KYBERNETIKA — VOLUME 50 (2014), NUMBER 6, PAGES 883–895

AN EFFICIENT ESTIMATOR FOR GIBBS RANDOM FIELDS

Martin Janžura

An efficient estimator for the expectation $\int f \, \mathrm{d}P$ is constructed, where P is a Gibbs random field, and f is a local statistic, i. e. a functional depending on a finite number of coordinates. The estimator coincides with the empirical estimator under the conditions stated in Greenwood and Wefelmeyer [6], and covers the known special cases, namely the von Mises statistic for the i.i.d. underlying fields and the case of one-dimensional Markov chains.

Keywords: Gibbs random field, efficient estimator, empirical estimator

Classification: 62F12, 62M40

1. INTRODUCTION

The expectation of a local statistic defined on a stationary random field can be estimated from observations in a large window with the aid of the empirical estimator, which is given as the average of the function values over all shifts inside the window. Under appropriate conditions on the underlying random field, the estimator is consistent and asymptotically normal. But, in general, it may not be efficient. For the exponential type distributions of random processes and fields, which coincide with the Gibbs random fields as studied within the frame of statistical physics as well as image processing, the problem of efficiency was addressed by Greenwood and Wefelmeyer [6]. Under the regularity condition, which only makes the efficiency study meaningful, and which for the Gibbs fields reads as the restriction to the Dobrushin's uniqueness region, they showed that the empirical estimator is efficient if and only if the considered local statistic belongs (in a special wide sense) to the same space as the potential of the underlying Gibbs random field.

In the present paper we adopt most of the results obtained by Greenwood and Wefelmeyer [6], but we shall extend them substantially by a general construction of the efficient estimator. Roughly speaking, the empirical estimate $\int f \, \mathrm{d}\widehat{P}$ must be substituted by the estimate in the form $\int f \, \mathrm{d}P^{\widehat{g}}$, where $\widehat{g}$ is the (efficient) estimate of the unknown potential. We shall also prove that these two estimates coalesce under the conditions stated in Greenwood and Wefelmeyer [6]. In addition, we shall show that the known special results concerning the i.i.d. and Markov cases, are also covered by our general result.

DOI: 10.14736/kyb-2014-6-0883

Besides the main reference, the efficiency issues are based on the pioneering Hájek's results [8] and the comprehensive book by Bickel et al. [1]. The regularity LAN condition is due to Janžura [10]. The results on Gibbs random fields are partly from Künsch' paper [14] and Georgii's book [4], the statistical analysis results from Janžura [12].

For the sake of simplicity, within this paper we consider only the finite state space and the finite-dimensional (parametric) case. With an appropriate effort, the generalization is possible.

2. GIBBS RANDOM FIELDS

Let X_0 be a finite set, and, for some $d \geq 1$, let $\mathbf{Z}^d$ be the d-dimensional integer lattice, and $\mathcal{X} = X_0^{\mathbf{Z}^d}$ the corresponding product space. For $V \subset \mathbf{Z}^d$ we denote by $\mathcal{F}_V$ the σ-algebra generated by the projection $\mathrm{Proj}_V : \mathcal{X} \to X_0^V$, and by $\mathcal{B}_V$ the set of (bounded) $\mathcal{F}_V$-measurable functions. Further, we shall write $x_V = \mathrm{Proj}_V(x)$ for $x \in \mathcal{X}$ and $V \subset \mathbf{Z}^d$, and we shall denote by $\mathcal{V} = \{W \subset \mathbf{Z}^d; 0 < |W| < \infty\}$ the system of finite non-void subsets of $\mathbf{Z}^d$.

We denote by $\mathcal{P}$ the set of all probability measures on $\mathcal{X}$ with the product σ-algebra $\mathcal{F}$. Further, by $\mathcal{P}_S$ we denote the set of *stationary* (translation invariant) probability measures, i. e. $\mathcal{P}_S = \{P \in \mathcal{P}; P = P \circ \tau_j^{-1} \text{for every } j \in \mathbf{Z}^d\}$, where $\tau_j : \mathcal{X} \to \mathcal{X}$ is for every $j \in \mathbf{Z}^d$ the corresponding shift operator defined by $[\tau_j(x)]_t = x_{j+t}$ for every $t \in \mathbf{Z}^d$, $x \in \mathcal{X}$. Finally, by $\mathcal{P}_E$ we denote the set of *ergodic* probability measures, i. e. $\mathcal{P}_E = \{P \in \mathcal{P}_S; P(F) \in \{0,1\}$ for every $F \in \mathcal{E}\}$ where $\mathcal{E} = \{F \in \mathcal{F}; F = \tau_j^{-1}(F)$ for every $j \in \mathbf{Z}^d\}$.

The functions from $\mathcal{B} = \bigcup_{V \in \mathcal{V}} \mathcal{B}_V$ will be quoted as (finite range) *potentials*. For a potential $\Phi \in \mathcal{B}_V$, $V \in \mathcal{V}$, we define the *Gibbs specification* as the family of probability kernels $\Pi^\Phi = \{\Pi_\Lambda^\Phi\}_{\Lambda \in \mathcal{V}}$ where

$$\Pi_\Lambda^\Phi(x_\Lambda | x_{\Lambda^c}) = Z_\Lambda^\Phi(x_{\Lambda^c}) \exp\left\{ \sum_{j \in \Lambda - V} \Phi \circ \tau_j(x) \right\}$$

with the normalizing constant

$$Z_\Lambda^\Phi(x_{\Lambda^c}) = \sum_{y_\Lambda \in X_0^\Lambda} \exp\left\{ \sum_{j \in \Lambda - V} \Phi \circ \tau_j(y_\Lambda, x_{\Lambda^c}) \right\}$$

for every $\Lambda \in \mathcal{V}$. Note that $\Lambda - V = \{j \in \mathbf{Z}^d; (j + V) \cap \Lambda \neq \emptyset\}$. We do not emphasize the range V in the notation. It is an inseparable part of the potential Φ, and it is only important that, due to the finiteness of V, the set $\Lambda - V$ is also finite. In fact, the set $\Lambda - V$ is the minimal set of indices for which the term $\Phi \circ \tau_j(x)$ actually depends on x_Λ, and we could sum over any larger set.

