PAPER • OPEN ACCESS

Using non-homogeneous point process statistics to find multi-species event clusters in an implanted semiconductor

View the article online for updates and enhancements.

Journal of Physics Communications

PAPER

Using non-homogeneous point process statistics to find multi-species event clusters in an implanted semiconductor

K Stockbridge[1] , S Chick[1] , E Crane[2] , A Fisher[2] and B N Murdin[1,3]

1   Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, United Kingdom
2   London Centre for Nanotechnology, UCL, London, WC1H 0AH, United Kingdom
3   Author to whom any correspondence should be addressed.

E-mail: b.murdin@surrey.ac.uk

Abstract

The Poisson distribution of event-to-ith-nearest-event radial distances is well known for homogeneous processes that do not depend on location or time. Here we investigate the case of a non-homogeneous point process where the event probability (and hence the neighbour configuration) depends on location within the event space. The particular non-homogeneous scenario of interest to us is ion implantation into a semiconductor for the purposes of studying interactions between the implanted impurities. We calculate the probability of a simple cluster based on nearest neighbour distances, and specialise to a particular two-species cluster of interest for qubit gates. We show that if the two species are implanted at different depths there is a maximum in the cluster probability and an optimum density profile.

Introduction

Individual interacting impurity atoms can be important for donor qubit gates, such as that proposed by Stoneham *et al* [1], while an important class of theoretical physics problems is produced by the Hubbard model, which relies on hopping and magnetic interactions between neighbours in chains [2]. In the case of donor impurities in a semiconductor, deterministic placement using scanning probe tips has improved greatly in recent years, but is currently limited to a small number of species of impurity (principally phosphorus and arsenic [3] in silicon [4, 5] and germanium [6], and Mn in GaAs [7]). Ion implantation methods can also be used to create impurity layers in semiconductors with merits including flexibility with regards to the numerous available implantable species and far faster device fabrication times which are less costly and more easily scalable. These merits clearly come at the cost of much less precision. Given the stochasticity of the donor placement it is important to look at the effects of the implant distribution on the neighbour-neighbour distances, and hence the probability of observable interactions. Contemporary work in this area [8] has focused on analytically understanding the interactions between donors, the dependence these interactions have on donor spacing and using the results of homogeneous Poisson point process statistics, optimising for these interactions. Here we generalise the statistics to include inhomogeneity of the impurities and optimise the event density profiles for a comparable event cluster definition. We show that in the case of a Gaussian distribution of events (which is a good approximation of the distribution of impurities after ion implantation) an analytic solution for the non-homogeneous nearest neighbour distribution exists. We also show that the numerical optimisation involved can be accelerated by introducing an appropriate heuristic.

Many physical problems involving stochastic probability have been studied which make use of point process statistics. They have been used to model distributions of events ranging from plants in a field [9], the locations of cellular network base stations[10] to the distribution of astronomical bodies[11]. The ability to model a distribution of points in an event space and be able to quantify irregularities such as clustering of points helps provide an insight into correlations between events and the consequences resulting from such a distribution. Using the well understood construct that is the Poisson point process [12, 13], the nearest neighbour

distribution of an impurity species has been used to model optical properties of donors in silicon[14] due to nearest neighbour interactions and also to calculate the probability of finding large clusters of donors[15] in homogeneously doped bulk semiconductors.

It is useful to discuss clusters comprising two different species so that we may expect to detect the effect that excitation of one of the species has on its interaction with the other. Species specific control/detection might make use of optical or electronic resonances. It seems reasonable to expect that if density profiles of species A and B are implanted at a different depth then we can control the distributions for the $A \rightarrow B$ separation, and by controlling the peak densities of each profile we might control the $A \rightarrow A$ and $B \rightarrow B$ separations. It is clear that if the homospecies separations (i.e. $A \rightarrow A$ or $B \rightarrow B$) are small compared with the width of the implant profile then the likelihood of an $A - B$ interaction will decrease relative to homospecies interactions and signal will be lost. This suggests low density sheets are ideal for inter-species interactions. Conversely, if the densities are very low then the $A \rightarrow B$ separation will be controlled by the density rather than the separation of the sheets. There is clearly an optimum to be found. The statistics developed in this work, though applied to the problem of interacting donor impurities for quantum technology, has been presented so that readers interested in structure between multiple species of events described by non-homogeneous density distributions may easily apply the ideas contained to their work.

In this work we analyse the distribution of nearest neighbour (NN) distances existing in a two-species point process where the density of those species are non-homogeneous in depth. When the event density varies as a Gaussian profile in one spatial dimension a solution to the distribution of $i$thnearest neighbours is shown to be analytically solvable. We have extended the calculation to the probability of finding particular multi-species cluster configurations and show that optimising the density profiles in favour of such clusters is a numerically feasible task which may be improved by an approximation. The optimised probabilities found with the aid of such an approximation were checked against those using the analytical solution which was also supported by simulation of the deemed optimal doping parameters.

Non-homogeneous Poisson point process

In this section we give some definitions of symbols useful later, and show the relationship between non-homogeneous and homogeneous Poisson distributions in order to indicate the method for investigating neighbour-distances.

