Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

On convergence rates of suprema in the presence of non–negligible trends

Valentin Konakov

submitted: 27th May 1994

Central Economics–Mathematical Institute
Academy of Sciences
Krasikova str., 32
117418 Moscow
Russia

Preprint No. 103
Berlin 1994

1991 Mathematics Subject Classification. Primary 62 G 07, Secondary 62 G 20, 62 M 40.
Key words and phrases. Kernel estimation, smoothing parameter, Gaussian fields, Laplace type integral, Leray–Gel'fand differential form.

This paper was written while the author was visiting the Institute of Applied Analysis and Stochastics and with support by the Deutsche Forschungsgemeinschaft, SFB 373.

Edited by
Institut für Angewandte Analysis und Stochastik (IAAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2004975
e-mail (X.400): c=de;a=d400;p=iaas-berlin;s=preprint
e-mail (Internet): preprint@iaas-berlin.d400.de

On convergence rates of suprema in the presence of non–negligible trends[1]

Valentin Konakov
Central Economics-Mathematical Institute,
Academy of Sciences, Russia.

To S.A. Aivazyan in his Sixtieth Anniversary

Abstract

We investigate the convergence rates for the maximal deviation distribution of kernel estimates from the true density. The convergence rates for related Gaussian fields are also investigated. We consider the optimal choice of the smoothing parameter in the sense of Konakov and Piterbarg (1994) and in doing so we take into account a non–negligible trend. It is shown that the convergence rates depend on the asymptotic behaviour of the Laplace type integrals over a small neighbourhood of the manifold of points at which the trend attains its maximal value. Using integration over the level sets (Leray–Gel'fand differential forms) it is proved that the convergence rates are tipically logarithmically slow, even if the rates are to be uniform over as few as three points. Some improved approximations with power rates of convergence are also obtained.

1 Introduction and Main Results

Throughout the paper

$$y_1, y_2, \ldots, y_n, \ldots$$

are independent and identically distributed random vectors with probability law μ which has a density f with respect to Lebesgue measure. We assume that

D: the density f is strictly positive function on the closed unit cube I^d.

We recall the Parzen-Rosenblatt estimator of the f

$$f_n(x) = (nh_n^d)^{-1} \sum_{i=1}^{n} K(h_n^{-1}(x - y_i)).$$

[1]This research was supported by Russian Foundation of Fundamental Researches grant 93-011-1437.

1991 Mathematics Subject Classification. Primary 62 G 07; Secondary 62 G 20, 62 M 40.
Key Words and phrases: Kernel estimation, smoothing parameter, Gaussian fields, Laplace type integral, Leray–Gel'fand differential form.

Also throughout we assume that the kernel K is twice continuously differentiable with compact support and the sequence of smoothing parameters $(h_n)_{n>0}$ satisfies the condition

S: $0 < h_n < 1$ for all $n \geq 1$, $h_n \downarrow 0$ and $nh_n^d \uparrow \infty$ if $n \uparrow \infty$.

We also introduce Rio (1994) condition:

R: $0 < \liminf_{n\to\infty} \log h_n^{-1}/\log n \leq \limsup_{n\to\infty} \log h_n^{-1}/\log n < 1/d$.

Let us introduce the normalized deviation field of the f_n

$$\xi_n(x) = \sqrt{nh_n^d} \cdot \frac{f_n(x) - f(x)}{\sigma\sqrt{f(x)}}, \quad x \in I^d,$$

where $\sigma^2 = \int K^2(y)dy$. Recall the condition **G** from Konakov and Piterbarg (1994).

Let $\mathcal{J}(\beta, L)$, $\beta = p + \alpha$, $0 < \alpha \leq 1$, $p \geq 0$, be the set of functions on I^d, having all partial derivatives up to the order p and such that

$$|(D^\nu f)(x_1) - (D^\nu f)(x_2)| \leq L|x_1 - x_2|^\alpha, \qquad |\nu| = p,$$

where $D^\nu = D_1^{\nu_1}...D_d^{\nu_d}$, $D_j = \frac{\partial}{\partial x_j}$. Denote $\mathcal{J}(\beta) = \cup_{L>0}\mathcal{J}(\beta, L)$. For $f \in \mathcal{J}(\beta)$ we introduce the function

$$G(x) = \frac{1}{\sqrt{f(x)}} \sum_{|\nu|=p} \frac{(D^\nu f)(x)}{\nu!} \int y^\nu K(y)dy, \tag{1}$$

and assume that

G: $G(x)$ does not identically equal zero and there exists a point of its maximum which lies inside the cube I^d:

$$G(x) \not\equiv 0, \ \exists x^+ \in (0,1)^d \text{ such that } |G(x^+)| = \max_{I^d} |G(x)|.$$

Konakov and Piterbarg (1994) investigated the maximal deviation distribution of kernel density estimates $f_n(x)$ when the smoothing parameter h is chosen in an optimal way, so that a "bias term" and a "random error" are balanced. They also studied the maximal deviation distribution for related Gaussian fields $X_T(t)$, $0 \leq t \leq T$. Their results may be roughly classified in two types. The results of the first type establish

(1) the convergence to double exponent $\exp(-e^{-z})$ of the probabilities

$$P_T(z) = P(a(T)(\sup_{0\leq t\leq T} X_T(t) - b(T)) \leq z),$$

as $T \to \infty$, where $X_T(t)$ is a sequence of Gaussian fields with non-vanishing trends in the limit, and

(2) the convergence to double exponent $\exp(-e^{-z})$ of the probabilities

$$P_n(z) = P(a_n(\max_{I^d} \xi_n(x) - b_n) \leq z), \quad \text{as } n \to \infty,$$

where $\xi_n(x)$ is the normalized deviation field $\xi_n(x)$ with an optimally choosen smoothing parameter h_n.

