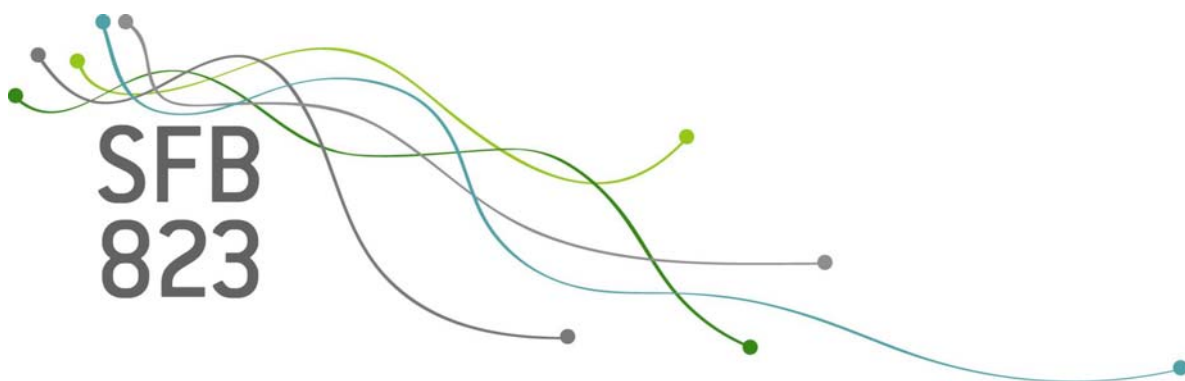


SFB  
823

# On quadratic expansions of log-likelihoods and a general asymptotic linearity result

Marc Hallin, Ramon van den Akker,  
Bas J.M. Werker

Nr. 34/2013



Discussion Paper



# On Quadratic Expansions of Log-Likelihoods and a General Asymptotic Linearity Result

Marc Hallin, Ramon van den Akker, and Bas J.M. Werker

**Abstract** Irrespective of the statistical model under study, the derivation of limits, in the Le Cam sense, of sequences of local experiments (see [7]-[10]) often follows along very similar lines, essentially involving differentiability in quadratic mean of square roots of (conditional) densities. This chapter establishes two abstract and very general results providing sufficient and nearly necessary conditions for (i) the existence of a quadratic expansion, and (ii) the asymptotic linearity of local log-likelihood ratios (asymptotic linearity is needed, for instance, when unspecified model parameters are to be replaced, in some statistic of interest, with some preliminary estimator). Such results have been established, for *locally asymptotically normal* (LAN) models involving independent and identically distributed observations, by, e.g., [1], [11] and [12]. Similar results are provided here for models exhibiting serial dependencies which, so far, have been treated on a case-by-case basis (see [4] and [5] for typical examples) and, in general, under stronger regularity assumptions. Unlike their i.i.d. counterparts, our results extend beyond the context of LAN experiments, so that non-stationary unit-root time series and cointegration models, for instance, also can be handled (see [6]).

---

Marc Hallin  
ECARES, Université Libre de Bruxelles, ORFE, Princeton University, and Econometrics Group,  
Tilburg University; e-mail: mhallin@ulb.ac.be

Ramon van den Akker  
Econometrics Group, Tilburg University; e-mail: R.vdnAkker@TilburgUniversity.edu

Bas J.M. Werker  
Econometrics Group and Finance Group, Tilburg University; e-mail:  
Werker@TilburgUniversity.edu

## 1 Main notation and some preliminary results

For each  $T \in \mathbb{N}$ , let  $(\Omega_T, \mathcal{F}_T)$  be a measurable space on which two probability measures,  $\tilde{\mathbb{P}}_T$  and  $\mathbb{P}_T$ , are defined. Let  $\mathcal{F}_{T0} \subset \dots \subset \mathcal{F}_{TT} \subset \mathcal{F}_T$  be a sequence of increasing  $\sigma$ -fields. Still for  $T \in \mathbb{N}$ , define the restrictions  $\tilde{\mathbb{P}}_T := \tilde{\mathbb{P}}_T|_{\mathcal{F}_{TT}}$  and  $\mathbb{P}_T := \mathbb{P}_T|_{\mathcal{F}_{TT}}$  of  $\tilde{\mathbb{P}}_T$  and  $\mathbb{P}_T$ , respectively, to  $\mathcal{F}_{TT}$ . Using obvious notation, similarly define, for  $t = 0, \dots, T$ , the restrictions  $\tilde{\mathbb{P}}_{Tt} := \tilde{\mathbb{P}}_T|_{\mathcal{F}_{Tt}}$  and  $\mathbb{P}_{Tt} := \mathbb{P}_T|_{\mathcal{F}_{Tt}}$ . The Lebesgue decomposition of  $\tilde{\mathbb{P}}_{Tt}$  on  $\mathbb{P}_{Tt}$  (with respect to  $\mathcal{F}_{Tt}$ ) takes the form

$$\tilde{\mathbb{P}}_{Tt}(A) = \int_A L_{Tt} d\mathbb{P}_{Tt} + \tilde{\mathbb{P}}_{Tt}(A \cap N_{Tt}) \quad A \in \mathcal{F}_{Tt},$$

where  $N_{Tt} \in \mathcal{F}_{Tt}$  is such that  $\mathbb{P}_{Tt}(N_{Tt}) = 0$  and  $L_{Tt}$  is the Radon-Nikodym derivative of that part of  $\tilde{\mathbb{P}}_{Tt}$  which is absolutely continuous with respect to  $\mathbb{P}_{Tt}$ .