A probability measure $P \in \mathcal{P}$ is a *Gibbs distribution* (*Gibbs random field*) with the potential $\Phi \in \mathcal{B}$ if

$$P_{\Lambda|\Lambda^c}(x_\Lambda | x_{\Lambda^c}) = \Pi_\Lambda^\Phi(x_\Lambda | x_{\Lambda^c}) \quad \text{a. s.} \quad [P]$$

for every $\Lambda \in \mathcal{V}$. The set of such P's will be denoted by $G(\Phi)$, while the set of stationary (resp. ergodic) Gibbs distributions will be denoted by $G_S(\Phi) = G(\Phi) \cap \mathcal{P}_S$ (resp.

$G_E(\Phi) = G(\Phi) \cap \mathcal{P}_E$). In general $G_E(\Phi) \neq \emptyset$. We may say that the (first-order) phase transition occurs if $|G(\Phi)| > 1$. Then, some elements are non-ergodic, and some even may not be translation invariant (stationary) although the specification is so. For a detailed treatment and the examples see, e. g., Georgii [4, Chapter 6.2]. Unfortunately, for the efficiency study the phase transitions mean the non-smoothness and non-regularity, and, therefore, must be excluded – see Section 5 below.

Let us end this section with the observation that, since for $\Phi \in \mathcal{B}_V$ we have $\Pi_\Lambda^\Phi \in \mathcal{B}_{\Lambda+V-V}$ for every $\Lambda \in \mathcal{V}$, the above defined Gibbs random fields obey the (spatial) Markov property.

3. EQUIVALENCE OF POTENTIALS

Besides the phase transitions, there is another kind of non-uniqueness that can complicate the treatment, namely the possible equivalence of potentials. Two potentials Φ, $\Psi \in \mathcal{B}$ are *equivalent*, we write $\Phi \approx \Psi$, if $G(\Phi) = G(\Psi)$. There is a couple of equivalent characterizations (see Georgii [4] or Janžura [11]), e. g., $\Phi \approx \Psi$ iff $\Pi_{\{0\}}^\Phi = \Pi_{\{0\}}^\Psi$. For our purposes, there will be important the following one:

$$\Phi \approx \Psi \quad \text{iff} \quad \int \Phi \,\mathrm{d}P = \int \Psi \,\mathrm{d}P + c \quad \text{for every } P \in \mathcal{P}_S \text{ with some fixed constant } c.$$

The equivalence can appear very easily, e. g.,

$$\Phi \approx \Phi + g - g \circ \tau_j + c \quad \text{for some } g \in \mathcal{B},\ j \in \mathbf{Z}^d, \text{ and a constant } c.$$

From the statistical analysis point of view the equivalence of potentials means breaking the basic identifiability condition, and, therefore, must be unambiguously avoided.

A rather standard way consists in dealing with the equivalence classes instead of the particular potentials. But we prefer to restrict our considerations to a rich enough subclass of mutually nonequivalent potentials. Such subclass should contain representatives of all equivalence classes. That can be arranged by dealing with the so called *vacuum* potentials. Let us fix some state $b \in X_0$, quoted as the vacuum state. For any $V \in \mathcal{V}$, let us denote $\mathcal{B}_V^b = \{\Phi \in \mathcal{B}_V; \Phi(x_V) = 0 \text{ if } x_t = b \text{ for some } t \in V\}$. Further, in order to avoid equivalence by shifting, let us introduce $\mathcal{V}^0 = \{V \in \mathcal{V}; \min_{t\in V}\{t\} = 0\}$ where the minimum is with respect to some fixed complete (e. g., the lexicographical) ordering. Now, for a finite subsystem $\mathcal{A} \subset \mathcal{V}^0$ we set $\mathcal{B}_\mathcal{A}^b = \{\Phi \in \mathcal{B}; \Phi = \sum_{A\in\mathcal{A}} \Phi_A, \Phi_A \in \mathcal{B}_A^b$ for every $A \in \mathcal{A}\}$, and, consequently, $\mathcal{B}^b = \bigcup_{\mathcal{A}\subset\mathcal{V}^0, |\mathcal{A}|<\infty} \mathcal{B}_\mathcal{A}^b$ will be our class of vacuum potentials.

Proposition 3.1. i) For Φ, $\Psi \in \mathcal{B}^b$ it holds:

$$\Phi \approx \Psi \quad \text{iff} \quad \Phi = \Psi.$$

ii) For every $\Phi \in \mathcal{B}$ there exists $\Psi \in \mathcal{B}^b$ such that $\Phi \approx \Psi$.

Proof. i) We may observe that for Φ, $\Psi \in \mathcal{B}^b$ we have $\Phi - \Psi \in \mathcal{B}^b$, and

$$\Phi \approx \Psi \quad \text{iff} \quad \Phi - \Psi \approx 0.$$

That simplifies a bit the tedious calculation. For some $\Phi \in \mathcal{B}^b_{\mathcal{A}}$ with $\overline{A} = \bigcup_{A\in\mathcal{A}} A$ we must deduce $\Phi \equiv 0$ from the condition $\sum_{j\in -\overline{A}} \Phi \circ \tau_j \in \mathcal{B}_{(\overline{A}-\overline{A})\setminus\{0\}}$ (that is equivalent to $\Phi \approx 0$) by a proper sequence of substituting. Let $A^1 \in \mathcal{A}$ be minimal in the sense: $(A^1)^c \cap (A+t) \neq \emptyset$ for every $A \in \mathcal{A} \setminus \{A^1\}$ and $t \in \mathbf{Z}^d$. Then

$$\textstyle\sum_{j\in -\overline{A}} \Phi \circ \tau_j(x_{A^1}, b_{(\overline{A}-\overline{A})\setminus A^1}) = \Phi_{A^1}(x_{A^1}),$$

and since by the assumption Φ_{A^1} must not depend on x_0 we may always substitute $x_0 = b$ and obtain $\Phi_{A^1} \equiv 0$. Then we repeat the consideration with $\mathcal{A} \setminus \{A^1\}$, etc., and finally we obtain $\Phi_A \equiv 0$ for every $A \in \mathcal{A}$.

ii) The statement follows from Theorem 2.35 b) in Georgii [4] or by direct calculations with the aid of Möbius formula for constructing the vacuum potentials. □

4. EMPIRICAL RANDOM FIELDS

For a fixed configuration $x \in \mathcal{X}$ and some $\Lambda \in \mathcal{V}$ we define a probability measure $\widehat{P}^\Lambda_x$ by

$$\int \Phi \, \mathrm{d}\widehat{P}^\Lambda_x = |\Lambda|^{-1} \sum_{t\in\Lambda} \Phi \circ \tau_t(x) \quad \text{for every } \Phi \in \mathcal{B}.$$

Such probability distributions will be quoted as empirical random fields. On the other hand, for fixed $\Phi \in \mathcal{B}_V$ we have $\int \Phi \, \mathrm{d}\widehat{P}^\Lambda_\bullet \in \mathcal{B}_{V+\Lambda}$, which means that for specifying the marginal distribution $\widehat{P}^\Lambda_x / \mathcal{F}_V$ we actually need to know only $x_{V+\Lambda} \in X_0^{V+\Lambda}$.