If the distribution of events is non-homogeneous, the density $n(\mathbf{x})$ and the expected number of events $\delta N = n(\mathbf{x})\delta\mathbf{x}$ in the infinitesimal volume $\delta\mathbf{x}$ varies with the location, $\mathbf{x}$. If δN is small then the probability of an event in $\delta\mathbf{x}$ is equal to δN. We may now find the probability that there are no events within a larger volume V by dividing it up into elemental volumes. The expected number of events in V is

$$N(V) = \sum_{\mathbf{x}_i \in V} \delta N_i = \int_{\mathbf{x} \in V} n(\mathbf{x}) d\mathbf{x}. \tag{1}$$

Assuming $n(\mathbf{x})$ is well-behaved, we may choose the size of the $i$thelement ($\delta\mathbf{x}_i$) so that the product $n(\mathbf{x}_i)\delta\mathbf{x}_i$ is a constant. The probability of an event within $\delta\mathbf{x}_i$ is then the same for every element, and it follows (see appendix A) that the probability of m events enclosed in the larger volume V is given by the probability mass function for a Poisson distribution

$$P_P\{m, N(V)\} = N(V)^m \exp(-N(V))/m!. \tag{2}$$

It is tempting to try to use the probability $n(\mathbf{x})\delta\mathbf{x}$ (of an event in the elemental volume $\delta\mathbf{x}$ around $\mathbf{x}$) to define a probability density function (PDF). This is best avoided because the probability of an event in a larger volume is not given by the integral over that volume $\int_{\mathbf{x} \in V} n(\mathbf{x}) d\mathbf{x} \neq P_P\{1, N\}$ from comparison of equations (1) and (2) (unless V and N are very small, obviously). This is due to the fact that the possibility of two or more events is non-negligible for a large volume. Later, we shall be concerned with both questions of counting types of events in some volume where an integration like equation (1) is needed, and of the probability of occupation by specific numbers of events (0, 1, 2 or more events etc) in some volume where equation (2) is needed.

Previously, clusters of impurities with homogeneous density have been discussed in terms of the distribution of neighbour-neighbour distance [15]. In order to put our discussion into this context we give the non-homogeneous case, which follows immediately from equation (2). We define the probability $p_{\mathbf{x} \rightarrow A_i}(r)\delta r$ that a point in 3D Euclidean space, $\mathbf{x} = (x, y, z)$ has its $i$thnearest event of species A at a radial distance between $r \rightarrow r + \delta r$. We shall refer to $p_{\mathbf{x} \rightarrow A_i}(r)$ as the nearest neighbour probability density function (NNPDF). In previous literature the NNPDF is the precursor to what is referred to as the 'void nearest neighbour distribution function' [12] which is simply the cumulative distribution of the NNPDF as defined here. The term 'void' is used since there is no event specified at the point $\mathbf{x}$ whose neighbour is being found. To calculate the NNPDF,

consider the sphere $V_{\text{sphere}}(r; \mathbf{x})$ centred on $\mathbf{x}$ of radius r, and the infinitesimal shell of thickness δr around it. The probability of finding the first nearest event within the shell is equal to the product of the probabilities of having no events within the sphere and one within the shell (since these two conditions are independent): $p_{\mathbf{x} \to A_i}(r) \delta r = P_P\{0, N_A(V_{\text{sphere}}(r; \mathbf{x}))\} \delta N_A$, where the probability of an event in the shell is $\delta N_A = \delta r \frac{d}{dr} N_A(V_{\text{sphere}}(r; \mathbf{x}))$. Hence the NNPDF for the ith nearest neighbour is

$$p_{\mathbf{x} \to A_i}(r) = P_P\{i - 1, N_A(V_{\text{sphere}}(r; \mathbf{x}))\} \frac{d}{dr} N_A(V_{\text{sphere}}(r; \mathbf{x})). \tag{3}$$

The distribution around a void can be extended to the distribution for the neighbours around an event by taking into account the density of events in the infinitesimal volume $\delta \mathbf{x}$ at $\mathbf{x}$.

We can recover the homogeneous results in 3D bulk doped or perfect 2D delta-doped layers in semiconductors, etc. For example, in 3D we replace $n_A(\mathbf{x}) \to n_A^{3D}$ and so $N_A(V_{\text{sphere}}(r; \mathbf{x})) \to \frac{4}{3} \pi r^3 n_A^{3D}$, and $\frac{d}{dr} N_A \to 4\pi r^2 n_A^{3D}$. Hence all terms in equations (2) and (3) are independent of $\mathbf{x}$ and we obtain the familiar 3D homogeneous neighbour-neighbour distributions, $p_{\mathbf{x} \to A_1}^{3D}(r) = 4\pi r^2 n_A^{3D} \exp\left(-\frac{4}{3} \pi r^3 n_A^{3D}\right)$. Similarly for 2D $p_{\mathbf{x} \to A_1}^{2D}(r) = 2\pi r n_A^{2D} \exp\left(-\pi r^2 n_A^{2D}\right)$.

Poisson point process with expectation varying in 1D

In this work we are particularly concerned with impurities that have been implanted from the surface. We therefore specialize to the case with inhomogeneity of event density in only one dimension, specifically when the density n_A is a function of z only. This is the case with broad-area ion implants which produce a finite spread of penetration perpendicular to the surface and are homogeneous in the plane. The expected number of events per unit area in $z \to z + \delta z$ is $n_A(z) \delta z$ and the total areal density (equivalent to implant dose) is $n_A^{2D} = \int n_A(z) dz$. To find the expected number of events in the sphere of radius r, one integrates over the thin discs perpendicular to z (which have constant density). For a sphere $V_{\text{sphere}}(r; z)$ of radius r centred on $\mathbf{x} = (0, 0, z)$ equation (1) becomes,

$$N_A(V_{\text{sphere}}(r; z)) = \int_{z' \in V_{\text{sphere}}(r;z)} n_A(z') \pi [r^2 - (z' - z)^2] dz' \tag{4}$$

where the limits of the integral are such that the area in the square brackets is positive, i.e. from $z - r$ to $z + r$. Differentiating under the integral sign,

$$\frac{d}{dr} N_A(V_{\text{sphere}}(r; z)) = 2\pi r \int_{z' \in V_{\text{sphere}}(r;z)} n_A(z') dz' \tag{5}$$

for use in equation (3).