The likelihood ratio statistic  $LR_T$  for  $\tilde{\mathbb{P}}_T$  with respect to  $\mathbb{P}_T$  is, by definition,  $L_{TT}$ . Put  $LR_{T0} := L_{T0}$ , and define the conditional likelihood ratio contribution of observation  $t$  as

$$LR_{Tt} := L_{Tt}/L_{T, t-1}, \quad t = 1, \dots, T,$$

with the convention  $0/0 = 1$ . Then, the likelihood ratio statistic  $LR_T$  factorizes into

$$LR_T = \prod_{t=0}^T LR_{Tt}, \quad \mathbb{P}_T\text{-a.s.}$$

This factorization follows from the fact that, under  $\mathbb{P}_T$ ,  $\{L_{Tt} : 0 \leq t \leq T\}$  is a supermartingale with respect to the filtration  $\{\mathcal{F}_{Tt} : 0 \leq t \leq T\}$  (which is easy to check) by repeated application of the following Lemma with  $X = L_{Tt}$ ,  $Y = L_{T, t-1}$ , and  $\mathcal{F} = \mathcal{F}_{T, t-1}$ , and  $t = 1, \dots, T$ .

**Lemma 1.** *Let  $X$  be a nonnegative, integrable random variable and  $Y$  a  $\mathcal{F}$ -measurable random variable satisfying  $Y \geq \mathbb{E}[X|\mathcal{F}]$ . Then,  $X \mathbb{1}_{\{Y=0\}} = 0$  a.s.*

*Proof.* This readily follows from the fact that

$$0 \leq \mathbb{E}X \mathbb{1}_{\{Y=0\}} = \mathbb{E}\mathbb{E}[X|\mathcal{F}] \mathbb{1}_{\{Y=0\}} \leq \mathbb{E}Y \mathbb{1}_{\{Y=0\}} = 0. \quad \square$$

We conclude this section with two lemmas that are needed in the sequel. The first one is a consequence of Theorem 2.23 and Corollary 3.1 in [3]. We refer to Lemma 2.2 in [2] for additional details.

**Lemma 2.** *If, for all  $T \in \mathbb{N}$ , the square-integrable process  $\{X_{Tt} : 1 \leq t \leq T\}$  is adapted to the filtration  $(\mathcal{F}_{Tt})_{0 \leq t \leq T}$  and satisfies  $\sum_{t=1}^T \mathbb{E}[X_{Tt}^2 | \mathcal{F}_{T, t-1}] = o_{\mathbb{P}}(1)$  as  $T \rightarrow \infty$ , then,*

$$\sum_{t=1}^T X_{Tt}^2 = o_{\mathbb{P}}(1) \quad \text{and} \quad \sum_{t=1}^T (X_{Tt} - \mathbb{E}[X_{Tt} | \mathcal{F}_{T, t-1}]) = o_{\mathbb{P}}(1)$$

as  $T \rightarrow \infty$ .

The second lemma follows by an application of a result due to Dvoretzky (see the proof of Theorem 2.23 in [3]).

**Lemma 3.** *If, for all  $T \in \mathbb{N}$ , the process  $\{X_{Tt} : 1 \leq t \leq T\}$  is adapted to the filtration  $(\mathcal{F}_{Tt})_{0 \leq t \leq T}$  and satisfies, for all  $\delta > 0$ ,*

$$\sum_{t=1}^T \mathbb{E} [X_{Tt}^2 \mathbf{1}_{\{|X_{Tt}| > \delta\}} \mid \mathcal{F}_{T, t-1}] = o_{\mathbb{P}}(1)$$

as  $T \rightarrow \infty$ , then  $\max_{t=1, \dots, T} |X_{Tt}| = o_{\mathbb{P}}(1)$  as  $T \rightarrow \infty$ .

## 2 Quadratic expansions of log likelihood ratios

The following proposition provides a general sufficient condition for the existence of a quadratic expansion of local log likelihood ratios. All limits,  $o_{\mathbb{P}}$ , and  $O_{\mathbb{P}}$  quantities are to be understood as  $T \rightarrow \infty$ .

**Proposition 1.** *Suppose that, for some  $k \in \mathbb{N}$ , there exist, for each  $T \in \mathbb{N}$ ,  $\mathcal{F}_{Tt}$ -measurable mappings  $S_{Tt} : \Omega_T \rightarrow \mathbb{R}^k$  and  $R_{Tt} : \Omega_T \rightarrow \mathbb{R}$ ,  $t = 1, \dots, T$ , such that the conditional likelihood ratio contribution  $LR_{Tt}$  can be written as*

$$LR_{Tt} = \left( 1 + \frac{1}{2} (h'_T S_{Tt} + R_{Tt}) \right)^2, \quad (1)$$

where

- (a)  $h_T$  is a bounded (deterministic) sequence in  $\mathbb{R}^k$ ,
- (b) for each  $T \in \mathbb{N}$ ,  $\{S_{Tt} : 1 \leq t \leq T\}$  is a  $\mathbb{P}_T$ -square integrable martingale difference array with respect to the filtration  $\{\mathcal{F}_{Tt} : 0 \leq t \leq T\}$ , satisfying the conditional Lindeberg condition and with tight squared conditional moments, i.e., such that, under  $\mathbb{P}_T$ ,

$$\mathbb{E}_{\mathbb{P}_T} [S_{Tt} \mid \mathcal{F}_{T, t-1}] = 0, \quad t = 1, \dots, T, \quad (2)$$

$$\sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} \left[ (h'_T S_{Tt})^2 \mathbf{1}_{\{|h'_T S_{Tt}| > \delta\}} \mid \mathcal{F}_{T, t-1} \right] = o_{\mathbb{P}}(1) \text{ for all } \delta > 0, \quad (3)$$

and

$$J_T := \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [S_{Tt} S'_{Tt} \mid \mathcal{F}_{T, t-1}] = O_{\mathbb{P}}(1),$$

