Bivalent nearest-available-neighbor distribution in n dimensions: A Monte Carlo calculation

J. R. Eggert
Department of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 6 September 1983; revised manuscript received 4 January 1984)

The bivalent nearest-available-neighbor distribution $P_{\mathrm{NAN},n}$ in n dimensions is of interest in the statistics of recombining particles. We generalize to arbitrary dimension a calculation of $P_{\mathrm{NAN},3}$ due to Dunstan. This derivation, however, is shown to be based on an invalid approximation. A Monte Carlo calculation of $P_{\mathrm{NAN},n}$ for $n=1$, 2, and 3 gives results significantly different from the Dunstan form. We propose the asymptotic form $P_{\mathrm{NAN},1}(r)\sim r^{-3/2}$ for large r.

I. INTRODUCTION

The statistics of recombining particles is of general interest in physics. In this paper we consider the case where there are two species of particles, A and B, in equal numbers, and each distributed randomly with a finite density in a given volume. Each particle is assumed to be localized and immobile, and to interact only by recombining with a particle of the opposite species. Furthermore, the recombination rate is assumed to be a sufficiently rapidly decreasing function of distance, so that each particle will recombine only with the nearest particle possible. In this limit, the kinetics of the recombination will be determined by the distribution of recombination distances. A possible example of the application of this model is the recombination of electrons and holes created by photoexcitation in an exactly compensated crystal.

It is tempting to suggest that the desired recombination distance distribution is simply the distribution of interspecies nearest-neighbor distances, $P_{\mathrm{NN},n}(r)$ in n dimensions. For this system $P_{\mathrm{NN},n}$ is of the form[1]

$$P_{\mathrm{NN},n}(r)=\rho n v_n r^{n-1}\exp(-\rho v_n r^n)\ , \tag{1}$$

where ρ is the density of each of the species, and v_n is the volume of the unit n-dimensional sphere,

$$v_n=\pi^{n/2}/\Gamma(\tfrac{1}{2}n+1)\ .$$

This distribution is not appropriate, however, because the nearest-neighbor relation that it describes is not commutative. If A_1 is the nearest neighbor of B_1, then B_1 is not necessarily the nearest neighbor of A_1. In fact, it is easy to show for this bivalent system that in any dimension exactly half of the particles are the nearest neighbors of their nearest neighbors.

It is this asymmetry that makes the nearest-neighbor distribution inappropriate for the description of recombining localized particles.[2] Recombination is inherently commutative: if A_1 recombines with B_1, then B_1 must recombine with A_1. A commutative nearest-neighbor relation is required; we call its associated distribution $P_{\mathrm{NAN},n}(r)$, the bivalent nearest-available-neighbor (NAN) distribution in n dimensions.

The NAN relation is commutative by construction. A particle A_1 and a particle B_1 are NAN's if each is available at the distance of the other. The definition of availability is recursive: A_1 is available at distance r if there is no particle B_0 at distance $r'<r$, which itself is available at distance r'. When applied to a finite collection of particles, the definition of NAN pairs is equivalent to sequentially selecting, without replacement, the pairs of particles with the least separation. The relation to localized particle recombination is now obvious. The NAN pairing process is unambiguous only for random distributions of particles, where any problems entailing equidistant particles exist only for a vanishingly small subset of all the particles. We also note here that the assumption of two species of particles, and the type of pairing allowed determine the form of the NAN distribution.

II. ANALYSIS

An ansatz used by Dunstan[3] to derive an analytic form $P_{D,n}(r)$ (Ref. 4) for the NAN distribution is based on the standard probabilistic derivation of $P_{\mathrm{NN},n}(r)$. In this method the probability of a particle A_0 finding a NAN at distance r is the product of three subprobabilities: the probability that A_0 has not found a NAN nearer than r (i.e., that A_0 is available at distance r), that there is a particle B_0 at r, and that B_0, in turn, has not found a NAN nearer than r. Thus P_D is written as

$$P_{D,n}(r)dr=\rho n v_n r^{n-1}dr\left[1-\int_0^r P_{D,n}(r')dr'\right]^2\ . \tag{2}$$