Now, let us consider a sequence of properly growing subsets $\{V_n\}_{n=1}^\infty$ in $\mathbf{Z}^d$, e. g., the cubes $V_n = [-n, n]^d$ for simplicity.

Then, for a fixed potential $\Phi \in \mathcal{B}$ let us denote by

$$q(\Phi) = \lim_{n\to\infty} |V_n|^{-1} \log Z^\Phi_{V_n}(x_{V_n{}^c})$$

the *pressure* corresponding to the potential Φ. The limit exists uniformly for every $x \in \mathcal{X}$, e. g., by Theorem 15.30 in Georgii [4].

In general, the pressure q is a convex continuous function on $\mathcal{B}$, strictly convex on $\mathcal{B}^b$, and even strongly convex on every compact subset of $\mathcal{B}^b$ (see Dobrushin and Nahapetian [3]).

Further, for every $P \in \mathcal{P}_S$ we may define the *entropy rate*

$$H(P) = \lim_{n\to\infty} |V_n|^{-1} \int [-\log P_{V_n}(x_{V_n})] \, \mathrm{d}P(x)$$

where the limit exists by Theorem 15.12 again in Georgii [4].

Proposition 4.1. For $P \in \mathcal{P}_S$, the following statements are equivalent:

i) $P \in G_S(\Phi)$;

ii) $|V_n|^{-1} \log P_{V_n}(x_{V_n}) - \int \Phi \, \mathrm{d}\widehat{P}^{V_n}_x + q(\Phi) \rightrightarrows 0$ for $n \to \infty$;

iii) $H(P) = q(\Phi) - \int \Phi \, \mathrm{d}P$.

Proof. While i) $\Rightarrow$ ii) and ii) $\Rightarrow$ iii) are rather straightforward, the proof of iii) $\Rightarrow$ i) needs a rather sophisticated construction (see, e. g., Georgii [4, Theorem 15.37] or Janžura [13]). □

5. UNIQUENESS REGION

For every $t \in \mathbf{Z}^d$ let us denote

$$\gamma_t(\Phi) = \frac{1}{2} \sup \left\{ \sum_{x_0 \in X_0} \left| \Pi^{\Phi}_{\{0\}}(x_0 | y_{\mathbf{Z}^d \setminus \{0\}}) - \Pi^{\Phi}_{\{0\}}(x_0 | z_{\mathbf{Z}^d \setminus \{0\}}) \right| ;\ y_s = z_s \text{ for } s \neq t \right\}.$$

If $\gamma(\Phi) = \sum_{t \in \mathbf{Z}^d} \gamma_t(\Phi) < 1$, the potential is said to satisfy the Dobrushin's uniqueness condition. Let us denote by $\mathcal{D} = \{\Phi \in \mathcal{B}; \gamma(\Phi) < 1\}$ the Dobrushin's uniqueness region. For every $\Phi \in \mathcal{D}$ it holds $G(\Phi) = G_S(\Phi) = G_E(\Phi) = \{P^{\Phi}\}$. For more details see Dobrushin [2], Künsch [14], or Georgii [4].

6. FINITE-DIMENSIONAL (PARAMETRIC) SPACE OF POTENTIALS

From now, we shall consider the linear subspace $\mathcal{H} \subset \mathcal{B}^b$ of potentials, where

$$\mathcal{H} = \mathit{Lin}\left(\Phi^1, \ldots, \Phi^K\right)$$

and $\Phi^1, \ldots, \Phi^K \in \mathcal{B}^b$ are vacuum potentials. Then the potentials are mutually non-equivalent, i. e.

$$\Phi^{\theta} = \sum_{i=1}^{K} \theta_i\, \Phi^i \approx 0 \quad \text{iff} \quad \theta = 0.$$

Let us denote $\Theta = \{\theta \in R^K,\ \Phi^{\theta} \in \mathcal{D}\}$ and $\mathcal{P}_{\Theta} = \{P^{\theta}\}_{\theta \in \Theta}$, where $P^{\theta} = P^{\Phi^{\theta}}$. Similarly, we shall write $q(\theta)$ for $q(\Phi^{\theta})$, or cov^{θ} for $\mathrm{cov}^{P^{\theta}}$, and denote

$$B_{\theta}(f^1, f^2) = \sum_{t \in \mathbf{Z}^d} \mathrm{cov}^{\theta}(f^1, f^2 \circ \tau_t)$$

for every $f^1,\ f^2 \in \mathcal{B}$.

Moreover, let us introduce the transform

$$U : \Theta \to R^K$$

defined by $U(\theta) = \int \mathbf{\Phi}\, \mathrm{d}P^{\theta}$ where $\mathbf{\Phi} = (\Phi^1, \ldots, \Phi^K)$.

Proposition 6.1. For $q : \Theta \to R$ and $U : \Theta \to R^K$ it holds:

i) q is a smooth, strongly convex function.

ii) $B_{\theta}(f^1, f^2)$ exists for any $f^1,\ f^2 \in \mathcal{B}$; $B_{\theta}(f, f) > 0$ iff $f \not\approx 0$.

iii) $\nabla q(\theta) = U(\theta)$,
$\nabla^2 q(\theta) = \nabla U(\theta) = B_{\theta}(\mathbf{\Phi}, \mathbf{\Phi})$.

iv) U is a regular mapping, $B_\bullet(f^1, f^2)$ is continuous for fixed $f^1, f^2 \in \mathcal{B}$.