Their results of the second type establish a rather simple sequence of approximating functions $\phi_T(z)$ (resp. $\phi_n(z)$) such that the difference

$$\sup_z |P_T(z) - \phi_T(z)|$$

(resp. $\sup_z |P_n(z) - \phi_n(z)|$) has the order $T^{-\delta}$, $\delta > 0$ (resp. $n^{-\delta}$, $\delta > 0$). The convergence rate to a double exponent $\exp(-e^{-z})$ was studied by Konakov and Piterbarg (1982, 1983, 1984) and later by Hall (1991). The first authors considered the case of vanishing trends in the limit and established that the rate of convergence of $P_T(z)$ and $P_n(z)$ to $\exp(-e^{-z})$ has the exact order $(\log T)^{-1}$ and $(\log n)^{-1}$ respectively. Hall (1991) recently proved that the supremum of centered kernel density estimator may converge to its limit no faster than $(\log T)^{-1}$, if the limit is to be achieved uniformly over three or more distinct points. But he also restricts himself by the unidimensional case and zero trends.

The case of non–vanishing trends in the limit being much more complicated reflects the essence of the matter. The aim of the present paper is to obtain the exact rates of convergence to $\exp(-e^{-z})$ in the situation when a "bias term" and a "random error" are balanced. Nonstationarity, expressed by the presence of non–vanishing trend, appears analytically as an additional multiplier which is a Laplace type integral (in stationary case this integral is equal to 1) over a small neighbourhood of the manifold at which the trend attains its maximal value. The asymptotic behavior of this integral may be obtained by using integration over the level sets (Leray–Gel'fand differential form). As a result we obtain exact rates of convergence to $\exp(-e^{-z})$ for broad spectrum of different cases.

For any $f \in C^2([0,T]^d)$ we denote

$$||f||_0 = \max_{[0,T]^d, \{e\}} |f(t)| \vee |f_e(t)| \vee |f_{ee}(t)|,$$

where the exterior maximum is taken over all $t \in [0,T]^d$ and all directions e in R^d, and $f_e(t)$, $f_{ee}(t)$ denote the first and second derivatives in the direction e.

Now we turn our attention to Gaussian fields. Let $X(t)$, $t \in R^d$, be a twice continuously differentiable Gaussian random field. Let a covariance function of the vector field

$$(X(t),\, X_i(t),\, X_{ij}(t),\, j \geq i),$$

where

$$X_i(t) = \frac{\partial}{\partial t_i} X(t), \qquad X_{ij}(t) = \frac{\partial^2}{\partial t_i \partial t_j} X(t),$$

be nondegenerated for all t. Following Adler (1981), we shall call such fields suitably regular [SR] fields. As in Konakov and Piterbarg (1994) much of our attention will be directed to the model

$$X(t) = \xi(t) + \mu(t), \tag{2}$$

where $\xi(t)$ is a homogeneous isotropic Gaussian centered SR field and $\mu(t)$ is a nonrandom function. Besides, statistical applications of the model lead us to let the trend $\mu(t)$ depend on some parameter T, which will finally appear as a function of sample size n. We assume that

$$\mu_T(\cdot) \in C^3(I^d), \tag{3}$$

and the correlation function $\rho(\tau) = \rho(||t||)$ of the $X(t)$, has the decomposition

$$\rho(\tau) = 1 - \frac{\lambda_2 \tau^2}{2} + \frac{\lambda_4 \tau^4}{4!} + h(\tau), \qquad \tau \to 0, \tag{4}$$

and the derivatives $h^{(k)}(\tau) = o(\tau^{5-k})$ for all $k = 0, 1, 2, 3, 4, 5$.

Let us denote $M_u^+(S)$ and $M_u(S)$ to be numbers of local maxima of the fields $X(t)$, $t \in S$ and $|X(t)|$, $t \in S$, respectively, above the level u. Assume that for the sequence of trends $\{\mu_T(\cdot)\}$ and for the sequence of levels $\{u_T\}$ the following conditions $\mathcal{M}$ and **U** hold true

$\mathcal{M}$: for some positive κ

$$||\mu_T(t) - \lambda_T\, G(T^{-1}t)||_\circ \; = \; O(T^{-\kappa}), \text{ as } T \to \infty$$

where G is defined in (1) and $\lambda_T \to \infty$ in such a way that

$$\lambda_T T^{-\varepsilon} \to 0$$

for any $\varepsilon > 0$.

U: $\lim_{T\to\infty} u_T T^{-\varepsilon} \; = \; 0$ for any $\varepsilon > 0$ and
$\liminf_{T\to\infty} (v_T - G_0) \geq 1$ where $G_0 = sup_{I^d} |G(x)|$, $v_T = u_T/\lambda_T$.

and $G(x)$ satisfies the condition **G**.

Definition 1. We say that a set A is a *maximizing set* for a function f defined on a set Ω if

$$A \; = \; \{x \,:\, x \in \Omega,\; f(x) = \max_{\Omega} f(x)\} \tag{5}$$

Example: The circumference $x_1^2 + x_2^2 = 1$ is the maximizing set for the function $f(x_1, x_2) = -(x_1^2 + x_2^2 - 1)^2$ defined on $\Omega = R^2$.

We consider two different cases corresponding to different groups of conditions **P** and **M**, namely

P
- $G(x) \in C([0,1]^d)$
- maximizing set for $G(x)$, $x \in [0,1]^d$, consists of a single point x^0
- $G(x) \in C^3$ in some neighbourhood of x^0
- x^0 is a non-degenerate point of maximum.

M
- $G(x) \in C([0,1]^d)$
- maximizing set for $G(x)$, $x \in [0,1]^d$, is a C^∞ manifold $M^{d-1} \subset [0,1]^d$
- $G(x) \in C^\infty$ in some neighbourhood of M^{d-1}
- $G''_{\nu\nu}(x) \neq 0$, $x \in M^{d-1}$, where $\partial/\partial\nu$ denote the derivatives in the direction of a normal vector to M^{d-1}.

