The Annals of Statistics
1991, Vol. 19, No. 2, 557–578

ASYMPTOTICS OF MAXIMUM LIKELIHOOD ESTIMATORS FOR THE CURIE–WEISS MODEL

BY FRANCIS COMETS[1] AND BASILIS GIDAS[2]

Brown University

We study the asymptotics of the ML estimators for the Curie–Weiss model parametrized by the inverse temperature β and the external field h. We show that if both β and h are unknown, the ML estimator of (β, h) does not exist. For β known, the ML estimator $\hat{h}_n$ of h exhibits, at a first order phase transition point, superefficiency in the sense that its asymptotic variance is half of that of nearby points. At the critical point ($\beta = 1$), if the true value is $h = 0$, then $n^{3/4}\hat{h}_n$ has a non-Gaussian limiting law. Away from phase transition points, $\hat{h}_n$ is asymptotically normal and efficient. We also study the asymptotics of the ML estimator of β for known h.

1. Introduction and main results. During the last few years a great deal of attention has been given [2, 15, 16, 14, 17, 19, 8, 10, 11, 20] to the estimation of parameters for Gibbs distributions–equivalently, Markov random fields (MRF). This statistical inference problem was primarily motivated by applications [6, 7, 9, 12, 13, 21] of the framework for image processing tasks formulated in [6].

In this paper we study the asymptotics of the maximum likelihood (ML) estimators for the Curie–Weiss model [4] which shares many qualitative properties with MRF. Our main results are stated in Theorems 1.1–1.4 below. Among other things, we prove (Theorem 1.1) that at a first order phase transition point, the ML estimator of the external magnetic field is asymptotically superefficient, in the sense that its asymptotic variance is half of the asymptotic variance of the estimators for nearby points. The ML estimates of these nearby points are asymptotically normal and their asymptotic variance saturates the Cramér–Rao lower bound. Preliminary work indicates that a similar superefficiency phenomenon occurs for the Ising model. The superefficiency property was conjectured in [11].

The Gibbs measure for the Curie–Weiss model is a probability distribution on $\{-1, 1\}^n$ whose density with respect to the counting measure is given by

$$P_{n,\beta,h}(x_1, \ldots, x_n) = \frac{1}{Z_n(\beta, h)} \exp\left\{n\left(\frac{\beta}{2}\overline{X}_n^2 + h\overline{X}_n\right)\right\}, \tag{1.1}$$

where $x_i \in \{-1, 1\}$, $i = 1, \ldots, n$, $\overline{X}_n = (x_1 + \cdots + x_n)/n$ and $Z_n(\beta, h)$ is a normalizing constant called the partition function. This distribution is

Received October 1988; revised March 1990.
[1]Partially supported by AFOSR Contract F-49620-86-C-0111.
[2]Partially supported by ARO DAAL03-86-K-0110 and ONR Contract N00014-88-K-0289.
AMS 1980 *subject classifications*. Primary 62F10, 62F12; secondary 62F99, 62P99.
Key words and phrases. Maximum likelihood estimators, phase transitions, superefficiency, consistency, asymptotic normality.

parametrized by two parameters: The inverse temperature $\beta > 0$, and the external magnetic field $h \in \mathscr{R}$. However, the distribution contains only one sufficient statistic, i.e., $\overline{X}_n$. This indicates that the pair(β, h) cannot be estimated simultaneously. In Section 2, we show that with probability 1, the ML estimator of the pair does not exist. In Theorems 1.1–1.3 below, β is assumed to be known and the ML estimator of h is studied. In Theorem 1.4, h is assumed known and the ML estimator of β is studied.

From (1.1), we obtain that the conditional distribution at site i is given by

$$P_{n,\beta,h}\left(x_i \middle| \{x_j\}_{j\neq i}\right) = \frac{\exp\{x_i(\beta(1/n)t_i + h)\}}{\exp\{\beta(1/n)t_i + h\} + \exp\{-\beta(1/n)t_i - h\}},$$

where $t_i = \sum_{j\neq i} x_j$. This shows the relation between the Curie–Weiss and Ising models, as well as the weak long-range interactions in the Curie–Weiss model.

In order to state our main results we need some elementary properties [4] of the Curie–Weiss model: For $h \neq 0$, let $m(\beta, h)$ be the unique solution of

$$\tanh(\beta m + h) = m \tag{1.2}$$

with $\operatorname{sign}(m) = \operatorname{sign}(h)$. For $\beta > 1$, let $m_+ = m_+(\beta)$ be the unique positive solution of

$$\tanh \beta m = m. \tag{1.3}$$

For fixed β, $m(\beta, h)$ is an odd increasing function of h, which is continuous for $h \neq 0$. For $\beta \leq 1$, $m(\beta, h)$ is continuous for all $h \in \Re$ and $m(\beta, 0) = 0$. For $\beta > 1$, $m(\beta, h) \to m_+(\beta)$ as $h \downarrow 0$, as $m(\beta, h) \to m_-(\beta) = -m_+(\beta)$ as $h \uparrow 0$ (see Figure 1). The Curie–Weiss model exhibits [4] a phase transition at $\beta = 1$, in the sense that for $h \neq 0$ or $\beta < 1$,

$$\overline{X}_n \to m(\beta, h) \tag{1.4}$$

in probability under $P_{n,\beta,h}$ and for $\beta > 1$, $h = 0$,

$$\overline{X}_n \to \tfrac{1}{2}\delta_{m_+(\beta)} + \tfrac{1}{2}\delta_{m_-(\beta)}$$

in distribution under $P_{n,\beta,0}$. These limits may be obtained as by-products of Proposition 2.1 in Section 2. The value $\beta = 1$ corresponds to the (inverse) critical temperature and the half line $\beta > 1$, $h = 0$, corresponds to the first order phase transition points. The term first order phase transition refers to the discontinuity of $m(\beta, h)$ across this half line. The point $\beta = 1$, $h = 0$ is referred to as the second order phase transition point. The following variances will enter our theorems: For $h \neq 0$, let

$$\sigma^2(\beta, h) = \frac{1 - m^2(\beta, h)}{1 - \beta[1 - m^2(\beta, h)]} \tag{1.5}$$

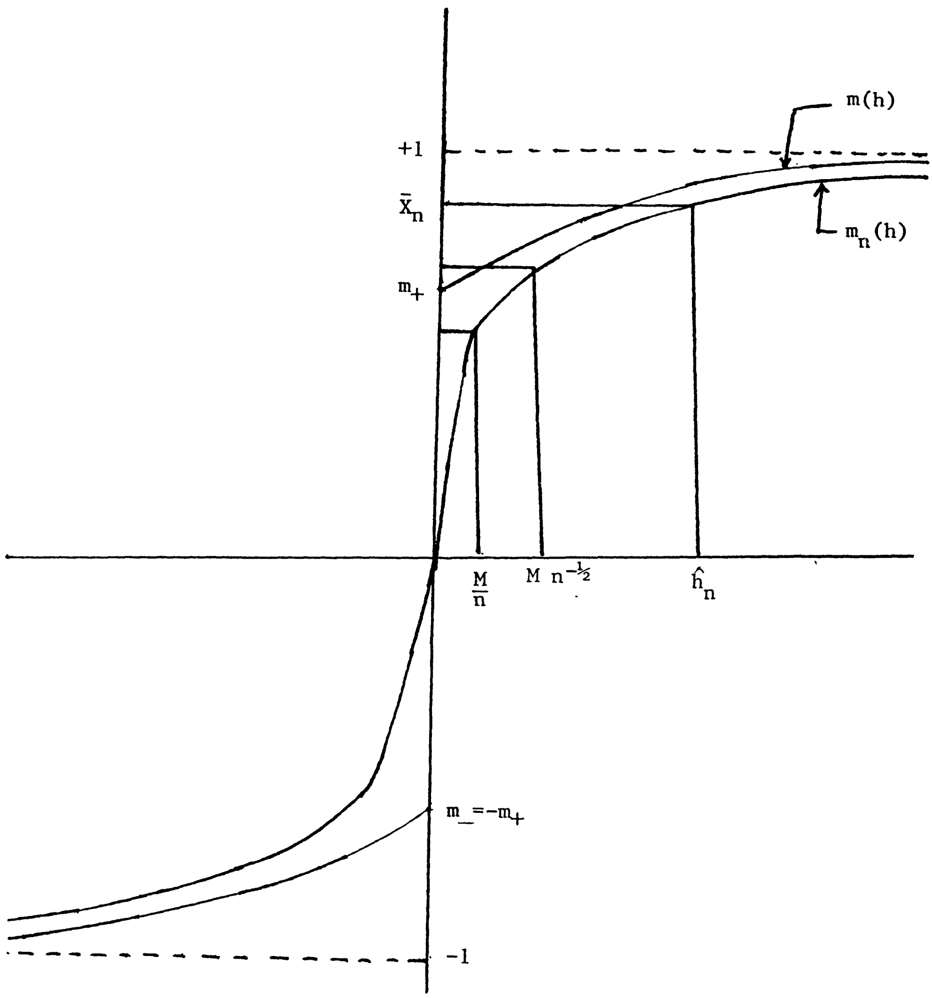

FIG. 1. *Graphs of* $m(h)$ *and* $m_n(h)$ *for* $\beta > 1$ *and large* n. *The behavior of* $m_n(h)$ *is deduced from* (2.36) *and* (2.33a).

and

$$\sigma^2(\beta, 0) = \lim_{h \to 0} \sigma^2(\beta, h)$$

(1.6a)
$$= \frac{1}{1-\beta}, \quad \text{if } \beta < 1,$$

(1.6b)
$$= \frac{1 - m_+^2}{1 - \beta(1 - m_+^2)}, \quad \text{if } \beta > 1.$$

Our main results are stated in Theorems 1.1–1.4, and are proven in the remaining sections. Throughout the paper, convergence in distribution will be indicated by $\to_D$. In Theorems 1.1–1.3, β is assumed to be known, and the β dependence of the various objects such as $P_{n,\beta,h}$, $m(\beta,h)$, $\sigma^2(\beta,h)$, will be suppressed.