- (c) the remainder terms  $R_{Tt}$  and the null-sets  $N_{Tt}$  from the Lebesgue decomposition of  $\mathbb{P}_T$  on  $\mathbb{P}_T$  are sufficiently small, i.e., under  $\mathbb{P}_T$ ,

$$\sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [R_{Tt}^2 \mid \mathcal{F}_{T, t-1}] = o_{\mathbb{P}}(1) \quad (4)$$

and

$$\sum_{t=1}^T (1 - \mathbb{E}_{\mathbb{P}_T} [LR_{Tt} \mid \mathcal{F}_{T, t-1}]) = o_{\mathbb{P}}(1), \quad (5)$$

(d) under  $\mathbb{P}_T$ ,  $\log LR_{T0} = o_{\mathbb{P}}(1)$ ,

then, under  $\mathbb{P}_T$ , the log likelihood ratio admits the quadratic expansion

$$\log LR_T = h'_T \sum_{t=1}^T S_{Tt} - \frac{1}{2} h'_T J_T h_T + o_{\mathbb{P}}(1). \quad (6)$$

*Proof.* Let  $r : 2x \mapsto r(2x) := 2(\log(1+x) - x + x^2/2)$ , and rewrite the log likelihood ratio statistic as

$$\begin{aligned} \log LR_T &= \sum_{t=0}^T \log LR_{Tt} = o_{\mathbb{P}}(1) + \sum_{t=1}^T h'_T S_{Tt} - \frac{1}{2} h'_T J_T h_T \\ &\quad + \frac{1}{4} \left( h'_T J_T h_T - \sum_{t=1}^T (h'_T S_{Tt})^2 \right) + \sum_{t=1}^T (R_{Tt} - \mathbb{E}_{\mathbb{P}_T} [R_{Tt} \mid \mathcal{F}_{T, t-1}]) \\ &\quad - \frac{1}{4} \sum_{t=1}^T R_{Tt}^2 - \frac{1}{2} \sum_{t=1}^T h'_T S_{Tt} R_{Tt} + \left( \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [R_{Tt} \mid \mathcal{F}_{T, t-1}] + \frac{1}{4} h'_T J_T h_T \right) \\ &\quad + \sum_{t=1}^T r(h'_T S_{Tt} + R_{Tt}), \end{aligned} \quad (7)$$

where we used Condition (d) to neglect the first term  $\log LR_{T0}$ . To establish (6), we show that the six remainder terms on the right-hand side of (7) all converge to zero in probability under  $\mathbb{P}_T$ .

By Theorem 2.23 in [3], Condition (a) and (2)-(1) we have

$$\sum_{t=1}^T (h'_T S_{Tt})^2 - h'_T J_T h_T = o_{\mathbb{P}}(1), \quad (8)$$

which shows that the first remainder term is indeed  $o_{\mathbb{P}}(1)$ .

Since  $(L_{Tt})_{0 \leq t \leq T}$  is a  $\mathbb{P}_T$ -supermartingale, we have  $\mathbb{E}_{\mathbb{P}_T} LR_{Tt} \leq 1$ . Since  $S_{Tt}$  is also  $\mathbb{P}_T$ -square integrable, it follows from (1) that  $R_{Tt}$  is  $\mathbb{P}_T$ -square integrable. From Lemma 2 and (4), we now immediately obtain

$$\sum_{t=1}^T (R_{Tt} - \mathbb{E}_{\mathbb{P}_T} [R_{Tt} \mid \mathcal{F}_{T, t-1}]) = o_{\mathbb{P}}(1) \text{ and } \sum_{t=1}^T R_{Tt}^2 = o_{\mathbb{P}}(1), \quad (9)$$

i.e. the second and third remainder terms also are negligible.

Next we show that the remainder term  $(1/2) \sum_{t=1}^T h'_T S_{Tt} R_{Tt}$  vanishes asymptotically. First note that Condition (a), (1) and (8) jointly imply  $\sum_{t=1}^T (h'_T S_{Tt})^2 = O_{\mathbb{P}}(1)$ . Combined with (9), an application of the Cauchy-Schwarz inequality thus yields the convergence of the fourth remainder term.

To prove the negligibility of the fifth remainder term in (7), observe that (1), (2), (1), (4), combined with the Cauchy-Schwarz inequality again, entail

$$\begin{aligned}
 \sum_{t=1}^T (\mathbb{E}_{P_T} [LR_{Tt} | \mathcal{F}_{T, t-1}] - 1) &= \sum_{t=1}^T \mathbb{E}_{P_T} [h'_T S_{Tt} | \mathcal{F}_{T, t-1}] + \sum_{t=1}^T \mathbb{E}_{P_T} [R_{Tt} | \mathcal{F}_{T, t-1}] \\
 &\quad + \frac{1}{4} \sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt})^2 | \mathcal{F}_{T, t-1}] + \frac{1}{4} \sum_{t=1}^T \mathbb{E}_{P_T} [R_{Tt}^2 | \mathcal{F}_{T, t-1}] \\
 &\quad + \frac{1}{2} \sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt}) R_{Tt} | \mathcal{F}_{T, t-1}] \\
 &= \sum_{t=1}^T \mathbb{E}_{P_T} [R_{Tt} | \mathcal{F}_{T, t-1}] + \frac{1}{4} h'_T J_T h_T + o_P(1).
 \end{aligned}$$

Now, the second part of (4) implies

$$\sum_{t=1}^T \mathbb{E}_{P_T} [R_{Tt} | \mathcal{F}_{T, t-1}] + \frac{1}{4} h'_T J_T h_T = o_P(1). \quad (10)$$

Thus, the fifth remainder term in (7) also is negligible.