For convenience of the reader we recall Theorem 2 from Konakov and Piterbarg (1994) which we formulate in a suitable form.

Theorem A (Konakov and Piterbarg (1994).)
Consider the model (2) with the sequence of trends $\{\mu_T(\cdot)\}$ satifying (3) and $\mathcal{M}$ and with the correlation function satisfying (4) and having a bounded support. Then there exist positive $q < 1$, $\gamma = \gamma(q) > 0$ and $C < \infty$ such that for any $a \in (0,1)$ and all

$$\mu \in \{\mu : ||\mu||_\circ \leq q u_T\}$$

we have

$$\begin{aligned} &\Big|P\,(\max_{[0,T]^d}|X(t)| \leq u_T) - \exp\big(-EM_{u_T}([0,T]^d)\big)\Big| \\ \leq\; & C\cdot T^d\cdot\exp\Big(-\frac{(u_T-m_T)^2}{2\sigma^2}\Big)\cdot\Big[u_T^{d-1}T^{-a} + u_T^{2d-2}\cdot T^{ad}\cdot\exp\Big(-\frac{(u_T-m_T)^2}{2\sigma^2}\Big) \\ &+ T^{ad}\exp\Big(-\frac{(\gamma u_T^2 - m_T^2)}{2\sigma^2}\Big) + u_T^{2d-1}T^{ad}\cdot\exp\Big(-\frac{(u_T-m_T)^2}{2\sigma^2}\cdot\frac{2-\rho}{\rho}\Big)\Big], \qquad (6) \end{aligned}$$

where

$$m_T = \max_{[0,T]^d}|\mu_T(t)|\,,\ \sigma^2 = \mathrm{var}X(t)\,,\ \rho = \max(1-\rho(t)) < 2\,.$$

Remark. If $(u_T - m_T) \sim \sqrt{2d\cdot\ln T}$ then the right–hand side of (6) has an order $O(T^{-\kappa})$, for some $\kappa > 0$, as $T \to \infty$.

We recall a definition from Konakov and Piterbarg (1994).

Definition 2. We say that a nonnegative function $\Psi(x)$, $x \in [0,1]^d$, is Laplace regular of indexes $R > 0$ and $r \leq 0$, if

$$\int_{[0,1]^d}\exp(-z\Psi(x))dx = Rz^r\,(1+o(1)) \qquad (7)$$

as $z \to +\infty$.

Now we formulate our first main result concerning the model (2), index T means that the trend $\mu(t)$ may depend on some parameter T. Let us introduce a notation

$$\Delta(z,a,b) \;=\; P(a(\max_{[0,T]^d} X_T(t) - \lambda_T G_0 - b) \;<\; z) - \exp(-e^{-z})\,,$$

where $G_0 = \max_{I^d} G(x)$, λ_T is the same as in the condition $\mathcal{M}$.
We need also the following definition.

Definition 3. We say that the function $R(x) \geq 0$ is the second order Laplace regular of indexes $R > 0$ and $r \leq 0$ if the condition **U** holds and

$$\begin{aligned} I_R \;&=\; \int_{I^d}\Big(1+\frac{R(x)}{v_T - G_0}\Big)^{d-1}\cdot\exp(-\lambda_T^2(v_T-G_0)R(x) - \frac{\lambda_T^2R^2(x)}{2})dx \\ &=\; R\cdot\lambda_T^{2r}\,(v_T-G_0)^r(1+O(\frac{1}{\lambda_T^2})) \end{aligned}$$

as $T \to \infty$.

Example: If $R(x) \equiv 0$ then $R(x)$ is the second order Laplace regular function of indexes $R = 1$ and $r = 0$. More interesting examples follow from Lemmas 3 and 4.

Theorem 1. Let $\mathcal{J}$ be any set containing three or more distinct points. Suppose that $G_0 - G(x)$ is Laplace regular function of some indexes $R > 0$ and $r \le 0$. Then

$$\liminf_{T\to\infty} (\log T) \inf_{-\infty<a,b<\infty} \sup_{z\in\mathcal{J}} |\Delta(z,a,b)| > 0. \tag{8}$$

Suppose that $G_0 - G(x)$ is the second order Laplace regular function and $\lambda_T \asymp \sqrt{\ln T}$. Then there exist a_T and b_T such that

$$\sup_{-\infty<z<\infty} |\Delta(z, a_T, b_T)| \le \frac{c}{\log T} \tag{9}$$

for some $c < \infty$. In particular, (9) holds if $G_0 - G(x)$ satisfies the conditions **P** or **M**. Analogous results hold for $\max_{[0,T]^d} |X(t)|$. The result formulated in terms of Laplace regularity of the function $\max_{I_d} |G(x)| - |G(x)|$.

Now we formulate corresponding result for density estimates. Recall that in this case our parameter T finally appears as a function of sample size n, namely, $T = h_n^{-1}$, $G(x)$ is defined in (1), and

$$\mu_n(x) = \left(\frac{nh_n^d}{\sigma^2 f(x)}\right)^{1/2} \cdot \int K(y)(f(x+h_n y) - f(x))dy.$$

Denote

$$\delta(z,a,b) = P(a(\max_{[0,1]^d} \xi_n(x) - \lambda_n \cdot G_0 - b) < z) - \exp(-e^{-z}),$$

where $G_0 = \max_{[0,1]^d} G(x)$, $\lambda_n = \left(\frac{nh_n^d}{\sigma^2}\right)^{1/2} \cdot h_n^p$.

To formulate our second theorem we recall the definition of a weak optimality.

Definition 4. (Konakov and Piterbarg (1994)). We say that a choice of $h = h_{n,w}$ is weakly optimal if

$$0 < \liminf_{n\to\infty} \frac{m_n}{\sqrt{2d \ln \frac{1}{h_n}}} \le \limsup_{n\to\infty} \frac{m_n}{\sqrt{2d\ln\frac{1}{h_n}}} \le \omega < \infty,$$

where $m_n = \max_{I^d} |\mu_n(x)|$.