THEOREM 1.1. *Assume that $\beta > 0$ is known and let $\hat{h}_n$ be the ML estimator of h. Then*

(a) *If $\beta > 0$, $h \neq 0$ or $\beta < 1$, $h = 0$, then under $P_{n,h}$,*

$$\sqrt{n}\left(\hat{h}_n - h\right) \to_D N\left(0, \frac{1}{\sigma^2(h)}\right). \tag{1.7}$$

(b) *If $\beta > 1$, $h = 0$, then under $P_{n,0}$,*

$$\sqrt{n}\,\hat{h}_n \to_D \frac{1}{2}\delta_0 + \frac{1}{2}N\left(0, \frac{1}{\sigma^2(0)}\right). \tag{1.8}$$

In Section 2, we will show that for $h \neq 0$ or $\beta < 1$, the Fisher information converges to $\sigma^2(h)$. Thus point (a) of Theorem 1.1 says that for $h \neq 0$ or $\beta < 1$, the ML estimator of h is consistent and asymptotically normal and efficient. While by part (b), the variance of the limiting distribution (i.e., the asymptotic variance) is $\frac{1}{2}(\sigma(0))^{-2}$. This means that at first order phase transition points, the ML estimator of $h = 0$ is asymptotically twice as efficient as the neighboring points (h close to zero, $\beta > 1$). This is the superefficiency phenomenon mentioned before. Part (b) of the theorem also shows that half of the time, the estimator converges to the true value $h = 0$ faster than $(1/\sqrt{n})$ [in fact at the speed $(\log n)/n$]. The intuitive origin of this lies in the fact (to be made precise in Section 2) that $\overline{X}_n$ visits, with probability one half for large n, each of the symmetric neighborhoods of m_+ and m_-. But for h slightly different from zero, $\overline{X}_n$, with probability tending to 1 as $n \to +\infty$, remains outside to the interval $[m_-, m_+]$.

The gap between $\frac{1}{2}(1/\sigma^2(0))$ and $1/\sigma^2(0)$ indicates a nonuniform convergence at the neighborhood of $h = 0 (\beta > 1)$. The next theorem fills the gap between these two values and provides an explanation of the underlying nonuniform convergence.

THEOREM 1.2. *Assume $\beta > 1$. For $\overline{h} \in \Re$, define*

$$\alpha(\overline{h}) = \frac{\exp\{\overline{h}m_+\}}{\exp\{\overline{h}m'_+\} + \exp\{-\overline{h}m_+\}}. \tag{1.9}$$

Let Y be a normal random variable with mean zero and variance $(\sigma(0))^{-2}$ and denote by $\mu_{\overline{h}}^+$ (resp., $\mu_{\overline{h}}^-$) the probability distribution of the truncated variable $Y1_{Y \geq \overline{h}}$ (resp., $Y1_{Y \leq \overline{h}}$). Then we have

(a) *Under $P_{n,\overline{h}/n}$ with $\overline{h} \in \Re$,*

$$\sqrt{n}\left(\hat{h}_n - \frac{\overline{h}}{n}\right) \to_D Z \sim \alpha(\overline{h})\mu_0^+ + \left[1 - \alpha(\overline{h})\right]\mu_0^-. \tag{1.10}$$

(b) *Under* $P_{n,\bar{h}n^{-1/2}}$ *with* $\bar{h} > 0$,

$$\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) \to_D Z \sim \mu^{+}_{-\bar{h}} \tag{1.11a}$$

and with $\bar{h} < 0$,

$$\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) \to_D Z \sim \mu^{-}_{\bar{h}}. \tag{1.11b}$$

Theorem 1.2 together with the (stochastic) monotonicity of $\hat{h}_n$ in h (see Section 2), yield all possible limits of $\sqrt{n}[\hat{h}_n - h(n)]$ for any sequence $h(n) \to 0$. For $h(n) = o(1/n)$, one obtains (1.10) with $\bar{h} = 0$. For $h(n) = o(n^{-1/2})$ and $1/n = o(h(n))$, one obtains (1.10) with $\bar{h} = \infty$ [equivalently, (1.11) with $\bar{h} = 0$]. Also, Theorem 1.2 together with certain uniform estimates (see Section 2), imply that if $h(n) = o(1)$ and $n^{-1/2} = o(h(n))$, then one obtains (1.11) with $\bar{h} = \infty$ [equivalently, (1.7) with $h = 0$]. Note also that (1.8) is equivalent to (1.10) with $\bar{h} = 0$.

The second moment of $\lim_n \sqrt{n}[\hat{h}_n - h(n)]$ reflects the risk of the ML estimator. The second moment of Z in (1.10) is independent of $\bar{h}$ and equal to $1/2\sigma^2(0)$, while the second moment of Z in (1.11) increases from $1/2\sigma^2(0)$ to $1/\sigma^2(0)$ as $|\bar{h}|$ ranges from 0 to $+\infty$, thus filling the gap between these two values. The mean of Z in (1.10) is equal to

$$\frac{\tanh \bar{h}m_+}{\sqrt{2\pi\sigma^2(0)}}, \tag{1.12}$$

which increases, as $\bar{h}$ ranges from 0 to $+\infty$, from 0 to $(2\pi\sigma^2(0))^{-1/2}$. A straightforward computation shows that the mean of Z in (1.11a) decreases from $(2\pi\sigma^2(0))^{-1/2}$ to zero as $\bar{h}$ ranges from 0 to $+\infty$. Hence, for large n, the worst behavior of the bias of $\hat{h}_n$ is $(2\pi\sigma^2(0))^{-1/2}n^{-1/2}$. Note that this goes to zero like $n^{-1/2}$, in contrast to the typical behavior n^{-1}. Also note that the median of the limiting variable Z is zero in all cases.

The next theorem treats the critical point case $\beta = 1$, $h = 0$.

THEOREM 1.3. *Suppose that* $\beta = 1$ (*known*) *and that the true value of* h *is zero. Then under* $P_{n,0}(= P_{n,1,0})$, *we have*

$$n^{3/4}\hat{h}_n \to_D H, \tag{1.13}$$

where

$$H(a) = F(g'(a)), \qquad a \in \mathfrak{R}, \tag{1.14a}$$

$$g(t) = \log \int \exp\left\{-\tfrac{1}{12}\xi^4 + t\xi\right\} d\xi, \tag{1.14b}$$

$$dF(\xi) = \exp\left\{-\tfrac{1}{12}\xi^4 - g(0)\right\} d\xi. \tag{1.14c}$$

Note that at the critical point $\beta = 1$, $h = 0$, *the ML estimator* $\hat{h}_n$ *approaches the true value* $h = 0$ *at a speed* $n^{-3/4}$ *which is faster than the typical speed* $n^{-1/2}$.

In Theorems 1.1–1.3, β was assumed to be known and we estimated h. In the next theorem, h is assumed to be known and we treat the asymptotic behavior of the ML estimator $\hat{\beta}_n$ of β. For simplicity, we suppress the h-dependence of various quantities such as $P_{n,\beta,h}$, $\sigma^2(\beta, h)$, etc.

THEOREM 1.4. *Suppose that h is known and let $\hat{\beta}_n$ be the ML estimator of β. Then*

(a) *If $h \neq 0$, then under $P_{n,\beta}$,*

$$\sqrt{n}\,(\hat{\beta}_n - \beta) \to_D N\big(0,(\sigma(\beta)m(\beta))^{-2}\big). \tag{1.15}$$

(b) *If $h = 0$ and the true β satisfies $0 < \beta < 1$, then*

$$\hat{\beta}_n \to_D 1 - (1 - \beta)\frac{1}{\chi_1^2}, \tag{1.16}$$

where χ_1^2 is the chi-square distribution. In particular, $\hat{\beta}_n$ is not consistent.

(c) *If $h = 0$ and the true β satisfies $\beta > 1$, then*

$$\sqrt{n}\,(\hat{\beta}_n - \beta) \to_D N\big(0,(\sigma(0)m_+(\beta))^{-2}\big). \tag{1.17}$$

(d) *If $h = 0$ and the true β is equal to 1, then for $a \in \Re$,*

$$P_{n,1,0}\big\{\sqrt{n}\,(\hat{\beta}_n - 1) < a\big\} \to_{n\to\infty} F\Big(\Big[-\sqrt{v(a)}\,,\sqrt{v(a)}\,\Big]\Big), \tag{1.18}$$

where F is defined by (1.14c) and $v(a)$ is the variance of the probability measure $c\exp\{-\frac{1}{12}\xi^4 + (a/2)\xi^2\}\,d\xi$ on $\Re$.