This equation can be solved by elementary methods (see Appendix A) to give

$$P_{D,n}(r)=\frac{\rho n v_n r^{n-1}}{(1+\rho v_n r^n)^2}\ . \tag{3}$$

We also define

$$Q_{\mathrm{NAN},n}(r)\equiv\int_r^\infty P_{\mathrm{NAN},n}(r')dr'\ , \tag{4}$$

in general, so that

$$Q_{D,n}(x)=\frac{1}{1+v_n x^n}\ , \tag{5}$$

where $x=\rho^{1/n}r$ is the normalized pair separation. One may also calculate the mean NAN pair separation $\langle r_{\mathrm{NAN},n}\rangle$ in n dimensions as

$$\langle r_{\mathrm{NAN},n}\rangle \equiv \int_0^\infty r P_{\mathrm{NAN},n}(r)dr = \rho^{-1/n}\int_0^\infty Q_{\mathrm{NAN},n}(x)dx \ . \quad (6)$$

In the Dunstan ansatz this becomes

$$\langle r_{D,n}\rangle = (\rho v_n)^{-1/n}(\pi/n)\csc(\pi/n) \ . \quad (7)$$

In this calculation $\langle r_{D,n}\rangle$ converges only for $n \geq 2$.

The Dunstan ansatz, although easily solved, is based on the critical assumption that the three subprobabilities used in Eq. (2) are independent. This is not the case, however, as can be seen in Fig. 1. For generality, we draw the counterexamples in one dimension. In Figs. 1(a) and 1(b) we show that the probability that A_0 has not yet found a NAN by the distance r is not independent of the existence of a particle B_0 at r. In Figs. 1(c) and 1(d) we demonstrate that the probability that A_0 has not yet found a NAN by the distance r is also not independent of the same probability for B_0. Similar examples do not exist for the nearest-neighbor relation; the second independence is not required by the nearest-neighbor calculation. Therefore, the assumption of independent subprobabilities in Eq. (2) is a decoupling approximation, whose accuracy we wish to investigate.

A more general approach not based on approximations, reproduced in Appendix B, allows one to write P_{NAN} in the form of an intractable integral. One may still write for sufficiently small r

$$P_{\mathrm{NAN},n}(r) \approx \rho n v_n r^{n-1} \ , \quad (8)$$

valid before any NAN competition becomes important. For intermediate and large r we propose a Monte Carlo calculation of $P_{\mathrm{NAN},n}$ using the following algorithm: Generate N random particles of each of two species, in a unit n-dimensional cube. Using periodic boundary conditions, find the pair (A_i,B_j) with the smallest pair separation r_s. Record r_s, delete A_i and B_j, and repeat until all the particles are exhausted. When N is sufficiently large, the distribution of r_s will give an adequate approximation to $P_{\mathrm{NAN},n}(r)$ with $\rho = N$. This algorithm is equivalent to the recursive definition for $P_{\mathrm{NAN},n}(r)$ given above.

