The Annals of Statistics
2022, Vol. 50, No. 4, 2021–2046
https://doi.org/10.1214/22-AOS2169

ASYMPTOTIC ACCURACY OF THE SADDLEPOINT APPROXIMATION FOR MAXIMUM LIKELIHOOD ESTIMATION

BY JESSE GOODMAN[a]

Department of Statistics, University of Auckland, [a]*jesse.goodman@auckland.ac.nz*

The saddlepoint approximation gives an approximation to the density of a random variable in terms of its moment generating function. When the underlying random variable is itself the sum of n unobserved i.i.d. terms, the basic classical result is that the relative error in the density is of order $1/n$. If instead the approximation is interpreted as a likelihood and maximised as a function of model parameters, the result is an approximation to the maximum likelihood estimate (MLE) that can be much faster to compute than the true MLE. This paper proves the analogous basic result for the approximation error between the saddlepoint MLE and the true MLE: subject to certain explicit identifiability conditions, the error has asymptotic size $O(1/n^2)$ for some parameters and $O(1/n^{3/2})$ or $O(1/n)$ for others. In all three cases, the approximation errors are asymptotically negligible compared to the inferential uncertainty.

The proof is based on a factorisation of the saddlepoint likelihood into an exact and approximate term, along with an analysis of the approximation error in the gradient of the log-likelihood. This factorisation also gives insight into alternatives to the saddlepoint approximation, including a new and simpler saddlepoint approximation, for which we derive analogous error bounds. As a corollary of our results, we also obtain the asymptotic size of the MLE approximation error when the saddlepoint approximation is replaced by the normal approximation.

1. Introduction. Let X be a random variable with density function $f(x)$, $x \in \mathbb{R}$. Define

$$M(s) = \mathbb{E}\big(e^{sX}\big), \qquad K(s) = \log M(s), \tag{1.1}$$

the moment generating function (MGF) and cumulant generating function (CGF), respectively, associated to X. Given $x \in \mathbb{R}$, let $\hat{s}$ be the solution to

$$K'(\hat{s}) = x \tag{1.2}$$

and set

$$\hat{f}(x) = \frac{\exp(K(\hat{s}) - \hat{s}x)}{\sqrt{2\pi K''(\hat{s})}}. \tag{1.3}$$

We call $\hat{f}(x)$ the *saddlepoint approximation* to the density function $f(x)$.

In the statistical context, an important use of the saddlepoint approximation has been to analyse sampling distributions, with X an estimator or a related statistic to be understood via its sampling density $f(x)$. In this setting, the natural way to assess the saddlepoint approximation is to measure how well $\hat{f}(x)$ approximates $f(x)$ as a function of x, for instance, by determining how fast the absolute error $|\hat{f}(x) - f(x)|$ or the relative error $|\hat{f}(x)/f(x) - 1|$ decay in a suitable limit, and whether this convergence is uniform.

The most prominent results of this kind concern what we will call the *standard asymptotic regime*, in which the observed value X is the sample average of n i.i.d. values. In this setup,

Received April 2021; revised December 2021.
MSC2020 subject classifications. Primary 62F12; secondary 41A60, 62F15.
Key words and phrases. Saddlepoint approximation, maximum likelihood estimation, approximation error.

the saddlepoint approximation has simple n-dependence and can be computed in constant time, whereas the true density $f(x)$ commonly becomes intractable. The classical error estimate [5] for the standard asymptotic regime states that $\hat{f}(x)/f(x) = 1 + O(1/n)$ as $n \to \infty$. Under stronger assumptions, the ratio $\hat{f}(x)/f(x)$ may remain uniformly bounded even in the tails (see [12], Theorem 4.6.1, and [2, 10]), whereas other density approximations such as normal approximations (see Appendix I) or Edgeworth expansions quickly lose relative accuracy away from the mean. For many applications based on densities, these error bounds are more than enough to justify using the saddlepoint approximation $\hat{f}(x)$ as a readily computable substitute for $f(x)$.

In this paper, we shift perspectives and consider (1.3) as an approximation to the *likelihood*. In this viewpoint, X represents the raw data obtained from an experiment, modelled by the parameter θ, and the observation x is fixed. We write the density and CGF as $f(x;\theta)$, $K(s;\theta)$ to emphasise their dependence on the parameter, and set $L(\theta;x) = f(x;\theta)$. Instead of (1.3), we form

$$\hat{L}(\theta;x) = \frac{\exp(K(\hat{s};\theta) - \hat{s}x)}{\sqrt{2\pi K''(\hat{s};\theta)}}, \tag{1.4}$$

the *saddlepoint approximation to the likelihood*, considered as a function of θ. Note that the saddlepoint $\hat{s} = \hat{s}(\theta, x)$ is a function of both the parameter θ and the observed value x, defined implicitly by

$$K'\big(\hat{s}(\theta,x);\theta\big) = x, \tag{1.5}$$

and the derivatives K', K'' are with respect to s.

Maximum likelihood inference involves maximising $L(\theta;x)$ with respect to θ to produce the maximum likelihood estimate (MLE)

$$\theta_{\mathrm{MLE}}(x) = \operatorname*{argmax}_{\theta} L(\theta;x), \tag{1.6}$$

if it exists. In cases where the true likelihood $L(\theta;x)$ is intractable, but the saddlepoint likelihood $\hat{L}(\theta;x)$ can be computed, it is natural to ask what are the consequences for inference in maximising the saddlepoint likelihood in place of the true likelihood. We term the resulting estimate the *saddlepoint MLE*,

$$\hat{\theta}_{\mathrm{MLE}}(x) = \operatorname*{argmax}_{\theta} \hat{L}(\theta;x), \tag{1.7}$$

if it exists. Our interest is in the error introduced by this approach: specifically, the error in the saddlepoint MLE as an approximation to the true MLE, $|\hat{\theta}_{\mathrm{MLE}}(x) - \theta_{\mathrm{MLE}}(x)|$. Investigation of approximation error has been a central issue when saddlepoint MLEs have been applied in the literature [7, 16, 21], see Examples 21–23 in Appendix J, but no general theory is currently available. The classical saddlepoint error estimate $\hat{L}(\theta;x)/L(\theta;x) = 1 + O(1/n)$ does not provide easily interpretable guidance about whether $\hat{\theta}_{\mathrm{MLE}}(x)$ is close to $\theta_{\mathrm{MLE}}(x)$. Instead, the key question is whether the *gradient* $\nabla_\theta \log \hat{L}(\theta;x)$ provides a good approximation to the true gradient $\nabla_\theta \log L(\theta;x)$ [14, 15]. The main theorems of this paper will provide sharp asymptotic bounds for $|\nabla_\theta \log \hat{L}(\theta;x) - \nabla_\theta \log L(\theta;x)|$, and hence for the size of the MLE approximation error $|\hat{\theta}_{\mathrm{MLE}}(x) - \theta_{\mathrm{MLE}}(x)|$ under general conditions.

Outline of the paper. Section 2.1 introduces further notation for the multivariate saddlepoint approximation, including conventions for row and column vectors and gradients. In Section 2.2, we formulate the standard asymptotic regime as an explicit limiting framework relating the distribution X_θ, its CGF K, the observed value x, and the parameter n. Section 2.3

introduces a class of examples in this asymptotic framework. The main results, Theorems 1–8, are stated in Section 2.4, and their interpretation is discussed in Section 2.5. Several examples from theory and the literature are discussed in detail in Appendix J; see Section 2.6 for a brief summary.

Section 3.1 expresses the saddlepoint procedure as a combination of an exact step (tilting) and an approximation step, and introduces a factorisation and reparametrisation of the likelihood that underlie the rest of the paper. As a by-product, we obtain in Section 3.2 a "lower-order" version of the saddlepoint approximation, satisfying analogues of Theorems 1–5 with a different power of n; see Theorem 9.

The proofs of Theorems 1–2 are given in Section 4, along with a summary in Section 4.1 of gradients of quantities related to the saddlepoint approximation. Further proofs, examples and technical details appear in the supplement [9], Appendices A–J.

Section 5 includes a summary, additional discussion and directions for further inquiry.

2. Main results.

2.1. *Setup and notation.*

2.1.1. *Moment and cumulant generating functions.* We consider a vector-valued random variable X of dimension d depending on a parameter θ of dimension p, and write $X = X_\theta$ when we wish to emphasise the dependence. We consider the values of X_θ and θ to be column vectors, that is, $d \times 1$ or $p \times 1$ matrices, which we express as $X_\theta \in \mathbb{R}^{d\times 1}$, $\theta \in \mathcal{R} \subset \mathbb{R}^{p\times 1}$, where $\mathcal{R}$ is an open subset of $\mathbb{R}^{p\times 1}$. The multivariate MGF and CGF are

$$M(s;\theta) = \mathbb{E}\big(e^{sX_\theta}\big), \qquad K(s;\theta) = \log M(s;\theta). \tag{2.1}$$

On those occasions when we consider several random variables, we will write M_X, M_Y and so on to distinguish the respective generating functions.

In (2.1), s is called the *dual variable* to X, and we interpret it as a row vector, a $1 \times d$ matrix, so that sX is a scalar or 1×1 matrix. This convention emphasises that s and X, despite being vectors of the same dimension, play quite different roles and are not interchangeable; rather, the space of row vectors is the natural dual space to the space of column vectors. This convention also avoids excessive use of transposes and explicit inner products.

We wish to consider M and K for complex-valued s, and to this end we must take care of convergence issues in (2.1). Let

$$\mathcal{S}_\theta = \big\{s \in \mathbb{R}^{1\times d} : \mathbb{E}\big(e^{sX_\theta}\big) < \infty\big\}, \qquad \mathcal{S} = \big\{(s,\theta) : s \in \mathcal{S}_\theta\big\}. \tag{2.2}$$

Writing $\mathrm{Re}(z)$ for the real part of the complex number z, we have $|e^z| = e^{\mathrm{Re}(z)}$ for $z \in \mathbb{C}$, so the expectation in (2.1) converges absolutely whenever $\mathrm{Re}(s) \in \mathcal{S}_\theta$. We take the domain of M to be $\{(s,\theta) \in \mathbb{C}^{1\times d} \times \mathcal{R} : \mathrm{Re}(s) \in \mathcal{S}_\theta\}$, where $\mathrm{Re}(s)$ is interpreted coordinatewise for each of the d complex entries of s. Note that for certain distributions X_θ, $\mathcal{S}_\theta$ may reduce to the single point 0 or otherwise become degenerate, but our assumptions will rule this out. As soon as the interior $\operatorname{int}\mathcal{S}_\theta$ is nonempty, $M(s;\theta)$ is analytic as a function of $s \in \operatorname{int}\mathcal{S}_\theta$.

2.1.2. *Gradients, Hessians, moments and cumulants.* We interpret gradient operators ∇_s, ∇_θ to have the shape of the transposes of the variables. Thus $\nabla_\theta K$ is a $1 \times p$ row vector and $\nabla_s K$, which we write as K', is a $d \times 1$ column vector. We apply this convention also to vector-valued functions; in particular, $\nabla_s \nabla_\theta K$ is the $(d \times p)$-matrix-valued function with i, j entry $\frac{\partial^2 K}{\partial s_i \partial \theta_j}$. We can write the Hessians of a scalar-valued function $f(s,\theta)$ as $\nabla_s \nabla_s^T f$ and $\nabla_\theta^T \nabla_\theta f$, with i, j entries $\frac{\partial^2 f}{\partial s_i \partial s_j}$ and $\frac{\partial^2 f}{\partial \theta_i \partial \theta_j}$. When $f = K$, we write the s-Hessian as $K'' = \nabla_s \nabla_s^T K$. For other derivative conventions, see Appendix B.

When $0 \in \operatorname{int} \mathcal{S}_\theta$, the derivatives $M'(0;\theta)$, $M''(0;\theta)$ and $K'(0;\theta)$, $K''(0;\theta)$ give moments and cumulants of X:

$$\begin{aligned} M'(0;\theta) &= \mathbb{E}(X_\theta), \qquad & M''(0;\theta) &= \mathbb{E}\big(X_\theta X_\theta^T\big), \\ K'(0;\theta) &= \mathbb{E}(X_\theta), \qquad & K''(0;\theta) &= \operatorname{Cov}(X_\theta, X_\theta). \end{aligned} \tag{2.3}$$

In particular, $K''(0;\theta)$ must be positive semidefinite. More generally, as we shall see in Section 3.1, $K''(s;\theta)$ has an interpretation as a covariance matrix for all $s \in \operatorname{int} \mathcal{S}_\theta$. It is natural to exclude the case where this covariance matrix is singular, and indeed our hypotheses will imply that

$$K''(s;\theta) \text{ is positive definite for all } s \in \operatorname{int} \mathcal{S}_\theta. \tag{2.4}$$

As a consequence, K is strictly convex as a function of s.

2.1.3. *Multivariate saddlepoint approximation.* With these preparations, we can state the multivariate saddlepoint approximation. For $x \in \mathbb{R}^{d\times 1}$, we form the *saddlepoint equation*

$$K'(\hat{s};\theta) = x \tag{SE}$$

for $\hat{s} \in \mathcal{S}_\theta$. The strict convexity of K implies that if equation (SE) has a solution, then the solution is unique and we call it the *saddlepoint* $\hat{s} = \hat{s}(\theta, x)$. We write

$$\begin{aligned} &\mathcal{X}_\theta = \big\{x \in \mathbb{R}^{d\times 1} \colon \exists s \in \mathcal{S}_\theta \text{ solving } K'(s;\theta) = x\big\}, \qquad \mathcal{X} = \big\{(x,\theta) \colon x \in \mathcal{X}_\theta\big\}, \\ &\mathcal{X}^o = \big\{(x,\theta) \in \mathcal{X} \colon \big(\hat{s}(\theta,x),\theta\big) \in \operatorname{int} \mathcal{S}\big\}. \end{aligned} \tag{2.5}$$

We will not discuss under what conditions the saddlepoint equation (SE) has a solution; see, for instance, [11], Section 2.1, or [1], Corollary 9.6. We merely remark that in many common examples, we can solve (SE) for all x in the interior of the convex hull of the support of X, but that this may fail if, for instance, X is nonnegative with finite mean and infinite variance.

For $x \in \operatorname{int} \mathcal{X}_\theta$, we can define the *saddlepoint approximation to the likelihood*,

$$\hat{L}(\theta;x) = \frac{\exp(K(\hat{s}(\theta,x);\theta) - \hat{s}(\theta,x)x)}{\sqrt{\det(2\pi K''(\hat{s}(\theta,x);\theta))}}, \tag{SPA}$$

the multivariate analogue of (1.4). As in (1.7), the saddlepoint MLE $\hat{\theta}_{\mathrm{MLE}}(x)$ is the value of θ that maximises $\hat{L}(\theta;x)$, if one exists, for a given observed vector x.