Comparing part (b) of Theorem 1.1 and part (c) of Theorem 1.4, we see that at first order phase transition points the ML estimator $\hat{h}_n$, when β is known ($\beta > 1$), is superefficient, while the ML estimator $\hat{\beta}_n$ when $h(= 0)$ is known, is asymptotically normal. A similar behavior is expected for the Ising model. The natural statistic of the parameter h which exhibits superefficiency, is the *mean magnetization* $\overline{X}_n$. In the physics literature, this statistic is called *order parameter* [18] and is intrisically related to the occurence of first order phase transition. The notion of order parameters extends to general Gibbs distribution. For the general case we expect that at first order phase transition points, the parameters of the distributions which are associated with the order parameters [18] exhibit superefficiency, but the other parameters have a more typical behavior. All our major results go through for the nonbinary Curie–Weiss models [4] [(1.2), (1.3), (1.13) and (1.14) must be suitably changed].

A basic tool in our technical estimates is Kac's Gaussian transform. This transform leads to simple expansions of expectations via the Laplace method. A large deviation technique [3] for evaluating expectations is also possible, but our procedure is simpler and more straightforward.

The organization of the paper is as follows: In Section 2 we derive various technical results which are used in the proof of Theorems 1.1 and 1.2. These theorems are proven in Section 3. Theorem 1.3 is proven in Section 4 and Theorem 1.4 in Section 5.

2. Technical estimates. The proof of Theorems 1.1 and 1.2 will be based on Proposition 2.1 and Corollary 2.1 derived in this section. The proof of the proposition is based on two technical lemmas (Lemmas 2.1 and 2.2). We start with some preliminary properties of the Curie–Weiss model and the Gaussian transform.

A. *Preliminaries and the Gaussian transform.* The log-likelihood function for the Curie–Weiss model (1.1) reads

$$l_n(\overline{X}_n;\beta,h) = p_n(\beta,h) - \tfrac{1}{2}\beta\overline{X}_n^2 - h\overline{X}_n, \tag{2.1}$$

where

$$p_n(\beta,h) = \frac{1}{n}\log Z_n(\beta,h) \tag{2.2}$$

is called the pressure of the model. The ML equations read

$$m_n(\beta,h) = \overline{X}_n, \tag{2.3a}$$

$$u_n(\beta,h) = \tfrac{1}{2}\overline{X}_n^2, \tag{2.3b}$$

where

$$m_n(\beta,h) = \frac{\partial p_n(\beta,h)}{\partial h} = E_{n,\beta,h}(\overline{X}_n), \tag{2.4a}$$

$$u_n(\beta,h) = \frac{\partial p_n(\beta,h)}{\partial \beta} = \frac{1}{2}E_{n,\beta,h}(\overline{X}_n^2). \tag{2.4b}$$

Applying Schwarz's inequality, one easily obtains that if both β and h are unknown, equations (2.3a, b) do not have a solution on a set of probability 1. Hence β and h cannot be estimated simultaneously.

The Gaussian transform is based on the simple identity

$$\exp\left\{\frac{1}{2}\alpha y^2\right\} = \sqrt{\frac{\alpha}{2\pi}}\int_{\Re}\exp\left\{-\frac{1}{2}\alpha\eta^2 + \alpha y\eta\right\}d\eta, \qquad \alpha > 0. \tag{2.5}$$

Let $E\{\cdot\}$ denote expectation with respect to the measure $\frac{1}{2}\delta(x-1) + \frac{1}{2}\delta(x+1)$. Using (2.5) we obtain

$$\begin{aligned} &E\left\{\exp\left\{n\left(\frac{1}{2}\beta\overline{X}_n^2 + h\overline{X}_n\right)\right\}1_{\overline{X}_n\in A}\right\} \\ &\quad = \sqrt{\frac{\beta n}{2\pi}}\exp\left\{-n\frac{1}{2}\frac{h^2}{\beta}\right\}\int_{\Re}d\xi\exp\left\{-n\left(\frac{1}{2}\beta\xi^2 - h\xi\right)\right\} \\ &\qquad \times E\left\{\exp\left[n\beta\xi\overline{X}_n\right]1_{\overline{X}_n\in A}\right\}, \end{aligned} \tag{2.6}$$

where we have used the change of variables $\beta\eta + h = \beta\xi$. From (2.6) with

$A = [-1, 1]$, we obtain

$$Z_n(\beta, h) = \sqrt{\frac{\beta n}{2\pi}} \exp\left\{-n\frac{h^2}{2\beta}\right\} \int_{\mathfrak{R}} \exp\{-n f_{\beta, h}(\xi)\}\, d\xi, \tag{2.7}$$

where

$$f_{\beta, h}(\xi) = \tfrac{1}{2}\beta\xi^2 - h\xi - \log\cosh \beta\xi. \tag{2.8}$$

The minima of $f_{\beta, h}(\xi)$ will play a crucial role in our analysis. They satisfy the equation

$$\beta\xi - h = \beta \tanh \beta\xi. \tag{2.9}$$

We will use the following well-known properties [4], some of which will be consequences of our estimates in the remaining subsections:

1. For $\beta \le 1$, $f_{\beta, h}$ has a unique minimum to be denoted by $\zeta(\beta, h)$. It satisfies $\zeta(\beta, 0) = 0$.
2. For $h \neq 0$, or $\beta \le 1$, $f_{\beta, h}$ has a unique global minimum to be denoted also by $\zeta(\beta, h)$.
3. For $\beta > 1$ and $|h|$ sufficiently small, $f_{\beta, h}$ has two minima $\zeta^+(\beta, h) > 0$ and $\zeta^-(\beta, h) < 0$. For $h > 0$ (resp., $h < 0$), ζ^+ (resp., ζ^-) is the global minimum. For $h = 0$, $\zeta^+(\beta, 0) = -\zeta^-(\beta, h) = m_+(\beta) = -m_-(\beta)$, where $m_\pm$ were defined via (1.3).
4. For $h \neq 0$, one easily derives, using (2.4a), (2.2) and (2.7), that $m_n(\beta, h)$ has a limit $m(\beta, h)$ as $n \to \infty$, which satisfies

   $$m(\beta, h) = -\frac{h}{\beta} + \zeta(\beta, h), \tag{2.10}$$

   where $\zeta(\beta, h)$ is the unique global minimum of $f_{\beta, h}$.
5. The $m(\beta, h)$ of (2.10) satisfies (1.2) and is monotone in h. For $\beta \le 1$, $m(\beta, 0) = 0$. For $\beta > 1$, $m(\beta, h) \to m_+(\beta)$ as $h \downarrow 0$ and $m(\beta, h) \to m_-(\beta)$ as $h \uparrow 0$.
6. By symmetry, $m_n(\beta, 0) = 0$. From (2.4a) we obtain

   $$\frac{\partial m_n(\beta, h)}{\partial h} = \frac{\partial^2 p_n(\beta, h)}{\partial h^2} = n \operatorname{Var}_{n, \beta, h}(\overline{X}_n) > 0 \tag{2.11}$$

   for finite n. Hence $m_n(\beta, h)$ [and $m(\beta, h)$] is monotone in h (see Figure 1). Using (2.7) one can easily show that for $h \neq 0$ or $\beta \le 1$, the right-hand side of (2.11) converges as $n \to +\infty$ to

   $$\sigma^2(\beta, h) = \left(f''_{\beta, h}(\zeta)\right)^{-1} - \beta^{-1}, \tag{2.12}$$

   where ζ is as in (2.10). This $\sigma^2(\beta, h)$ is the same as in (1.5).
7. From (2.1), we obtain that the Fisher information (per observation), when β is known, is given by $(\partial^2 pn(\beta, h))/\partial h^2$. Hence for $h \neq 0$ or $\beta \le 1$, the Fisher information has a limit equal to $\sigma^2(\beta, h)$. For $\beta \ge 1$, $h = 0$, the Fisher information converges to $+\infty$ as $n \to +\infty$.

B. *Two technical lemmas.* In the rest of this section we will suppress the β-dependence of the various objects that occur.

LEMMA 2.1. *Let g be a measurable function on $\Re$ such that $|g(x)| \le \exp\{c|x|\}$ for some $c > 0$. Let I be an interval such that $f_h(\xi)$ has a unique minimum $\zeta(h)$ on the closure $\bar{I}$ of I with $\zeta \in \dot{I}$ ($\dot{I}$ = interior of I). If $f_h''(\zeta) > 0$ (which holds for all $\beta \ne 1$ or $h \ne 0$), then*

$$(2.13)\quad \begin{aligned}&\sqrt{n}\int_I d\xi\, g(\sqrt{n}(\xi - \zeta))\exp\{-n[f_h(\xi) - f_h(\zeta)]\}\\ &\qquad = \int_{\Re} d\xi\, g(\xi)\exp\{-\tfrac{1}{2}f_h''(\zeta)\xi^2\} + O_{c,h}(n^{-1/2}).\end{aligned}$$

