ON THE LIMIT EXPERIMENTS OF RANDOMLY THINNED GARCH(1,1) IN DEFICIENCY *

By Boris Buchmann and Gernot Müller

Monash University and Munich University of Technology

GARCH is the most prominent nonlinear time series model, both widely applied and thoroughly studied. Aggregating its only one-dimensional innovations, it is a well-known oddity that GARCH converges in law to a diffusion driven by a two-dimensional Brownian motion. As a result, the convergence is not in Le Cam's sense of deficiency and, thus, a passage from discrete to continuous time is impossible in all plausible decision problems. Recently an intuitively appealing continuous time version of GARCH (called COGARCH) was introduced which is driven by an only one-dimensional Lévy process, thus, maintaining one of the key features of GARCH. Further investigations have shown that COGARCH occurs as a limit of GARCH models in law when the innovations are randomly thinned. In this paper we investigate the validity of the corresponding approximations in Le Cam's framework of deficiency. We identify the limit experiments for two kinds of sampling schemes. If the corresponding volatilities are unobservable, we show that the limit experiment is not equivalent to COGARCH in deficiency. Otherwise, if, in addition, full information observations about the volatility processes is available, then we show that the limiting experiment is generically equivalent to COGARCH.

1. Introduction. Since the seminal papers by Engle (1982, [6]) and Bollerslev (1986, [3]) GARCH has proved itself to be useful in the modeling of heteroscedasticity in discrete-time financial times series. On the other hand, continuous-time models are more useful, for instance, in option pricing as shown by Black & Scholes (1973, [2]) and Merton (1973, [16]), the analysing tick-by-tick data and observations from irregularly spaced time series. In the 1990's researchers tried to bridge the gap between continuous and discrete time. Most notable, Nelson (1990, [17]) showed that an appropriately parametrised GARCH can be seen as a discrete-time approximation of a bivariate diffusion model on an approximating time grid. At the first glance this is an appealing result as the estimates of the underlying model parameters can be easily obtained by the time series formulation and

*This research was partially supported by ARC grants DP0664603 and CE0348217.

AMS 2000 subject classifications: Primary 62B15, 60J25; secondary 91B84, 91B70

Keywords and phrases: Time series, GARCH, COGARCH, Statistical equivalence, Le Cam's deficiency distance, Random thinning

plugged into the continuous-time limit for other purposes. However, convergence in law, such as studied by Nelson, does usually not allow a passage from discrete to continuous time in all plausible decision problems. This is a more general obstacle in asymptotic statistics and, in order to resolve it, Le Cam advocated the concept of convergence in deficiency [cf. Le Cam (1986, [10]), Le Cam & Young (1990, [11]) and Strasser (1985, [19])]. Since then many authors investigated Le Cam's approach in various statistical problems [Brown & Low (1996, [4]), Nussbaum (1996, [18]) and Grama & Neumann (2006, [7]) and references therein; cf. Milstein & Nussbaum (1998, [12]) with potential applications to time series analysis.]

In the more specific context of Nelson's diffusion limit, Wang (2002, [21]) and Brown et al. (2003, [5]) showed that Nelson's diffusion approximation of GARCH is not in place in deficiency since the innovations encounters both models in an intrinsically different way: whereas GARCH is driven by one-dimensional innovations, its diffusion limit is driven by planar Brownian motion. In Klüppelberg et al. (2004, [13]) an intuitively appealing version of a continuous-time GARCH (COGARCH) was introduced. Among other things, Kallsen & Vesenmayer [9] and Maller et al. [14] have identified COGARCH as a limit of GARCH (in distribution and in probability) provided the innovations are randomly thinned.

As argued above, convergence in distribution may not be satisfactory from a statistician's point of view and, thus, the validity of the corresponding approximations in Le Cam's framework of deficiency are investigated in this paper. In particular, we aim to give an answer to the following questions: does there exist a limit experiment of randomly thinned GARCH in deficiency? If so, is the corresponding limit experiment of COGARCH-type? How does the underlying sampling scheme encounter this scenario?

Dealing with Le Cam's distance in deficiency is a challenging task for a mathematical statistician and little is known. Besides this, we have to cope with innovations in this paper, not aggregated, but randomly thinned. Hence, it is meaningless to compare the corresponding experiments to the usual paradigm in the above-mentioned literature, i.e. *signal recovery with Gaussian white noise*. Similar to [21], further obstacles arise from the intrinsic heteroscedasticity of (CO)GARCH in our analysis.

The paper is organised as follows. Section 2 contains our main results. To be more specific, we introduce the experiments and sampling schemes in Subsection 2.1. In Subsection 2.2 we construct a limiting experiment for randomly thinned GARCH with conditionally variances unobserved. As shown in Subsection 2.3, using both theoretical and numerical methods, this experiment is generically not equivalent to COGARCH. If, however, the conditional variances are observable in full, all experiments are generically (asymptotically) equivalent to COGARCH. This is shown in Subsection 2.4. We conclude in Section 3. Section 4 contains all proofs. In the Appendix we review some of the necessary facts of Le Cam's distance in deficiency.

2. Main results.

2.1. *Garch-type experiments in discrete and continuous time.* We fix a probability measure Q on the Borel field $\mathcal{B}(\mathbb{R})$. For all $n \in \mathbb{N}$ let $p_n \in (0,1)$ and $Z_n = (Z_{n,k})_{1\le k\le n}$ be an n-dimensional vector with distribution

$$\mathcal{L}(Z_n) \;=\; \big((1-p_n)\varepsilon_0 + p_n Q\big)^{\otimes n}. \tag{2.1}$$

The parameter p_n modulates our random thinning. In accordance with the law of rare events we assume that the following limit exists in $(0,\infty)$:

$$\gamma \;=\; \lim_{n\to\infty} np_n \in\; (0,\infty)\,. \tag{2.2}$$

In the sequel we will encounter several GARCH-type processes, all of them indexed by $\theta \in [0,\infty)^4$. In discrete time, processes will be indexed additionally by $n \in \mathbb{N}$ and a suitable parametrisation. Throughout this paper a parametrisation is a pair $(\Theta, (H_n)_{n\in\mathbb{N}})$ where Θ is a nonempty subset of $[0,\infty)^4$ and, for all $n \in \mathbb{N}$, H_n is a mapping $H_n = (h_{0,n}, \beta_n, \alpha_n, \lambda_n) : \Theta \to [0,\infty)^4$. In this sense h_0 $(h_{0,n}(\theta))$ denotes the unknown initial value of the volatility h_0 which is contrived as an additional unknown parameter in this paper. For the corresponding continuous time limits, β/α and α are the mean level and the mean reversion parameter of the volatility processes, respectively; λ is a scaling parameter for the corresponding jumps of the volatility processes.

For a parametrization $(\Theta, (H_n)_{n\in\mathbb{N}})$ we consider the sequence of partial sums corresponding to a *randomly thinned* GARCH model, indexed by $\theta \in \Theta$ and $n \in \mathbb{N}$, defined by

$$\begin{aligned} G_n(k) &= G_n(k-1) + h_n^{1/2}(k-1)\, Z_{n,k}\,, \quad G_n(0) \;=\; 0\,, \\ h_n(k) &= \beta_n(\theta) + \alpha_n(\theta) h_n(k-1) + \lambda_n(\theta)\, h_n(k-1)\, Z_{n,k}^2\,, \\ h_n(0) &= h_{0,n}(\theta)\,, \qquad 1 \le k \le n\,, \quad \theta \in \Theta\,, \end{aligned} \tag{2.3}$$

where $H_n(\theta) = (h_{0,n}(\theta), \beta_n(\theta), \alpha_n(\theta), \lambda_n(\theta))$ for all $\theta \in \Theta$. Note that the specification of a GARCH does not quite follow the traditional one, but enumerating the indices generates the same processes. Also, observe that the definition of (G_n, h_n) in (2.3) depends on the choice of $(\Theta, (H_n)_{n\in\mathbb{N}})$.

The limit in (2.2) sets up convergence in distribution of $\sum_{k=1}^{[n\cdot]} Z_{n,k}$ to a compound Poisson process with rate $\gamma \in (0,\infty)$ and jump distribution Q as $n \to \infty$. For a choice of $(\Theta, (H_n)_{n\in\mathbb{N}})$ it is, thus, natural to ask whether the limit of $(G_n([nt]), h_n([nt]))_{0\le t\le 1}$ in distribution exists along $H_n(\theta)$ as $n \to \infty$ for fixed $\theta \in \Theta$. In [9] and [14] such parametrisations have been successfully constructed. Moreover, the corresponding continuous time limit equals COGARCH driven by a compound Poisson process.

COGARCH is a process $(G,h) = (G(t), h(t))_{0\le t\le 1}$ that is indexed by $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$ and determined as the unique pathwise solution of the system of the following integral equations:

$$
\begin{aligned}
(2.4)\quad G(t) &= \int_{[0,t]\times\mathbb{R}} h^{1/2}(s-)z\, N(ds,dz)\,,\\
h(t) &= h_0 + \int_{[0,t]} \beta - \alpha h(s-)\, ds + \lambda \int_{[0,t]\times\mathbb{R}} h(s-)z^2\, N(ds,dz)\,,
\end{aligned}
$$

where N is a Poisson point measure on $[0,1]\times\mathbb{R}$ with an intensity $\gamma\ell\otimes Q$.

In the sequel we restrict or analysis to the following two sampling schemes:

- *incomplete observations:* only G and G_n ($n \in \mathbb{N}$) are observable in full whereas the corresponding volatility processes h and h_n ($n \in \mathbb{N}$) are unobservable.
- *complete observations:* both processes (G,h) and (G_n,h_n) are observable in full.

We are dealing with both sampling schemes in the separate Subsections 2.2–2.3 and Subsection 2.4, respectively. Not surprisingly, a simpler theory is in place in case of complete observations. In the more realistic scenario, where observations of the volatility processes are not available, results are more difficult due to the nonlinearity of (CO)GARCH.

Throughout the whole paper, the space of right-continuous functions $g : [0,1] \to \mathbb{R}^d$ with left limits on $[0,1]$ is denoted by D_d. We endow D_d with the σ-algebra $\mathcal{D}_d$, generated by the point evaluations [cf. Billingsley [1]]. Furthermore, let $\mathbb{M}_d$ be the space of all nonnegative point measures on $[0,1]\times\mathbb{R}^d$ with finite support. We equip this space with the σ-algebra $\mathcal{M}_d$ generated by the point evaluations [cf. Reiss (1993), [20], pages 5–6].

The trace of the Borel field in $\overline{\mathbb{R}}^d = (\mathbb{R}\cup\{-\infty,\infty\})^d$ with respect to $A \subseteq \overline{\mathbb{R}}^d$ is denoted by $\mathcal{B}(A)$. The Lebesgue measure on $\mathcal{B}(\mathbb{R})$ and the Dirac measure with total mass in x are denoted by ℓ and ϵ_x, respectively. If $(E,\mathcal{A})$ is a measurable space and X is a random element taking values in $(E,\mathcal{A})$

then its distribution is denoted by $\mathcal{L}(X)$. Whenever this distribution depends on a parameter θ we employ the notation $\mathcal{L}_\theta(X)$. If $(E_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces and $X : E_1 \to E_2$ is $\mathcal{A}_1/\mathcal{A}_2$ measurable then μ^X denotes the image of a measure μ under X.

2.2. *Limit experiments of GARCH (incomplete observations).* In this subsection we assume that the volatility processes are unobservable. To pursue our programme we introduce another class of processes. Therefore let $(\widehat{G}, \widehat{h}) = (\widehat{G}(t), \widehat{h}(t))_{0 \le t \le 1}$ be the unique pathwise solution of the following system of integral equations:

$$
\begin{aligned}
(2.5)\quad \widehat{G}(t) &= \int_{[0,t]\times\mathbb{R}} \widehat{h}^{1/2}(s-) z \, N(ds, dz) \,, \\
\widehat{h}(t) &= h_0 + \int_{[0,t]} \beta - \alpha \widehat{h}(s-) \, dT_N \, (s) + \lambda \int_{[0,t]\times\mathbb{R}} \widehat{h}(s-) z^2 \, N(ds, dz) \,,
\end{aligned}
$$

where $\theta = (h_0, \beta, \alpha, \lambda) \in [0, \infty)^4$. Here $T : \mathbb{M}_1 \to D_1$, $\sigma \mapsto T_\sigma$ is defined as follows: if, for some $m \in \mathbb{N}$, $0 = t_0 < t_1 < ... < t_m < 1$ and $x_1, \dots, x_m \in \mathbb{R}$, $\sigma \in \mathbb{M}$ admits a representation of form $\sigma = \sum_{k=1}^m \varepsilon_{(t_k, x_k)}$ where $0 = t_0 < t_1 < ... < t_m < 1$, then we set

$$
\begin{aligned}
(2.6)\quad T_\sigma(t) &= \frac{t - t_k}{m(t_k - t_{k-1})} + \frac{k}{m} \,, \qquad t \in [t_{k-1}, t_k) \,,\ \ 1 \le k \le m \,, \\
T_\sigma(t) &= \frac{t - t_m}{m(t_m - t_{m-1})} + 1 \,, \qquad t \in [t_m, 1] \,.
\end{aligned}
$$

If such a representation does not exist, then we set $T_\sigma(t) = t$ for all $t \in [0, 1]$.

Let us call $(\widehat{G}, \widehat{h})$ the MCOGARCH, an acronym referring to *Modified* COGARCH. To illustrate the difference between COGARCH and MCOGARCH, we consider a simpler representation of $\widehat{G}$ next. (We will return to (2.5) in our analysis in Subsection 2.4.)

To this end, let $\nu = (\nu(t))_{0 \le t \le 1}$ be a Poisson process with rate $\gamma \in (0, \infty)$ and $(Z_k)_{k \in \mathbb{N}}$ be a sequence of independent random variables, independent of ν. By solving the integral equations for $\hat{h}$ in (2.5) we observe that

$$
(2.7)\qquad \mathcal{L}_\theta(\widehat{G}) = \mathcal{L}_\theta \left(\sum_{k=1}^{\nu(\cdot)} \widehat{h}^{1/2}_{\nu(1),k,\theta} \, Z_k \right) ,
$$

where, for $k, m \in \mathbb{N}$, $k \ge 2$, we set

$$
\begin{aligned}
(2.8)\qquad \widehat{h}_{m,k,\theta} &= \frac{\beta}{\alpha} \left(1 - e^{-\alpha/m} \right) + e^{-\alpha/m} \widehat{h}_{m,k-1} [1 + \lambda \, Z_{k-1}^2] \,, \\
\widehat{h}_{m,1,\theta} &= \frac{\beta}{\alpha} \left(1 - e^{-\alpha/m} \right) + e^{-\alpha/m} h_0 \,,
\end{aligned}
$$

for $\theta = (h_0, \beta, \alpha, \lambda) \in [0, \infty)^4$, $\alpha > 0$, with the convention $\sum_\emptyset = 0$. Here we extend the definition of $\widehat{h}_{m,k,\theta}$ to $\theta = (h_0, \beta, 0, \lambda) \in [0, \infty)^4$ by taking $\alpha \downarrow 0$ in (2.8).

In view of (2.8) note that the magnitudes of the jumps of $\widehat{G}$ (in space) depend on their multiplicity and the innovations (Z_n), but not on their arrival times. This attribute is not shared by COGARCH. To some extend it is, thus, justified to speak of $\widehat{G}$ and G as experiments driven by two and three sources of randomness, respectively: the number of jumps, the innovations, and the arrival times.

As no information about the volatility processes is assumed in this subsection we consider the following experiment of MCOGARCH type:

$$\widehat{\mathcal{E}} \;=\; \big(D_1, \mathcal{D}_1, \big(\mathcal{L}_\theta(\widehat{G})\big)_{\theta \in [0,\infty)^4}\big)\,. \tag{2.9}$$

For a parametrisation $(\Theta, (H_n)_{n\in\mathbb{N}})$ we consider the corresponding GARCH experiments in discrete time by

$$\mathcal{E}_{n,H_n}(\Theta) \;=\; \big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \big(\mathcal{L}_\theta(G_n)\big)_{\theta\in\Theta}\big)\,, \qquad n \in \mathbb{N}\,, \tag{2.10}$$

where, for $n \in \mathbb{N}$, $G_n = (G_n(k))_{1\le k\le n}$ is defined by (2.3) via the parametrisation $(\Theta, (H_n)_{n\in\mathbb{N}})$. We write $\mathcal{E}_{n,H_n} = \mathcal{E}_{n,H_n}(\Theta)$, provided we have $\Theta = [0, \infty)^4$ in (2.10).

