HOPPING CONDUCTIVITY IN LIGHTLY DOPED SEMICONDUCTORS

by

MARK R. A. SHEGELSKI

B.Sc. University of Calgary, 1979

M.Sc. The University of British Columbia, 1982

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Physics)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November, 1986

© Mark. R. A. Shegelski, 1986
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of  

PHYSICS  

The University of British Columbia  
1956 Main Mall  
Vancouver, Canada  
V6T 1Y3  

Date  

November 17, 1986
ABSTRACT

In lightly doped semiconductors (LDSs), electrons can exist in localized states around impurities and dc electronic conduction can occur by electrons hopping between localized states. Such hopping is the dominant mechanism for conduction if the temperature is so low that the contribution from band electrons is negligible.

According to theories of hopping conduction, at low enough temperature $T$, the conductivity $\sigma$ will be $\sigma = \sigma_0 e^{- (T_0/T)^{1/4}}$ where $T_0$ is a temperature which depends on the material. Experimental work on doped semiconductors which exhibits this form of $\sigma$ is scarce. Recently, however, conductivities which were clearly of this form were reported for lightly doped n-GaAs and lightly doped n-InP. The experimental results were surprising in that the temperature ranges were well above, and the $T_0$ values well below, the limits set by the theories.

To understand these experimental results, hopping in LDSs is modelled in this dissertation using a resistor network. This dissertation is unique in that the conductivity of the unabridged resistor network is examined in a temperature range (called "the high temperature regime") where $kT$ is comparable to the spread $\Delta \varepsilon$ in the energies of localized electrons. A numerical simulation is performed and an analytic theory based on percolation methods is presented. In this dissertation, an analytic approach is developed for the first time for studying how, in the high temperature regime, the conductivity of the unabridged resistor network depends on the density of localized states.
It is found that, in either two or three dimensions, if the density of states is flat, \( \sigma = \sigma_0 e^{-\frac{\epsilon}{kT}} \). The activation energies are found to be \( \epsilon_a = 0.28\Delta \epsilon \) in two dimensions and \( \epsilon_a = 0.20\Delta \epsilon \) in three dimensions. These values are considerable improvements over the estimates of previous workers, who used the low temperature asymptotic form of the resistance in the high temperature regime.

It is also revealed that \( \sigma \) can be \( \sigma = \sigma_0 e^{-\frac{(T_0/T)^{1/4}}{\epsilon}} \) in the high temperature regime if the density of states decreases with \( |\epsilon - \mu_0| \) for energy \( \epsilon \) far enough away from the zero temperature chemical potential \( \mu_0 \). These results are in accord with the experimental results described above.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xiii</td>
</tr>
<tr>
<td>NOTE REGARDING THE FORMAT OF THIS DISSERTATION</td>
<td>xiv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>I. Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>II. Aims, Assumptions, and Strategies</td>
<td>5</td>
</tr>
<tr>
<td>III. Review of the Literature; Basic Concepts in VRH</td>
<td>10</td>
</tr>
<tr>
<td>(i) The Basic Idea of VRH</td>
<td>11</td>
</tr>
<tr>
<td>(ii) Percolation Approach to Calculating the Conductivity</td>
<td>12</td>
</tr>
<tr>
<td>(iii) Experimental Studies</td>
<td>14</td>
</tr>
<tr>
<td>(iv) Numerical Studies</td>
<td>16</td>
</tr>
<tr>
<td>(v) Analytic Studies</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>19</td>
</tr>
</tbody>
</table>
**PAPER [I]: Two-dimensional Numerical Simulation at High Temperatures**

Abstract ................................................................. 22

I. Introduction .......................................................... 24
II. Comparison of Theory and Experiment for VRH in DSs........... 32
III. The Numerical Simulation ........................................... 47
   (i) General Description ............................................. 47
   (ii) Computational Details ......................................... 51
IV. Results and Discussion ............................................. 53
Acknowledgments .................................................................. 64
References ......................................................................... 65

**PAPER [II]: Two-dimensional Analytic Theory at High Temperatures**

Abstract ................................................................. 68

I. Introduction .......................................................... 70
II. Formulation as a Percolation Problem ............................. 73
III. \( X_{\text{min}} (\xi) \) and \( X_{\text{max}} (\xi) \) ......................... 76
IV. The Boundary Curve .................................................. 79
V. Expressions for \( p(X;\xi) \) ............................................ 82
VI. Temperature Dependence; The Extremal Boundary Curve ...... 86
VII. Summary of Equations ............................................... 90
VIII. Form Chosen for \( B_c (t) \) ......................................... 93
IX. Comparison of Analytic Theory and Numerical Simulation ..... 100
X. Summary ............................................................................. 108
Appendix ....................................................................................... 110
Acknowledgment ......................................................................... 113
References ..................................................................................... 115

PAPER [III]: Three-dimensional, High Temperature Analytic Theory
for Flat Density of States ......................................................... 116
Abstract ....................................................................................... 116
I. Introduction .............................................................................. 118
II. Essential Elements of the Theory in Three Dimensions ...... 123
III. Form Chosen for B (t); Overall Resistance for Flat
Density of States in Three Dimensions ................................. 127
IV. Effect of the Q ij Factor ....................................................... 136
V. Summary .................................................................................. 148
Acknowledgment ......................................................................... 150
References ..................................................................................... 151

PAPER [IV]: High Temperature T⁻¹/₄ Behavior With Small
T₀ Values .................................................................................... 153
Abstract ....................................................................................... 153
I. Introduction .............................................................................. 155
II. The Densities of States ......................................................... 157
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>III. Theory</td>
<td>161</td>
</tr>
<tr>
<td>IV. Results</td>
<td>167</td>
</tr>
<tr>
<td>V. Discussion</td>
<td>173</td>
</tr>
<tr>
<td>VI. Summary and Recommendations for Further Study</td>
<td>176</td>
</tr>
<tr>
<td>Appendix</td>
<td>178</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>180</td>
</tr>
<tr>
<td>References</td>
<td>181</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>183</td>
</tr>
<tr>
<td>I. Summary and Highlights</td>
<td>183</td>
</tr>
<tr>
<td>II. Suggestions for Further Work</td>
<td>190</td>
</tr>
<tr>
<td>References</td>
<td>193</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>194</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>197</td>
</tr>
</tbody>
</table>


LIST OF TABLES

PAPER [I]

<table>
<thead>
<tr>
<th>TABLE I:</th>
<th>Comparison of the Theoretical Values $\xi_{\text{c min}}$ to the Experimental Values $\xi_{\text{c exp}}$ for Lightly Doped n-GaAs</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE II:</td>
<td>Comparisons of the Theoretical Values $\tilde{T}_0^{\text{min}}$ to the Experimental Values $T_0^{\text{exp}}$ and of the Theoretical Values $\tilde{T}_c^{\text{max}}$ to the Experimental Ranges of Temperature for Lightly Doped n-GaAs</td>
<td>40</td>
</tr>
<tr>
<td>TABLE III:</td>
<td>Values of the Localization Lengths $a$ and Bandwidths $\Delta \varepsilon$ Required to Give Agreement Between the Theoretical and Experimental Values of $\xi_{\text{c}}$ and $T_0$ for Lightly Doped n-GaAs</td>
<td>41</td>
</tr>
</tbody>
</table>

PAPER [III]

| TABLE I: | Effect of the $Q_{ij}$ Factor on the Slopes of the $-\xi_{\text{c}}$ vs. $t^{-1}$ Curves | 145 |
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>PAPER [I]</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fig. 1:</strong> Temperature Dependence of $\frac{R_{\text{full}}}{R_{\text{low}}}^{i,j}$ for Typical $\tilde{\varepsilon}_j$ and $\tilde{\varepsilon}_i$</td>
<td>30</td>
</tr>
<tr>
<td><strong>Fig. 2:</strong> Plots of the Average Currents Obtained Numerically vs. $t^{-1/3}$ for 1000-, 2000-, and 2800-site Samples</td>
<td>54</td>
</tr>
<tr>
<td><strong>Fig. 3:</strong> Comparison of Average Currents for the 1000-site Systems and 2000-site Systems, Average Current for the 2800-site Systems and Current for one 2000-site Rectangular System</td>
<td>55</td>
</tr>
<tr>
<td><strong>Fig. 4:</strong> Plots of the Average Currents for the 1000-, 2000-, and 2800-site Systems vs. $t^{-1}$</td>
<td>58</td>
</tr>
<tr>
<td><strong>Fig. 5:</strong> Comparison of Average Current for the Square 2000-site Systems and Currents for two Rectangular 2000-site Systems</td>
<td>61</td>
</tr>
<tr>
<td><strong>Fig. 6:</strong> Comparison of Average Current for the 2000-site Systems and $R_{i,j}^{-1}$</td>
<td>63</td>
</tr>
<tr>
<td>Fig.</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>The Triangle and the Boundary Curve</td>
</tr>
<tr>
<td>2</td>
<td>Four Possible Relative Positions of the Triangle and the Boundary Curve</td>
</tr>
<tr>
<td>3</td>
<td>The Critical Number ( B_c ) of Bonds per Site Required for Percolation as a Function of Temperature</td>
</tr>
<tr>
<td>4</td>
<td>Comparison Between Average Overall Current in Numerical Simulation and Analytic Theory Result Using ( t^{-1} ) Abscissa</td>
</tr>
<tr>
<td>5</td>
<td>Comparison Between Average Overall Current in Numerical Simulation and Analytic Theory Result Using ( t^{-1/3} ) Abscissa</td>
</tr>
<tr>
<td>6</td>
<td>Temperature Dependence of ( X_{max} )</td>
</tr>
<tr>
<td>7</td>
<td>Temperature Dependence of the Typical Hopping Energy</td>
</tr>
</tbody>
</table>
Fig. 8: Comparison Between Exact and Approximate Analytic Theory Curves for $\xi_c$ ........................................... 114

PAPER [III]

Fig. 1: The $-\xi_c$ Curve When Plotted Against $t^{-0.3}$ ................. 132

Fig. 2: Plot Showing Linearity of $-\xi_c$ in $t^{-1}$ ......................... 134

Fig. 3: The Triangle and Various Boundary Curves ......................... 137

Fig. 4: Effect of the $Q_{ij}$ Factor on the Resulting Curve
for $-\xi_c$ vs. $t^{-1/4}$ ................................................. 142

Fig. 5: Effect of the $Q_{ij}$ Factor on the Resulting Curve
for $-\xi_c$ vs. $t^{-1}$ .................................................... 144

PAPER [IV]

Fig. 1: Various Densities of States ................................. 159

Fig. 2: Plots of $-\xi_c$ vs. $t^{-1}$ for the Cases $\lambda = +\frac{1}{2}$, $\lambda = 0$,
$\lambda = -\frac{1}{2}$, Using the "Flat Density of States" $B_c(t)$ ..... 168
Fig. 3: Plots of $-\xi_c$ vs. $t^{-1/4}$ for the Cases $\lambda = \frac{1}{2}$, $\lambda = 0$, $\lambda = -\frac{1}{2}$ ................................................................. 169

Fig. 4: $B_c$ Plotted Against $t^{-1/4}$ ............................................. 171
ACKNOWLEDGMENTS

It is with great pleasure that I thank my supervisor, Professor Robert Barrie, for excellent supervision, many helpful suggestions, and especially for the highly skilled manner in which he provided me with enough guidance that I could find my way yet also with sufficient freedom that I can fully enjoy the successes of this dissertation. It would be an unfortunate oversight if I did not also express my appreciation of Professor Barrie's deep concern for me as his student, and of his unwavering commitment to perform his duties as a supervisor to the utmost of his abilities - a very high level of proficiency, indeed.

I have benefitted tremendously from discussions with numerous professors, postdoctoral fellows, and graduate students at UBC. I wish especially to acknowledge many illuminating discussions with Dr. Peter Holdsworth; fellow students Matthew Choptiuk, Peter Mulhern and Dan Zimmerman must also be mentioned by name.

I was fortunate to receive essential moral support from several fine individuals, most notably Stephanie Mundle.

The list of people who helped me significantly in obtaining my Ph.D. would fill several pages; I regret I cannot mention each of you by name.

This work was supported financially by NSERC Postgraduate Scholarship, NSERC Operating Grant A713, and by the UBC University Graduate Fellowship Programme.

The work comprising the four papers in this dissertation was divided between Professor Barrie and myself approximately as follows:

M.R.A. Shegelski: 95%; R. Barrie: 5%.
Note Regarding the Format of this Dissertation

This dissertation consists of seven major divisions: an Introduction, four Papers, a Conclusion and an Appendix. The four Papers are referred to sequentially as "[I]", "[II]", "[III]", and "[IV]". References in each major division are given at the end of that division.
I. Statement of the Problem

Lightly doped semiconductors can have electrons in localized states around the impurities. A possible mechanism for dc electronic conduction is for these electrons to undergo transitions between localized states. Such transitions are referred to as "hops" and the associated conductivity is called "hopping conductivity". Hopping is the dominant mechanism for conduction if the temperature is so low that the number of conduction band electrons is very much smaller than the number of localized electrons.

Established theories\(^2,^3,^4\) of hopping conduction predict that, at low enough temperature \(T\), the conductivity \(\sigma\) in a lightly doped semiconductor is of the form

\[
\sigma = \sigma_0 e^{-\left(T_0/T\right)^{1/4}},
\]

where the prefactor \(\sigma_0\) depends weakly on \(T\), and \(T_0\) is given by

\[
T_0 = \frac{C}{kg_p a^3},
\]

where \(k\) is Boltzmann's constant, \(g_F\) is the density of localized states at
the Fermi energy, \( a \) is the electronic localization length and \( C \) is a dimensionless constant approximately equal to 20. Hopping conductivity of this form is almost always called "variable range hopping" (VRH) conductivity. The term "variable range hopping" is used because, according to the theories\(^2,3,4\) of hopping conduction, the conductivity has the form (1.1) when the range in position space of typical hops increases as the temperature decreases. The basic ideas of VRH will be developed more fully in Sec. III.

Published experimental work which exhibits clear VRH conduction in doped semiconductors is scarce, as will be discussed in Sec. III. In response to this shortage of data, Benzaquen and Walsh\(^5\) and Benzaquen et al.\(^6\) published experimental accounts of hopping conduction in lightly doped n-GaAs and lightly doped n-InP, respectively. For the temperature range \( 1.4K < T \leq 3-7K \), they found

\[
\sigma = \sigma_0 e^{-\left(\frac{T_0}{T}\right)^s}, \tag{1.3}
\]

with \( s \) close to \( \frac{1}{4} \), \( T_0 \) of the order of \( 10^3K \) in GaAs and of the order of \( 10^4K \) in InP. At first sight, these results appear to support existing theories\(^2,3,4\) of VRH. However, the experimental temperature range was well above the range where the theories predict VRH to be observed in these materials, and the experimental \( T_0 \) values were far too small to be accounted for on the basis of these theories. The theories predict VRH with \( T_0 \) values
of the order of $10^5$K to be observed in these materials for temperatures less than 0.2-0.3K. The predicted behavior was observed by Emel'yanenko et al.\(^7\) in n-GaAs — the impurity concentrations were essentially the same as those in ref. [5] — thereby constituting evidence in favor of the theories. Moreover, although Emel'yanenko et al. did not themselves say so — a point to be fully discussed in [I] — the Emel'yanenko et al. data seemed to indicate a second temperature regime in which $\sigma$ was of the form (I.1). The temperature range for this regime was $1.2K < T < 4.2K$, and the associated $T_0$ values were about $10^3$K, just as reported by Benzaquen and Walsh.

These three experimental accounts\(^5,6,7\) suggest the existence of two $T^{-1/4}$ regimes. The low temperature ($0.1K < T < 1K$) regime is well understood on the basis of the established theories of VRH, but the high temperature ($1K < T < 3-7K$) $T^{-1/4}$ behavior cannot be understood on the basis of these theories. A full discussion of the inapplicability of the theories to the high temperature $T^{-1/4}$ behavior will be given in [I].