We will compare $\hat{L}(\theta;x)$ and $\hat{\theta}_{\mathrm{MLE}}(x)$ with the true likelihood $L(\theta;x)$ and true MLE $\theta_{\mathrm{MLE}}(x)$. We are assuming that X has an absolutely continuous distribution, so that the true likelihood $L(\theta;x)$ should be taken to coincide with the density function for X. Complications can arise if there is ambiguity in the choice of density function, for instance, if the density function has jumps, and later we will impose decay bounds on M that will imply that X has a continuous and, therefore, essentially unique, density function. Note, however, that the saddlepoint approximation can be applied whether or not X has a density function, and indeed even when X has a discrete distribution; see Theorem 8.

Throughout the paper, we will use the symbol $\hat{}$ to indicate saddlepoint approximations, rather than estimators based on observations. Thus θ_{MLE} and $\hat{\theta}_{\mathrm{MLE}}$ denote two deterministic functions of the formal argument x, whose nature depends on our chosen parametric model. Although we will continue to describe x as the observed value of X, we will think of x as the arbitrary input value to the functions θ_{MLE} and $\hat{\theta}_{\mathrm{MLE}}$, rather than as the result of a random experiment or sampling procedure. When we turn to sampling distributions in Theorems 4–5, we will introduce further notation to encode any randomness in the observed value.

2.2. *The standard asymptotic regime.* Often an approximation is given theoretical justification by proving that the approximation error becomes negligible in some relevant limit. For the saddlepoint approximation, the most commonly treated and mathematically tractable limiting framework is to assume that X is the sum of n unobserved i.i.d. terms,

$$X = \sum_{i=1}^{n} Y^{(i)}, \tag{2.6}$$

where $Y^{(1)}, Y^{(2)}, \ldots$ are i.i.d. copies of a random variable Y_θ whose parametric distribution does not depend on n. We also scale the observed value x in a matching way. Thus we take

$$M(s;\theta) = M_0(s;\theta)^n, \qquad K(s;\theta) = nK_0(s;\theta), \qquad x = ny, \quad n \to \infty, \tag{SAR}$$

where M_0 and K_0, the MGF and CGF corresponding to Y_θ, are fixed. Throughout the paper, we will assume the relation $x = ny$ implicitly. It can be helpful to interpret the value $y = x/n$ as the sample mean implied by an observed value x. We think of y as fixed (or varying within a small neighbourhood) in the limit $n \to \infty$, so that both x and X will be of order n, and constraints on the observed value x will usually be expressed as constraints on y. We refer to this limiting framework as the *standard asymptotic regime* for the saddlepoint approximation.

In the standard asymptotic regime, the saddlepoint equation (SE) simplifies to

$$K_0'(\hat{s};\theta) = y. \tag{$\mathrm{SE}_{\mathrm{SAR}}$}$$

Write $\hat{s}_0(\theta, y)$ for the function that maps $y \in \mathcal{Y}_\theta$ to the solution of ($\mathrm{SE}_{\mathrm{SAR}}$), where $\mathcal{Y}_\theta$, $\mathcal{Y}$, $\mathcal{Y}^o$ are defined as in (2.5) with x, K replaced by y, K_0. Note that when the relations (SAR) hold, the saddlepoint $\hat{s}(\theta, x)$ does not depend on n, with

$$\hat{s}(\theta, x) = \hat{s}_0(\theta, y). \tag{2.7}$$

We sometimes write $\hat{s}$ for the common value in (2.7) when the distinction is immaterial. The domains of $\hat{s}$ and $\hat{s}_0$ are related by $\mathcal{X}_\theta = \{ny\colon\ y \in \mathcal{Y}_\theta\}$, and similarly for $\mathcal{X}$, $\mathcal{X}^o$, $\mathcal{Y}$, $\mathcal{Y}^o$.

In the standard asymptotic regime, the saddlepoint approximation (SPA) becomes

$$\hat{L}(\theta; x) = \frac{\exp(n[K_0(\hat{s}_0(\theta, y); \theta) - \hat{s}_0(\theta, y)y])}{\sqrt{\det(2\pi n K_0''(\hat{s}(\theta, y); \theta))}}. \tag{$\mathrm{SPA}_{\mathrm{SAR}}$}$$

The basic error estimate for the saddlepoint approximation states that, in the standard asymptotic regime and subject to certain technical assumptions, the relative error in the likelihood is of order $1/n$:

$$\frac{\hat{L}(\theta; x)}{L(\theta; x)} = 1 + O(1/n) \quad \text{as } n \to \infty, \text{ for fixed } (y, \theta) \in \mathcal{Y}^o. \tag{2.8}$$

Recall the convention that x and y are implicitly related as in (SAR), and note that the n-dependence of L and $\hat{L}$ is omitted from the notation. See (4.11) for a more precise statement. See also Appendix I, where we compare (2.8) to its analogue for normal approximations.

REMARK. The standard asymptotic regime supposes that the observed value X is the sum of n unobserved i.i.d. terms. That is, the likelihoods $L(\theta; x)$, $\hat{L}(\theta; x)$ and MLEs $\theta_{\mathrm{MLE}}(x)$, $\hat{\theta}_{\mathrm{MLE}}(x)$ pertain to a single observation, $X = x$, rather than n observations of the summands $Y^{(1)}, \ldots, Y^{(n)}$. For this reason, the parameter n *should not be interpreted as a sample size* in the usual sense.

A model with k i.i.d. observations can be adapted to the framework of (SAR) by giving each observation its own entry in the vector X; this setup is discussed further in Section 2.5.5 and Appendix J, Example 20. Note, however, that the results of this paper apply for $n \to \infty$, with the sample size k fixed; the joint limit $n \to \infty$, $k \to \infty$ is excluded from consideration.

2.3. *A class of examples.* One application of saddlepoint MLEs has been to analyse certain *latent identity* models [21]. Each individual in a population of size N is assigned one of ℓ latent identities, with U_j the number of individuals having latent identity j. The counts $U_1, \dots, U_\ell$ are not observed directly, and instead we observe d partial totals $X_1, \dots, X_d$, $d < \ell$. The latent identities are constructed sufficiently richly to make the relationship between latent identities and partial totals deterministic, and we can define a $d \times \ell$ matrix A by $A_{ij} = 1$ if individuals with latent identity j contribute to the total X_i, and $A_{ij} = 0$ otherwise. Thus the vectors of counts are related by $X = AU$, and the MGF for X can be computed simply by $M_X(s;\theta) = M_U(sA;\theta)$. However, the likelihood is difficult to compute because the value of U cannot be recovered from observing the value of X.

In many cases, there is no natural model for the observed count vector X itself. However, we can naturally model the latent vector U by assuming that $N \sim \text{Poisson}(\lambda)$ and that, given N, each individual's latent identity is chosen independently. The main parameter of interest is λ, the population size intensity, and the probabilities of different latent identities are determined by other model-specific parameters.

To match this model to the setup of (SAR), we make the change of variables $\lambda = n\tilde{\lambda}$. Note that $N \sim \text{Poisson}(\lambda)$ has the same distribution as the sum of n i.i.d. Poisson$(\tilde{\lambda})$ random variables, so that U and hence X can be written as a sum of n i.i.d. terms. The scaling parameter n is chosen so that $y = x/n$ is of order 1 (if the observed data vector x is given, in the limit of large observed counts) or so that $\tilde{\lambda} = \lambda/n$ is of order 1 (if we consider sampling distributions, in the limit of large population size intensities). In either case, x, n and λ will be all of the same asymptotic order. For further details, see Examples 19 and 21 in Appendix J.

2.4. *Main results.* Theorem 1 states a general asymptotic error bound for the gradients of the true and saddlepoint log-likelihoods. The error bound for the MLE depends on the structure of the model, and we distinguish two cases, which we call fully identifiable (Theorems 2–5) and partially identifiable (Theorem 6). Theorem 7 states the corresponding error bounds for normal approximations. Almost the same results apply to integer-valued random variables, and Theorem 8 shows how the assumptions should be modified for this setting.

2.4.1. *Approximation error in the log-gradient.* Because we wish to understand the true and approximate likelihoods as functions of θ, our central objects of study will be $\nabla_\theta \log L$ and $\nabla_\theta \log \hat{L}$ rather than L and $\hat{L}$. We therefore prove a general bound analogous to (2.8).

As we will see in Section 3.1, L and its derivatives can be expressed as integrals involving $M_0(s + \mathrm{i}\varphi;\theta)$ and its derivatives, where s is fixed and φ is integrated over $\mathbb{R}^{1\times d}$. To ensure that these integrals converge, we make the following technical assumptions on the growth or decay of M_0 and its derivatives:

there is a continuous function $\delta\colon \operatorname{int}\mathcal{S} \to (0,\infty)$ such that

$$\left|\frac{M_0(s+\mathrm{i}\varphi;\theta)}{M_0(s;\theta)}\right| \le \bigl(1+\delta(s,\theta)|\varphi|^2\bigr)^{-\delta(s,\theta)} \quad \text{for all } \varphi \in \mathbb{R}^{1\times d},\ (s,\theta) \in \operatorname{int}\mathcal{S}, \tag{2.9}$$

there is a continuous function $\gamma\colon \operatorname{int}\mathcal{S} \to (0,\infty)$ such that

$$\left|\frac{\partial^{k+\ell} M_0}{\partial\theta_{i_1}\cdots\partial\theta_{i_k}\partial s_{j_1}\cdots\partial s_{j_\ell}}(s+\mathrm{i}\varphi;\theta)\right| \le \gamma(s,\theta)\bigl(1+|\varphi|\bigr)^{\gamma(s,\theta)} \tag{2.10}$$

for all $\varphi \in \mathbb{R}^{1\times d}$ and $(s,\theta) \in \operatorname{int}\mathcal{S}$,

for $k \in \{0,1\}$, $1 \le k+\ell \le 6$ and for $k = 2$, $0 \le \ell \le 2$,

and each of these partial derivatives is continuous in all its variables.

These assumptions are relatively mild: (2.9) asserts that $\varphi \mapsto |M_0(s+\mathrm{i}\varphi;\theta)/M_0(s;\theta)|$, which always attains its maximum at $\varphi = 0$, has a nondegenerate critical point at $\varphi = 0$ (and in particular, (2.9) implies (2.4)) and decays at least polynomially as $|\varphi| \to \infty$. Similarly, (2.10) asserts that the partial derivatives grow at most polynomially in φ. Note that the expression in (2.9) means $\frac{\partial^\ell M_0}{\partial s_1 \cdots \partial s_\ell}$ when $k = 0$, and similarly when $\ell = 0$.

THEOREM 1 (Gradient error bound). *If* (SAR) *and* (2.9)–(2.10) *hold, then*

$$\nabla_\theta \log \hat{L}(\theta; x) = \nabla_\theta \log L(\theta; x) + O(1/n) \quad \text{as } n \to \infty \tag{2.11}$$

for $(y, \theta) \in \operatorname{int} \mathcal{Y}^o$ *fixed. Moreover, given a compact subset* $C \subset \operatorname{int} \mathcal{Y}^o$, *there exists* $n_0 \in \mathbb{N}$ *such that the bound in the term* $O(1/n)$ *is uniform over* $n \geq n_0$ *and* $(y, \theta) \in C$.

Note that Theorem 1 includes the implicit assertion that the likelihoods and their gradients exist for n sufficiently large. However, the likelihoods may be ill-behaved for small n: see Example 32 in Appendix J. Note also our convention from (SAR) that $x = ny$.

In the rest of our results, we apply the gradient error bound from Theorem 1 (or rather, its more precise analogues: see Corollary 12 and equation (4.10)) in the neighbourhood of a local maximiser.

2.4.2. *MLE error, posterior and sampling distributions in the fully identifiable case.* In the limit $n \to \infty$, the asymptotic behaviour of the MLE, and of other quantities derived from the likelihood, depends on the asymptotic shape of the likelihood function near its maximum. The simplest case occurs when this maximum occurs due a nondegenerate maximum for the leading-order exponential factor in ($\mathrm{SPA}_{\mathrm{SAR}}$). Concretely, if we set $x = ny_0$, $y = y_0$ and take the limit $n \to \infty$ with y_0 fixed, then the leading-order behaviour of the log-likelihood comes from the function

$$\theta \mapsto K_0\big(\hat{s}_0(\theta, y_0); \theta\big) - \hat{s}_0(\theta, y_0) y_0. \tag{2.12}$$

The results in this section apply when (2.12) has a nondegenerate local maximum at θ_0.

THEOREM 2 (MLE error bound—fully identifiable case). *Let* s_0, θ_0, y_0 *be related by*

$$\begin{aligned} &y_0 = K_0'(s_0; \theta_0) \quad \text{with } (s_0, \theta_0) \in \operatorname{int} \mathcal{S} \quad \text{or equivalently} \\ &s_0 = \hat{s}_0(\theta_0, y_0) \quad \text{with } (y_0, \theta_0) \in \mathcal{Y}^o, \end{aligned} \tag{2.13}$$

and suppose that (SAR) *and* (2.9)–(2.10) *hold. Suppose also that*

$$\nabla_\theta K_0(s_0; \theta_0) = 0 \quad \text{and that} \tag{2.14}$$

$$\begin{aligned} H = {} & \nabla_\theta^T \nabla_\theta K_0(s_0; \theta_0) \\ & - \big(\nabla_s \nabla_\theta K_0(s_0; \theta_0)\big)^T K_0''(s_0; \theta_0)^{-1} \big(\nabla_s \nabla_\theta K_0(s_0; \theta_0)\big) \\ & \text{is negative definite.} \end{aligned} \tag{2.15}$$

Then there exist $n_0 \in \mathbb{N}$ *and neighbourhoods* $U \subset \mathcal{R}$ *of* θ_0 *and* $V \subset \mathbb{R}^{d \times 1}$ *of* y_0 *such that, for all* $n \geq n_0$ *and* $y \in V$, *the functions* $\theta \mapsto \hat{L}(\theta; x)$ *and* $\theta \mapsto L(\theta; x)$ *have unique local maximisers in* U. *Moreover, writing these local maximisers as* $\hat{\theta}_{\text{MLE in } U}(x)$ *and* $\theta_{\text{MLE in } U}(x)$,

$$\big|\hat{\theta}_{\text{MLE in } U}(x) - \theta_{\text{MLE in } U}(x)\big| = O\big(1/n^2\big) \quad \text{as } n \to \infty, \tag{2.16}$$

uniformly over $n \geq n_0$, $y \in V$.

