NUMERICAL SOLUTION OF AN INTEGRAL EQUATION FROM POINT PROCESS THEORY

R.S. ANDERSSEN, A.J. BADDELEY, F.R. de HOOG AND G.M. NAIR

ABSTRACT. We propose and analyze methods for the numerical solution of an integral equation which arises in statistical physics and spatial statistics. Instances of this equation include the Mean Field, Poisson-Boltzmann and Emden equations for the density of a molecular gas, and the Poisson saddlepoint approximation for the intensity of a spatial point process. Conditions are established under which the Picard iteration and the under relaxation iteration converge. Numerical validation is included.

1. Introduction. Several important techniques in statistical physics and spatial statistics require the solution of integral equations of the general form

$$(1) \qquad f(x) = g(x) \exp\left(-\int_{\mathbb{D}} c(x-y)f(y)\,\mathrm{d}y \right), \quad x \in \mathbb{D},$$

where $\mathbb{D}$ denotes a bounded domain in $\mathbb{R}^k$ on which the function g is strictly positive. The function $c \in L_2(\mathbb{D})$ is defined on the extended domain $\mathbb{E} = \{ z \in \mathbb{R}^k \mid z = x - y, \ x \in \mathbb{D}, \ y \in \mathbb{D} \}$. In addition, c is assumed to be non-negative on $\mathbb{E}$ and is symmetric, that is, $c(z) = c(-z)$. Clearly, the solution f, if it exists, is also strictly positive on $\mathbb{D}$.

For notational simplicity, we use $c * f$ as the convolution of c and f. That is,

$$(c * f)(x) = \int_{\mathbb{D}} c(x-y)f(y)\,\mathrm{d}y.$$

This is consistent with the usual Fourier convolution on defining f and c to be zero on $\mathbb{R}^k \backslash \mathbb{D}$ and $\mathbb{R}^k \backslash \mathbb{E}$, respectively. In addition, we will occasionally drop the dependent variable from equations when the meaning is clear from the context as a means of simplifying the algebra in the convergence analysis.

Instances of equation (1) include the Mean Field [9] and Poisson-Boltzmann-Emden [6, 14] approximations for the density of a molecular gas, and the Poisson-saddlepoint [2, 3] approximation for the intensity of a spatial point process. Our goal in this paper is to develop stable numerical methods for solving the integral equation (1).

The paper is organized as follows. Background about the mathematical structure of statistical physics and spatial statistics applications is surveyed in Section 2. The convergence analysis is the focus of Section 3. Sufficient conditions for convergence are derived in subsection 3.1 for the Picard iterative solution of equation (1), while, in subsection 3.2, sufficient conditions are derived for the application of an under relaxation iterative procedure. Numerical examples are given in Section 4.

Overall, it is concluded that, though conditions are established that guarantee global convergence for the Picard iteration, they are so constraining that they limit the utility of this form of iteration. As a more flexible alternative, under relaxation is proposed, and conditions for its convergence are established that are much less restrictive. These results are analogous to the known numerical performance of the Picard and under relaxation iterations for scalar problems.

2. Background. The integral equation examined here arises in the analysis of *spatial point processes* in statistical physics and in spatial statistics. A spatial point process [8] is a stochastic model for a random spatial configuration of points in $\mathbb{R}^k$. In statistical physics, the points may represent the locations of individual molecules in a gas at some instant of time, and the point process is a stochastic model of the fine-scale behavior of the gas [9]. In spatial statistics, the points may represent observed locations of trees, road accidents, bird nests, crystal impurities and so on, and the point process is a statistical model of the pattern of these locations [10, 15].

We restrict attention to the class of *pairwise-interaction Gibbs* point

process models, which is the mainstay of classical statistical mechanics of gases, as well as playing an important role in spatial statistics [**10**]. In brief, a configuration of n molecules or particles, situated at locations $x_1, \ldots, x_n$ inside a container $\mathbb{D} \subset \mathbb{R}^k$, is assumed to have probability density

$$(2) \qquad p(x_1, \ldots, x_n) = \alpha \left[\prod_{i=1}^{n} g(x_i) \prod_{i<j} h(||x_i - x_j||) \right],$$

where α is a normalizing constant, and $g : \mathbb{R}^k \to [0, \infty)$ and $h : [0, \infty) \to [0, \infty)$ are known functions specifying the model. The function g influences the propensity for points to be found at some locations more frequently than other locations, while the function h introduces stochastic dependence or interaction between the points. In statistical physics, $-\log p(x_1, \ldots, x_n)$ represents the total potential energy of the configuration; $-\log g(x)$ is an external field representing the energy required to position a point at location x, while $-\log h(||x - y||)$ is the pair potential between two points at locations x and y. In many, but not all, examples of interest, $0 \le h(\cdot) \le 1$. The number of points n in the configuration may be fixed, or may also be allowed to vary. In the latter case, the function g also influences the overall abundance of points.