Turning to the sixth and last remainder term, let us first show that

$$\max_{t=1, \dots, T} |h'_T S_{Tt} + R_{Tt}| = o_P(1) \quad \text{and} \quad \sum_{t=1}^T |h'_T S_{Tt} + R_{Tt}|^3 = o_P(1). \quad (11)$$

As (3) and (4) yield, for  $\delta > 0$ ,

$$\begin{aligned}
 &\sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt} + R_{Tt})^2 \mathbb{1}_{\{|h'_T S_{Tt} + R_{Tt}| > \delta\}} | \mathcal{F}_{T, t-1}] \\
 &\leq 4 \sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt})^2 \mathbb{1}_{\{|h'_T S_{Tt}| > \delta/2\}} | \mathcal{F}_{T, t-1}] + 4 \sum_{t=1}^T \mathbb{E}_{P_T} [R_{Tt}^2 | \mathcal{F}_{T, t-1}] \\
 &= o_P(1),
 \end{aligned}$$

the first part of (11) follows as an application of Lemma 3. The second part is obtained from the latter by taking out the maximum (which tends to zero) and by observing that the remaining quadratic term is bounded in probability. In view of the first part of (11), indeed, it is sufficient to study the behavior of the final remainder term on the event  $\{|h'_T S_{Tt} + R_{Tt}| \leq 1\}$ . On this set, this remainder term is bounded: using the fact that

$$\left| \log(1+x) - x + \frac{1}{2}x^2 \right| \leq \frac{2}{3}x^3 \quad \text{for} \quad |x| \leq \frac{1}{2},$$

indeed, we obtain

$$\left| \sum_{t=1}^T r (h'_T S_{Tt} + R_{Tt}) \right| \leq \frac{4}{3} \sum_{t=1}^T (h'_T S_{Tt} + R_{Tt})^3.$$

Convergence to zero is now obtained from the second part of (11). This completes the proof of the proposition.  $\square$

### 3 Asymptotic linearity: general result

This section provides a sufficient condition for the asymptotic linearity of a fairly general class of statistics, extending and generalizing Proposition A.10 in [11] to the case of serially dependent observations under possibly non-LAN limit experiments.

All limits are taken as  $T \rightarrow \infty$ .

**Proposition 2.** *Let, for each  $T \in \mathbb{N}$ ,  $\{\tilde{Z}_{Tt} : 1 \leq t \leq T\}$  and  $\{Z_{Tt} : 1 \leq t \leq T\}$  be a  $\tilde{\mathbb{P}}_T$ -square integrable martingale difference array, and a  $\mathbb{P}_T$ -square integrable martingale difference array, respectively. Suppose that Conditions (a)-(d) in Proposition 1 hold, as well as the following Conditions (e)-(h):*

- (e)  $(\sum_{t=1}^T S_{Tt}, J_T)$  converges in distribution, under  $\mathbb{P}_T$ , to a limit  $(\Delta, J)$  that satisfies, for all  $a \in \mathbb{R}^k$ ,  $E \exp(a' \Delta - \frac{1}{2} a' J a) = 1$ ;
- (f)  $\sum_{t=1}^T E_{\mathbb{P}_T} \left[ (\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt})^2 \mid \mathcal{F}_{T, t-1} \right] = o_{\mathbb{P}}(1)$  under  $\mathbb{P}_T$ ;
- (g)  $\sum_{t=1}^T E_{\tilde{\mathbb{P}}_T} [\tilde{Z}_{Tt}^2 \mid \mathcal{F}_{T, t-1}] = O_{\mathbb{P}}(1)$  under  $\tilde{\mathbb{P}}_T$ , and  $\sum_{t=1}^T E_{\mathbb{P}_T} [Z_{Tt}^2 \mid \mathcal{F}_{T, t-1}] = O_{\mathbb{P}}(1)$  under  $\mathbb{P}_T$ ;
- (h) the conditional Lindeberg condition holds for  $\{\tilde{Z}_{Tt} : 1 \leq t \leq T\}$  under  $\tilde{\mathbb{P}}_T$ , namely, for all  $\delta > 0$ ,  $\sum_{t=1}^T E_{\tilde{\mathbb{P}}_T} \left[ \tilde{Z}_{Tt}^2 \mathbb{1}_{\{|\tilde{Z}_{Tt}| > \delta\}} \mid \mathcal{F}_{T, t-1} \right] = o_{\mathbb{P}}(1)$  under  $\tilde{\mathbb{P}}_T$ .

Then, letting  $\tilde{I}_T := \sum_{t=1}^T \tilde{i}_{Tt} := \sum_{t=1}^T E_{\mathbb{P}_T} [(h'_T S_{Tt}) Z_{Tt} \mid \mathcal{F}_{T, t-1}]$ , we have, under  $\mathbb{P}_T$ ,

$$\sum_{t=1}^T \tilde{Z}_{Tt} = \sum_{t=1}^T Z_{Tt} - \tilde{I}_T + o_{\mathbb{P}}(1). \quad (12)$$

*Proof.* The proof decomposes into four parts. In Part 1, we show that (12) holds if, under  $\mathbb{P}_T$ ,

$$\sum_{t=1}^T \tilde{Z}_{Tt} (1 - \sqrt{LR_{Tt}}) + \frac{1}{2} \tilde{I}_T = o_{\mathbb{P}}(1). \quad (13)$$

In Part 2, we show that (13) holds provided that, still under  $\mathbb{P}_T$ ,



$$\sum_{t=1}^T \tilde{Z}_{Tt}(h'_T S_{Tt}) - \tilde{I}_T = o_{\mathbb{P}}(1). \quad (14)$$

In Part 3, we introduce a new sequence of probability measures  $(\mathbb{P}'_T)$  and show that it is contiguous to  $(\mathbb{P}_T)$ . In Part 4, we establish that (14) holds under the new sequence  $(\mathbb{P}'_T)$ . In view of contiguity, it also holds under  $(\mathbb{P}_T)$ , which concludes the proof.