Neighbour separations with Gaussian density profiles

In the case of Gaussian density profiles which are reasonable approximations for typical mono-energetic implants, $n_A(z) = n_A^{2D} \pi^{-1/2} d_A^{-1} \exp(-z^2/d_A^2)$, where $d_A/\sqrt{2}$ is the r.m.s. thickness of the density profile. When substituted into (3) for $i = 1$ this gives the analytical solution for the probability of a first nearest event within $r \to r + \delta r$ measured from a starting point at depth z and hence the NNPDF:

$$p_{z \to A_i}(r) = n_A^{2D} \pi S_\zeta r \exp \frac{n_A^{2D} \pi^{1/2}}{4} [-2d_A^2(\zeta_+ e^{-\xi_-^2} + \zeta_- e^{-\xi_+^2}) + S_\zeta \sqrt{\pi}(d_A^2 - 2r^2 + 2z^2)] \tag{6}$$

where

$$S_\zeta = \text{erf}(\zeta_+) + \text{erf}(\zeta_-)$$

$$\zeta_\pm = \frac{r \pm z}{d_A}.$$

Having found the NNPDF we can further give the number of A atoms (per unit area) between $z \to z + \delta z$ with an ith nearest neighbour A within the range $r \to r + \delta r$, which is

$$n_{A \to A_i}(r, z) \delta r \delta z = p_{z \to A_i}(r) \delta r \times n_A(z) \delta z \tag{7}$$

since the density of A's at z and the probability density for the surrounding A's are independent. This is easily generalized to a multiple species situation: the density of A's (per unit area) at z with an ith nearest species B neighbour at r is

$$n_{A \to B_i}(r, z) = p_{z \to B_i}(r) n_A(z). \tag{8}$$

While the NNPDF $p_{\mathbf{x} \to X_i}(r)$ is a 1D function, $n_{A \to X_i}(r, z)$ is a two-dimensional surface; the ith nearest neighbour density surface (NNDS).

Figure 1. Nearest Neighbour Density Functions (NNDFs) describing the distribution of A's ith nearest B neighbours. The impurity density profiles for both A and B species are Gaussian (lower panel). In the upper panel, the colour scale gives the normalized density $n_{A \to B_i}(r, z)/n_A^{2D}$, such that the number of A's within the depth range $z \to z + \delta z$ with an ith nearest B at a radial distance $r \to r + \delta r$ is $n_{A \to B_i}(r, z)\delta r \delta z$. The centre of mass of the NNDF for $i = 1$ has a z position close to the peak of the A density and an r position defined by the depth separation between the A and B layers. As expected, the 20th nearest neighbours are further away than the 1st nearest. Vertical slices may be used to find the expected distance to a B for A's of specific depth, and integrating horizontal slices (projecting onto the r axis) produces the total probability density irrespective of the depth of A.

Since $n_A(z)$ is analytical, the NNDS $n_{A \to B_i}(r, z)$ is also analytical. It may be normalized easily since $\int_{z=-\infty}^{\infty} \int_{r=0}^{\infty} p_{z \to B_i}(r) n_A(z) dr dz = \int_{z=-\infty}^{\infty} n_A(z) dz = n_A^{2D}$: figure 1 shows $n_{A \to B_i}(r, z)/n_A^{2D}$ whereby the two Gaussian density profiles for $n_{A,B}(z)$ have a unit separation in depth, and each has an r.m.s width of $1/\sqrt{2}$ i.e. $d_A = d_B = 1$ unit of length, and unit height i.e. $n_A^{2D} \pi^{-1/2} d_A^{-1} = n_B^{2D} \pi^{-1/2} d_B^{-1} = 1$ inverse volume units.

From figure 1 the fraction of A's with an ith nearest B between $r \to r + \delta r$ is $\int_{z=-\infty}^{\infty} n_{A \to B_i}(r, z) dz \delta r / n_A^{2D}$. Referring to figure 1 this is equivalent to integrating over horizontal slices or projecting onto the r-axis. The probability of, say, the first nearest neighbour B having a separation of 1 unit from an A can be optimised by varying the separation and density of the layers. In the next section we consider the effect of adding constraints on the next nearest neighbours.

Density of specific cluster configurations

The total number of useful clusters, N_{good}, is a question of counting events and may be found from an integration like equation (1). The number of useful clusters in the elemental volume $\delta \mathbf{x}$ around $\mathbf{x}$ is given by the number of A's in the elemental volume multiplied by the probability that each is part of a useful cluster, $n_A(\mathbf{x})\delta \mathbf{x} P_{A\mathbf{x}}\{\text{Good Cluster}\}$. Hence the total is

$$N_{\text{good}} = \int_{\mathbf{x}} P_{A\mathbf{x}} \{\text{Good Cluster}\} n_A(\mathbf{x}) d\mathbf{x}. \tag{9}$$

We have used the shorthand notation $P_{A\mathbf{x}}\{\text{Good Cluster}\}$ for the conditional probability of a useful cluster configuration around $\mathbf{x}$ when it is given that a central A exists at point $\mathbf{x}$. This probability depends on the probability of specified numbers of events in regions of space around the A, and must be found from computations like equation (2). In the case of a 1D non-homogeneous density variation along z, the number of useful clusters (per unit area) is

$$N_{\text{good}}^{2D} = \int_z P_{Az} \{\text{Good Cluster}\} n_A(z) dz. \tag{10}$$

Cluster probability for donor qubit gates

Here we investigate the calculation of $P_{Az}\{\text{Good Cluster}\}$ for the example case of a simple qubit gate made with donors in silicon.