The *intensity* of a point process is the expected ('average') number of points per unit volume. In the problems of interest, the intensity is a spatially-varying function $\lambda(x)$ of location x in $\mathbb{R}^k$. Given a specification of a point process, it is often difficult or impossible to determine the intensity function $\lambda(x)$ exactly.

Various methods for approximating the intensity have been proposed [**2, 3, 6, 9, 14**] each requiring us to solve an integral equation (*closure relation*) of the general form (1), where $f(x)$ is the proposed approximation to the intractable intensity $\lambda(x)$. The classical Mean Field approximation [**9, 16**] to the intensity function $\lambda(x)$ is the solution $\lambda_{\mathrm{MF}}(x)$ of the integral equation

$$\lambda_{MF}(x) = g(x) \exp \left(\int_{\mathbb{D}} \lambda_{MF}(y) \log h(x - y) \, \mathrm{d}y \right), \quad x \in \mathbb{D}.$$

In [**2, 3**], an alternative 'Poisson saddlepoint' approximation $\lambda_{\mathrm{PS}}(x)$ is

proposed which is the solution of

$$(3) \quad \lambda_{PS}(x) = g(x) \exp\left(-\int_{\mathbb{D}} \lambda_{PS}(y)[1 - h(x - y)] \, dy\right), \quad x \in \mathbb{D}.$$

The latter is also known as the Poisson-Boltzmann, Emden or generalized Poisson-Boltzmann-Emden equation [6, 14]. Both of these equations are of the form highlighted in equation (1) with $f(x) = \lambda_{MF}(x)$, $c(x) = -\log h(x)$ for the Mean Field approximation, and $f(x) = \lambda_{PS}(x)$, $c(x) = 1 - h(x)$ for the Poisson saddlepoint approximation. In either case, $0 \leq h(x) \leq 1$ implies $c(x) \geq 0$.

Our goal is to compute, using equation (1), approximations of the intensity function $\lambda(x)$ of the pairwise interaction process, for situations where the trend function $g(x)$ and the interaction function $h(x)$ are representative of statistical physics and spatial statistics applications. Even though equation (1) determines useful approximations for λ, exact solutions are not known in general. It is therefore necessary to turn to computational procedures to determine information about $f(x)$ with respect to given choices for g and c.

In a typical application of these methods in spatial statistics [4, 5], the function g would be of the form $g(x) = \exp(\beta^\top Z(x))$ where $Z(x)$ is a p-dimensional vector valued function containing "covariate" data (such as soil acidity values, terrain altitude, the Cartesian coordinate functions, or a function that is equal to 1 inside some specified region and 0 outside), and β is a p-dimensional vector of coefficients that has been estimated from data by some appropriate statistical technique. Thus, it can be assumed that g is positive and bounded over $\mathbb{D}$, but it may not be continuous. The function h would usually be more tightly prescribed, often having the form $h(x) = \exp(-\theta V(x))$, where V is one of a small class of "interaction potential" functions the properties of which are well studied. Again, θ is a parameter that has been estimated from data by some appropriate statistical technique. The goal of the application is to compute the predicted intensity $\lambda(x)$ of the point process model, approximately, by solving the integral equations above. Numerical solution of (1) would be vastly more efficient than the currently accepted practice of running very large Monte Carlo experiments and estimating the intensity by averaging the densities of the simulated point patterns.

3. Convergence analysis. In general, it is not possible to determine the solution of equation (1) analytically and it is therefore necessary to determine an approximate numerical solution computationally. Although the structure of the equation is relatively simple, it is a k-dimensional problem with $k = 2$ or $k = 3$ and can result in a large number of non-linear equations when discretized. For this reason, iterative methods are an attractive option and, because the details of the discretization method used are of secondary importance, we focus here on iterative schemes based on the continuous, rather than the discretized, problem.