Next we give a GARCH parametrisation such that the randomly thinned GARCH converges strongly to the MCOGARCH experiment $\widehat{\mathcal{E}}$ in deficiency: therefore pick $\theta = (h_0, \beta, \alpha, \lambda) \in \Theta$ and $n \in \mathbb{N}$. If $\alpha > 0$ then we set

$$\begin{aligned} h_{0,n}^{(0)}(\theta) \;&=\; h_0 e^{-\alpha/n} + \frac{\beta}{\alpha}(1-e^{-\alpha/n})\,, & \beta_n^{(0)}(\theta) \;&=\; \frac{\beta}{\alpha}\,(1-e^{-\alpha/n})\,,\\ \alpha_n^{(0)}(\theta) \;&=\; e^{-\alpha/n}\,, & \lambda_n^{(0)}(\theta) \;&=\; \lambda e^{-\alpha/n}\,, \end{aligned} \tag{2.11}$$

and, otherwise, if $\alpha = 0$ then we set

$$h_{0,n}^{(0)}(\theta) = h_0 + \frac{\beta}{n}\,, \quad \beta_n^{(0)}(\theta) \;=\; \frac{\beta}{n}\,, \qquad \alpha_n^{(0)}(\theta) \;=\; 1\,, \quad \lambda_n^{(0)}(\theta) = \lambda\,. \tag{2.12}$$

Let $([0,\infty)^4, (H_n^{(0)}))$ be the corresponding parametrization and $(G_n^{(0)})$ be the corresponding partial sum processes of GARCH in (2.3).

Although the parametrisation in (2.11)–(2.12) is quite elaborated, we show that the corresponding GARCH experiments converges to the experiment of MCOGARCH-type, with no restrictions on Q assumed [cf. Subsection 4.1 for a proof].

THEOREM 2.1. *If (2.2) is in place for some $\gamma \in (0,\infty)$ and $p_n \in (0,1)$, $n \in \mathbb{N}$, then $\mathcal{E}_{n,H_n^{(0)}}$ converges strongly to $\widehat{\mathcal{E}}$ in deficiency as $n \to \infty$.*

If Q is absolutely continuous with respect to the Lebesgue measure, then Theorem 2.1 extends partially to other GARCH parametrisations [cf. Section 4.3 for a proof of the following theorem].

THEOREM 2.2. *Suppose that* (2.2) *holds for some* $\gamma \in (0,\infty)$ *and* $p_n \in (0,1)$, $n \in \mathbb{N}$, *and* $Q \ll \ell$. *Let* $\Theta \neq \emptyset$ *with compact closure* $\overline{\Theta}$ *in* $(0,\infty) \times [0,\infty)^3$. *For* $n \in \mathbb{N}$, *let* $H_n = (h_{0,n}, \beta_n, \alpha_n, \lambda_n) : \Theta \to [0,\infty)^4$ *be a GARCH parametrisation and* G_n *be the corresponding GARCH model in* (2.3).

If there exist $n_0 \in \mathbb{N}$ *and* $C > 0$ *such that, for all* $n \geq n_0$, *both*

$$\sup_{\theta=(h_0,\beta,\alpha,\lambda)\in\Theta} \max\left\{ \left|h_{0,n}(\theta) - h_0\right|, \ \left|\lambda_n(\theta) - \lambda\right| \right\} \leq \frac{C}{n}, \tag{2.13}$$

and

$$\sup_{\theta=(h_0,\beta,\alpha,\lambda)\in\Theta} \max\left\{ \left|n\beta_n(\theta) - \beta\right|, \ \left|n(\alpha_n(\theta) - 1) + \alpha\right| \right\} \leq C, \tag{2.14}$$

then

$$\lim_{n\to\infty} \sup_{\theta\in\Theta} \left\| \mathcal{L}_\theta(G_n) - \mathcal{L}_\theta(G_n^{(0)}) \right\| = 0, \tag{2.15}$$

and $\mathcal{E}_{n,H_n}(\Theta)$ *converges strongly to* $\widehat{\mathcal{E}}(\Theta)$ *in deficiency as* $n \to \infty$.

In Kallsen and Vesenmayer [9] and Maller et al. [14] the following GARCH parametrisations $(\Theta, (H_n^{(KV)})_{n\in\mathbb{N}})$ and $(\Theta, (H_n^{(M)})_{n\in\mathbb{N}})$ have been considered where, for $\theta = (h_0, \beta, \alpha, \lambda) \in (0,\infty)^3 \times [0,\infty)$, in obvious notation, $\Theta = (0,\infty)^3 \times [0,\infty)$ and

$$\begin{aligned}
h_{0,n}^{(KV)}(\theta) &= h_{0,n}^{(M)}(\theta) = h_0, \\
\beta_n^{(KV)}(\theta) &= \beta_n^{(M)}(\theta) = \frac{\beta}{n}, \\
\alpha_n^{(KV)}(\theta) &= \alpha_n^{(M)}(\theta) = e^{-\alpha/n}, \\
\lambda_n^{(KV)}(\theta) &= \lambda, \quad \lambda_n^{(M)}(\theta) = e^{-\alpha/n}\lambda.
\end{aligned} \tag{2.16}$$

Kallsen and Vesenmayer [9] have shown that $(G_n[n\cdot], h_n[n\cdot])$, as defined in (2.3) by $H_n(\theta) = H_n^{(KV)}(\theta)$, converge to COGARCH with parameter θ in (2.4) in law with respect to the Skorokhod topology, as $n \to \infty$, for all $\theta \in \Theta$.

Maller et al. [14] have encountered a slightly different scenario. For $\theta \in \Theta$ they have embedded a sequence of GARCH models into a given COGARCH and obtained the convergence with respect to the same topology, now driven by a general Lévy process, even in probability. If the driving process is a compound Poisson process with rate $\gamma > 0$ and jump size distribution Q then their analysis comprises a situation where the corresponding partial sums have the same law as $(G_n[n\cdot], h_n[n\cdot])$ under the parametrisation $H_n^{(M)}(\theta)$, $\theta \in \Theta$, $n \in \mathbb{N}$.

In short, it follows from the analysis in [9] and [14] that the partial sum processes of GARCH converge to COGARCH with parameter θ in law along both parametrisations, along $H_n^{(KV)}(\theta)$ and $H_n^{(M)}(\theta)$, respectively, as $n \to \infty$, with respect to the Skorokhod topology, for all $\theta \in \Theta$. On the other hand, both parametrisations fall into the framework of Theorem 2.2. Hence, if the distribution of the innovations admits a Lebesgue density the limiting experiment is given by MCOGARCH $\widehat{\mathcal{E}}(\Theta)$ rather than COGARCH $\mathcal{E}(\Theta)$.

2.3. *COGARCH vs. MCOGARCH (incomplete observations).* In this subsection we investigate whether the experiments induced by GARCH and MOCOGARCH are of the same type. Here we again assume that the volatility processes are unobservable. Therefore recall (2.4) and consider the experiment

$$\mathcal{E} \;=\; \left(D_1, \mathcal{D}_1, \left(\mathcal{L}_\theta(G)\right)_{\theta\in[0,\infty)^4}\right). \tag{2.17}$$

Note that both experiments $\mathcal{E}$ and $\widehat{\mathcal{E}}$ depend on the intensity measure $\gamma\ell\otimes Q$ which enters (2.4)– (2.5) via N. In this subsection we include this dependence into our notation by writing $\mathcal{E}_{\gamma,Q}$ and $\widehat{\mathcal{E}}_{\gamma,Q}$ instead of $\mathcal{E}$ and $\widehat{\mathcal{E}}$, respectively. Next we are concerned how COGARCH relates to MCOGARCH in deficiency [cf. Subsection 4.4 for a proof].

THEOREM 2.3. *Let $\emptyset \neq \Theta \subseteq (0,\infty)\times[0,\infty)^3$ and Q be the standard normal distribution.*

Suppose that $\mathcal{E}_{\gamma,Q}(\Theta)$ is equivalent to $\widehat{\mathcal{E}}_{\gamma,Q}(\Theta)$ for all $\gamma > 0$. Then we have:

(i) If $(h_{0,1},\beta,\alpha,\lambda),(h_{0,2},\beta,\alpha,\lambda) \in \Theta$ then $h_{0,1} = h_{0,2}$.

(ii) If $(h_0,\beta_1,\alpha,\lambda),(h_0,\beta_2,\alpha,\lambda) \in \Theta$ then $\beta_1 = \beta_2$.

(iii) If $(h_0,\beta,\alpha_1,\lambda),(h_0,\beta,\alpha_2,\lambda) \in \Theta$ then $\alpha_1 = \alpha_2$.

Theorem 2.3 indicates that equivalence of MCOGARCH and COGARCH is restricted to parameter sets that are of considerably lower dimensions and have nonempty interior. Hence, for Gaussian innovations we have generically no equivalence in deficiency. A generalization of Theorem 2.3 to other distributions seems to be not obvious. However, there is strong evidence that

Table 1
Choices of θ_0 and $\theta = \theta_{ij}$ in Equation (2.18).

θ_0	2	1	1	0.1
θ_{11}	0.4	1	1	0.1
θ_{12}	10	1	1	0.1
θ_{21}	2	0.2	1	0.1
θ_{22}	2	5	1	0.1
θ_{31}	2	1	0.2	0.1
θ_{32}	2	1	5	0.1
θ_{41}	2	1	1	0.02
θ_{42}	2	1	1	0.5

the non-equivalence of COGARCH and MCOGARCH holds in more general situations.

Therefore recall that statistical equivalence of the experiments $\mathcal{E}$ and $\widehat{\mathcal{E}}$ is implied [cf. [19], Theorem 53.10] when for all finite subsets $\Theta \subseteq [0,\infty)^4$ and all $\theta_0 \in \Theta$ we have

$$\mathcal{L}_{\theta_0}\left(\left(\frac{d\mathcal{L}_\theta(G)}{d\mathcal{L}_{\theta_0}(G)}\right)_{\theta\in\Theta}\right) \;=\; \mathcal{L}_{\theta_0}\left(\left(\frac{d\mathcal{L}_\theta(\widehat{G})}{d\mathcal{L}_{\theta_0}(\widehat{G})}\right)_{\theta\in\Theta}\right). \tag{2.18}$$

We generated samples from these two distributions according to the recursion (4.37) in the proof of Theorem 2.3 in Subsection 4.4. To this end, we first restricted the parameter space to a set with two elements, θ_0 and θ. While fixing θ_0 to $(2,1,1,0.1)$, we have chosen eight vectors θ_{ij}, $i = 1,\dots,4$, $j = 1,2$, for the parameter vector θ, which differ from θ_0 in only one component, cf. Table 1. Secondly, we checked the distributional equality (2.18) for three different jump distributions: the standard normal one (for comparison), the standard Cauchy distribution Cauchy$(0,1)$, and the normal mixture distribution

$$\frac{1}{2}N(-0.5, 0.75) + \frac{1}{2}N(0.5, 0.75),$$

which has mean 0 and variance 1. The intensity γ was always fixed to 4.

For each of the eight pairs (θ_0, θ_{ij}) and each of the three jump distributions, we generated 10^6 samples of the two distributions referring to the COGARCH and MCOGARCH in Equation (2.18). Table 2 reports in the left column the choice of θ_{ij}, whereas the other three columns report, for each of the three jump distributions, the 25% quantile, the median, and the 75% quantile of the distribution in Equation (2.18).

Next, we applied the Wilcoxon rank sum test (also known as Mann-Whitney test) to investigate the null hypothesis *the median of the likelihood*

TABLE 2

Estimated 25% quantiles, medians, and 75% quantiles for the distributions in (2.18).

jumps quantiles	$N(0,1)$ 25% median 75%	Cauchy$(0,1)$ 25% median 75%	mixed N 25% median 75%
	COGARCH MCOGARCH	COGARCH MCOGARCH	COGARCH MCOGARCH
θ_{11}	0.1081 0.5560 1.3888 0.1785 0.6977 1.3495	0.5521 0.7775 1.1767 0.5884 0.8226 1.1811	0.0909 0.5329 1.3918 0.1558 0.6743 1.3543
θ_{12}	0.1505 0.3152 0.6449 0.1637 0.3377 0.6768	0.4173 0.8127 1.4573 0.4412 0.8335 1.4505	0.1436 0.3008 0.6136 0.1575 0.3264 0.6559
θ_{21}	0.8326 1.0168 1.1711 0.7605 1.0114 1.2459	0.9273 0.9761 1.0393 0.9051 0.9566 1.0539	0.8307 1.0201 1.1766 0.7560 1.0155 1.2512
θ_{22}	0.4883 0.7071 1.0086 0.4201 0.6077 1.0000	0.7765 1.0229 1.2130 0.7010 1.0247 1.2676	0.4797 0.6956 1.0000 0.4100 0.5988 0.9798
θ_{31}	0.6928 0.8543 1.0621 0.6304 0.7841 1.0629	0.8497 1.0000 1.1506 0.8029 1.0000 1.1881	0.6863 0.8476 1.0530 0.6248 0.7757 1.0524
θ_{32}	0.0053 0.1702 1.1056 0.0010 0.0590 0.9129	0.3853 0.6449 1.1172 0.3093 0.5650 1.1090	0.0028 0.1392 1.0856 0.0005 0.0437 0.8703
θ_{41}	0.9864 1.0104 1.0735 0.9884 1.0100 1.0693	0.8265 1.0000 1.0798 0.8357 1.0000 1.0779	0.9863 1.0114 1.0762 0.9884 1.0109 1.0722
θ_{42}	0.6851 0.8870 1.0000 0.6963 0.8942 1.0000	0.6217 0.9328 1.0418 0.6281 0.9360 1.0388	0.6750 0.8802 1.0000 0.6865 0.8874 1.0000

ratio for the COGARCH experiment equals the median of the likelihood ratio for the MCOGARCH experiment. Table 3 reports the values of the Wilcoxon test statistic W, together with the corresponding p-values. For each jump distribution, the first column corresponds to a sample size of 10^4, the second row to 10^5, and the third column to a sample size of 10^6 per experiment. Obviously, the p-values tend to 0 as the sample size increases. Based on 10^6 samples, the null hypothesis is most significantly rejected, for all three jump distributions and for all eight parameter vectors θ_{ij}. In other words, there is strong evidence that, in the case of uncomplete observations, the randomly thinned GARCH and the COGARCH experiment are not statistically equivalent for these jump distributions. This confirms our conjecture, that Theorem 2.3 holds in a much more general formulation for quite arbitrary jump distributions.

2.4. *Complete Observations.* In the last subsections we have investigated both convergence and equivalence in deficiency of a variety of GARCH-type experiments under the assumption that their volatility processes h_n, h and $\widehat{h}$ are unobservable. In this subsection we are dealing with the less realistic situation where the corresponding volatility processes are observable in full, though. To this end, consider the following GARCH-type experiments in

Table 3
Wilcoxon rank sum test: Values of Wilcoxon test statistic W and corresonding p-values.

jumps sample size	$N(0,1)$ 10^4 10^5 10^6	Cauchy$(0,1)$ 10^4 10^5 10^6	mixed N 10^4 10^5 10^6
	W statistic p-value	W statistic p-value	W statistic p-value
θ_{11}	-7.10 -24.12 -73.91 0.0000 0.0000 0.0000	-8.82 -25.46 -73.81 0.0000 0.0000 0.0000	-8.11 -24.01 -71.91 0.0000 0.0000 0.0000
θ_{12}	-3.04 -14.90 -47.21 0.0024 0.0000 0.0000	-0.35 -2.73 -9.90 0.7245 0.0064 0.0000	-6.04 -15.37 -48.40 0.0000 0.0000 0.0000
θ_{21}	-1.56 -3.20 -12.52 0.1189 0.0014 0.0000	8.71 31.90 98.90 0.0000 0.0000 0.0000	-0.45 -3.45 -14.05 0.6545 0.0006 0.0000
θ_{22}	12.10 44.13 136.09 0.0000 0.0000 0.0000	-1.92 -2.69 -8.28 0.0546 0.0070 0.0000	14.17 45.16 141.15 0.0000 0.0000 0.0000
θ_{31}	12.38 37.96 116.09 0.0000 0.0000 0.0000	1.76 2.16 10.30 0.0788 0.0311 0.0000	12.07 38.89 119.95 0.0000 0.0000 0.0000
θ_{32}	11.34 39.48 126.96 0.0000 0.0000 0.0000	11.63 33.04 100.34 0.0000 0.0000 0.0000	13.59 42.66 131.75 0.0000 0.0000 0.0000
θ_{41}	0.83 1.41 5.79 0.4054 0.1572 0.0000	-1.52 -2.85 -4.98 0.1280 0.0044 0.0000	2.83 3.64 5.35 0.0047 0.0003 0.0000
θ_{42}	-2.96 -3.71 -13.73 0.0031 0.0002 0.0000	-1.29 -2.35 -2.94 0.1963 0.0189 0.0032	-1.63 -4.81 -15.26 0.1041 0.0000 0.0000

continuous time with fully observed volatilities by

$$\mathcal{E}_h = \left(D_2, \mathcal{D}_2, \left(\mathcal{L}_\theta(G,h)\right)_{\theta\in[0,\infty)^4}\right), \quad \widehat{\mathcal{E}}_h = \left(D_2, \mathcal{D}_2, \left(\mathcal{L}_\theta(\widehat{G},\widehat{h})\right)_{\theta\in[0,\infty)^4}\right),$$

where $\widehat{h}$ is defined by the specification in (2.5) and (2.6). Similar to Subsections 2.1–2.2, where we dealt with the continuous time, both experiments $\mathcal{E}_h$ and $\widehat{\mathcal{E}}_h$ depend upon Q and $\gamma > 0$ as well. In this subsection we will suppress this dependence in our notations.