Proof. The results follow by Künsch [14], together with Dobrushin and Nahapetian [3].

Theorem 6.2. For $\theta \in \Theta$, $f \in \mathcal{B}$ it holds:

i) $\int f \,\mathrm{d}\widehat{P}_\bullet^{V_n} \to \int f \,\mathrm{d}P^\theta$, for $n \to \infty$ a.s. $[P^\theta]$.

ii) $|V_n|^{\frac{1}{2}} \left[\int f \,\mathrm{d}\widehat{P}_\bullet^{V_n} - \int f \,\mathrm{d}P^\theta\right] \Rightarrow \mathcal{N}(0,\, B_\theta(f,f))$ for $n \to \infty$ in distribution $[P^\theta]$.

iii) $|V_n|^{\frac{1}{2}} \left[\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n} - \int \mathbf{\Phi} \,\mathrm{d}P^\theta - \ell_n^\theta\right] \to 0$ for $n \to \infty$ in probability $[P^\theta]$,

where $\ell_n^\theta = \nabla\left[|V_n|^{-1} \log P_{V_n}^\theta\right]$ is the score function.

iv) $\log \dfrac{P_{V_n}^{\theta + |V_n|^{-\frac{1}{2}}\alpha}}{P_{V_n}^\theta} - |V_n|^{\frac{1}{2}} \left[\int \Phi^\alpha \,\mathrm{d}\widehat{P}_\bullet^{V_n} - \int \Phi^\alpha \,\mathrm{d}P^\theta\right] + \dfrac{1}{2} B_\theta(\Phi^\alpha, \Phi^\alpha) \to 0$

for $n \to \infty$ in probability $[P^\theta]$.

Proof. While i) is just the ergodic theorem, ii) is the CLT due to Künsch[14]. For iii) and iv) (local asymptotic normality) see Janžura [10]. □

7. EFFICIENT ESTIMATOR

For a fixed local statistic $f \in \mathcal{B}_W$, $W \in \mathcal{V}$, let us set

$$T_n = \int f \,\mathrm{d}P^{\widehat{\theta}_n} \quad \text{with} \quad \widehat{\theta}_n = U^{-1}\left(\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n}\right)$$

whenever $\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n} \in U(\Theta)$. We claim that T_n is an *efficient estimator* for $\int f \,\mathrm{d}P^\theta$, $\theta \in \Theta$.

Since U is regular and therefore $U(\Theta) \subset R^K$ is open, and $\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n} \to \int \mathbf{\Phi} \,\mathrm{d}P^\theta$ a.s. $[P^\theta]$ by the ergodic theorem, we have $\widehat{\theta}_n$, and, consequently, T_n defined with a probability tending to 1. For $\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n} \notin U(\Theta)$ we may set T_n arbitrary. In fact, due to Proposition 6.1, we have

$$\widehat{\theta}_n = \arg\min_{\theta \in R^K} \left\{ q(\theta) - \int \mathbf{\Phi}^\theta \,\mathrm{d}\widehat{P}_\bullet^{V_n} \right\}$$

which is well defined whenever $\int \mathbf{\Phi} \,\mathrm{d}\widehat{P}_\bullet^{V_n} \in \bigcup_{\alpha \in R^K} \left\{\int \mathbf{\Phi} \,\mathrm{d}P; P \in G_S(\Phi^\alpha)\right\}$, and we need the formal definition only outside that set.

Remark 7.1. For every $i = 1, \dots, K$ we have $\Phi^i \in \mathcal{B}^b_{\overline{A}_i}$ with $\overline{A}_i = \bigcup_{A \in \mathcal{A}^i} A$. Thus $\Phi^\alpha \in \mathcal{B}_{\overline{\overline{A}}}$ with $\overline{\overline{A}} = \bigcup_{i=1}^K \overline{A}_i$, and, in order to have $\int \mathbf{\Phi}\, \mathrm{d}\widehat{P}^{V_n}_\bullet$ well defined, we need the data to be observed from the region $V_n + \overline{\overline{A}} = \overline{V}_n$.

Theorem 7.2. For every $\theta \in \Theta$ we have

$$|V_n|^{\frac{1}{2}} \left[T_n - \int f \,\mathrm{d}P^\theta - \int g^\theta_f \,\mathrm{d}\widehat{P}^{V_n}_\bullet \right] \to 0 \quad \text{for } n \to \infty \text{ in probability } [P^\theta],$$

where

$$g^\theta_f = B_\theta(f, \mathbf{\Phi})\, B_\theta(\mathbf{\Phi}, \mathbf{\Phi})^{-1} \left(\mathbf{\Phi} - \int \mathbf{\Phi}\, \mathrm{d}P^\theta \right).$$

P r o o f. Since $T_n = \int f\, \mathrm{d}P^{U^{-1}(\int \mathbf{\Phi}\, \mathrm{d}\widehat{P}^{V_n}_\bullet)}$ and $\int f\, \mathrm{d}P^\theta = \int f\, \mathrm{d}P^{U^{-1}(\int \mathbf{\Phi}\, \mathrm{d}P^\theta)}$, we obtain by the first order expansion with some (random) $\widetilde{\varepsilon}_n \in [0,1]$:

$$\begin{aligned} T_n - \int f\, \mathrm{d}P^\theta &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int f\, \mathrm{d}P^{U^{-1}(\int f(\varepsilon\, \mathrm{d}\widehat{P}^{V_n}_\bullet + (1-\varepsilon)\, \mathrm{d}P^\theta))} \Big|_{\varepsilon = \widetilde{\varepsilon}_n} \\ &= B_{\widetilde{\theta}_n}(f, \mathbf{\Phi})\, B_{\widetilde{\theta}_n}(\mathbf{\Phi}, \mathbf{\Phi})^{-1} \left(\int \mathbf{\Phi}\, \mathrm{d}\widehat{P}^{V_n}_\bullet - \int \mathbf{\Phi}\, \mathrm{d}P^\theta \right) \end{aligned}$$

with $\widetilde{\theta}_n = U^{-1}\left(\int f(\widetilde{\varepsilon}_n\, \mathrm{d}\widehat{P}^{V_n}_\bullet + (1-\widetilde{\varepsilon}_n)\, \mathrm{d}P^\theta) \right)$. Thus