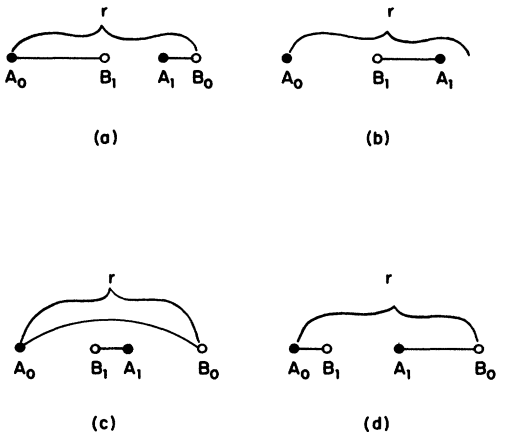

FIG. 1. Nearest-available-neighbor complexes showing the correlation of the subprobabilities used in Eq. (2). (a) In this configuration A_0 and B_1 are NAN's, as are A_1 and B_0. A_0 is not available at distance r. (b) Same as (a), but now B_0 is missing. Now B_1 and A_1 are NAN's, and A_0 *is* available at distance r. For this pair of cases, the existence of B_0 is linked to the availability of A_0 at the distance of B_0. (c) In this configuration A_0 and B_0 are NAN's, as are B_1 and A_1. A_0 is available at distance r, as is B_0. (d) Same as (c), but B_1 is shifted towards A_0. Now A_0 is no longer available at distance r, and neither is B_0. For this pair of cases, the availability of A_0 at the distance of B_0 is linked, via the exact position of B_1, to the availability of B_0 at the distance of A_0.

III. RESULTS

We wrote a FORTRAN program that performs the $P_{\mathrm{NAN},n}(r)$ algorithm with execution time $\sim N^2$. As a test of the algorithm, the program also generated the nearest-neighbor distribution as an intermediate step. In all cases this Monte Carlo nearest-neighbor distribution converged well to the form of Eq. (1). The data generated for $P_{\mathrm{NAN},n}(r)$ converged well at small r to the form required by Eq. (8). 10 000 particles of each species were used for the one-dimensional calculation, and 5000 for the two- and three-dimensional calculations.

In Figs. 2, 3, and 4 are presented the results for the NAN distribution in one, two, and three dimensions, respectively, plotted as $\log_{10}Q_{\mathrm{NAN},n}(x)$ versus $\log_{10}x$. For comparison the predictions of Eq. (5) are also shown. In all three cases the large-x dependence of $Q_{\mathrm{NAN},n}(x)$ is markedly shallower than that of the analytic form Q_D. For $x \gtrsim 4$, the analytic form is too low by at least a half an order of magnitude for $n=2$ and 3. In Fig. 5 the ratios of the Monte Carlo results to Q_D for $n=1$, 2, and 3 are shown. It is clear that the decoupling approximation ansatz consistently underestimates the number of NAN pairs with large separations and becomes increasingly inaccurate for higher dimensionalities.

The large-x dependences of $Q_{\mathrm{NAN},n}(x)$ in Figs. 2–4 all fit well to a power law

$$Q_{\mathrm{NAN},n}(x) \sim a_n x^{-b_n} \quad (\text{large } x) \ . \quad (9)$$

The deviations from this power law at the largest x seen in the figures can be attributed to the finite number of

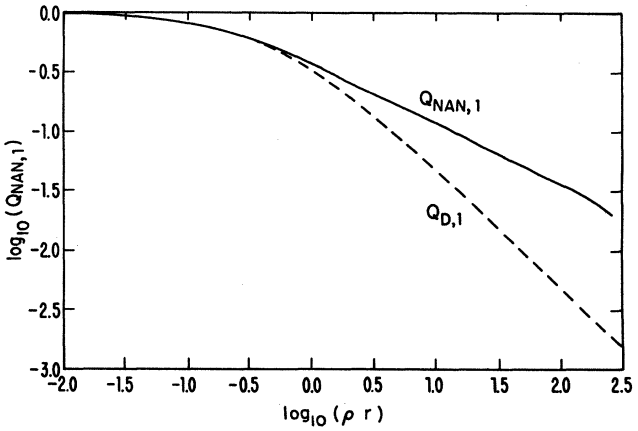

FIG. 2. One-dimensional nearest-available-neighbor distribution. Analytic theory (dashed line) and Monte Carlo calculation (solid line). $N = 10^4$ particles of each type were used in the calculation.