We need to specify a set $\Theta_e \subseteq [0,\infty)^4$ of *exceptional points* in the parameter space $[0,\infty)^4$ by

$$\Theta_e = \{\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4 : \ h_0\alpha = \beta\}. \tag{2.19}$$

Observe that Θ_e is closely connected to the fixpoint of the affine differential equation $h'(t) = \beta - \alpha h(t)$. Indeed, if $\theta = (h_0, \beta, \alpha, \lambda) \in \Theta_e$ then we have $h(t) = \widehat{h}(t) \equiv h_0$ for all $t \in [0,T)$ where T is the first jump of (M)COGARCH. It is impossible to recover the parameters β, α, λ in full within the time horizon $[0,T)$. Otherwise, if h_0 is not the fixpoint of this differential equation then it is always possible to recover parts of θ by taking appropriate derivatives. In the next proposition we formalize this idea and show that both $\mathcal{E}_h$ and $\widehat{\mathcal{E}}_h$ are equivalent to a simple reference experiment [cf. Subsection 4.5.1 for a formal proof].

Proposition 2.1. *If $Q(\{0\})=0$ then both $\mathcal{E}_h$ and $\widehat{\mathcal{E}}_h$ are equivalent to $\mathcal{F}=([0,\infty]^4,\mathcal{B}([0,\infty]^4),(Q_\theta)_{\theta\in[0,\infty)^4})$ where, for $\theta=(h_0,\beta,\alpha,\lambda)\in[0,\infty)^4$, $\gamma>0$, we set*

(2.20)

$$Q_\theta \;=\; \begin{cases} e^{-\gamma}\,\varepsilon_{(h_0,\beta,\alpha,\infty)} + (1-e^{-\gamma})\,\varepsilon_\theta\,, & \theta\notin\Theta_e\,,\\ e^{-\gamma}\,\varepsilon_{(h_0,\infty,\infty,\infty)} + (1-e^{-\gamma})\,\varepsilon_\theta\,, & \theta\in\Theta_e\,,\\ & h_0>0\,,\ \lambda>0\,,\\ e^{-\gamma}\,\varepsilon_{(h_0,\infty,\infty,\infty)} + (1-e^{-\gamma})\,\varepsilon_{(h_0,\infty,\infty,0)}\,, & \theta\in\Theta_e\,,\\ & h_0>0\,,\ \lambda=0\,,\\ \varepsilon_{(0,\infty,\infty,\infty)}\,, & \theta\in\Theta_e\,,\ h_0=0\,, \end{cases}$$

and Θ_e is the set as defined in (2.19).

Remark 2.1. In the situation of Proposition 2.1 we require Q to satisfy $Q(\{0\})=0$. Indeed, if $Q=\varepsilon_0$ then it is easy to see that both $\mathcal{E}_h$ and $\widehat{\mathcal{E}}_h$ are equivalent to $\mathcal{F}$ where we formally set $\gamma=0$ in (2.20). Otherwise, if $Q(\{0\})\in[0,1)$ then we may adjust the intensity measures of the driving Poisson measure accordingly, to see that both $\mathcal{E}_h$ and $\widehat{\mathcal{E}}_h$ are equivalent to $\mathcal{F}$, but with γ replaced by $\gamma\,Q(\mathbb{R}\backslash\{0\})$ in the definition of Q_θ. Analogously, one can adjust the discrete-time experiments that we consider in Proposition 2.2. We leave the details to the reader. □

Next we investigate the discrete time experiments. Note that the initial value of h is observable in continuous time. As a result, it is always possible to recover the parameter h_0 in full. To account for this phenomenon in discrete time we shall introduce the following sequence of experiments $\mathcal{E}_{h,n,H_n}$, indexed by $n\in\mathbb{N}$, where we set

$$\mathcal{E}_{h,n,H_n} \;=\; \big([\mathbb{R}^{n+1}]^2,\mathcal{B}([\mathbb{R}^{n+1}]^2),\big(\mathcal{L}_\theta(G_n,h_n)\big)_{\theta\in[0,\infty)^4}\big)\,,\qquad n\in\mathbb{N}\,. \tag{2.21}$$

Here $([0,\infty)^4,(H_n))$ is a parametrisation of the full parameter space $[0,\infty)^4$; both $G_n=(G_{n,k})_{0\le k\le n}$ and $h_n=(h_{n,k})_{0\le k\le n}$ are defined by (2.3) via $H_n(\theta)=(h_{0,n}(\theta),\beta_n(\theta),\alpha_n(\theta),\lambda_n(\theta))$ for $n\in\mathbb{N}$ and $\theta\in[0,\infty)^4$ [by a slight abuse of the previous notations]. Now we are in the position to state an analogon of Proposition 2.1 in the discrete time [cf. Subsection 4.5.2 for a formal proof].

Proposition 2.2. *Suppose that (2.2) is satisfied for some $\gamma\in(0,\infty)$ and $p_n\in(0,1)$, $n\in\mathbb{N}$. Let $([0,\infty)^4,H_n)_{n\in\mathbb{N}}$ be the parametrisation in (2.11)–(2.12). Also, let $([0,\infty)^4,H_n^{(KV)})_{n\in\mathbb{N}}$ and $([0,\infty)^4,H_n^{(M)})_{n\in\mathbb{N}}$ be the parametrisations in (2.16), respectively.*

If $Q(\{0\})=0$ then the following assertions are in place as $n\to\infty$, both in deficiency:

(i) $\mathcal{E}_{h,n,H_n}$ *converges strongly to* $\mathcal{F}$.

(ii) Both $\mathcal{E}_{h,n,H_n^{(KV)}}$ *and* $\mathcal{E}_{h,n,H_n^{(M)}}$ *are asymptotically equivalent to* $\mathcal{F}_n = ([0,\infty]^4, \mathcal{B}([0,\infty]^4), (Q_{\theta,n})_{\theta\in[0,\infty)^4})$, *where for* $n\in\mathbb{N}$ *and* $\theta=(h_0,\beta,\alpha,\lambda)\in[0,\infty)^4$ *we define* $Q_{\theta,n}$ *as* Q_θ *in* (2.20), *but with* Θ_e *replaced by*

$$\Theta_{e,n} \;=\; \{\theta=(h_0,\beta,\alpha,\lambda)\in[0,\infty)^4:\; h_0 n\,(1-e^{-\alpha/n})=\beta\}\,.$$

Finally we are concerned with the relationships between the experiments $\mathcal{F}$ and $\mathcal{F}_n$, $n\in\mathbb{N}\cup\{\infty\}$ [cf. Subsection 4.5.3 for a formal proof].

PROPOSITION 2.3. *Let* $\gamma>0$ *and* $\emptyset\neq\Theta\subseteq[0,\infty)^4$. *Let* $\mathcal{F},\mathcal{F}_n$, $n\in\mathbb{N}$, *be the experiments in Propositions 2.1 and 2.2.*
Let $\widehat{\mathcal{F}} = ([0,\infty]^4, \mathcal{B}([0,\infty]^4), (\widehat{Q}_\theta)_{\theta\in[0,\infty)^4})$ *be the experiment where for* $\theta=(h_0,\beta,\alpha,\lambda)\in[0,\infty)^4$ *we define* $\widehat{Q}_\theta$ *as* Q_θ *in* (2.20), *but with* Θ_e *replaced by*

$$\widehat{\Theta}_e \;=\; \{0\}^2\times(0,\infty)\times[0,\infty)\cup[0,\infty)\times\{0\}^2\times[0,\infty)\,.$$

Then the following assertions hold:

(i) Always $\delta(\widehat{\mathcal{F}}(\Theta),\mathcal{F}(\Theta))=\delta(\widehat{\mathcal{F}}(\Theta),\mathcal{F}_n(\Theta))=0$ *for all* $n\in\mathbb{N}$.

(ii) Further, $\delta(\mathcal{F}(\Theta),\widehat{\mathcal{F}}(\Theta))=0$ *if and only if, for all* $h_0>0$,

$$\begin{aligned}(2.22)\quad &\Big\{(\beta,\alpha,\lambda)\in[0,\infty)^3:\;(h_0,\beta,\alpha,\lambda)\in\Theta\cap\Theta_e\cap\widehat{\Theta}_e^C\Big\}\neq\emptyset\\ &\Rightarrow\quad \#\Big\{(\beta\,,\alpha)\in[0,\infty)^2\;:\;\exists\lambda\geq0\;(h_0,\beta,\alpha,\lambda)\in\Theta_e\cap\Theta\Big\}=1\,.\end{aligned}$$

(iii) Further, $\lim_{n\to\infty}\delta(\mathcal{F}_n(\Theta),\widehat{\mathcal{F}}(\Theta))=0$ *if and only if there exists* n_0 *such that, for all* $n\geq n_0$ *and* $h_0>0$, (2.22) *is in place, but with* Θ_e *replaced by* $\Theta_{e,n}$. *In particular,* $\mathcal{F}_n$ *converges weakly to* $\widehat{\mathcal{F}}$ *as* $n\to\infty$ *in deficiency.*

Let us rephrase our results in terms of the GARCH experiments, with the volatility processes fully observed in both continuous and discrete time. In contrast to the situation in Theorem 2.3 it follows from Proposition 2.1 that the continuous-time experiments induced by (M)COGARCH are mutually equivalent in deficiency. Depending on the parametrisation, (M)COGARCH occurs also as the limit in deficiency of discrete-time GARCH, in particular, this is the case for the parametrisation in Proposition 2.2. In contrast to Theorem 2.3, for a large class of parameter sets Θ, all of these discrete-time experiments, i.e. $\mathcal{E}_{h,n,H_n^{(0)}}(\Theta)$, $\mathcal{E}_{h,n,H_n^{(KV)}}(\Theta)$, $\mathcal{E}_{h,n,H_n^{(M)}}(\Theta)$, are asymptotically equivalent to (M)COGARCH $\mathcal{E}_h(\Theta)$ and $\widehat{\mathcal{E}}_h(\Theta)$, in deficiency, as $n \to \infty$, for instance, this happens if $\Theta \subseteq [0,\infty)^4$ does not contain an open neighbourhood of Θ_e. Since the set Θ_e is of lower dimension than $[0,\infty)^4$ it is, thus, justified to say that the randomly thinned GARCH is *generically* equivalent to COGARCH in deficiency, as $n \to \infty$.

3. Conclusion. As was shown in Maller et al. [14] and Kallsen and Vesenmayer [9], for a suitable parametrisations the GARCH model converges to the COGARCH model in probability and in distribution, respectively. These papers are dealing with a general Lévy process as driving process of the COGARCH. In this paper we study an important special case in Le Cam's framework of statistical experiments, namely, we assume that the driving process of COGARCH is a compound Poisson process and the innovations of GARCH are randomly thinned. Then GARCH converges generically to COGARCH, even in deficiency, provided that the volatility processes are observed in full. Otherwise, GARCH might still have a limiting experiment, but this will usually not be equivalent to COGARCH in deficiency. In Le Cam's framework Wang [21] and Brown et al. [5] investigated GARCH and Nelson's diffusion limit. They deal with aggregated Gaussian innovations. It would be interesting to extend the analysis to more general Lévy processes, rather than Brownian motion and compound Poisson processes. However, this needs substantial investigations on the approximation and randomisations of Lévy processes themselves and seems far out of reach at the present stage of research. In any case, we believe that our analysis offers a first important step into this direction.

4. Proofs.

4.1. *Proof of Theorem 2.1.* For $n \in \mathbb{N}$ define a point measure $N_{1,n}$ on $[0,1] \times \mathbb{R}$ by

$$N_{1,n} = \sum_{k=1}^{n} 1_{Z_{n,k} \neq 0} \ \varepsilon_{(k/n\,,\,Z_{n,k})}\,, \quad n \in \mathbb{N}\,. \tag{4.1}$$

Using $N_{1,n}$ we pass from discrete to continuous time. For $n \in \mathbb{N}$ define $\mathcal{E}^\star_{1,n} = \{D_1, \mathcal{D}_1, (\mathcal{L}_\theta(G_{1,n})\}_{\theta\in[0,\infty)^4})$, where, for all $0 \le t \le 1$, $n \in \mathbb{N}$ and $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$, $(G_{1,n}, h_{1,n})$ is the unique pathwise solution of the following system of integral equations ($t \in [0,1]$):

$$
\begin{aligned}
G_{1,n}(t) &= \int_{[0,t]\times\mathbb{R}} h_{1,n}^{1/2}(s-)\, z\, N_{1,n}(ds,dz)\,, \\
h_{1,n}(t) &= h_0 + \int_{[0,t]} \beta - \alpha h_{1,n}(s-)\, ds \\
&\quad + \lambda \int_{[0,t]\times\mathbb{R}} h_{1,n}(s-) z^2\, N_{1,n}(ds,dz)\,.
\end{aligned} \tag{4.2}
$$

Fix $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$ with $\alpha \neq 0$. By solving the linear ode for $h_{1,n}$ in (4.2) observe that

$$
h_{1,n}(t) = \frac{\beta}{\alpha}\left[1 - e^{-\alpha[t-(k-1)/n]}\right] + e^{-\alpha[t-(k-1)/n]}\, h_{1,n}\left(\frac{k-1}{n}\right), \tag{4.3}
$$

for $(k-1)/n \le t < k/n$, $1 \le k \le n$ and $n \in \mathbb{N}$. It, thus, follows from (2.11) and (4.3) that, for all $n \in \mathbb{N}$,

$$
\begin{aligned}
h_{1,n}(1/n-) &= h_0 e^{-\alpha/n} + \frac{\beta}{\alpha}[1 - e^{-\alpha/n}] = h_{0,n}(\theta)\,. \\
h_{1,n}\left(\frac{k}{n}-\right) &= \beta_n(\theta) + h_{1,n}\left(\frac{k-1}{n}-\right)\left[\alpha_n(\theta) + \lambda_n(\theta)\, Z_{n,k-1}^2\right],\ 2 \le k \le n\,.
\end{aligned}
$$

In view of (2.3) and the identities in the last display, we, thus, have $h_n(k) = h_{1,n}\left(((k+1)/n)-\right)$ for all $n \in \mathbb{N}$, $0 \le k \le n-1$ and $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$ with $\alpha > 0$. A similar argument is applicable to (2.12) and $\theta = (h_0, \beta, 0, \lambda) \in [0,\infty)^4$. It, thus, follows from (2.3) and (4.2) that

$$
\mathcal{L}_\theta\big((G_{1,n}(k/n))_{1\le k\le n}\big) = \mathcal{L}_\theta\big((G_n(k))_{1\le k\le n}\big)\,, \quad n \in \mathbb{N}\,,\ \theta \in [0,\infty)^4\,.
$$

Note that $G_{1,n}$ is constant on $[(k-1)/n, k/n)$, $1 \le k \le n$ and $n \in \mathbb{N}$. Hence $\mathcal{E}_{n,H_n^{(0)}}$ is equivalent to $\mathcal{E}_{1,n}^\star$ in deficiency for all $n \in \mathbb{N}$ by (A.2) and the monotonicity theorem for Markov kernels [cf. [20], Lemma 1.4.2(i)].

Next we randomize the deterministic time grid. Therefore let $(U_k)_{k\in\mathbb{N}}$ be an iid sequence of random variables independent of the vector Z_n, where U_k is uniformly distributed on $[0,1]$. Set

$$
V_{n,k} = ((k-1)+U_k)/n\,, \qquad 1 \le k \le n \tag{4.4}
$$

and define a point process $N_{2,n}$ by

$$
N_{2,n} = \sum_{k=1}^{n} 1_{Z_{n,k}\neq 0}\, \varepsilon_{(V_{n,k}, Z_{n,k})}\,, \qquad n \in \mathbb{N}\,. \tag{4.5}
$$

Let T be as in (2.6). For $n \in \mathbb{N}$ let $\mathcal{E}^\star_{2,n} = (D_1, \mathcal{D}_1, (\mathcal{L}_\theta(G_{2,n}))_{\theta\in[0,\infty)^4})$, where, for all $0 \le t \le 1$, $n \in \mathbb{N}$ and $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$, $(G_{2,n}, h_{2,n})$ is the pathwise unique solution of the following system of integral equations:

$$
\begin{aligned}
G_{2,n}(t) &= \int_{[0,t]\times\mathbb{R}} h_{2,n}^{1/2}(s-) z \, N_{2,n}(ds, dz)\,, \\
h_{2,n}(t) &= h_0 + \int_{[0,t]} \beta - \alpha h_{2,n}(s-)\, dT_{N_{2,n}}\,(s) \\
&\quad + \lambda \int_{[0,t]\times\mathbb{R}} h_{2,n}(s-) z^2 \, N_{2,n}(ds, dz)\,.
\end{aligned}
\tag{4.6}
$$

To proceed we need the following lemma [cf. Subsection 4.2 for a proof].