We are interested in multi-species clusters that have specifications on the separations. We imagine a pair of qubits that carry quantum information in their spin. The gate operation is performed by controlling the entanglement between two impurity electrons. By changing the state of the impurity of species A which we name the control, it will get entangled with a nearby impurity electron of species B which we name the target. This controlled change of state might be by optical excitation or use of electric fields etc. To facilitate the controlled interaction they must be separated by an appropriate radial distance, equivalent to defining a radial interval rather than an infinitesimal shell. We imagine such changes in interaction range occur in other fields such as in ecological networks of consumer species and resource species where there is e.g. a seasonal change in the interaction strength. We now add a specification on the next-nearest neighbours because the control and target impurities should be sufficiently isolated from the environment, i.e. other impurities, that they do not decohere. In ecology this might be analogous to the effects of competition.

With the results of the previous section we can calculate the probability density for an A donor having its first nearest (target) B at the optimal distance using equations (8) and (6) (or (3) for non Gaussian event density profiles). We now examine the combination of this condition with specification that the second nearest B and first nearest A are out of range. We allow for some tolerance on the useful target B distance.

In the simplest specification of our useful cluster for which the probability is $P_{A\mathbf{x}}\{\text{Good Cluster}\}$, we define a useful cluster as one in which the A control atom has:

1. its nearest A outside the range $r_{A_1} > r_{A \to A_1}^{\min}$, i.e. in the region labelled '1' on figure 2. This minimum ensures that when all As are in their excited state, they do not interact with each other, and the target B only feels its central controlling A;

2. its nearest B within the range $r_{A \to B_1}^{\min} < r_{B_1} < r_{A \to B_1}^{\max}$, i.e. in the region labelled '3' on figure 2. This ensures there is a target atom within range of the central control atom A when it is excited, but not so close that it is in range when the control is not excited;

3. its second nearest B outside the range $r_{B_2} > r_{A \to B_1}^{\max} + r_{B \to B_{1'}}^{\min}$ i.e. in the regions labelled '1' or '2' on figure 2, ensuring that the central A and its corresponding B_1 cannot interact with any other Bs.

These conditions may be recast in terms of the Poisson probabilities from above, $P_P\{i, N_X\}$ for i, X events within regions that contain an expectation N_X (which may be calculated from equation (4) for spherical regions in which the densities vary in 1D only):

i. there are no nearest A's within the complement of region '1' (regions '2'-'5') on figure 2. We shall refer to this region as $V_{A\mathbf{x}}^A$ to indicate the volume around the A control atom at $\mathbf{x}$ from which other A's are excluded. In it, the expected number of A's is $N_A(V_{A\mathbf{x}}^A)$.

ii. region '3' on figure 2 includes exactly one B. This region will be referred to as $V_{A\mathbf{x}}^{\text{target}}$, the volume around the A at $\mathbf{x}$ in which there is a B target atom. In it, the expectation number of B's is $N_B(V_{A\mathbf{x}}^{\text{target}})$.

iii. regions '4' and '5' on figure 2 contain no B's. The next equation becomes more compact if we define the combined volume of these regions with region '3' of the previous condition, $V_{A\mathbf{x}}^{B'}(\mathbf{x}')$, i.e. the total region for which there are conditions on the number of B's. $V_{A\mathbf{x}}^{B'}(\mathbf{x}')$ is the expected number of B's in this region.

It is relatively straightforward to specify this good cluster, though less easy to calculate its probability. The useful cluster, given an A at $\mathbf{x}$ and B at $\mathbf{x}'$, requires there are no A's in $V_{A\mathbf{x}}^A$ and there are no B's in $V_{A\mathbf{x}}^{B'}(\mathbf{x}')$ (other

Figure 2. Specification of a useful cluster for a qubit gating scheme with a central control atom of species *A* (red dots) with a target atom of species *B* (blue dots) for which there are restrictions on the distance from each other as described in the text.
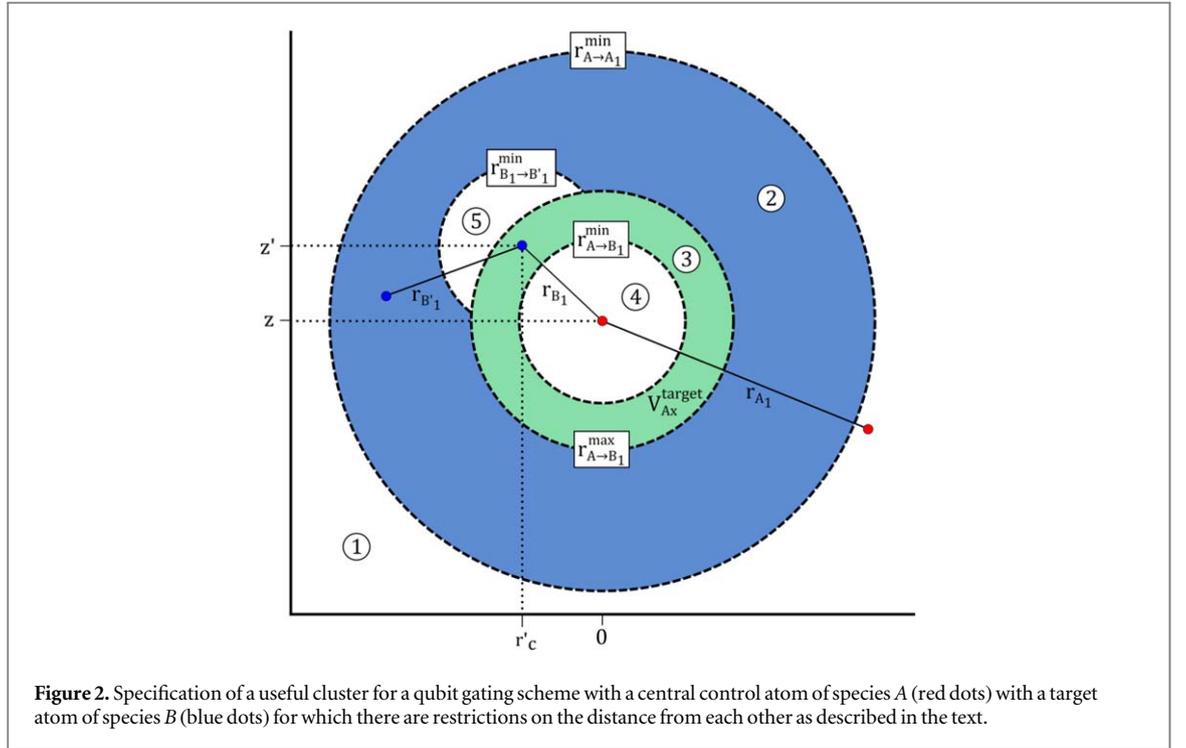
than those at $\mathbf{x}$ and $\mathbf{x}'$, which we consider to be elemental volumes so small that it does not affect the required expectation). Since a useful cluster having its *B* at $\mathbf{x}'$ is mutually exclusive with a useful cluster having its *B* at $\mathbf{x}''$ we can add these probabilities, i.e. integrate over the allowed range of $\mathbf{x}'$.