Theorem 2 Let $\mathcal{J}$ be any set containing three or more distinct points. Suppose that $G_0 - G(x)$ is Laplace regular function of some indexes $R > 0$ and $r \le 0$. Then

$$\liminf_{n\to\infty} (\log n) \inf_{-\infty<a,b<\infty} \sup_{z\in\mathcal{J}} |\delta(z,a,b)| > 0. \tag{10}$$

Suppose that $G_0 - G(x)$ is second order Laplace regular function. Then there exists a weak optimal choice of $h_{h,\omega}$, a_n, b_n such that

$$\sup_{-\infty<z<\infty} |\delta(z, a_n, b_n)| \le \frac{c}{\log n} \tag{11}$$

for some $c < \infty$. In particular, (11) holds if $G_0 - G(x)$ satisfies the conditions **P** or **M**. Analogous results hold for $\max_{[0,1]^d} |\xi_n(x)|$.

2 Auxiliary lemmas

In this section we collect necessary preliminary results.

Lemma 1. Consider four sequences of real numbers $\{\alpha_T\},\{\beta_T\},\{\gamma_T\}$ and $\{\delta_T\}$ depending on some parameter T. Suppose that $\alpha_T > 0$ for all T and

$$\alpha_T z_i^2 + \beta_T z_i + \gamma_T - \ln(1+\delta_T z_i) \longrightarrow 0 \text{ as } T \to \infty \tag{12}$$

for three distinct points z_i, $i = 1,2,3$. Then $\alpha_T,\beta_T,\gamma_T$ and δ_T tend to zero as $T \to \infty$.

Proof. Obviously, that the equation

$$\alpha z^2 + \beta z + \gamma = \ln(1+\delta_T z),\ \alpha > 0$$

has at most two distinct roots. If $\delta_T \to 0$ then $\ln(1+\delta_T z_i) \to 0$, $i = 1,2,3$, and, hence, $\alpha_T,\beta_T,\gamma_T \to 0$ as $T \to \infty$. If $\delta_T \not\to 0$ then there exists a subsequence $\{T_k\}$ for which $|\delta_{T_k}| > \delta_0$ for some positive δ_0 and we obtain that the function $\ln(1+\delta_{T_k} z)$ has maximal oscillation over the set $\{z_1\}\cup\{z_2\}\cup\{z_3\}$ separated from zero for all k. By (12) the same assertion holds true for the corresponding sequence of parabolas $\{\alpha_{T_k} z^2 + \beta_{T_k} z + \gamma_{T_k}\}$. Suppose that $\delta_{T_k} \not\to \infty$. Then there exists a subsequence $\{T_k'\} \subset \{T_k\}$, such that $\delta_{T_k'} \to \delta_1 > 0$ as $k \to \infty$. It easily follows that $\alpha_{T_k'},\beta_{T_k'}$ and $\gamma_{T_k'}$ tend to some limits $\alpha_1,\beta_1,\gamma_1$ as $k \to \infty$ and

$$\alpha_1 z_i^2 + \beta_1 z_i + \gamma_1 = \ln(1+\delta_1 z_i),\ i = 1,2,3.$$

This contradicts our previous conclusion about the maximal number of different roots of this equation. The case $\delta_{T_k} \to \infty$ may be handled similary. This completes the proof.

Denote

$$\mathcal{I}_i(T) = \lambda_2^{d/2} \cdot \int_{[0,T]^d} \varphi_d(\lambda_2^{-1/2} \mathrm{grad}\mu_T(t))(u_T - \mu_T(t))^i \varphi_1(u_T - \mu_T(t))dt,$$

where φ_d and φ_1 are d-dimensional and 1-dimensional standard Gaussian densities, respectively, and λ_2 is the coefficient in expansion (4).

Lemma 2. Let the condition $\mathcal{M}$ be fulfilled. Then there exist positive γ,κ and $\mathcal{C}$ such that

$$EM_{u_T}^+([0,T]^d) = \mathcal{I}_{d-1}(T)\cdot[1 + O(T^{-\kappa}) + O(e^{-\gamma u_T^2})] + \mathcal{I}_{d-3}(T)\cdot[\mathcal{C} + o(1)],$$

as $T \to \infty$

Proof. The complete proof is rather tedious and we give here only a sketch of the proof.

Step one. Assume that $\lambda_2 = 1$. It can be shown analogously to Konakov (1992) (see also Hasofer (1976) for the case $\mu_T(\cdot) \equiv 0$) that

$$EM_{u_T}^+([0,T]^d) = (-1)^d \cdot \int_{[0,T]^d} dt\, \varphi_d(\mathrm{grad}\,\mu_T(t)) \cdot \int_{u_T}^{\infty} \varphi_1(x - \mu_T(t))dx \cdot \int_{U(x,t)} \det|W - (x-\mu_T(t))\cdot I| \cdot \varphi(\mathbf{w})d\mathbf{w},$$

where $W = ||w_{ij}||$ and I are symmetric and identity matrices respectively, $\mathbf{w} = (w_{11}, ..., w_{1d}, w_{22}, ..., w_{2d}, ..., w_{dd})$, $\varphi(\mathbf{w})$ is a Gaussian density with the covariance matrix $\sum_t$ not depending on x, and with the mean vector

$$\left(\frac{\partial^2 \mu_T(t)}{\partial t_1^2}, ..., \frac{\partial^2 \mu_T(t)}{\partial t_1 \partial t_d}, \frac{\partial^2 \mu_T(t)}{\partial t_2^2}, ..., \frac{\partial^2 \mu_T(t)}{\partial t_2 \partial t_d}, ..., \frac{\partial^2 \mu_T(t)}{\partial t_d^2}\right),$$

$U(x,t) = \{\mathbf{w} \in R^{d(d+1)/2} : W - (x - \mu_T(t))I$ is negatively definite $\}$.