In practice, the function c is often quite peaked around the origin, and somewhat resembles a delta function. Then

$$f(x) \approx g(x) \exp\left(-f(x)\,(c * 1)\,(x)\right),$$

and we can obtain an approximate solution $\widehat{f}$ by solving

$$\widehat{f}(x) = g(x) \exp\left(-\widehat{f}(x)(c * 1)(x)\right). \tag{4}$$

Note that (4) is a scalar equation for each $x \in \mathbb{D}$. It has the solution

$$\widehat{f}(x) = \frac{W\left(g(x)(c * 1)(x)\right)}{(c * 1)(x)}, \quad x \in \mathbb{D} \tag{5}$$

called the *pseudostationary solution* [2], where W is the Lambert W function [7], the inverse function of $t \mapsto t \exp(t)$. In addition to providing a useful starting approximation for iterative methods, (4) provides considerable insight about iteration schemes applied to (1). To illustrate this, consider the Picard iteration

$$\widehat{f}_{n+1}(x) = g(x) \exp\left(-\widehat{f}_n(x)(c * 1)(x)\right), \quad n = 0, 1, 2, \ldots,$$

for which a sufficient condition for convergence is $g(x)(c * 1)(x) < 1$ and a necessary condition is $W(g(x)(c * 1)(x)) < 1$. Both are violated when $g(x)(c * 1)(x)$ becomes sufficiently large and $\widehat{f}_n(x)$ then oscillates between a fixed lower and a fixed upper bound for $\widehat{f}(x)$. This suggests that the Picard iteration is over-correcting, and that under relaxation, namely,

$$\widehat{f}_{n+1}(x) = (1 - \alpha)\widehat{f}_n(x) + \alpha g(x) \exp\left(-\widehat{f}_n(x)(c * 1)(x)\right), \tag{6}$$

$$0 < \alpha \leq 1,$$

may be more appropriate. In fact, it can be shown that if $\alpha < 2/(1 + W(g(x)(c * 1)(x)))$ and the initial iterate $\widehat{f}_0(x)$ is sufficiently close to $\widehat{f}(x)$, then the under relaxation iteration (6) converges. As W is a monotonically increasing function satisfying $W(0) = 0$ and $W(z) \approx \ln(z)$ for large z, it follows that, for most choices of g and c, we can choose α sufficiently small to guarantee convergence provided our initial error is sufficiently small. Note also that, if $g(x)(c * 1)(x)$ is sufficiently small, a choice of $\alpha > 1$, which guarantees that $\alpha(1 + W(g(x)(c * 1)(x))) < 2$, will also lead to convergence.

3.1. Convergence analysis for the Picard iteration. The integral equation (1) can be rewritten as

$$(7) \qquad f = \mathcal{A}(f), \qquad \mathcal{A}(f) = g \exp(-c * f),$$

from which it follows that, with respect to the solution f,

$$(8) \qquad \mathcal{A}(f + q) = f \exp(-c * q).$$

For equation (1), the Picard iteration takes the form

$$(9) \qquad f_{n+1} = \mathcal{A}(f_n), \quad n = 0, 1, 2, \dots.$$

Subtracting equation (9) from equation (7) gives, with $\epsilon_n = f - f_n$,

$$(10) \qquad \epsilon_{n+1} = \mathcal{A}(f) - \mathcal{A}(f_n),$$

$$= -g\left(\int_0^1 \exp(-c * ((1 - s)f + sf_n)\, \mathrm{d}s) \right)(c * \epsilon_n).$$

It follows that if, for some n, ϵ_n is either strictly positive or strictly negative, then, for $m \geq n$, the signs of ϵ_m and ϵ_{m+1} will oscillate. This implies that the corresponding Picard iterates f_n generate upper and lower bounds for f. In addition, since f and f_n, for $n > 0$, are non-negative, it follows from (10) that

$$|\epsilon_{n+1}| \leq gc * |\epsilon_n|.$$

We can generalize this by introducing a weighted error $\widehat{\epsilon}_n = \epsilon_n/u$, where u is an arbitrary strictly positive function. Consequently,

$$|\widehat{\epsilon}_{n+1}| \leq \frac{g}{u}c * (|\widehat{\epsilon}_n|u),$$

and hence

(11) $$\|\widehat{\epsilon}_{n+1}\|_\infty \leq \|\frac{g}{u}(c * u)\|_\infty \|\widehat{\epsilon}_n\|_\infty.$$

This provides a sufficient condition for the global convergence of the Picard iteration, namely:

Theorem 1. *A sufficient condition for the convergence of the Picard iteration* (9) *is that* $\|(g/u)(c*u)\|_\infty < 1$, *where* u *is any strictly positive function defined on* $\mathbb{D}$.