In the case of the 1D non-homogeneous problem the probability of a *B* in an elemental ring at cylindrical coordinates z' and r'_c (from the vertical axis containing the central *A* control atom) is $2\pi r'_c n_B(z') \delta r'_c \delta z'$, and so

$$P_{Az}\{\text{Good cluster}\} = \exp[-N_A(V^A_{Az})] \times \iint_{z', r'_c \in V^{\text{target}}_{Az}} 2\pi r'_c n_B(z') \exp[-N_B(V^{B'}_{Az}(z', r'_c))] dr'_c dz' \quad (11)$$

where the expected number of *B*'s within the region $V^{B'}_{Az}(z', r'_c)$ around an *A* at *z* and a *B* at z', r'_c is

$$N_B(V^{B'}_{Az}(z', r'_c)) = \int_{z'' \in V^{B'}_{Az}(z', r'_c)} S(V^{B'}_{Az}(z', r'_c), z'') n_B(z'') dz'' \quad (12)$$

and $S(V, z)$ is the area of a horizontal slice at height *z* through *V*.

The area $S(V^{B'}_{Az}(z', r'_c), z'')$ is a slice through the intersection of two spheres, which is surprisingly complicated but may be written analytically. Even so, the integral in equation (11) is a nested triple integral with complicated bounds. In cases where many calculations of $P_z\{\text{Good cluster}\}$ are required, such as in our problem of optimising the species density profiles, it is helpful to produce a heuristic method that accelerates the numerical calculation of this probability.

Heuristic method to approximate the best case cluster probability

So long as it is given that there is only B_1 within the region V^{target}_{Az}, then the probability of finding B_1 between $z' \to z' + \delta z'$ is proportional to $n(z') S(V^{\text{target}}_{Az}, z') \delta z'$. We can use this to find the location of B_1 within V^{target}_{Az} with the most important contribution to $P_z\{\text{Good cluster}\}$. Let us call the coordinates of this location Z'_{Az} and R'_{Az}, and let the regions '3'-'5' in figure 2 around this particular configuration be $V^{Bave'}_{Az}(R'_{Az}, Z'_{Az})$ (as usual the subscript indicates it is given that there is an *A* control atom at *z*). We may now use an approximate version of equation (11):

$$P_{Az}\{\text{Good cluster}\} \approx N_B(V^{\text{target}}_{Az}) \exp[-N_A(V^A_{Az}) - N_B(V^{Bave'}_{Az}(R'_{Az}, Z'_{Az}))] \quad (13)$$

There are a number of reasonable but different choices for calculating the most important location of the target Z'_{Az} and R'_{Az} for use in equation (13). We tried finding the expectation radius using (6)

$$R'_{Az} = \langle r' \rangle_{Az} = \frac{\int_{r' \in V^{\text{target}}_{Az}} r' p_{z \to B_1}(r') dr'}{\int_{r' \in V^{\text{target}}_{Az}} p_{z \to B_1}(r') dr'} \quad (14)$$

and the expectation depth given this spherical radius

$$Z'_{Az} = \langle z' \rangle_{Az} = \frac{\int_{z-R'_{Az}}^{z+R'_{Az}} z' \frac{n_A(z')}{n_A^{2D}} C(z') dz'}{\int_{z-R'_{Az}}^{z+R'_{Az}} \frac{n_A(z')}{n_A^{2D}} C(z') dz'}$$

(15)

where $C(z')$ is the circumference of the small circle of the sphere R'_{Az} through z'. We also tried finding the expectation depth and then the expectation spherical radius given this depth. Finally, we tried looking for the most likely depth (where the maximum $n_B(z)$ occurs) and most likely radius (at the outer edge of V_{Az}^{target}). The latter is the easiest to find and requires no integration yet dramatically underestimates $P_{Az}\{Good\ cluster\}$. We found that calculating R'_{Az} first gave the best agreement with (11) for the good cluster configuration we were interested in.

An approximate solution to this probability which is less computationally intensive accelerates the process of numerically optimising that probability. The closer the approximate solution is to the optimum found using the vigorous method, the more efficiently one can converge to an optimum Gaussian doping profile.

Results of optimising cluster probability

We provide a numerical example of the cluster optimization in the case of the silicon donor qubit gate using separation tolerances estimated from consideration of exchange interactions [8]. The separation range for the control to target distance is from $r_{A \to B_1}^{min} = 15$nm to $r_{A \to B_1}^{max} = 28$ nm. The exclusion radius for control to control is $r_{A \to A_1}^{min} = 60$ nm, and for target to target is $r_{B_1 \to B_1'}^{min} = 15$ nm.