Matters become even more interesting upon consideration of other experimental investigations\(^8-11\) into the hopping conductivity of lightly doped n-GaAs and lightly doped n-InP. In contrast to the experimental work discussed above, these experiments revealed a conductivity of the so called "activated" form,

$$\sigma = \sigma_3 e^{-e_3/kT}$$

(1.4)
in the high temperature range, \(1K \leq T \leq 3-7K\). \(^{12}\) (The notations \(\sigma_3\) and \(\epsilon_3\) are used for historical reasons\(^{12}\).) In these various experiments\(^{5-11}\), strikingly different results for the conductivity have been reported for the same materials, even though the temperature ranges and the impurity concentrations were essentially the same. Moreover, theories have been advanced\(^{13}\) which have predicted the activated conductivity (I.4) for the high temperature regime, but such theories have not suggested the alternative possibility of a conductivity of the form (I.1).

Some intriguing questions emerge. Why, for the same material, and in the same temperature range, have some investigators found the conductivity to be of the form (I.3) with \(s \approx \frac{1}{4}\) while others have found it to be of form (I.4)? The conductivity (I.3) with \(s \approx \frac{1}{4}\) has unequivocally been observed in ref.s [5] and [6], and possibly also in ref. [7] (see [1]). However, the experimental temperature range was well in excess of, and the experimental \(T_0\) values were well below, the limits set by established theories of VRH. How can such experimental results be explained? Why is it that these experimental results were not admitted as a possibility by those theories\(^{13}\) which predicted the conductivity (I.4)?
II. Aims, Assumptions, and Strategies

The principal aim of this dissertation is to answer the questions raised at the end of the previous section. The investigations to be undertaken will be based on an extension of the model underlying almost all those theories which have had success in accounting for various experimentally observed features of hopping conduction. The model referred to is the resistor network model for hopping conduction, first proposed by Miller and Abrahams. In this model, a lightly doped semiconductor (LDS) is represented by a resistor network wherein every pair of donor impurity sites \(i\) and \(j\) is mapped into a resistor \(R_{ij}\) (for definiteness, a compensated, n-type semiconductor is considered). The overall conductivity of the LDS is given in terms of the overall resistance of the resistor network. In the Appendix, the mapping of a LDS into a resistor network is discussed. That a resistor network can model the hopping conduction is readily seen as follows. The hopping of electrons from donor to donor can be described in terms of currents flowing between the donors. Evaluating these currents gives the resistances associated with the hops. A derivation is presented in the Appendix of the following form of the resistance \(R_{ij}\):

\[
R_{ij} = R_0 Q_{ij} f_{ij} X_{ij}^{-2} e^{X_{ij}}, \quad (II.1)
\]
where

\[ R_0 = \frac{6\pi d s^5 \hbar^4}{e^2 E_d^2} \left( \frac{3\kappa a}{2k_0 e^2} \right)^2, \]  
(II.2)

\[ Q_{ij} = \left[ 1 + \left( \frac{\varepsilon_i - \varepsilon_j}{3} \right) a^2 \right]^4, \]  
(II.3)

and

\[ f_{ij} = \frac{kT}{\varepsilon_j - \varepsilon_i} \left( 1 + \frac{1}{2} \varepsilon_i \right) \left( 1 + \frac{1}{2} \varepsilon_j \right) \left( 1 + \frac{1}{2} \varepsilon_i \right) \left( 1 + \frac{1}{2} \varepsilon_j \right) - 1. \]  
(II.4)

d, s, E_d, and \( \kappa \) are, respectively, the semiconductor's density, its speed of sound, its deformation potential constant, and its dielectric constant; \( \hbar = \frac{\hbar}{2\pi} \), where \( \hbar \) is Planck's constant, \( k_0 \) is the Coulomb force constant, \( \mu \) is the chemical potential, \( \varepsilon_i \) and \( \varepsilon_j \) are the electron energies corresponding to the localized states for majority impurity sites \( i \) and \( j \), with \( \varepsilon_j > \varepsilon_i \) chosen for definiteness, and \( X_{ij} = \frac{2r_{ij}}{a} \) where \( r_{ij} \) is the distance separating sites \( i \) and \( j \) and \( a \) is the electron localization length. The major assumptions involved in deriving (II.1) - (II.4), as discussed in the Appendix, are as follows.

The Coulomb interaction between localized electrons and charged impurities is taken into account in only a very simple way. At low
temperature, an n-type semiconductor of unit volume with \( n_D \) donors and \( n_A \) acceptors \( (n_D > n_A) \) will have \( n_D - n_A \) electrons localized about \( n_D \) donors. Each donor site will have a particular and unique environment of negatively charged acceptors, positive donors and neutral donors. The combination of the Coulomb interaction and the unique environment for each donor site implies that the energy level for an electron localized about a donor will differ from site to site. This very complex situation is dealt with in the resistor network model by simply treating the electrons as independent and bringing into play a one-electron density of states \( g(\varepsilon) \) for the localized electrons.

Electron-electron correlations are ignored: the resistance \( R_{ij} \), which models the rate at which electrons hop between sites \( i \) and \( j \), is taken to depend only on quantities pertaining to sites \( i \) and \( j \). No consideration is afforded to the way in which electrons on surrounding sites will affect the hopping rate.

A time-averaged rate equation approach is employed. Furthermore, the donor and acceptor positions, as well as the donor energy levels for electrons, are assumed to be uncorrelated. Finally, the electron-phonon interaction is treated via the deformation potential approximation.

The above assumptions were common to most investigations in hopping conduction. In these investigations, the \( T \to 0 \) asymptotic form of the resistor network (II.1) - (II.4) was almost always used. In some investigations, where the full form for \( R_{ij} \) was used (to be discussed in
Sec. III), the temperature dependence of the conductivity was examined only for $kT$ much smaller than the bandwidth $\Delta \varepsilon$ of localized states. The work in this dissertation is unique in that, for the first time, the hopping conductivity is examined using the full form of the resistor network (II.1) - (II.4), and in a temperature range where $kT$ is of the order of $\Delta \varepsilon$.

The reasons for using the full resistor network, and not the low $T$ asymptotic form, and for taking $kT$ to be of order $\Delta \varepsilon$, are as follows. The bandwidth $\Delta \varepsilon$ of localized states is of the order of the mean electrostatic potential energy shared by a localized electron and a nearby charged impurity, or about 1 meV in the GaAs and InP samples of ref.s [5] and [6]. For the experimental range of temperatures (Sec. I), $kT$ lies in the range $0.1 \text{ meV} \lesssim kT \lesssim 0.6 \text{ meV}$. As such, the assumption $kT \ll \Delta \varepsilon$, made by other workers, is not valid for these experiments. It will become evident in this dissertation that $kT$ must be a much smaller fraction of $\Delta \varepsilon$ than 0.6 if the $T \to 0$ form of the resistor network is to be trusted.

It is necessary to use the full form of the resistance in order to perform a meaningful investigation of the conductivity if the resistor network model is used when $kT$ is of order $\Delta \varepsilon$. As such, it is a principal aim of this dissertation to develop a suitable means of evaluating the net resistance of the resistor network (II.1) - (II.4) for $kT$ of order $\Delta \varepsilon$. It will then be feasible to investigate why it is possible to obtain a high temperature experimental hopping conductivity of the form (I.3) with $s = \frac{1}{4}$, and why some investigators found $\sigma$ to be given by (I.4) instead.
The work in this dissertation is divided into four sequential studies. In [I], a two-dimensional numerical simulation is carried out to calculate the net resistance of the full resistor network. A two-dimensional analytic theory for evaluating the overall resistance of the network is presented in [II], and the results are compared to the numerical results of [I]. Based on the knowledge gained for the two-dimensional case, the analytic theory is extended to three dimensions in [III]. A flat density of states is assumed, for simplicity, in [I], [II], and [III]. The high temperature conductivity is calculated, using the analytic theory, for various fundamentally different densities of states in [IV].

The most important of the results of this work are as follows. A means is devised for dealing with the full resistor network at high temperature. The conductivity is found to be of the form (1.4) for a flat density of states, in either two or three dimensions. In three dimensions, \( \sigma \) can very plausibly be of the form (1.3) with \( s = \frac{1}{4} \) if the density of states decreases with \( |\varepsilon - \mu_0| \) for \( \varepsilon \) far enough away from the zero temperature chemical potential \( \mu_0 \). These results help us to understand why Benzaquen and co-workers observed \( T^{-1/4} \) behavior whereas other investigators obtained \( T^{-1} \) behavior, as will be discussed fully in [IV].

Before proceeding to the main body of this dissertation, a review is presented of the relevant literature, and the basic ideas involved in VRH are discussed.
III. Review of the Literature; Basic Concepts in VRH

The main topic of this dissertation is VRH at high temperatures. As such, the literature reviewed will be predominantly concerned with VRH. Studies of VRH were, however, preceded by investigations into activated hopping conduction. It is therefore best to briefly review the major works in activated hopping conduction before turning to those concerning VRH.

Hopping conduction was first predicted by Gudden and Schottky\textsuperscript{15} in 1935. The first experimental data obtained was the 1946 work of Bush and Labhart\textsuperscript{16} on SiC. In 1954, Hung and Gliessman\textsuperscript{17} studied Ge and Si. Many other experimental investigations of hopping conductivity appeared in the last half of the 1950s and in the early 1960s. The hopping conductivity in these experiments was found to be of the activated form (I.4). More recent experimental studies which showed $\sigma$ of this form were those of ref.s [8]-[11] on n-GaAs and n-InP.

In 1960, Miller and Abrahams\textsuperscript{14} developed the resistor network model of hopping conduction. Their approach to calculating the net resistance of the network was, however, subsequently criticized. Improved approaches were based on percolation theory. Percolation-based theories which calculated $\varepsilon_3$ in terms of the bandwidth $\Delta\varepsilon$, for intermediate compensations, were given by Skal, Shklovskii and Efros\textsuperscript{18} and by Hayden and Butcher\textsuperscript{19}. These theoretical accounts will be further discussed in [I] and [III], where a similar but improved calculation for $\varepsilon_3$ in terms of $\Delta\varepsilon$ will be presented.
(i) The Basic Idea of VRH

The possibility of VRH was first pointed out by Mott in 1968. His derivation of the conductivity (I.1) was essentially as follows. The rate at which an electron will hop from site $i$ to site $j$ is proportional to the quantity $p_{ij} = e^{-\left(\frac{2r_{ij}}{a} + \frac{\epsilon_{ij}}{kT}\right)}$, where $\epsilon_{ij}$ is a particular energy difference associated with the hop. For example, if $\epsilon_j > \epsilon_i$, a phonon is absorbed in the hop, and $\epsilon_{ij} = \epsilon_j - \epsilon_i$. If $kT$ is of order $\Delta\epsilon$ and the average distance between sites is much larger than $a$, then for almost all pairs $i$ and $j$, $2r_{ij}/a \gg \epsilon_{ij}/kT$. In this case, electrons will hop between nearest neighbors in order to minimize $2r_{ij}/a + \epsilon_{ij}/kT$ and thereby maximize $p_{ij}$. However, if $kT \ll \Delta\epsilon$, then for a site $j$ which is a nearest neighbor of site $i$, it will typically be that $\epsilon_{ij} \gg kT$ and $\epsilon_{ij}/kT \gg 2r_{ij}/a$. Under these conditions, an electron on site $i$ could very well find it more favorable to hop to a site $j'$ which is further away from $i$ than $j$ is. Such would be the case if $\epsilon_{ij'}$ was sufficiently smaller than $\epsilon_{ij}$ that, even though $r_{ij'} > r_{ij}$, $p_{ij'} > p_{ij}$.

Consider this situation more fully. First, note that $p_{ij}$ can vary over an enormous range. As such, not all of the pairs of sites in the system are important. Which pairs of sites will be the dominant pairs and will thereby determine the overall conductivity? To answer this, restrict attention to the set of sites whose energies lie within $\epsilon_0$ of $\mu$. Sites with energies close to $\mu$ are focused on because $kT \ll \Delta\epsilon$ implies sites far from $\mu$ are
either completely full or completely empty. The aim is to determine the value of $\varepsilon_0$ that corresponds to the set of sites yielding the largest conductivity. To do so, consider a randomly chosen site $i$ where $|\varepsilon_i - \mu| < \varepsilon_0$. Imagine a sphere centered on $i$ which is just large enough that there is a second site $j$, for which $|\varepsilon_j - \mu| < \varepsilon_0$, which also lies within the sphere. Define $r_0$ as the radius of such a sphere averaged over all sites $i$ with $|\varepsilon_i - \mu| < \varepsilon_0$. Notice that $r_0$ will be a measure of the hopping distance, and $\varepsilon_0$ will be a measure of the hopping energy. $\varepsilon_0$ is easily seen to be related to $r_0$ by $\varepsilon_0 \sim \frac{1}{g_F r_0^3}$, where $g_F$ is the density of states at the Fermi energy (ignore all numerical factors in this heuristic argument). The typical value of $p_{ij}$ will be $p \sim e^{-\frac{r_0}{a + \frac{1}{g_F kT r_0^3}}}$. Maximizing $p$ with respect to $r_0$ gives $r_0 \sim a (g_F kT a^3)^{-1/4}$, which is a measure of the typical hopping distance, $\varepsilon_0 \sim kT (g_F kT a^3)^{-1/4}$, which is a measure of the typical hopping energy, and $p \sim e^{-\frac{r_0}{a + \frac{1}{g_F kT r_0^3}}}$. which implies a conductivity of the form (I.1). $T_0$ is given by (I.2), and $C$ is an undetermined numerical factor. Note that $r_0$ increases as $T$ decreases, whence the name "variable range hopping" conductivity.

(ii) Percolation Approach to Calculating the Conductivity

The above is only a heuristic argument for VRH. Subsequent investigators were able to provide more rigorous derivations. The best
known of these works are those of Ambegaokar, Halperin, and Langer\textsuperscript{3} (AHL) and Pollak\textsuperscript{4}. Both works were based on a percolation theory approach to calculating the net resistance of the Miller and Abrahams resistor network. It was recognized in these theories that, because of the enormous range over which the resistance $R_{ij}$ could vary, only a small fraction of the huge number of pairs of sites could be of any importance in determining the overall conductivity. In both theories, it was assumed that $kT$ was so much smaller than $A\varepsilon$ that the resistance $R_{ij}$ could be replaced by its $T \to 0$ asymptotic form for essentially all pairs of sites of importance. In ref. [3], the following approach was presented. Imagine removing all the resistors. Put back those for which $R_{ij} < R$. Start with $R$ arbitrarily small; what happens as $R$ increases? At first, only isolated clusters of resistors appear, and there is no continuous path from one end of the network to the opposite end. As $R$ increases, the clusters grow in size. At some critical value $R_C$ of $R$, the clusters merge, being connected by resistors of magnitude $R_C$, and the set of resistors present "percolates" from one end of the system to the other. Since the resistors vary over an exponential range, the resistors in the clusters are exponentially small compared to the critical resistors. Moreover, as further resistors are introduced, they are shorted out by the network formed by the condition $R_{ij} < R_C$. As such, the overall conductivity is determined by the critical resistors. Writing

$$R_C = R_0 e^{\xi_C} \quad (III.1)$$
it follows that

\[ \sigma = \sigma_0 e^{-\xi_C}. \]  

(AHL) found

\[ \xi_C = \left( \frac{T_0}{T} \right)^{1/4}, \]  

with \( T_0 \) given by (I.2); they estimated \( C \) to be about 16. Pollak's analysis also gave equations (III.2), (III.3) and (I.2), with \( C \) estimated to be about 11.5. Other workers\(^{20,21}\) have also provided estimates for \( C \). The various estimates available range from about 10 to about 28. The most recent calculation\(^{22}\) gives

\[ C = 21.2 \pm 1.2. \]  

The AHL derivation of (III.3) will be presented in [I].

(iii) Experimental Studies

A summary of investigations into hopping conductivity in doped semiconductors will be given next. Before doing so, it is important to emphasize that this dissertation is concerned almost exclusively with VRH in lightly doped semiconductors (LDSs). The only experimental work on VRH in
LDSs has already been discussed. Nevertheless, so as to convey the extent to which VRH has been studied in doped semiconductors, the following brief summary is presented.