LEMMA 4.1. *Let N be a Poisson measure with intensity measure $\gamma\ell \otimes Q$ and $N_{2,n}$ as in (4.5). If (2.2) holds then* $\lim_{n\to\infty} \|\mathcal{L}(N_{2,n}) - \mathcal{L}(N)\| = 0$.

Let N be a Poisson measure with intensity measure $\gamma\ell\otimes Q$. It follows from (2.5) and (4.6) that there exists a family of deterministic Markov kernels $K_\theta : \mathbb{M}_1 \times \mathcal{D}_1 \to [0,1]$, indexed by $\theta \in [0,\infty)^4$, such that both $\mathcal{L}_\theta(G_{2,n}) = K_\theta \mathcal{L}(N_{2,n})$ and $\mathcal{L}_\theta(\widehat{G}) = K_\theta \mathcal{L}(N)$ for all $n \in \mathbb{N}$ and $\theta \in \Theta$. Since we assumed (2.2) the assertion of Lemma 4.1 is in place, we, thus, get from (A.4) and the monotonicity theorem for Markov kernels (cf. [20], Lemma 1.4.2(i)) that as $n \to \infty$,

$$
\Delta(\widehat{\mathcal{E}}, \mathcal{E}^\star_{2,n}) \le \sup_{\theta\in[0,\infty)^4} \|\mathcal{L}_\theta(\widehat{G}) - \mathcal{L}_\theta(G_{2,n})\| \le \|\mathcal{L}(N) - \mathcal{L}(N_{2,n})\| \to 0\,.
$$

Consequently, $\mathcal{E}^\star_{2,n}$ converges (strongly) to $\widehat{\mathcal{E}}$ in deficiency as $n \to \infty$. Recall that $\mathcal{E}_{n,H_n}$ is equivalent to $\mathcal{E}^\star_{1,n}$ in deficiency for all $n \in \mathbb{N}$. To complete the proof of the theorem it, thus, suffices to show that $\mathcal{E}^\star_{1,n}$ is equivalent to $\mathcal{E}^\star_{2,n}$.

Therefore let $\mathbb{M}_0$ be the space of all nonnegative point measures on $[0,1]$ with finite support. We equip this space with the σ-algebra $\mathcal{M}_0$ generated by the point evaluations [cf. Reiss (1993), [20], pages 5–6]. Let $\mathbb{M}_{0,1} \subseteq \mathbb{M}_0$ be the subset of point measures $\sigma \in \mathbb{M}_0$ such there exist $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_m < 1$ with $\sigma = \sum_{k=1}^m \varepsilon_{t_k}$. For $\sigma \in \mathbb{M}_0$, we define mappings $T_{1,\sigma}, T_{2,\sigma} : [0,1] \to [0,\infty)$ and $T_{3,\sigma}, T_{4,\sigma} : [0,1]\times\mathbb{R} \to [0,\infty)\times\mathbb{R}$ as follows: if $\sigma \in \mathbb{M}_0\backslash\mathbb{M}_{0,1}$ then for all $t \in [0,1]$ and $x \in \mathbb{R}$, we set $T_{1,\sigma}(t) = T_{2,\sigma}(t) = t$ and $T_{3,\sigma}(t,x) = T_{4,\sigma}(t,x) = (t,x)$. Otherwise, if $\sigma \in \mathbb{M}_{0,1}$ then there exist $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_m < 1$ with $\sigma = \sum_{k=1}^m \varepsilon_{t_k}$ and we set

$$
\begin{aligned}
T_{1,\sigma}(t) &= \frac{t - t_k}{m(t_k - t_{k-1})} + \frac{k}{m}\,, \qquad t \in [t_{k-1}, t_k)\,,\ \ 1 \le k \le m\,, \\
T_{1,\sigma}(t) &= \frac{t - t_m}{m(t_m - t_{m-1})} + 1\,, \qquad t \in [t_m, 1]\,.
\end{aligned}
$$

In this case, define $T_{4,\sigma}:[0,1]\times\mathbb{R}\to[0,1]\times\mathbb{R}$ by $T_{4,\sigma}=(T_{1,\sigma}(t),x)$. Then $T_{1,\sigma}:[0,t_m]\to[0,1]$ and $T_{4,\sigma}:[0,t_m]\times\mathbb{R}\to[0,1]\times\mathbb{R}$ are bijections and we let $T_{2,\sigma}:[0,1]\to[0,t_m]$ and $T_{3,\sigma}:[0,1]\times\mathbb{R}\to[0,t_m]\times\mathbb{R}$ to be their corresponding inverses.

Let $n\in\mathbb{N}$. Recall (4.4) and set

$$\begin{aligned} M_{1,n} &= \sum_{k=1}^{n}\varepsilon_{V_{n,k}}\,1_{G_{1,n}(k/n)-G_{1,n}((k-1)/n)\neq 0}\,,\\ M_{2,n} &= \sum_{0\le t\le 1}\varepsilon_{([tn]+1)/n}\,1_{G_{2,n}(t)-G_{2,n}(t-)\neq 0}\,. \end{aligned}$$

For $n\in\mathbb{N}$ and $i=1,2$, it follows from the transformation theorem that

$$\begin{aligned} G_{i,n}\circ T_{i,M_{i,n}}(t) &= \int_{[0,t]\times\mathbb{R}}(h_{n,i}\circ T_{M_{i,n}})^{1/2}(s-)\,z\,N_{i,n}^{T_{i+2,M_{i,n}}}(ds,dz)\\ h_{i,n}\circ T_{i,M_{i,n}}(t) &= h_0+\int_{[0,t]}\beta-\alpha(h_{i,n}\circ T_{i,M_{i,n}})(s-)\,dT_{i,M_{i,n}}(s)\\ &\quad+\lambda\int_{[0,t]\times\mathbb{R}}(h_{i,n}\circ T_{i,M_{i,n}})(s-)\,z^2\,N_n^{T_{i+2,M_{i,n}}}(ds,dz) \end{aligned} \tag{4.7}$$

for all $t\in[0,1]$ and $\theta=(h_0,\beta,\alpha,\gamma)\in[0,\infty)^4$.

Let $\theta=(h_0,\beta,\alpha,\gamma)\in[0,\infty)^4$. If $h_0=\beta=0$ then it follows from (4.2) and (4.6) and (4.7) that $h_{i,n}=h_{i,n}\circ T_{i,M_{i,n}}\equiv 0$, $i=1,2$, a.s., and, thus,

$$\mathcal{L}_\theta(G_{i,n}) = \mathcal{L}_\theta(G_{i,n}\circ T_{i,M_{i,n}})) = \varepsilon_0\,,\qquad n\in\mathbb{N}\,,\ i=1,2\,.$$

Otherwise, if $h_0+\beta>0$ then it follows from (4.2) and (4.6) that $h_{i,n}(t)>0$ for all $t\in(0,1]$ a.s., $i=1,2$. In this case we have $M_{1,n}=N_{2,n}$, $M_{2,n}=N_{1,n}$, $N_{1,n}^{T_{3,M_{1,n}}}=N_{n,2}$ and $N_{2,n}^{T_{4,M_{2,n}}}=N_{n,1}$ and, thus, we get from (4.7) that both

$$\mathcal{L}_\theta(G_{1,n}) = \mathcal{L}_\theta(G_{2,n}\circ T_{2,M_{2,n}})\quad\text{and}\quad \mathcal{L}_\theta(G_{2,n})=\mathcal{L}_\theta(G_{1,n}\circ T_{1,M_{1,n}}))\,.$$

for $n\in\mathbb{N}$. In other words, for all $n\in\mathbb{N}$ there are Markov kernels $K_{1,2,n}: D_1\times\mathcal{D}_1\to[0,1]$ and $K_{2,1,n}:D_1\times\mathcal{D}_1\to[0,1]$, not depending on $\theta\in[0,\infty)^4$, such that $K_{1,2,n}\mathcal{L}_\theta(G_{2,n})=\mathcal{L}_\theta(G_{1,n})$ and $K_{2,1,n}\mathcal{L}_\theta(G_{1,n}) = \mathcal{L}_\theta(G_{2,n})$ for all $\theta\in[0,\infty)^4$. Hence $\mathcal{E}^\star_{1,n}$ is equivalent to $\mathcal{E}^\star_{2,n}$ in deficiency by (A.2) for all $n\in\mathbb{N}$. This completes the proof of the theorem. □

4.2. *Proof of Lemma 4.1.* Suppose that (2.2) is satisfied for $n\in\mathbb{N}$, $p_n\in(0,1)$ and $\gamma\in(0,\infty)$. Let $B_{n,1},\dots,B_{n,n}$ be independent Bernoulli variables with parameter p_n. Suppose that $(U_k,Z_k)_{k\in\mathbb{N}}$ is an iid sequence of random vectors with independent components where U_k is uniformly distributed on $(0,1)$ and $\mathcal{L}(Z_k) = Q$. Suppose that $B_{n,1}, \dots, B_{n,n}$ and $(U_k, Z_k)_{k\in\mathbb{N}}$ are independent. Observe that

$$\mathcal{L}(N_{2,n}) \;=\; \mathcal{L}\left(\sum_{k=1}^{n} B_{n,k}\varepsilon_{(V_{n,k},Z_k)}\right)\,,$$

with $V_{n,k} = (k-1+U_k)/n$ for all $n \in \mathbb{N}$ and $1 \le k \le n$.

Let $\hat{N}_n$ be a Poisson measure on $[0,1]\otimes\mathbb{R}$ with intensity measure $np_n\ell\otimes Q$ and define

$$\hat{N}_{n,k}(B) = N\left(B \cap \left(\left(\frac{k-1}{n}, \frac{k}{n}\right] \times \mathbb{R}\right)\right)\,, \qquad B \in \mathcal{B}([0,1]\times\mathbb{R})\,.$$

Then $\widehat{N}_{n,1}, .., \widehat{N}_{n,n}$ are independent Poisson point processes where, for all $n \in \mathbb{N}$, $1 \le k \le n$, $\widehat{N}_{n,k}$ has intensity measure

$$np_n[\ell\otimes Q]\left(B \cap \left(\left(\frac{k-1}{n}, \frac{k}{n}\right] \times \mathbb{R}\right)\right)\,, \qquad B \in \mathcal{B}([0,1]\times\mathbb{R})\,.$$

By the monotonicity theorem of Markov kernels (cf. [20], Lemma 1.4.2(i)), observe that, for all $n \in \mathbb{N}$,

$$\|\mathcal{L}(N_{2,n}) - \mathcal{L}(\hat{N}_n)\| \;\le\; \left\|\bigotimes_{k=1}^{n}\mathcal{L}\left(B_{n,k}\varepsilon_{(V_k,Z_k)}\right) - \bigotimes_{k=1}^{n}\mathcal{L}\left(\widehat{N}_{n,k}\right)\right\|\,. \tag{4.8}$$

Denote the Hellinger's distance between two probability measures P_1 and P_2 by $H(P_1, P_2)$. This gives us the following upper bound [cf. [20], Section 1.3, Equation (1.23) and Section 1.3, Equation (1.25)]:

$$\begin{aligned}
&\left\|\bigotimes_{k=1}^{n}\mathcal{L}\left(B_{n,k}\varepsilon_{(V_k,Z_k)}\right) - \bigotimes_{k=1}^{n}\mathcal{L}\left(\widehat{N}_{n,k}\right)\right\| \\
&\quad\le\; H\left(\bigotimes_{k=1}^{n}\mathcal{L}\left(B_{n,k}\varepsilon_{(V_k,Z_k)}\right), \bigotimes_{k=1}^{n}\mathcal{L}\left(\widehat{N}_{n,k}\right)\right) \\
&\quad\le\; \left(\sum_{k=1}^{n} H^2\left(\mathcal{L}\left(B_{n,k}\varepsilon_{(V_k,Z_k)}\right), \mathcal{L}\left(\widehat{N}_{n,k}\right)\right)\right)^{1/2}\,.
\end{aligned} \tag{4.9}$$

Fix $n \in \mathbb{N}$ and $1 \le k \le n$. Let $(V_{k,l}, Z_{k,l})$ be an iid sequence of random vectors with $\mathcal{L}(V_{k,l}, Z_{k,l}) = \mathcal{L}(V_k)\otimes Q$, $l \in \mathbb{N}$. Suppose that $(V_{k,l}, Z_{k,l})$ are independent of $B_{n,k}$ and $\tau_{n,k}$ where $\tau_{n,k}$ is a Poisson variable with parameter p_n. Then we have the following identities:

$$\mathcal{L}\left(B_{n,k}\varepsilon_{(V_k,Z_k)}\right) \;=\; \mathcal{L}\left(\sum_{l=1}^{B_{n,k}}\varepsilon_{(V_{k,l},Z_{k,l})}\right)\,, \quad \mathcal{L}(\widehat{N}_{n,k}) \;=\; \mathcal{L}\left(\sum_{l=1}^{\tau_{n,k}}\varepsilon_{(V_{k,l},Z_{k,l})}\right)\,.$$

By Lemma 1.4.2(ii) in [20], for $n \in \mathbb{N}$ and $1 \le k \le n$, we must have

$$H\left(\mathcal{L}\left(B_{n,k}\, \varepsilon_{(V_k,Z_k)}\right), \mathcal{L}\left(\widehat{N}_{n,k}\right)\right) \le H(\mathcal{L}(B_{n,k}), \mathcal{L}(\tau_{n,k}))\,. \tag{4.10}$$

As $H(\mathcal{L}(B_{n,k}), \mathcal{L}(\tau_{n,k})) \le 3^{1/2} p_n$ [cf. [20], Theorem 1.3.1(ii)], it follows from (4.8)–(4.10), and (2.2) that

$$\limsup_{n\to\infty} \|\mathcal{L}(\tilde{N}_n) - \mathcal{L}(\hat{N}_n)\| \le \limsup_{n\to\infty} (3np_n^2)^{1/2} = 0\,. \tag{4.11}$$

In view of a well-known upper bound of the laws of a Poisson point measures in terms of the corresponding intensity measures [cf. [20], Section 3.2, Equation (3.8)], it follows from (2.2) and $\|\ell \otimes Q\| = 1$ that

$$\|\mathcal{L}(\hat{N}_n) - \mathcal{L}(N)\| \le 3|np_n - \gamma| \|\ell \otimes Q\| \le 3|np_n - \lambda| \to 0\,, \quad n \to \infty\,. \tag{4.12}$$

By means of (4.11) and (4.12), this completes the proof of the lemma. $\square$

4.3. *Proof of Theorem 2.2.* Recall that Le Cam's distance is a pseudometric. In view of (A.4) and Theorem 2.1, it, thus, suffices to show (2.15). Therefore we need some notations: for $n \in \mathbb{N}$ let $Z_n = (Z_{n,k})_{1\le k\le n}$ be a random vector with a distribution as in (2.1). Let N_n be as in (4.1) and set $\|N_n\| = N_n([0,1] \times \mathbb{R})$, $n \in \mathbb{N}$. Let Θ be as in the assertion of the theorem. Suppose that $H_{1,n} = H_n = (h_{0,1,n}, \beta_{1,n}, \alpha_{1,n}, \lambda_{1,n}) : \Theta \to [0,\infty)^4$ satisfies the assumptions of the theorem. Further, let $H_{2,n} = (h_{0,2,n}, \beta_{2,n}, \alpha_{2,n}, \lambda_{2,n}) = H_n^{(0)} : \Theta \to [0,\infty)^4$ be defined by the identities in (2.11)–(2.12).

For $\theta \in \Theta$ and $i = 1, 2$, let us define $X_{i,n} = (X_{i,n}(k))_{1\le k\le n}$ by

$$\begin{aligned} X_{i,n}(k) &= h_{i,n}^{1/2}(k-1)\, Z_{n,k}\,, \qquad X_{i,n}(0) = 0\,, \\ h_{i,n}(k) &= \beta_{i,n}(\theta) + h_{i,n}(k-1)\left[\alpha_{i,n}(\theta) + \lambda_{i,n}(\theta)\, Z_{n,k}^2\right], \\ h_{i,n}(0) &= h_{0,i,n}(\theta)\,, \qquad n \in \mathbb{N},\ 1 \le k \le n\,. \end{aligned} \tag{4.13}$$

Let

$$M_{n,k} = \left\{\sigma = (\sigma_l)_{1\le l\le k} \in \mathbb{N}^k : \sum_{l=1}^k \sigma_l \le n\right\}, \quad 1 \le k \le n\,,\ n \in \mathbb{N}. \tag{4.14}$$

By employing the conventions $0^0 = 1$ and $\sum_{l=k}^m = 0$ for $m < k$, we set

$$\begin{aligned} \eta_{i,n,1,l,\sigma}(\theta) &= \beta_{i,n}(\theta) \sum_{m=0}^{\sigma_{l+1}-1} [\alpha_{i,n}(\theta)]^m\,, \\ \eta_{i,n,2,l,\sigma}(\theta) &= [\alpha_{i,n}(\theta)]^{\sigma_{l+1}}\,, \\ \eta_{i,n,3,l,\sigma}(\theta) &= \lambda_{i,n}(\theta)\, [\alpha_{i,n}(\theta)]^{\sigma_{l+1}-1}\,. \end{aligned} \tag{4.15}$$

for $\sigma=(\sigma_l)_{1\le l\le k}\in M_{n,k}$, $1\le k\le n$, $0\le l\le k-1$, $i=1,2$, and $n\in\mathbb{N}$.