To optimise the number of good clusters we allow for four independent parameters for the Gaussian density profiles—the two areal densities $n_{A,B}^{2D}$, the width of both density profiles $d_A = d_B = d$, and the separation of the two layers μ. Here we make the assumption that both profiles can be implanted at different depths with the same width. In practise, independent control of layer depth and width is not achievable with ion implantation and the depth profile for a particular implant species and target depends principally on the implant energy. The profile of impurities will also change during necessary post-processing such as diffusion during annealing and whether each species is annealed as implanted or if the system is annealed only once after all implants. As these specifics are dependent on the species and target of interest we leave possible simplifications to the reader. It has been shown previously that Bi implanted at a high energy into silicon (therefore creating many lattice defects) can be electrically activated to a high enough quality that the donor electron spin states are measurable[16]. This demonstrates the feasibility of ion implantation as the means creating the qubits of interest here. The following optimisation can be used to determine a best case implant profile if the final donor profiles are expected to be Gaussian. If the final active donor profile can be measured, it may be used directly with equations (10) and (11) to determine the final viable cluster yield.

We first maximised the total density of good clusters, N_{good}^{2D} for a given combination of width and separation using the heuristic as detailed in the previous section. This is shown in figure 3(a), in which the values of $n_{A,B}^{2D}$ were varied (the resulting optimum values of $n_{A,B}^{2D}$ are not shown in this figure) for various combination of μ and d to find the optimum good cluster density N_{good}^{2D}.

For sufficiently large layer widths layer separation clearly has no effect on either N_{good}^{2D} or N_{good}^{2D}/n_A^{2D}. In this limit the optimum values of $n_A^{2D} d_A = n_B^{2D} d_B$ tend to the optimum homogeneous bulk densities [8]. We see that for layers spaced far apart the density of viable clusters tends to zero as the layer width is reduced, as expected as it becomes increasingly difficult to obtain an $A \to B$ distance within the allowed range. For very narrow layers, there is an obvious optimum layer separation of 15 nm which can be deduced from geometrical configurations relating to the definition of our 'good configuration'. This distance is the same as the minimum separation of control from target. For intermediate layer widths of around 10 nm the lines cross, and the optimum is now obtained for layers of zero separation, i.e. the target and control layers should be at the same depths.

The areal density N_{good}^{2D} is not necessarily the most useful optimization objective function. Even with a small density, the total number of good clusters may be increased simply by increasing the sample size. The ratio of signal to background in an experiment might be improved if instead we maximise the fraction of donors that are involved in a good cluster. For example, we imagine an experiment where we detect the effects of the interaction by measuring the effect on the spin of the A's after exciting the A's (which produces an effect only for those A's that are part of a good cluster). In our simple cluster of interest there is only one A per cluster, so the number of A's involved in a good cluster is just equal to the number of good clusters, and the signal to background will be optimized by maximising the fraction N_{good}^{2D}/n_A^{2D}. By simply maximising this fraction, the optimum occurs when n_A^{2D} is as small as possible. In this limit the condition on the control's nearest A ($r_{A \to A_1} > 60$nm) becomes guaranteed, and the only conditions that need to be satisfied are the ones on the nearest and next-nearest B's. Alternatively, we might imagine a different experiment where we detect the effect of the interaction by
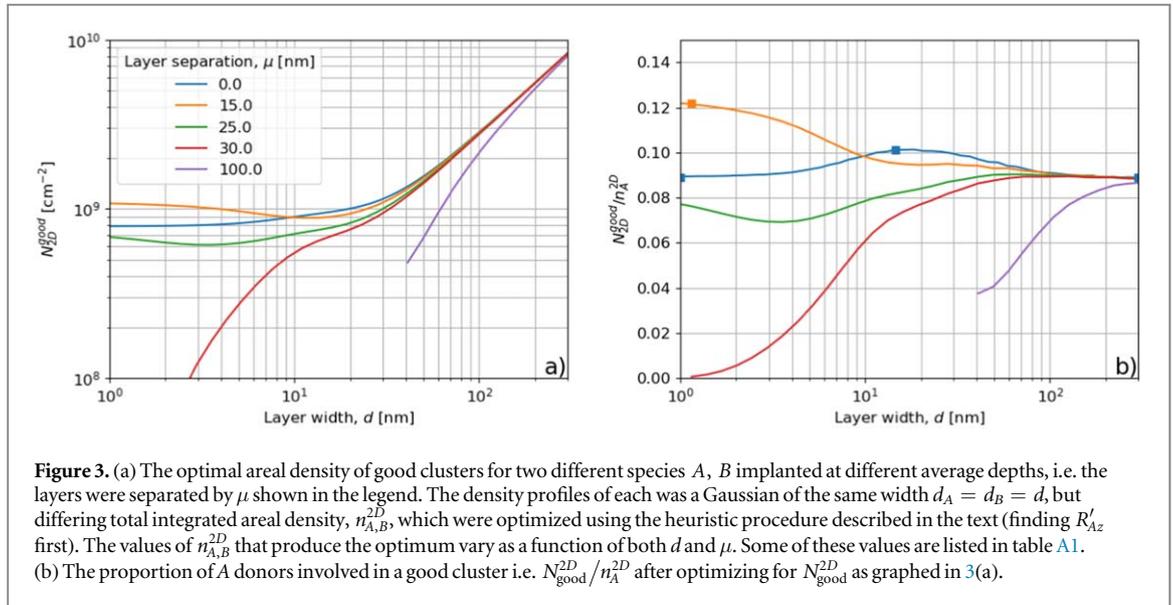
Figure 3. (a) The optimal areal density of good clusters for two different species A, B implanted at different average depths, i.e. the layers were separated by μ shown in the legend. The density profiles of each was a Gaussian of the same width $d_A = d_B = d$, but differing total integrated areal density, $n_{A,B}^{2D}$, which were optimized using the heuristic procedure described in the text (finding R'_{Az} first). The values of $n_{A,B}^{2D}$ that produce the optimum vary as a function of both d and μ. Some of these values are listed in table A1. (b) The proportion of A donors involved in a good cluster i.e. $N_{\text{good}}^{2D}/n_A^{2D}$ after optimizing for N_{good}^{2D} as graphed in 3(a).
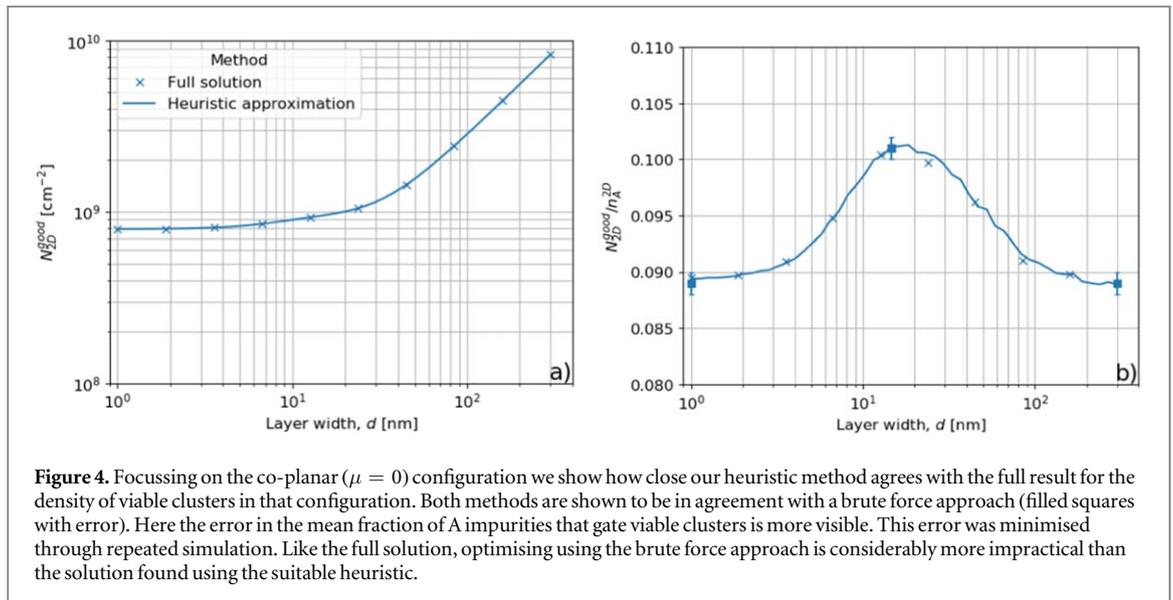
Figure 4. Focussing on the co-planar ($\mu = 0$) configuration we show how close our heuristic method agrees with the full result for the density of viable clusters in that configuration. Both methods are shown to be in agreement with a brute force approach (filled squares with error). Here the error in the mean fraction of A impurities that gate viable clusters is more visible. This error was minimised through repeated simulation. Like the full solution, optimising using the brute force approach is considerably more impractical than the solution found using the suitable heuristic.