The reports of Benzaquen and Walsh, Benzaquen et al. and Emel'yanenko et al. constitute the only experimental investigations of VRH in LDSs. Allen and co-workers presented evidence for VRH between 0.2K and 1K in heavily doped n-Ge with intermediate compensation \((n_A/n_D \text{ not too close to } 0 \text{ or } 1)\). Shlimak and Nikulin examined p-Ge at intermediate compensation for \(0.1K \leq T \leq 1K\); they found VRH for the sample with the lightest doping, but claimed that the value of \(s\) in equation (1.3) for \(\sigma\) increased with impurity concentration. Yaremenko reported VRH for heavily doped n-InSb with strong compensation \((n_A/n_D \text{ just less than unity})\) for \(4K \leq T \leq 10K\), although the data presented revealed some scatter. On the other hand, Redfield, Zabrodskii, and Tokumoto et al. obtained conductivities of form (1.3) with \(s = \frac{1}{2}\), Redfield for heavily doped, strongly compensated n-GaAs with \(2K \leq T \leq 20K\), Zabrodskii for heavily doped, strongly compensated n-Ge with \(1K \leq T \leq 10K\), and Tokumoto et al. for heavily doped n-InSb with intermediate compensation and \(0.03K \leq T \leq 0.3K\) (in the last case, it was the magnetoresistance that was studied). Investigations have also been carried out on hopping conductivity in semiconductors doped so heavily as to be barely on the insulating side of the metal-insulator transition. Kobayashi and co-workers, as well as Shafarman and Castner, reported VRH in uncompensated n-Si, the former authors for
2K \lesssim T \lesssim 10K, the latter for 0.15K \lesssim T \lesssim 10K. Ionov et al.\textsuperscript{32} observed the conductivity (I.3) with \( s = \frac{1}{2} \) in intermediately compensated n-Ge for 0.5K \lesssim T \lesssim 2K.

It is evident that significant diversity exists in the experimental literature, and that the study of VRH in doped semiconductors will be of great interest for some time to come. In this dissertation, only the discrepancies for high temperature hopping conduction in LDSs will be investigated. As such, the only experimental results which relate to this dissertation are those of ref.s [5] - [11], all of which are for LDSs.

It must be mentioned that theoretical work exists\textsuperscript{33} which predicts equation (I.3) with \( s = \frac{1}{2} \) for the low temperature conductivity in doped semiconductors. This theoretical work is not, however, related in a major way to the work in this dissertation.

(iv) Numerical Studies

Another area of the literature of direct interest concerns certain numerical simulations. These simulations were designed to determine the net resistance of a resistor network like the network (II.1) - (II.4). Prior to this dissertation, no thorough study of \( \sigma(T) \) at high \( T \) (\( kT \) of order \( \Delta \varepsilon \)) for the network (II.1) - (II.4) existed. However, some work has been done which is related to such a study.

Numerical works concerning the resistor network
include those by Ambegaokar, Cochran, and Kurkijarvi\textsuperscript{35}, and Seager and Pike\textsuperscript{20}. The simulation carried out by Ambegaokar et al. established the validity of the percolation approach to calculating the net resistance of a random resistor network with exponentially varying resistances. These works will be referred to at the appropriate places in this dissertation.

The exponential part of the low temperature asymptotic form of the resistor network (II.1) is

\[
R_{ij} = R_0 e^{2r_{ij}/a} \quad (III.5)
\]

\[
R_{ij} = R_0 e^{\frac{2r_{ij}}{a} + \frac{|\varepsilon_i - \mu| + |\varepsilon_j - \mu| + |\varepsilon_i - \varepsilon_j|}{2kT}}. \quad (III.6)
\]

Maschke, Overhof, and Thomas have performed computer simulations for the network (III.6) for one dimension\textsuperscript{36} and two dimensions\textsuperscript{37}. They verified that \( \ln \sigma \) goes as \( T^{-1/3} \) at low temperatures in two dimensions. For high temperatures and a flat density of states, they stated that \( \ln \sigma \) varies as \( T^{-1} \), but offered no further information.

Some numerical work has been done, by Seager and Pike\textsuperscript{20} and by McInnes and Butcher\textsuperscript{38}, using the full Miller and Abrahams resistor network. Except to say that \( \ln \sigma \) varied as \( T^{-1} \) at high temperatures if the density of states was flat, these works dealt exclusively with the low temperature regime. Again, these works will be referred to as required.
(v) Analytic Studies

Some work was done by Butcher and co-workers\textsuperscript{39} in an effort to construct an analytic theory for the full Miller and Abrahams resistor network. The theory was constructed for a flat density of states, was developed in only an approximate form, and was valid for only low temperatures. No attempt was made to determine the high temperature conductivity, or the dependence of the conductivity on the density of states.
References

1 For definiteness, the case of a compensated n-type semiconductor is considered. See, for example, N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, Philadelphia, 1976), ch. 28 for an introduction to semiconductors.


The prefactor and energy in equation (1.4) are denoted $\sigma_3$ and $\epsilon_3$ respectively for historical reasons. In the early days of hopping conductivity, it was found for decreasing temperature that there were three regimes for which $\ln \sigma \sim T^{-1}$; the associated activation energies were denoted $\epsilon_1$, $\epsilon_2$, and $\epsilon_3$, respectively.


A. Miller and E. Abrahams, Phys. Rev. 120, 745 (1960).


29 See, for example, N.F. Mott, Metal-Insulator Transitions (Taylor and Francis, London, 1974).
Hopping conductivity in lightly doped semiconductors - I: 
Two-dimensional numerical simulation at high temperatures

Mark R.A. Shegelski* and Robert Barrie
Department of Physics, University of British Columbia
Vancouver, B.C., CANADA V6T 2A6

Abstract

We point out the existence of experimental results for lightly doped n-GaAs and n-InP in which the conductivity $\sigma$ is given by $\sigma = \sigma_0 e^{-(T_0/T)^s}$ with $s$ close to $\frac{1}{4}$, indicating variable range hopping. We show that the experimental range of temperatures $T$, $1K \lesssim T \lesssim 7K$, in which this behaviour holds, is well in excess of the temperature range predicted for these materials by established theories of variable range hopping, and that the experimental $T_0$ values are too small to be understood on the basis of these theories. As a first step toward understanding these experimental results, we present in this first paper the results of a numerical simulation of a two-dimensional lightly doped semiconductor. We choose a flat density of states with width $\Delta \varepsilon$. We model the semiconductor as a Miller and Abrahams type resistor network; we use the full form of the resistance and do not take the low temperature asymptotic form because we carry out the simulation
at temperatures for which $kT$ is of order $\Delta \varepsilon$. We find that, for a flat density of states and two dimensions, there is no extensive temperature range in which $\sigma = \sigma_0 e^{-(T_0/T)^s}$ with $T_0$ substantially smaller than the theoretical $T_0$ for variable range hopping, and with $s$ close to $\frac{1}{3}$ (in two dimensions, $s$ for variable range hopping is $\frac{1}{3}$). Instead, we find a wide temperature range for which $\sigma = \sigma_0 e^{-\varepsilon_3/kT}$ with $\varepsilon_3 = 0.28\Delta \varepsilon$. This value of $\varepsilon_3$ is considerably smaller than the value found by Hayden and Butcher. We believe that the difference between our result and theirs may be attributed to their use of the low temperature form of the resistance ($kT \ll \Delta \varepsilon$) in a temperature range in which $kT$ is of order $\Delta \varepsilon$.

PACS numbers: 72.20.-i
1. Introduction

In disordered systems such as amorphous or doped semiconductors, electrons can exist in localized states and conduction can occur by electrons hopping from one localized state to another. Mott\textsuperscript{1} was first to suggest that, for a density of states which is flat in the vicinity of the Fermi level and for a low enough temperature $T$, the d.c. electric conductivity $\sigma$ of such a system of random localized electrons would be given by

\[ \sigma = \sigma_0 e^{-\left(\frac{T_0}{T}\right)^{1/4}}, \]  

(1)

with

\[ T_0 = \frac{C}{k g_F a^3}; \]  

(2)

the prefactor $\sigma_0$ is weakly dependent on $T$, $k$ is Boltzmann's constant, $g_F$ is the density of states at the Fermi level, $a$ is the electronic localization length and $C$ is a dimensionless constant of order unity. Subsequent investigators\textsuperscript{2,3} were able to obtain the same result and also estimate $C$. The physical idea underlying equation (1) is that, as $T$ is lowered, the range of electron hops increases. In consequence, systems whose conductivities are given by equation (1) are usually said to exhibit "variable range hopping" (VRH).
VRH has been experimentally observed in amorphous semiconductors (ASs) by many investigators. Although one cannot use theoretical means to accurately estimate $a$ and $g_F$ for an AS, one can obtain values for $a$ and $g_F$ by equating the theoretical value for $T_Q$ to the experimentally observed value. Since the values of $a$ and $g_F$ obtained in this manner are very reasonable, it must be said that the theoretical models$^{1-3}$ for VRH agree very well with the experimental results for ASs, at least so far as $T_Q$ goes. Agreement between the theoretical and experimental values for $\sigma_0$ was not obtained with much success until fairly recently, when two independent models$^{4,5}$ were presented. Recent experimental work$^6$ selects one of the two models as providing a better description of the experimental results than the other. It would therefore appear that the experimental results for VRH in ASs are well understood on the basis of established theories.

Hopping conduction has also been studied experimentally in doped semiconductors (DSs). Except for some fairly recent experimental results$^7,8$ which show unequivocal VRH behavior, it is for only a few of the experimental studies on DSs that the conductivity has been found to be given by equation (1). (For a brief review of the experimental situation, see ref. [7]). These recent experimental results, however, have the intriguing feature that the experimental value for $T_0$ is far too small to be understood on the basis of the existent theoretical models. Moreover, the experimental temperature range is well above the highest temperature $T_C$ for which one
would expect, on the basis of existent theories, to have \( \sigma \) given by equation (1). We discuss these claims in detail in Sec. II.

We have set upon a program to understand how, in a temperature range well in excess of the theoretical temperature \( T_c \), the experimental conductivity of a DS can be of the form shown in equation (1), but also have an experimental value of \( T_0 \) which is well below the theoretical value given in equation (2). Our program is based on an extension of those theories which have been so successful in explaining the experimentally observed conductivity in ASs. These theories calculate the conductivity of a system of localized electrons by reducing the problem to that of determining the overall resistance of a random resistor network. Such a network is formed by associating a resistor \( R_{ij} \) with each pair of one-electron localized states \( \psi_i \) and \( \psi_j \). The ends of \( R_{ij} \) are located at the centers \( \hat{r}_i \equiv \langle \psi_i | \hat{r} | \psi_i \rangle \) and \( \hat{r}_j \equiv \langle \psi_j | \hat{r} | \psi_j \rangle \) of the two states. \( R_{ij} \) is given by

\[
R_{ij} = R_0^{ij} e^{\xi_{ij}} \quad (i)
\]

\[
\xi_{ij} = \frac{2r_{ij}}{a} + \left( \frac{|\epsilon_i - \mu| + |\epsilon_j - \mu| + |\epsilon_i - \epsilon_j|}{2kT} \right) \quad (ii)
\]

where \( R_0^{ij} \) is a resistance which depends algebraically on \( r_{ij} \), \( \epsilon_i \) and \( \epsilon_j \), \( r_{ij} \equiv |\hat{r}_i - \hat{r}_j| \), \( \epsilon_i \) and \( \epsilon_j \) are the one-electron energies associated with
states $\psi_i$ and $\psi_j$, and $\mu$ is the chemical potential. This expression for $R_{ij}$ is actually a simplified, low-temperature form of a more general expression first obtained by Miller and Abrahams. Since our investigations are for a higher temperature range, we base our program of study on a resistor network of a more general form:

$$R_{ij} = R_0 Q_{ij} f_{ij} X_{ij}^{-2} e^{X_{ij}}, \quad (4)$$

where

$$R_0 = \frac{4\pi d s^5 h^4}{e^2 E_d^2} \left( \frac{3k a}{2k_0 e^2} \right)^2, \quad (5)$$

$$Q_{ij} = \left[ 1 + \frac{(\varepsilon_i - \varepsilon_j)a^2}{2\pi s} \right]^4, \quad (6)$$

and

$$f_{ij} = \frac{kT}{\varepsilon_j - \varepsilon_i} \left( 1 + \frac{1}{2} e^{(\varepsilon_i - \mu)/kT} \right) \left( 1 + \frac{1}{2} e^{-(\varepsilon_j - \mu)/kT} \right) \left( 1 + \frac{1}{2} e^{(\varepsilon_j - \varepsilon_i)/kT} \right) \left( 1 + \frac{1}{2} e^{-(\varepsilon_j - \varepsilon_i)/kT} \right) - 1. \quad (7)$$

d is the density of the DS, $s$ its speed of sound, $\pi = \frac{h}{2\pi}$, $h$ is Planck's constant, $e$ is the electronic charge, $E_d$ is the deformation potential.
constant, \( \kappa \) is the dielectric constant, \( k_0 \) is the Coulomb force constant, \( \varepsilon_j > \varepsilon_i \) in these equations, and \( X_{ij} \equiv 2r_{ij}/a \). This form of \( R_{ij} \) is obtained by using the deformation potential approximation for the electron-phonon interaction. It is easy to verify that \( R_{ij} \) as given by equations (3) results by taking the low \( T \) limit of equation (4).

Our program to understand the small \( T_0 \) values observed experimentally is based on equations (4) to (7), and takes the form of four sequential investigations:

1. We perform a numerical simulation in two dimensions to calculate the overall resistance of the resistor network (4). To simplify matters, we replace \( Q_{ij} \) by unity and we take the number density of donors \( n_D \) to be twice the number density of acceptors \( n_A \). \( n_D = 2n_A \) leads to \( \mu = \mu_0 - kT \ln 2 \) (\( \mu_0 \) is the chemical potential at \( T = 0 \)). We choose a flat density of states with bandwidth \( \Delta \varepsilon \) and take \( \mu_0 = 0 \). The network simplifies to

\[
R_{ij} = R_0 f(E_i, E_j) X_{ij}^2 e^{X_{ij}},
\]

where \( R_0 \) is a constant resistance and

\[
f(E_i, E_j) = \frac{E_i - E_j}{(1+e^{-E_i})(1+e^{-E_j})(e^{E_j-E_i}-1)},
\]
\[ E_i \equiv \frac{\varepsilon_i}{t}; \tilde{E}_i \equiv \frac{\varepsilon_i}{\Delta \varepsilon}; t \equiv \frac{kT}{\Delta \varepsilon}; X_{ij} \equiv \frac{2\tau_{ij}}{a}. \]  

(10)

Again, \( E_j > E_i \). It is essential to note that \( t \) will be of order unity in our investigations. As such, it is crucial that we use the full form (8) for \( R_{ij} \), and not the low temperature asymptotic form. This is because the ratio of these two expressions, \( \frac{R_{ij}^{\text{full}}}{R_{ij}^{\text{low}}} \), deviates appreciably from unity unless \( t \) is very small. We show this in Fig. 1 by plotting \( \frac{R_{ij}^{\text{full}}}{R_{ij}^{\text{low}}} \) vs. \( t^{-1/3} \) for typical \( \tilde{\varepsilon}_i \) and \( \tilde{\varepsilon}_j \) and \( 1 < t^{-1/3} < 3 \). (The reasons for choosing this specific \( t \)-range and for choosing the abscissa to be \( t^{-1/3} \) will become evident in Sec. IV of this paper.) The upper curve is a typical curve for \( \tilde{\varepsilon}_i \tilde{\varepsilon}_j < 0 \); the lower curve is a typical curve for \( \tilde{\varepsilon}_i \tilde{\varepsilon}_j > 0 \).

2. We develop an analytic theory to calculate the overall resistance of the resistor network (8) for \( t \) of order unity. We first develop the theory for two dimensions and compare the results to those of the numerical simulation.

3. We extend the analytic theory to three dimensions, bringing back into consideration the \( Q_{ij} \) factor.
Fig. 1: Temperature dependence of $R_{ij}^{\text{full}}/R_{ij}^{\text{low}}$ for typical $\tilde{\varepsilon}_j$ and $\tilde{\varepsilon}_i$. The upper curve is for $\tilde{\varepsilon}_j = 0.1$, $\tilde{\varepsilon}_i = -0.1$; the lower curve is for $\tilde{\varepsilon}_j = 0.2$, $\tilde{\varepsilon}_i = 0.1$. 
4. We examine the predictions of the analytic theory for various densities of states and compare our findings to the experimental results.