$$\begin{aligned} & |V_n|^{\frac{1}{2}} \left[T_n - \int f\, \mathrm{d}P^\theta - \int g^\theta_f\, \mathrm{d}\widehat{P}^{V_n}_\bullet \right] \\ = & \left[B_{\widetilde{\theta}_n}(f, \mathbf{\Phi})\, B_{\widetilde{\theta}_n}(\mathbf{\Phi}, \mathbf{\Phi})^{-1} - B_\theta(f, \mathbf{\Phi})\, B_\theta(\mathbf{\Phi}, \mathbf{\Phi})^{-1} \right] \cdot \\ & \cdot |V_n|^{\frac{1}{2}} \left(\int \mathbf{\Phi}\, \mathrm{d}\widehat{P}^{V_n}_\bullet - \int \mathbf{\Phi}\, \mathrm{d}P^\theta \right), \end{aligned}$$

where, with the aid of results of the preceding section, namely Proposition 4.1 and Theorem 6.2, the first term tends to zero a. s. and in probability $[P^\theta]$ due to the ergodic theorem and continuity of B_θ. The second term tends to $\mathcal{N}(0, B_\theta(\mathbf{\Phi}, \mathbf{\Phi}))$ in distribution by the CLT, which completes the proof. □

Remark 7.3. We may observe

$$B_\theta(f - g^\theta_f, \mathbf{\Phi}) = B_\theta(f, \mathbf{\Phi}) - B_\theta(f, \mathbf{\Phi})\, B_\theta(\mathbf{\Phi}, \mathbf{\Phi})^{-1} B_\theta(\mathbf{\Phi}, \mathbf{\Phi}) = 0.$$

Thus, g^θ_f is the canonical gradient, i. e. the projection of the local statistic f into the space $\mathcal{H}$ equipped with the scalar product $B_\theta(\cdot, \cdot)$ (cf. Greenwood and Wefelmeyer [6] or Bickel et al. [1, Section 3.3]).

Remark 7.4. Due to Theorem 6.2 iii) we can obtain the same limit theorem with the *efficient influence function*

$$B_\theta(f,\mathbf{\Phi})\,B_\theta(\mathbf{\Phi},\mathbf{\Phi})^{-1}\,\ell_n^\theta$$

(as required, e. g., in Bickel et al. [1, Section 2.3]) instead of $\int g_f^\theta \,\mathrm{d}\widehat{P}_\bullet^{V_n}$.

Before proving the efficiency of the estimator, let us introduce the relevant definitions.

Definition 7.5. The estimator $\widetilde{T}_n$ is called *regular*, if

$$|V_n|^{\frac{1}{2}}\left(\widetilde{T}_n - \int f\,\mathrm{d}P^{\theta+|V_n|^{-\frac{1}{2}}\alpha}\right) \Rightarrow \mathcal{L}(\theta) \quad \text{for } n\to\infty \text{ in distribution } \left[P^{\theta+|V_n|^{-\frac{1}{2}}\alpha}\right],$$

where $\mathcal{L}(\theta)$ is some limiting distribution.

Proposition 7.6. The estimator

$$T_n = \int f\,\mathrm{d}P^{\widehat{\theta}_n} \quad \text{with} \quad \widehat{\theta}_n = U^{-1}\left(\int \mathbf{\Phi}\,\mathrm{d}\widehat{P}_\bullet^{V_n}\right)$$

is regular with $\mathcal{L}(\theta) = \mathcal{N}\left(0, B_\theta(f,\mathbf{\Phi})\,B_\theta(\mathbf{\Phi},\mathbf{\Phi})^{-1}B_\theta(\mathbf{\Phi},f)\right)$.

P r o o f. For $\theta\in\Theta_{\mathcal{D}}$ let us denote $L_n^\theta = \log\dfrac{P_{V_n}^{\theta+|V_n|^{-\frac{1}{2}}\alpha}}{P_{V_n}^\theta}$ and $T_n^\theta = |V_n|^{\frac{1}{2}}\left[T_n - \int f\,\mathrm{d}P^\theta\right]$. Then due to Theorem 6.2 iv) and Theorem 7.2 we obtain

$$(T_n^\theta, L_n^\theta)^\top \Rightarrow \mathcal{N}_2(\mu,\Sigma) \quad \text{in distribution} \quad [P^\theta],$$

where

$$\mu = \begin{pmatrix} 0 \\ -\frac{1}{2}B_\theta(\Phi^\alpha,\Phi^\alpha)\end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} B_\theta(g_f^\theta,g_f^\theta) & B_\theta(g_f^\theta,\Phi^\alpha) \\ B_\theta(\Phi^\alpha,g_f^\theta) & B_\theta(\Phi^\alpha,\Phi^\alpha)\end{pmatrix}.$$

And by the third LeCam's lemma (see, e. g., Bickel et al. [1, Lemma A.9.3.]) it also holds

$$(T_n^\theta, L_n^\theta)^\top \Rightarrow \mathcal{N}_2(\mu^*,\Sigma) \quad \text{in distribution} \quad [P_{V_n}^{\theta+|V_n|^{-\frac{1}{2}}\alpha}],$$

where

$$\mu^* = \begin{pmatrix} B_\theta(\Phi^\alpha,g_f^\theta) \\ \frac{1}{2}B_\theta(\Phi^\alpha,\Phi^\alpha)\end{pmatrix} \quad .$$

Now, we observe $B_\theta(\Phi^\alpha,g_f^\theta) = B_\theta(\Phi^\alpha,f)$, and

$$|V_n|^{\frac{1}{2}}\left(T_n - \int f\,\mathrm{d}P^{\theta+|V_n|^{-\frac{1}{2}}\alpha}\right) = T_n^\theta + |V_n|^{\frac{1}{2}}\left(\int f\,\mathrm{d}P^\theta - \int f\,\mathrm{d}P^{\theta+|V_n|^{-\frac{1}{2}}\alpha}\right)$$

$$= \quad T_n^\theta - B_{\theta+\epsilon_n|V_n|^{-\frac{1}{2}}\alpha}(\Phi^\alpha,g_f^\theta) \quad \Rightarrow \quad \mathcal{N}_1\left(B_\theta(\Phi^\alpha,f), B_\theta(g_f^\theta,g_f^\theta)\right) - B_\theta(\Phi^\alpha,f)$$

$= \mathcal{N}_1\left(0, B_\theta(g_f^\theta,g_f^\theta)\right)$, where the limit is in distribution $[P_{V_n}^{\theta+|V_n|^{-\frac{1}{2}}\alpha}]$, and $\epsilon_n\in[0,1]$.