Step two. As in Konakov (1992) (and Hasofer (1976) for the case $\mu_T(\cdot) \equiv 0$) we prove that there exists $\varepsilon > 0$ such that the set $U(x,t)$ contains the ball of a radius $\varepsilon \cdot u_T$. Hence, integration over $U(x,t)$ may be replaced by the integration over $R^{d(d+1)/2}$ with exponentially small error.

Step three. Using the condition $\mathcal{M}$ and the step two we get

$$\int_{u_T}^{\infty} \varphi_1(x - \mu_T(t))dx \int_{U(x,t)} \det|W - (x - \mu_T(t)) \cdot I| \cdot \varphi(\mathbf{w})d\mathbf{w}$$
$$= \sum_{i=1}^{d} a_i \cdot \int_{u_T - \mu_T(t)}^{\infty} y^i \cdot \varphi_1(y)dy + R, \tag{13}$$

where

$$R = (u_T - \mu_T(t))^{d-1} \cdot \varphi_1(u_T - \mu_T(t)) \cdot O(T^{-\kappa} + e^{-\gamma u_T^2}), \text{ as } T \to \infty$$

for some $\gamma > 0$, $\kappa > 0$, $a_d = (-1)^d$, $a_{d-1} = 0$, $a_{d-2} = (-1)^{d-2} \times$ (sum of all main minors of the second order of $\sum_t$).

Note, that SR–condition imply that $(-1)^{d-2} \cdot a_{n-2}$ is strictly positive for $t \in [0,T]^d$.

Step four. Using Watson's lemma (see e. g. Copson (1965)) for the analysis of integrals in (13) we obtain

$$(-1)^d \cdot \int_{u_T}^{\infty} \varphi_1(x - \mu_T(t))dx \int_{U(x,t)} \det|W - (x - \mu_T(t)) \cdot I| \cdot \varphi(\mathbf{w})d\mathbf{w}$$
$$= (u_T - \mu_T(t))^{d-1} \cdot \varphi_1(u_T - \mu_T(t)) \cdot \Bigg[1 + \frac{(d-1) + (-1)^{d-2} \cdot a_{d-2} + o(1)}{(u_T - \mu_T(t))^2}$$
$$+ O(T^{-\kappa} + e^{-\gamma u_T^2}) \Bigg], \text{ as } T \to \infty \tag{14}$$

Clearly (14) imply the assertion of the lemma for the case $\lambda_2 = 1$. To prove the general case note, that the field $\widetilde{X}(t) = X(\lambda_2^{-1/2} t)$ has covariance matrix of grad $\widetilde{X}(t)$ equal to I. This completes the proof.

Repeating arguments from Konakov and Piterbarg (1994), pp. 19 - 21, we obtain that

$$\begin{aligned} \mathcal{I}_{d-i}(T) &= \lambda_2^{d/2}(2\pi)^{-\frac{d+1}{2}} T^d \cdot \lambda_T^{d-i} \cdot (v_T - G_0)^{d-i} \cdot \exp\left(-\frac{\lambda_T^2}{2}(v_T - G_0)^2\right) \\ &\quad \times \mathcal{I}_R \cdot (1 + O(T^{-\kappa})), \text{ as } T \to \infty, i = 1, 3, \end{aligned} \tag{15}$$

where

$$\mathcal{I}_R = \int_{I^d} \left(1 + \frac{R(x)}{v_T - G_0}\right)^{d-1} \cdot \exp\left(-\lambda_T^2 (v_T - G_0) \cdot R(x) - \frac{\lambda_T^2}{2} R^2(x)\right) dx,$$

$R(x) = G_0 - G(x)$.

Note that the multiplier $\mathcal{I}_R$ in (15) reflects the effect of non–stationarity (in stationary case $R(x) \equiv 0$ and $\mathcal{I}_R = 1$). Note also that if the condition **U** is fulfilled then we have the following estimates

$$\int_{I^d} \left(1 + \frac{R(x)}{v_T - G_0}\right)^{d-1} \cdot e^{\lambda_T^2 (v_T - G_0)\cdot S_1(x)}\, dx \le \mathcal{I}_R$$
$$\le \int_{I^d} \left(1 + \frac{R(x)}{v_T - G_0}\right)^{d-1} \cdot e^{\lambda_T^2 (v_T - G_0)\cdot S_2(x)}\, dx \tag{16}$$

where

$$S_1(x) = -R(x),\ S_2(x) = -R(x) - \frac{R^2(x)}{2}.$$

The integrals in (16) are estimated in the following lemma.

Lemma 3. Let the condition **P** be fulfilled. Then

$$\int_{I^d} \left(1 + \frac{R(x)}{v_T - G_0}\right)^{d-1} \cdot e^{\lambda_T^2 (v_T - G_0)\cdot S_i(x)}\, dx$$
$$= \lambda_T^{-d} \cdot (v_T - G_0)^{-d/2} \cdot \left(b_0 + \frac{b_i + o(1)}{\lambda_T^2 (v_T - G_0)}\right),\ i = 1, 2,$$

as $\lambda_T \to \infty$. The coefficients b_i, $i = 1, 2$, depend only on the local structure of $G(x)$ in the neighbourhood of x^0.