Proof. From (11), $\|\widehat{\epsilon}_{n+1}\|_\infty < \|\widehat{\epsilon}_n\|_\infty$ and, hence $\|\widehat{\epsilon}_n\|_\infty \to 0$ as $n \to 0$, independently of the choice of the starting solution f_0. $\qquad\square$

Note that, when $u = 1$, Theorem 1 yields exactly the same sufficient condition derived in the previous section for the Picard iteration applied to (10).

Although Theorem 1 provides a sufficient condition for convergence, the Picard iteration will converge locally under much weaker conditions. To investigate this, (8) can be used to obtain the following alternative formulation of (9)

$$f_{n+1} = g \exp(-c * (f - (f - f_n))) = f \exp(c * \epsilon_n),$$

from which it follows that

(12) $$\epsilon_{n+1} = \mathcal{A}(f) - \mathcal{A}(f_n) = -f\left(\int_0^1 \exp(sc * \epsilon_n)\, ds\right)(c * \epsilon_n)$$

$$= -(1 + E_n)f(c * \epsilon_n),$$

where

(13) $$E_n = (c * \epsilon_n)\int_0^1 s \int_0^1 \exp(stc * \epsilon_n)\, dt\, ds.$$

For an arbitrary positive function u, it follows that

$$\widehat{\epsilon}_{n+1} = -(1 + E_n)\frac{f}{u}(c * u\widehat{\epsilon}_n),$$

and, hence, that

$$\|\widehat{\epsilon}_{n+1}\|_\infty \leq \|(1 + E_n)\frac{f}{u}(c * u)\|_\infty \|\widehat{\epsilon}_n)\|_\infty.$$

Thus, if $\|(1 + E_n)f/u(c * u)\|_\infty < 1$, then $\|\widehat{\epsilon}_n\|_\infty \to 0$ as $n \to 0$ and the Picard iteration converges. Consequently, for a starting solution f_0 sufficiently close to f, $\epsilon_n \sim 0$ for suitably large n, and, hence, from (13), it follows that $E_n \approx 1/2c * \epsilon_n$ is arbitrarily small and thus the Picard iteration converges locally if $\|(f/u)(c * u)\|_\infty < 1$. This will be explored in more detail in the analysis of the under relaxation iteration.

3.2. Convergence analysis for the under relaxation iteration.
The under relaxation iteration for equation (1) takes the form

$$(14) \qquad f_n = (1 - \alpha)f_{n-1} + \alpha\mathcal{A}(f_{n-1}),$$
$$\mathcal{A}(f) = g\exp(-c * f), \quad 0 < \alpha < 1,$$

where the starting solution f_0 is given. Assuming that the solution f of equation (1) exists, an analogous derivation to that used in the construction of (12) yields

$$(15) \qquad \epsilon_{n+1} = (1 - \alpha)\epsilon_n - \alpha f(c * \epsilon_n) - \alpha E_n f(c * \epsilon_n),$$

where E_n is given in (13).

For the analysis of the under relaxation iteration, it is convenient to work with a weighted error; namely, $\widetilde{\epsilon}_n = \epsilon_n/\sqrt{f}$. The choice of the weighting will become clear in the sequel. Equation (15) then takes the form

$$(16) \qquad \widetilde{\epsilon}_{n+1} = \left[(1 - \alpha)\mathcal{I} - \alpha\widetilde{\mathcal{B}} - \alpha E_n\widetilde{\mathcal{B}}\right]\widetilde{\epsilon}_n,$$

where $\mathcal{I} : L_2(\mathbb{D}) \mapsto L_2(\mathbb{D})$ is the identity operator and, given that $c \in L_2(\mathbb{D})$, the integral operator $\widetilde{\mathcal{B}} : L_2(\mathbb{D}) \mapsto L_2(\mathbb{D})$ defined by

$$(\widetilde{\mathcal{B}}q)(x) := \int_{\mathbb{D}} \sqrt{f(x)}c(x - y)\sqrt{f(y)}q(y)\,\mathrm{d}y, \quad x \in \mathbb{D}$$

is compact. Taking norms on both sides of (16) yields

$$(17) \qquad \|\widetilde{\epsilon}_{n+1}\|_2 \leq \left(\|(1 - \alpha)\mathcal{I} - \alpha\widetilde{\mathcal{B}}\|_2 + \alpha\|E_n\|_\infty\|\widetilde{\mathcal{B}}\|_2\right)\|\widetilde{\epsilon}_n\|_2,$$

and, hence, a sufficient condition for convergence of the over-relaxation iteration is

$$(18) \qquad \|(1 - \alpha)\mathcal{I} - \alpha\widetilde{\mathcal{B}}\|_2 + \alpha\|E_n\|_\infty\|\widetilde{\mathcal{B}}\|_2 < 1.$$

