]

REFERENCES