Proof. It is sufficiently to consider our integrals over a small neighbourhood of x^0. We choose this neighbourhood in the form $\{x : -\delta_T < G(x) - G_0 \le 0\}$, $\delta_T = (\lambda_T^2(v_T - G_0))^{\gamma - 1}$, $\gamma > 0$. Using integration over the level sets we obtain

$$\int_{I^d} e^{-\lambda_T^2 (v_T - G_0) R(x)}\, dx = \int_0^{\delta_T} e^{-\lambda_T^2 (v_T - G_0)\cdot c} \cdot \Phi_\omega(c) dc$$
$$+ O\Big(\exp[-(\lambda_T^2 (v_T - G_0))^{\gamma'}]\Big),\quad 0 < \gamma' < \gamma, \tag{17}$$

where

$$\Phi_\omega(c) = \int_{G(x) - G_0 = -c} \omega_G(x),$$

$\omega_G(x)$ is the Leray–Gel'fand differential form

$$\omega_G(x) = \sum_{j=1}^{d} (-1)^{j-1} \cdot \frac{\partial G(x)}{\partial x_j} \cdot \left|\nabla G(x)\right|^{-2} dx_1 \wedge \ldots dx_{j-1} \wedge dx_{j+1} \wedge \ldots \wedge dx_d .$$

We suppose that $G(x) \in C^5$ in some neighbourhood of x^0, hence $D(y) = \det \varphi'(y) \in C^2$ in some neighbourhood of the point $y = 0$. Here $\varphi(y)$ is a diffeomorphism in Morse lemma (see, e.g., Fedorjuk (1977), p.68) and $\varphi'(y)$ is a corresponding Jacobian. Arguing as in Fedorjuk (1977), Proposition 3.3, p. 70, we obtain

$$\Phi_\omega(c) = c^{d/2 - 1} \cdot (a_0 + a_1 c + o(c)),\ \text{as } c \to +0 \tag{18}$$

where

$$a_0 = \frac{1}{2}\int_{|y|=1} \omega_G(y), \quad a_1 = \sum_{i,j=1}^{d} \frac{\partial^2 D(\mathrm{o})}{\partial y_i \partial y_j} \cdot \int_{|y|=1} y_i y_j \omega_G(y).$$

Substituting (18) in (17) and using Watson's lemma (Fedorjuk (1977), p. 31) we get

$$\int_{I^d} e^{-\lambda_T^2(v_T-G_0)R(x)}dx = \lambda_T^{-d}\cdot(v_T-G_0)^{-d/2}\cdot\left(a_0\Gamma(\frac{d}{2}) + \frac{a_1\cdot\Gamma(\frac{d}{2}+1)+\mathrm{o}(1)}{\lambda_T^2(v_T-G_0)}\right),$$

as $\lambda_T \to +\infty$.

The integral

$$\int_{I^d} e^{-\lambda_T^2(v_T-G_0)\left(R^2(x)+\frac{R^2(x)}{2}\right)}dx$$

$$= \int_0^{\delta_T} e^{-\lambda_T^2(v_T-G_0)(c+\frac{c^2}{2})}\cdot\Phi_\omega(c)dc + O(\exp[-(\lambda_T^2(v_T-G_0))^{\gamma'}]),\ 0<\gamma'<\gamma,$$

may be estimated quite analogously if we notice that for $c\in[0,\delta_T]$

$$\lambda_T^2(v_T-G_0)\cdot\frac{c^2}{2} \longrightarrow 0, \quad T\to\infty$$

and use the decomposition

$$e^{-x} = 1 - x + O(x^2), \quad x\to 0.$$

Using binomial decomposition for

$$(1+\frac{R(x)}{v_T-G_0})^{d-1}$$

and estimating the integrals

$$\int_{I^d} R^k(x)\, e^{\lambda_T^2(v_T-G_0)S_i(x)}\, dx$$

exactly by the same way as it was discribed for the case $k=0$ we easily obtain the conclusion of the lemma.

Lemma 4. Let the conditions **M** be fulfilled. Then

$$\int_{I^d}\left(1+\frac{R(x)}{v_T-G_0}\right)^{d-1}\cdot e^{\lambda_T^2(v_T-G_0)S_i(x)}\,dx$$

$$= (\lambda_T^2(v_T-G_0))^{-1/2}\cdot\left(c_0+\frac{c_i+\mathrm{o}(1)}{\lambda_T^2(v_T-G_0)}\right),\ i=1,2,\ \text{as}\ \lambda_T\to\infty. \tag{19}$$

Proof. As usual it is sufficiently to consider integrals in (19) over a small neighbourhood of M^{d-1}. Let M_0^{d-1} be a connected component of M^{d-1} and $x^0\in M_0^{d-1}$. If $\delta>0$ is sufficiently small then a set M_δ^{d-1} : $-\delta<\sqrt{-S_1(x)}<\delta$ contains M_0^{d-1} and doesn't intersects with the other components of the set M^{d-1}. Arguing as in Fedorjuk (1977), Theorem 4.8, p. 89, we reduce our integrals to the form

$$\int_{M_\delta^{d-1}} \left(1+\frac{R(x)}{v_T-G_0}\right)^{d-1} \cdot e^{\lambda_T^2(v_T-G_0)S_i(x)}\,dx$$
$$= \int_{-\delta}^{\delta} e^{-\lambda_T^2(v_T-G_0)\cdot g_i(t)} \cdot \left(1+\frac{t^2}{v_T-G_0}\right)^{d-1} \cdot \Psi_R(t)dt\,,\ i=1,2,$$

where
$\Psi_R(t) = \int_{\sqrt{R(x)}=t} \omega_R$, ω_R is a Leray–Gel'fand differential form, $g_1(t) = t^2$, $g_2(t) = t^2+\frac{t^4}{2}$. The function $\Psi_R(t) \in C^\infty([-\delta,\delta])$ and the conclusion of the lemma follows from Theorem 1.3 (Fedorjuk (1977), p. 39).

We conclude this section by the following simple lemma. Consider a parabola $az^2 + bz + c$, $a > 0$, and three distinct points z_1, z_2 and z_3.

Lemma 5. There exists $\varepsilon > 0$ depending only on z_1, z_2, z_3 such that

$$|az_i^2 + bz_i + c| > \varepsilon \cdot a$$

for at least one of z_i, $i = 1, 2, 3$.