Because the kernel $\sqrt{f(x)}c(x-y)\sqrt{f(y)}$ is symmetric with respect to x and y, the operator $\widetilde{\mathcal{B}}$ is self-adjoint. As a consequence, $\|\widetilde{\mathcal{B}}\|_2$ and $\|(1-\alpha)\mathcal{I}-\alpha\widetilde{\mathcal{B}}\|_2$ can be determined from the greatest lower and lowest upper bounds of the eigenvalues of $\widetilde{\mathcal{B}}$. This is the reason for choosing the weighted error $\widetilde{\epsilon}_n$. Specifically, let $\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots$ be the eigenvalues of $\widetilde{\mathcal{B}}$ in order of decreasing absolute value, $\widetilde{\beta} = \max_m\{\widetilde{\lambda}_m\}$ and $\widetilde{\mu} = -\inf_m\{\widetilde{\lambda}_m\}$. From the Perron-Frobenius theorem, it follows that $\widetilde{\beta} \geq \widetilde{\mu}$ since the kernel of $\widetilde{\mathcal{B}}$ is non-negative [11]. Then

$$(19) \qquad \|\widetilde{\mathcal{B}}\|_2 = \widetilde{\beta}$$

and

$$(20) \qquad \|(1-\alpha)\mathcal{I}-\alpha\widetilde{\mathcal{B}}\|_2 = \max\left\{1-\alpha(1-\widetilde{\mu}), \alpha(1+\widetilde{\beta})-1\right\}.$$

In addition, it follows from (13) and the inequality $\|\epsilon\|_1 \leq \sqrt{\|f\|_1}\|\widetilde{\epsilon}\|_2$ that

$$(21) \qquad \|E_n\|_\infty \leq \frac{1}{2}\|c\|_\infty\|\sqrt{\|f\|_1}\|\widetilde{\epsilon}_n\|_2 \exp(\|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_n\|_2).$$

The inequalities (19), (20) and (21), when applied to (18), yield sufficient conditions for the convergence of the under relaxation iteration. Specifically:

Theorem 2. *Sufficient conditions for the convergence of the over-relaxation iteration* (14) *are:*

$$(22) \qquad \widetilde{\mu} < 1 - \frac{\widetilde{\beta}}{2}\|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_0\|_2 \exp(\|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_0\|_2)$$

and

$$(23) \qquad \alpha < \frac{2}{1 + \widetilde{\beta}\left(1 + \frac{1}{2}\|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_0\|_2 \exp(\|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_0\|_2)\right)}.$$

Proof. It follows by induction from (17), (19), (20), (21), (22) and (23) that $\|\widehat{\epsilon}_{n+1}\|_2 < \|\widehat{\epsilon}_n\|_2$, and hence $\|\widehat{\epsilon}_n\|_2 \to 0$ as $n \to 0$. $\square$

If we let

$$(24) \qquad m = \|c\|_\infty\sqrt{\|f\|_1}\|\widetilde{\epsilon}_0\|_2,$$

the inequalities (22) and (23) can be rewritten as

$$(25) \qquad \widetilde{\mu} + \frac{\widetilde{\beta} m \exp(m)}{2} \leq 1$$

and

$$(26) \qquad \alpha < \frac{2}{1 + \widetilde{\beta}\left(1 + (m \exp(m))/2\right)}.$$

For inequalities (25) and (26) to hold, it is necessary that $\widetilde{\mu} < 1$ and $\alpha < 2/(1 + \widetilde{\beta})$. In fact, it can be shown that these are also necessary conditions for the under relaxation iteration to converge. Furthermore, it follows from Theorem 2 that these are also sufficient conditions for local convergence – that is, convergence provided that the initial iterate f_0 is sufficiently close to the solution f. As the inequality $\alpha < 2/(1 + \widetilde{\beta})$ can always be satisfied by taking α sufficiently small, the hard constraint on local convergence for a relaxation of the Picard iteration is $\widetilde{\mu} < 1$. However, note that, for the Picard iteration ($\alpha = 1$), local convergence requires $\widetilde{\beta} < 1$, a substantially stronger constraint than $\widetilde{\mu} < 1$.