The first of these investigations will be reported in this paper; the other three will be the subjects of forthcoming papers.

In Sec. II of this paper, we discuss briefly the experimental situation for hopping conduction in DSs and explain why the experimental results of refs. [7] and [8] may not be understood on the basis of existent theories. In Sec. III, we outline how we did our numerical simulation of the resistor network (8). We present and discuss our results in Sec. IV.
II. Comparison of Theory and Experiment for VRH in DSs

In this section, we compare the theoretical value for $T_0$ to the value observed experimentally by Benzaquen and Walsh\textsuperscript{7} for lightly doped GaAs.

The basic method underlying the theoretical $T_0$ value is the percolation method\textsuperscript{11}. To calculate the overall conductivity of the resistor network (3) using the percolation method, one imagines removing all of the resistors except those which satisfy

$$\xi_{ij} < \xi$$

(11)

for some arbitrarily chosen value of $\xi$. The remaining network of resistors will percolate only if $\xi$ is equal to or exceeds a critical value $\xi_c$. According to percolation theory, the overall conductivity $\sigma$ of the original network is given by

$$\sigma = \sigma_0 e^{-\xi_c \xi}.$$  

(12)

From equations (3(ii)) and (11), it follows that all pairs of sites connected by a resistor satisfy

$$r_{ij} < r_{\text{max}} \equiv \frac{a}{2} \xi,$$

(13)

and all sites in the percolating piece have energy $\varepsilon_i$ within $\varepsilon_{\text{max}}$ of $\mu$: 
Pairs of sites which satisfy relation (13) are "spatially eligible" for connection and a site which satisfies relation (14) is "energetically eligible" for connection to another site, but the criterion for connection is the inequality (11).

For a density of states which is flat within $\frac{1}{2} \Delta \varepsilon$ above and below $\mu$, the number density $\rho_\varepsilon$ of energetically eligible states is

$$\rho_\varepsilon = 2g_F \varepsilon_{\text{max}}$$

so long as $\varepsilon_{\text{max}} < \frac{1}{2} \Delta \varepsilon$. If $\rho_\varepsilon \varepsilon_{\text{max}}^3$ is large enough, the set of resistors will percolate: the condition for percolation ($\xi = \xi_c$) is that

$$\rho_\varepsilon \varepsilon_{\text{max}}^3 = n_c,$$

where $n_c$ is the critical number of sites required for percolation. This leads easily to

$$\xi_c = \left( \frac{T_0}{T} \right)^{1/4},$$

where
Equations (12), (17) and (18) are equivalent to equations (1) and (2).

Before comparing the theoretical $T_0$, equation (18), to the experimental value, we note some very important consequences of the above derivation.

First, $\varepsilon_{\text{max}} < \frac{1}{2} \Delta \varepsilon$ implies $T < T_c$ where

$$T_c = \frac{\Delta \varepsilon}{k} \left( \frac{g_F \Delta \varepsilon}{4n_c} \right)^{1/3}.$$  \hfill (19)

This means that when $T$ begins to exceed $T_c$, the curve for $\ln \frac{\sigma}{\sigma_0}$ vs. $T^{-1/4}$ will begin to deviate from linearity: strictly speaking, the theory predicts $\ln \frac{\sigma}{\sigma_0} = - \left( \frac{T_0}{T} \right)^{1/4}$ for $T < T_c$. However, as pointed out by Pollak\textsuperscript{3} and again by Pollak et al.\textsuperscript{12}, because the dominant hops involve energy differences considerably smaller than $\varepsilon_{\text{max}}$, it is to be expected that the curve for $\ln \frac{\sigma}{\sigma_0}$ vs. $T^{-1/4}$ will continue to be linear even for $T > T_c$. In Sec. IV, we will see that, when the full resistor network (4) is used, this linearity extends, for two dimensions, no further than about $2T_c$. In a numerical simulation similar to ours but for three dimensions, and using the full Miller and Abrahams resistor network, Seager and Pike\textsuperscript{13} also found that the linearity extended to about $2T_c$. Pollak et al.\textsuperscript{12} found the linear regime,
for three dimensions, to extend to even higher temperatures, but their results were obtained using the simplified form (3) for the resistors. We will, therefore, take the VRH regime to be confined to \( T < \alpha T_c \) with \( \alpha \approx 2 \).

Second, we must have \( n_D \Delta \varepsilon < n_D \), which, together with \( T < \alpha T_c \), implies that

\[
\xi_c > \frac{1}{\alpha^{1/4}} \frac{2n_1^{1/3}}{n_D^{1/3}} \equiv \frac{1}{\alpha^{1/4}} \xi_c \min .
\]  

(20)

We also have

\[
T_c < \frac{n_D^{1/3}}{k} \frac{1}{4n_c^{1/3}} \frac{\Delta \varepsilon}{\alpha} = \frac{\Delta \varepsilon}{2k \xi_c \min} \equiv T_c \max ,
\]  

(21)

as well as

\[
T_0 > \frac{\Delta \varepsilon}{2k} (\xi_c \min)^3 \equiv T_0 \min .
\]  

(22)

In summary, the theory predicts that, for a flat density of states and \( T < \alpha T_c \max \), \( \sigma = \sigma_0 e^{-(T_0/T)^{1/4}} \) and \( \xi_c > \frac{1}{\alpha^{1/4}} \xi_c \min \) with \( T_0 > T_0 \min \) and \( \alpha \approx 2 \).

We will next use information from the Benzaquen and Walsh experiment to calculate the theoretical values of \( \xi_c \min , T_c \max \) and \( T_0 \min \). We will compare these to the actual experimental values.
From relation (20), we see that $\xi_{\text{c}}^{\text{min}}$ can be calculated if we are given the experimental value of $n_D$. We will also require theoretical values for $n_c$ and $a$.

For the three samples reported by Benzaquen and Walsh, $n_D \times 10^{-15}$ cm$^3$ had values of 6.4, 8.7 and 9. The estimated uncertainty in these values was less than about 20%.

Various attempts have been made to estimate $n_c$ with results ranging between 2 and 7; we choose the most recent value available:

$$n_c = 5.3 \pm 0.3.$$  \hspace{1cm} (23)

The effective mass approximation may be used with considerable confidence to calculate a theoretical value for $a$ for lightly doped GaAs. In the effective mass approximation, one calculates the Bohr radius $a_B$ and the donor energy level $\varepsilon_d$ for an isolated impurity level. The theoretical and experimental values for $\varepsilon_d$ are found to be in good agreement for donor impurities for GaAs. As such, the theoretical value for $a_B$ for GaAs can be considered to be extremely reliable. For light doping, the typical localization length $a$ for the actual localized states in GaAs will differ only slightly from $a_B$. From the effective mass approximation, we calculate $a_B = 100\text{Å}$. In the Benzaquen and Walsh experiments on GaAs, $n_D < 10^{16}$ cm$^{-3}$, i.e. light doping. This means we can be confident that $a$ will differ from 100Å by no more than perhaps a few percent.
Writing \( a = \overline{a} \times 100\text{Å} \) and

\[
\xi_c \min = \xi_c \min / \overline{a},
\]

we may calculate \( \xi_c \min \) from relation (20). We may compare this theoretical value to the experimental value \( \xi_c \exp \) for \( \xi_c \), the latter of which is simply \((T_0 \exp /T)^{1/4}\) where \( T_0 \exp \) is the experimental value for \( T_0 \). By estimating the \( T \)-range in which VRH was observed (see last column of Table II), we can calculate upper and lower bounds for \( \xi_c \exp \). We compare in Table I \( \xi_c \min \) and \( \xi_c \exp \); we see that for all three samples, the experimental values for \( \xi_c \) do not go well with equation (13): \( \xi_c \lesssim 5.3 \) implies \( r_{\max} \lesssim 260\text{Å} \) (using \( a = 100\text{Å} \)) whereas \( n_D^{-1/3} \) (a measure of nearest neighbor distance) is 500Å.

Recall that \( r_{\max} \) is a measure of the separation between critical pairs of sites and is expected to be well in excess of \( n_D^{-1/3} \).

We may also calculate the theoretical values \( T_c \max \) and \( T_0 \min \), provided we can get an estimate for \( \Delta \varepsilon \). An estimate for \( \Delta \varepsilon \) is easily obtained by noting that Benzaquen and Walsh reported values of 3.55 meV to 3.95 meV for \( \varepsilon_c - \mu \), where \( \varepsilon_c \) is the energy at the bottom of the conduction band.

Writing \( \Delta \varepsilon = \overline{\Delta \varepsilon} \times 1\text{meV} \), we have

\[
T_c \max = \overline{T_c} \max \overline{\Delta \varepsilon} \overline{a}
\]

(25)
<table>
<thead>
<tr>
<th>Sample</th>
<th>$n_D(a)$ (in $10^{15}$ cm$^{-3}$)</th>
<th>$n_A/n_D(a)$</th>
<th>$\xi_c^{\text{min}}$</th>
<th>Range of $\xi_c^{\text{exp}}[a]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.4</td>
<td>0.64</td>
<td>18.7</td>
<td>3.52 to 5.26</td>
</tr>
<tr>
<td>2</td>
<td>8.7</td>
<td>0.43</td>
<td>16.9</td>
<td>3.51 to 5.06</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>0.31</td>
<td>16.7</td>
<td>3.51 to 4.83</td>
</tr>
</tbody>
</table>

TABLE I: Comparison of the theoretical minimum values of $\xi_c$, $\xi_c^{\text{min}}$, for localization length $a = 100\AA$, to the ranges of experimental values of $\xi_c$, $\xi_c^{\text{exp}}$, for lightly doped n-GaAs. Values of $n_D$ and $n_D/n_A$ for the experimental samples are also given.

(a) From Ref. [7].
and

$$T_0^{\text{min}} = T_0^{\text{min}} \frac{\Delta \varepsilon}{a^3}. \quad (26)$$

We record in Table II the values of $T_c^{\text{max}}$, $T_0^{\text{min}}$ as well as the values of $T_0^{\text{exp}}$ and the T-range in which VRH was observed.

Since $\bar{a}$ is close to unity, we see from this table that there is no way to choose $\Delta \varepsilon$ to get agreement between theory and experiment: if one chooses $\Delta \varepsilon$ small enough to get $T_0^{\text{min}}$ to agree with $T_0^{\text{exp}}$, then the theoretical value $T_c^{\text{max}}$ becomes even further removed from the experimental T-range. Even taking into account that linearity is expected for $T$ as high as $\alpha T_c$ will not, for a reasonable choice of $\alpha$, bring the theoretical T-range into line with the experimental T-range.

Ad hoc agreement between theory and experiment may be obtained by making a particular choice for $\alpha$, $\bar{a}$ and $\Delta \varepsilon$. For example, we list in Table III the values of $\bar{a}$, $\Delta \varepsilon$ and $\alpha T_c^{\text{max}}$ which result, for each of the 3 samples, upon taking $\alpha = 2$. For the values shown, equations (13) and (14) make sense, given the experimental values for $\xi_c$, and the theoretical quantities $T_0^{\text{min}}$ and $\alpha T_c^{\text{max}}$ agree with $T_0^{\text{exp}}$ and the experimental T-range. However, this ad hoc agreement is not easily justified, for two reasons: (i) since the experimental value agrees so well with the effective mass value for $\varepsilon_d$ in GaAs, $a_B = 100\text{Å}$ is a very reliable result; thus, $n_D a_B^3$ has values
### TABLE II

<table>
<thead>
<tr>
<th>Sample</th>
<th>$T_0 \text{ min}$ (in K)</th>
<th>$T_0^{\text{exp}}[a]$ (in K)</th>
<th>$T_c^{\text{max}}$ (in K)</th>
<th>experimental $T$-range$^{(a)}$ (in K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>38,500</td>
<td>1,073</td>
<td>0.31</td>
<td>1.4 to = 7</td>
</tr>
<tr>
<td>2</td>
<td>28,300</td>
<td>915</td>
<td>0.35</td>
<td>1.4 to = 6</td>
</tr>
<tr>
<td>3</td>
<td>27,300</td>
<td>764</td>
<td>0.35</td>
<td>1.4 to = 5</td>
</tr>
</tbody>
</table>

**TABLE II**: Comparisons of the theoretical minimum values of $T_0$, $T_0^{\text{min}}$, to the experimental values of $T_0$, $T_0^{\text{exp}}$, and of the theoretical maximum values of $T_c$, $T_c^{\text{max}}$, to the experimental ranges of temperature for lightly doped n-GaAs. For the theoretical values, the localization length $a$ was taken as 100Å and the bandwidth $\Delta E$ of localized states as 1 meV.

(a) From Ref. [7].
TABLE III

<table>
<thead>
<tr>
<th>Sample</th>
<th>Value of $\bar{a}$ giving $\xi_c$ agreement</th>
<th>Value of $\Delta\varepsilon$ giving $T_0$ agreement</th>
<th>$\alpha T_c^{\text{max}}$ (in K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.5</td>
<td>2.5</td>
<td>7.0</td>
</tr>
<tr>
<td>2</td>
<td>4.0</td>
<td>2.1</td>
<td>5.9</td>
</tr>
<tr>
<td>3</td>
<td>4.0</td>
<td>1.8</td>
<td>5.0</td>
</tr>
</tbody>
</table>

TABLE III: Values of the localization lengths $a$ ($a = \bar{a} \times 100\text{Å}$) and bandwidths $\Delta\varepsilon$ ($\Delta\varepsilon = \Delta\varepsilon \times 1\text{meV}$) required to give agreement between the theoretical and experimental values of $\xi_c$ and $T_0$ for lightly doped n-GaAs. Also given are the resulting values of $\alpha T_c^{\text{max}}$, the theoretical maximum temperature above which variable range hopping will no longer be observed. All quantities are calculated based on the choice $\alpha = 2$. 
of 0.0064, 0.0087, 0.009 (to within 20% or so\(^{14}\)) for the three samples; as discussed at the end of this section, under these conditions, we are in the strongly localized regime: it is difficult to envisage how a could be 400Å; (ii) if a were 400Å, we would have extended states near the middle of the impurity band; the experimental results, however, indicate localized states.

There is no reasonable choice for \(\alpha, \bar{a}\) and \(\Delta \epsilon\) which will bring theory in line with experiment. We conclude that Benzaquen and Walsh observed VRH which cannot adequately be described by the models which successfully describe VRH in ASs. The T-range of the data is well above the T-range one would expect, from a theoretical perspective, to see VRH, and the experimental \(T_0\) values are much smaller than those one would predict theoretically. The only other possibility, that \(a \approx 400\text{Å}\), is very difficult to envisage justifying; we do not consider this a viable alternative. (See below.)

Benzaquen et al. reported\(^8\) VRH in lightly doped n-InP. A similar comparison between theory and these experimental data can be made. The conclusion is the same: the experimental T-range is well above the T-range for VRH predicted by the theory, the experimental \(T_0\) values are too small to be accounted for by the theory, and since the doping was very light in the n-InP samples studied, there is no reasonable way to choose \(a, a\) and \(\Delta \epsilon\) to get agreement between theory and experiment; in short: the VRH observed in
both GaAs and InP cannot be adequately described by those models which so successfully describe VRH in ASs.

The observations of VRH by Benzaquen and Walsh and by Benzaquen et al. become even more interesting when we compare them to the findings of Emel'yanenko et al., who also investigated hopping conduction in lightly doped GaAs. Emel'yanenko et al. observed VRH in 4 samples in temperature ranges 0.15-0.5 K to 1 K. For $n_D$ values of $2-7.2 \times 10^{15}$ cm$^{-3}$, they found $T_0$ values between about $2 \times 10^4$ K and $5 \times 10^4$ K. These findings are in good agreement with the predictions of the existent theories. The Emel'yanenko et al. data, as presented in a log$_{10}\rho$ vs. $T^{-1/4}$ plot, also suggests VRH in their samples for the temperature range 1.8K to 4.2K, but with considerably smaller $T_0$ values. For example, for one of their samples (sample 4), $T_0$ for the 0.15K to 1K temperature range is about $2 \times 10^4$ K whereas $T_0$ for 1.2K to 4.2K is about $9 \times 10^2$ K. This higher temperature VRH is like the VRH reported by Benzaquen and Walsh and by Benzaquen et al.