Also, we define recursively functions from $\mathbb{R}^k\to\mathbb{R}$ by setting

$$
\begin{aligned}
(4.16)\quad \hat g_{i,n,0,\sigma,\theta} &\equiv h_{0,i,n}(\theta)\,\alpha_{i,n}(\theta)^{\sigma_1-1}+\beta_{i,n}(\theta)\sum_{m=0}^{\sigma_1-2}\alpha_{i,n}^m(\theta)\,,\\
\hat g_{i,n,l,\sigma,\theta}(y) &= \eta_{i,n,1,l,\sigma}(\theta)+\eta_{i,n,2,l,\sigma}(\theta)\,\hat g_{i,n,l-1,\sigma,\theta}(y)+\eta_{i,n,3,l,\sigma}(\theta)\,y_l^2\,,
\end{aligned}
$$

for $y\in\mathbb{R}^k$, $\sigma=(\sigma_l)_{1\le l\le k}\in M_{n,k}$, $1\le k\le n$, $0\le l\le k-1$, $i=1,2$ and $n\in\mathbb{N}$.

Let $n\in\mathbb{N}$ and $1\le k\le n$. On $\{\|N_n\|=k\}$ we consider the following stopping times

$$
\tau_0=0\,,\quad \tau_m=\min\{\nu\in\{\tau_{m-1}+1,\dots,n\}:Z_{n,\nu}\neq 0\}\,,\quad 1\le m\le k\,.
$$

Using these stopping times let $\Delta\tau=((\Delta\tau_m)_{1\le m\le k})\in M_{n,k}$ be the random vector defined componentwise by $\Delta\tau_m=\tau_m-\tau_{m-1}$ for $1\le m\le k$.

Let $i=1,2$, $n\in\mathbb{N}$, $1\le k\le n$ and $\theta\in\Theta$. On $\{\|N_n\|=0\}$ we set $Y_{i,n}=0$. Otherwise, let

$$
(4.17)\qquad Y_{i,n}=(Y_{i,n}(l))_{1\le l\le\|N_n\|}=(X_{i,n}(\tau_l))_{1\le l\le\|N_n\|}\,.
$$

In the notations of (4.15) and (4.16), $Y_{i,n}$ satisfies the following recursion on $\{\|N_n\|=k\}$:

$$
\begin{aligned}
(4.18)\quad Y_{i,n}(l) &= g_{i,n}^{1/2}(l-1)\,Z_{n,\tau_l}\,,\quad Y_{i,n}(0)=0\,,\quad 1\le l\le k\,,\\
g_{i,n}(l) &= \eta_{i,n,1,l,\Delta\tau}(\theta)+g_{i,n}(l-1)\,\eta_{i,n,2,l,\Delta\tau}(\theta)\\
&\quad+\eta_{i,n,3,l,\Delta\tau}(\theta)\,g_{i,n}(l-1)\,Z_{n,\tau_l}^2\,,\quad 1\le l\le k-1\,,\\
g_{i,n}(0) &= \hat g_{i,n,0,\Delta\tau,\theta}\,.
\end{aligned}
$$

Recall (4.14). For all $n\in\mathbb{N}$, $1\le k\le n$ and $\sigma=(\sigma_l)_{1\le l\le k}\in M_{n,k}$ let

$$
(4.19)\qquad A_{n,k,\sigma}=\Big\{\|N_n\|=k\,,\ \Delta\tau=\sigma\Big\}\,.
$$

For future purposes we collect some useful inequalities into the next lemma.

Lemma 4.2. *Suppose that $(\Theta,(H_n)_{n\in\mathbb{N}})$ satisfies the assumption of Theorem 2.2. Let $S\in(0,\infty)$ and suppose that $Q([-S,S])=1$.*

Then there exists $C=C(S,\Theta)\in(1,\infty)$ and $n_0=n_0(S,\Theta)\in\mathbb{N}$ such that the following three inequalities are in place

$$
\begin{aligned}
(4.20)\qquad |\hat g_{1,n,0,\sigma,\theta}-\hat g_{2,n,0,\sigma,\theta}| &\le \frac{C}{n}\\
(4.21)\qquad \hat g_{i,n,l,\sigma,\theta}(y) &\ge C^{-1}\,,
\end{aligned}
$$

$$E_\theta\big[|\hat g_{1,n,l,\sigma,\theta}(Y_{1,n})-\hat g_{2,n,l,\sigma,\theta}(Y_{1,n})|\,\big|\,A_{n,k,\sigma}\big]\;\le\;\frac{C^k}{n}\,. \tag{4.22}$$

for all $n\ge n_0$, $1\le k\le n$, $0\le l\le k-1$, $\sigma\in M_{n,k}$, $i=1,2$, $\theta\in\Theta$, $y\in\mathbb{R}^k$ *and* $i=1,2$.

Proof of Lemma 4.2. Let $(\Theta,(H_n)_{n\in\mathbb{N}}$ be as in in Theorem 2.2. Note that $(\Theta,(H_{1,n})_{n\in\mathbb{N}})=(\Theta,(H_n)_{n\in\mathbb{N}})$ satisfies the assumption in Theorem 2.2. Also, recall that $(\Theta,(H_{2,n})_{n\in\mathbb{N}})=(\Theta,(H_n^{(0)})_{n\in\mathbb{N}})$ is defined in (2.11)–(2.11). In particular, observe that $\alpha_{i,n}(\theta)\to 1$ uniformly for all $\theta=(h_0,\beta,\alpha,\lambda)\in\Theta$ as $n\to\infty$, $i=1,2$ and, thus, there is a $n_1=n_1(\Theta)\in\mathbb{N}$ satisfying

$$\frac{e^{-1}}{2}\le[\alpha_{i,n}(\theta)]^n=\exp(n\log[n+n(\alpha_{i,n}(\theta)-1)]-n\log n)\;\le\;2e\,, \tag{4.23}$$

for all $n\ge n_1$, $i=1,2$ and $\theta=(h_0,\beta,\alpha,\lambda)\in\Theta$.

It follows from our assumptions on $(\Theta,(H_n)_{n\in\mathbb{N}})$ that there exist $n_0=n_0(\Theta)\ge n_1$ and $C_1=C_1(\Theta)\in(1,\infty)$ such that

$$\begin{aligned}\hat g_{i,n,l,\sigma,\theta}(y)\;&\ge\;h_{0,i,n}(\theta)[\alpha_{i,n}(\theta)]^{-1+\sum_{m=1}^{l+1}\sigma_k}\;\ge\;\frac{e^{-1}}{2}\frac{h_{0,i,n}(\theta)}{\alpha_{i,n}(\theta)}\\ &\ge\;\frac{e^{-1}}{4}\inf_{(h_0,\beta,\alpha,\lambda)\in\overline{\Theta}}h_0\;\ge\;C_1^{-1}\,,\end{aligned} \tag{4.24}$$

and

$$\begin{aligned}\max\left\{h_{0,i,n}(\theta)\,,\;\beta_{i,n}(\theta)\,,\;[\alpha_{i,n}(\theta)]^n\,,\;\frac{h_{0,i,n}(\theta)}{\alpha_{i,n}(\theta)}\right\}\;&\le\;C_1\,,\\ \max\{|h_{0,1,n}(\theta)-h_{0,2,n}(\theta)|\,,\;|\beta_{1,n}(\theta)-\beta_{2,n}(\theta)|\,,&\\ |\alpha_{1,n}(\theta)-\alpha_{2,n}(\theta)|\,,\;|\lambda_{1,n}(\theta)-\lambda_{2,n}(\theta)|\}\;&\le\;\frac{C_1}{n}\,,\end{aligned} \tag{4.25}$$

for all $n\ge n_0$, $1\le k\le n$, $0\le l\le k-1$ $\sigma=(\sigma_l)_{1\le l\le k}\in M_{n,k}$, $i=1,2$, $\theta\in\Theta$ and $y\in\mathbb{R}^k$.

Recall (4.15) and (4.16). It follows from (4.25) that we have

$$\begin{aligned}\max\{\eta_{i,n,2,l,\sigma}(\theta)\,,\;\eta_{i,n,3,l,\sigma}(\theta)\}\;&\le\;C_1^2\,,\\ \max\{\eta_{i,n,1,l,\sigma}(\theta)\,,\;\hat g_{i,n,0,\sigma,\theta}\}\;&\le\;(k+1)\,C_1^2\,,\\ \max\{|[\alpha_{1,n}(\theta)]^m-[\alpha_{1,n}(\theta)]^m|\}\;&\le\;\frac{C_1^2m}{n}\,,\\ \max\{|\eta_{1,n,j,l,\sigma}(\theta)-\eta_{2,n,j,l,\sigma}(\theta)|\,:\;j=1,2,3\}\;&\le\;\frac{(2e^2C_1^3)^k}{n}\,,\\ |\hat g_{1,n,0,\sigma,\theta}-\hat g_{2,n,0,\sigma,\theta}|\;&\le\;\frac{(4e^2C_1^3)^k}{n}\,,\end{aligned} \tag{4.26}$$

for all $n \geq n_0$, $1 \leq k \leq n$, $0 \leq l \leq k-1$, $\sigma = (\sigma_l)_{1\leq l\leq k} \in M_{n,k}$, $i = 1, 2$, $m \in \mathbb{N}_0$ and $\theta \in \Theta$.

Recall (4.18). Let $S > 1$ such that $Q([-S, S]) = 1$ and set $C_2 = C_2(S, \theta) = e^2(1 + S)^2 C_1^4$ and $C_3 = C_3(S, \theta) = S^2 C_2$. It follows from an induction and the inequalities in (4.26) that

$$
\begin{aligned}
E_\theta[g_{i,n}(l)|A_{n,k,\sigma}] &\leq C_1^2(k+1) + C_1^2(1+S^2)E_\theta[g_{i,n}(l-1)|A_{n,k,\sigma}] \\
&\leq (k+1) \sum_{m=0}^{l} (1+S^2)^m C_1^{2(1+m)} \leq C_2^k ,
\end{aligned}
\tag{4.27}
$$

and, thus,

$$
E_\theta[Y_{i,n}^2(l)|A_{n,k,\sigma}] \leq C_3^k , \tag{4.28}
$$

for all $n \geq n_0$, $1 \leq k \leq n$, $0 \leq l \leq k-1$, $\sigma = (\sigma_l)_{1\leq l\leq k} \in M_{n,k}$, $i = 1, 2$ and $\theta \in \Theta$.

Finally, let $C = C(S, \theta) = 12e^3 C_1^3 C_3$. By an induction it follows from (4.26)–(4.28) that

$$
\begin{aligned}
&= E_\theta\left[|\hat{g}_{1,n,l,\sigma,\theta}(Y_{1,n}) - \hat{g}_{2,n,l,\sigma,\theta}(Y_{1,n})|\middle|A_{n,k,\sigma}\right] \\
&\leq C_3^k \sum_{j=1}^{3} |\eta_{1,n,j,l}(\theta) - \eta_{2,n,j,l}(\theta)| \\
&\quad + E_\theta\left[|\hat{g}_{1,n,l-1,\sigma,\theta}(Y_{1,n}) - \hat{g}_{2,n,l-1,\sigma,\theta}(Y_{1,n})|\middle|A_{n,k,\sigma}\right] \\
&\leq |\hat{g}_{1,n,0,\sigma,\theta} - \hat{g}_{2,n,0,\sigma,\theta}| + C_4^k \sum_{l=1}^{k-1} \sum_{j=1}^{3} |\eta_{1,n,j,l}(\theta) - \eta_{2,n,j,l}(\theta)| \leq \frac{C^k}{n} ,
\end{aligned}
$$

for all $n \geq n_0$, $1 \leq k \leq n$, $0 \leq l \leq k-1$ $\sigma = (\sigma_l)_{1\leq l\leq k} \in M_{n,k}$ and $\theta \in \Theta$. This completes the proof in view of (4.24) and (4.26). □

We provide an upper bound for conditional laws and their total variation norm in the next lemma.

LEMMA 4.3. *Suppose that Q admits a Lebesgue density f where f is globally Lipschitz and has a compact support $\overline{\{f > 0\}}$.*

If $(\Theta, (H_n)_{n\in\mathbb{N}})$ satisfies the assumptions of the Theorem 2.2 then there exist $n_0 = n_0(f, \Theta) \in \mathbb{N}$ and $C = C(f, \Theta) \in (0, \infty)$ such that

$$
\left\| \mathcal{L}_\theta(Y_{1,n}|A_{n,k,\sigma}) - \mathcal{L}_\theta(Y_{2,n}|A_{n,k,\sigma}) \right\| \leq \frac{C^k}{n} , \tag{4.29}
$$

for all $\theta \in \Theta$, $n \geq n_0$, $1 \leq k \leq n$ and $\sigma \in M_{n,k}$.

Proof of Lemma 4.3. By assumption we have $f(x) = 0$ for all $|x| \geq S$ and some $S > 0$. Hence there are $n_0 = n_0(f, \theta) \in \mathbb{N}$ and $C_1 = C_1(f, \theta) \in (1, \infty)$ such that, for C replaced by C_1, the assertion of Lemma 4.2 is in place.

Let $n \geq n_0$, $i = 1, 2$, $\theta \in \Theta$, $1 \leq k \leq n$ and $\sigma \in M_{n,k}$. Recall (4.16). In view of (4.21), $\Psi_{i,n,\theta} : \mathbb{R}^k \to \mathbb{R}^k$ is a well-defined C^∞-diffeomorphism, where $\Psi_{i,n,\sigma,\theta} = (\psi_{i,n,l,\sigma,\theta})_{1 \leq l \leq k} : \mathbb{R}^k \to \mathbb{R}^k$ is defined by

$$\psi_{i,n,l,\sigma,\theta}(y) = \frac{y_l}{\hat{g}^{1/2}_{i,n,l-1,\sigma,\theta}(y)}\,, \tag{4.30}$$

for $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$ and $1 \leq l \leq k$. For all $n \geq n_0$, $\theta \in \Theta$, $n \geq n_0$, $1 \leq k \leq n$ and $\sigma \in M_{n,k}$ we define

$$\tilde{f}_{i,n,k,\sigma,\theta}(y) \;=\; \prod_{l=1}^{k} \frac{f(\psi_{i,n,l,\sigma,\theta}(y))}{\hat{g}^{1/2}_{i,n,l-1,\sigma,\theta}(y)}\,, \quad y \in \mathbb{R}^k\,,\ i = 1, 2\,.$$

It follows from (4.16) and (4.18) and (4.30) that $\tilde{f}_{i,n,k,\sigma,\theta}$ is a density of the probability measure $\mathcal{L}_\theta(Y_{i,n}|A_{n,k,\sigma})$ with respect to the Lebesgue measure $\ell^{\otimes k}$ on $\mathcal{B}(\mathbb{R}^k)$. In particular, we must have

$$\begin{aligned} &\|\mathcal{L}_\theta(Y_{1,n}|A_{n,k,\sigma}) - \mathcal{L}_\theta(Y_{2,n}|A_{n,k,\sigma})\| \\ &\qquad = \; \frac{1}{2}\int_{\mathbb{R}^k} |\tilde{f}_{1,n,k,\sigma,\theta}(y) - \tilde{f}_{2,n,k,\sigma,\theta}(y)|\, dy\,, \end{aligned} \tag{4.31}$$

for all $\theta \in \Theta$, $n \geq n_0$, $1 \leq k \leq n$ and $\sigma \in M_{n,k}$.