measuring the effect on the spin of the B's after exciting the A's, and therefore we optimise the fraction $N_{\text{good}}^{2D}/n_B^{2D}$, which occurs when n_B^{2D} is as small as possible for a similar argument. To avoid these cases where the optimum density of a species tends to zero we optimised for the absolute number of good clusters which optimises the total signal (figure 3(a)), and subsequently calculated the corresponding fraction $N_{\text{good}}^{2D}/n_A^{2D}$ describing the ratio of signal to background shown in figure 3(b).

3(b) shows that if one were able to fabricate atomically flat layers separated in depth then this fraction is optimised when such layers are separated by 15 nm. An interesting situation arises if there is a lower limit on the possible width of the density profiles. This is the case when ion implanting species into a lattice. Depending on the implantation specifics, it is difficult to make very thin layers due to ion straggle. Figure 3 shows that if the layers cannot be fabricated with widths less than 10 nm, the optimum configuration is to have the two species co-planar ($\mu = 0$) with widths of ≈ 15 nm (not to be as thin as possible) achieving a reasonable 10%. The fractions in both of these cases are an improvement over the optimised, bulk dope case achieving a good cluster fraction of 9%. The results of 3 doping configurations were simulated using a brute force (Monte Carlo) approach (filled squares with error) confirming the densities as calculated by (10).

The quality of the optimization is shown in figure 4, in which we compare it with optimization using the full solution of equations 11–12, and also using a brute force (Monte Carlo) method. Agreement is excellent for cases examined here where the two species profiles overlap in depth. We found that the heuristic method does not agree with the full solution when the layers are thinner than the separation between layers as shown in figure 5. Under these circumstances the optimum B density can differ by as much as 2 orders of magnitude.

Figure 5. The density of viable clusters plotted as a function of the areal density of *B* donors. The same optimum *B* is found using either the heuristic (dashed lines) of full solution (solid lines). These methods where compared for different layer widths (coloured lines) and different layer separations (separate subplots). The discrepancy between methods is only apparent for layers whose profile width is smaller than their separation.

Conclusion

This work has demonstrated how the ideas used in non-homogeneous Poison point process statistics may be used not only to describe the distribution of *i*th nearest neighbour distances between events that exist non-homogeneously in space but also to calculate the probabilities of events forming more complex structures existing in such a process. When the event species density inhomogeneity is described by a Gaussian probability density function of depth and with complete spatial randomness in the *xy*-plane an analytic solution to the NNPDF is yielded. This density surface gives a good representation of event-event separations existing in non-homogeneous processes. for the purposes of research into solid state qubit gates, we calculate the frequency of appearance of suitable structures for qubit interactions in random doped samples. The introduction of a heuristic for calculating the probabilities of such structures improves the ability to optimise the dopant profiles for these interactions. Such an optimisation will influence the fabrication requirements of test-bed solid state quantum devices whereby distance dependant interactions with nearest neighbours are of paramount importance.