We note that Emel'yanenko et al. did not claim to have observed VRH for the temperature range 1.2 K to 4.2 K. Instead, they interpreted this higher temperature data as impurity conduction with a constant activation energy. However, based solely on a comparison of their log$_{10}\rho$ vs. $T^{-1}$ plots to their log$_{10}\rho$ vs. $T^{-1/4}$ plots, we must claim that their samples do not show impurity conduction with a constant activation energy in this higher temperature range any more clearly than they show VRH.

These three experimental reports together suggest that there are two VRH regimes, one below $T_c$ and one above. The VRH below $T_c$ is well
understood on the basis of established theories\textsuperscript{1-3}. While the Emel'yanenko et al. data for $T > T_c$ is only as convincing as their data for $T < T_c$, the data presented by Benzaquen and Walsh and by Benzaquen et al. for $T > T_c$ show unequivocal VRH. This higher temperature VRH, associated with such small $T_0$ values, cannot be understood on the basis of existent theories.

Several other experimental investigations of hopping conduction in DSs have been carried out. These are reviewed in ref. [7], where it is pointed out that the only other work in which unequivocal VRH was observed is reported in two papers\textsuperscript{18,19} by Allen and co-workers. In the investigations of Allen and co-workers, heavily doped n-Ge was used. In their samples, there was considerable overlap, extended states existed in the middle of the impurity band and the value of $a$ for the localized states was not known. Allen and co-workers deduced values for $a$ and $g_F$ by comparing their experimental results to theoretical expressions.

Other experimental studies which show some, albeit equivocal, evidence for VRH in DSs are those of Shlimak and Nikulin\textsuperscript{20} and Yaremenko\textsuperscript{21}. Unfortunately, the materials studied were heavily doped and so the value of $a$ for these materials was unknown.

Since the value of $a$ cannot be calculated by theoretical methods for the n-Ge studies of Allen and co-workers, ref.s [7], [8] and [17] constitute the only experimental data available to date which allow for a genuine comparison between theory and experiment for VRH in three-dimensional DSs. We have shown in this section that certainly two and possibly all three of
these investigations reported VRH which cannot be understood on the basis of existent theories.

Before proceeding to describe our numerical simulation, we point out that Shafarman and Castner\textsuperscript{22} have recently reported VRH with small $T_0$ values in uncompensated Si:As. They found $T_0$ to decrease with increasing $n_D$. As their investigation was close to the Metal-Insulator (MI) transition, the localization length $\alpha$ would be expected to be considerably larger than the effective Bohr radius for shallow donors in Si. A large localization length explains, at least for the samples very close to the MI transition, the small $T_0$ values observed by Shafarman and Castner.

We do not believe that a similar explanation can account for the small $T_0$ values observed by Benzaquen and Walsh or by Benzaquen et al. Our objections to such a notion are fourfold. First, the samples studied were compensated; as pointed out by Fritzsche\textsuperscript{23}, compensation tends to increase the critical concentration of donor impurities marking the MI transition. The donor concentrations of the samples studied ($6.4-9 \times 10^{15}$ cm$^{-3}$ in GaAs and $2.1-8.4 \times 10^{15}$ cm$^{-3}$ in InP) are such that the samples are well removed from the MI transition, as is verified by two experimental reports of Emel'yanenko et al.\textsuperscript{24,25} Second, as reported by Shafarman and Castner, near the MI transition the localization length increases and $T_0$ decreases as $n_D$ increases. While $T_0$ was found to decrease with increasing $n_D$ in the GaAs samples\textsuperscript{7}, the opposite trend of increasing $T_0$ with increasing $n_D$ was observed in the InP samples\textsuperscript{8}. Third, if one were to claim that the
localization length for the Benzaquen and Walsh GaAs samples was three to four times the Bohr radius, so as to explain the small $T_0$ values, one would then be at a loss to account for the VRH reported by Emel'yanenko et al., which is well understood via the established theories of VRH. Finally, strong supporting evidence that — in compensated GaAs and InP samples with the donor concentrations specified above — the localization length is indeed $a_B$, is presented in the experimental reports of Kahlert and Landwehr, Lemoine et al. and Emel'yanenko et al. (In these papers, it is shown that the resistivity $\rho_3$ varies as $\rho_0 \exp\left(\frac{-\alpha}{a_B} \right)$ with $\alpha$ close to $n_D a_B$ the theoretical value of 1.73 afforded by percolation theory.)

In view of these facts, we conclude that — instead of attempting to force the established theories of VRH to account for the VRH of ref.s [7] and [8] — it is much more reasonable to take $a=a_B$ and proceed in the manner outlined in Sec. I to understand the small $T_0$ values reported in ref.s [7], [8] and possibly also in ref. [17].
III. The Numerical Simulation
(i) General Description

We have carried out a two-dimensional numerical simulation for the purpose of calculating the overall resistance of the resistor network described in Sec. I. In this section, we outline how the numerical simulation was done. We started with an $L_x \times L_y$ square lattice of boxes, where $L_x$ and $L_y$ were the numbers of boxes in the $x$- and $y$-directions, respectively. We chose a fixed number $F$, where $0 < F < 1$, and for each box generated and compared to $F$ a random number $\phi$ where, again, $0 < \phi < 1$. The random number $\phi$ was generated using a uniform $(0,1)$ random number generator (available in the UBC APMATH library). If $\phi > F$, the box was left empty but if $\phi < F$, the box had placed at its center a "site". Each time a site was produced, another random value for $\phi$ was generated; this value of $\phi$ was used to assign an energy $\tilde{\varepsilon} = \phi - \frac{1}{2}$ to the site. After having done this for every box, we ended up with a set of $N$ sites distributed randomly in position-space and having energies randomly distributed between $-\frac{1}{2}$ and $\frac{1}{2}$. The $N$ sites were labelled $1, 2, 3, \ldots, i, j, k, \ldots, N$. Site "i" had energy $\tilde{\varepsilon}_i$ and coordinates $(x_i, y_i)$, where $x_i$ and $y_i$ were the coordinates of the box within which the site was placed. We repeated this process several times, thereby generating a set of samples, each sample having approximately the same number of sites, and we calculated the average resistance for the sample set. Had we produced an extremely large number of samples each having the same $L_x, L_y$ and $F$, the average number $\overline{N}$ of sites per sample would
have been $\bar{N} = F L_x L_y$; for a given $L_x$ and $L_y$, $F$ was chosen to give a particular value for $\bar{N}$.

Each of the samples in the collection was a member of the grand canonical ensemble for a doped semiconductor. $L_x L_y$ was the two-dimensional volume while $F$ was related to the average density of donor impurities. To see this more clearly, suppose we had produced three-dimensional samples, each starting from an $L_x$ by $L_y$ by $L_z$ cubic lattice of boxes, and suppose that we could have taken $L_x$, $L_y$, and $L_z$ to be large enough that $F$ could be taken to be the ratio of donor impurities to total number of host crystal atoms in a lightly doped semiconductor (so $F$ would have been about $10^{-7}$, for example). Then each box would have represented a host crystal atom and each site would have represented a donor atom.

$\bar{N}$ had to be chosen large enough to do a meaningful simulation ($\bar{N} > 1000$, for example). Since the number of boxes was limited to about $10^6$, we were forced to choose $F$ to be much larger than $10^{-7}$. $L_x L_y \lesssim 10^6$ means that $\bar{N} = 3000$ implies $F = 0.003$. As such, each empty box represented an aggregate of many host-crystal atoms and each box with a site in it represented an aggregate of host atoms with a donor impurity at the center of the aggregate. One advantage to having generated sites in the manner described — as opposed to having randomly generated coordinates for a fixed number of sites (which would correspond to having worked in the canonical ensemble) — was that we were able to avoid having donor atoms
which would have been so close as to have been effectively a single site. The reasons for having worked in two dimensions instead of three was related to the limitation on the number of boxes: placing 3000 sites in $10^6$ boxes gives a larger site-to-site distance and hence a more realistic modelling of spatial randomness in two dimensions than in three. This is not to say, however, that a three-dimensional simulation would not produce meaningful results.

In a given sample, the N sites represented N nodes for the resistor network model of the doped semiconductor. The resistance between any two sites i and j was given, in terms of $\mathbf{r}_i \equiv (x_i, y_i)$, $\mathbf{r}_j \equiv (x_j, y_j)$, $e_i$, $e_j$, a and the temperature t, by equations (8) - (10).

To determine the overall resistance of the network, we began by extending the network as follows: we treated the $L_x$ by $L_y$ set of boxes as a "unit cell" which was periodically repeated in two dimensions. A potential drop of 1 unit was placed between any given site and its next image in one of the two directions. This arrangement resulted in an overall current flow.

We obviously did not need to include all the resistors in order to calculate the overall resistivity of the network, because resistors which connected sites that were very far apart would have been very large, and those resistors would have been shorted out by smaller resistors which connected sites that were closer. However, we needed to be sure that we did include all resistors that mattered, and to do so, we proceeded as follows:
We chose an integer $d_{\text{max}}$ and constructed a resistor network based on the value of $d_{\text{max}}$. For a given site $i$, we included in this network all resistors $R_{ij}$ for which two conditions were satisfied: $|x_i - x_j| < d_{\text{max}}$, $|y_i - y_j| < d_{\text{max}}$. We did this for each of the $N$ sites. Our original choices of $L_x$ and $L_y$ were made large enough to ensure that $d_{\text{max}} \ll L_x$, $d_{\text{max}} \ll L_y$ for all $d_{\text{max}}$ of interest.

For a given $d_{\text{max}}$, we next checked to see if the resistor network so formed would "percolate" in the direction of the potential drop, which we took to be the $y$-direction. The network percolated in the $y$-direction if there was a continuous chain of resistors going between the edge of the unit cell at $y = 0$ and the edge at $y = L_y$. If we did not have percolation, the overall conductivity was zero and we would proceed to the next (larger) value of $d_{\text{max}}$. If the network percolated, we used Kirchoff's laws to calculate the potential $V_i$ at each site $i$ in the network. The $V_i$ were obtained by employing the incomplete Cholesky conjugate gradient (ICCG) method. We then calculated the current $I_{ij} = (V_i - V_j)/R_{ij}$ flowing from site $i$ to site $j$ through resistor $R_{ij}$. For each of several planes at constant $y$, we calculated the net current crossing the plane. Since they were all found to be the same to within computational accuracy, we had thereby ensured ourselves of having found the correct answer. At this stage, we had calculated the current $I(d_{\text{max}}, t)$ for the values chosen for $d_{\text{max}}$ and $t$.

We repeated the calculation of $I(d_{\text{max}}, t)$ for several values of $d_{\text{max}}$ for a given $t$. As $d_{\text{max}}$ increased, so did $I(d_{\text{max}}, t)$, eventually levelling off.
when \( d_{\text{max}} \) became so large that an increase in \( d_{\text{max}} \) resulted in adding to the network resistors that were very large and hence made no contribution to the overall current. For the largest values of \( d_{\text{max}} \) tried, \( I(d_{\text{max}}, t) \) was essentially constant; we denote this value of the current by simply \( I(t) \).

We calculated \( I(t) \) for several values of \( t \); \( t \) was taken to lie in the range \( 1 > t > 0.012 \). The result was a set of values of \( I(t) \) for several values of \( t \).

We obtained such a set of values of \( I(t) \) for each of a collection of samples with the same \( L_x \), \( L_y \) and \( F \). The total number of sites \( N \) differed from sample to sample, as did the positions of energies of the various sites. As such, each sample represented a member of the ensemble of samples for the resistor network model of a doped semiconductor. By calculating \( I(t) \) for each of the values of \( t \) for each sample, we were able to determine an average for \( I(t) \), denoted \( \langle I(t) \rangle \) and estimate its error.

(ii) Computational Details

We obtained \( I(t) \) values for 4 "1000-site" systems (with \( N \) values of 932, 947, 992 and 1015), 4 "2000-site" systems (with \( N \) values of 1920, 1926, 1952 and 1986), 2 "2000-site, rectangular" systems (\( N = 1991 \) for each) and 7 "2800-site, rectangular" systems (with \( N \) values of 2653, 2756, 2758, 2788, 2800, 2804 and 2820). The "1000-site" and "2000-site" systems were all square with \( L_x = L_y = 700 \). The "2800-site, rectangular" systems and one of the two "2000-site, rectangular" systems had \( L_x = 500 \), \( L_y = 1040 \); the other "2000-site, rectangular" system had \( L_x = 400 \), \( L_y = 1300 \). For all 17
systems, we chose \( r_{s-s}/a = 6 \) where \( r_{s-s} \equiv (L_x L_y / N)^{1/2} \).

We chose \( d_{\text{max}} \) to vary from an initial value \( d_{\text{in max}} \) close to \( r_{s-s} \) to a final value \( d_{\text{fin max}} \) which was roughly \( 3d_{\text{in max}} \). For the first few systems studied, \( d_{\text{max}} \) was taken to increase in steps of unity from \( d_{\text{in max}} \) to \( d_{\text{fin max}} \).

Once we had clearly established that \( I(d_{\text{max}}; t) \) increased then levelled off with increasing \( d_{\text{max}} \) in a regular fashion, we needed only a few \( d_{\text{max}} \) values to get \( I(t) \) for the systems subsequently studied.

The computing time required to calculate the net current for given \( d_{\text{max}} \) and \( t \) was dominated by the step where Kirchoff's laws were solved for the voltages \( V_i \). Large computing times were required despite the use of the ICCG method. The computing time increased with \( N \) roughly as \( N^{3/2} \). For \( N = 1000 \), the average computing time corresponding to a single \( d_{\text{max}} \) value and a given \( t \) was roughly 17 seconds; for \( N = 2000 \), the time was about 50 seconds; for \( N = 2800 \), the time was roughly 90 seconds. The computer time was generally larger for a smaller value of \( t \). In consequence, we chose the minimum \( t \) value to increase slightly with the number of sites (as can be seen from Fig. 2).

Because of the large computing times required, we used the array processor at UBC (the FPS-164/MAX, with one megaword of memory). The results for the 17 systems were collected within twenty calendar days.
IV. Results and Discussion

In Fig. 2, we show plots of $\ln \langle I(t) \rangle$ vs. $t^{-1/3}$ for the 4 1000-site systems, the 4 2000-site systems and the 7 2800-site systems. (For two-dimensional systems, the $1/4$ in equation (1) is replaced by $1/3$.) The ordinates for each case are staggered in the figure to avoid clutter. The error bars shown reflect the standard deviations in $I(t)$ (not $\ln I(t)$), as inferred from the set of samples. The dimensionless $I(t)$ values reflect our choosing the potential drop to be 1 unit of voltage and our taking $R_0$ to be 1 unit of resistance.

Recall that $r_s/a = 6$ for every system. This means that each system may be regarded as a small section of an infinite system with $r_s/a = 6$. As such, assuming the number of sites in each system is large enough, we should expect close agreement between the $\ln \langle I(t) \rangle$ vs. $t^{-1/3}$ curves for the systems with different numbers of sites.

In Fig. 3, we compare the plots for the 1000-site and 2000-site (square systems) and also for the 2800-site systems and the single 2000-site system with $L_x = 500$, $L_y = 1040$. Again, the ordinates are staggered to allow ease of comparison. We see that in both cases, the plots agree very well, indicating that 1000 or more sites are enough.