Finally, since

$$B_\theta(g_f^\theta,g_f^\theta) = B_\theta(f,\mathbf{\Phi})\,B_\theta(\mathbf{\Phi},\mathbf{\Phi})^{-1}B_\theta(\mathbf{\Phi},f),$$

we have the claimed result. □

Definition 7.7. A regular estimator T_n will be called *efficient* if

$$|V_n|^{-\frac{1}{2}} \left(T_n - \int f \, \mathrm{d}P^\theta \right) \Rightarrow \mathcal{N}\left(0, B_\theta(f, \mathbf{\Phi}) \, B_\theta(\mathbf{\Phi}, \mathbf{\Phi})^{-1} B_\theta(\mathbf{\Phi}, f)\right)$$
$$\text{for } n \to \infty \text{ in distribution } [P^\theta]$$

(see also Definition 1 in Section 2.3 in Bickel et. al. [1]).

Remark 7.8. The above definition is fully justified by the Hájek's convolution theorem (cf., e. g. Theorem 2.3.1 in Bickel et. al. [1]), which guaranties the maximum possible concentration of the estimate in the vicinity of the true value.

In fact, the *asymptotic linearity* of the proposed estimator, as proved in Theorem 7.2, is crutial for both the regularity and the efficiency of the estimator (see also Proposition 7.6 in Greenwood and Wefelmeyer [6]).

Theorem 7.9. The estimator T_n is efficient for $\int f \, \mathrm{d}P^\theta$, $\theta \in \Theta$.

Proof. Directly from Theorem 7.2, Proposition 3.1, and the above definitions. □

8. EXAMPLES

8.1. Efficiency of the empirical estimator

By the definition of the transform U we have

$$\int \Phi \, \mathrm{d}P^{\widehat{\theta}_n} = \int \Phi \, \mathrm{d}\widehat{P}_\bullet^{V_n}$$

for every $\Phi \in \{\Phi^1, \ldots, \Phi^K\}$ and, consequently, for every $\Phi \in \mathcal{H}$. Therefore, whenever $f \in \mathcal{H}$, the empirical estimator $\int f \, \mathrm{d}\widehat{P}_\bullet^{V_n}$ is efficient, which fully agrees with the main Theorem in Greenwood and Wefelmeyer[6]. But, moreover, suppose only $f \approx \Phi_0$ for some $\Phi_0 \in \mathcal{H}$. Then, by Section 3, we have $\int (f - \Phi_0) \, \mathrm{d}P = c$ for every $P \in \mathcal{P}_S$. Moreover, since $B_\theta(f - \Phi_0, f - \Phi_0) = 0$, by the CLT (Theorem 6.2 iii)) we now have

$$|V_n|^{\frac{1}{2}} \int (f - \Phi_0 - c) \, \mathrm{d}\widehat{P}_\bullet^{V_n} \to 0 \quad \text{for } n \to \infty \text{ in probability } [P^\theta].$$

Then we may write

$$\begin{aligned} \int f \, \mathrm{d}\widehat{P}_\bullet^{V_n} &= \int (f - \Phi_0 - c) \, \mathrm{d}\widehat{P}_\bullet^{V_n} + c + \int \Phi_0 \, \mathrm{d}\widehat{P}_\bullet^{V_n} \\ &= \int (f - \Phi_0 - c) \, \mathrm{d}\widehat{P}_\bullet^{V_n} + c + \int \Phi_0 \, \mathrm{d}P^{\widehat{\theta}_n} = \int (f - \Phi_0 - c) \, \mathrm{d}\widehat{P}_\bullet^{V_n} + \int f \, \mathrm{d}P^{\widehat{\theta}_n} \end{aligned}$$

and since $\int (f - \Phi_0 - c) \, \mathrm{d}\widehat{P}_\bullet^{V_n} = o_{P^\theta}\left(|V_n|^{-\frac{1}{2}}\right)$, the empirical estimator $\int f \, \mathrm{d}\widehat{P}_\bullet^{V_n}$ still remains efficient. The correction term occurs due to the "non-stationarity" of the empirical distribution $\widehat{P}_\bullet^{V_n}$. By a slight modification it could be made stationary, but we would loose the unbiasness: $\int \int \Phi \, \mathrm{d}\widehat{P}_\bullet^{V_n} \mathrm{d}P^\theta = \int \Phi \, \mathrm{d}P^\theta$ which is crucial for the limit theorems.

Remark 8.1. Thanks to the above result we have also directly the efficiency of the parameter estimate

$$\widehat{\theta}_n = U^{-1}\left(\int \mathbf{\Phi}\,\mathrm{d}\widehat{P}_{\bullet}^{V_n}\right)$$

which is a regular transform of the efficient empirical estimate $\int \mathbf{\Phi}\,\mathrm{d}\widehat{P}_{\bullet}^{V_n}$. We may observe that the estimate $\widehat{\theta}_n$ is secondary, derived from the primary estimate $\int \mathbf{\Phi}\,\mathrm{d}\widehat{P}_{\bullet}^{V_n}$. The definition of the estimate is natural, it corresponds to the relation between the theoretical terms, and it also agrees with the (approximate) maximum likelihood estimate (see, e. g., Janžura [12]). Anyhow, also the efficiency study for the estimate $\widehat{\theta}_n$ would be hardly possible without studying the efficiency of the estimate $\int \mathbf{\Phi}\,\mathrm{d}\widehat{P}_{\bullet}^{V_n}$.

8.2. von Mises statistic for i.i.d. variables

For $X_0 = \{x_0, \dots, x_K\}$ with $b = x_0$ we consider

$$\Phi_i \in \mathcal{B}_{\{0\}}^b \quad \text{for } i = 1, \dots, K,$$

defined by $\Phi_i(x) = \delta(x, x_i)$ for $x \in X_0$.