The assumptions of Theorem 2 can be understood as follows. We shall show that, when (2.13) holds, the expressions in (2.14)–(2.15) are the gradient and Hessian, respectively, of (2.12); see the differentiation formulas in (4.1)–(4.3) and [9], (B.13)–(B.14). Thus (2.14)–(2.15) state that θ_0 should be a nondegenerate local maximiser of (2.12).

More directly, we can interpret θ_0 as the limiting local maximiser for the likelihood of an observation $x = ny_0$, in the limit $n \to \infty$ with y_0 fixed. In such a limit, an observed value $x = ny_0$ for X amounts to an observed value y_0 for the true mean of the summands $Y^{(1)}, \dots, Y^{(n)}$ from (2.6). Thus (2.13)–(2.15) state that we should be able to recover the asymptotic (local) MLE θ_0 based solely on the implied sample mean y_0. We might describe (2.13)–(2.15) as saying that the model is *fully identifiable at the level of the sample mean*.

In particular, Theorem 2 applies in the well-specified case where the observed value is itself drawn according to the model with true parameter θ_0, where $(0, \theta_0) \in \operatorname{int} \mathcal{S}$. Then, setting $s_0 = 0$ and $y_0 = K_0'(0; \theta_0) = \mathbb{E}(Y_{\theta_0})$, the conditions (2.13)–(2.14) hold identically, while (2.15) reduces to the condition that the $d \times p$ matrix $\nabla_s \nabla_\theta K_0(0; \theta_0)$ should have rank p. Because $X_{\theta_0}/n \to y_0$ and $\mathbb{P}(X_{\theta_0}/n \in V) \to 1$ as $n \to \infty$, the conclusions of Theorem 2 will hold with high probability with x replaced by X_{θ_0}.

To place the MLE approximation error $O(1/n^2)$ in context, we can compare it to the inferential uncertainty inherent in the model. The next three theorems give the asymptotics, either Bayesian or frequentist, that result from taking $n \to \infty$.

In a Bayesian framework, let the parameter Θ be drawn according to a prior π_Θ on $\mathcal{R}$. For a given observed value x, we will consider the posterior distribution $\pi_{\Theta|U,x}$ on a neighbourhood $U \subset \mathcal{R}$ with $\pi_\Theta(U) > 0$, defined by the Radon–Nikodym derivative

$$\frac{d\pi_{\Theta|U,x}}{d\pi_\Theta}(\theta) = \frac{L(\theta; x)\mathbb{1}_{\{\theta \in U\}}}{C}, \quad C = C_{U,x} = \int_U L(\theta; x)\, d\pi_\Theta(\theta). \tag{2.17}$$

We construct the *saddlepoint posterior distribution* on U, $\hat{\pi}_{\Theta|U,x}$, by replacing L with $\hat{L}$:

$$\frac{d\hat{\pi}_{\Theta|U,x}}{d\pi_\Theta}(\theta) = \frac{\hat{L}(\theta; x)\mathbb{1}_{\{\theta \in U\}}}{\hat{C}}, \quad \hat{C} = \int_U \hat{L}(\theta; x)\, d\pi_\Theta(\theta). \tag{2.18}$$

THEOREM 3 (Posterior distributions). *Let (s_0, θ_0, y_0) be related as in (2.13)–(2.14), and suppose that (SAR), (2.9)–(2.10) and (2.15) hold. Suppose also that the prior distribution π_Θ has a probability density function that is continuous and positive at θ_0. Fix $y = y_0$, $x = ny_0$. Then there exists a neighbourhood $U \subset \mathcal{R}$ of θ_0 such that*

$$\text{under } \pi_{\Theta|U,x} \text{ or } \hat{\pi}_{\Theta|U,x}, \quad \sqrt{n}(\Theta - \theta_0) \xrightarrow{d} \mathcal{N}\big(0, -H^{-1}\big) \quad \text{as } n \to \infty, \tag{2.19}$$

where H is the negative definite matrix from (2.15).

In particular, Theorem 3 shows that both the true and saddlepoint likelihoods lead to the same asymptotic posterior. The proof will follow from a stronger statement, Proposition 18 in Appendix F, that removes the assumption $y = y_0$.

Theorems 2 and 3 concern the deterministic functions that map an observed value x to the corresponding MLE or posterior distribution, via either the true likelihood or the saddlepoint approximation. In this description, the observed value x has been treated as deterministic, separate from any consideration of the random process that might have generated this observation. The next theorem describes the sampling distribution when the observation is itself a random variable χ_n.

THEOREM 4 (Sampling distributions). *Let (s_0, θ_0, y_0) be related as in (2.13)–(2.14), and suppose that (SAR), (2.9)–(2.10) and (2.15) hold. Let U be the neighbourhood of θ_0*

given by Theorem 2. Suppose also that $\chi_n \in \mathbb{R}^{d\times 1}$ *are random variables satisfying*

$$\frac{\chi_n - n y_0}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \Sigma) \quad \text{as } n \to \infty, \tag{2.20}$$

where $\Sigma \in \mathbb{R}^{d\times d}$ *is a positive semidefinite matrix. Then*:

(a) *The joint sampling distribution of the true and saddlepoint MLEs satisfies*

$$\big(\sqrt{n}\big(\theta_{\text{MLE in } U}(\chi_n) - \theta_0\big), \sqrt{n}\big(\hat{\theta}_{\text{MLE in } U}(\chi_n) - \theta_0\big)\big) \xrightarrow{d} (Z, Z) \quad \text{as } n \to \infty \tag{2.21}$$

with

$$Z \sim \mathcal{N}\big(0, H^{-1} B^T A^{-1} \Sigma A^{-1} B H^{-1}\big), \tag{2.22}$$

where we have abbreviated $A = K_0''(s_0; \theta_0)$, $B = \nabla_s \nabla_\theta K_0(s_0; \theta_0)$, *and* H *is the negative definite matrix from* (2.15).

(b) *If in addition* $s_0 = 0$, $y_0 = K_0'(0; \theta_0)$ *and* $\Sigma = K_0''(0; \theta_0)$, *then the limiting distribution has*

$$Z \sim \mathcal{N}\big(0, -H^{-1}\big). \tag{2.23}$$

Theorems 2 and 4(b) apply in particular in the well-specified case where the observed data are drawn according to the model distribution X_{θ_0}, with $s_0 = 0$ and $y_0 = \mathbb{E}(Y_{\theta_0})$.

THEOREM 5 (Sampling distribution in the well-specified case). *Let* $\theta_0 \in \mathcal{R}$ *be such that* $(0, \theta_0) \in \operatorname{int} \mathcal{S}$, *and set* $s_0 = 0$. *Suppose that* (SAR) *and* (2.9)–(2.10) *hold and that*

$$B = \nabla_s \nabla_\theta K_0(0; \theta_0) \text{ has rank } p. \tag{2.24}$$

Then:

(a) *The matrix* H *from* (2.15) *reduces to* $H = -B^T K_0''(0; \theta_0)^{-1} B$ *and is negative definite.*

(b) *With observed data* X_{θ_0}, *both* $\theta_{\text{MLE in } U}(X_{\theta_0})$ *and* $\hat{\theta}_{\text{MLE in } U}(X_{\theta_0})$ *are consistent and asymptotically normal estimators of* θ_0 *in the limit* $n \to \infty$, *with*

$$\big(\sqrt{n}\big(\theta_{\text{MLE in } U}(X_{\theta_0}) - \theta_0\big), \sqrt{n}\big(\hat{\theta}_{\text{MLE in } U}(X_{\theta_0}) - \theta_0\big)\big) \xrightarrow{d} (Z, Z) \quad \text{as } n \to \infty, \tag{2.25}$$

where $Z \sim \mathcal{N}(0, -H^{-1})$. *Moreover*,

$$\big|\hat{\theta}_{\text{MLE in } U}(X_{\theta_0}) - \theta_{\text{MLE in } U}(X_{\theta_0})\big| = O_{\mathbb{P}}(1/n^2) \quad \text{as } n \to \infty. \tag{2.26}$$

(c) *With* $\Theta = \theta_{\text{MLE in } U}(X_{\theta_0})$ *or* $\Theta = \hat{\theta}_{\text{MLE in } U}(X_{\theta_0})$, *the Hessians* $\frac{1}{n}\nabla_\theta^T \nabla_\theta \log L(\Theta; X_{\theta_0})$, $\frac{1}{n}\nabla_\theta^T \nabla_\theta \log \hat{L}(\Theta; X_{\theta_0})$ *and* $-(\nabla_s \nabla_\theta K_0(0; \Theta))^T K_0''(0; \Theta)^{-1} (\nabla_s \nabla_\theta K_0(0; \Theta))$ *are consistent estimators of* H.

A key conclusion from these results is that the approximation error in using the saddlepoint MLE in place of the true MLE is negligible, in the limit $n \to \infty$ as in (SAR), compared to the underlying inferential uncertainty. Namely, according to Theorem 2, the difference between the true and saddlepoint MLEs is of order $1/n^2$. Asymptotically, this approximation error is much smaller than the spatial scale $1/\sqrt{n}$ corresponding to either sampling variability of the MLE (in the frequentist setup of Theorems 4–5) or posterior uncertainty of the parameter (in the Bayesian setup of Theorem 3). To the extent that the assumptions of (SAR) and Theorems 1–5 apply in a given application, the saddlepoint likelihood and saddlepoint MLE may therefore be appropriate as readily calculated substitutes for the true likelihood and MLE.

2.4.3. *MLE error in the partially identifiable case.* Theorems 2–5 apply to models that are fully identifiable at the level of the sample mean. However, many reasonable models lack this property, notably when some parameters affect the variance only. Then lower-order contributions to ($\mathrm{SPA}_{\mathrm{SAR}}$) become relevant, and the scaling of the MLE approximation error changes. The following theorem is the analogue of Theorem 2 in this case, for the well-specified setting (see Section 2.5.4) where $s_0 = 0$.

THEOREM 6 (MLE error bound—partially identifiable case). *Suppose we can split the parameter vector as*

$$\theta = \begin{pmatrix} \omega \\ \nu \end{pmatrix} \quad \text{such that } K_0'(0;\theta) = \mathbb{E}(Y_\theta) \text{ depends only on } \omega, \tag{2.27}$$

where $\omega \in \mathbb{R}^{p_1 \times 1}$, $\nu \in \mathbb{R}^{p_2 \times 1}$, $p_1 + p_2 = p$. *Let* $\theta_0 = \binom{\omega_0}{\nu_0}$ *be such that* $(0, \theta_0) \in \operatorname{int} \mathcal{S}$, *and suppose that* (SAR) *and* (2.9)–(2.10) *hold. Suppose further that the* $d \times p_1$ *matrix* $B_\omega = \nabla_s \nabla_\omega K_0(0;\theta_0)$ *has rank* p_1.

Introduce the "partially linearised" model in which $\Xi_{w,\nu} \in \mathbb{R}^{d\times 1}$ *is normally distributed with mean vector* $B_\omega w$ *and covariance matrix* $K_0''(0; \binom{\omega_0}{\nu})$, *where the parameters are* $w \in \mathbb{R}^{p_1\times 1}$ *and* $\nu \in \mathbb{R}^{p_2\times 1}$. *Consider* $\xi_0 \in \mathbb{R}^{d\times 1}$, *and suppose that* $(w, \nu) = (w_0, \nu_0)$ *is a nondegenerate local MLE for the observation* $\Xi_{w,\nu} = \xi_0$. *Make the change of variables*

$$x = nK_0'\left(0; \begin{pmatrix} \omega' \\ \nu_0 \end{pmatrix}\right) + \sqrt{n}\xi, \qquad y = x/n = K_0'\left(0; \begin{pmatrix} \omega' \\ \nu_0 \end{pmatrix}\right) + \xi/\sqrt{n}, \tag{2.28}$$

where $\omega' \in \mathbb{R}^{p_1\times 1}$, $\xi \in \mathbb{R}^{d\times 1}$. *Then*:

(a) *There exist* $n_0 \in \mathbb{N}$ *and neighbourhoods* $U, U' \subset \mathcal{R}$ *of* θ_0 *and* $\tilde{V} \subset \mathbb{R}^{d\times 1}$ *of* ξ_0 *such that, whenever* $n \ge n_0$, $\binom{\omega'}{\nu_0} \in U'$, $\xi \in \tilde{V}$ *and* (2.28) *holds, the functions* $\theta \mapsto \hat{L}(\theta; x)$ *and* $\theta \mapsto L(\theta; x)$ *have unique local maximisers in* U. *Moreover, writing these local maximisers as* $\hat{\theta}_{\mathrm{MLE\ in\ } U}(x) = \binom{\hat{\omega}_{\mathrm{MLE\ in\ } U}(x)}{\hat{\nu}_{\mathrm{MLE\ in\ } U}(x)}$ *and* $\theta_{\mathrm{MLE\ in\ } U}(x) = \binom{\omega_{\mathrm{MLE\ in\ } U}(x)}{\nu_{\mathrm{MLE\ in\ } U}(x)}$,

$$\begin{aligned} \bigl|\hat{\omega}_{\mathrm{MLE\ in\ } U}(x) - \omega_{\mathrm{MLE\ in\ } U}(x)\bigr| &= O\bigl(1/n^{3/2}\bigr), \\ \bigl|\hat{\nu}_{\mathrm{MLE\ in\ } U}(x) - \nu_{\mathrm{MLE\ in\ } U}(x)\bigr| &= O(1/n) \end{aligned} \quad \text{as } n \to \infty, \tag{2.29}$$

uniformly over $n \ge n_0$, $\xi \in \tilde{V}$.

(b) *Suppose in addition that*

$$\tilde{B}_\omega^T \tilde{A}^{-1} \frac{\partial K_0''}{\partial \nu_j}(0;\theta)\tilde{J} = 0 \quad \text{for all } \theta \text{ and for } j = 1, \dots, p_2, \tag{2.30}$$

where we have abbreviated $\tilde{A} = K_0''(0;\theta)^{-1}$, $\tilde{B}_\omega^T = \nabla_s \nabla_\omega K_0''(0;\theta)$ *and* $\tilde{J} = \tilde{A}^{-1} - \tilde{A}^{-1}\tilde{B}_\omega(\tilde{B}_\omega^T \tilde{A}^{-1}\tilde{B}_\omega)^{-1}\tilde{B}_\omega^T\tilde{A}^{-1}$. *Suppose also that* (2.10) *holds for* $k \le 2$, $1 \le k + \ell \le 7$ *and for* $k = 3$, $\ell \le 4$. *Then*

$$\bigl|\hat{\omega}_{\mathrm{MLE\ in\ } U}(x) - \omega_{\mathrm{MLE\ in\ } U}(x)\bigr| = O\bigl(1/n^2\bigr) \quad \text{as } n \to \infty, \tag{2.31}$$

uniformly over $n \ge n_0$, $\xi \in \tilde{V}$.