Suppose that $C_f \in (0, \infty)$ is a global Lipschitz constant of f. By means of simple substitutes, for all $\epsilon > 0$ and $w, v \geq \epsilon$, observe

$$\frac{1}{2}\int \left| \frac{f(x/v)}{v} - \frac{f(x/w)}{w} \right| dx \;\leq\; \frac{1}{\epsilon}\,(S^2 C_f + 1)\,|v - w|\,.$$

Consequently, for all $\epsilon > 0$, we find a $\kappa_1 = \kappa_1(f, \epsilon) \in (1, \infty)$ such that

$$\frac{1}{2}\int \left| \frac{f(x/v)}{v} - \frac{f(x/w)}{w} \right| dx \;\leq\; \kappa_1(\epsilon)|v - w|\,, \quad v, w \geq \epsilon\,.$$

In view of (4.21), there, thus, exists $\kappa_2 = \kappa_2(f, \Theta) \in (1, \infty)$ such that

$$\begin{aligned} &\frac{1}{2}\int \left| \frac{f(y_l/\hat{g}^{1/2}_{1,n,l-1,\sigma,\theta}(y))}{\hat{g}^{1/2}_{1,n,l-1,\sigma,\theta}(y)} - \frac{f(y_l/\hat{g}^{1/2}_{2,n,l-1,\sigma,\theta}(y)}{\hat{g}^{1/2}_{2,n,l-1,\sigma,\theta}(y))} \right| dy_l \\ &\qquad\qquad \leq \;\; \kappa_2 |\hat{g}_{1,n,l-1,\sigma,\theta}(y) - \hat{g}_{2,n,l-1,\sigma,\theta}(y)|\,, \end{aligned} \tag{4.32}$$

for all $n \geq n_0$, $1 \leq k \leq n$, $1 \leq l \leq k$, $\sigma \in M_{n,k}$, $y \in \mathbb{R}^k$ and $\theta \in \Theta$. By integrating over y_k, we get from (4.32) that

$$
\begin{aligned}
(4.33) \qquad & \frac{1}{2}\int_{\mathbb{R}^k} |\tilde{f}_{1,n,k,\sigma,\theta}(y) - \tilde{f}_{2,n,k,\sigma,\theta}(y)|\, dy \\
\leq \;& \kappa_2 \int_{\mathbb{R}^{k-1}} \prod_{l=1}^{k-1} \frac{f(\psi_{1,n,l,\sigma,\theta}(y))}{\hat{g}^{1/2}_{1,n,l-1,\sigma,\theta}(y)} \,|\hat{g}_{1,n,k-1,\sigma,\theta}(y) - \hat{g}_{2,n,k-1,\sigma,\theta}(y)|\; dy \\
& + \frac{1}{2}\int_{\mathbb{R}^{k-1}} \left| \prod_{l=1}^{k-1} \frac{f(\psi_{1,n,l,\sigma,\theta}(y))}{\hat{g}^{1/2}_{1,n,l-1,\sigma,\theta}(y)} - \prod_{l=1}^{k-1} \frac{f(\psi_{2,n,l,\sigma,\theta}(y))}{\hat{g}^{1/2}_{2,n,l-1,\sigma,\theta}(y)} \right| dy\,,
\end{aligned}
$$

for all $n \geq n_0$, $1 \leq k \leq n$, $\sigma \in M_{n,k}$ and $\theta \in \Theta$. It follows from (4.22) that

$$
\begin{aligned}
(4.34) \qquad & \int_{\mathbb{R}^{k-1}} \prod_{l=1}^{k-1} \frac{f(\psi_{1,n,l,\sigma,\theta}(y))}{g^{1/2}_{1,n,l-1,\sigma,\theta}(y)} \,|\hat{g}_{1,n,k-1,\sigma,\theta}(y)) - \hat{g}_{2,n,k-1,\sigma,\theta}(y))|\; dy \\
= \;& E_\theta\left[\left|\hat{g}_{1,n,k-1,\theta}(Y_{1,n}) - \hat{g}_{2,n,k-1,\theta}(Y_{1,n})\right| \Big| A_{n,k,\sigma}\right] \leq \frac{C_1^k}{n}\,.
\end{aligned}
$$

for all $n \geq n_0$, $1 \leq k \leq n$, $\sigma \in M_{n,k}$ and $\theta \in \Theta$, .

Let $C = e\,\kappa_2\, C_1$. By an induction we, thus, get from (4.20) and (4.33)–(4.34) that

$$
\|\mathcal{L}_\theta(Y_{1,n}|A_{n,k,\sigma}) - \mathcal{L}_\theta(Y_{2,n}|A_{n,k,\sigma})\| \leq \frac{C^k}{n}\,,
$$

uniformly for all $n \geq n_0$, $1 \leq k \leq n$, $\sigma \in M_{n,k}$ and $\theta \in \Theta$. This completes the proof of the lemma. □

Let f be a Lebesgue density of Q and Θ be as in Theorem 2.2. We denote the positive part of a function $g : \mathbb{R} \to \mathbb{R}$ by g^+. Let C_C^∞ be the space of infinitely often continuously differentiable functions $g : \mathbb{R} \to \mathbb{R}$ with compact support $\overline{\{g > 0\}}$. As C_C^∞ is dense in L^1 we find a sequence of $g_m \in C_C^\infty$, $m \in \mathbb{N}$, such that $\int |g_m - f| d\ell \to 0$ as $m \to \infty$. It is immediate that both, $\int |g_m^+ - f| d\ell \to 0$ and $\int g_m^+\, d\ell \to 1$ as $m \to \infty$. Without loss of generality, we may, thus, assume that $\int g_m^+\, d\ell > 0$ for all $m \in \mathbb{N}$. Then $f_m := g_m^+ / \int g_m^+ d\ell$ defines a sequence of globally Lipschitz continuous probability densities with a compact support $\overline{\{g_m > 0\}}$ such that $\int |f_m - f| d\ell \to 0$.

For $m \in \mathbb{N}$ let $Z_n^{(m)} = (Z_{n,k}^{(m)})_{1 \leq k \leq n}$ be a random vector with distribution

$$
\mathcal{L}(Z_n^{(m)})(B) \;=\; \big((1-p_n)\varepsilon_0(B) + p_n \int_B f_m d\ell\big)^{\otimes n}\,,
$$

with $B \in \mathcal{B}(\mathbb{R}^n)$, $m, n \in \mathbb{N}$, $1 \le k \le n$. If we replace $Z_{n,k}$ by $Z_{n,k}^{(m)}$ in (4.13) then we get yet another family of GARCH models $X_{i,n}^{(m)} = (X_{i,n}^{(m)}(k))_{1\le k\le n}$, say, indexed by $\theta \in \Theta$, $i = 1, 2$ and m, $n \in \mathbb{N}$.

It follows from the monotonicity theorem for Markov kernels and a well-known upper bound for product measures [cf. [20], Lemma 1.4.2(i) & page 23] that, for all $i = 1, 2$,

$$\begin{aligned} \sup_{\theta\in\Theta_0} \|\mathcal{L}_\theta(X_{i,n}) - \mathcal{L}_\theta(X_{i,n}^{(m)})\| \;&\le\; \|\mathcal{L}(Z_n) - \mathcal{L}(Z_n^{(m)})\| \\ \le\;\; n\|\mathcal{L}(Z_{n,1}) - \mathcal{L}(Z_{n,1}^{(m)})\| &= \frac{np_n}{2} \int |f_m - f|d\ell \,. \end{aligned} \tag{4.35}$$

As f_m is globally Lipschitz with a compact support $\overline{\{f_m > 0\}}$ for all $m \in \mathbb{N}_0$ the assumptions of Lemma 4.3 are in place. For all $m \in \mathbb{N}$ there, thus, exist $n_m \in \mathbb{N}$ and $C_m = C(f_m, \Theta) \in (0, \infty)$ such that, for all $n \ge n_m$, we get by conditioning and the monotonicity theorem for Markov kernels that

$$\sup_{\theta\in\Theta} \|\mathcal{L}_\theta(X_{1,n}^{(m)}) - \mathcal{L}_\theta(X_{2,n}^{(m)}\| \;\le\; \frac{1}{n}\, E\Big[C_m^{\|N_n\|}\Big]\,, \tag{4.36}$$

for N_n as defined in (4.1). Recall (4.13). By combining (4.35) and (4.36) we get from the triangular inequality that

$$\sup_{\theta\in\Theta} \Big\|\mathcal{L}_\theta(G_n) - \mathcal{L}_\theta(G_n^{(0)})\Big\| \;\le\; np_n \int |f_m - f|d\ell + \frac{2}{n}\, E\Big[C_m^{\|N_n\|}\Big]\,,$$

for all $m \in \mathbb{N}$ and $n \ge n_m$. As (2.2) is in place, we have $\lim_{n\to\infty} EC_m^{\|N_n\|} = e^{\lambda(C_m-1)}$ and, thus,

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta} \Big\|\mathcal{L}_\theta(G_n) - \mathcal{L}_\theta(G_n^{(0)})\Big\| \le \lambda\, \limsup_{m\to\infty} \int |f_m - f|d\ell = 0\,,$$

giving (2.15). This completes the proof of Theorem 2.2. □

4.4. *Proof of Theorem 2.3.* We need some preparations. Let $Z = (Z_n)_{n\in\mathbb{N}}$ and $U = (U_n)_{n\in\mathbb{N}}$ be independent sequences of iid random variables such that Z_1 and U_1 are standard normal and uniformly distributed on $(0, 1)$, respectively. For $d \in \mathbb{N}$, we denote the order statistics of $0, U_1, \ldots, U_d$ by $0 =: U_{d,0} < U_{d,1} \le \ldots \le U_{d,d}$ $[d \in \mathbb{N}]$. For each $\gamma > 0$ let $\nu = \nu(\gamma)$ be a Poisson random variable with parameter $\gamma > 0$, independent of Z and U. [We suppress the dependence of γ in our notations in the sequel.]

In both (2.4) and (2.5), N admits a representation $N = \sum_{k=1}^{\nu} \varepsilon_{(U_{\nu,k}, Z_k)}$, since N is a Poisson measure with the intensity $\gamma\ell \otimes Q$. On $\{\nu = 0\}$ let $\Delta U_\nu = \Delta G_\nu = \Delta \widehat{G}_\nu = 0$, whereas, on $\{\nu > 0\}$, we set

$$\begin{aligned} \Delta U_\nu &= (U_{\nu,k} - U_{\nu,k-1})_{1\le k\le \nu}\,, \\ \Delta G_\nu &= (G(U_{\nu,k}) - G(U_{\nu,k}-))_{1\le k\le \nu}\,, \\ \Delta \widehat{G}_\nu &= (\widehat{G}(U_{\nu,k}) - \widehat{G}(U_{\nu,k}-))_{1\le k\le \nu}\,. \end{aligned}$$

Let $S_0 = \mathbb{R}^0 = \{0\}$ and $\widetilde{\mathbb{R}} = \bigcup_{d=0}^{\infty} \{d\} \times S_d \times \mathbb{R}^d$ where, for $d \in \mathbb{N}$, S_d equals the set of all $w = (w_1, \ldots, w_d)' \in (0,1)^d$ such that $\sum_{i=1}^d w_i \le 1$. We endow S_d and $\widetilde{\mathbb{R}}$ with the Borel trace field $\mathcal{B}(S_d)$ $[d \ge 0]$ and the σ-algebra $\widetilde{\mathcal{B}}$, respectively, where $\widetilde{\mathcal{B}}$ is the set of all $B \subseteq \widetilde{\mathbb{R}}$ such that $B \cap (\{d\} \times S_d \times \mathbb{R}^d) \in \{\emptyset, \{d\}\} \otimes \mathcal{B}(S_d) \otimes \mathcal{B}(\mathbb{R}^d)$ for all $d \in \mathbb{N}_0$.

Since we assumed that $\Theta \subseteq (0,\infty) \times [0,\infty)^3$, and since G and $\widehat{G}$ jump always at the same time as N does, all arrival times are observed in full and, thus, $\mathcal{E}_{\gamma,Q}(\Theta)$ and $\widehat{\mathcal{E}}_{\gamma,Q}(\Theta)$ are equivalent to $\mathcal{F}_\gamma$ and $\widehat{\mathcal{F}}_\gamma$ in deficiency, respectively, in view of (A.2), where, for all $\gamma > 0$, we set

$$\mathcal{F}_\gamma = (\widetilde{\mathbb{R}}, \widetilde{\mathcal{B}}, (\mathcal{L}_\theta(\nu, \Delta U_\nu, \Delta G_\nu))_{\theta\in\Theta})\,,\ \widehat{\mathcal{F}}_\gamma = (\widetilde{\mathbb{R}}, \widetilde{\mathcal{B}}, (\mathcal{L}_\theta(\nu, \Delta U_\nu, \Delta \widehat{G}_\nu))_{\theta\in\Theta}).$$

Let $\widehat{w}_0 = 0$ and, for $d > 0$, set $\widehat{w}_d = (1/d, \ldots, 1/d) \in \mathbb{R}^d$. Recall that $\Theta \subseteq (0,\infty) \times [0,\infty)^3$ and pick $d \in \mathbb{N}_0$, $\theta = (h_0, \beta, \alpha, \lambda) \in \Theta$, $w = (w_1, \ldots, w_d) \in S_d \cup \{\widehat{w}_d\}$. We define a diffeomorphism $\Psi_{d,w,\theta} : \mathbb{R}^d \to \mathbb{R}^d$ as follows: if $d = 0$ then let $\Psi_{d,w,\theta} = 0$, otherwise, if $d > 0$ then let

$$\Psi_{d,w,\theta}(z) = \left(h_{d,w,\theta,k}^{1/2}(z)\, z_k\right)_{1\le k\le d}\,, \qquad z = (z_1, \ldots, z_d) \in \mathbb{R}^d\,,$$

where, for $2 \le k \le d$, recursively, we define

$$\begin{aligned} h_{d,w,\theta,k}(z) &= \frac{\beta}{\alpha}(1 - e^{-\alpha w_k}) + e^{-\alpha w_k}(1 + \lambda z_{k-1}^2)\, h_{d,w,\theta,k-1}(z)\,, \\ h_{d,w,\theta,1}(z) &\equiv h_{d,w,\theta,1} = \frac{\beta}{\alpha}(1 - e^{-\alpha w_1}) + e^{-\alpha w_1} h_0\,, \end{aligned} \tag{4.37}$$

provided $\alpha > 0$, and, otherwise, if $\alpha = 0$ then we set

$$\begin{aligned} h_{d,\theta,w,k}(z) &= \beta w_k + h_{d,\theta,w,k-1}(z)(1 + \lambda z_{k-1}^2)\,, \\ h_{d,\theta,w,1}(z) &\equiv h_{d,w,\theta,1} = \beta w_1 + h_0\,. \end{aligned} \tag{4.38}$$

With $f = dQ/dx$ being the Lebesgue density of a standard normal distribution we set

$$\mathcal{H}_{d,\theta_1,\theta_2,w}(z) = \int_{\mathbb{R}^d} \left(|J_{\Psi^{-1}_{d,w,\theta_1}}(x)|\, f^{\otimes d}(\Psi^{-1}_{d,w,\theta_1}(x))\right)^z \left(|J_{\Psi^{-1}_{d,w,\theta_2}}(x)|\, f^{\otimes d}(\Psi^{-1}_{d,w,\theta_2}(x))\right)^{1-z} dx\,,$$

for all $\theta_1, \theta_2 \in \Theta$, $0<z<1$, $w \in S_d \cup \{\widehat{\mu}_d\}$.

Finally, we recall that the Hellinger transformations of two experiments are the same, provided they are equivalent in deficiency. We have already shown that the assumption, that $\mathcal{E}_{\gamma,Q}(\Theta)$ is equivalent to $\widehat{\mathcal{E}}_{\gamma,Q}(\Theta)$ for all $\gamma > 0$, is identical to the assumption, that $\mathcal{F}_\gamma$ is equivalent to $\widehat{\mathcal{F}}_\gamma$ for all $\gamma > 0$. By solving the linear odes in (2.4) and (2.5), we, thus, arrive at the following identity:

$$\sum_{d=1}^{\infty} \frac{\gamma^d e^{-\gamma}}{d!} \, \mathcal{H}_{d,\theta_1,\theta_2,\widehat{w}_d}(z) \;=\; \sum_{d=1}^{\infty} \frac{\gamma^d e^{-\gamma}}{d!} \int_{S_d} \mathcal{H}_{d,\theta_1,\theta_2,w}(z) \, \frac{dw}{\ell^{\otimes d}(S_d)} \,,$$

for all $\theta_1, \theta_2 \in \Theta$, $0<z<1$. In the last display the functions are analytical in γ; hence, we must have

$$\mathcal{H}_{d,\theta_1,\theta_2,\widehat{w}_d}(z) \;=\; \int_{S_d} \mathcal{H}_{d,\theta_1,\theta_2,w}(z) \, \frac{dw}{\ell^{\otimes d}(S_d)} \,, \tag{4.39}$$

for all $d \in \mathbb{N}$, $\theta_1, \theta_2 \in \Theta$, $0 < z < 1$. Now we return to the proof of the theorem:

(i) and (ii) For $i = 1, 2$ let $\theta_i = (h_{0,i}, \beta_i, \alpha, \lambda) \in \Theta$. For $0 < z < 1$ let $g_z : (0, \infty) \to \mathbb{R}$ be defined by $g_z(h) := h^{z/2}/(1-z+hz)^{1/2}$, $h > 0$. Let $H_{\theta_1,\theta_2} : [0,1] \to (0,\infty)$ be defined by $H_{\theta_1,\theta_2}(w) := h_{1,w,\theta_1,1}(1)/h_{1,w,\theta_2,1}(1)$. Observe that $\mathcal{H}_{1,\theta_1,\theta_2,w}(z) = g_z(H_{\theta_1,\theta_2}(w))$ for all $0<z<1$ and $0 \le w \le 1$ since we assumed that Q is the standard normal law. Further, g_z is strictly increasing on $(0,1]$ and strictly decreasing on $[1,\infty)$; also,

$$\frac{d}{dw} H_{\theta_1,\theta_2}(w) \;=\; e^{-\alpha w} \frac{\beta_1 h_{0,2} - \beta_2 h_{0,1}}{h^2_{1,w,\theta_2,1}(1)} \,, \quad 0 \le w \le 1 \,.$$

Firstly, suppose that $\beta_1 = \beta_2$ and $h_{0,1} < h_{0,2}$. Then H_{θ_1,θ_2} is strictly increasing with $H_{\theta_1,\theta_2}(1) \le 1$ contradicting (4.39) for $d = 1$. If $\beta_1 = \beta_2$ and $h_{0,1} > h_{0,2}$, then H_{θ_1,θ_2} is strictly decreasing with $H_{\theta_1,\theta_2}(1) \ge 1$, again a contradiction to (4.39). Secondly, if $h_{0,1} = h_{0,2}$ and $\beta_2 < \beta_1$ (without loss of generality) then H_{θ_1,θ_2} is strictly decreasing with $H_{\theta_1,\theta_2}(0) = 1$, again contradicting (4.39). This completes the proof of (i) and (ii).