Clearly, the greatest lower and lowest upper bounds of the eigenvalues of $\widetilde{\mathcal{B}}$ are important quantities for convergence of the over-relaxation iteration and we briefly examine these. A bound for the lowest upper bound $\widetilde{\beta}$ is relatively straightforward and follows from

$$(27) \qquad \widetilde{\beta} = \|\widetilde{\mathcal{B}}\|_2 \leq \|\widetilde{\mathcal{B}}\|_\infty = \|\sqrt{f}(c * \sqrt{f})\|_\infty \leq \|f\|_\infty \|c\|_1.$$

Estimates for lower bounds are somewhat more difficult. Note, however, that when $\widetilde{\mathcal{B}}$ is non-negative definite $\widetilde{\mu} < 0$ and over-relaxation will converge locally provided α is sufficiently small.

If $\widetilde{\mu} > 0$, bounds can be obtained by considering convolution equations whose spectral properties have been widely studied. Let $\mathcal{B} : L_2(\mathbb{D}) \mapsto L_2(\mathbb{D})$ be the integral operator defined by $\mathcal{B}q := c * q$, $\lambda_1, \lambda_2, \ldots$ denote the eigenvalues of $\mathcal{B}$ in order of decreasing absolute value, $\beta = \max_m\{\lambda_m\}$ and $\mu = -\inf_m\{\lambda_m\}$. In addition, let $C(\zeta) = \int_{\mathbb{E}} c(x) \exp(i\zeta \cdot x)\, dx$ be the Fourier transform of c. Then, it follows from Widom [19] that

$$\beta \leq \sup_\zeta \{C(\zeta)\} = \|c\|_1,$$

and

$$(28) \qquad \mu \leq - \inf_{\zeta} \{C(\zeta)\}.$$

This last bound is useful when the Fourier transform can be evaluated as is the case, for example, for a Strauss point process. Bounds for $\widetilde{\beta}$ and $\widetilde{\mu}$ now follow on noting that $\widetilde{\beta} \leq \|f\|_{\infty}\beta$ and $\widetilde{\mu} \leq \|f\|_{\infty}\mu$.

4. Numerical performance. Two simple examples of the under relaxation iteration in $\mathbb{R}^2$ have been analyzed in [**2**]. A comparison of the intensity determined by the Poisson saddlepoint approximation and by simulation can be found in [**2, 3**]. In this section, we give an example which is somewhat more challenging and realistic.

4.1. Data. Figure 1 shows part of a survey of gold deposits in the Murchison region of Western Australia [**18**]. Crosses show the gold deposits; solid lines are geological faults; grey shading is greenstone outcrop.
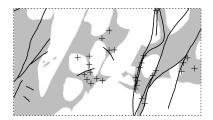

FIGURE 1. Data from a geological survey of the Murchison region.

Typically the gold deposits would be modeled as a Poisson point process [**1, 13**] with intensity function of the form

$$(29) \qquad g(x) = \exp(\beta_0 + \beta_1 I(x) + \beta_2\, d(x)),$$

where $\beta_0, \beta_1, \beta_2$ are parameters to be estimated, $d(x)$ is the distance from x to the nearest geological fault, and $I(x)$ is the indicator that equals 1 if x falls inside the greenstone outcrop, and equals 0 otherwise. Typical values of the fitted parameters for this dataset are $\beta_0 =$

$\log 0.005$, $\beta_1 = \log 2$ and $\beta_2 = -0.1$. The function with these parameter values is shown in Figure 2.

FIGURE 2. Poisson intensity function, or Gibbs forcing function, $g(x)$ for the Murchison data example.

4.2. Application to Strauss process. A possible alternative model for the above data is a Strauss point process [**12, 17**], a particular case of the pairwise interaction Gibbs process (2) with interaction term

$$h(x) = \left\{ \begin{array}{ll} \gamma & \text{if } \|x\|_2 \leq R, \\ 1 & \text{otherwise.} \end{array} \right.$$

We take the first order term ('forcing function') $g(x)$ in (2) to be of the form (29). The intensity function of the Strauss process is not known analytically as a function of the model parameters, so we shall use the Poisson-saddlepoint approximation, the solution of (3), which is of the form (1) with $c(x) = 1 - h(x)$.

Typical parameter values for the Strauss interaction in this example are radius $R = 10$ km and strength $\gamma = 0.5$. This gives $\|c\|_\infty = (1 - 0.5) = 0.5$. Since $c \geq 0$, $f * c \geq 0$ and so (1) gives $f(x) \leq g(x)$ for all x. Repeating this argument gives inequalities

(30) $g_1(x) \leq f(x) \leq g_2(x),$

where $g_1(x) = g(x) \exp(-(g * c)(x))$ and $g_2(x) = g(x) \exp(-(g_1 * c)(x))$. The initial state $f_0(x)$ for the iterations will be the *pseudostationary solution* (5).