In Theorem 6, we are assuming, roughly, that ω is identifiable at the level of the sample mean but ν is not. Thus with $\nu = \nu_0$ fixed, the function $(s, \omega) \mapsto K_0\bigl(s; \binom{\omega}{\nu_0}\bigr)$ satisfies the assumptions of Theorems 2–5 with $s_0 = 0$, $y_0 = K_0'(0;\theta_0)$, whereas for the function $(s, \nu) \mapsto K_0\bigl(s; \binom{\omega_0}{\nu}\bigr)$ the analogue of the matrix H from (2.15) vanishes identically. As a result, the inferential uncertainty for ν need not decrease as $n \to \infty$; see Example 28 in Appendix J

for an instance of this. However, Theorem 6 shows that the MLE approximation error is still negligible in the limit $n \to \infty$.

The change of variables (2.28) is the same one used in Edgeworth expansions, with a separation of scales between the mean, parametrised by ω', and lower-order fluctuations, parametrised by ξ. The partially linearised model $\Xi_{w,\nu}$ arises from X_θ under this rescaling in the limit $n \to \infty$, and in [9], (H.4)–(H.5), we give explicit conditions equivalent to the MLE assumption from Theorem 6.

2.4.4. *Normal approximations.* Given a complicated parametric model X_θ, a different and more elementary approach is to replace X_θ by the normal random vector $\tilde{X}_\theta$ having the same mean and variance as functions of θ. Note that the combination of (SAR) and (2.28) is the standard asymptotic setting in which to apply the (local) central limit theorem, suggesting that this approximation is reasonable. We might therefore expect the MLE for observing $\tilde{X}_\theta = x$ to be a reasonable approximation to the MLE for observing $X_\theta = x$. The following theorem states the asymptotic size of the resulting MLE approximation error.

THEOREM 7 (MLE error bound—normal approximations). *Let $\tilde{L}(\theta; x)$ denote the likelihood function for the normal approximation model $\tilde{X}_\theta \sim \mathcal{N}(K'(0;\theta), K''(0;\theta))$. Then, under the hypotheses of Theorem 6*(a)*, there exist $n_0 \in \mathbb{N}$ and neighbourhoods $U, U' \subset \mathcal{R}$ of θ_0 and $\tilde{V} \subset \mathbb{R}^{d\times 1}$ of ξ_0 such that, whenever $n \geq n_0$, $\binom{\omega'}{\nu_0} \in U'$, $\xi \in \tilde{V}$ and (2.28) holds, the function $\theta \mapsto \tilde{L}(\theta; x)$ has a unique local maximiser in U. Moreover, writing this local maximiser as $\tilde{\theta}_{\text{MLE in } U}(x) = \binom{\tilde{\omega}_{\text{MLE in } U}(x)}{\tilde{\nu}_{\text{MLE in } U}(x)}$,*

$$\begin{aligned} &\left|\tilde{\omega}_{\text{MLE in } U}(x) - \omega_{\text{MLE in } U}(x)\right| = O(1/n), \\ &\left|\tilde{\nu}_{\text{MLE in } U}(x) - \nu_{\text{MLE in } U}(x)\right| = O(1/\sqrt{n}) \end{aligned} \qquad \text{as } n \to \infty, \tag{2.32}$$

uniformly over $n \geq n_0$, $\xi \in \tilde{V}$.

In both Theorems 6 and 7, the change of variables (2.28) requires the implied sample mean y to lie in a narrow "tube" of diameter of order $1/\sqrt{n}$ around the p_1-dimensional surface of possible model means, parametrised by $\omega' \mapsto K_0'(0; \binom{\omega'}{\nu_0})$. In Theorem 6, this restriction does not seem entirely essential, and we conjecture that a similar result holds for the case $s_0 \neq 0$. Theorem 7, by contrast, can have no such analogue, even in the fully identifiable case $p_1 = p$, $p_2 = 0$. Indeed, away from their shared surface of possible model means, the models $\theta \mapsto X_\theta$ and $\theta \mapsto \tilde{X}_\theta$ may be completely different, and there is no reason to expect any relationship between their MLEs; see, for instance, Example 24 in Appendix J.

2.4.5. *The integer-valued case.* Finally, all of these results apply to integer-valued random variables. As we shall discuss in Section 2.5.2, however, it would be natural to make different and more flexible assumptions in the integer-valued case.

THEOREM 8. *Let X_θ have values in $\mathbb{Z}^{d\times 1}$ and set $L(\theta; x) = \mathbb{P}(X_\theta = x)$, with the restriction $x \in \mathbb{Z}^{d\times 1}$. Then the results of Theorems 1–7 hold, with the assumption (2.9) replaced by the assumption that $|M_0(s + \mathrm{i}\varphi; \theta)| < M_0(s; \theta)$ for all $(s, \theta) \in \operatorname{int}\mathcal{S}$ and $\varphi \in [-\pi, \pi]^{1\times d} \setminus \{0\}$.*

In Theorem 8, the assumption $|M_0(s + \mathrm{i}\varphi; \theta)| < M_0(s; \theta)$ can be interpreted as a "nonlattice" condition on the distribution Y_θ. For instance, the condition fails if Y_θ has an entry with only odd values. On the other hand, the condition holds if, for some $n \in \mathbb{N}$, $\mathbb{P}(X_\theta = x) > 0$ for all $x \in \{0, 1\}^{d\times 1}$. Meanwhile, because $M_0(s + \mathrm{i}\varphi; \theta)$ is periodic in φ, (2.10) can be verified by showing that the relevant partial derivatives are continuous.

2.5. *Discussion*.

2.5.1. *Application and scope of the results*. The results in this paper show that the saddlepoint MLE offers a high degree of asymptotic accuracy, considered as a substitute for the true MLE. Most notably, when we consider the limit $n \to \infty$, the approximation error in using the saddlepoint MLE, of size $O(1/n^2)$, $O(1/n^{3/2})$ or $O(1/n)$, is negligible compared to the inferential uncertainty, of order $1/\sqrt{n}$ or larger, inherent in the MLE.

The scope for applying saddlepoint methods is reasonably wide: it is enough to know the moment generating function M (either exactly or numerically to high precision on a computer). Provided we know M, the saddlepoint approximation can be computed quickly, uniformly in n, whereas the true likelihood often becomes increasingly intractable for larger n. The model need not even fall into the standard asymptotic regime described here, as in [16], see Example 23 from Appendix J; indeed, we have followed Butler in thinking of (SPA_{SAR}) as the special case of (SPA) where X is a sum of n i.i.d. terms, rather than thinking of (SPA) as the special case of (SPA_{SAR}) where $n = 1$; see [4], Section 2.2.2. In all cases, the usual suite of likelihood-based approaches can be applied, using the saddlepoint likelihood as a substitute for the true likelihood.

The encouraging asymptotic results from Theorems 1–8 should, like many limiting statements, be interpreted with some caution in practice. Whereas the theorems apply when, for instance, we have a sequence of observed values x_n with $y_n = x_n/n \to y_0$ as $n \to \infty$, a typical application yields a single observed value x, with n fixed. Even if we interpret the observed value as being part of an infinite sequence, we may not have access to the limiting implied sample mean y_0 (as in Theorems 2–5) or the limiting rescaled deviation ξ_0 (as in Theorems 6–7), nor the corresponding parameter θ_0. Thus, in practice, rather than verifying that the Hessians from (2.15) or (H.5) are nonsingular at a specified base point θ_0, it may be more relevant to enquire whether these Hessians are nearly singular near the computed saddlepoint MLE $\hat{\theta}_{\text{MLE}}(x)$. However, such complications are not specific to the saddlepoint approximation; the same dilemma applies whenever we appeal to an asymptotic result to interpret a fixed data set.

Note also that this paper is concerned primarily with approximation accuracy. Theorems 1–8 guarantee under broad conditions that the saddlepoint log-likelihood, its gradient, and the resulting MLE are close to the true values, provided only that n is large. However, large n does not guarantee that the true or saddlepoint MLE values will be close to an underlying true parameter (if one exists) in the partially identifiable case from Sections 2.4.3–2.4.4. Such a guarantee would typically come from having k i.i.d. observations, with k large. Section 2.5.5 outlines how repeated observations can be implemented in our notation, but the theorems consider the number of i.i.d. observations to be fixed. This paper excludes from consideration the double limit $n \to \infty$, $k \to \infty$.

Finally, we note that our results concern local maxima and local neighbourhoods in parameter space. The true likelihood $L(\theta; x)$ may have a complicated global structure as a function of θ, with multiple local maxima, and the saddlepoint approximation cannot do better than faithfully replicating this complicated structure. Moreover, the saddlepoint approximation might have greater error in distant parts of parameter space, so that the saddlepoint MLE might fail to exist globally even if the true likelihood has a global maximum; see Example 30 in Appendix J. This possibility does not usually arise in practice but seems difficult to rule out *a priori*. We note, however, that the uniformity assertions in Theorems 1–8 robustly handle variability in the observed values: if we assume that x is drawn from the distribution X_{θ_0}, the law of large numbers means that $y = X_{\theta_0}/n$ will lie in a neighbourhood of $y_0 = \mathbb{E}(Y_{\theta_0})$ for n large enough.

2.5.2. *Integer-valued versus continuous distributions.* On the face of it, the discrete Theorem 8 parallels quite closely the continuous Theorems 1–7; the conclusions are identical, and the hypotheses are quite similar. There is however an important contextual difference: in the discrete case, the restriction to $\operatorname{int}\mathcal{S}$ and $\mathcal{Y}^o$ excludes values of interest. For instance, suppose X represents count data, $X \in \mathbb{Z}_+^{d\times 1}$. If we observe a count of 0 (or an observed vector including one or more zero counts), then we will be unable to find the saddlepoint $\hat{s}$. Even if we circumvent this issue by interpreting $\hat{s}$ as the limit $s \to -\infty$, the resulting saddlepoint approximation will still diverge, and no MLE can be computed.

This distinction arises in part because, as we shall discuss in Section 3.1, the saddlepoint approximation is fundamentally a (normal) density approximation, and it can be problematic when applied as an approximation for a probability mass function. It is the author's intention to return to this topic in future research.

2.5.3. *Heuristic for the true MLE and saddlepoint MLE.* We give a heuristic argument for how the size of the gradient error from Theorem 1 leads to the size of the MLE error in Theorem 2. Namely, fix $x = ny_0$ and assume as a simplification that

- the function $\theta \mapsto K_0(\hat{s}_0(\theta, y_0); \theta) - \hat{s}_0(\theta, y_0)y_0$ from (2.12)—which by (2.14)–(2.15) has a nondegenerate local maximum at $\theta = \theta_0$—is purely quadratic around its maximum value, say $\theta \mapsto a + \frac{1}{2}(\theta - \theta_0)^T H(\theta - \theta_0)$ where H is negative definite;
- the function $\theta \mapsto -\frac{1}{2}\log\det K_0''(\hat{s}(\theta, y_0); \theta)$ is purely affine, say $\theta \mapsto b + u(\theta - \theta_0)$ for some fixed $u \in \mathbb{R}^{1\times p}$; and
- the difference $\nabla_\theta \log\hat{L}(\theta; x) - \nabla_\theta \log L(\theta; x)$ from Theorem 1 has the form $\frac{1}{n}v$ for some fixed $v \in \mathbb{R}^{1\times p}$.

Under these assumptions,

$$\log\hat{L}(\theta; x) = na + \frac{1}{2}n(\theta - \theta_0)^T H(\theta - \theta_0) - \frac{d}{2}\log(2\pi n) + b + u(\theta - \theta_0), \tag{2.33}$$

where d is the dimension of the vectors X, x. We can complete the square to find

$$\begin{aligned}\log\hat{L}(\theta; x) = {}& \frac{1}{2}n\left(\theta - \theta_0 + \frac{1}{n}H^{-1}u^T\right)^T H\left(\theta - \theta_0 + \frac{1}{n}H^{-1}u^T\right)\\ & - \frac{1}{2n}uH^{-1}u^T + na - \frac{d}{2}\log(2\pi n) + b.\end{aligned} \tag{2.34}$$

Thus the saddlepoint MLE comes to

$$\hat{\theta}_{\mathrm{MLE}}(x) = \theta_0 - \frac{1}{n}H^{-1}u^T. \tag{2.35}$$

For the true likelihood, define the constants $c_n = \log\hat{L}(\theta_0; x) - \log L(\theta_0; x)$. Then

$$\begin{aligned}\log L(\theta; x) = {}& na + \frac{1}{2}n(\theta - \theta_0)^T H(\theta - \theta_0) - \frac{d}{2}\log(2\pi n)\\ & + b + u(\theta - \theta_0) - c_n - \frac{1}{n}v(\theta - \theta_0).\end{aligned} \tag{2.36}$$

Comparing (2.36) with (2.33), we see that changing from $\hat{L}$ to L amounts to replacing u by $u - \frac{1}{n}v$ and subtracting a constant term c_n. We can again complete the square to find

$$\begin{aligned}&\log L(\theta; x)\\ &= \frac{1}{2}n\left(\theta - \theta_0 + \frac{1}{n}H^{-1}u^T - \frac{1}{n^2}H^{-1}v^T\right)^T H\left(\theta - \theta_0 + \frac{1}{n}H^{-1}u^T - \frac{1}{n^2}H^{-1}v^T\right)\\ &\quad - \frac{1}{2n}\left(u - \frac{1}{n}v\right)H^{-1}\left(u - \frac{1}{n}v\right)^T + na - \frac{d}{2}\log(2\pi n) + b - c_n\end{aligned} \tag{2.37}$$

leading to

$$\theta_{\mathrm{MLE}}(x) = \theta_0 - \frac{1}{n} H^{-1} u^T + \frac{1}{n^2} H^{-1} v^T, \tag{2.38}$$

with an extra term of order $1/n^2$ in accordance with Theorem 2.