PROOF. Let $J_n(I)$ denote the left-hand side of (2.13). Let $J_n(a) = J_n([\zeta - a, \zeta + a])$ for positive a (to be chosen next) and

$$(2.14)\qquad A = I \cap \{\xi: a \le |\xi - \zeta|\}.$$

Then

$$(2.15)\qquad J_n(I) = J_n(a) + J_n(A).$$

By Taylor's formula,

$$f_h(\xi) = f_h(\zeta) + \tfrac{1}{2}f_h''(\zeta)(\xi - \zeta)^2 + R$$

with $|R| \le c_1|\xi - \zeta|^3$. Changing $\sqrt{n}(\xi - \zeta)$ to ξ, we obtain

$$(2.16)\qquad J_n(a) = \int_{|\xi| \le a\sqrt{n}} d\xi\, g(\xi)\exp\{-\tfrac{1}{2}f_h''(\zeta)\xi^2 + R_n\}$$

with $|R_n| \le c_1|\xi|^3 n^{-1/2}$. Let

$$I_n(a) = \int_{|\xi| \le a\sqrt{n}} d\xi\, g(\xi)\exp\{-\tfrac{1}{2}f_h''(\zeta)\xi^2\}.$$

Using the inequality $|e^u - 1| \le |u|e^{|u|}$, we obtain

$$(2.17)\qquad |J_n(a) - I_n(a)| \le K_n(a),$$

with

$$(2.18)\qquad K_n(a) \le \int_{|\xi| \le a\sqrt{n}} d\xi |R_n| g(\xi)\exp\{-\tfrac{1}{2}f_h''(\zeta)\xi^2 + |R_n|\}.$$

On the set $\{\xi: |\xi| \le a\sqrt{n}\}$, we have $|R_n| \le ac_1|\xi|^2$. Now we choose a so that $ac_1 \le \frac{1}{4}f_h''(\zeta)$ and $[\zeta - a, \zeta + a] \subset I$. Then

$$K_n(a) \le c_1 n^{-1/2}\int_{\Re}|\xi|^3\exp\{-\tfrac{1}{4}f_h''(\zeta)\xi^2 + c|\xi|\}\,d\xi.$$

Hence $K_n(a) = O_{h,c}(n^{-1/2})$. By a standard estimate on the tail of the normal distribution, we have that

$$\int_{|\xi| \geq a\sqrt{n}} d\xi \exp\{-\tfrac{1}{2} f_h''(\zeta)\xi^2 + c|\xi|\}$$

is exponentially small. Hence (2.16) implies

$$J_n(a) = \int_{\mathfrak{R}} d\xi\, g(\xi) \exp\{-\tfrac{1}{2} f_h''(\zeta)\xi^2\} + O_{h,\zeta}(n^{-1/2}). \tag{2.19}$$

Next we show that $J_n(A)$ in (2.15) is exponentially small. To this end, we establish the following general fact: For all $\gamma > 0$, $d > 0$, there exists a $\delta = \delta(\gamma, d)$ such that for any Borel set A with $d \leq \operatorname{dist}(\zeta, A)$ and

$$\gamma \leq \inf\left\{ \frac{1}{(\xi - \zeta)^2} [f_h(\xi) - f_h(\zeta)] : \xi \in A \right\}, \tag{2.20}$$

we have

$$J_n(A) \leq \delta^{-1} \exp\{-n\delta\}. \tag{2.21}$$

Indeed, changing variables, we obtain

$$\begin{aligned} J_n(A) &\leq \sqrt{n} \int_A d\xi \exp\{-n\gamma(\xi - \zeta)^2 + c\sqrt{n}|\xi - \zeta|\} \\ &\leq 2 \int_{\xi \geq d\sqrt{n}} d\xi \exp\left\{ -\gamma\left(\xi - \frac{c}{2\gamma}\right)^2 + \frac{c^2}{4\gamma} \right\}, \end{aligned}$$

which is exponentially small. Now, from the definition of f_h we have

$$\lim_{\xi \to \infty} \frac{1}{(\xi - \zeta)^2} [f_h(\xi) - f_h(\zeta)] = \frac{\beta}{2} > 0.$$

Hence, if A is chosen to be of the form

$$A = I \cap \{\xi : a \leq |\xi - \zeta| \leq a'\}, \qquad a' > a,$$

then the infimum in (2.20) is strictly positive (since ζ is the unique minimum of f_h on $\bar{A}$). But this is true for all $a' > a$. Hence it is true for the set A of (2.14). Thus (2.21) together with (2.19) yields (2.13). □

REMARK. Although not needed in the proofs of our theorems, we note that (2.13) may be strengthened in two ways: (a) if g is even, then the error in (2.13) is the order $O_{c,h}(n^{-1})$. (b) If $\beta > 1$, $I = \mathfrak{R}^+$ (resp., $\mathfrak{R}^-$), and $\xi = \xi^+(h)$ [resp., $\xi^-(h)$], then there exists $h_0 > 0$ such that estimate (2.13) is uniform in h for $|h| \leq h_0$. The proof of these refinements is not included here, since we do not need these improvements.

LEMMA 2.2. *Let g be as in Lemma 2.1. Suppose that $\beta > 1$, $h = 0$, and let m_ε, $\varepsilon = \{+1, -1\}$ be the two global minima of $f_0(\xi)$. Then*

$$\begin{aligned} &\sqrt{n} \int_{\mathfrak{R}} d\xi\, g(\sqrt{n}(\xi - m_\varepsilon)) \exp\{-n\tfrac{1}{2}\beta\xi^2\} E\{\exp[n\beta\xi\bar{X}_n] 1_{\operatorname{sign}(\bar{X}_n) = \varepsilon}\} \\ &\quad = \exp\{-nf_0(m_+)\} \left\{ \int_{\mathfrak{R}} d\xi\, g(\xi) \exp[-\tfrac{1}{2} f_0''(m_+)\xi^2] + O_c(n^{-1/2}) \right\}, \end{aligned} \tag{2.22}$$

where

$$\operatorname{sign}(x) = \begin{cases} +1, & \text{if } x > 0, \\ -1, & \text{if } x \le 0. \end{cases}$$

PROOF. Let J_n denote the left-hand side of (2.22). We write

$$J_n = I_n + R_{n,1} - R_{n,2}, \tag{2.23a}$$

where

$$I_n = \sqrt{n}\int_{\varepsilon\xi>0} d\xi\, g\big(\sqrt{n}\,(\xi - m_\varepsilon)\big)\exp\{-n\tfrac{1}{2}\beta\xi^2\}E\{\exp[n\beta\xi\overline{X}_n]\}, \tag{2.23b}$$

$$\begin{aligned} R_{n,1} = \sqrt{n}\int_{\varepsilon\xi\le 0} & d\xi\, g\big(\sqrt{n}\,(\xi - m_\varepsilon)\big)\exp\{-n\tfrac{1}{2}\beta\xi^2\} \\ & \times E\{\exp[n\beta\xi\overline{X}_n]1_{\operatorname{sign}(\overline{X}_n)=\varepsilon}\}, \end{aligned} \tag{2.23c}$$

$$\begin{aligned} R_{n2} = \sqrt{n}\int_{\varepsilon\xi>0} & d\xi\, g\big(\sqrt{n}\,(\xi - m_\varepsilon)\big)\exp\{-n\tfrac{1}{2}\beta\xi^2\} \\ & \times E\{\exp[n\beta\xi\overline{X}_n]1_{\operatorname{sign}(\overline{X}_n)=-\varepsilon}\}. \end{aligned} \tag{2.23d}$$

Now, I_n can be written as

$$I_n = \exp\{-nf_0(m_+)\}\sqrt{n}\int_{\varepsilon\xi>0} d\xi\, g\big(\sqrt{n}\,(\xi - m_\varepsilon)\big)\exp\{-n[f_0(\xi) - f_0(m_\varepsilon)]\}$$

and by Lemma 2.1,

$$I_n = \exp\{-nf_0(m_+)\}\left\{\int_{\Re} d\xi\, g(\xi)\exp\left[-\tfrac{1}{2}f_0''(m_+)\xi^2\right] + O_c(n^{-1/2})\right\}. \tag{2.24}$$

The first error term $R_{n,1}$ is bounded as

$$\begin{aligned} |R_{n,1}| &\le \sqrt{n}\int_{\varepsilon\xi\le 0} d\xi\, \exp\{-n\tfrac{1}{2}\beta\xi^2 + c\sqrt{n}\,|\xi - m_\varepsilon|\} \\ &= O_c\big(\exp\{c\sqrt{n}\,m_+\}\big) \end{aligned} \tag{2.25}$$

and the second error term $R_{n,2}$ as

$$\begin{aligned} |R_{n,2}| &\le \sqrt{n}\int_{\varepsilon\xi> 0} d\xi\, \exp\{-n\tfrac{1}{2}\beta\xi^2 + c\sqrt{n}\,|\xi - m_\varepsilon|\} \\ &= O_c\big(\exp\{c\sqrt{n}\,m_+\}\big). \end{aligned} \tag{2.26}$$

Combining (2.23)–(2.26) we obtain (2.22). □

C. *Asymptotics of the magnetization and central limit theorems for* $\overline{X}_n$.

PROPOSITION 2.1. *Let* $\bar{h} \in \Re$. *Then* (a) *For* $h \neq 0$ *or* $\beta < 1$, *we have under* $P_{n,h+\bar{h}/\sqrt{n}}$,

$$\sqrt{n}\,(\overline{X}_n - m(h)) \to_D N(\sigma^2(h)\bar{h}, \sigma^2(h)). \tag{2.27}$$

(b) *For $\beta > 1$ and $h = 0$, we have under $P_{n,\bar{h}/\sqrt{n}}$*,

$$\left(\sqrt{n}\,(\bar{X}_n - m_+)\middle|\bar{X}_n > 0\right) \to_D N(\sigma^2(0)\bar{h}, \sigma^2(0)), \tag{2.28a}$$

$$\left(\sqrt{n}\,(\bar{X}_n - m_-)\middle|\bar{X}_n \le 0\right) \to_D N(\sigma^2(0)\bar{h}, \sigma^2(0)), \tag{2.28b}$$

and under $P_{n,\bar{h}/n}$,

$$\left(\sqrt{n}\,(\bar{X}_n - m_+)\middle|\bar{X}_n > 0\right) \to_D N(0, \sigma^2(0)), \tag{2.29a}$$

$$\left(\sqrt{n}\,(\bar{X}_n - m_-)\middle|\bar{X}_n \le 0\right) \to_D N(0, \sigma^2(0)). \tag{2.29b}$$

REMARK 1. In (2.28) and (2.29) we use the notation: For a sequence of random variables Y_n, a sequence of events A_n and a probability distribution F on $\mathfrak{R}$, we write $(Y_n|A_n) \to_D F$ to mean $P(Y_n \in dy|A_n) \to dF(y)$.