Also shown in Fig. 2 are extrapolations of both the high $t$ ($t \gg 1$) and low $t$ ($t \ll t_c = kT_c/\Delta\epsilon$) theoretical curves. At high $t$, the resistor network (8) becomes independent of $t$ and the problem reduces to the so-called $r$-percolation problem. For low $t$, the resistor network reduces to the form (3) and the theoretical curve corresponds to the Mott VRH Law for
Fig. 2: Plots of the average currents obtained numerically vs. $t^{-1/3}$ for:
1000-site samples ($\ln I_1$ ordinate, O symbols), 2000-site samples
($\ln I_2$ ordinate, • symbols) and 2800-site samples ($\ln I_3$ axis,
$\triangle$ symbols). The dashed curves are theoretical high $t$ and low $t$
curves which are described in the text.
Fig. 3: Comparison of average currents for the 1000-site systems (o) and 2000-site systems (•) ($\ln I_1$ ordinate), average current for the 2800-site systems (Δ) and current for one 2000-site rectangular system (+) ($\ln I_2$ ordinate). Error bars (not shown) are given in Fig. 2.
two dimensions. In comparing the numerical results and the theoretical curves, one must bear in mind that the theoretical estimate for $\sigma_0$ is not precise: the uncertainty in $\sigma_0$ is of order unity (see ref. [30] for details). As such, we have agreement at high $t$ between the theoretical curves ($a_1$, $a_2$ and $a_3'$) and the limiting numerical values, and at low $t$ between the magnitude of the current for the theoretical curves ($b_1$, $b_2$ and $b_3$) and the magnitude of the numerical points. (The theoretical curves $a_3$ and $a_3'$ differ in that $a_3'$ takes into account that the 2800-site systems are rectangular and equates the resistance of a sample to $R_{cY}/L_x$; curves $a_1$, $a_2$ and $a_3$ take the resistance to be simply $R_c$. This apparent dependence on shape of the overall resistance is discussed more fully near the end of this section). At low $t$ we see that, for the 1000- and 2000-site systems, we could draw a straight line through the final few $ln <I(t)>$ points. However, the largeness of the error bars prevents us from doing so. We would have preferred smaller error bars in this $t$-range, but because of the long computing times for small $t$, because our interest is in the higher $t$-range and because the low $t$-range is well understood, we chose not to perform further runs so as to reduce these error bars. Even so, we can conclude that $ln <I(t)>$ is not linear in $t^{-1/3}$, for our two-dimensional study, at temperatures in excess of about $2t_c$ (for our samples, $t_c^{-1/3} \approx 4.0$). As pointed out in Sec. II, this estimate agrees with the results of Seager and Pike for three dimensions.

Quite apart from their relation to experimental data on hopping conduction in DSs, our numerical simulations are important in their own
right because they allow us to test and thereby establish an analytic theory for percolation at temperatures when \( t \) is of order unity. Even so, it is essential that we compare the results of our numerical investigations to the experimental results discussed in Sec. II.

We see from Fig. 2 that there exists no extensive temperature range above \( t_c \) for which \( \ln \langle I(t) \rangle \) is linear in \( t^{-1/3} \) with a slope significantly smaller than the slope for VRH at \( t \lesssim t_c \). However, our numerical results do explain the general behavior observed by Emel'yanenko et al.\(^7\) in three-dimensional lightly doped GaAs (see Sec. II). Emel'yanenko et al. observed the following general behavior: that although the conductivity continued to rise with \( T \) above \( T_c \), its rate of increase with \( T \) was smaller than for \( T \) below \( T_c \). This general behavior, questioned by Benzaquen and Walsh\(^7\), is understandable on the basis of our two-dimensional numerical simulations, as our results reproduce this general behavior. Moreover, recall from Sec. II that Emel'yanenko et al. claimed that \( \ln \sigma \) was linear in \( T^{-1} \) above \( T_c \). In Fig. 4, we have plotted our numerical results as \( \ln \langle I(t) \rangle \) vs. \( t^{-1} \). Indeed, we find a linear region above \( t_c \). In our next paper, we will present our analytic theory and show that it agrees with our numerical simulation for two dimensions and a flat density of states. At that point, we will be able to use the analytic theory to examine the behavior of the conductivity in three dimensions for various chosen densities of states. We will compare the results of our theory for these various densities of states to the experimental results of Benzaquen and Walsh and Benzaquen et al. For a flat density of states in three dimensions, we will see that there is
Fig. 4: Plots of the average currents for the 1000-site systems (O),
2000-site systems (●) and 2800-site systems (Δ) vs. $t^{-1}$. The
linear regions are indicated by solid lines. The slopes,
determined by a $\chi^2$ minimization routine, are as follows:
-0.271±0.009 (O), -0.278±0.014 (●), -0.278±0.012 (Δ).
again a temperature range above \( T_c \) in which \( \ln \sigma \) is linear in \( t^{-1} \). If the Emel'yanenko et al. claim that \( \ln \sigma \) for their samples is linear in \( T^{-1} \) above \( T_c \) is correct, it may be understood on the basis of our model: for either two or three dimensions and a flat density of states, our model predicts the Mott law below \( T_c \) and \( \ln \sigma \) linear in \( T^{-1} \) above \( T_c \). Such linearity of \( \ln \sigma \) in \( T^{-1} \) associated with hopping conduction in three-dimensional DSs has often been observed experimentally. (For example, see ref. s [24] - [27].)

Fowler and Hartstein\(^{31,32} \) observed experimentally a transition in two-dimensional systems from \( \ln \sigma \) linear in \( T^{-1/3} \) at low temperatures to \( \ln \sigma \) linear in \( T^{-1} \) at higher temperatures. Hayden and Butcher\(^{33} \) have presented a theory for activated hopping in two dimensions and compared their theory to the Fowler and Hartstein results\(^{31} \). Unfortunately, in their theory, the low temperature form of the resistance \( R_{ij} \), as given by equation (3), was used in a temperature range where \( t \) was of order unity. As shown in Fig. 1, the low temperature form of \( R_{ij} \) deviates appreciably from the full form, equation (8), for \( t > 2t_c \). This very important difference between the two forms of \( R_{ij} \) is revealed by comparing the slopes found in Fig. 4 to those reported by Hayden and Butcher for their model: Hayden and Butcher reported an activation energy of \( 0.38\Delta \epsilon \), whereas from Fig. 4, we find the considerably different value \( 0.28\Delta \epsilon \). This difference is due to the use by Hayden and Butcher of the low temperature form of \( R_{ij} \) in the high temperature region.

A somewhat unexpected result of our numerical investigations comes in the form of preliminary indications that \( I(t) \) depends on the shape of the
system. Notice that the compared plots in Fig. 3 correspond to systems having the same shape. A comparison of curves for systems with different shapes does not reveal the agreement seen in Fig. 2. For example, in Fig. 2, we see that the curve for the 2800-site systems — all of which had \( L_x = 500, L_y = 1040 \) — appears to drop, with increasing \( t \), below the coinciding curves for the 1000- and 2000- site systems — all of which were square. In Fig. 5, we plot \( \ln I(t) \) for the 2 2000-site rectangular systems and \( \ln \langle I(t) \rangle \) for the 4 2000-site square systems (with a single ordinate) against \( t^{-1/3} \). We see the same trend as in Fig. 2, namely that, for the higher \( t \) values, the greater the ratio \( L_y/L_x \), the smaller the value of \( I(t) \).

One might be inclined to think that this is simply reflecting that the systems are "macroscopic", i.e. that the resistance \( R \) of a given system is given by \( R = \rho L_y/L_x \) where \( \rho \) is the macroscopic resistivity. Indeed, the curve \( a_3' \) of Fig. 2 was obtained by writing \( R = R_c L_y/L_x \). However, according to percolation theory\(^{30,34}\), these sample systems are small enough that the overall resistance of a system is expected to be due to a single resistor. As such, we cannot expect the shape dependence to be due to macroscopic behavior of the systems. Nevertheless, that curve \( a_3' \) agrees better with the numerical result than the curve \( a_3 \) does, suggests there is a dependence of \( R \) on the shape of the system.

We will not discuss this shape dependence further in this paper. We plan to perform runs for several differently-shaped samples and discuss this topic in a future publication.
Fig. 5: Comparison of average current for the square 2000-site systems (•), current for one 2000-site system with $L_x = 500$, $L_y = 1040$ (+) and current for one 2000-site system with $L_x = 400$, $L_y = 1300$ (□).
Finally, one may well ask what the shape of the curve $\ln I(t)$ vs. $t^{-1/3}$ tells us about the typical hopping distance and typical hopping energy as a function of temperature. With relation to this point, we have compared the results of our numerical simulations to a simple plot based on equation (8). We replaced $r_{ij}$ by $1.2 r_{s-s}$ (which is in accord with the r-percolation problem) and plotted $\ln (R^{-1})$ vs. $t^{-1/3}$ for various values of $\tilde{c}_i$ and $\tilde{c}_j$ (see equations (9) and (10) as well). In Fig. 6, we compare one such plot to the average for the 2000-site systems. The $R_{ij}$ curve agrees well with the numerical results at higher $t$ but deviates substantially at lower $t$. The figure indicates that, if we think in terms of a single resistor dominating, we have nearest neighbor hopping down to $t \approx 0.07$ but for lower temperatures, the values of $\tilde{c}_i$, $\tilde{c}_j$ and $r_{ij}$ entering the typical resistance $R_{ij}$ vary with temperature, corresponding to hops of larger range. Comparing Figs. 4 and 6, we see that, for the 2000-site systems, essentially all of the temperature range for which $\ln I$ is linear in $t^{-1}$ is a nearest-neighbor hopping regime. Moreover, Fig. 6 indicates that, for the 2000-site systems, VRH should be seen in a temperature range for which $t < 0.07$. Fig. 2 indicates this to be so.

A fuller understanding of the shape of the $\ln I(t)$ vs. $t^{-1/3}$ curve is obtained via the analytic theory we have developed to calculate the overall resistance of the resistor network (8). In the next paper in this series, we report the analytic theory and compare the results for our theory to the numerical results presented here.
Fig. 6: Comparison of average current for the 2000-site systems (●) and $R_{ij}^{-1}$ (equation (8)) with $r_{ij} = 1.2$ $r_{s-s}$, $\bar{\xi}_j = 0.225$, $\bar{\xi}_i = -0.225.$
ACKNOWLEDGMENTS

We wish to thank Peter C.W. Holdsworth for many useful discussions, Matthew Choptuik, Jim Glosli and Tom Nicol for assistance with computational details, Dan Zimmerman for many useful discussions and Neb Duric for supplying the $\chi^2$ minimization routine. This work was supported financially by NSERC Operating Grant A713 and by the Graduate Student Fellowship Programme at UBC.
REFERENCES

*Present address: Department of Physics, Dalhousie University, Halifax, N.S., CANADA B3H 3J5

10 A derivation of this form of $R_{ij}$ may be found either in B.I. Shklovskii and A.L. Efros, Electronic Properties of Doped Semiconductors (Springer-Verlag, Berlin, 1984), pp. 82-88, or in M.R.A. Shegelski, Ph.D. thesis (unpublished); this form of $R_{ij}$ differs slightly from the form derived by Miller and Abrahams.


14 D. Walsh, private communication.


34 H. Overhof, Festkorperprobleme XVI, 239 (1976).
Hopping conductivity in lightly doped semiconductors - II: Two-dimensional analytic theory at high temperatures

Mark R.A. Shegelski* and Robert Barrie

Department of Physics, University of British Columbia, Vancouver, B.C., CANADA V6T 2A6

Abstract

In a previous paper, we performed a numerical simulation of a two-dimensional lightly doped semiconductor. We chose a flat density of states with width $\Delta \varepsilon$ and modelled the semiconductor as a Miller and Abrahams type resistor network. In this paper, we again determine the overall resistance of the resistor network, but this time by developing an analytic theory. As before, we use the full form of the resistance and do not take the low temperature asymptotic form because we are interested in temperatures where the reduced temperature $t = \frac{kT}{\Delta \varepsilon}$ is of order unity. We use percolation theory to obtain a set of equations which specify the overall resistance $R(t)$ of the network as a function of $t$. This theoretical result for $R(t)$ depends on the form chosen for the temperature dependence of the critical number $B_c(t)$ of bonds (resistors) per site required for percolation. We obtain agreement between the analytic theory and the
numerical simulation by choosing a very reasonable form for \( B_c(t) \) for \( t \) of order unity; this form for \( B_c(t) \) is in accord with the high \( t \) limit and the low \( t \) asymptotic form of \( B_c(t) \). In particular, we confirm the result of our numerical simulation that there exists a \( t \)-range where \( R(t) = R_0 e^{\epsilon_3/kt} \), with \( \epsilon_3 = 0.28 \Delta \epsilon \). The analytic theory also provides us with the temperature dependences of the typical hopping distance and the typical energy associated with a hop in the temperature range where \( t \) is of order unity.

PACS numbers: 72.20.-i
I. Introduction

Variable range hopping (VRH) in disordered systems, such as amorphous semiconductors or doped semiconductors, is characterized by a conductivity \( \sigma \) of the form

\[
\sigma = \sigma_0 e^{-(T_0/T)^{1/h}},
\]  

(I.1)

where \( \sigma_0 \) depends weakly on the temperature \( T \). Mott \(^1\) and others \(^2,3\), by assuming a density of states flat in the vicinity of the Fermi level, have provided theoretical derivations of equation (I.1), finding

\[
T_0 = \frac{C}{k g_F a^3},
\]  

(I.2)

where \( C \) is a dimensionless constant approximately equal to 20, \( k \) is Boltzmann's constant, \( g_F \) is the density of states at the Fermi level and \( a \) is the electron localization length.

In a previous paper \(^4\), hereafter referred to as [I], we reported that there exist experimental results \(^5,6\) showing unequivocal VRH in lightly doped semiconductors (LDSs) and that this experimental VRH has at present no suitable theoretical explanation. The experimental VRH was observed in a temperature range well above the temperature range where existent theories predict VRH and the experimental \( T \) values were found to be much smaller than the theoretical values.
This paper represents the second of four sequential investigations which comprise a program designed to understand these experimental results. The program is based on the resistor network model of a LDS, which was presented in [I]. In this model, the LDS is replaced by a network of resistors. Every pair of donors in the LDS is mapped into a resistor $R_{ij}$. In [I], we performed a numerical simulation in two dimensions to calculate the overall resistance of the network. To simplify matters, we took the number density $n_D$ of donors to be twice the number density $n_A$ of acceptors, which simplified the chemical potential $\mu$ to $\mu = \mu_0 - kT \ln 2$. We chose a flat density of states with bandwidth $\Delta \varepsilon$ and took $\mu_0 = 0$. The resistor network simplified to

$$R_{ij} = R_0 f(E_i, E_j) \frac{X_{ij}^x}{X_{ij}^x e^{\frac{X_{ij}}{a}}} \quad (I.3)$$

where $R_0$ was a constant resistance and

$$f(E_i, E_j) = \frac{(1 + e^{\frac{E_i}{\Delta \varepsilon}})(1 + e^{\frac{-E_j}{\Delta \varepsilon}})(e^{\frac{E_j - E_i}{\Delta \varepsilon}} - 1)}{E_j - E_i} \quad (I.4)$$

$$E_i = \frac{\varepsilon_i}{t}, \quad \varepsilon_i = \frac{\varepsilon_i}{\Delta \varepsilon}, \quad t = \frac{kT}{\Delta \varepsilon}, \quad X_{ij} = \frac{2r_{ij}}{a} \quad (I.5)$$

where $E_j > E_i$. Since $t$ was of order unity, we did not make the usual
simplification and replace the $R_{ij}$ by their low temperature asymptotic form.

In this paper, we again calculate the overall resistance of the resistor network (1.3), but this time analytically. In Sec. II, we formulate the calculation in terms of a percolation problem. Although the calculation is conceptually simple, it is rather involved. We therefore obtain the necessary set of equations step by step in Secs. III through VI, summarizing in Sec. VII. Our calculation requires a knowledge of the temperature dependence of the number of bonds (resistors) per site $B_c(t)$ required for percolation. In Sec. VIII, we discuss how we choose a form for $B_c(t)$. We show in Sec. IX that our analytic theory results are in good agreement with the numerical results we presented in [I]. We summarize our results in Sec. X. In an appendix, we outline a much simplified version of our theory which can be used more easily and gives results within a few percent of the more exact theory.
II. Formulation as a Percolation Problem

In this paper, we describe an analytic approach which enables us to determine the overall resistance of the resistor network given by equation (I.3).