Figure 3 shows transects of these functions along a diagonal line from bottom left to top right in the survey region depicted in Figure 1. Dotted lines show the forcing function g, and thick solid lines show the pseudostationary approximation $f_0(x)$. The grey shading is delimited
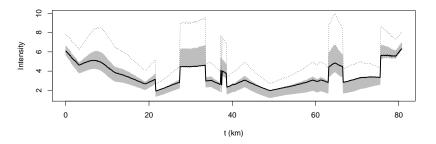
FIGURE 3. Diagonal transects (bottom left to top right) of functions in spatial domain of Figure 1. Thick solid line: pseudostationary approximation $f_0(x)$. Grey shading: interval between upper and lower bounds $g_1(x), g_2(x)$ for the true solution f. Dotted lines: forcing function $g(x)$.

by the upper and lower bounds $g_1(x), g_2(x)$ for the true solution f. (Note that it is not necessarily true that the pseudostationary approximation lies between the limits g_1, g_2 as it happens to do in this example).

The inequalities (30) give

$$(31) \qquad 0.007503 = \|g_1\|_\infty \le \|f\|_\infty \le \|g_2\|_\infty = 0.00854,$$

where the numerical values were computed using a pixel approximation. This gives an upper bound on $\widetilde{\beta}$,

$$(32) \qquad \widetilde{\beta} \le \|f\|_\infty \beta \le \|f\|_\infty \|c\|_\infty \le 0.00854 \times 0.5 = 0.00427.$$

For $\widetilde{\mu}$, we have

$$\widetilde{\mu} \le \|f\|_\infty \mu \le \|f\|_\infty (-\inf_\zeta \mathcal{C}(\zeta)),$$

where $\mathcal{C}$ is the Fourier transform of c. We have $\mathcal{C}(\zeta) = (1 - \gamma)\mathcal{F}(\zeta)$, where $\mathcal{F}$ is the Fourier transform of the indicator function of a disc of radius R,

$$\mathcal{F}(\zeta) = 2\pi R^2 \frac{J_1(R\|\zeta\|)}{R\|\zeta\|},$$

where J_1 is the Bessel function of the first kind, of order 1. Noting that $\inf_{t \ge 0}(J_1(t)/t) \approx -0.07$, we have

$$-\inf_\zeta \mathcal{C}(\zeta) \approx 0.14\pi(1 - \gamma)R^2 = 21.99,$$

yielding $\widetilde{\mu} \leq 0.1878$.

Similarly, the sandwich inequalities for f give bounds on $\sqrt{\|f\|_1}$:

$$2.642 = \sqrt{\|g_1\|_1} \leq \sqrt{\|f\|_1} \leq \sqrt{\|g_2\|_1} = 3.213.$$

Now $\widetilde{\epsilon}_0 = \epsilon_0/\sqrt{f}$ where $\epsilon_0 = f - f_0$ is the discrepancy between the initial starting function and the correct answer f. We need an upper bound on $\|\widetilde{\epsilon}_0\|_2$. Using the sandwich inequalities for f, we get $|\epsilon_0(x)| \leq U(x)$ and $|\widetilde{\epsilon}_0(x)| \leq U(x)/\sqrt{g_1(x)}$ where $U(x) = \max\{|f_0(x) - g_1(x)|, |f_0(x) - g_2(x)|\}$. This yields $\|\widetilde{\epsilon}_0\|_2 \leq 1.01$. The results above yield an upper bound for m in (24) as $m \leq 1.623$, and this gives an upper bound for the left hand side of the condition in (25) as

$$0.1878 + 0.00427/2 \times 1.623 \exp(1.623) = 0.2054.$$

Hence, condition (25) is satisfied.

In condition (26), if we replace the denominator of the right hand side of (26) by an upper bound for the denominator, we obtain a lower bound for the right hand side of (26). Any value of α less than this bound is guaranteed to satisfy (26). The bound, using (32), is

$$\alpha_{\max} = \frac{2}{1 + 0.00427(1 + 1.623 \exp(1.623)/2)} = 1.957.$$

Consequently, *any* value of α in $[0, 1]$ would suffice.

4.3. Picard iteration. We shall use under-relaxation with $\alpha = 0.9$. Figure 4 shows the L_∞ discrepancies between successive iterations, i.e., $\|\lambda_{i+1}(\cdot) - \lambda_i(\cdot)\|_\infty$ where $\lambda_i(\cdot)$ is the i-th iterate. Convergence is clearly occurring at an exponential rate.