Note that Lemma 1, Condition (e), and Le Cam's first lemma imply that  $(\tilde{\mathbb{P}}_T)$  and  $(\mathbb{P}_T)$  are contiguous. It follows that  $o_{\mathbb{P}}$ 's and  $O_{\mathbb{P}}$ 's under  $(\tilde{\mathbb{P}}_T)$  and  $(\mathbb{P}_T)$  coincide; therefore, in the sequel, we safely can write  $o_{\mathbb{P}}$  and  $O_{\mathbb{P}}$  without specifying whether  $(\tilde{\mathbb{P}}_T)$  or  $(\mathbb{P}_T)$  is the underlying sequence of probability measures.

*Part 1.* Recalling that  $\tilde{I}_T := \sum_{t=1}^T \tilde{I}_{Tt} := \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [(h'_T S_{Tt}) Z_{Tt} \mid \mathcal{F}_{T, t-1}]$ , we have

$$\begin{aligned} \sum_{t=1}^T \{\tilde{Z}_{Tt} - Z_{Tt} + \tilde{I}_{Tt}\} &= \sum_{t=1}^T \tilde{Z}_{Tt} (1 - \sqrt{LR_{Tt}}) + \frac{1}{2} \tilde{I}_T \\ &\quad + \sum_{t=1}^T \{\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt} - \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}} \mid \mathcal{F}_{T, t-1}]\} \\ &\quad + \sum_{t=1}^T \left\{ \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}} \mid \mathcal{F}_{T, t-1}] + \frac{1}{2} \tilde{I}_{Tt} \right\}; \end{aligned}$$

hence, (13) implies (12) in case

$$\sum_{t=1}^T \{\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt} - \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}} \mid \mathcal{F}_{T, t-1}]\} = o_{\mathbb{P}}(1) \quad (15)$$

and

$$\sum_{t=1}^T \left\{ \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}} \mid \mathcal{F}_{T, t-1}] + \frac{1}{2} \tilde{I}_{Tt} \right\} = o_{\mathbb{P}}(1). \quad (16)$$

As (15) is implied by Condition (f) and Lemma 2 (recall that  $\mathbb{E}_{\mathbb{P}_T} [Z_{Tt} \mid \mathcal{F}_{T, t-1}] = 0$ ), we only need to show that (16) holds in order to complete Part 1. We have

$$\begin{aligned} \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}} \mid \mathcal{F}_{T, t-1}] &= \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [Z_{Tt} (1 - \sqrt{LR_{Tt}}) \mid \mathcal{F}_{T, t-1}] \\ &\quad + \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [(\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt}) (1 - \sqrt{LR_{Tt}}) \mid \mathcal{F}_{T, t-1}] \\ &\quad + \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} LR_{Tt} \mid \mathcal{F}_{T, t-1}] \\ &= -\frac{1}{2} \tilde{I}_T - \frac{1}{2} r_T^{(1)} + r_T^{(2)} + r_T^{(3)}, \end{aligned}$$

with

$$\begin{aligned} r_T^{(1)} &= \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [Z_{Tt} R_{Tt} \mid \mathcal{F}_{T, t-1}], \\ r_T^{(2)} &= \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} \left[ (\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt})(1 - \sqrt{LR_{Tt}}) \mid \mathcal{F}_{T, t-1} \right], \text{ and} \\ r_T^{(3)} &= \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} LR_{Tt} \mid \mathcal{F}_{T, t-1}]. \end{aligned}$$

Starting with  $r_T^{(1)}$ ,

$$\begin{aligned} |r_T^{(1)}|^2 &\leq \left( \sum_{t=1}^T \sqrt{\mathbb{E}_{\mathbb{P}_T} [Z_{Tt}^2 \mid \mathcal{F}_{T, t-1}]} \sqrt{\mathbb{E}_{\mathbb{P}_T} [R_{Tt}^2 \mid \mathcal{F}_{T, t-1}]} \right)^2 \\ &\leq \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [Z_{Tt}^2 \mid \mathcal{F}_{T, t-1}] \sum_{t=1}^T \mathbb{E}_{\mathbb{P}_T} [R_{Tt}^2 \mid \mathcal{F}_{T, t-1}], \end{aligned}$$

so that (4) and Condition (g) imply  $r_T^{(1)} = o_{\mathbb{P}}(1)$ . In the same way, (1), (4) and Condition (f) yield  $r_T^{(2)} = o_{\mathbb{P}}(1)$ . As for  $r_T^{(3)}$ , since  $\mathbb{E}_{\tilde{\mathbb{P}}_T} [\tilde{Z}_{Tt} \mid \mathcal{F}_{T, t-1}] = 0$ , we obtain, using (4) and Condition (g) again,

$$\begin{aligned} |r_T^{(3)}|^2 &= \left| \sum_{t=1}^T \mathbb{E}_{\tilde{\mathbb{P}}_T} [\tilde{Z}_{Tt} \mathbb{1}_{N_{Tt}} \mid \mathcal{F}_{T, t-1}] \right|^2 \\ &\leq \sum_{t=1}^T \mathbb{E}_{\tilde{\mathbb{P}}_T} [\tilde{Z}_{Tt}^2 \mid \mathcal{F}_{T, t-1}] \sum_{t=1}^T (1 - \mathbb{E}_{\mathbb{P}_T} [LR_{Tt} \mid \mathcal{F}_{T, t-1}]) = o_{\mathbb{P}}(1). \end{aligned}$$