Proof. Obviously,

$$az^2+bz+c = a(z+\frac{b}{2a})^2 + c - \frac{b^2}{4a}$$

and, hence, it is sufficiently to prove that $\mathcal{R} > \varepsilon$, where $\mathcal{R}$ is a maximal oscillation of the function $(z+\frac{b}{2a})^2$ over the set $\{z_1\} \cup \{z_2\} \cup \{z_3\}$. We have

$$\mathcal{R} = |z_1-z_2|\cdot|z_1+z_2+\frac{b}{a}| \vee |z_1-z_3|\cdot|z_1+z_3+\frac{b}{a}| \vee |z_2-z_3|\cdot|z_2+z_3+\frac{b}{a}|$$
$$> (|z_1-z_2| \wedge |z_1-z_3| \wedge |z_2-z_3|)^2 = \varepsilon(z_1,z_2,z_3)\,.$$

This completes the proof.

3 Proofs of Main Results

Proof of the Theorem 1.

Take $u_T = \lambda_T G_0 + b_T + a_T^{-1} z$, where a_T, b_T are choosen so that

$$\exp(-M_{u_T}^+([0,T]^d)) \longrightarrow \exp(-e^{-z})$$

uniformly over $z \in \mathcal{J}$, as $T \to \infty$. Using Lemma 2 and (15) we obtain after simple calculations

$$\alpha_T z^2 + \beta_T z + \gamma_T - (d-1+r)\cdot\ln(1+\frac{z}{a_T b_T}) - \ln(1+\rho_T) \to 0$$

uniformly over $z \in \mathcal{J}$ as $T \to \infty$, where

$$\alpha_T = \frac{1}{2a_T^2}\,,\ \beta_T = a_T^{-1}\cdot b_T - 1\,,\ \gamma_T = \frac{b_T^2}{2} - \ln(cT^d\lambda_T^r) - (d-1+r)\ln b_T\,,\ c = (2\pi)^{-\frac{d+1}{2}}\cdot R\cdot\lambda_2^{d/2}$$

and ρ_T is a remainder term in the relation

$$EM_{u_T}^+([0,T]^d) = c\cdot T^d\cdot\lambda_T^r\cdot(u_T-\lambda_T G_0)^{d-1+r}\cdot\exp\Big(-\frac{(u_T-\lambda_T G_0)^2}{2}\Big)(1+\rho_T) \quad (20)$$

as $T \to \infty$.
Note that we use only that $\rho_T \to 0$ as $T \to \infty$. More precise estimates for ρ_T follow from Lemmas 2, 3, 4 and (15). By Lemma 1 $\alpha_T, \beta_T, \gamma_T \to 0$ as $T \to \infty$. Using the expansion

$$\log(1+x) = x + O(x^2) \quad \text{as } x \to 0$$

for $\ln(1+\rho_T)$ and $\ln(1+\frac{z}{a_T b_T}) = \ln(1+\frac{2\alpha_T}{1+\beta_T}\cdot z)$, and the expansion $e^x = 1 + x + O(x^2)$ as $x \to 0$ we easily obtain

$$\begin{aligned}\exp\Big(-EM_{u_T}^+([0,T]^d)\Big) &= \exp(-e^{-z})\cdot\Big[1 + e^{-z}\cdot(\alpha_T z^2 + \tilde{\beta}_T z + \tilde{\gamma}_T + O(\alpha_T^2+\rho_T^2)\Big) \\ &\quad + O\Big((\alpha_T z^2 + \tilde{\beta}_T z + \tilde{\gamma}_T + O(\alpha_T^2+\rho_T^2))^2\Big) \qquad (21)\end{aligned}$$

where

$$\tilde{\beta}_T = \beta_T - \frac{(d-1+r)\cdot 2\alpha_T}{1+\beta_T}\,,\ \tilde{\gamma}_T = \gamma_T - \rho_T.$$

It follows from (21) and Lemma 5 that

$$\sup_{z\in\mathcal{J}}\Big|\exp\Big(-EM_{u_T}^+([0,T]^d)\Big) - \exp(-e^{-z})\Big| \geq \varepsilon\cdot\alpha_T \quad (22)$$

for some positive ε depending only on $\mathcal{J}$. Taking into account that

$$b_T = \frac{1+\beta_T}{\sqrt{2\alpha_T}} \sim \frac{1}{\sqrt{2\alpha_T}}\,,\ b_T^2 \sim (u_T-\lambda_T G_0)^2 = 2z + 2\ln(CT^d\lambda_T^r) + 2(d-1+r)\ln b_T + o(1)\,,$$

we have $2\ln(CT^d\lambda_T^r) \sim b_T^2 \sim \frac{1}{4\alpha_T^2}$ and hence

$$\alpha_T \sim \frac{1}{4\ln(CT^d\lambda_T^r)}, \quad u_T - \lambda_T G_0 \sim \sqrt{2\ln(CT^d\lambda_T^r)} \quad (23)$$

The first assertion of the Theorem 1 follows from (22), (23) and Theorem A (see also Remark after Theorem A). To prove the second assertion we take $u_T = \lambda_T G_0 + l_T + \frac{Z}{l_T}$, where l_T is the maximal root of the equation

$$CT^d\lambda_T^r\cdot l^{d-1+r}e^{-l^2/2} = 1$$

It is easy to see that $l_T \sim \sqrt{2d\ln T}$ as $T \to \infty$. The second assertion of Theorem 1 follows from Lemma 2, (15) and Definition 3. In particular, if the conditions **P** and **M** are fulfilled then by Lemmas 3, 4 and (16) we obtain the second order Laplace regularity for $R(x)$ and, hence, (9). This completes the proof.

Proof of the Theorem 2.