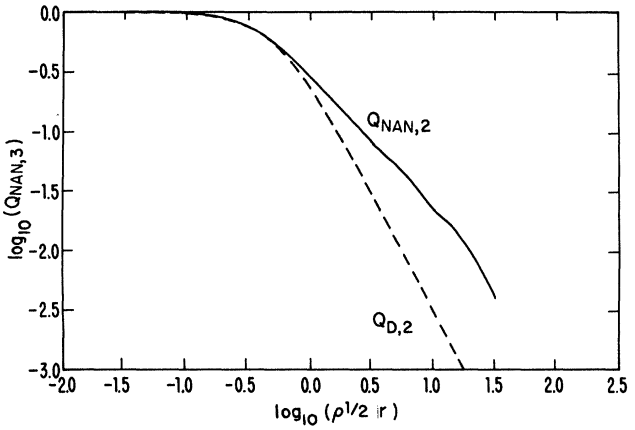

FIG. 3. Two-dimensional nearest-available-neighbor distribution. Analytic theory (dashed line) and Monte Carlo calculation (solid line). $N=5\times10^3$ particles of each type were used in the calculation.

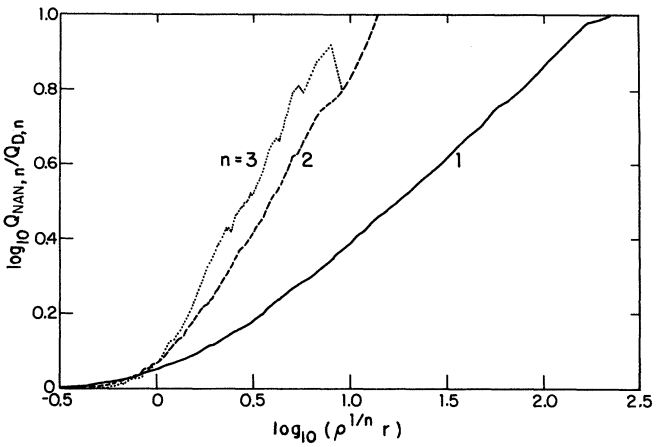

FIG. 5. Ratio of Monte Carlo results to analytic calculation of nearest-available-neighbor distribution for $n=1$ (solid line), 2 (dashed line), and 3 (dotted line).

particles used in the calculation. The analytic theory would predict $a_{D,n}=v_n^{-1}$ and $b_{D,n}=n$. In Fig. 6 we plot the results of a least-squares fit of Eq. (9) to the Monte Carlo data, as a_n and b_n versus n. The statistical uncertainties of the fit are smaller than the size of the symbols. For the one-dimensional case the asymptotic power law is quite good, and is of the form

$$Q_{\mathrm{NAN},1}(x)\sim 0.38x^{-1/2} \quad (\text{large } x) \tag{10}$$

or, equivalently,

$$P_{\mathrm{NAN},1}(r)\sim 0.19\rho^{-1/2}r^{-3/2} \quad (\text{large } r)\,. \tag{11}$$

Because higher-dimensional space is more difficult to sample using a Monte Carlo approximation, the asymptotic form for $n=2$ and 3 is less certain than the one-dimensional case. In two dimensions the algebraic exponent for Q is near -1, so it is uncertain whether $\langle r_{\mathrm{NAN},2}\rangle$ converges. We can calculate the mean NAN separation for $n=3$ by numerical integration of the data of Fig. 4 by using Eq. (6). The result is $\langle r_{\mathrm{NAN},3}\rangle \approx 0.9\rho^{-1/3}$, compared with $\langle r_{D,3}\rangle \approx 0.75\rho^{-1/3}$ for the analytic form.

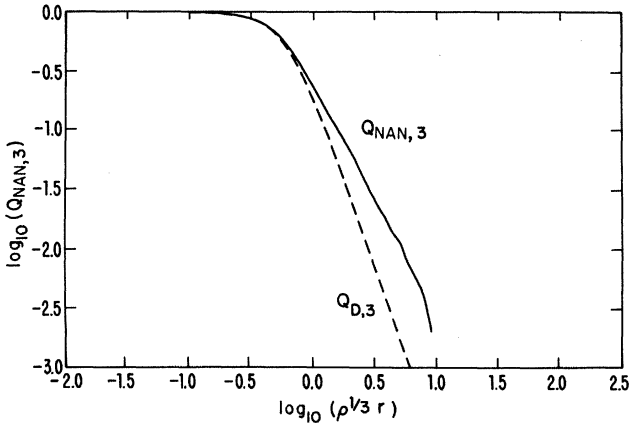

FIG. 4. Three-dimensional nearest-available-neighbor distribution. Analytic theory (dashed line) and Monte Carlo calculation (solid line). $N=5\times10^3$ particles of each type were used in the calculation.