*Part 2.* We have

$$\left| \sum_{t=1}^T \tilde{Z}_{Tt} (1 - \sqrt{LR_{Tt}}) + \frac{1}{2} \sum_{t=1}^T \tilde{Z}_{Tt} (h'_T S_{Tt}) \right| = \frac{1}{2} \left| \sum_{t=1}^T \tilde{Z}_{Tt} R_{Tt} \right| \leq \frac{1}{2} \sqrt{\sum_{t=1}^T \tilde{Z}_{Tt}^2} \sqrt{\sum_{t=1}^T R_{Tt}^2}.$$

Now, by (9),  $\sum_{t=1}^T R_{Tt}^2 = o_{\mathbb{P}}(1)$  and, by Conditions (g) and (h) and an application of [3, Theorem 2.23],  $\sum_{t=1}^T \tilde{Z}_{Tt}^2 = O_{\mathbb{P}}(1)$ . Hence, (13) follows from (14).

*Part 3.* For all  $T \in \mathbb{N}$ , define the new sequence of probability measures  $(\mathbb{P}'_{Tt})_{t=1}^T$  on  $\mathcal{F}_{Tt}$ , absolutely continuous with respect to  $\mathbb{P}_{Tt}$ , with density

$$\frac{d\mathbb{P}'_{Tt}}{d\mathbb{P}_{Tt}} := \prod_{s=1}^t \sqrt{LR_{Ts}} c_{Ts}$$

with, for  $s = 1, \dots, T$ ,  $c_{Ts}^{-1} := \mathbb{E}_{\mathbb{P}_T} [\sqrt{LR_{Ts}} \mid \mathcal{F}_{T, s-1}]$ . Note that the probability that all  $c_{Ts}^{-1}$  are strictly positive tends to one, since (4) implies

$$\lim_{T \rightarrow \infty} \mathbb{P}_T [\exists s \in \{1, \dots, T\} : c_{Ts}^{-1} = 0] \leq \lim_{T \rightarrow \infty} \mathbb{P}_T \left[ \sum_{t=1}^T (1 - \mathbb{E}_{\mathbb{P}_T} [LR_{Tt} | \mathcal{F}_{T, t-1}]) \geq 1 \right] = 0.$$

In the sequel, we thus safely can ignore the event  $\{\exists s \in \{1, \dots, T\} : c_{Ts}^{-1} = 0\}$ . Defining  $\mathbb{P}'_T := \mathbb{P}'_{TT}$ , note that  $\mathbb{P}'_{Tt}$  is the restriction of  $\mathbb{P}'_T$  to  $\mathcal{F}_{T,t}$ . Because of (2), we have  $c_{Ts}^{-1} = 1 + \frac{1}{2} \mathbb{E}_{\mathbb{P}_T} [R_{Ts} | \mathcal{F}_{T, t-1}]$ . This yields, using an expansion of  $\log(1+x)$ , (4), and (10),

$$\sum_{t=1}^T \log c_{Tt}^{-1} = -\frac{1}{8} h'_T J_T h_T + o_{\mathbb{P}}(1).$$

Moreover, an application of Lemma 3 and (4) yields  $\max_{t=1, \dots, T} |c_{Tt}^{-1} - 1| = o_{\mathbb{P}}(1)$ , and thus also

$$\max_{t=1, \dots, T} |c_{Tt} - 1| = o_{\mathbb{P}}(1). \quad (17)$$

Inserting (6) and recalling that  $\log LR_{T0} = o_{\mathbb{P}}(1)$ , we obtain, under  $\mathbb{P}_T$ ,

$$\log \frac{d\mathbb{P}'_T}{d\mathbb{P}_T} = \frac{1}{2} \sum_{t=1}^T \log LR_{Tt} - \sum_{t=1}^T \log c_{Tt}^{-1} + o_{\mathbb{P}}(1) = \frac{1}{2} \sum_{t=1}^T h'_T S_{Tt} - \frac{1}{8} h'_T \tilde{I}_T h_T + o_{\mathbb{P}}(1).$$

Condition (e) and Le Cam's first lemma entail that the sequences  $(\mathbb{P}'_T)$  and  $(\mathbb{P}_T)$  are mutually contiguous. This completes Part 3 of the proof.

*Part 4.* Let us show that, under the measures  $(\mathbb{P}'_T)$ ,

$$\sum_{t=1}^T \mathbb{E}_{\mathbb{P}'_T} [\tilde{Z}_{Tt}(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}] = \tilde{I}_T + o_{\mathbb{P}}(1) \quad (18)$$

and

$$\sum_{t=1}^T \tilde{Z}_{Tt}(h'_T S_{Tt}) = \sum_{t=1}^T \mathbb{E}_{\mathbb{P}'_T} [\tilde{Z}_{Tt}(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}] + o_{\mathbb{P}}(1). \quad (19)$$

Since  $o_{\mathbb{P}}(1)$ 's under  $(\mathbb{P}'_T)$  are  $o_{\mathbb{P}}(1)$ 's under the contiguous  $(\mathbb{P}_T)$  too, a combination of these two results yields (14) and concludes the proof.