REMARK 2. For $\bar{h} = 0$, (2.28) is proven in [5] and (2.27) in [4]. The proofs in [4] and [5] use the characteristic function and can be extended to $\bar{h} \neq 0$. Our proofs use the Laplace transform and are easy consequences of Lemma 2.1 and 2.2. The Laplace transform yields also convergence of moments needed in the proofs of the theorems.

PROOF OF PROPOSITION 2.1. We will work with the Laplace transform instead of the characteristic function

(a)

$$\begin{aligned} &E_{n,h+\bar{h}n^{-1/2}}\left\{\exp\left\{t\sqrt{n}\,(\bar{X}_n - m(h))\right\}\right\} \\ &\quad = \exp\left\{-t\sqrt{n}\,m(h)\right\}\frac{Z_n\left(h + (t+\bar{h})/\sqrt{n}\right)}{Z_n\left(h + \bar{h}/\sqrt{n}\right)}. \end{aligned} \tag{2.30a}$$

Using (2.7) and Lemma 2.1 with $g(\xi) = \exp\{(t+\bar{h})\xi\}$, $I = \mathfrak{R}$, one easily obtains

$$\begin{aligned} Z_n\left(h + \frac{t+\bar{h}}{\sqrt{n}}\right) &= \sqrt{\frac{\beta}{f_h''(\xi)}}\exp\left\{-n\frac{h^2}{2\beta} - nf_h(\zeta) + \sqrt{n}\,(t+\bar{h})m(h)\right\} \\ &\quad \times\exp\left\{\frac{1}{2}\sigma^2(h)(t+\bar{h})^2\right\}\left[1 + O(n^{-1/2})\right], \end{aligned} \tag{2.30b}$$

where we have used $m(h) = -h/\beta + \zeta(h)$ and $\sigma^2(h) = (f_h''(\zeta))^{-1} - \beta^{-1}$. From this we deduce that the left-hand side of (2.30a) is equal to

$$\exp\left\{\tfrac{1}{2}\sigma^2(h)(t^2 + 2t\bar{h})\right\} + O(n^{-1/2}),$$

which yields (2.27).

(b) We prove only (2.28a). The proof of (2.28b) is identical. We write

$$\begin{aligned} & E_{n,\bar{h}n^{-1/2}}\left\{\exp\left\{t\sqrt{n}\,(\bar{X}_n - m_+)\right\}\middle|\bar{X}_n > 0\right\} \\ & \quad = \exp\left\{-t\sqrt{n}\,m_+\right\}\frac{Z_n^+\left((t+\bar{h})/\sqrt{n}\right)}{Z_n^+\left(\bar{h}/\sqrt{n}\right)}, \end{aligned} \tag{2.31a}$$

where

$$Z_n^+\left(\frac{t+\bar{h}}{\sqrt{n}}\right) = E\left\{\exp\left\{n\left(\frac{1}{2}\beta\bar{X}_n^2 + \frac{t+\bar{h}}{\sqrt{n}}\bar{X}_n\right)\right\}\middle|1_{\bar{X}_n>0}\right\}. \tag{2.31b}$$

Using (2.6), we have

$$\begin{aligned} Z_n^+\left(\frac{t+\bar{h}}{\sqrt{n}}\right) &= \sqrt{\frac{\beta}{2\pi}}\exp\left\{-\frac{(t+\bar{h})^2}{2\beta}\right\}\exp\left\{\sqrt{n}\,(t+\bar{h})m_+\right\} \\ &\quad \times\sqrt{n}\int d\xi\,\exp\left\{\sqrt{n}\,(t+\bar{h})(\xi - m_+)\right\} \\ &\quad \times\exp\left\{-n\frac{1}{2}\beta\xi^2\right\} \\ &\quad \times E\left\{\exp\left[n\beta\xi\bar{X}_n\right]1_{\bar{X}_n>0}\right\}. \end{aligned}$$

Applying Lemma 2.2 with $g(\xi) = \exp\{(t+\bar{h})\xi\}$, we obtain

$$\begin{aligned} Z_n^+\left(\frac{t+\bar{h}}{\sqrt{n}}\right) &= \sqrt{\frac{\beta}{f_0''(m_+)}}\exp\left\{-nf_0(m_+) + \sqrt{n}\,(t+\bar{h})m_+\right\} \\ &\quad \times\exp\left\{\frac{1}{2}\sigma^2(0)(t+\bar{h})^2\right\} + O_{t,\bar{h}}(n^{-1/2}), \end{aligned} \tag{2.32}$$

where we have used $\sigma^2(0) = (f_0''(m_+))^{-1} - \beta^{-1}$. This implies that the left-hand side of (2.31) is equal to

$$\exp\left\{\tfrac{1}{2}\sigma^2(0)(t^2 + 2t\bar{h})\right\}\left[1 + O_{t,\bar{h}}(n^{-1/2})\right],$$

which yields (2.28a). The proof of (2.29) can be given along a similar line and we do not spell out the details [notice that (2.29) is formally obtained by letting $\bar{h}\to 0$]. □

Since convergence of the Laplace transform implies convergence of moments, we obtain the following corollary which will be used in the proof of Theorem 1.1.

COROLLARY 2.1. (a) *For $h \neq 0$ or $\beta < 1$, we have*

$$m_n\left(h + \frac{\bar{h}}{\sqrt{n}}\right) = m(h) + \sigma^2(h)\frac{\bar{h}}{\sqrt{n}} + o(n^{-1/2}). \tag{2.33a}$$

(b) *For $\beta > 1$, we have*

$$(2.33b) \quad m_n\left(h + \frac{\bar{h}}{\sqrt{n}}\right) = m_+ + \sigma^2(0)\frac{\bar{h}}{\sqrt{n}} + o(n^{-1/2}), \qquad \bar{h} > 0,$$

$$(2.33c) \quad m_n\left(h + \frac{\bar{h}}{\sqrt{n}}\right) = m_- + \sigma^2(0)\frac{\bar{h}}{\sqrt{n}} + o(n^{-1/2}), \qquad \bar{h} < 0.$$

PROOF. Part (a) is a straightforward consequence of (2.27). The proofs of (2.33b) and (2.33c) are identical and we treat only the case $\bar{h} > 0$. The convergence of the expectation in (2.29a) shows that

$$(2.34) \qquad E_{n,\bar{h}/\sqrt{n}}\left(\bar{X}_n \middle| \bar{X}_n > 0\right) = m_+ + \sigma^2(0)\frac{\bar{h}}{\sqrt{n}} + o(n^{-1/2}).$$

By $Z_n^+(\bar{h} = Z_n^-(-\bar{h}))$ and (2.32), we have

$$(2.35) \qquad P_{n,\bar{h}/\sqrt{n}}\left(\bar{X}_n > 0\right) = \frac{Z_n^+(\bar{h}/\sqrt{n})}{Z_n^+(\bar{h}/\sqrt{n}) + Z_n^-(\bar{h}/\sqrt{n})} \to 1$$

as $n \to \infty$. This together with (2.34) yields (2.33b). □

REMARKS. (1) The convergence of the Laplace transform also implies that for $h \neq 0$ or $\beta < 1$, the Fisher information $n\,\mathrm{Var}_{n,\beta,h}(\bar{X}_n)$ converges to $\sigma^2(\beta, h)$ given by (1.5) [(see vii) of Section 2A].

(2) We have a uniform expansion of the finite size magnetization for β and h near first order phase transition points, which implies (2.33b) and (2.33c): For $\beta > 1$, we have

$$(2.36) \qquad m_n(h) = m_+ \tanh(nhm_+) + \sigma^2(o)h + O\left(\max\left(\frac{1}{n}, h^2\right)\right)$$

uniformly in h form small $|h|$. The proof of this expansion is lengthy and involves a uniform (and refined) version of Lemma 2.1. Since it is not used in this paper and in order to reduce the size of the paper, we do not provide the proof of (2.36) here. The chief interest of this expansion is in providing a rigorous analysis of the *rounding* of a first order phase transition [18]. The analogues expansion for the Ising model is an interesting and important open question. The interest of (2.36) in our context, is in providing the quantitative behavior of $m_n(h)$ as is shown in Figure 1 at the end of the paper.