Note that order $1/n^2$ is better than what follows from the likelihood error bound (2.8) alone; knowing only that $\log \hat{L}$ and $\log L$ differ by $O(1/n)$, we could conclude at best that $\hat{\theta}_{\mathrm{MLE}}$ and θ_{MLE} differ by $O(1/n)$, since that is the size of the region in which the functions remain within $O(1/n)$ of their maximum even in the fully identifiable case. In the heuristic calculation above, however, the size of c_n (the log-likelihood approximation error at $\theta = \theta_0$) was irrelevant to the size of the MLE error. So indeed was the term $-\frac{d}{2}\log(2\pi n)$ in $\log \hat{L}$. In fact, even if we drop all terms arising from the factor $(\det(2\pi K''(\hat{s}(\theta, x); \theta)))^{-1/2}$ in the saddlepoint approximation $\hat{L}$, the heuristic suggests that the resulting MLE approximation would still be within $O(1/n)$ of the true MLE. This intuition is correct; see Theorem 9.

2.5.4. *The well-specified case.* In Theorems 2–4, some simplification occurs if

$$y_0 = \mathbb{E}(Y_{\theta_0}), \tag{2.39}$$

that is, if the limiting implied sample mean matches with the model mean for some parameter value. In this case, we might say that the model is "well-specified at the level of the mean." By the law of large numbers, this condition holds under the usual assumption of well-specifiedness where the observed value x is itself drawn randomly with the distribution X_{θ_0}, because then $X_{\theta_0}/n \to \mathbb{E}(Y_{\theta_0})$ as $n \to \infty$.

If (2.39) holds and $(0, \theta_0) \in \operatorname{int} \mathcal{S}$, we see from (2.3) that $s_0 = 0$ is the solution of the saddlepoint equation $K_0'(s_0) = y_0$. Since $K_0(0; \theta) = 0$ for all θ, it follows that

$$\nabla_\theta K_0(0; \theta_0) = 0, \qquad \nabla_\theta^T \nabla_\theta K_0(0; \theta_0) = 0. \tag{2.40}$$

Thus the condition (2.14) holds automatically, and the matrix H from (2.15) simplifies to

$$H = -\big(\nabla_s \nabla_\theta K_0(0; \theta_0)\big)^T K_0''(0; \theta_0)^{-1} \big(\nabla_s \nabla_\theta K_0(0; \theta_0)\big). \tag{2.41}$$

The matrix $K_0''(0; \theta_0)$ is already positive definite by assumption (see (2.4)), so the condition (2.15) reduces to (2.24). Note that the $d \times p$ matrix $\nabla_s \nabla_\theta K_0(0; \theta_0)$ from (2.24) is the gradient of the mapping

$$\theta \mapsto \mathbb{E}(Y_\theta) \tag{2.42}$$

(evaluated at θ_0) so the condition (2.15)/(2.24) is equivalent to saying that the linear approximation (at θ_0) to the mapping $\theta \mapsto \mathbb{E}(Y_\theta)$ is one-to-one.

Heuristically, if the implied sample mean y_0 matches with the model at parameter value θ_0, and if the mapping $\theta \mapsto \mathbb{E}(Y_\theta)$ is one-to-one, then by the law of large numbers, y_0 is an unlikely observation under any other parameter value $\theta \neq \theta_0$. This matches the observation that the function (2.12) will vanish at $\theta = \theta_0$ and must be strictly negative elsewhere.

Conversely, if the gradient of $\theta \mapsto \mathbb{E}(Y_\theta)$ has rank $p_1 < p$, there will be a p_2-dimensional hyperplane, $p_2 = p - p_1$, along which the mapping $\theta \mapsto \mathbb{E}(Y_\theta)$ is constant to first order. Hence, for all θ along this surface, $\hat{s}(\theta, y_0) \approx 0$ continues to be an approximate solution of the saddlepoint equation, the leading-order coefficient $K_0(\hat{s}(\theta, y_0); \theta) - \hat{s}(\theta, y_0) y_0$ remains zero to first order, and the heuristic from Section 2.5.3 fails. In such a case, it may still be possible to apply Theorem 6, with the coordinate ν representing the position within p_2-dimensional level surfaces along which $\mathbb{E}(Y_\theta)$ is constant.

2.5.5. *Observing multiple samples.* As remarked in Section 2.2, the parameter n should not be interpreted as a sample size in the traditional sense since the summands $Y^{(1)}, \dots, Y^{(n)}$ of (2.6) are not observed. A model with i.i.d. observations $X^{(1)}, \dots, X^{(k)} \in \mathbb{R}^{d_0 \times 1}$ can be fit into the framework of this paper by concatenating them into a vector $\vec{X}$ of dimension $d = kd_0$. Likewise, the observed data values $x^{(1)}, \dots, x^{(k)} \in \mathbb{R}^{d_0 \times 1}$ are concatenated into a vector $\vec{x} \in \mathbb{R}^{kd_0 \times 1}$. We then study the likelihood $L_{\vec{X}}(\theta; \vec{x})$ and MLE $\theta_{\mathrm{MLE}}(\vec{x}) = \operatorname{argmax}_\theta L_{\vec{X}}(\theta; \vec{x})$ for a single combined observation $\vec{x}$ by applying Theorems 1–8 to $\vec{X}, \vec{x}$ instead of X, x.

When we apply the saddlepoint approximation to the concatenated vector $\vec{X}$ of dimension kd_0, the covariance matrix $K''_{\vec{X}}(\vec{s})$ will be block-diagonal. Other quantities appearing in Theorems 1–8 also take a special form; see Example 20 in Appendix J for further details. For present purposes, we note that the saddlepoint approximation for $\vec{X}$ factors as a product of k d_0-dimensional saddlepoint approximations for $X^{(1)}, \dots, X^{(k)}$, each of which is a saddlepoint approximation applied to the distribution X_θ.

This paper considers only the limit $n \to \infty$ along which the saddlepoint approximation becomes more accurate. In particular, the number k of i.i.d. observations is considered to be fixed throughout.

2.5.6. *Sufficient statistics of exponential families.* Consider the case where X is a sufficient statistic for a full exponential family of distributions with natural parameter $\eta \in \mathbb{R}^{1 \times d}$. That is, we assume that $p = d$ and that $\eta = \eta(\theta)$ is a reparametrisation of $\theta \in \mathcal{R} \subset \mathbb{R}^{d \times 1}$, that is, the mapping $\theta \mapsto \eta(\theta)$ and its inverse are smooth, and η varies over an open subset of $\mathbb{R}^{1 \times d}$. Then we can write

$$L(x; \theta) = f(x; \theta) = h(x) \exp\bigl(\eta x - \rho(\eta)\bigr), \qquad K_X(s; \theta) = \rho(\eta + s) - \rho(\eta), \tag{2.43}$$

for scalar-valued functions h, ρ with ρ convex. The saddlepoint equation (SE) reduces to

$$\rho'(\eta + \hat{s}) = x. \tag{2.44}$$

In particular, the quantity $\hat{\eta} = \eta + \hat{s}$ depends on x alone and is fixed as a function of η, provided (2.44) has a solution. The saddlepoint approximation can be written as

$$\hat{f}(x; \theta) = \frac{\exp(\rho(\hat{\eta}) - \hat{\eta} x)}{\sqrt{\det(2\pi \rho''(\hat{\eta}))}} \exp\bigl(\eta x - \rho(\eta)\bigr). \tag{2.45}$$

The first factor need not coincide with $h(x)$, so the saddlepoint approximation need not be exact, but because the first factor depends on x only, *the saddlepoint MLE is exact for an exponential family* provided that the saddlepoint approximation itself is well-defined. Indeed, in terms of the natural parameter η, the MLE is precisely the quantity $\hat{\eta} = \eta + \hat{s}$ solving (2.44), which we already find in the course of computing the saddlepoint.

2.6. *Guide to examples.* Appendix J contains a number of examples based on theory and on the literature. Example 19 discusses the case $X = AU$, where A is a constant matrix and U is a random vector with K_U known. Example 20 gives the details for the construction of Section 2.5.5, in which $\vec{X}$ is a concatenation of k independent random vectors each formed in accordance with (2.6) and (SAR).

Examples 21–23 examine applications of saddlepoint MLEs in the literature [7, 16, 21]. Example 21 gives further detail for certain models mentioned in Section 2.3; it illustrates the fully identifiable case of Theorems 2–5, and also the setup of Example 19. Example 22 illustrates the partially identifiable case of Theorems 6–7, and also the setup of Example 20. Example 23 shows a model that falls outside the setup of (SAR).

Examples 24–33 explore families of distributions for which some direct calculations are possible, including the normal, Poisson and Gamma families; families where the true likelihood has different global behaviour than the saddlepoint likelihood; and ill-behaved distributions showing how the regularity conditions (2.9)–(2.10) may hold or fail.

3. Structure of the saddlepoint approximation. In Section 3.1, we break down the saddlepoint approximation into two steps: an exact step based on tilting, and an approximation step based on the normal distribution. Understanding the saddlepoint approximation via tilting is not a new idea (cf. for instance [17], Section 2) but here we use it to motivate a novel factorisation of the likelihood into an exact factor, which encodes the effect of tilting and is shared between the true and approximate likelihoods; and a correction term, a normal approximation of which leads to the saddlepoint likelihood. This factorisation establishes the framework in which the proofs will take place.

As a natural by-product of the factorisation, we introduce in Section 3.2 a simpler but less accurate alternative to the saddlepoint approximation, which satisfies results similar to Theorems 1–8.

3.1. *Tilting and the saddlepoint approximation.* Calculating the saddlepoint approximation splits naturally into two steps. For given θ, x, we first compute the saddlepoint $\hat{s}(\theta, x)$ by solving (SE), and then this value is substituted into the expression from (SPA). As we now explain, the first step can be understood in terms of tilting, and this will clarify the nature of the approximation made in the second step.

If X has density function f, the MGF $M(\mathrm{i}\varphi;\theta)$ along the imaginary axis gives the Fourier transform of f. Consequently, we can use the inverse Fourier transform to recover f:

$$f(x;\theta) = \int_{\mathbb{R}^{1\times d}} M(\mathrm{i}\varphi;\theta) e^{-\mathrm{i}\varphi x} \frac{d\varphi}{(2\pi)^d} = \int_{\mathbb{R}^{1\times d}} \exp\bigl(K(\mathrm{i}\varphi;\theta) - \mathrm{i}\varphi x\bigr) \frac{d\varphi}{(2\pi)^d}. \tag{3.1}$$

In practice, either of the factors $M(\mathrm{i}\varphi;\theta)$ or $e^{-\mathrm{i}\varphi x}$ may be highly oscillatory, and will not typically cancel with each other. To make the integral more manageable, we *exponentially tilt* the distribution of X. Define

$$f_{s_0}(x;\theta) = \frac{e^{s_0 x} f(x;\theta)}{M(s_0;\theta)} \tag{3.2}$$

for all $s_0 \in \mathcal{S}_\theta$. Then f_{s_0} is still a density function, corresponding to a tilted distribution $X_\theta^{(s_0)}$ having the same support as X_θ, and we compute

$$M_{X^{(s_0)}}(s;\theta) = \frac{M(s_0+s;\theta)}{M(s_0;\theta)}, \qquad K_{X^{(s_0)}}(s;\theta) = K(s_0+s;\theta) - K(s_0;\theta). \tag{3.3}$$

Recalling (2.3), we note that $K'(s_0;\theta) = \mathbb{E}(X_\theta^{(s_0)})$ and $K''(s_0;\theta) = \mathrm{Cov}(X^{(s_0)}, X^{(s_0)})$ can themselves be interpreted as means and covariance matrices, respectively. We remark also that, since X_θ and $X_\theta^{(s_0)}$ have the same support, (2.4) is unaffected by the choice of $s \in \operatorname{int}\mathcal{S}_\theta$.

If we apply the inversion formula (3.1) to $X_\theta^{(s)}$, we can solve to find

$$\begin{aligned} f(x;\theta) &= M(s;\theta) e^{-sx} f_s(x;\theta) \\ &= \exp\bigl(K(s;\theta) - sx\bigr) \int_{\mathbb{R}^{1\times d}} \exp\bigl(K(s+\mathrm{i}\varphi;\theta) - K(s;\theta) - \mathrm{i}\varphi x\bigr) \frac{d\varphi}{(2\pi)^d}. \end{aligned} \tag{3.4}$$

Note that the assumption (2.9) implies that the integral in (3.4) converges absolutely, and thus defines a continuous density function $x \mapsto f(x;\theta)$ whenever $(s,\theta) \in \operatorname{int}\mathcal{S}$ and $n\delta(s,\theta) > 1$, and this justifies our use of the Fourier inversion formula. Subject to this condition, we can choose $s \in \mathcal{S}_\theta$ arbitrarily. To make the integral more tractable, we wish to choose s so that the linear term $-\mathrm{i}\varphi x$ cancels with $K(s+\mathrm{i}\varphi;\theta) - K(s;\theta)$ to leading order. That is, we choose $s = \hat{s}(\theta;x)$, the solution of (SE), provided that $x \in \mathcal{X}_\theta$. Replacing x by $K'(\hat{s}(\theta,x);\theta)$,

$$\begin{aligned} f(x;\theta) &= \exp\bigl(K(\hat{s};\theta) - \hat{s}K'(\hat{s};\theta)\bigr) \\ &\quad \cdot \int_{\mathbb{R}^{1\times d}} \exp\bigl(K(\hat{s}+\mathrm{i}\varphi;\theta) - K(\hat{s};\theta) - \mathrm{i}\varphi K'(\hat{s};\theta)\bigr) \frac{d\varphi}{(2\pi)^d}. \end{aligned} \tag{3.5}$$

At this point, explicit x dependence has been virtually eliminated from (3.5); all x dependence on the right-hand side is now carried implicitly by $\hat{s} = \hat{s}(\theta, x)$. We introduce new notation to exploit this feature. Along with this new notation, we again switch from considering the density f to the likelihood L.

Define

$$
\begin{aligned}
L^*(s,\theta) &= \exp\bigl(K(s;\theta) - sK'(s;\theta)\bigr),\\
P(s,\theta) &= \int_{\mathbb{R}^{1\times d}} \exp\bigl(K(s+\mathrm{i}\varphi;\theta) - K(s;\theta) - \mathrm{i}\varphi K'(s;\theta)\bigr)\frac{d\varphi}{(2\pi)^d},
\end{aligned}
\tag{3.6}
$$

for all $s \in \operatorname{int}\mathcal{S}_\theta$. Setting $x = K'(s;\theta)$, (3.5) can be reformulated as

$$
L\bigl(\theta; K'(s;\theta)\bigr) = L^*(s,\theta)P(s,\theta). \tag{3.7}
$$

We can recognise $\log L^*(s,\theta)$ as the negative of the relative entropy (or Kullback–Leibler divergence) for the distribution of $X_\theta^{(s)}$ relative to X_θ. We can also give a probabilistic interpretation of $P(s,\theta)$: it is the density of the tilted distribution $X_\theta^{(s)}$ at its mean $K'(s;\theta)$.