(iii) can be shown analogously to (i). □

4.5. *Proofs of the results in Subsection 2.4.*

4.5.1. *Proof of Proposition 2.1.* For $f \in D_d[0,1]$ we write $\Delta f = f(t) - f(t-)$, $0 \le t \le 1$, with the convention $\Delta f(0) = 0$. With the usual convention $\inf \emptyset = \infty$, define $T(f) = \inf\{t \in [0,1] : \Delta f_1(t) \neq 0\} \wedge 1$ for all $f =$

$(f_1, f_2) \in D_2$. Let S be the set of all functions $f \in D_2$ with $T(f) \in (0,1)$. Let $D_0'' \subseteq D_2[0,1]$ be the set all functions $f = (f_1, f_2)$ such that the right-hand derivatives $f_1'(0+)$ and $f_2''(0+)$ exist in $\mathbb{R}$. Further, let $D_{0,T}'' \subseteq S \cap D_0''$ be the set of all functions $f = (f_1, f_2)$ such that the right-hand derivatives $f_2'(T(f)+)$ and $f_2''(T(f)+)$ exist in $\mathbb{R}$.

Let $f \in D_2$ with $T = T(f)$. If $f \in D_{0,T}''$ and $f_2'(0+) \neq 0$ then we set

$$X(f) \;=\; \left(|f_2(0)|, \left|\frac{(f_2'(0+))^2 - f_2(0) f_2''(0+)}{f_2'(0+)}\right|, \left|\frac{f_2''(0+)}{f_2'(0+)}\right|, \frac{|\Delta f_2(T))|}{(\Delta f_1(T))^2}\right).$$

If $f \in D_{0,T}''$ and $f_2'(0+) = 0$ and $f_2'(T+) \neq 0$ then we set

$$X(f) \;=\; \left(|f_2(0)|, \left|\frac{(f_2'(T+))^2 - f_2(T) f_2''(T+)}{f_2'(T+)}\right|, \left|\frac{f_2''(T+)}{f_2'(T+)}\right|, \frac{|\Delta f_2(T))|}{(\Delta f_1(T))^2}\right).$$

If $f \in D_{0,T}''$ and $f'(0+) = f_2'(T+) = 0$ and $\Delta f_2(T) \neq 0$ then we set

$$X(f) \;=\; \left(|f_2(0)|, 0, 0, \frac{|\Delta f_2(T))|}{(\Delta f_1(T))^2}\right).$$

If $f \in D_{0,T}''$ and $f'(0+) = f_2'(T+) = 0$ and $\Delta f_2(T) = 0$ then we set $X(f) = (|f_2(0)|, \infty, \infty, 0)$. If $f \in D_0'' \backslash S$ and $f_2'(0+) \neq 0$ then define

$$X(f) \;=\; \left(|f_2(0)|, \left|\frac{(f_2'(0+))^2 - f_2(0) f_2''(0+)}{f_2'(0+)}\right|, \left|\frac{f_2''(0+)}{f_2'(0+)}\right|, \infty\right).$$

For the remaining cases we set $X(f) = (|f_2(0)|, \infty, \infty, \infty)$. Then $X : D_2 \to [0,\infty]^4$ is a $\mathcal{D}_2$-$\mathcal{B}([0,\infty]^4)$-measurable mapping. Since $Q(\{0\}) = 0$ it follows from (2.4) that $\mathcal{L}_\theta^X((G,h)) = Q_\theta$ for all $\theta \in [0,\infty)^4$ and, thus, $\delta(\mathcal{E}_h, \mathcal{F}) = 0$ by (A.2), where $\mathcal{F}$ is the experiment as defined in the assertion of the proposition.

Next we show that $\delta(\mathcal{F}, \mathcal{E}_h) = 0$. To this end we define $\xi = (\xi_1, \ldots, \xi_3) : [0,\infty]^4 \to [0,\infty)^3$ as follows: let $\omega = (\omega_1, \ldots, \omega_4) \in [0,\infty]^4$. If $(\omega_1, \ldots, \omega_3) \in [0,\infty)^3$ then we set $\xi(\omega) = (\omega_1, \omega_2, \omega_3)$; if $\omega_1 \in [0,\infty)$ and either $\omega_2 = \infty$ or $\omega_3 = \infty$ then we set $\xi(\omega) = (\omega_1, 0, 0)$; otherwise, we set $\xi(\omega) = 0$.

In the notations of the Introduction we define $\Psi : [0,\infty)^3 \times \mathbb{M}_2 \to D_2$ where, for $0 \leq t \leq 1$, $\omega = (\omega_1, \omega_2, \omega_3) \in [0,\infty)^3$ and $\sigma \in \mathbb{M}_2$, $(f_1(t), f_2(t)) = \Psi[\omega, \sigma](t)$ is defined to be the unique solution of the system of the following integral equations

$$
\begin{aligned}
(4.40) \qquad f_1(t) &= \int_{[0,t]\times\mathbb{R}^2} f_2^{1/2}(s-) z_1\, \sigma(ds, dz_1, dz_2)\,, \\
f_2(t) &= \omega_1 + \int_{[0,t]} (\omega_2 - \omega_3 f_2(s-))\, ds \\
&\quad + \int_{[0,t]\times\mathbb{R}\times(0,\infty)} f_2(s-) z_2^2\, \sigma(ds, dz_1, dz_2)\,,
\end{aligned}
$$

Clearly, Ψ is $(\mathcal{B}([0,\infty)^3) \otimes \mathcal{M}_2)/\mathcal{D}_2$ measurable and, thus, defines a deterministic Markov kernel $K_2 : ([0,\infty)^3 \times \mathbb{M}_2) \times \mathcal{D}_2 \to [0,1]$.

Let ν_0 be the zero measure on $\mathcal{B}([0,1]\times\mathbb{R}^2)$. For $\lambda \geq 0$ let M_λ be a Poisson measure on $[0,1] \times \mathbb{R}^2$ with the intensity measure $\gamma\ell \otimes \mathcal{L}(Z, \lambda^{1/2} Z)$, where $\mathcal{L}(Z) = Q$ and $\gamma > 0$ is the intensity parameter of N in (2.4). Consider the Markov kernel $K_1 : [0,\infty]^4 \times (\mathcal{B}([0,\infty)^3) \otimes \mathcal{M}_2) \to [0,1]$ defined by

$$
K_1\big[(\omega_1,\omega_2,\omega_3,\omega_4), \cdot\,\big] \;=\; \varepsilon_{\xi(\omega)} \otimes \begin{cases} \varepsilon_{\nu_0}\,, & \omega_4 = \infty\,, \\ \mathcal{L}(M_{\omega_4} | M_{\omega_4} \neq \nu_0)\,, & \omega_4 < \infty\,. \end{cases}
$$

Observe that $K_2 K_1 Q_\theta = \mathcal{L}_\theta(G, h)$ for all $\theta \in [0,\infty)^4$ in view of (2.4). Hence $\delta(\mathcal{F}, \mathcal{E}_h) = 0$ by (A.2).

To summarize, we have shown that $\mathcal{E}_h$ is equivalent to $\mathcal{F}$ in deficiency. By the similar arguments one can show that $\Delta(\mathcal{F}, \widehat{\mathcal{E}}_h) = 0$. $\square$

4.5.2. *Proof of Proposition 2.2.* (i) Let $H_n = H_n^{(0)} : [0,\infty)^4 \to [0,\infty)^4$ be as defined in (2.11)–(2.12) and define $\bar{H}_n : [0,\infty)^3 \to M := \{(x_1,x_2,x_3) \in [0,\infty)^2 \times (0,1] \,:\, x_1 \geq x_2\}$ by

$$
\bar{H}_n(h_0,\beta,\alpha) = \big(h_{0,n}(h_0,\beta,\alpha,0), \beta_n(h_0,\beta,\alpha,0), \alpha_n(h_0,\beta,\alpha,0)\big)\,,
$$

$h_0, \beta, \alpha \in [0,\infty)$. Then $H_n : [0,\infty)^3 \to M \times [0,\infty)$ and $\bar{H}_n : [0,\infty)^3 \to M$ are both bijections with corresponding inverse functions $H_n^{-1} : M \times [0,\infty) \to [0,\infty)^4$ and $\bar{H}_n^{-1} : M \to [0,\infty)^3$, respectively. Define $\widetilde{H}_n : \mathbb{R}^3 \to [0,\infty)^3$ and $\widehat{H}_n : \mathbb{R}^4 \to [0,\infty)^4$ by $\widetilde{H}_n(x_1,x_2,x_3) = \bar{H}_n^{-1}(|x_1| \vee |x_2|, |x_2|, |x_3| \wedge 1)$ and $\widehat{H}_n(x_1,x_2,x_3,x_4) = H_n^{-1}(|x_1| \vee |x_2|, |x_2|, |x_3| \wedge 1, |x_4|)$ for $x_1,x_2,x_3,x_4 \in \mathbb{R}$ with $x_3 \neq 0$.

In the sequel we write $x = (x(k))_{0 \leq k \leq n}$ for a generical element of $\mathbb{R}^{n+1}$. Fix $n \geq 5$. Let $M_{0,n} \subseteq [\mathbb{R}^{n+1}]^2$ be the set of all (x,y) such that both $y(0) \neq y(1)$ and $y(1) \neq y(2)$ are in place. By employing the convention $\inf \emptyset = \infty$ define $T_n : [\mathbb{R}^{n+1}]^2 \to \{1, \ldots, n+1\}$ by

$$
T_n(x,y) \;=\; \inf\{1 \leq k \leq n :\; x(k) \neq x(k-1)\} \wedge 1 \qquad (x,y) \in [\mathbb{R}^{n+1}]^2\,.
$$

Let S_n be the set of all $(x,y) \in [\mathbb{R}^{n+1}]^2$ with $3 \le T(x,y) \le n-2$ such that $x(T) = x(T+1) = x(T+2)$. Consider the subset $M_{T,n} \subseteq S_n$ of all $(x,y) \in [\mathbb{R}^{n+1}]^2$ such that both $y(T) \neq y(T+1)$ and $y(T+1) \neq y(T+2)$ are satisfied.

For all $n \ge 5$ we define a mapping $X_n : [\mathbb{R}^{n+1}]^2 \to [0,\infty]^4$ as follows: fix $(x,y) \in [\mathbb{R}^{n+1}]^2$ and set $T = T_n(x,y)$. If $(x,y) \in S_n \cap M_{0,n}$ then set

$$\begin{aligned} X_n(x,y) &= \widehat{H}_n\Big(y(0), \frac{y(1)^2 - y(0)y(2)}{y(1)-y(0)}, \frac{y(2)-y(1)}{y(1)-y(0)}, \\ &\qquad \frac{y(T)}{[x(T)-x(T-1)]^2} - \frac{y(1)^2 - y(0)y(2) + y(T-1)[y(2)-y(1)]}{[y(1)-y(0)][x(T)-x(T-1)]^2}\Big). \end{aligned}$$

If $(x,y) \in M_{T,n} \backslash M_{0,n}$ then set

$$\begin{aligned} &X_n(x,y) = \\ &\widehat{H}_n\Big(y(0)\,,\ \frac{y(T+1)^2 - y(T)y(T+2)}{y(T+1)-y(T)}\,,\ \frac{y(T+2)-y(T+1)}{y(T+1)-y(T)}, \\ &\quad \frac{y(T)}{[x(T)-x(T-1)]^2} - \frac{y(T+1)^2 - y(T)y(T+2) + y(T-1)[y(T+2)-y(T+1)]}{[y(T+1)-y(T)][x(T)-x(T-1)]^2}\Big). \end{aligned}$$

If $(x,y) \in S_n \backslash (M_{0,n} \cup M_{T,n})$ and $y(T) \neq y(T-1)$

$$X_n(x,y) = \Big(y(0), 0, 0, \frac{|y(T)-y(T-1)|}{(x(T)-x(T-1))^2}\Big).$$

If $(x,y) \in S_n \backslash (M_{0,n} \cup M_{T,n})$ and $y(T) = y(T-1)$ then set $X_n(x,y) = (|y(0)|, \infty, \infty, 0)$. If $(x,y) \in M_{0,n} \backslash S_n$ and $T = n+1$ then set

$$X_n(x,y) = \Big(\widetilde{H}_n\Big[y(0), \frac{y(1)^2 - y(0)y(2)}{y(1)-y(0)}, \frac{y(2)-y(1)}{y(1)-y(0)}\Big], \infty\Big),$$

Otherwise, set $X_n(x,y) = (|y(0)|, \infty, \infty, \infty)$.

Recall that both $G_n = (G_{n,k})_{0\le k\le n}$ and $h_n = (h_{n,k})_{0\le k\le n}$ are defined by (2.3) via (2.11)–(2.12). For $n \ge 5$ the mapping $X_n : [\mathbb{R}^{n+1}]^2) \to [0,\infty]^4$ is well-defined and $\mathcal{B}([\mathbb{R}^{n+1}]^2)/\mathcal{B}([0,\infty]^4)$-measurable. Recall that $Q(\{0\}) = 0$ and, thus,

$$\mathcal{L}_\theta^{X_n}(G_n, h_n) = \begin{cases} q_{1,n}\varepsilon_{(h_0,\beta,\alpha,\infty)} + q_{2,n}\,\varepsilon_\theta & \\ \quad + (1-q_{1,n}-q_{2,n})\varepsilon_{(h_{0,n}(\theta),\infty,\infty,\infty)}\,, & \theta \notin \Theta_e\,, \\ (1-q_{2,n})\,\varepsilon_{(h_0,\infty,\infty,\infty)} + q_{2,n}\varepsilon_\theta\,, & \theta \in \Theta_e\,, \\ & h_0 > 0\,,\ \lambda > 0\,, \\ (1-q_{2,n})\,\varepsilon_{(h_0,\infty,\infty,\infty)} + q_{2,n}\varepsilon_{(h_0,\infty,\infty,0)}\,, & \theta \in \Theta_e\,, \\ & h_0 > 0\,,\ \lambda = 0\,, \\ \varepsilon_{(0,\infty,\infty,\infty)}\,, & \theta \in \Theta_e\,,\ h_0 = 0\,, \end{cases}$$

for all $n \geq 5$, $\theta = (h_0, \beta, \alpha, \lambda) \in [0,\infty)^4$, where we set $q_{1,n} = (1-p_n)^n$ and $q_{2,n} = (1-p_n)^2[1-p_n - p_n(1-p_n)][1-(1-p_n)^{n-4}]$.

On the other hand, define a mapping $\xi_n = (\xi_{1,n}, \ldots, \xi_{4,n}) : [0,\infty]^4 \to [0,\infty)^4$ as follows: let $\omega = (\omega_1, \ldots, \omega_4) \in [0,\infty]^4$. If $\omega \in [0,\infty)^4$ then set $\xi_n(\omega) = H_n(\omega)$. If $\omega \in [0,\infty)^3 \times \{\infty\}$ then set $\xi_n(\omega) = (\bar{H}_n(\omega_1, \omega_2, \omega_3), 0)$. If $\omega \in [0,\infty) \times (\{\infty\} \times [0,\infty] \cup [0,\infty] \times \{\infty\}) \times [0,\infty)$ then set $\xi_n(\omega) = (\omega_1, 0, 1, \omega_4)$. If $\omega \in [0,\infty) \times (\{\infty\} \times [0,\infty] \cup [0,\infty] \times \{\infty\}) \times \{\infty\}$ then set $\xi_n(\omega) = (\omega_1, 0, 1, 0)$. Otherwise, set $\xi(\omega) = 0$. Define a Markov kernel $K_{1,n} : [0,\infty]^4 \times \mathcal{B}([0,\infty)^3 \times [\mathbb{R}^n]^2) \to [0,1]$ by

$$\begin{aligned} K_{1,n}[\omega, \cdot\,] \;&=\; \varepsilon_{(\xi_{n,1}(\omega),\, \xi_{n,2}(\omega),\, \xi_{n,3}(\omega))} \\ &\quad\otimes \begin{cases} \varepsilon_0\,, & \omega_4 = \infty\,, \\ \mathcal{L}((Z_{n,k})_k, (\xi_4(\omega) Z_{n,k}^2)_k | (Z_{n,k})_k \neq 0)\,, & \omega_4 < \infty\,, \end{cases} \end{aligned}$$

for $\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \in [0,\infty]^4$, where $Z_n = (Z_{n,k})_k$ is the random vector with the distribution as specified by (2.1).