3. Proofs of Theorems 1.1 and 1.2.

PROOF OF THEOREM 1.1. (a) Since $m_n(h)$ is an increasing function of h, the ML equation (2.3a) yields for any $a \in \mathfrak{R}$,

$$P_{n,h}\{\sqrt{n}(\hat{h}_n - h) < a\} = P_{n,h}\left\{\bar{X}_n < m_n\left(h + \frac{a}{\sqrt{n}}\right)\right\}.$$

Thus, by Corollary 2.1(a),

$$P_{n,h}\{\sqrt{n}(\hat{h}_n - h) < a\} = P_{n,h}\left\{\sqrt{n}(\bar{X}_n - m(h)] < \sigma^2(h)a + O_{h,a}\left(\frac{1}{\sqrt{n}}\right)\right\}.$$

The right-hand side of this converges, by Proposition 2.1(a), to

$$\Pr\{N(0,\sigma^2(h)) < \sigma^2(h)a\} = \mathrm{Prob}\left\{N\left(0, \frac{1}{\sigma^2(h)}\right) < a\right\},$$

which proves (1.7).

(b) Let $a < 0$. By the monotonicity of $m_n(h)$ and the ML equation (2.3a), we have

$$P_{n,0}\{\sqrt{n}\,\hat{h}_n < a\} = P_{n,0}\left\{\bar{X}_n < m_n\left(\frac{a}{\sqrt{n}}\right)\middle|\bar{X}_n \le 0\right\}P_{n,0}(\bar{X}_n \le 0)$$

and by (2.33c),

$$P_{n,0}\{\sqrt{n}\,\hat{h}_n < a\} = P_{n,0}\left\{\sqrt{n}(\bar{X}_n - m_-) < \sigma^2(0)a + O_a\left(\frac{1}{\sqrt{n}}\right)\middle|\bar{X}_n \le 0\right\}P_{n,0}\{\bar{X}_n \le 0\}. \tag{3.1}$$

By Proposition 2.1(b), the conditional probability in (3.1) converges to $\Pr\{N(0, 1/\sigma^2(0)) < a\}$. Now $P_{n,0}(\bar{X}_n \le 0)$ and $P_{n,0}(\bar{X}_n > 0)$ sum to 1 and by symmetry, differ only by $P_{n,0}(\bar{X}_n = 0)$ which is exponentially small as $n \to \infty$. Hence, both $P_{n,0}(\bar{X}_n \le 0)$ and $P_{n,0}(\bar{X}_n > 0)$ converge to $\frac{1}{2}$. Thus, for $a < 0$,

$$\lim_{n\to\infty} P_{n,0}\{\sqrt{n}\,\hat{h}_n < a\} = \frac{1}{2}\Pr\left\{N\left(0, \frac{1}{\sigma^2(0)}\right) < a\right\}. \tag{3.2}$$

Now let $a > 0$ and write

$$P_{n,0}\{\sqrt{n}\,\hat{h}_n < a\} = P_{n,0}\{\hat{h}_n \le 0\} + P_{n,0}\left\{0 < \bar{X}_n < m_n\left(\frac{a}{\sqrt{n}}\right)\middle|\bar{X}_n > 0\right\}P_{n,0}(\bar{X}_n > 0).$$

Using (2.33b) and (2.28a), we obtain (for $a > 0$),

$$\begin{aligned} P_{n,0}\{\sqrt{n}\,\hat{h}_n < a\} &= P_{n,0}\{\bar{X}_n \le 0\} \\ &\quad + P_{n,0}\left\{-\sqrt{n}\,m_+ \le \sqrt{n}(\bar{X}_n - m_+) < \sigma^2(0)a + O_a\left(\frac{1}{\sqrt{n}}\right)\middle|\bar{X}_n > 0\right\}P_{n,0}\{\bar{X}_n > 0\} \\ &\to_{n\to\infty} \frac{1}{2} + \frac{1}{2}\Pr\left\{N\left(0, \frac{1}{\sigma^2(0)}\right) < a\right\}. \end{aligned} \tag{3.3}$$

This together with (3.2) yields (1.8). □

PROOF OF THEOREM 1.2. We proceed as in the proof of Theorem 1.1: (a) For $a < 0$, we have

$$\begin{aligned} &P_{n,\bar{h}/n}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{n}\right) < a\right\} \\ &\quad = P_{n,\bar{h}/n}\left\{\bar{X}_n < m_n\left(\frac{a}{\sqrt{n}} + \frac{\bar{h}}{n}\right)\right\} \\ &\quad = P_{n,\bar{h}/n}\left\{\sqrt{n}\,(\bar{X}_n - m_-) \right. \\ &\qquad\qquad \left. < \sigma^2(0)a + O_{a,\bar{h}}\left(\frac{1}{\sqrt{n}}\right) \middle| \bar{X}_n < 0\right\} P_{n,\bar{h}/n}(\bar{X}_n < 0). \end{aligned} \tag{3.4}$$

Now

$$P_{n,\bar{h}/n}(\bar{X}_n < 0) = \frac{Z_n^-(\bar{h}/n)}{Z_n(\bar{h}/n)},$$

where $Z_n^-(h)$ [and $Z_n^+(h)$] are defined as in (2.31b). Using the Gaussian transform as in (2.31b), one can easily show that

$$P_{n,\bar{h}/n}(\bar{X}_n < 0) \to_{n\to\infty} 1 - \alpha(\bar{h}). \tag{3.5}$$

By (2.29b), the conditional probability in (3.4) converges to

$$\Pr\{Y < a\}.$$

Hence, for $a < 0$,

$$P_{n,\bar{h}/n}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{n}\right) < a\right\} \to_{n\to\infty} (1 - \alpha(\bar{h}))\Pr\{y < a\}. \tag{3.6}$$

Now, for $a > 0$, we write

$$\begin{aligned} &P_{n,\bar{h}/n}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{n}\right) < a\right\} \\ &\quad = P_{n,\bar{h}/n}\{\bar{X}_n \le 0\} \\ &\qquad + P_{n,\bar{h}/n}\left\{-\sqrt{n}\,m_+ \le \sqrt{n}\,(\bar{X}_n - m_+) \right. \\ &\qquad\qquad \left. < \sigma^2(0)a + O_{a,\bar{h}}\left(\frac{1}{\sqrt{n}}\right) \middle| \bar{X}_n > 0\right\} P_{n,\bar{h}/n}(\bar{X}_n > 0). \end{aligned}$$

As in (3.5), we have

$$P_{n,\bar{h}/n}(\bar{X}_n > 0) \to_{n\to\infty} \alpha(\bar{h}).$$

This together with (3.5) and (2.29a) yield for $a > 0$,

$$(3.7)\quad P_{n,\bar{h}/n}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{n}\right) < a\right\} \to_{n\to\infty} [1 - \alpha(\bar{h})] + \alpha(\bar{h})\Pr\{Y < a\}.$$

This together with (3.6) is equivalent to (1.10).

(b) We will prove only (1.11a). The proof of (1.11b) is similar. Thus let $\bar{h} > 0$. For $a < -\bar{h}$, we have

$$P_{n,\bar{h}/\sqrt{n}}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) < a\right\} = P_{n,\bar{h}/\sqrt{n}}\left\{\bar{X}_n < m_n\left(\frac{a+\bar{h}}{\sqrt{n}}\right)\right\}$$
$$\le P_{n,\bar{h}/\sqrt{n}}\{\bar{X}_n \le 0\},$$

which tends to zero by (2.35). Hence, for $a < -\bar{h}$, $\bar{h} > 0$,

$$(3.8)\qquad P_{n,\bar{h}/\sqrt{n}}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) < a\right\} \to_{n\to\infty} 0.$$

For $a > -\bar{h}$, we have

$$P_{n,\bar{h}/\sqrt{n}}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) < a\right\}$$
$$= P_{n,\bar{h}/\sqrt{n}}\left\{\bar{X}_n < m_n\left(\frac{a+\bar{h}}{\sqrt{n}}\right)\right\}$$
$$(3.9)\quad = P_{n,\bar{h}/\sqrt{n}}(\bar{X}_n \le 0)$$
$$+ P_{n,\bar{h}/\sqrt{n}}\left\{-\sqrt{n}\,m_+ \le \sqrt{n}(\bar{X}_n - m_+) < \sigma^2(0)(a+\bar{h}) + O\left(\frac{1}{\sqrt{n}}\right)\middle|\bar{X}_n > 0\right\}P_{n,\bar{h}/\sqrt{n}}(\bar{X}_n > 0).$$

Now (2.35) together with (2.28a) and (3.9) yield (for $a > -\bar{h}$, $\bar{h} > 0$),

$$(3.10)\qquad P_{n,\bar{h}/\sqrt{n}}\left\{\sqrt{n}\left(\hat{h}_n - \frac{\bar{h}}{\sqrt{n}}\right) < a\right\} \to_{n\to\infty} \Pr\{Y < a\}.$$

This together with (2.9) is equivalent to (1.11a). □

4. Proof of Theorem 1.3. The proof of Theorem 1.3 proceeds along the lines of the proof of Theorem 1.1. In place of Proposition 2.1, we have the following limit theorem at the critical point $\beta = 1$, $h = 0$.

PROPOSITION 4.1. *For $\beta = 1$ and $h = 0$, we have under $P_{n,\bar{h}/n^{3/4}}$, $\bar{h} \in \mathfrak{R}$,*

$$(4.1)\qquad n^{1/4}\bar{X}_n \to_D F_{\bar{h}},$$

where

$$dF_{\bar{h}}(x) = \exp\{-\tfrac{1}{12}x^4 + \bar{h}x - g(\bar{h})\}\,dx$$

with $g(t)$ defined in (1.14b).