We emphasise at this point that all the calculations so far are exact. The factor L^* measures the probabilistic "cost" of shifting from the original distribution X_θ to the tilted distribution $X_\theta^{(s)}$, but no distributional information is lost in the tilting step.

The natural next step will be to make an approximation to $P(s,\theta)$. Define

$$
\hat{P}(s,\theta) = \frac{1}{\sqrt{\det(2\pi K''(s;\theta))}}. \tag{3.8}
$$

Then the saddlepoint approximation (SPA) becomes

$$
\hat{L}\bigl(\theta; K'(s;\theta)\bigr) = L^*(s,\theta)\hat{P}(s,\theta). \tag{3.9}
$$

We recognise $\hat{P}(s,\theta)$ as the density (at its mean) of a normal random variable with covariance matrix $K''(s;\theta)$. Comparing (3.7) and (3.9), we see that the saddlepoint approximation amounts to approximating $P(s,\theta)$, the density at its mean of the tilted distribution, by $\hat{P}(s,\theta)$, the density at its mean of the normal random variable with the same covariance matrix.

In the remainder of the paper, we will study the two-variable functions $L^*(s,\theta)$, $P(s,\theta)$ and $\hat{P}(s,\theta)$ and their gradients. Since the factor $L^*(s,\theta)$ appears in both (3.7) and (3.9), it will suffice to compare the gradients of P and $\hat{P}$. Ultimately, our interest will be in the case $s = \hat{s}(\theta, x)$, and we will therefore rewrite (3.7) and (3.9) as

$$
\begin{aligned}
L(\theta;x) &= L^*\bigl(\hat{s}(\theta,x),\theta\bigr)P\bigl(\hat{s}(\theta,x),\theta\bigr),\\
\hat{L}(\theta;x) &= L^*\bigl(\hat{s}(\theta,x),\theta\bigr)\hat{P}\bigl(\hat{s}(\theta,x),\theta\bigr)
\end{aligned}
\qquad \text{for } x \in \operatorname{int}\mathcal{X}_\theta. \tag{3.10}
$$

REMARK. A key advantage to studying $L^*(s,\theta)$, $P(s,\theta)$ and $\hat{P}(s,\theta)$ is that, because P and $\hat{P}$ both represent densities at the mean, they vary less dramatically as a function of their arguments than the full likelihoods $L(\theta;x)$ and $\hat{L}(\theta;x)$, even in the limit $n \to \infty$ from (SAR). As we shall see, both P and $\hat{P}$ and their gradients behave asymptotically as powers of n. The only factor that decays exponentially in n is L^*, and since it is a common factor of $L(\theta;x)$ and $\hat{L}(\theta;x)$ we can circumvent its effects.

The representation (3.10) is also useful for numerical calculation. Recent work by Lunde, Kleppe and Skaug [13] calculates densities and likelihoods to high relative precision, even in the tails, using saddlepoint methodology. Expressed in our notation, they evaluate $P(s,\theta)$ by applying a quadrature rule to the integral in (3.6). Because $P(s,\theta)$ is a density at the mean,

the integral defining $P(s,\theta)$ is much more amenable to numerical integration, in general, than the inverse Fourier integral (3.1).

Likewise, the fact that P and $\hat{P}$ both represent densities at the mean helps to explain why the saddlepoint approximation can be so accurate even in the tails. Given x close to the boundary of $\mathcal{X}_\theta$ (or tending to infinity), the corresponding saddlepoint $\hat{s}$ will be close to the boundary of $\mathcal{S}_\theta$ (or will tend to infinity). In such a limit, there is no reason in general to expect the ratio $\hat{P}/P$ to converge to 1; but often the tilted density at the tilted mean remains on the order of the inverse of the tilted standard deviation, so that $\hat{P}/P$ may remain bounded. By contrast, normal approximations and other similar techniques rely on extrapolating a density far from the mean, and such an extrapolation can only be accurate if the true density happens to have the same tail behaviour far from the mean; see Appendix I.

L^* and P in the integer-valued case. When $X \in \mathbb{Z}^{d\times 1}$, the analogue of (3.1) is

$$\mathbb{P}(X_\theta = x) = \int_{[-\pi,\pi]^{1\times d}} M(\mathrm{i}\varphi;\theta) e^{-\mathrm{i}\varphi x} \frac{d\varphi}{(2\pi)^d} \quad \text{for } x \in \mathbb{Z}^{d\times 1}. \tag{3.11}$$

We can repeat the tilting argument above, leading us to define

$$P_{\mathrm{int}}(s,\theta) = \int_{[-\pi,\pi]^{1\times d}} \exp\bigl(K(s+\mathrm{i}\varphi;\theta) - K(s;\theta) - \mathrm{i}\varphi K'(s;\theta)\bigr) \frac{d\varphi}{(2\pi)^d}. \tag{3.12}$$

At integer values $x \in \mathbb{Z}^{d\times 1}$, we take the likelihood function to be $L(\theta;x) = \mathbb{P}(X_\theta = x)$, and we will have

$$L(\theta;x) = L^*\bigl(\hat{s}(\theta,x),\theta\bigr) P_{\mathrm{int}}\bigl(\hat{s}(\theta,x),\theta\bigr). \tag{3.13}$$

However, $P_{\mathrm{int}}(s,\theta)$ is defined whenever $s \in \operatorname{int}\mathcal{S}_\theta$. Thus we can use (3.13) as a definition of $L(\theta;x)$ even when x is noninteger, although $L(\theta;x)$ may not represent a probability in that case.

3.2. *A lower-order saddlepoint approximation.* Approximating P using first and second moments is a natural and reasonable step, but is not the only possible approach. As we have seen, the standard saddlepoint approximation selects a density from the family of normal distributions after matching first and second moments. An alternative is to use a different reference family of distributions; this is one approach to non-Gaussian saddlepoint approximations, originally developed by Wood, Booth and Butler [20] in a different context and with different tools.

An even simpler alternative, however, is to ignore P altogether. Define

$$\hat{L}^*(\theta;x) = L^*\bigl(\hat{s}(\theta,x),\theta\bigr) \tag{3.14}$$

for $x \in \mathcal{X}_\theta$. We could describe $\hat{L}^*(\theta;x)$ as the "zeroth-order" saddlepoint approximation to the likelihood. We saw the quantity $\frac{1}{n}\log \hat{L}^*(\theta;ny_0)$ and its derivatives in (2.12) and (2.14)–(2.15), and $I_\theta(x) = -\log \hat{L}^*(\theta;x)$ are the large deviations rate function from Cramér's theorem applied to X_θ. Equivalently, $-\log \hat{L}^*(\theta;\cdot)$ is the Legendre transform of $K(\cdot;\theta)$; see, for instance, [8], Sections I.4 and V.1, or [11], Section 1.2.

Form the corresponding maximum likelihood estimator,

$$\hat{\theta}^*_{\mathrm{MLE}}(x) = \operatorname*{argmax}_\theta \hat{L}^*(\theta;x) \tag{3.15}$$

when it exists, and likewise define $\hat{\pi}^*_{\Theta|U,x}$ as in (2.18) with $\hat{L}$ replaced by $\hat{L}^*$. We remark that in the standard asymptotic regime (SAR) where $x = ny$, the MLE $\hat{\theta}^*_{\mathrm{MLE}}(x)$ depends on y but not on n; the same is true if we maximise over θ restricted to a neighbourhood U.

Applied to $\hat{L}^*$, the conclusions of Theorems 1–5 and 8 are almost unchanged, except that error terms have powers of n changed by one.

THEOREM 9 (Zeroth-order saddlepoint approximation—fully identifiable case).

(a) *Under the hypotheses of Theorem 1,*

$$\nabla_\theta \log \hat{L}^*(\theta; x) = \nabla_\theta \log L(\theta; x) + O(1) \quad \text{as } n \to \infty \tag{3.16}$$

for $(y, \theta) \in \mathcal{Y}^o$, *with uniformity if* (y, θ) *is restricted to a compact subset of* $\mathcal{Y}^o$.

(b) *Under the hypotheses of Theorem 2, there exist* $n_0 \in \mathbb{N}$ *and neighbourhoods* $U \subset \mathcal{R}$ *of* θ_0 *and* $V \subset \mathbb{R}^{d\times 1}$ *of* y_0 *such that, for all* $n \geq n_0$ *and* $y \in V$, *the function* $\theta \mapsto \hat{L}^*(\theta; x)$ *has a unique local maximiser in* U *and, writing this local maximiser as* $\hat{\theta}^*_{\text{MLE in } U}(x)$,

$$\left|\hat{\theta}^*_{\text{MLE in } U}(x) - \theta_{\text{MLE in } U}(x)\right| = O(1/n) \quad \text{as } n \to \infty. \tag{3.17}$$

(c) *Under the hypotheses of Theorem 3, with* $x = ny_0$, *the distribution of the rescaled parameter* $\sqrt{n}(\Theta - \theta_0)$ *under* $\hat{\pi}^*_{\Theta|U,x}$ *converges as* $n \to \infty$ *to the same limiting distribution* $\mathcal{N}(0, -H^{-1})$, *where* H *is the negative definite matrix from* (2.15).

(d) *Under the hypotheses of Theorems 4 or 5, the sampling distributions of the rescaled approximate MLEs* $\sqrt{n}(\hat{\theta}^*_{\text{MLE in } U}(\chi_n) - \theta_0)$ *or* $\sqrt{n}(\hat{\theta}^*_{\text{MLE in } U}(X_{\theta_0}) - \theta_0)$, *respectively, converge as* $n \to \infty$ *to* Z, *where* Z *is as in Theorems 4 or 5, respectively. These convergences occur jointly with the convergences from Theorems 4 or 5, with the same limiting random variable* Z *in each case. Furthermore, Theorem 5*(*c*) *applies also with* $\Theta = \hat{\theta}^*_{\text{MLE in } U}(X_{\theta_0})$.

(e) *The above results also apply under the hypotheses of Theorem 8.*

(f) *Suppose* X *is the sufficient statistic for an exponential family indexed by the natural parameter* η. *Let* $\hat{\eta}^*_{\text{MLE}}(x)$ *be the parameter obtained by maximising* $\hat{L}^*$, *provided the maximiser exists. Then* $\hat{\eta}^*_{\text{MLE}}(x)$ *coincides with the true MLE.*

It is notable that maximizing $\hat{L}^*$ results in an approximation to the MLE whose error as $n \to \infty$ is still smaller than the spatial scale $1/\sqrt{n}$, even though $\hat{L}^*$ is no longer an adequate approximation to the likelihood itself in the sense that $\hat{L}^*/L \to \infty$ as $n \to \infty$.

The proofs of (a)–(e) are given along with the corresponding proofs of Theorems 1–5 and 8. Part (f) holds by the same reasoning as in Section 2.5.6: in the first factor in (2.45), we now remove the denominator, and the resulting factor still does not depend on η.

4. Proofs of Theorems 1–2. The proof of Theorem 1 is based on two key results, Proposition 10 and Corollary 12, that refine the statement of Theorem 1 using the factorisation from Section 3.1. Theorem 2 follows using a scaling analysis and the implicit function theorem. We begin with some derivative formulas; see Appendix B for their derivation.

4.1. *Summary of saddlepoint derivatives.* Under the scaling of (SAR), the quantities $\log L^*(s, \theta)$, $\log L^*(\hat{s}(\theta, x), \theta)$ and their gradients are proportional to n. Define

$$\begin{aligned} L^*_0(s, \theta) &= \exp\bigl(K_0(s; \theta) - s K_0'(s; \theta)\bigr), \\ \hat{L}^*_0(\theta; y) &= L^*_0\bigl(\hat{s}_0(\theta, y), \theta\bigr) = \exp\bigl(K_0\bigl(\hat{s}_0(\theta, y); \theta\bigr) - \hat{s}_0(\theta, y) y\bigr), \end{aligned} \tag{4.1}$$

so that

$$\log L^*(s, \theta) = n \log L^*_0(s, \theta), \qquad \nabla_\theta \log \hat{L}^*(\theta; x) = n \nabla_\theta \log \hat{L}^*_0(\theta; y) \tag{4.2}$$

and so on, where

$$\begin{aligned} \nabla_\theta \log L^*_0(s, \theta) &= \nabla_\theta K_0(s; \theta) - s \nabla_s \nabla_\theta K_0(s; \theta), \\ \nabla_s \log L^*_0(s, \theta) &= -K_0''(s; \theta) s^T, \\ \nabla_\theta \log \hat{L}^*_0(\theta; y) &= \nabla_\theta K_0\bigl(\hat{s}_0(\theta, y); \theta\bigr), \\ \nabla^T_\theta \nabla_\theta \log \hat{L}^*_0(\theta; y) &= \nabla^T_\theta \nabla_\theta K_0(\hat{s}; \theta) - \bigl(\nabla_s \nabla_\theta K_0(\hat{s}; \theta)\bigr)^T K_0''(\hat{s}; \theta)^{-1} \nabla_s \nabla_\theta K_0(\hat{s}; \theta). \end{aligned} \tag{4.3}$$

By contrast, $\hat{s}$, $\nabla_\theta \hat{s}^T$ and $\frac{\partial}{\partial t}\log \hat{P}$ do not depend on n:

$$\begin{aligned}
\hat{s}(\theta, x) &= \hat{s}_0(\theta, y),\\
\nabla_\theta \hat{s}^T(\theta, x) &= \nabla_\theta \hat{s}_0^T(\theta, y) = -K_0''\big(\hat{s}(\theta, x); \theta\big)^{-1} \nabla_s \nabla_\theta K_0\big(\hat{s}(\theta, x); \theta\big),\\
\frac{\partial}{\partial t}\log \hat{P}(s, \theta) &= -\frac{1}{2}\operatorname{tr}\left(K_0''(s; \theta)^{-1}\frac{\partial K_0''}{\partial t}(s; \theta)\right).
\end{aligned} \tag{4.4}$$

For the remainder of the proofs, we will emphasise that $P(s, \theta)$, $P_{\text{int}}(s, \theta)$, $L(\theta; x)$, $\hat{L}(\theta; x)$, $\theta_{\text{MLE}}(x)$, $\hat{\theta}_{\text{MLE}}(x)$ depend on n by writing them as $P_n(s, \theta)$, $P_{\text{int},n}(s, \theta)$, $L_n(\theta; x)$, $\hat{L}_n(\theta; x)$, $\theta_{\text{MLE}}(x, n)$, $\hat{\theta}_{\text{MLE}}(x, n)$. The n-dependences of $M(s; \theta)$, $K(s; \theta)$, $\log L^*(s, \theta)$ and $\log \hat{L}^*(\theta; x)$ are simple in form and we will handle them by directly substituting the formulas in (SAR) and (4.2). We remark that the expression in (2.12) reduces to $\log \hat{L}_0^*(\theta; y_0)$, and (2.14)–(2.15) amount to the assertion that θ_0 is a nondegenerate local maximum for $\theta \mapsto \log \hat{L}_0^*(\theta; y_0)$. Note that $\hat{P}_n(s, \theta)$ also depends on n in a simple way,

$$\hat{P}_n(s, \theta) = \frac{1}{\sqrt{\det(2\pi n K_0''(s; \theta))}} = \frac{n^{-d/2}}{\sqrt{\det(2\pi K_0''(s; \theta))}}, \tag{4.5}$$

but by (4.4) the gradients $\nabla_s \log \hat{P}(s, \theta)$, $\nabla_\theta \log \hat{P}(s, \theta)$ do not depend on n and we will omit the subscript in those cases.