IV. CONCLUSION

We have generalized to arbitrary dimension the Dunstan ansatz for the nearest-available-neighbor distribution, $P_{\mathrm{NAN},n}(r)$. This derivation, however, is shown to be based on the inaccurate approximation of independent probabilities. By a Monte Carlo technique we approximate $P_{\mathrm{NAN},n}(r)$ and show that the number of NAN pairs with large separation is consistently larger than previously calculated, and that the deviation is larger for higher dimensionality. In the present calculation the asymptotic form of $P_{\mathrm{NAN},1}$ is shown to be $r^{-3/2}$ in one dimension. The

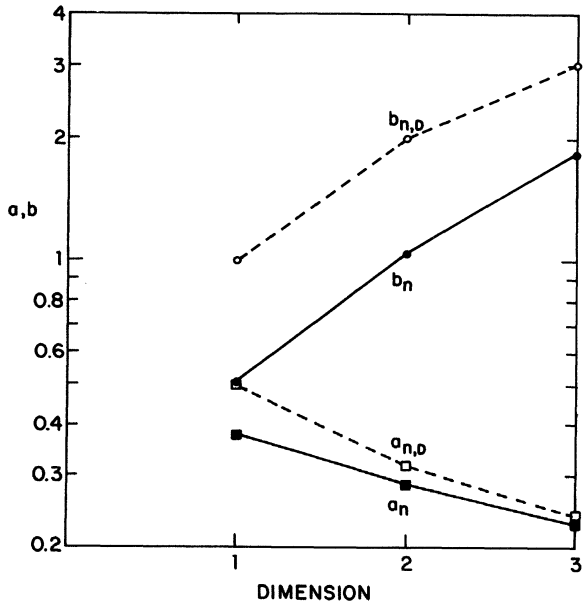

FIG. 6. Parameters a and b of a least-squares fit $Q=ax^{-b}$ to the asymptotic form of $Q_{\mathrm{NAN},n}$ for $n=1$, 2, and 3, compared with the predictions of the analytic theory.

data for higher dimensions indicate that $P_{\mathrm{NAN},2}$ and $P_{\mathrm{NAN},3}$ behave approximately as r^{-2} and r^{-3}, respectively, for large r. The large-r dependence is very important for the recombination kinetics of localized particles, for there it is assumed that the recombination time is a rapidly increasing function of the distance. Even a small error in the estimation of a typical recombination distance would lead to a large error in the recombination time estimate. For a system with a large distribution of recombination times, the present calculation demonstrates that the density of quasistable distant pairs is significantly larger than previously predicted. Furthermore, because the shape of the distribution of pair separations is different from the former approach, the kinetics of recombination must be of a different form. In light of these results, any theory that depends on the form of $P_{\mathrm{NAN},n}$ must be reevaluated.

ACKNOWLEDGMENTS

I wish to thank W. Paul for seminal conversations and for a careful reading of the manuscript. This research was supported by the National Science Foundation under Grant No. DMR81-08327.

APPENDIX A

We generalize the integral equation used in the Dunstan ansatz:

$$P_G(r)dr=g(r)dr\left[1-\int_0^r P_G(r')dr'\right]^m \tag{A1}$$

and assume $m\neq 1$ for now. Now using the definition of Q,

$$Q_G(r)=\int_r^\infty P_G(r')dr' , \tag{A2}$$

the integral becomes

$$-\frac{d}{dr}Q_G(r)=g[Q_G]^m . \tag{A3}$$

This differential equation has the solution

$$Q_G(r)=\left[(1-m)\left[h-\int_0^r g(r')dr'\right]\right]^{1/(1-m)} , \tag{A4}$$

where h is a constant determined by the normalization of P_G,

$$\int_0^\infty P_G(r)dr=Q_G(0)=1 , \tag{A5}$$

so that for suitably reasonable $g(r)$, $h=1/(1-m)$. Hence,

$$Q_G(r)=\left[1-(1-m)\int_0^r g(r')dr'\right]^{1/(1-m)} . \tag{A6}$$

Now for $m=2$ and $g(r)=\rho n v_n r^{n-1}$, this becomes

$$Q_{D,n}(r)=(1+\rho v_n r^n)^{-1} , \tag{A7}$$

the solution of the Dunstan ansatz.