It follows from the integral representation for $EM_u^+([0,T]^d)$ and (13) that

$$\left|EM_{u'_n}^+([0,T]^d) - EM_{u''_n}^+([0,T]^d)\right| = O(n^{-\kappa})$$

if $|u'_n - u''_n| = O(n^{-\kappa})$ as $n \to \infty$. The proof of Theorem 2 follows now from Theorem 1 and the Rio's estimate (see Konakov and Piterbarg (1994))

$$P\Big(\sup_{I^d} |\xi_n(x) - \mu_n(x) - \sigma^{-1} h_n^{-d/2} W_n(K_{h_n}(\cdot - x))| > Cn^{-\gamma}\Big) \leq Cn^{-\gamma}$$

for some $C < \infty$ and $\gamma > 0$.

References

[1] Copson, E. (1965). *Asymptotic expansions.* Cambridge Univ. Press.

[2] Fedorjuk, M. (1977). *The saddle–point method.* Nauka, Moscow (in Russian).

[3] Hall, P. (1991). *On convergence rates of suprema.* Probab. Th. Rel. Fields. **89**, 447 - 455.

[4] Hasofer, A. (1976). *The mean number of maxima above high levels in Gaussian random fields.* J. Appl. Prob. **13**, 377 - 379.

[5] Konakov, V. (1992). *Asymptotic Methods of Theory of Random Processes and Fields in Problems of Nonparametric Statistics.* Thesis. Doct. of Sci. Moscow State University (in Russian).

[6] Konakov, V. and Piterbarg, V. (1982). *Convergence rates of maximal deviations distributions of Gaussian processes and empirical densities.* I. Probab. Th. and its Appl. **27**, 759 - 779.

[7] Konakov, V. and Piterbarg, V. (1983). *Convergence rates of maximal deviations distributions of Gaussian processes and empirical densities.* II. Probab. Th. and Appl. **28**, 172 - 178.

[8] Konakov, V. and Piterbarg, V. (1984). *On the convergence rate of maximal deviations distribution for kernel regression estimates.* J. Multivar. Anal., **15**, 279 - 294.

[9] Konakov, V. and Piterbarg, V. (1994). *High level excursions of Gaussian fields and the optimal choice of the smoothing parameter.* Preprint ISRN: LUNFD6/NFMS - 3148 - SE, Univ. of Lund.

[10] Rio, E. (1994). *Local invariance principles and their application to density estimation.* Prob. Th. and Rel. Fields. **98**, 21 - 46.

Recent publications of the Institut für Angewandte Analysis und Stochastik

Preprints 1993

74. Roger Tribe: The long term behavior of a Stochastic PDE.

75. Annegret Glitzky, Konrad Gröger, Rolf Hünlich: Rothe's method for equations modelling transport of dopants in semiconductors.

76. Wolfgang Dahmen, Bernd Kleemann, Siegfried Prößdorf, Reinhold Schneider: A multiscale method for the double layer potential equation on a polyhedron.

77. Hans-Günter Bothe: Attractors of non invertible maps.

78. Gregori Milstein, Michael Nussbaum: Autoregression approximation of a nonparametric diffusion model.

Preprints 1994

79. Anton Bovier, Véronique Gayrard, Pierre Picco: Gibbs states of the Hopfield model in the regime of perfect memory.

80. Roland Duduchava, Siegfried Prößdorf: On the approximation of singular integral equations by equations with smooth kernels.

81. Klaus Fleischmann, Jean-François Le Gall: A new approach to the single point catalytic super-Brownian motion.

82. Anton Bovier, Jean-Michel Ghez: Remarks on the spectral properties of tight binding and Kronig-Penney models with substitution sequences.

83. Klaus Matthes, Rainer Siegmund-Schultze, Anton Wakolbinger: Recurrence of ancestral lines and offspring trees in time stationary branching populations.

84. Karmeshu, Henri Schurz: Moment evolution of the outflow-rate from nonlinear conceptual reservoirs.

85. Wolfdietrich Müller, Klaus R. Schneider: Feedback stabilization of nonlinear discrete-time systems.

86. Gennadii A. Leonov: A method of constructing of dynamical systems with bounded nonperiodic trajectories.

87. Gennadii A. Leonov: Pendulum with positive and negative dry friction. Continuum of homoclinic orbits.

88. Reiner Lauterbach, Jan A. Sanders: Bifurcation analysis for spherically symmetric systems using invariant theory.

89. Milan Kučera: Stability of bifurcating periodic solutions of differential inequalities in $\mathbb{R}^3$.

90. Peter Knabner, Cornelius J. van Duijn, Sabine Hengst: An analysis of crystal dissolution fronts in flows through porous media Part I: Homogeneous charge distribution.

91. Werner Horn, Philippe Laurençot, Jürgen Sprekels: Global solutions to a Penrose–Fife phase–field model under flux boundary conditions for the inverse temperature.

92. Oleg V. Lepskii, Vladimir G. Spokoiny: Local adaptivity to inhomogeneous smoothness. 1. Resolution level.

93. Wolfgang Wagner: A functional law of large numbers for Boltzmann type stochastic particle systems.

94. Hermann Haaf: Existence of periodic travelling waves to reaction–diffusion equations with excitable–oscillatory kinetics.

95. Anton Bovier, Véronique Gayrard, Pierre Picco: Large deviation principles for the Hopfield model and the Kac–Hopfield model.

96. Wolfgang Wagner: Approximation of the Boltzmann equation by discrete velocity models.

97. Anton Bovier, Véronique Gayrard, Pierre Picco: Gibbs states of the Hopfield model with extensively many patterns.

98. Lev D. Pustyl'nikov, Jörg Schmeling: On some estimations of Weyl sums.

99. Michael H. Neumann: Spectral density estimation via nonlinear wavelet methods for stationary non–Gaussian time series.

100. Karmeshu, Henri Schurz: Effects of distributed delays on the stability of structures under seismic excitation and multiplicative noise.

101. Jörg Schmeling: Estimates of Weyl sums over subsequences of natural numbers.

102. Grigori N. Milstein, Michael V. Tret'yakov: Mean–square approximation for stochastic differential equations with small noises.