Starting with (18), we have

$$\begin{aligned} \sum_{t=1}^T \mathbb{E}_{\mathbb{P}'_T} [\tilde{Z}_{Tt}(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}] &= \sum_{t=1}^T c_{Tt} \mathbb{E}_{\mathbb{P}_T} [\tilde{Z}_{Tt} \sqrt{LR_{Tt}}(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}] \\ &= \tilde{I}_T + \sum_{t=1}^T (c_{Tt} - 1) \mathbb{E}_{\mathbb{P}_T} [Z_{Tt}(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}] \\ &\quad + \sum_{t=1}^T c_{Tt} \mathbb{E}_{\mathbb{P}_T} [(\tilde{Z}_{Tt} \sqrt{LR_{Tt}} - Z_{Tt})(h'_T S_{Tt}) | \mathcal{F}_{T, t-1}]. \end{aligned}$$

Condition (f) and (17) imply (18) since  $\sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt})^2 \mid \mathcal{F}_{T, t-1}] = O_P(1)$  (see (1)) and  $\sum_{t=1}^T \mathbb{E}_{P_T} [Z_{Tt}^2 \mid \mathcal{F}_{T, t-1}] = O_P(1)$  (see Condition (g)).

Turning to (19), first note that  $\sum_{t=1}^T (h'_T S_{Tt})^2 = O_P(1)$  and  $\sum_{t=1}^T \tilde{Z}_{Tt}^2 = O_P(1)$  by an application of [3, Theorem 2.23] and (3), (1), Condition (g) and Condition (h), respectively. Hence,

$$\sum_{t=1}^T |\tilde{Z}_{Tt}| |h'_T S_{Tt}| = O_P(1) \quad \text{and} \quad \sum_{t=1}^T \mathbb{E}_{P_T} [|\tilde{Z}_{Tt}| |h'_T S_{Tt}| \mid \mathcal{F}_{T, t-1}] = O_P(1).$$

Let  $\varepsilon, \delta > 0$ . In view of the previous remarks, we can find  $B$  and  $T_1$  such that, for  $T \geq T_1$ ,

$$P'_T(\mathcal{A}_\delta^{(T)}) \leq \delta/6$$

with

$$\mathcal{A}_\delta^{(T)} := \left\{ \sum_{t=1}^T \left| (h'_T S_{Tt}) \tilde{Z}_{Tt} - \mathbb{E}_{P_T} [(h'_T S_{Tt}) \tilde{Z}_{Tt} \mid \mathcal{F}_{T, t-1}] \right| > B \right\}.$$

Setting  $\eta := \min\{1, \sqrt{\delta\varepsilon}(108(B+2))^{-1/2}\}$  and

$$\mathcal{A}_{\eta, Tt} := \left\{ |Z_{Tt}| \leq \eta \right\} \cap \left\{ |h'_T S_{Tt}| \leq \eta \right\},$$

decompose

$$\sum_{t=1}^T \tilde{Z}_{Tt} (h'_T S_{Tt}) - \sum_{t=1}^T \mathbb{E}_{P_T} [\tilde{Z}_{Tt} (h'_T S_{Tt}) \mid \mathcal{F}_{T, t-1}] = p_T^{(1)} - p_T^{(2)} + p_T^{(3)},$$

with

$$\begin{aligned} p_T^{(1)} &:= \sum_{t=1}^T \tilde{Z}_{Tt} (h'_T S_{Tt}) \mathbb{1}_{\mathcal{A}_{\eta, Tt}^c}, \\ p_T^{(2)} &:= \sum_{t=1}^T \mathbb{E}_{P_T} \left[ \tilde{Z}_{Tt} (h'_T S_{Tt}) \mathbb{1}_{\mathcal{A}_{\eta, Tt}^c} \mid \mathcal{F}_{T, t-1} \right], \text{ and} \\ p_T^{(3)} &:= \sum_{t=1}^T \tilde{Z}_{Tt} (h'_T S_{Tt}) \mathbb{1}_{\mathcal{A}_{\eta, Tt}} - \sum_{t=1}^T \mathbb{E}_{P_T} \left[ \tilde{Z}_{Tt} (h'_T S_{Tt}) \mathbb{1}_{\mathcal{A}_{\eta, Tt}} \mid \mathcal{F}_{T, t-1} \right]. \end{aligned}$$

Let us show that there exists  $T^*$  such that, for all  $T \geq T^*$ ,  $P'_T \left( |p_T^{(i)}| > \varepsilon/3 \right) \leq \delta/3$ , which, as  $\varepsilon > 0$  and  $\delta > 0$  can be taken arbitrarily small, yields (19). Applying Theorem 2.23 in [3], (1), (3), Condition (g) and Condition (h), we obtain

$$\sum_{t=1}^T \tilde{Z}_{Tt}^2 \mathbb{1}_{\{|\tilde{Z}_{Tt}| > \eta\}} + \sum_{t=1}^T (h'_T S_{Tt})^2 \mathbb{1}_{\{|h'_T S_{Tt}| > \eta\}} = o_P(1).$$

This yields, using (1) and Condition (g) again,

$$\begin{aligned}
 |p_T^{(1)}| &\leq \sqrt{\sum_{t=1}^T (h'_T S_{Tt})^2 \mathbf{1}\{|h'_T S_{Tt}| > \eta\}} \sqrt{\sum_{t=1}^T \tilde{Z}_{Tt}^2} \\
 &\quad + \sqrt{\sum_{t=1}^T (h'_T S_{Tt})^2} \sqrt{\sum_{t=1}^T \tilde{Z}_{Tt}^2 \mathbf{1}\{|\tilde{Z}_{Tt}| > \eta\}} = o_P(1).
 \end{aligned}$$

From (3), (1), Condition (g) and Condition (h), we also obtain

$$\begin{aligned}
 |p_T^{(2)}| &\leq \sqrt{\sum_{t=1}^T c_{Tt}^2 \mathbb{E}_{\tilde{P}_T} [\tilde{Z}_{Tt}^2 \mathbf{1}\{|\tilde{Z}_{Tt}| > \eta\} \mid \mathcal{F}_{T, t-1}]} \sqrt{\sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt})^2 \mid \mathcal{F}_{T, t-1}]} \\
 &\quad + \sqrt{\sum_{t=1}^T c_{Tt}^2 \mathbb{E}_{\tilde{P}_T} [\tilde{Z}_{Tt}^2 \mid \mathcal{F}_{T, t-1}]} \sqrt{\sum_{t=1}^T \mathbb{E}_{P_T} [(h'_T S_{Tt})^2 \mathbf{1}\{|h'_T S_{Tt}| > \eta\} \mid \mathcal{F}_{T, t-1}]} \\
 &= o_P(1).
 \end{aligned}$$

Hence, there exists  $T_2$  such that, for all  $T \geq T_2$ ,  $P'_T \left( |p_T^{(j)}| > \varepsilon/3 \right) \leq \delta/3$  for  $j = 1, 2$ .