Then $P^\theta \in G(\Phi^\theta)$ is the product distribution with $P_{\{t\}}^\theta(x_i) = \frac{e^{\theta_i}}{1+\sum_{j=1}^K e^{\theta_j}}$ for $i \neq 0$ and $P_{\{t\}}^\theta(x_0) = \frac{1}{1+\sum_{j=1}^K e^{\theta_j}}$ for every $t \in \mathbf{Z}^d$.

With a given collection of data $\widehat{x}_{V_n} \in X_0^{V_n}$ we have $\int \Phi_i\,\mathrm{d}\widehat{P}_{\bullet}^{V_n} = \frac{1}{|V_n|}\sum_{t\in V_n}\delta(x_i, \widehat{x}_t)$ for $i = 1, \dots, K$. Providing $\int \Phi^i \mathrm{d}\widehat{P}_{\bullet}^{V_n} > 0$ for $i = 1, \dots, K$, and $\sum_{i=1}^K \int \Phi^i \mathrm{d}\widehat{P}_{\bullet}^{V_n} < 1$ we obtain

$$\widehat{\theta}_n^i = \log \frac{\int \Phi^i \mathrm{d}\widehat{P}_{\bullet}^{V_n}}{1 - \sum_{j=1}^K \int \Phi^j \mathrm{d}\widehat{P}_{\bullet}^{V_n}} \quad \text{for } i = 1, \dots, K.$$

Now, for a local statistic $f \in \mathcal{B}_\Lambda$, $\Lambda \in \mathcal{V}$, we obtain by Theorem 6.2 the efficient estimate of $\int f\,\mathrm{d}P^\theta$ in the form

$$\begin{aligned}\int f\,\mathrm{d}\bigotimes_{t\in\Lambda} P_{\{t\}}^{\widehat{\theta}_n} &= \sum_{y_\Lambda \in X_0^\Lambda} f(y_\Lambda) \prod_{t\in\Lambda} \frac{1}{|V_n|} \sum_{\ell\in V_n} \delta(y_t, \widehat{x}_\ell) \\ &= \frac{1}{|V_n|^{|\Lambda|}} \sum_{z_\Lambda \in (\widehat{x}_{V_n})^\Lambda} f(z_\Lambda)\end{aligned}$$

which is nothing else but the von Mises statistic, i. e. the average over all data permutations.

8.3. Markov chains

Now, suppose $d = 1$, $X_0 = \{x_0, x_1, \dots, x_L\}$, $K = L(L+1)$, and

$$\mathbf{\Phi} = \left\{\Phi^i, \Phi^{ik}; i,\, k = 1, \dots, L\right\}$$

where

$$\begin{aligned}\Phi^i(x) &= \delta(x,x_i) \qquad \text{for } x\in X_0,\\ \Phi^{ik}(x,y) &= \delta(x,x_i)\cdot\delta(y,x_k) \quad \text{for } x\in X_0,\ y\in X_1.\end{aligned}$$

Then, providing the empirical distribution is again positive, we have $P^{\widehat{\theta}_n}$ given as the distribution of an ergodic Markov chain with the initial distribution $P^{\widehat{\theta}_n}_{\{0\}}(x_i) = \int \Phi^i \,\mathrm{d}P_0^{V_n}$ for $i = 1,\dots,L$, and the transition probabilities $P^{\widehat{\theta}_n}_{\{1|0\}}(y_j|x_i) = \frac{\int \Phi^{ij}\mathrm{d}\widehat{P}_\bullet^{V_n}}{\int \Phi^{i}\mathrm{d}\widehat{P}_\bullet^{V_n}}$ for $i,\, j = 1,\dots,L$. Thus, the result agrees with that of Greenwood and Wefelmeyer [5].

8.4. Reversible Markov chains

Let us suppose, moreover, that there is a finite group of transforms $\Gamma = \{\gamma : \mathcal{H}\to\mathcal{H}\}$ where again $\mathcal{H} = Lin(\Phi^i,\Phi^{ik}; i,\, k = 1,\dots,L)$. Let us denote by

$$\mathcal{H}_\Gamma = \{\Phi\in\mathcal{H}; \gamma\circ\Phi = \Phi \text{ for every } \gamma\in\Gamma\}$$

the set of invariant potentials with respect to the group Γ. For all P^Φ, $\Phi\in\mathcal{H}_\Gamma$, suppose, in addition, $\int\Psi\,\mathrm{d}P^\Phi = \int \gamma\circ\Psi\,\mathrm{d}P^\Phi$ for every $\gamma\in\Gamma$, $\Psi\in\mathcal{H}$. (This is fulfilled, e. g., if $\gamma\circ\Phi = \Phi\circ\widetilde{\gamma}$ with some $\widetilde{\gamma}:\mathcal{X}\to\mathcal{X}$.) Then the efficient estimator for $\int\Psi\,\mathrm{d}P^\Phi$,$\Psi\in\mathcal{H}$, is given by

$$\int \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\gamma\circ\Psi\,\mathrm{d}\widehat{P}_\bullet^{V_n}.$$

Namely, we may write

$$\int\Psi\,\mathrm{d}P^{\widehat{\Phi}^n} = \int\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\gamma\circ\Psi\,\mathrm{d}P^{\widehat{\Phi}^n} = \int\frac{1}{|\Gamma|}\sum\gamma\circ\Psi\,\mathrm{d}\widehat{P}_\bullet^{V_n}$$

since $\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\gamma\circ\Psi\in\mathcal{H}_\Gamma$. Here $\widehat{\Phi}^n\in\mathcal{H}_\Gamma$ is given by $\int\widetilde{\Phi}\,\mathrm{d}P^{\widehat{\Phi}^n} = \int\widetilde{\Phi}\,\mathrm{d}\widehat{P}_\bullet^{V_n}$ for every $\widetilde{\Phi}\in\mathcal{H}_\Gamma$.