PROOF. We estimate the Laplace transform

$$E_{n,\bar{h}n^{-1/4}}\{\exp\{tn^{1/4}\bar{X}_n\}\} = \frac{Z_n((t+\bar{h})n^{-3/4})}{Z_n(\bar{h}n^{-3/4})}. \tag{4.2}$$

From (2.7) we have

$$Z_n(\bar{h}n^{-3/4}) = \sqrt{\frac{n}{2\pi}}\exp\left\{-n^{-1/2}\frac{\bar{h}^2}{2}\right\}\int d\xi\exp\{-nf_0(\xi)+n^{1/4}\bar{h}\xi\}. \tag{4.3}$$

For $\beta = 1$, $f_0(\xi)$ has only one minimum $m(0) = 0$. At the minimum $\xi = 0$, we have

$$f_0(0) = f_0'(0) = f_0''(0) = f_0'''(0) = 0,$$
$$f_0^{(4)}(0) = 2.$$

Thus near the minimum $\xi = 0$,

$$f_0(\xi) = \frac{2}{4!}\xi^4 + R.$$

Using this and following the proof of Lemma 2.1, one can show that

$$\begin{aligned} n^{1/4}\int_{\Re} d\xi\exp\{-nf_0(\xi)+\bar{h}n^{1/4}\xi\} \\ = \int_{\Re} d\xi\exp\{-\tfrac{1}{12}\xi^4+\bar{h}\xi\} + o(n^{-1/4}) \\ = \exp\{g(\bar{h})\} + o(n^{-1/4}). \end{aligned} \tag{4.4}$$

From this and (4.2) we obtain

$$E_{n,\bar{h}n^{-3/4}}\{\exp\{tn^{1/4}\bar{X}_n\}\} \to_{n\to\infty} \exp\{g(t+\bar{h}) - g(\bar{h})\}, \tag{4.5}$$

which yields (4.1). □

REMARK. For $\bar{h} = 0$, Proposition 4.1 is contained in [4].

The following corollary is an easy consequence of the proof of Proposition 4.1 and will be used in the proof of Theorem 1.3.

COROLLARY 4.1. *For $\beta = 1$, we have*

$$m_n(\bar{h}n^{-3/4}) = g'(\bar{h})n^{-1/4} + o(n^{-1/4}).$$

PROOF OF THEOREM 1.3. For $a \in \Re$,

$$\begin{aligned} P_{n,0}\{n^{3/4}\hat{h}_n < a\} &= P_{n,0}\{\bar{X}_n < m_n(an^{-3/4})\} \\ &= P_{n,0}\{n^{1/4}\bar{X}_n < g'(a) + o(1)\}. \end{aligned} \tag{4.6}$$

By Proposition 4.1 with $\bar{h} = 0$, the right-hand side of (4.6) converges to $F_0(g'(a))$, which implies the theorem. □

5. Proof of Theorem 1.4. The proof of Theorem 1.4 proceeds along lines similar to the proofs of Theorems 1.1 and 1.3. For each case (a)–(d) of Theorem 1.4, we need the analogues of Proposition 2.1 and Corollary 2.1 (or of Proposition 4.1 and Corollary 4.1). We shall see that the technical estimates for the present theorem reduce to the technical estimates of Section 2. Next, we sometimes suppress the h-dependence for convenience.

(a) For $h \neq 0$, we have from Proposition 2.1(a),

$$\begin{aligned}&\sqrt{n}\,(\bar{X}_n^2 - m^2(\beta,h))\\ &\quad= \sqrt{n}\,(\bar{X}_n - m(\beta,h))(\bar{X}_n + m(\beta,h))\\ &\quad\to_D N(0, 4\sigma^2(\beta,h)m^2(\beta,h))\end{aligned} \tag{5.1}$$

under $P_{n,\beta,h}$. We will now show that the $u_n(\beta) = u_n(\beta,h)$ defined by (2.4b) satisfies for $\bar{\beta} \in \Re$,

$$u_n\left(\beta + \frac{\bar{\beta}}{\sqrt{n}}\right) = \frac{1}{2}m^2(\beta) + \sigma^2(\beta)m^2(\beta)\frac{\bar{\beta}}{\sqrt{n}} + o(n^{-1/2}). \tag{5.2}$$

To prove this we consider the Laplace transform

$$\begin{aligned}&E_{n,\beta+\bar{\beta}n^{-1/2},h}\left\{\exp\left\{t\sqrt{n}\,[\bar{X}_n - m(\beta,h)]\right\}\right\}\\ &\quad= \exp\{-t\sqrt{n}\,m(\beta,h)\}\frac{Z_n(\beta + (\bar{\beta}/\sqrt{n}), h + (t/\sqrt{n}))}{Z_n(\beta + (\bar{\beta}/\sqrt{n}), h)},\end{aligned} \tag{5.3}$$

where

$$\begin{aligned}&Z_n\left(\beta + \frac{\bar{\beta}}{\sqrt{n}}, h + \frac{t}{\sqrt{n}}\right)\\ &\quad= E\left\{\exp\left\{n\left[\frac{1}{2}\beta\bar{X}_n^2 + \left(h + \frac{t}{\sqrt{n}}\right)\bar{X}_n\right]\right\}\exp\left\{\sqrt{n}\,\frac{1}{2}\bar{\beta}\bar{X}_n^2\right\}\right\}\\ &\quad= \exp\left\{-\sqrt{n}\,\frac{1}{2}\bar{\beta}m^2(\beta,h)\right\}\\ &\qquad\times E\left\{\exp\left\{n\left[\frac{1}{2}\beta\bar{X}_n^2 + \left(h + \frac{t+\bar{\beta}m(\beta,h)}{\sqrt{n}}\right)\bar{X}_n\right]\right\}\right.\\ &\qquad\qquad\left.\times\varphi_{n,\bar{\beta}}\left(\sqrt{n}\,[\bar{X}_n - m(\beta,h)]\right)\right\}\\ &\quad= \exp\left\{-\sqrt{n}\,\frac{1}{2}\bar{\beta}m^2(\beta,h)\right\}Z_n\left(\beta, h + \frac{t+\bar{\beta}m(\beta,h)}{\sqrt{n}}\right)\\ &\qquad\times E_{n,\beta,h+(t+\bar{\beta}m)/\sqrt{n}}\left\{\varphi_{n,\bar{\beta}}\left(\sqrt{n}\,[\bar{X}_n - m]\right)\right\},\end{aligned} \tag{5.4}$$

with

$$\varphi_{n,\bar{\beta}}(x) = \exp\left[\frac{1}{2\sqrt{n}}\bar{\beta}x^2\right].$$

The random variable $\varphi_{n,\bar\beta}(\sqrt{n}[\bar X_n - m(\beta,h)])$ converges as $n \to +\infty$ to 1 in probability under $P_{n,\beta,h+(\bar\beta m(\beta,h)/\sqrt{n}}$ and it is uniformly integrable with respect to the same measure, since it is bounded above by $\exp\{|\bar\beta\sqrt{n}[\bar X_n - m(\beta,h)]|\}$. Hence, the expectation in (5.4) tends to 1 as $n \to +\infty$ and therefore, we have

$$\frac{Z_n\big(\beta + (\bar\beta/\sqrt{n}), h + (t/\sqrt{n})\big)}{Z_n\big(\beta + (\bar\beta/\sqrt{n}), h\big)} \sim \frac{Z_n\big(\beta, h + (t + \bar\beta m(\beta,h)/\sqrt{n})\big)}{Z_n\big(\beta, h + (\bar\beta m(\beta,h))/\sqrt{n}\big)} \tag{5.5}$$

for large n. The right-hand side of (5.5) [and hence of (5.3)] can be estimated as that of (2.30a). Then, (5.2) is obtained from the convergence of the Laplace transform [as in Corollary 2.1 (a)].

Combining (5.1), (5.2) and the strict monotonicity of $u_n(\beta)$, we see that the ML estimates $\hat\beta_n$ given by (2.3b) satisfy [compare with the proof of Theorem 1.1(a)]

$$\begin{aligned} P_{n,\beta,h}\{\sqrt{n}(\hat\beta_n - \beta) \le a\} &\sim P_{n,\beta,h}\{\sqrt{n}[\bar X_n^2 - m^2(\beta)] \le 2\sigma^2(\beta)m^2(\beta)a\} \\ &\sim \Pr\{N(0,(\sigma(\beta)m(\beta))^{-2}) \le a\}, \end{aligned}$$

which proves (1.15).

(b) From the proof of Proposition 2.1(a), we obtain (for $0 < \beta < 1$ and $h = 0$) as $n \to \infty$,

$$n\bar X_n^2 \to_D \sigma^2(\beta)\chi_1^2, \quad \text{under } P_{n,\beta,0}$$

and

$$u_n(\beta) = \frac{\sigma^2(\beta)}{2n} + o\left(\frac{1}{n}\right).$$

Thus for $0 < a < b < +\infty$, we have

$$\begin{aligned} \lim_{n\to+\infty} P_{n,\beta}\{\hat\beta_n \in (a,b)\} &= \lim_{n\to+\infty} P_{n,\beta}\left\{\frac{n}{2}\bar X_n^2 \in (nu_n(a), nu_n(b))\right\} \\ &= \Pr\{\sigma^2(\beta)\chi_1^2 \in (\sigma^2(a), \sigma^2(b))\}. \end{aligned}$$

This proves (1.16), since $\sigma^2(\beta) = (1-\beta)^{-1}$.