Acknowledgments

Appendix A. Poissonian statistics of events in non-homogeneous point process

The sizes of the ith elements ($\delta \mathbf{x}_i$) that make up the finite domain V in which there are an expected $N(V)$ events can be varied such that the expected number of events within those elements is constant. $\delta N_i = N/M = \delta N$, where M is the number of elemental volumes. As N is finite, we can make M large enough and δN small enough that we may neglect the possibility of two or more events in any element; hence the probability of no events is $1 - \delta N$ for each element. Since the incidence of events in each element is independent from the others, the probability of no events in V is $P_P\{0, N\} = (1 - \delta N)^M$. For large M and small δN this tends to $P_P\{0, N\} = \exp(-M\delta N) = \exp(-N)$. This is the same as the usual Poisson probability of zero events when the expectation is N, in spite of the fact that $n(\mathbf{x})$ is non-homogeneously distributed within V.

We may further find the probability of one event in the volume V by considering one additional elemental volume added to it; $P_P\{1, N + \delta N\} = P_P\{1, N\}(1 - \delta N) + P_P\{0, N\}\delta N$, where the first term is the probability of no events in the additional element and one event in the original volume, and the second term is for the opposite scenario (the two scenarios are mutually exclusive so the probabilities add). Rearranging, $[P_P\{1, N + \delta N\} - P_P\{1, N\}]/\delta N = P_P\{0, N\} - P_P\{1, N\}$, in which the LHS is $\frac{d}{dN}P_P\{1, N\}$. Making use of the result just obtained for $P_P\{0, N\}$, we see the solution is $P_P\{1, N\} = N\exp(-N)$. Continuing similar arguments for higher numbers of events, m, we again recover the standard Poisson distribution (2) with expectation N (even though $n(\mathbf{x})$ is non-homogeneous).

Table A1. Numerically found solutions for the optimum integrated impurity densities for particular points along the curves displayed in figures 3 and 4.

Separation μ [nm]	Width d [nm]	Optimum A density n_A [cm^{-2}]	Optimum B density n_B [cm^{-2}]
0	1.0	4.99×10^9	2.14×10^{10}
	14.5	3.64×10^8	1.97×10^9
	300.0	1.76×10^8	1.66×10^9
15	1.2	4.33×10^9	2.62×10^{10}
	16.0	3.34×10^8	2.28×10^9
	281.2	2.23×10^8	1.74×10^9
30	14.5	3.63×10^8	2.88×10^{10}
	300.0	1.75×10^8	1.65×10^9

Appendix B. Optimum impurity densities for configurations of interest

ORCID iDs

K Stockbridge ⬤ https://orcid.org/0000-0002-5361-3404
S Chick ⬤ https://orcid.org/0000-0001-8599-6609

References

[1] Stoneham A M, Fisher A J and Greenland P T 2003 Optically driven silicon-based quantum gates with potential for high-temperature operation *J. Phys. Condens. Matter* **15** L447

[2] Le N H, Fisher A J and Ginossar E 2017 Extended Hubbard model for mesoscopic transport in donor arrays in silicon *Phys. Rev.* B **96** 245406

[3] Stock T, van Loon M, Warschkow O, Hofmann E, Crane E, Schofield S and Curson N 2019 Atomic scale patterned arsenic in silicon *APS Meeting Abstracts* 11.013

[4] Schofield S R, Studer P, Hirjibehedin C F, Curson N J, Aeppli G and Bowler D R 2013 Quantum engineering at the silicon surface using dangling bonds *Nat. Commun.* **4** 1649

[5] Crane E, Kölker A, Stock T Z, Stavrias N, Saeedi K, van Loon M A W, Murdin B N and Curson N J 2018 Hydrogen resist lithography and electron beam lithography for fabricating silicon targets for studying donor orbital states *Journal of Physics: Conf. Series* **1079** 012010

[6] Scappucci G, Capellini G, Lee W C T and Simmons M Y 2009 Atomic-scale patterning of hydrogen terminated Ge(001) by scanning tunneling microscopy *Nanotechnology* **20** 3–9

[7] Koenraad P M and Flatté M E 2011 Single dopants in semiconductors *Nat. Mater.* **10** 91–100

[8] Crane E, Crane T, Schuckert A, Le N H, Stockbridge K, Chick S and Fisher A J 2019 Optically controlled entangling gates in randomly doped silicon *Phys. Rev.* B **100** 064201

[9] Thompson H R 1956 Dirstribution of distance to nth neightbour in a population of randomly distributed individuals *Ecology* **37** 391–4

[10] Al-Hourani A, Evans R J and Kandeepan S 2016 Nearest neighbor distance distribution in hard-core point processes *IEEE Commun. Lett.* **20** 1872–5

[11] Babu G J and Feigelson E D 1996 Spatial point processes in astronomy *J. Stat. Plan. Inference* **50** 311–26

[12] Torquato S, Lu B and Rubinstein J 1990 Nearest-neighbor distribution functions in many-body systems *Phys. Rev. A* **41** 2059-2075

[13] Moltchanov D 2012 Distance distributions in random networks *Ad Hoc Networks* **10** 1146–66

[14] Thomas G a, Capizzi M, DeRosa F, Bhatt R N and Rice T M 1981 Optical study of interaction donors in semiconductors *Phys. Rev.* B **28** 5472

[15] Edgal U F and Wiley J D 1983 Near-neighbor configuration and impurity-cluster size distribution in a Poisson ensemble of monovalent impurity atoms in semiconductors *Phys. Rev.* B **27** 4997-5006

[16] Peach T, Homewood K, Lourenco M, Hughes M, Saeedi K, Stavrias N, Li J, Chick S, Murdin B and Clowes S 2018 The effect of lattice damage and annealing conditions on the hyperfine structure of ion implanted bismuth donors in silicon *Advanced Quantum Technologies* **1** 1800038