Next, define the martingales

$$\left\{ A_{Tt} := \sum_{s=1}^t \left\{ \tilde{Z}_{Ts} (h'_T S_{Ts}) \mathbf{1}_{\mathcal{A}_{\eta, Ts}} - \mathbb{E}_{P'_T} [\tilde{Z}_{Ts} (h'_T S_{Ts}) \mathbf{1}_{\mathcal{A}_{\eta, Ts}} \mid \mathcal{F}_{T, s-1}] \right\} : 1 \leq t \leq T \right\},$$

the stopping times  $\mathcal{S}^{(T)} := \inf \{ t \in \mathbb{N} \mid \sum_{s=1}^t |\Delta A_{Ts}| > B \}$ , and the processes

$$\left\{ M_{Tt} := A_{T, t \wedge \mathcal{S}^{(T)}} : 1 \leq t \leq T \right\},$$

namely, the stopped versions of the martingales  $\{A_{Tt} : 1 \leq t \leq T\}$ —which thus also are martingales. Note that  $|\Delta A_{Tt}| \leq 2\eta^2$ . We obtain

$$\begin{aligned}
 \mathbb{E}_{P'_T} M_{TT}^2 &= \sum_{t=1}^T \mathbb{E}_{P'_T} (M_{Tt} - M_{T, t-1})^2 \leq \mathbb{E}_{P'_T} \left[ \sum_{t=1}^{\mathcal{S}^{(T)}} (\Delta A_{Tt})^2 \right] \\
 &\leq 2\eta^2 \mathbb{E}_{P'_T} \left[ \sum_{t=1}^{\mathcal{S}^{(T)}} |\Delta A_{Tt}| \right] \leq 2\eta^2 (B + 2\eta^2).
 \end{aligned}$$

So, for  $T \geq T_1$ , we have

$$\begin{aligned}
 P'_T \left( |p_T^{(3)}| > \varepsilon/3 \right) &= P'_T (|A_{TT}| > \varepsilon/3) \leq P'_T (M_{TT} \neq A_{TT}) + P'_T (|M_{TT}| > \varepsilon/3) \\
 &\leq P'_T (\mathcal{S}^{(T)} \leq T) + P'_T (|M_{TT}| > \varepsilon/3) \\
 &\leq P'_T (\mathcal{A}_\delta^{(T)}) + P'_T (|M_{TT}| > \varepsilon/3) \leq \frac{\delta}{6} + \frac{18\eta^2(B+2)}{\varepsilon^2} \leq \frac{\delta}{3}.
 \end{aligned}$$

Letting  $T^* := \max\{T_1, T_2\}$  completes the proof.  $\square$

**Acknowledgements** The research of Marc Hallin is supported by the Sonderforschungsbereich “Statistical modelling of nonlinear dynamic processes” (SFB 823) of the Deutsche Forschungsgemeinschaft, a Discovery Grant of the Australian Research Council, and the IAP research network grant P7/06 of the Belgian government (Belgian Science Policy).

## References

1. BICKEL, P.J., C.A.J. KLAASSEN, Y. RITOV, AND J.A. WELLNER (1993). Efficient and Adaptive Estimation for Semiparametric Models. Johns Hopkins University Press, Baltimore, MD.
2. DROST, F.C., C.A.J. KLAASSEN, AND B.J.M. WERKER (1997). Adaptive estimation in time series models, *Ann. Statist.* **25**, 786–818.
3. HALL, P. AND C.C. HEYDE (1980). Martingale limit theory and its application. Academic Press, New York.
4. HALLIN, M. AND PAINDAVEINE, D. (2005). Asymptotic linearity of serial and nonserial multivariate signed rank statistics. *Journal of Statistical Planning and Inference* **136**, 1–32.
5. HALLIN, M. AND PURI, M.L. (1984). Aligned rank tests for linear models with autocorrelated errors, *Journal of Multivariate Analysis* **50**, 175–237.
6. HALLIN, M., VAN DEN AKKER, R., AND WERKER, B.J.M. (2013). Optimal Pseudo-Gaussian and Rank-based Tests of the Cointegrating Rank in Semiparametric Error-Correction Models, available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2320810](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2320810).
7. JEGANATHAN, P. (1995). Some aspects of asymptotic theory with applications to time series models, *Econometric Theory* **11**, 818–887.
8. LE CAM, L. (1986). Asymptotic methods in statistical decision theory. New York: Springer-Verlag.
9. LE CAM, L. AND G.L. YANG (1990). Asymptotics in statistics - some basic concepts. New York: Springer.
10. STRASSER, H. (1985). Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory. Walter de Gruyter.
11. VAN DER VAART, A.W. (1988). Statistical Estimation in Large Parameter Spaces. CWI tract 44. Amsterdam: CWI.
12. VAN DER VAART, A.W. (2000). Asymptotic statistics. Cambridge: Cambridge University Press.