If, on the other hand, we take the limit of Eq. (A6) as $m\rightarrow 1$, and use the definition of the exponential, then for the same $g(r)$ the result is the familiar nearest-neighbor distribution

$$Q_{\mathrm{NN},n}(r)=\exp(-\rho v_n r^n) . \tag{A8}$$

APPENDIX B

The n-dimensional NAN distribution may be written as a limit of a multidimensional integral. Let there be N particles of each of two species, distributed randomly in a volume V. Let $r_{A_1}\cdots r_{A_N}$ and $r_{B_1}\cdots r_{B_N}$ be the n-dimensional coordinates of the N particles of type A and B, respectively. We choose to number the particles in NAN pairs, in order of increasing NAN pair separation. Thus the $2Nn$-dimensional space of volume V^{2N}, which describes the positions of all the particles, is divided into $(N!)^2$ equivalent subvolumes, each with the pair ordering

$$|R_1|<|R_2|<|R_3|<\cdots<|R_N| , \tag{B1}$$

where $R_k\equiv r_{A_k}-r_{B_k}$. Now the other restriction inherent in the NAN selection procedure is that the kth NAN pair is that pair with the least separation, after the removal of the NAN pairs $1,\ldots,k-1$. This restriction may be written

$$|R_k|\leq|r_{A_i}-r_{B_j}| \quad \text{for all } i,j\geq k . \tag{B2}$$

Then we may write

$$P_{\mathrm{NAN},n}(r)=\lim_{\substack{N\rightarrow\infty,\\ V\rightarrow\infty,\\ N/V=\rho}}\frac{(N!)^2}{V^{2N}}\int_{V_R}dr_{A_1}\cdots dr_{A_N}dr_{B_1}\cdots dr_{B_N}\times\sum_{k=1}^N\delta(r-|R_k|) \tag{B3}$$

where V_R is the $2Nn$-dimensional integration volume subject to the restrictions (B1) and (B2). Alternatively, the geometric constraints (B1) and (B2) may be converted into multiplicative factors in the integrand to give

$$P_{\mathrm{NAN},n}(r)=\lim_{\substack{N\rightarrow\infty,\\ V\rightarrow\infty,\\ N/V=\rho}}\frac{(N!)^2}{V^{2N}}\int_{V_u}dr_{A_1}\cdots dr_{A_N}dr_{B_1}\cdots dr_{B_N}\left[\sum_{k=1}^N\delta(r-|R_k|)\right]\left[\prod_{k=2}^N\Theta(|R_k|-|R_{k-1}|)\right]$$

$$\times\prod_{k=1}^N\prod_{i=k}^N\prod_{j=k}^N\Theta(|r_{A_i}-r_{B_j}|-|R_k|) , \tag{B4}$$

where

$$\Theta(x)=\begin{cases}1, & x\geq 0\\ 0, & x<0\end{cases}$$

and V_u is the unrestricted $2Nn$-dimensional integration volume of volume V^{2N}. It is the intractability of this exact expression that leads to the use of a Monte Carlo approximation technique.

---

[1]P. Hertz, Math. Ann. 67, 387 (1909). See also Appendix A.
[2]D. J. Dunstan, Philos. Mag. B 46, 579 (1982).
[3]D. J. Dunstan, in Proceedings of the 16th International Conference on the Physics of Semiconductors, Montpellier, 1982 [Physica 117&118B + C, 902 (1983)].
[4]We use the subscript D to denote any variable which is evaluated according to the Dunstan ansatz, and the subscript NAN for the correct evaluation.