Now, for the above Markov chains example, let us set $\Gamma = \{\gamma_0,\gamma_1\}$ where γ_0 is the identity, i. e., $\gamma_0\circ\widetilde{\Phi} = \widetilde{\Phi}$ for every $\widetilde{\Phi}\in\mathcal{H}$, and

$$\begin{aligned}\gamma_1\circ\widetilde{\Phi}(x_0) &= \widetilde{\Phi}(x_0) \qquad \text{for } \widetilde{\Phi}\in\mathcal{B}_{\{0\}},\\ \gamma_1\circ\widetilde{\Phi}(x_0,x_1) &= \widetilde{\Phi}(x_1,x_0) \qquad \text{for } \widetilde{\Phi}\in\mathcal{B}_{\{0,1\}} \text{ is the reversion.}\end{aligned}$$

Then all the assumptions are satisfied and we obtain the efficient estimator for $f\in\mathcal{B}_{\{0,1\}}$ in the form

$$\frac{1}{2}\,\frac{1}{n-1}\sum_{i=1}^{n-1}[f(\widehat{x}_i,\widehat{x}_{i+1}) + f(\widehat{x}_{i+1},\widehat{x}_i)]$$

just as in Greenwood and Wefelmeyer [5].

9. A NOTE ON THE PERFORMANCE

Unfortunately, for the Gibbs random fields, there is a lack of analytic formulas for calculating the expectations $\int f \, \mathrm{d}P^{\Phi}$. The only feasible way consists of substituting the theoretic terms by their simulated counterparts.

The problem occurs both for calculating the estimate $\widehat{\theta}_n = U^{-1}\left(\int \mathbf{\Phi} \, \mathrm{d}\widehat{P}_{\bullet}^{V_n}\right)$, and the final evaluation of the expectation $\int f \, \mathrm{d}P^{\widehat{\theta}_n}$.

The estimate $\widehat{\theta}_n$ can be calculated by an iterative procedure

$$\widehat{\theta}_n(k+1) = \widehat{\theta}(k) - C_k \left(\int \mathbf{\Phi} \, \mathrm{d}P^{\widehat{\theta}_n(k)} - \int \mathbf{\Phi} \, \mathrm{d}\widehat{P}_{\bullet}^{V_n} \right)$$

where $C_k \to 0$ for $k \to \infty$, but the term $\int \mathbf{\Phi} \, \mathrm{d}P^{\widehat{\theta}^n(k)}$ must be for every actual k substituted by $\int \mathbf{\Phi} \, \mathrm{d}\widehat{P}^M_{x_M(k)}$, where $M \in \mathcal{V}$ is as large as possible and $x_M(k) \in X_0^M$ is simulated with the distribution $P^{\widehat{\theta}_n(k)}$, e. g., by a Markov Chain Monte Carlo procedure (cf., e. g., Younes [15]).

At the stopping time k^*, the procedure returns both the estimate $\widehat{\theta}_n(k^*)$ and the simulated data $x_M(k^*)$. Thus, the expectation $\int f \, \mathrm{d}P^{\widehat{\theta}_n}$ is finally substituted by

$$\int f \, \mathrm{d}\widehat{P}^M_{x_M(k^*)}.$$

The whole procedure remains optimal from the asymptotic point of view, but for a fixed size data $x_{\overline{V}_n}$ it is an open question whether it gives really a "better" estimate than the empirical estimator $\int f \, \mathrm{d}\widehat{P}_{\bullet}^{V_n}$. Apparently, if $f \in \mathcal{B}_V$, with $V \in \mathcal{V}$ much "larger" to compare with A where $\mathbf{\Phi} \in \mathcal{B}_A$, or, at least, with much larger diameter, our approach may be strongly recommended since the empirical mean $\int f \, \mathrm{d}\widehat{P}^M_{x_M(k^*)}$ is an average of much larger number of values than $\int f \, \mathrm{d}\widehat{P}^{V_n}_{x_{\overline{V}_n}}$, and would be very likely more "precise". And the whole approach can be understood as a bootstrap-like method.

ACKNOWLEDGEMENT

Supported by the grant No. P402/12/G097.

(Received January 14, 2014)

REFERENCES

[1] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, and J. A. Wellner: Efficient and Adaptive Estimation for Semiparametric Models. Johns Hopkins University Press, Baltimore 1993.

[2] R. L. Dobrushin: Prescribing a system of random variables by conditional distributions. Theor. Probab. Appl. *15* (1970), 458–486.

[3] R. L. Dobrushin and B. S. Nahapetian: Strong convexity of the pressure for lattice systems of classical physics (in Russian). Teoret. Mat. Fiz. *20* (1974), 223–234.

[4] H. O. Georgii: Gibbs Measures and Phase Transitions. De Gruyter Studies in Mathematics *9*, De Gruyter, Berlin 1988.

[5] P. E. Greenwood and W. Wefelmeyer: Efficiency of empirical estimators for Markov Chains. Ann. Statist. *23* (1995), 132–143.

[6] P. E. Greenwood and W. Wefelmeyer: Characterizing efficient empirical estimators for local interaction Gibbs fields. Stat. Inference Stoch. Process. *2* (1999), 119–134.

[7] K. Gross: Absence of second-order phase transitions in the Dobrushin uniqueness region. J. Statist. Phys. *25* (1981), 57–72.

[8] J. Hájek: A characterization of limiting distributions of regular estimates. Wahrsch. Verw. Gebiete *14* (1970), 323–330.

[9] M. Janžura: Statistical analysis of Gibbs random fields. In: Trans. 10th Prague Conference on Inform. Theory, Stat. Dec. Functions, Random Processes, Prague 1984, pp. 429–438.

[10] M. Janžura: Local asymptotic normality for Gibbs random fields. In: Proc. Fourth Prague Symposium on Asymptotic Statistics (P. Mandl and M. Hušková, eds.), Charles University, Prague 1989, pp. 275–284.

[11] M. Janžura: Asymptotic behaviour of the error probabilities in the pseudo-likelihood ratio test for Gibbs–Markov distributions. In: Prof. Asymptotic Statistics (P. Mandl and M. Hušková, eds.), Physica–Verlag, Heidelberg 1994, pp. 285–296.

[12] M. Janžura: Asymptotic results in parameter estimation for Gibbs random fields. Kybernetika *33* (1997), 2, 133–159.

[13] M. Janžura: On the concept of the asymptotic Rényi distances for random fields. Kybernetika *35* (1999), 3, 353–366.

[14] H. Künsch: Decay of correlations under Dobrushin's uniqueness condition and its applications. Comm. Math. Phys. *84* (1982), 207–222.

[15] L. Younes: Parameter inference for imperfectly observed Gibbsian fields. Probab. Theory Rel. Fields *82* (1989), 625–645.

Martin Janžura, Institute of Information Theory and Automation — Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: janzura@utia.cas.cz