4.2. *Proof of Theorem 1.* Unlike L^* and $\hat{P}_n$, the quantity P_n has no closed form and is instead given as an integral as in (3.6). In the standard asymptotic regime given by (SAR), we can substitute $K = nK_0$ to obtain

$$P_n(s, \theta) = \int_{\mathbb{R}^{1\times d}} \exp\big(n\big[K_0(s + \mathrm{i}\varphi; \theta) - K_0(s; \theta) - \mathrm{i}\varphi K_0'(s; \theta)\big]\big)\frac{d\varphi}{(2\pi)^d}. \tag{4.6}$$

Note that the integrand in (4.6) has the form $h(\varphi)e^{ng(\varphi)}$ with $g(0) = 0$ and $g'(0) = 0$. This is the standard setup for applying the multivariate Laplace method; see [3, 19]. Thus the limiting framework of (SAR), in which X is the sum of n i.i.d. terms, leads us to expect that $P_n(s, \theta)$, after suitable rescaling by a power of n, will have an asymptotic series expansion in powers of $1/n$.

Let the scalar t denote one of the coordinates θ_i or s_j. As we will show in the proof of Proposition 10, for large enough n we may differentiate under the integral sign:

$$\begin{aligned}
\frac{\partial P_n}{\partial t}(s, \theta) = \int_{\mathbb{R}^{1\times d}} &\exp\big(n\big[K_0(s + \mathrm{i}\varphi; \theta) - K_0(s; \theta) - \mathrm{i}\varphi K_0'(s; \theta)\big]\big)\\
&\cdot n\left[\frac{\partial K_0}{\partial t}(s + \mathrm{i}\varphi; \theta) - \frac{\partial K_0}{\partial t}(s; \theta) - \mathrm{i}\varphi\frac{\partial K_0'}{\partial t}(s; \theta)\right]\frac{d\varphi}{(2\pi)^d}.
\end{aligned} \tag{4.7}$$

A key observation motivating Theorem 1 is that $\frac{\partial P_n}{\partial t}(s, \theta)$ scales according to the same n-dependent factor as $P_n(s, \theta)$. Indeed, the values of both integrals arise primarily from the region where φ is of order $1/\sqrt{n}$, and the extra factor in (4.7) is of order 1 in that region. We could summarise by saying that taking gradients of $P_n(s, \theta)$ with respect to θ or s does not substantially change the nature of the n-dependence.

When we turn to MLEs, we will need to control the dependence of $P_n(s, \theta)$ (and its gradients) on s, θ and n simultaneously. Specifically, the proof of Theorem 2 is based on the implicit function theorem, which requires continuous differentiability. We will therefore prove the following result, which is more precise than is necessary for Theorem 1.

PROPOSITION 10. *Under the hypotheses of Theorem 1, with t being one of the entries θ_i or s_j, there are continuously differentiable functions $q_1(s,\theta,\varepsilon)$, $q_2(s,\theta,\varepsilon)$, defined on an open set $\mathcal{Q}$ containing $\{(s,\theta,0)\colon (s,\theta)\in \operatorname{int}\mathcal{S}\}$, such that $q_1(s,\theta,0)=q_2(s,\theta,0)=0$ and*

$$\begin{aligned} P_n(s,\theta) &= \hat{P}_n(s,\theta)\bigl(1+q_1(s,\theta,1/n)\bigr),\\ \frac{\partial P_n}{\partial t}(s,\theta) &= \hat{P}_n(s,\theta)\biggl(\frac{\partial}{\partial t}\log\hat{P}(s,\theta)+q_2(s,\theta,1/n)\biggr) \end{aligned} \tag{4.8}$$

whenever n is large enough that $(s,\theta,1/n)\in\mathcal{Q}$.

As with Theorem 1, Proposition 10 includes the implicit assertion that all the quantities in (4.8) exist when $(s,\theta,1/n)\in\mathcal{Q}$.

An almost identical statement holds in the integer-valued case.

PROPOSITION 11. *Under the hypotheses of Theorem 8, the conclusions of Proposition 10 hold with $P_n(s,\theta)$ replaced by $P_{\mathrm{int},n}(s,\theta)$.*

The proofs of Propositions 10–11 use many of the elements of standard proofs of Laplace's method. Additional care is needed to ensure that q_1, q_2 are continuously differentiable, and we defer the details to Appendix C.

Everything we will use from Propositions 10–11 can be encapsulated in the following corollary, or its analogue for $P_{\mathrm{int},n}(s,\theta)$.

COROLLARY 12. *Under the hypotheses of Theorem 1, there is a continuously differentiable function $q_3(\theta,y,\varepsilon)$, with values in $\mathbb{R}^{1\times p}$, defined on an open set $\mathcal{Q}'$ containing $\{(\theta,y,0)\colon (y,\theta)\in\mathcal{Y}^o\}$, such that $q_3(\theta,y,0)=0$ and*

$$\nabla_\theta\bigl(\log P_n\bigl(\hat{s}_0(\theta,y),\theta\bigr)\bigr)=\nabla_\theta\bigl(\log\hat{P}\bigl(\hat{s}_0(\theta,y),\theta\bigr)\bigr)+q_3(\theta,y,1/n) \tag{4.9}$$

whenever $(\theta,y,1/n)\in\mathcal{Q}'$.

The proof is again deferred to Appendix C. Theorem 1 now follows immediately.

PROOF OF THEOREM 1. Use (3.10) and cancel the term $\log L^*$ to obtain

$$\begin{aligned} \nabla_\theta\log L(\theta;x)-\nabla_\theta\log\hat{L}(\theta;x) &= \nabla_\theta\bigl(\log P_n\bigl(\hat{s}_0(\theta;y),\theta\bigr)\bigr)-\nabla_\theta\bigl(\log\hat{P}\bigl(\hat{s}_0(\theta;y),\theta\bigr)\bigr)\\ &= q_3(\theta,y,1/n). \end{aligned} \tag{4.10}$$

By Corollary 12, q_3 is continuously differentiable and $q_3(\theta,y,0)=0$ in a neighbourhood of (θ_0,y_0). An application of the mean value theorem (see, for instance, [18], Theorem 5.10) completes the proof. □

Although it is not necessary to our study of MLEs, we note that the same argument yields

$$\log L(\theta;x)-\log\hat{L}(\theta;x)=\tilde{q}(\theta,y,1/n), \tag{4.11}$$

where $\tilde{q}(\theta,y,\varepsilon)$ is a continuously differentiable function defined on $\mathcal{Q}'$ with $\tilde{q}(\theta,y,0)=0$. This is a more precise form of the basic saddlepoint error estimate (2.8). Theorem 9(a) also follows easily.

PROOF OF THEOREM 9(a). Since Theorem 1 already gives a bound on $\nabla_\theta\log\hat{L}-\nabla_\theta\log L$, it suffices to show that $\nabla_\theta\log\hat{L}-\nabla_\theta\log\hat{L}^*=O(1)$. From (3.10) and (4.4), $\nabla_\theta\log(\hat{L}(\theta;x)/\hat{L}^*(\theta;x))=\nabla_\theta\log\hat{P}(\hat{s}_0(\theta,y),\theta)$ is constant with respect to n, so it is $O(1)$ in the limit $n\to\infty$. Since it also depends continuously on θ and y, uniformity follows. □

4.3. *Proof of Theorem 2.* To study the MLE $\theta_{\text{MLE in } U}(x,n)$, we will show that the function $\theta \mapsto \log L_n(\theta; x)$ has a unique maximum when θ is restricted to lie in a suitably chosen neighbourhood U. In fact, it will be more convenient to consider the rescaled function

$$R_{x,n}(\theta) = \frac{1}{n}\log L_n(\theta; x) = \log L_0^*\bigl(\hat{s}_0(\theta, y), \theta\bigr) + \frac{1}{n}\log P_n\bigl(\hat{s}_0(\theta, y), \theta\bigr), \tag{4.12}$$

where we have substituted (3.10) and (4.2). Use (4.4) and Corollary 12 to compute

$$\begin{aligned}\nabla_\theta R_{x,n}(\theta) &= \nabla_\theta K_0\bigl(\hat{s}_0(\theta, y); \theta\bigr) + \frac{1}{n}q_3(\theta, y, 1/n) + \frac{1}{n}\nabla_\theta \log \hat{P}\bigl(\hat{s}_0(\theta, y), \theta\bigr)\\ &\quad - \frac{1}{n}\nabla_s^T \log \hat{P}\bigl(\hat{s}_0(\theta, y), \theta\bigr) K_0''\bigl(\hat{s}_0(\theta, y); \theta\bigr)^{-1}\nabla_s\nabla_\theta K_0\bigl(\hat{s}_0(\theta, y); \theta\bigr).\end{aligned} \tag{4.13}$$

We will define a function $F(s^T, \theta; y, \varepsilon)$ such that, with the substitution $\varepsilon = 1/n$, a solution of $F = 0$ corresponds to a critical point of $R_{x,n}$.

PROOF OF THEOREM 2. Define the functions

$$\begin{aligned}F\bigl(s^T, \theta; y, \varepsilon\bigr) &= \begin{pmatrix} F_1\bigl(s^T, \theta; y\bigr) \\ F_2\bigl(s^T, \theta; y, \varepsilon\bigr)\end{pmatrix}, \qquad F_1\bigl(s^T, \theta; y\bigr) = K_0'(s; \theta) - y,\\ F_2\bigl(s^T, \theta; y, \varepsilon\bigr) &= \nabla_\theta^T K_0(s; \theta) + \varepsilon q_3(\theta, y, \varepsilon)^T\\ &\quad + \varepsilon\bigl(\nabla_\theta^T \log \hat{P}(s, \theta) - \bigl(\nabla_s\nabla_\theta K_0(s; \theta)\bigr)^T K_0''(s; \theta)^{-1}\nabla_s \log \hat{P}(s, \theta)\bigr).\end{aligned} \tag{4.14}$$

We think of F_1, F_2, F as column-vector-valued functions of column-vector arguments, with $(s^T, \theta; y, \varepsilon)$ and $F(s^T, \theta; y, \varepsilon)$ interpreted as column vectors expressed in block form, of sizes $(2d + p + 1) \times 1$ and $(d + p) \times 1$, respectively. We will show that we can solve $F = 0$ to define θ and s implicitly as functions of y and ε; to indicate this, we will merge the column vectors s^T and θ and write $\binom{s^T}{\theta} = G(y, \varepsilon) \in \mathbb{R}^{(d+p)\times 1}$ such that $F(G(y, \varepsilon); y, \varepsilon) = 0$.

By Corollary 12, the function F is continuously differentiable with respect to all its parameters. Our assumptions imply that $F(s_0^T, \theta_0; y_0, 0) = 0$ and $\nabla_{s^T,\theta} F(s_0^T, \theta_0; y_0, 0)$ is nonsingular; see Appendix E. We can therefore apply the implicit function theorem (see, for instance, [18], Theorem 9.28) to find neighbourhoods U, V, W of θ_0, y_0, s_0, a neighbourhood $[-1/n_0, 1/n_0]$ of 0 and a continuously differentiable function $G(y, \varepsilon)$ defined on $V \times [-1/n_0, 1/n_0]$ such that, for all $y \in V$, $\varepsilon \in [-1/n_0, 1/n_0]$, the point $\binom{s^T}{\theta} = G(y, \varepsilon)$ is the unique solution in $W \times U$ of $F(s^T, \theta; y, \varepsilon) = 0$.

As outlined above, when $\varepsilon = 1/n$, the solution of $F = 0$ corresponds to the MLE.

LEMMA 13. *Possibly after shrinking U, V and increasing n_0, we have*

$$G(y, 1/n) = \begin{pmatrix} \hat{s}_0^T\bigl(\theta_{\text{MLE in } U}(x,n), y\bigr) \\ \theta_{\text{MLE in } U}(x,n)\end{pmatrix} \tag{4.15}$$

for all $y \in V$ and $n \geq n_0$, including the assertion that the maximum of $L_n(\theta; x)$, restricted to $\theta \in U$, is attained uniquely.

We defer the proof to Appendix E.

We now turn to the saddlepoint MLE, which amounts to omitting the term εq_3:

$$\begin{aligned}\hat{F}_2\bigl(s^T, \theta; y, \varepsilon\bigr) &= \nabla_\theta^T K_0(s; \theta) + \varepsilon\nabla_\theta^T \log \hat{P}(s, \theta)\\ &\quad - \varepsilon\bigl(\nabla_s\nabla_\theta K_0(s; \theta)\bigr)^T K_0''(s; \theta)^{-1}\nabla_s \log \hat{P}(s, \theta),\\ \hat{F}\bigl(s^T, \theta; y, \varepsilon\bigr) &= \begin{pmatrix} F_1\bigl(s^T, \theta; y\bigr) \\ \hat{F}_2\bigl(s^T, \theta; y, \varepsilon\bigr)\end{pmatrix}.\end{aligned} \tag{4.16}$$

Then $\hat{F}$ and its gradients agree with F and its gradients at $(s_0^T, \theta_0; y_0, 0)$, so that (after shrinking U, V, W and increasing n_0 if necessary) the implicit function theorem again produces a function $\hat{G}(y, \varepsilon)$ giving the unique solution $\binom{s^T}{\theta}$ in $U \times W$ of $\hat{F}(s^T, \theta; y, \varepsilon) = 0$. Moreover, the analogue of Lemma 13 applies, with the same proof, so that

$$\hat{G}(y, 1/n) = \begin{pmatrix} \hat{s}_0^T\big(\hat{\theta}_{\text{MLE in } U}(x, n), y\big) \\ \hat{\theta}_{\text{MLE in } U}(x, n) \end{pmatrix}. \tag{4.17}$$

For later convenience, write $\hat{G}$ in block form as $\hat{G} = \binom{\hat{G}_s^T}{\hat{G}_\theta}$.