Also, let $K_{2,n} : [0,\infty)^3 \times [\mathbb{R}^n]^2 \times \mathcal{B}([\mathbb{R}^{n+1}]^2) \to [0,1]$ be the Markov kernel defined by the deterministic mapping $(\xi_1, \xi_2, \xi_3, z_1, z_2) \mapsto (x, y)$ where, recursively, we set $x(0) = 0$ and $y(0) = \xi_1$ and, for $1 \leq k \leq n$,

$$x(k) = x(k-1) + y^{1/2}(k-1)\, z_1(k)\,, \quad y(k) = \xi_2 + y(k-1)(\xi_3 + z_2(k))\,.$$

For $n \geq 5$ let $\mathcal{F}_n = ([0,\infty]^4, \mathcal{B}([0,\infty]^4), (\mathcal{L}_\theta^{X_n}(G_n, h_n))_{\theta \in [0,\infty)^4})$. By construction we have $\delta(\mathcal{E}_{h,n}, \mathcal{F}_n) = 0$ by means of (A.2). For all $n \geq 5$, observe that

$$\begin{aligned} \delta(\mathcal{F}_n, \mathcal{E}_{h,n}) \;&\leq\; \sup_{\theta \in [0,\infty)^3} \|\mathcal{L}_\theta(G_n, h_n) - K_{2,n} K_{1,n} \mathcal{L}_\theta^{X_n}(G_n, h_n)\| \\ &\leq\; |1 - q_{1,n} - q_{2,n}| + |1 - (1-p_n)^n - q_{2,n}|\,, \end{aligned}$$

and, thus, $\mathcal{E}_{h,n}$ is strongly asymptotically equivalent to $\mathcal{F}_n$ as $n \to \infty$, by means of (A.2) and (2.2). By (A.4), $\mathcal{F}_n$ converges strongly to the experiment $\mathcal{F}$ in the assertion of Proposition 2.1, completing the proof of (i).

(ii) This follows from the same arguments as in (i). □

4.5.3. *Proof of Proposition 2.3.* (i) Define $X, X_n : [0,\infty]^4 \to [0,\infty]^4$ as follows: if $\omega = (\omega_1, \ldots, \omega_4) \in [0,\infty)^3 \times \{\infty\}$ such that $\omega_1 \omega_3 = \omega_2$ then set $X(\omega) = (\omega_1, \infty, \infty, \infty)$; otherwise, set $X(\omega) = \omega$. If $\omega = (\omega_1, \ldots, \omega_4) \in [0,\infty)^3 \times \{\infty\}$ such that $\omega_1 n(1-e^{-\omega_3/n}) = \omega_2$ then set $X_n(\omega) = (\omega_1, \infty, \infty, \infty)$; otherwise, set $X_n(\omega) = \omega$, $n \in \mathbb{N}$.

By definition, the deficiency is nondecreasing in the parameter set with respect to set-inclusions. Further, we have $\widehat{Q}_\theta^X = Q_\theta$ and $\widehat{Q}_\theta^{X_n} = Q_{\theta,n}$ for all $n \in \mathbb{N}$ and, thus, by (A.2), that $\delta(\widehat{\mathcal{F}}(\Theta), \mathcal{F}(\Theta)) \le \delta(\widehat{\mathcal{F}}, \mathcal{F}) = 0$ and $\delta(\widehat{\mathcal{F}}(\Theta), \mathcal{F}_n(\Theta)) \le \delta(\widehat{\mathcal{F}}, \mathcal{F}_n) = 0$ for all $n \in \mathbb{N}$, completing the proof of (i).

(ii) Firstly, assume that Θ satisfies (2.22) for all $x > 0$. Without generality we may assume that $\Theta \subseteq [0,\infty)^4$ is a finite set [cf. Theorem 51.4 in [19]]. Define Ω_Θ to be the set of all $\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \in (0,\infty) \times \{\infty\}^2 \times \{0,\infty\}$ such that $(\omega_1, \beta, \alpha, \lambda) \in (\Theta \cap \Theta_e) \backslash \widehat{\Theta}_e$ for some $(\beta, \alpha, \lambda) \in [0,\infty)^3$. If $\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega_\Theta$ then it follows from (2.22) that the corresponding pair $(\beta, \alpha) = (\beta(\omega_1), \alpha(\omega_1)) \in [0,\infty)^2$ is uniquely determined by ω_1. Hence we may define a mapping $Y : [0,\infty]^4 \to [0,\infty]^4$ as follows: if $\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega_\Theta$ then we set $Y(\omega) = (\omega_1, \beta(\omega_1), \alpha(\omega_1), \omega_4)$; otherwise, if $\omega \in [0,\infty]^4 \backslash \Omega_\Theta$ then we set $Y(\omega) = \omega$. As both Θ and, thus, Ω_Θ are finite sets the mapping Y is $\mathcal{B}([0,\infty]^4)/\mathcal{B}([0,\infty]^4)$-measurable. In view of (2.22), note that $Q_\theta^Y = \widehat{Q}_\theta$ for all $\theta \in \Theta$ and, thus, $\delta(\mathcal{F}(\Theta), \widehat{\mathcal{F}}(\Theta)) = 0$ by (A.2).

Secondly assume that (2.22) is violated. Then there exist $h_0 > 0$ and $\theta_1 = (h_0, \beta_1, \alpha_2, \lambda_1) \in \Theta \cap \Theta_e \cap \widehat{\Theta}_e^C$ and $\theta_2 = (h_0, \beta_2, \alpha_2, \lambda_2) \in \Theta \cap \Theta_e$ such that $(\beta_1, \alpha_1) \ne (\beta_2, \alpha_2)$.

Consider $\Theta_0 = \{\theta_1, \theta_2\}$ and the decision space $D = \{(\beta_1, \alpha_1), (\beta_2, \alpha_2)\}$, endowed with the discrete topology. For $\theta = (h_0, \beta, \alpha, \lambda) \in \Theta$ consider (continuous and bounded) loss functions $W_\theta : D \to \mathbb{R}$, where, for $x = (x_1, \ldots, x_4) \in [0,\infty]^4$, we set $W_\theta(x) = 1 - 1_{\{(\beta,\alpha)\}}(x_2, x_3)$. Further, we define a Markov kernel $\widehat{\rho} : [0,\infty]^4 \times \mathcal{B}(D) \to [0,1]$, where, for $x \in [0,\infty]^4$ and $B \in \mathcal{B}(D)$, we set

$$\widehat{\rho}(x, B) = \begin{cases} \epsilon_{(\beta_1,\alpha_1)}(B), & \text{if } x \in (0,\infty) \times \{\beta_1\} \times \{\alpha_1\} \times [0,\infty), \\ \epsilon_{(\beta_2,\alpha_2)}(B), & \text{otherwise.} \end{cases}$$

Then we have $\int W_{\theta_i}(x) \widehat{\rho}(\omega, dx)\, \widehat{Q}_{\theta_i}(d\omega) = 0$ for $i = 1, 2$. On the other hand, any Markov kernel $\rho : [0,\infty]^4 \times \mathcal{B}(D) \to [0,1]$ is of form $\rho(\omega, B) = p(\omega)\epsilon_{(\beta_1,\alpha_1)}(B) + (1 - p(\omega))\epsilon_{(\beta_2,\alpha_2)}(B)$ where $p : [0,\infty]^4 \to [0,1]$ is Borel and $\omega \in [0,\infty]^4$ and $B \in \mathcal{B}(D)$. It is easy to see that for such a Markov kernel ρ there exists a Markov kernel $\bar{\rho} : [0,\infty]^4 \times \mathcal{B}(D) \to [0,1]$ such that, both

$$\begin{aligned} \int W_{\theta_1}(x) \rho(\omega, dx)\, Q_{\theta_1}(d\omega) &\ge e^{-\gamma}(1 - p(h_0, \infty, \infty, \infty)), \quad \text{and} \\ \int W_{\theta_2}(x) \rho(\omega, dx)\, Q_{\theta_2}(d\omega) &\ge e^{-\gamma} p(h_0, \infty, \infty, \infty). \end{aligned}$$

In view of (A.1) we, thus, have $\delta(\mathcal{F}(\Theta), \widehat{\mathcal{F}}(\Theta)) \ge \delta(\mathcal{F}(\Theta_0), \mathcal{F}(\Theta_0)) \ge e^{-\gamma}/2$, which completes the proof of (ii).

(iii) This follows by the same arguments as in (ii). □

APPENDIX

We collect necessary facts regarding Le Cam's distance in deficiency. The reader is referred to Le Cam [10] and Le Cam & Young [11] and Strasser's monograph [19] for unexplained notations not encountered in this section. Let Θ be a nonempty set and $(E, \mathcal{A})$ be a measurable space and $(P_\theta)_{\theta\in\Theta}$ be a family of probability measures on $\mathcal{A}$. Then the triplet $\mathcal{E} = (E, \mathcal{A}, (P_\theta)_{\theta\in\Theta})$ is called a (statistical) experiment. Consider two experiments $\mathcal{E}_i = (E_i, \mathcal{A}_i, (P_{i,\theta})_{\theta\in\Theta})$, $i = 1, 2$, indexed by Θ. A decision problem is a triple (Θ, D, W) where D is a topological space and $W = (W_\theta)_{\theta\in\Theta}$ is a loss function $W_\theta : D \to \mathbb{R}$, $\theta \in \Theta$. Let $\|W_\theta\|_\infty = \sup_{d\in D} |W_\theta(d)|$. Also, let $\epsilon \geq 0$. Then $\mathcal{E}_1$ is called ϵ-deficient with respect to $\mathcal{E}_2$, shortly $\mathcal{E}_1 \supseteq_\epsilon \mathcal{E}_2$, iff for all decision problems (Θ, D, W), with W being continuous and bounded, and all $\beta_2 \in \mathcal{B}(\mathcal{E}_2, D)$ there exists $\beta_1 \in \mathcal{B}(\mathcal{E}_1, D)$ such that

$$\beta_1(W_\theta, P_{1,\theta}) \leq \beta_2(W_\theta, P_{2,\theta}) + \epsilon\|W_\theta\|_\infty\,, \qquad \theta \in \Theta\,,$$

where $\mathcal{B}(\mathcal{E}_i, D)$ (i=1, 2) is the space of generalized decision functions [cf. [19], Definition 42.2]. The deficiency of $\mathcal{E}_1$ with respect to $\mathcal{E}_2$ is the number

$$\delta(\mathcal{E}_1, \mathcal{E}_2) \;=\; \inf\{\epsilon > 0 \;:\; \mathcal{E}_1 \supseteq_\epsilon \mathcal{E}_2\}\,. \tag{A.1}$$

The relation $\mathcal{E}_1 \supseteq_\epsilon \mathcal{E}_2$ is interpreted in the following sense: we have $\mathcal{E}_1 \supseteq_\epsilon \mathcal{E}_2$ if $\mathcal{E}_1$ is more informative than $\mathcal{E}_2$ uniformly over all decision problems with continuous and bounded loss functions up to some error ϵ. Two experiments $\mathcal{E}_1$ and $\mathcal{E}_2$ are called *equivalent in deficiency* iff $\mathcal{E}_1 \supseteq_0 \mathcal{E}_2$ and $\mathcal{E}_2 \supseteq_0 \mathcal{E}_1$.

Recall that [cf. [19], Lemma 55.4 & Remark 55.6(2)]

$$\delta(\mathcal{E}_1, \mathcal{E}_2) \;=\; \inf_K \sup_{\theta\in\Theta} \|P_{2,\theta} - KP_{1,\theta}\|\,, \tag{A.2}$$

with an infimum now taken over all Markov kernels $K : E_1 \times \mathcal{E}_2 \to [0, 1]$.

Le Cam's distance between $\mathcal{E}_1$ and $\mathcal{E}_2$ is a pseudometric on the space of all experiments indexed by Θ [cf. [19], Corollary 59.6], defined by setting,

$$\Delta(\mathcal{E}_1, \mathcal{E}_2) = \max\{\delta(\mathcal{E}_1, \mathcal{E}_2), \delta(\mathcal{E}_2, \mathcal{E}_1)\}. \tag{A.3}$$

If $(E_1, \mathcal{A}_1) = (E_2, \mathcal{A}_2)$, then we have [cf. [19], Corollary 59.6]:

$$\Delta(\mathcal{E}_1, \mathcal{E}_2) \leq \sup_{\theta\in\Theta} \|P_{1,\theta} - P_{2,\theta}\|\,. \tag{A.4}$$

Clearly, if $\mathcal{E}_1$ and $\mathcal{E}_2$ are two experiments indexed by the same Θ then $\mathcal{E}_1$ is equivalent to $\mathcal{E}_2$ in deficiency if and only if $\Delta(\mathcal{E}_1, \mathcal{E}_2) = 0$. Let $\mathcal{E}$, $\mathcal{E}_n$,

$\mathcal{F}_n$, $n \in \mathbb{N}$, be experiments, all indexed by Θ. Then we say that $\mathcal{E}_n$ converges (strongly) in deficiency, or, $\mathcal{E}_n$ and $\mathcal{F}_n$ are (strongly) asymptotically equivalent in deficiency iff $\Delta(\mathcal{E}_n, \mathcal{E}) \to 0$ and $\Delta(\mathcal{E}_n, \mathcal{F}_n) \to 0$ as $n \to \infty$.

For $\emptyset \neq \Theta_0 \subseteq \Theta$ we employ the notation $\mathcal{E}(\Theta_0) = (E, \mathcal{A}, (P_\theta)_{\theta \in \Theta_0})$ for corresponding subexperiments of $\mathcal{E} = (E, \mathcal{A}, (P_\theta)_{\theta \in \Theta})$. We refer to weak convergence and weak asymptotically equivalence in deficiency iff, for all nonempty and finite $\Theta_0 \subseteq \Theta$, the corresponding subexperiments converges strongly and are strongly asymptotically equivalent in deficiency, respectively.

REFERENCES

[1] Billingsley, P. (2004) *Convergence of Probability Measures.* Wiley, New York.

[2] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Political Economy* **81**, 637–659.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *J. Econ.* **31**, 307–327.

[4] Brown, L.D., and Low, M. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24**, 2384–2398.

[5] Brown, L.D., Wang, Y. and Zhao, L.H. (2003). On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion model. *Statistica Sinica* **13**, 993–1013.

[6] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the United Kingdom inflation. *Econometrica* **50**, 987–1007.

[7] Grama, I.O. and Neumann, M.H. (2006). Asymptotic equivalence of nonparametric autoregression and nonparametric regression. *Ann. Statist.* **34**, 1701–1732.

[8] Janssen, A., Milbrodt, H. and Strasser, H. (1985) *Infinitely Divisible Statistical Experiments.* Lecture Notes in Statistics 27. Springer, Berlin.

[9] Kallsen, J. and Vesenmayer, B. (2007). COGARCH as a continuous time limit of GARCH(1,1). *Preprint. Munich University of Technology.*

[10] Le Cam, L. (1986) *Asymptotic Methods in Statistical Decision Theory.* Springer, New York.

[11] Le Cam, L. and Yang, G.L. (1990) *Asymptotics in Statistics. Some Basic Concepts.* Springer, New York.

[12] Milstein, G. and Nussbaum, M. (1998). Diffusion approximation for nonparametric autoregression. *Prob. Theor. Rel. Fields* **112**, 535–543.

[13] Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *J. Appl. Probab.* **41**, 601–622.

[14] Maller, R., Müller, G. and Szimayer, A. (2008). GARCH Modelling in Continuous Time for Irregularly Spaced Time Series Data. *Bernoulli* **14**, 519–542.

[15] Maller, R., Müller, G. and Szimayer, A. (2008). Ornstein-Uhlenbeck Processes and Extensions. To apear in: *T.G. Andersen, R.A. Davis, J.-P. Kreiss, T. Mikosch (Eds.) Handbook of Financial Time Series*, Springer, New York.

[16] Merton, R. (1973). The theory of rational option pricing. *Bell J. Econom. Management Sci.* **4**, 141–183.

[17] Nelson, D.B. (1990). ARCH models as diffusion approximations. *Journal of Econometrics* **45**, 7–38.

[18] Nussbaum, M. (1996). Asymtotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24**, 2399–2430.

[19] Strasser, H. (1985) *Mathematical Theory of Statistics.* deGruyter, Berlin.

[20] Reiss, R.-D. (1993) *A Course on Point Processes.* Springer, New York.

[21] Wang, Y. (2002). Asymptotic nonequivalence of GARCH models and diffusions. *Ann. Statist.* **30**, 754–783.

Boris Buchmann
School of Mathematics
Monash University
Clayton Campus, VIC 3800
Australia
Phone (61-3) 9905 9760
FAX (61-3) 9905 9520
E-mail: Boris.Buchmann@sci.monash.edu.au

Gernot Müller
Centre of Mathematics
Munich University of Technology
85747 Garching
Germany
Phone (49-89) 289 17441
FAX (49-89) 289 17435
E-mail: mueller@ma.tum.de