(c) For $\beta > 1$ and $h = 0$, Proposition 2.1(b) yields

$$\sqrt{n}(\bar X_n^2 - m_+^2(\beta)) \to_D N(0, 4\sigma^2(\beta)m_+^2(\beta)) \tag{5.6}$$

under $P_{n,\beta,0}$. We will now show that for $\bar\beta \in \Re$, we have [compare with (5.2)]

$$u_n\left(\beta + \frac{\bar\beta}{\sqrt{n}}\right) = \frac{1}{2}m_+^2(\beta) + \sigma^2(\beta)m_+^2(\beta)\frac{\bar\beta}{\sqrt{n}} + o(n^{-1/2}). \tag{5.7}$$

Then the proof of (1.17) proceeds as in part (a) of the present theorem. To prove (5.7), we consider

$$\begin{aligned} &E_{n,\beta+\bar\beta n^{-1/2},0}\{\exp\{t\sqrt{n}[\bar X_n - m_+(\beta)]\}|\bar X_n > 0\} \\ &\quad = \exp\{-t\sqrt{n}\,m_+(\beta)\}\frac{Z_n^+(\beta + (\bar\beta/\sqrt{n}),(t/\sqrt{n}))}{Z_n^+(\beta + (\bar\beta/\sqrt{n}),0)}, \end{aligned}$$

where Z_n^+ is defined in (2.31b). As in (5.4), $Z_n^+(\beta + (\bar\beta/\sqrt{n}),(t/\sqrt{n}))$ can be expressed in terms of $Z_n^+(\beta,(t+\bar{\beta}m_+)/\sqrt{n})$ and a conditional expectation which tends to 1 as $n \to +\infty$. This gives an estimate similar to (5.5). Then as for (5.2), we obtain

$$\frac{1}{2}E_{n,\beta+\bar{\beta}n^{-1/2},0}\{\bar{X}_n^2|\bar{X}_n \leq 0\} = \frac{1}{2}m_+^2(\beta) + \sigma^2(\beta)m_+^2(\beta)\frac{\bar{\beta}}{\sqrt{n}} + o(n^{-1/2})$$

and

$$\frac{1}{2}E_{n,\beta+\bar{\beta}n^{-1/2},0}\{\bar{X}_n^2|\bar{X}_n \leq 0\} = \frac{1}{2}m_-^2(\beta) + \sigma^2(\beta)m_-^2(\beta)\frac{\bar{\beta}}{\sqrt{n}} + o(n^{-1/2}).$$

Since these are the same, they yield (5.7).

(d) For $\beta = 1$ and $h = 0$, we will use Proposition 4.1. We will also show that

$$u_n\left(1+\frac{\bar{\beta}}{\sqrt{n}}\right) = \frac{1}{2}v(\bar{\beta})\frac{1}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right), \tag{5.8a}$$

where

$$v(\bar{\beta}) = \int \xi^2 \frac{\exp\left[-\frac{1}{12}\xi^4 + \frac{1}{2}\bar{\beta}\xi^2\right]}{\int d\xi \exp\left[-\frac{1}{12}\xi^4 + \frac{1}{2}\bar{\beta}\xi^2\right]}\, d\xi. \tag{5.8b}$$

Then for $a \in \mathfrak{R}$, we have

$$\begin{aligned} P_{n,1,0}\{\sqrt{n}(\hat{\beta}_n - 1) > a\} &= P_{n,1,0}\left\{\frac{1}{2}\bar{X}_n^2 > u_n\left(1+\frac{a}{\sqrt{n}}\right)\right\} \\ &\sim P_{n,1,0}\{\sqrt{n}\,\bar{X}_n^2 > v(a)\} \\ &\to_{n\to+\infty} 1 - F\left(\left[-\sqrt{v(a)}, \sqrt{v(a)}\right]\right), \end{aligned}$$

which proves (1.18). To prove (5.8), we consider

$$E_{n,1+\bar{\beta}n^{-1/2},0}\{\exp\{tn^{1/4}\bar{X}_n\}\} = \frac{Z_n(1+\bar{\beta}n^{-1/2}, tn^{-3/4})}{Z_n(1+\bar{\beta}n^{-1/2},0)}. \tag{5.9}$$

As in (4.3),

$$Z_n(1+\bar{\beta}n^{-1/2}, tn^{-3/4}) \sim \sqrt{\frac{n}{2\pi}}\int d\xi \exp\{-nf_{1+\bar{\beta}n^{-1/2},0}(\xi) + n^{1/4}t\xi\}. \tag{5.10}$$

Using

$$f_{1+\bar{\beta}n^{-1/2},0}(\xi) = \frac{2}{4!} - \frac{\bar{\beta}}{2\sqrt{n}}\xi^2 + O\left(\frac{\xi}{\sqrt{n}}\right) + O(\xi^6),$$

one finds that the right-hand side of (5.10) is equivalent to

$$(2\pi)^{-1/2}n^{1/4}\int d\xi \exp\left\{-\frac{1}{12}\xi^4 + \frac{\bar{\beta}}{2}\xi^2 + t\xi\right\}.$$

Then this yields that the left-hand side of (5.9) converges to

$$\frac{\int d\xi \exp\{-\frac{1}{12}\xi^4 + \frac{1}{2}\bar{\beta}\xi^2 + t\xi\}}{\int d\xi \exp\{-\frac{1}{12}\xi^4 + \frac{1}{2}\bar{\beta}\xi^2\}},$$

which implies (5.8).

REFERENCES

[1] BAHADUR, R. R. (1964). On Fisher's bound for asymptotic variances. *Ann. Math. Statist.* **35** 1545–1552.

[2] BESAG, J. (1986). On the statistical analysis of dirty pictures (with discussion). *J. Roy. Statist. Soc. Ser. B* **48** 259–302.

[3] BOLTHAUSEN, E. (1986, 1987). Laplace approximations for sums of independent random vectors. *Probab. Theory Related Fields* **72** 305–318; **76** 167–206.

[4] ELLIS, R. S. (1985). *Entropy, Large Deviations, and Statistical Mechanics.* Springer, New York.

[5] ELLIS, R. S., NEWMAN, C. M. and ROSEN, J. S. (1980). Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **51** 153–169.

[6] GEMAN, S. and GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. and Machine Intelligence* **6** 721–741.

[7] GEMAN, D., GEMAN, S., GRAFFIGNE, C. and DONG, P. (1990). Boundary detection by constraint optimization. *IEEE Trans. Pattern Anal. and Machine Intelligence* **12** 609–628.

[8] GEMAN, S. and GRAFFIGNE C. (1987). Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians* (A. M. Gleason, ed.) **2** 1496–1517. Amer. Math. Soc., Providence, R.I.

[9] GEMAN, S. and MCCLURE, D. E. (1985). Bayesian image analysis: An application to single photon emission tomography. In *Proc. Amer. Statist. Assoc. Statistical Computing Section* 12-18. Amer. Statist. Soc., Washington, D.C.

[10] GIDAS, B. (1987). Consistency of maximum-likelihood and pseudo-likelihood estimations for Gibbs distribution. In *Proc. Workshop Stochastic Differential Systems with Applications in Electrical/Computer Engineering, Control Theory and Operation Research.* IMS, Springer, New York.

[11] GIDAS, B. (1991). Parameter estimation for Gibbs distribution, I: Fully observed data. In *Markov Random Fields: Theory and Applications* (R. Chellappa and A. Jain, eds.). Academic, New York.

[12] GIDAS, B. (1989). A renormalization group approach to image processing problems. *IEEE Trans. Pattern Anal. and Machine Intelligence* **2** 164–180.

[13] GRENANDER, U. (1983). Tutorial in pattern theory. Div. Applied Math., Brown Univ.

[14] GUYON, X. (1985). Estimation d'un champ par pseudo-raisemblance conditionelle: etude asympotique et application an cas Markovien. Actes de la 6^{eme} Rencontre Franco-Belge de Staticiens, Bruxelles.

[15] KÜNSCH, H. (1982). Asymptotically unbiased inference for Ising models. *J. Appl. Probab.* **19A** 345–357.

[16] PICKARD, D. K. (1987). Inference for discrete Markov fields: The simplest nontrivial case. *J. Amer. Statist. Assoc.* **82** 90–96.

[17] POSSOLO, A. (1986). Estimation of binary Markov random fields. Unpublished manuscript.

[18] PRIVMAN, V. and FISHER, M. E. (1983). Finite-size effects at first-order transitions. *J. Statist. Phys.* **33** 385–417.

[19] RIPLEY, B. D. (1987). Statistical inference for spatial processes. Unpublished manuscript.

[20] YOUNÈS, L. (1988). Estimation and annealing for Gibbsian fields. *Ann. Inst. H. Poincaré Probab. Statist.* **24** 269–294.

[21] YOUNÈS, L. (1988). Problèmes d'estimation parametrique par des champs de Gibbs Markoviens. Application en traitenent d'image. Thesis, Université de Paris-Sud, Orsay.

LABORATOIRE DE STATISTIQUE APPLIQUÉ
MATHEMATIQUES BÂT. 425
UNIVERSITÉ PARIS-SUD
F-91405 ORSAY CÉDEX
FRANCE

DIVISION OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RHODE ISLAND 02912