To compare $G(y, \varepsilon)$ with $\hat{G}(y, \varepsilon)$, note that F_2 and $\hat{F}_2$ are close:

$$F_2\big(s^T, \theta; y, \varepsilon\big) = \hat{F}_2\big(s^T, \theta; y, \varepsilon\big) + \varepsilon q_3(\theta, y, \varepsilon). \tag{4.18}$$

In particular, $F(\hat{G}(y, \varepsilon); y, \varepsilon)$ is almost zero:

$$F\big(\hat{G}(y, \varepsilon); y, \varepsilon\big) = \begin{pmatrix} F_1\big(\hat{G}(y, \varepsilon); y, \varepsilon\big) \\ \hat{F}_2\big(\hat{G}(y, \varepsilon); y, \varepsilon\big) \end{pmatrix} + \begin{pmatrix} 0 \\ \varepsilon q_3\big(\hat{G}_\theta(y, \varepsilon), y, \varepsilon\big) \end{pmatrix}. \tag{4.19}$$

The first term in the right-hand side vanishes by definition. In the second term, note that $q_3(\hat{G}_\theta(y, \varepsilon), y, \varepsilon)$ is a continuously differentiable function of (y, ε) that vanishes whenever $\varepsilon = 0$ (since q_3 has the same property by Corollary 12). We can therefore conclude that

$$F\big(\hat{G}(y, \varepsilon); y, \varepsilon\big) = \varepsilon^2 q_4(y, \varepsilon), \tag{4.20}$$

where $|q_4(y, \varepsilon)| \le C$ for (y, ε) in a suitable neighbourhood of $(y_0, 0)$.

To make use of (4.20), we define an augmented version of F that is locally invertible. Let

$$\begin{aligned} \widetilde{F}\big(s^T, \theta; y, \varepsilon\big) &= \begin{pmatrix} F\big(s^T, \theta; y, \varepsilon\big) \\ y \\ \varepsilon \end{pmatrix} \quad \text{so that} \\ \nabla_{s^T, \theta, y, \varepsilon} \widetilde{F} &= \begin{pmatrix} \nabla_{s^T, \theta} F & \nabla_y F & \nabla_\varepsilon F \\ 0 & I_{m \times m} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{4.21}$$

in block form. Thus $\nabla_{s^T, \theta, y, \varepsilon} \widetilde{F}(s_0^T, \theta_0; y_0, 0)$ is an invertible $(2d + p + 1) \times (2d + p + 1)$ matrix, and by the inverse function theorem [18], Theorem 9.24, after shrinking the domain of $\widetilde{F}$ if necessary, $\widetilde{F}$ has a continuously differentiable inverse function $\widetilde{G}(u; y, \varepsilon)$. As above, we may shrink the domain further to make the partial derivatives of $\widetilde{G}$ uniformly bounded.

The inverse function $\widetilde{G}$ is related to the implicit function G by $\widetilde{G}(0; y, \varepsilon) = \binom{G(y, \varepsilon)}{\substack{y \\ \varepsilon}}$. From (4.20) and the definition of $\widetilde{F}$, we have

$$\widetilde{F}\big(\hat{G}(y, \varepsilon); y, \varepsilon\big) = \begin{pmatrix} \varepsilon^2 q_4(y, \varepsilon) \\ y \\ \varepsilon \end{pmatrix} \quad \text{and so} \quad \widetilde{G}\big(\varepsilon^2 q_4(y, \varepsilon); y, \varepsilon\big) = \begin{pmatrix} \hat{G}(y, \varepsilon) \\ y \\ \varepsilon \end{pmatrix}. \tag{4.22}$$

Thus, setting $\varepsilon = 1/n$ for n sufficiently large,

$$\begin{aligned} \big|\hat{\theta}_{\text{MLE in } U}(x, n) - \theta_{\text{MLE in } U}(x, n)\big| &\le \big|\hat{G}(y, 1/n) - G(y, 1/n)\big| \\ &= \big|\widetilde{G}\big(n^{-2} q_4(y, 1/n); y, 1/n\big) - \widetilde{G}(0; y, 1/n)\big|. \end{aligned} \tag{4.23}$$

The boundedness of q_4 and of the partial derivatives of $\widetilde{G}$ imply that this upper bound is $O(1/n^2)$, uniformly over y in a suitable neighbourhood of y_0. $\square$

The corresponding assertion from Theorem 9 has a similar proof.

PROOF OF THEOREM 9(b). Note that $\theta = \hat{\theta}^*_{\text{MLE in } U}(x, n)$ and $s = \hat{s}(\theta, x)$ are the solutions of $F(s^T, \theta; y, 0) = 0$. (The proof is the same as for Lemma 13; see Appendix E.) Then

$$|\hat{\theta}^*_{\text{MLE in } U}(x, n) - \theta_{\text{MLE in } U}(x, n)| \le |G(y, 0) - G(y, 1/n)| \tag{4.24}$$

and the conclusion follows from the fact that G is continuously differentiable. □

5. Conclusion. A long-established result tells us that the saddlepoint approximation gives a relative error of order $1/n$. That is, applied to a random variable given as a sum of n i.i.d. terms, the saddlepoint approximation estimates the values of the density (or likelihood) up to a factor of the form $1 + O(1/n)$: see (2.8) or Proposition 10. Very commonly, however, we are not interested in the likelihood for its own sake but rather as a step towards computing the MLE. This paper gives the analogous basic results, Theorems 2 and 6, for the approximation error between the true MLE and saddlepoint MLE: under certain explicit identifiability conditions, it is $O(1/n)$, $O(1/n^{3/2})$, or $O(1/n^2)$.

It is worth noting that this MLE error estimate is sharper than what we obtain from the basic likelihood error estimate. In the fully identifiable case, knowing that $\log \hat{L}(\theta; x)$ has a maximum at $\theta = \hat{\theta}_{\text{MLE}}(x)$ and knowing that the true log-likelihood satisfies $|\log L(\theta; x) - \log \hat{L}(\theta; x)| = O(1/n)$, we conclude only that $|\theta_{\text{MLE}}(x) - \hat{\theta}_{\text{MLE}}(x)| = O(1/n)$; see Section 2.5.3. Although an MLE error bound of size $O(1/n)$ is small compared to the scale of the inferential uncertainty in estimating θ (see Theorems 3–5 and the remarks at the end of Section 2.4.2), it is still a significant overestimate compared to the true MLE error $O(1/n^2)$. The results in this paper help to explain why saddlepoint MLEs in practice often turn out to be so much more accurate than expected.

A key point in the analysis is to ask how well the saddlepoint approximation captures the *shape* of the log-likelihood as a function of the parameter θ. Specifically, it is error bounds on the *gradient* of the log-likelihood, as in Theorem 1 and Corollary 12, that control the size of the MLE approximation error and lead to sharp results (see [14, 15] for a similar finding in another context). The same logic underpins the finding of Section 2.5.6 that saddlepoint MLEs are exact for exponential families; errors in the saddlepoint approximation to the log-likelihood are irrelevant if the sizes of the errors do not depend on the parameter.

Seen in this light, it is less surprising that a lower-order saddlepoint approximation, $\hat{L}^*(\theta; x)$ from Section 3.2, can give a good approximation to the MLE (see Theorem 9) despite being a poor approximation to the likelihood with $\hat{L}^*(\theta; x)/L(\theta; x) \to \infty$ as $n \to \infty$. We remark that since $\hat{L}^*(\theta; x)$ is even less computationally demanding than the usual saddlepoint approximation, it may be a useful tool for high-dimensional and computationally intensive applications, or for initialising the search for a true or saddlepoint MLE.

In the other direction, refinements of the saddlepoint approximation that improve likelihood accuracy may be less effective than anticipated when applied to MLEs. For instance, normalising the saddlepoint approximation $\hat{f}(x; \theta)$ to make it a density (as a function of x) often brings the saddlepoint density values closer to the true density (see, for instance, [4]). However, this operation is slow and is often not pursued; cf. [6]. Using the viewpoint developed in this paper, we can reframe the issue by asking "does normalising bring the saddlepoint log-gradient closer to the true log-gradient"? The general answer is far from clear. Indeed, for an exponential family such as the Poisson family, Example 24 in Appendix J, for which the saddlepoint MLE is already exact, normalising will actually make the MLE worse.

This paper has considered likelihoods and MLEs in a particularly tractable limiting framework, the standard asymptotic regime (SAR) in which the observation is a sum of n i.i.d.

terms. In particular, we have seen that the basic likelihood accuracy estimate does not directly lead to the correct MLE error estimate, which is markedly better. It would be of interest to extend these results in several directions, including non-Gaussian saddlepoint approximations, as in [20]: non-i.i.d. sums, integer-valued random variables, uniform upper bounds, large-sample-size limits and applications based on approximating tail probabilities rather than likelihoods.

Acknowledgements. The author thanks Rachel Fewster and Joey Wei Zhang, whose use of saddlepoint MLEs sparked the author's interest in this question; Rachel Fewster for helpful comments on drafts of this paper; and Godrick Oketch for suggesting the model of Example 26 in Appendix J as a useful one for explicit comparisons. Useful comments and questions from the anonymous referees helped to improve this paper.

Funding. This work was supported in part by funding from the Royal Society of New Zealand.

SUPPLEMENTARY MATERIAL

Appendices A–J (DOI: 10.1214/22-AOS2169SUPP; .pdf). Proofs, further technical details and examples.

REFERENCES

[1] BARNDORFF-NIELSEN, O. (2014). *Information and Exponential Families in Statistical Theory*. *Wiley Series in Probability and Statistics*. Wiley, Chichester. MR3221776 https://doi.org/10.1002/9781118857281

[2] BARNDORFF-NIELSEN, O. E. and KLÜPPELBERG, C. (1999). Tail exactness of multivariate saddlepoint approximations. *Scand. J. Stat.* **26** 253–264. MR1707583 https://doi.org/10.1111/1467-9469.00148

[3] BLEISTEIN, N. and HANDELSMAN, R. A. (1986). *Asymptotic Expansions of Integrals*, 2nd ed. Dover, New York. MR0863284

[4] BUTLER, R. W. (2007). *Saddlepoint Approximations with Applications*. *Cambridge Series in Statistical and Probabilistic Mathematics* **22**. Cambridge Univ. Press, Cambridge. MR2357347 https://doi.org/10.1017/CBO9780511619083

[5] DANIELS, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25** 631–650. MR0066602 https://doi.org/10.1214/aoms/1177728652

[6] DANIELS, H. E. (1982). The saddlepoint approximation for a general birth process. *J. Appl. Probab.* **19** 20–28. MR0644416 https://doi.org/10.2307/3213912

[7] DAVISON, A. C., HAUTPHENNE, S. and KRAUS, A. (2021). Parameter estimation for discretely observed linear birth-and-death processes. *Biometrics* **77** 186–196. MR4229731 https://doi.org/10.1111/biom.13282

[8] DEN HOLLANDER, F. (2000). *Large Deviations*. *Fields Institute Monographs* **14**. Amer. Math. Soc., Providence, RI. MR1739680 https://doi.org/10.1007/s00440-009-0235-5

[9] GOODMAN, J. (2022). Supplement to "Asymptotic accuracy of the saddlepoint approximation for maximum likelihood estimation." https://doi.org/10.1214/22-AOS2169SUPP

[10] JENSEN, J. L. (1988). Uniform saddlepoint approximations. *Adv. in Appl. Probab.* **20** 622–634. MR0955507 https://doi.org/10.2307/1427038

[11] JENSEN, J. L. (1995). *Saddlepoint Approximations*. *Oxford Statistical Science Series* **16**. The Clarendon Press, Oxford University Press, New York. MR1354837

[12] KOLASSA, J. E. (2006). *Series Approximation Methods in Statistics*, 3rd ed. *Lecture Notes in Statistics* **88**. Springer, New York. MR1295242 https://doi.org/10.1007/978-1-4757-4275-6

[13] LUNDE, B. Å. S., KLEPPE, T. S. and SKAUG, H. J. (2018). Saddlepoint adjusted inversion of characteristic functions. Preprint. Available at arXiv:1811.05678 [stat.CO].

[14] OGDEN, H. E. (2017). On asymptotic validity of naive inference with an approximate likelihood. *Biometrika* **104** 153–164. MR3626485 https://doi.org/10.1093/biomet/asx002

[15] OGDEN, H. E. (2018). On the error in Laplace approximations of high-dimensional integrals. Preprint. Available at arXiv:1808.06341 [math.ST].

[16] PEDELI, X., DAVISON, A. C. and FOKIANOS, K. (2015). Likelihood estimation for the INAR(p) model by saddlepoint approximation. *J. Amer. Statist. Assoc.* **110** 1229–1238. MR3420697 https://doi.org/10.1080/01621459.2014.983230

[17] REID, N. (1988). Saddlepoint methods and statistical inference. *Statist. Sci.* **3** 213–227. MR0968390

[18] RUDIN, W. (1976). *Principles of Mathematical Analysis*, 3rd ed. *International Series in Pure and Applied Mathematics*. McGraw-Hill, New York. MR0385023

[19] WONG, R. (2001). *Asymptotic Approximations of Integrals. Classics in Applied Mathematics* **34**. SIAM, Philadelphia, PA. MR1851050 https://doi.org/10.1137/1.9780898719260

[20] WOOD, A. T. A., BOOTH, J. G. and BUTLER, R. W. (1993). Saddlepoint approximations to the CDF of some statistics with nonnormal limit distributions. *J. Amer. Statist. Assoc.* **88** 680–686. MR1224395 https://doi.org/10.2307/2290351

[21] ZHANG, W., BRAVINGTON, M. V. and FEWSTER, R. M. (2019). Fast likelihood-based inference for latent count models using the saddlepoint approximation. *Biometrics* **75** 723–733. MR4012079 https://doi.org/10.1111/biom.13030