Figure 5 shows the initial pseudostationary approximation and the result of 30 successive iterates of the under-relaxation operator. Figure 6 shows diagonal transects of the initial state and the 30th iteration.
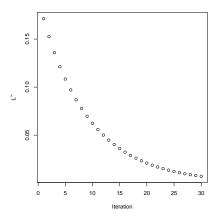
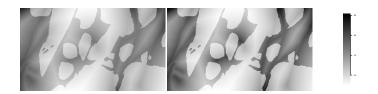FIGURE 4. L_∞ discrepancies between successive iterations.

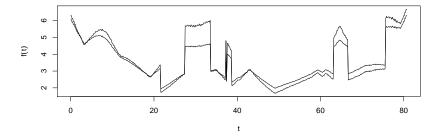FIGURE 5. Initial state (*Left*) and state after 30 iterations (*Right*).

FIGURE 6. Diagonal transects of the initial state and the 30th iteration.

REFERENCES

1. A. Baddeley, M. Berman, N.I. Fisher, A. Hardegen, R.K. Milne, D. Schuhmacher and R. Turner, *Spatial logistic regression and change-of-support for Poisson point processes*, Electr. J. Stat. **4** (2010), 1151–1201.

2. A. Baddeley and G. Nair, *Approximating the moments of a spatial point process*, Stat. **1** (2012), 18–30.

3. _____ , *Fast approximation of the intensity of Gibbs point processes*, Electr. J. Stat. **6** (2012), 1155–1169.

4. A. Baddeley and R. Turner, *Spatstat*: *An* R *package for analyzing spatial point patterns*, J. Stat. Software **12** (2005), 1–42.

5. _____ , *Modelling spatial point patterns in* R, in *Case studies in spatial point pattern modelling*, Lect. Notes Stat. **185**, Springer-Verlag, New York, 2006.

6. F. Bavaud, *Equilibrium properties of the Vlasov functional*: *The generalized Poisson-Boltzmann-Emden equation*, Rev. Modern Physics **63** (1991), 129–149.

7. R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey and D.E. Knuth, *On the Lambert W function*, Comp. Math. **5** (1996), 325–359.

8. D.J. Daley and D. Vere-Jones, *An Introduction To The Theory Of Point Processes*, Volume I: *Elementary theory and methods*, Springer, New York, 2003.

9. R.P. Feynman, *Statistical mechanics. A set of lectures*, Benjamin, Reading, MA, 1972.

10. J. Illian, A. Penttinen, H. Stoyan and D. Stoyan, *Statistical analysis and modelling of spatial point patterns*, John Wiley and Sons, Chichester, 2008.

11. R. Jentzsch, *Integral equations with positive nucleus*, J. reine angew. Math. **141** (1912) 235–244.

12. F.P. Kelly and B.D. Ripley, *A note on Strauss's model for clustering*, Biometrika **63** (1976), 357–360.

13. J.F.C. Kingman, *Poisson Processes*, Oxford University Press, New York, 1993.

14. J.G. Kirkwood and E. Monroe, *Statistical mechanics of fusion*, J. Chem. Phys. **9** (1941), 514–526.

15. J. Møller and R.P. Waagepetersen, *Statistical inference and simulation for spatial point processes*, Chapman and Hall/CRC, Boca Raton, 2004.

16. D. Ruelle, *Statistical mechanics: Rigorous results*, W.A. Benjamin, Reading, MA, 1969.

17. D.J. Strauss, *A model for clustering*, Biometrika **63** (1975), 467–475.

18. K.P. Watkins and A.H. Hickman, *Geological evolution and mineralization of the Murchison Province, Western Australia*, Bulletin 137, Geological Survey of Western Australia, 1990. Published by Department of Mines, Western Australia, 1990. Available online from Department of Industry and Resources, State Government of Western Australia, `www.doir.wa.gov.au`.

19. H, Widom, *Extreme eigenvalues of N-dimensional convolution operators*, Trans. Amer. Math. Soc. **106** (1963), 391–414.

CSIRO Digital Productivity, GPO Box 664, Canberra, ACT 2601
Email address: **Bob.Anderssen@csiro.au**

Centre for Exploration Targeting (M006), University of Western Australia, 35 Stirling Hwy., Crawley WA 6009
Email address: **Adrian.Baddeley@uwa.edu.au**

CSIRO Digital Productivity, GPO Box 664, Canberra, ACT 2601
Email address: **Frank.deHoog@csiro.au**

School of Mathematics and Statistics (M019), University of Western Australia, 35 Stirling Hwy., Crawley WA 6009
Email address: **Gopalan.Nair@uwa.edu.au**