The basic method underlying our theory is the percolation method. In the percolation method, one imagines removing all of the resistors except those which satisfy

\[ R_{ij} < R_0 e^\xi \]  

(II.1)

for some arbitrarily chosen value of \( \xi \). The remaining network of resistors will percolate only if \( \xi \) is equal to or exceeds a critical value \( \xi_c \). According to percolation theory, the overall conductivity \( \sigma \) of the original network is given by

\[ \sigma = \sigma_0 e^{-\xi_c} \]  

(II.2)

where \( \sigma_0 \) depends only weakly on the temperature \( t \).

Combining equation (I.3) for \( R_{ij} \) with the inequality (II.1) gives

\[ f(E_i, E_j) x_{ij}^{-2} e^{x_{ij}} < e^{\xi}. \]  

(II.3)

Our objective is to find the critical value \( \xi_c \) of \( \xi \) for which the network of
resistors specified by inequality (II.3) just percolates. Since the $E_i$ and $E_j$ in inequality (II.3) involve the temperature $t$, $\xi^c$ will be a function of $t$. Once we know $\xi^c(t)$, we have, via equation (II.2), the dominant temperature dependence of the overall conductivity $\sigma$.

To determine $\xi^c(t)$, we begin by recognizing that inequality (II.3) specifies, in terms of $\xi$, which pairs of sites will be connected by a resistor. As such, we can determine $\xi^c(t)$ by proceeding as follows. We imagine choosing a site at random. The probability that there will be another site between $X$ and $X + dX$ away from the chosen site is

$$p(X) = \frac{\pi}{2} \rho_s a^2 X dX,$$

where $\rho_s$ is the number of sites per unit area. Given that there is indeed a site between $X$ and $X + dX$ away from the chosen site, let $p(X;\xi)$ denote the probability that the energies of the two sites are such that the two sites are connected. We will use inequality (II.3) to determine $p(X;\xi)$. Once we know $p(X;\xi)$, we can easily calculate the average number $B(\xi)$ of resistors, or "bonds", that are connected to a site. By $B(\xi)$, we mean the ratio of twice the total number of resistors admitted by inequality (II.3) to the total number of sites. It is clear that

$$B(\xi) = \frac{\pi}{2} \rho_s a^2 \int_{X_{\min}(\xi)}^{X_{\max}(\xi)} X p(X;\xi) dX,$$

where $X_{\min}(\xi)$ and $X_{\max}(\xi)$ are the minimum and maximum values of $X$ for the energy range considered in inequality (II.3).
where $X_{\min}(\xi)$ and $X_{\max}(\xi)$ are, respectively, the closest and farthest apart two sites can be in accordance with inequality (II.3). If we know $X_{\min}(\xi)$, $X_{\max}(\xi)$ and $p(X;\xi)$, we can get $\xi_c(t)$. The way to do so is to note that if $B(\xi)$ is too small, there will be no percolation; percolation will occur, at temperature $t$, only if $B(\xi) \geq B_c(t)$, where $B_c(t)$ is the critical number of bonds per site required for percolation at temperature $t$. Solving

$$B(\xi) = B_c(t) \quad (\text{II.6})$$

will give us $\xi_c(t)$.

We will turn next to a discussion of how to determine $X_{\min}(\xi)$, $X_{\max}(\xi)$ and $p(X;\xi)$. We will see that calculating $p(X;\xi)$ is an involved process. Several equations involving $\xi$ will need to be considered.

Although the calculation of $\xi_c$, as outlined above, is a conceptually simple task, there are many details involved; the reader will appreciate that the forthcoming calculation is necessarily quite complicated. To make the reading easier, we have split the paper into short sections.

After we have derived all the equations required to calculate $\xi_c(t)$ given $B_c(t)$, we will describe how $B_c(t)$ may be specified.
III. \( X_{\text{min}}(\xi) \) and \( X_{\text{max}}(\xi) \)

It will be helpful to introduce a quantity \( b = E_j - E_i \) and to work with the three variables \( E_i, E_j \) and \( b \) even though only two of them are independent. Rewriting \( f(E_i, E_j) \) in terms of \( b \) and \( E_j \) gives

\[
f(b, E_j) = 2\left( \frac{\sinh b}{b} + \frac{\sinh (b/2)}{b/2} \cosh(E_j - \frac{b}{2}) \right).
\]

(III.1)

Recalling that \( E_i < E_j \) and \( -\frac{1}{2t} < E_j < \frac{1}{2t} \), we see that the region of the \( E_i, E_j \) plane in which we are interested is the triangle formed by the lines \( E_i = -\frac{1}{2t}, E_j = \frac{1}{2t}, b = 0 \) (see Fig. 1). We will refer to this region, from here on in, simply as "the triangle". From equation (III.1), we see that \( f(b, E_j) \) is symmetrical about the line \( E_j = -E_i \) or \( E_j = \frac{b}{2} \). Moreover, \( f(b, E_j) \) increases along a line of fixed \( b \) as we move away from \( E_j = \frac{b}{2} \). Finally, \( f(b, E_j) \) increases along the line \( E_j = \frac{b}{2} \) as we move away from the origin. We see then that the minimum value of \( f(E_i, E_j) \) is 4, which occurs when \( E_i = E_j = 0 \). This implies, via inequality (II.3), that \( X_{ij} \) is restricted according to

\[
X_{ij} < \frac{1}{4} e^{\xi}.
\]

(III.2)

Provided \( \xi > 2 \), the inequality (III.2) is satisfied for \( X_{ij} \) such that
Fig. 1: The triangle and the boundary curve. The triangle is the region specified by \(-\frac{1}{2\tau} < E_i < \frac{1}{2\tau}\) and \(-\frac{1}{2\tau} < E_j < \frac{1}{2\tau}\). The boundary curve is symmetrical about the line \(E_j = -E_i\) (dashed line). The quantities \(b_{\text{max}}\), \(E_j^{b_{\text{max}}}\), \(E_j^{0}\), associated with the boundary curve, are described in the text.
$X_{\text{min}} < X_{ij} < X_{\text{max}}$, where $X_{\text{min}}$ and $X_{\text{max}}$ are the two solutions of $X^2 e^{X - X} = 4$ ($X_{\text{min}} < 2$ and $X_{\text{max}} > 2$). Recall that $X_{ij} = 2$ corresponds to two sites a distance a apart. Because it is unreasonable to be thinking, in the case of lightly doped semiconductors, in terms of localized electronic states with centers less than about a apart, and since the function $X^{-2} e^X$ is minimized at $X = 2$, we take $X_{\text{min}} (\xi) \equiv 2$ for all $\xi > 2$ and $2 < X_{ij} < X_{\text{max}} (\xi)$, where $X_{\text{max}} (\xi)$ is given by

$$X_{\text{max}}^2 e^{X_{\text{max}} - X} = 4.$$ (III.3)

Equation (II.5) reduces to

$$B(\xi) = \frac{\pi}{2} \rho_s a^2 \int_2^{X_{\text{max}}} X p(X;\xi) dX.$$ (III.4)

Hereafter, we shall always take $\xi > 2$ and $2 < X < X_{\text{max}}$. 
IV. The Boundary Curve

We want to eventually obtain an expression for \( p(X;\xi) \). We begin by observing that, according to inequality (II.3), two sites separated by a distance \( X \) will be connected only if their "energies" \( E_i \) and \( E_j \) satisfy

\[
f(E_i, E_j) < X^2 \ e^{\xi - X}.
\]  

(IV.1)

This means that \( E_i \) and \( E_j \) must lie not only within "the triangle" but also within a "boundary curve" given by

\[
f(b, E_j) = X^2 \ e^{\xi - X}.
\]  

(IV.2)

For a given \( X \) and \( \xi \), equation (IV.2) specifies \( E_j \) in terms of \( b \). We denote this relationship by \( E_j^b(X;\xi) \) or, so as to avoid clutter in notation, simply by \( E_j^b \). We call the relationship \( E_j^b \) between \( E_j \) and \( b \) the "boundary curve". From equations (III.1) and (IV.2),

\[
\cosh \left( E_j^b - \frac{b}{2} \right) = \frac{b/2}{\sinh (b/2)} \left( \frac{X^2 \ e^{\xi - X}}{2} - \frac{\sinh b}{b} \right).
\]  

(IV.3)

A solution for \( E_j^b \) exists provided \( 0 < b < b_{\text{max}} \), where
\[
\frac{\sinh b_\text{max}}{b_\text{max}} + \frac{\sinh (b_\text{max}/2)}{b_\text{max}/2} = \frac{x^2 e^{\xi-X}}{2};
\]  
(IV.4)

i.e.: as \( b \to b_\text{max}, E_j^b \to \frac{b_\text{max}}{2} \). Note that \( b_\text{max} \) depends on \( X \) and \( \xi \). As \( b \to 0, E_j^b \to E_j^0 \) where

\[
\cosh E_j^0 = \frac{x^2 e^{\xi-X}}{2} - 1.
\]  
(IV.5)

Thus, as shown in Fig. 1, the boundary curve goes from \( E_i = E_j = E_j^0 \) when \( b = 0 \) to \( E_j = -E_i = -\frac{b_\text{max}}{2} \) when \( b = b_\text{max} \). One can readily show that the slope \( \frac{\partial E_i^b}{\partial E_i} \) of the boundary curve approaches \(-1\) as \( b \to 0 \) and approaches \(+1\) as \( b \to b_\text{max} \). This implies that \( E_j^b \) has a maximum value, denoted \( E_j^{b,\text{max}} \), at some value of \( b \) between 0 and \( b_\text{max} \). This particular value of \( b \) and \( E_j^{b,\text{max}} \) are found by simultaneously solving equation (IV.3) and the equation

\[
\frac{\partial E_i^b}{\partial b} = 0,
\]

which is

\[
b\left(\frac{x^2 e^{\xi-X}}{2} - \frac{\sinh b}{b}\right)\left(\sinh \frac{b}{2} - \frac{b}{2} \cosh \frac{b}{2}\right) + \sinh \frac{b}{2}(\sinh b - b \cosh b) \\
+ b \sinh^2 \frac{b}{2} \sinh (E_j^b - \frac{b}{2}) = 0.
\]  
(IV.6)

Summarizing, the essential features of the boundary curve, all of which
are displayed in Fig. 1, are as follows. $E_i = E_j = E_j^0$ and $\frac{\partial E_j^b}{\partial E_i} = -1$ at $b = 0$ where $E_j^0$ is given by equation (IV.5); as $b$ increases, $E_j^b$ increases until, at a value of $b$ determined by equations (IV.3) and (IV.6), $E_j^b$ reaches a maximum value $E_j^{b_{\text{max}}}$ also determined by equations (IV.3) and (IV.6); $b$ can further increase to a maximum value $b_{\text{max}}$ given by equation (IV.4), where $E_j^b = -E_i = \frac{b_{\text{max}}}{2}$ and $\frac{\partial E_j^b}{\partial E_i} = +1$. The boundary curve is symmetrical about the line $E_j = -E_i$. One can also show that $E_j^0 > \frac{b_{\text{max}}}{2}$: the value of $E_j^b$ at $b = 0$ exceeds the value of $E_j^b$ at $b = b_{\text{max}}$ (we refer here to the positive value of $E_j^0$ which satisfies equation (IV.5)).
V. Expressions for \( p(X;\xi) \)

Based on the above discussion of the salient features of the boundary curve \( E^b_j \), we may proceed to write down an expression for \( p(X;\xi) \). A flat density of states implies that the probability that one site has energy \( E_i \) and the other \( E_j \) is the same for all values of the pair \( E_i, E_j \) that lie within the triangle. This means that \( p(X;\xi) \) is simply the fraction of the area of the triangle which is enclosed by the boundary curve. Depending on the values of \( X, \xi \) and \( t \), the boundary curve will be positioned relative to the triangle in one of four possible arrangements, as displayed in Fig. 2. We discuss each of these arrangements in turn.

(i) The boundary curve may be wholly inside the triangle. In this case,

\[
p(X;\xi) = I(0,b_{\text{max}})
\]

where

\[
I(b_1,b_2) \equiv 4t^2 \int_{b_1}^{b_2} \frac{db}{\sinh (b/2)} \left( \frac{\xi - X}{2} - \sinh \frac{b}{2} \right).
\]

(ii) The boundary curve can cut through the triangle at two values of \( b \). This occurs when \( E^b_{j_{\text{max}}} > \frac{1}{2t} \) and \( E^o_j < \frac{1}{2t} \). The two values of \( b \) are denoted \( b_\lambda \) and \( b_h \), where \( 0 < b_\lambda < b_h < b_{\text{max}} \). \( p(X;\xi) \) is given, in terms of \( b_\lambda \) and \( b_h \), by
Fig. 2: Four possible relative positions of the triangle and the boundary curve. (i) The boundary curve lies completely inside the triangle. (ii) The boundary curve cuts through the triangle at two distinct values of $b = E_j - E_i$. (iii) The boundary curve cuts through the triangle at one value of $b$. (iv) The boundary curve encloses the triangle.
\[ p(X;\xi) = I(0,b^0_\lambda) + I(b_{h},b_{\text{max}}) + 2t(b_{h} - b^0_{\lambda}) - t^2(b_{h}^2 - b^0_{\lambda}^2). \quad (V.3) \]

The equation which determines \( b^0_{\lambda} \) and \( b_{h} \) is obtained by putting \( E_j^b = \frac{1}{2t} \) in equation \( (IV.3) \), and is

\[ 2\cosh \left( \frac{1}{2t} - \frac{b}{2} \right) \sinh \frac{b}{2} + \sinh b - \frac{b}{2} X e^{\xi-X} = 0. \quad (V.4) \]

Note that, like \( b_{\text{max}} \), \( b^0_{\lambda} \) and \( b_{h} \) depend on \( X \) and \( \xi \), but \( b^0_{\lambda} \) and \( b_{h} \) also involve \( t \).

(iii) The boundary curve may cut through the triangle at only one value of \( b \). This occurs when \( E_j^0 > \frac{1}{2t} \) and \( b_{\text{max}} < \frac{1}{t} \). Denoting this single value of \( b \) by \( b_1 \), we write

\[ p(X;\xi) = I(b_1, b_{\text{max}}) + t b_1 (2 - t b_1). \quad (V.5) \]

\( b_1 \) is given by equation \( (V.4) \).

(iv) If the boundary curve wholly encloses the triangle, then

\[ p(X;\xi) = 1. \quad (V.6) \]

This occurs when \( b_{\text{max}} > \frac{1}{t} \).

The final step is to use the above information to calculate \( \xi_c \) as a function of \( t \). Although this is a conceptually simple task, there are
several equations which are involved. To see this, recall that $\xi_c(t)$ is
determined by equations (II.6), (III.3), and (III.4) with $p(X;\xi)$ given by
one of equations (V.1), (V.3), (V.5) or (V.6). Equations (V.1), (V.3) and
(V.5) involve $b_{\max}$, given by equation (IV.4), and possibly also either $b_\lambda$
and $b_\h$ or $b_1$, given by equation (V.4). This is not all, however, because
the temperature $t$ also comes into play: which of the four relative
arrangements of the boundary curve and the triangle are possible will depend
on the value of $t$. We turn next to a detailed discussion of this issue.
VI. Temperature Dependence; The Extremal Boundary Curve

We wish to ensure that the discussion of the temperature dependence is clear. To this end, we phrase the discussion not in terms of an arbitrary value of $\xi$, but in terms of the critical value $\xi_c(t) \equiv \xi_c$. 

We start by observing that, for a given $\xi_c$, the boundary curve $E^b_j(X,\xi_c)$ closes in on the origin as $X$ increases from 2 to $X_{max}$. For $X = X_{max}'$, the boundary curve collapses completely, whereas for $X = 2$, the boundary curve encloses the maximum area possible. We will refer to the $X = 2$ boundary curve as "the extremal boundary curve". 

At low temperature, hopping will occur primarily between sites whose energy difference $|\epsilon_i - \epsilon_j|$ is comparable to $kT$. If $t$ is very small ($t << 1$), the extremal boundary curve must lie well inside the triangle. As the temperature rises, the triangle becomes smaller. This in itself is not important. What is important is the relation between the extremal boundary curve and the triangle. If $t$ is very large ($t >> 1$), then hopping will occur essentially just between nearest neighbors: the energy of a site will be unimportant. Thus, as $t$ increases, the extremal boundary curve will move out from the origin and approach the triangle — this is a relative statement, because it turns out that both the triangle and the extremal boundary curve shrink as $t$ increases, but the former shrinks more quickly with increasing $t$ than the latter. 

As $t$ increases, the extremal boundary curve will approach the triangle and will eventually, at $t = t_{top}$, "top out" and just touch the triangle. As
such, when \( t < t_{\text{top}} \), each boundary curve \( E_j^b(X; \xi_c) \), where \( 2 < X < X_{\text{max}} \), will lie wholly inside the triangle, corresponding to case (i) of Fig. 2.

When \( t = t_{\text{top}} \), the point \( E_{\text{j}}^0 \) of the extremal boundary curve lies below the top of the triangle: \( E_{\text{j}}^0 < \frac{1}{2t} \). Increasing \( t \) will eventually bring this point up to the top of the triangle. We call \( t_{\text{mid}} \) the value of \( t \) for which \( E_{\text{j}}^0 = \frac{1}{2t} \). At \( t = t_{\text{mid}} \), the point of the extremal boundary curve at \( b = b_{\text{max}} \) will have \( E_j^b = \frac{b_{\text{max}}}{2} \) and will lie below the top of the triangle: \( \frac{b_{\text{max}}}{2} < \frac{1}{2t} \). Increasing \( t \) will eventually result in this point moving to the upper left corner of the triangle. \( t = t_{\text{full}} \) is the temperature at which the extremal boundary curve just fully encloses the triangle.

Now we need to discuss the various boundary curves which correspond to values of \( X \) exceeding 2. Recall that, as \( X \) increases from 2 to \( X_{\text{max}} \), the boundary curve closes toward the origin.

For \( t < t_{\text{top}} \), each boundary curve for each \( X \) such that \( 2 < X < X_{\text{max}} \) will lie wholly within the triangle, corresponding to case (i). When \( t_{\text{top}} < t < t_{\text{mid}} \), there is a range of \( X \), \( X_{\text{top}} < X < X_{\text{max}} \), where the boundary curve lies wholly inside the triangle. \( X_{\text{top}} \) is the value of \( X \) for which the boundary curve \( E_j^b(X; \xi_c) \) just tops out. If \( 2 < X < X_{\text{top}} \), the boundary curve will cut through the triangle as in case (ii).

Things are even more complicated for \( t_{\text{mid}} < t < t_{\text{full}} \). Once again, there exists an \( X_{\text{top}} \) such that for \( X_{\text{top}} < X < X_{\text{max}} \), the boundary curves are case (i). But now there is also a value \( X_{\text{mid}} \) such that the boundary curve
$E_b(X_{\text{mid}}; \xi_c)$ has $E_j^0 = \frac{1}{2t}$. That is, for $X_{\text{mid}} < X < X_{\text{top}}$, the boundary curves are case (ii), while for $2 < X < X_{\text{mid}}$, the curves are case (iii).

Finally, if $t_{\text{full}} < t$, we have an $X_{\text{top}}$ and an $X_{\text{mid}}$ as before, but now also an $X_{\text{full}}$. For $X_{\text{top}} < X < X_{\text{max}}$, the boundary curves are case (i); for $X_{\text{mid}} < X < X_{\text{top}}$, the curves are case (ii); for $X_{\text{full}} < X < X_{\text{mid}}$, the curves are case (iii); for $2 < X < X_{\text{full}}$, the curves are case (iv). The boundary curve $E_j^b(X_{\text{full}}, \xi_c)$ just wholly encloses the triangle.

$X_{\text{top}}$ is given in terms of $\xi_c$ via equations (IV.3) and (IV.6) by replacing $E_j^b$ by $\frac{1}{2t}$ and $\xi$ by $\xi_c$. The two equations resulting may be written as follows:

\[
2\cosh\left(\frac{1}{2t} - \frac{b}{2}\right)(\sinh \frac{b}{2} - \frac{b}{2}\cosh \frac{b}{2}) + \sinh b - b \cosh b + b \sinh \frac{b}{2} \sinh \left(\frac{1}{2t} - \frac{b}{2}\right) = 0, \quad \text{(VI.1a)}
\]
\[
2\cosh\left(\frac{1}{2t} - \frac{b}{2}\right) \sinh \frac{b}{2} + \sinh b - b \frac{2}{X_{\text{top}}} e^{-\xi_c-X_{\text{top}}} = 0. \quad \text{(VI.1b)}
\]

$b$ is solved for in terms of $t$ via equation (VI.1a) and then used in equation (VI.1b) to give $X_{\text{top}}$.

We use equation (IV.5) to obtain the equation for $X_{\text{mid}}$:

\[
X_{\text{mid}}^2 e^{\xi_c-X_{\text{mid}}} = 2 + 2\cosh \frac{1}{2t}. \quad \text{(VI.2)}
\]
$X_{\text{full}}$ is given by putting $\xi = \xi_c$ and $b_{\text{max}} = \frac{1}{t}$ into equation (IV.4):

$$X_{\text{full}}^2 e^{\xi_c - X_{\text{full}}} = 2t \sinh \frac{1}{t} + 4t \sinh \frac{1}{2t}.$$  \quad (VI.3)
VII. Summary of Equations

In summary, $c_c(t)$ is specified by the following equations:

\[ B_c(t) = B(c_c), \quad (VII.1) \]
\[ B(c_c) = \frac{\pi}{2} \rho a^2 \int_0^{X_{\text{max}}} X p(X; c_c) dX, \quad (VII.2) \]

and

\[ X_{\text{max}}^2 e^{c_c - X_{\text{max}}} = 4, \quad (VII.3) \]

where $p(X; c_c)$ depends upon the $t$-range and the $X$-value for that $t$-range as follows:
\[ p(X; \xi_c) = \]

1, if \( t > t_{\text{full}} \) and \( 2 < X < X_{\text{full}} \);

\[ t_{\text{mid}} < t < t_{\text{full}} \) and \( 2 < X < X_{\text{mid}} \);

\[ I(b_1, b_{\text{max}}) + tb_1(2-tb_1), \] if \( t > t_{\text{full}} \) and \( X_{\text{full}} < X < X_{\text{mid}} \);

\( t < t_{\text{top}} \) and \( 2 < X < X_{\text{max}} \);

\[ t_{\text{top}} < t < t_{\text{mid}} \) and \( X_{\text{top}} < X < X_{\text{max}} \);

\[ I(0, b_{\text{max}}) + I(b_h, b_{\text{max}}) + 2t(b_h - b_h) - t^2(b_h^2 - b_h^2), \] if \( t_{\text{mid}} < t < t_{\text{full}} \) and \( X_{\text{mid}} < X < X_{\text{top}} \);

\( t > t_{\text{full}} \) and \( X_{\text{full}} < X < X_{\text{mid}} \);

\( t < t_{\text{top}} \) and \( 2 < X < X_{\text{max}} \);

\( t_{\text{top}} < t < t_{\text{mid}} \) and \( X_{\text{top}} < X < X_{\text{max}} \);

\[ I(0, b_{\text{max}}), \] if \( t_{\text{mid}} < t < t_{\text{full}} \) and \( X_{\text{top}} < X < X_{\text{max}} \);

\( t > t_{\text{full}} \) and \( X_{\text{top}} < X < X_{\text{max}} \);

(VII.4)
I(b_1,b_2) is given by equation (V.2), b_{\text{max}} by equation (IV.4), b, b_h and b_l by equation (V.4), X_{\text{top}} by equations (VI.1), X_{\text{mid}} by equation (VI.2) and X_{\text{full}} by equation (VI.3). t_{\text{top}}, t_{\text{mid}} and t_{\text{full}} are given by:

\begin{equation}
E_{j}^{b} \mid_{X=2, \xi = \xi_c(t_{\text{top}})} = \frac{1}{2t_{\text{top}}}, \quad (VII.5)
\end{equation}

\begin{equation}
E_{j}^{0} \mid_{X=2, \xi = \xi_c(t_{\text{mid}})} = \frac{1}{2t_{\text{mid}}}, \quad (VII.6)
\end{equation}

and

\begin{equation}
b_{\text{max}} \mid_{X=2, \xi = \xi_c(t_{\text{full}})} = \frac{1}{t_{\text{full}}}, \quad (VII.7)
\end{equation}

These equations may be solved numerically for not only \( \xi_c(t) \) but also, if desired, \( t_{\text{top}}, t_{\text{mid}}, t_{\text{full}}, X_{\text{max}}(t), X_{\text{top}}(t), X_{\text{mid}}(t) \) and \( X_{\text{full}}(t) \), provided one has an expression for \( B_c(t) \). We next describe how we obtained such an expression.
VIII. Form Chosen for $B_c(t)$

Our basic strategy for obtaining an expression for $B_c(t)$ with $t$ of order unity is to first determine high $t$ and low $t$ asymptotic forms for $B_c(t)$, then to interpolate between the two.

If $t \gg 1$, the resistor network (1.3) reduces to

$$R_{ij} = 4R_0 X_{ij}^{-2} e^{i j}, \quad (VIII.1)$$

and the temperature dependence essentially disappears. For the resistor network $R_{ij} = R_0 e^{i j}$, $B_c$ is approximately 4.5 in two dimensions. The factor $X_{ij}^{-2}$ in the resistor network (VIII.1) will be dominated by the exponential factor $e^{i j}$. As such, we take $B_c$ to be approximately 4.5 when $t \gg 1$.

If $t \ll 1$, the resistor network (1.3) may be replaced by a reduced form which has been studied by many investigators:

$$R_{ij} = R_0 e^{i j}, \xi_{ij} = f_{ij} + X_{ij}, \quad (VIII.2)$$
where

\[ f_{ij} = \frac{\varepsilon_{ij}}{kT}, \quad \varepsilon_{ij} = \varepsilon_j - \varepsilon_i \text{ if } \varepsilon_j > 0 > \varepsilon_i, \quad (VIII.3) \]

\[ -\varepsilon_i \text{ if } 0 > \varepsilon_j > \varepsilon_i. \]

(The zero of energy has once again been chosen to coincide with the chemical potential.) Imposing \( R_{ij} < R_0 e^\varepsilon \) and looking for percolation leads directly to the Mott law for variable range hopping. For two dimensions, \( \xi_c \) may be written in the form

\[ \xi_c(t) = \left( \frac{2n_c(2)^{1/3}}{\rho_s a^2 t} \right), \quad (VIII.4) \]

where \( n_c^{(2)} \) was recently given as

\[ n_c^{(2)} = 6.9 \pm 0.4. \quad (VIII.5) \]

It is a straightforward matter to proceed from \( \xi_{ij} < \xi_c(t) \) to derive an
expression for \( B_c(t) \). We first note that \( 0 < X_{ij} < \xi_c \). We also see that two sites a distance \( X \) apart are connected if

\[
\bar{f}_{ij} < \xi_c - X. \tag{VIII.6}
\]

From equation (VIII.3) and relation (VIII.6), we find that the boundary curve here is described by

\[
\begin{align*}
E_j &= \xi_c - X & \text{for } & 0 < E_i < \xi_c - X, \\
E_j &= E_i + \xi_c - X & \text{for } & -(\xi_c - X) < E_i < 0, \tag{VIII.7} \\
E_i &= -(\xi_c - X) & \text{for } & -(\xi_c - X) < E_j < 0.
\end{align*}
\]

These three line segments together with the line segment

\[
E_j = E_i \text{ for } -(\xi_c - X) < E_i < \xi_c - X
\]

determine a trapezoid. Since \( t \ll 1 \), the trapezoid will lie wholly inside the triangle, whence

\[
p(X;\xi_c) = 3t^2(\xi_c - X)^2 \tag{VIII.8}
\]

and, in turn,
\[ B_c(t) = \frac{\pi}{8} \rho_s a^2 t^2 \xi_c^4(t). \]  

(VIII.9)

The extremal trapezoid, corresponding to \( X = 0 \), reaches the triangle, or "tops out", at a critical temperature \( t_c \) given by

\[ \xi_c(t_c) = \frac{1}{2t_c}. \]  

(VIII.10)

\( t_c \) is given by

\[ t_c = \left( \frac{\rho_s a^2}{16n_c(2)} \right)^{1/2}. \]  

(VIII.11)

The standard percolation theory based on resistor network (VIII.2) is valid for \( t \leq 2t_c \) (see [I]). Typically, \( \rho_s a^2 \) will be small, whence \( t_c \) is also small. In this \( t \)-range, we have from equations (VIII.4), (VIII.9) and (VIII.11),

\[ B_c(t) = \frac{\pi}{8} n_c(2) \left( \frac{t}{t_c} \right)^{2/3}. \]  

(VIII.12)

For our purposes, we take \( B_c(t) \) to be given by equation (VIII.12) for \( t \leq t_c \).

In Fig. 3, we plot \( B_c \) vs. \( t^{-1/3} \) for \( \rho_s a^2 = \frac{1}{36} \) (\( t_c = 0.0159 \)).
Fig. 3: The critical number $B_c$ of bonds per site required for percolation as a function of temperature. The high $t$ limit of $B_c$ is approximately 4.5. The low $t$ asymptotic form for $B_c$ is indicated by the dashed curve: the curve for $B_c$ will merge onto this low $t$ curve when $t$ is very small. The curves labelled "1", "2" and "3" are described in the text.
We choose this value for $t_a^2$ to coincide with the value used in [I]. The high temperature limit of $B_c (\approx 4.5)$ and the low temperature asymptotic form of $B_c$, shown in Fig. 3, suggest a rough form for $B_c$ in the intermediate temperature range. However, we will need to appeal to the numerical work of [I] to determine a suitable curve for $B_c (t)$.

Fig. 6 of [I] suggests that the hopping is essentially nearest neighbor hopping for $t^{-1/3} \lesssim 2.4$. As such, the curve for $B_c (t)$ should be fairly flat and close to the high $t$ limiting value ($\approx 4.5$) for $t^{-1/3} \lesssim 2.4$. For $2.4 \lesssim t^{-1/3} \lesssim t_c^{-1/3}$, the curve for $B_c (t)$ must decrease with increasing $t^{-1/3}$ and, at low enough $t$, eventually merge onto the curve for the low $t$ asymptotic form. Three possible $B_c$ curves having these features are shown in Fig. 3 for $1 < t^{-1/3} < 3.47$ ($1 > t > 0.024$). The curve labelled "1" consists of a line segment which is nearly tangent to the low $t$ asymptotic form at $t = t_c$ and extends from $(t_c^{-1/3}, B_c) = (3.47, 3.26)$ to $(2.3, 4.5)$, plus a line segment with $B_c = 4.5$ from this point to $t^{-1/3} = 1$. The curve labelled "2" consists of a segment going from $(3.47, 3.6)$ to $(2.4, 4.3)$, plus a second segment going from this point to $(1, 4.4)$. Curve "3" consists of five line segments: one from $(3.47, 3.26)$ to $(3.1, 3.7)$, another extending to $(2.7, 4.05)$, another going to $(2.4, 4.25)$, another proceeding to match up to curve "2" at $t^{-1/3} = 2.1$ and a final segment which coincides with curve "2" for $1 < t^{-1/3} < 2.1$.

We determined $\xi_c (t)$ for each of these three curves for $B_c (t)$. 
Satisfactory agreement between the analytic theory and the numerical results was obtained with curve "3". We next present the results of our analytic investigation and compare the analytic theory to the numerical simulation described in [I].
IX. Comparison of Analytic Theory and Numerical Simulation

In [1], we performed a numerical simulation to determine the overall current in a random resistor network with resistances (1.3). We calculated the average current $I(t)$ for several $t$ values for samples of roughly 1000, 2000 and 2800 sites. In Figs. 4 and 5, we compare the numerical results for $\ln I$ to the analytic result for $-\xi_C$. The current $I$ in these figures is expressed in units where the voltage drop across the sample is one unit of voltage and $R_0$ is the unit of resistance. $\ln I$ and $-\xi_C$ are plotted against $t^{-1}$ in Fig. 4 and against $t^{-1/3}$ in Fig. 5. It was revealed in [1] that the numerical results for the 1000- and 2000-site systems agree within the standard deviations in their $I(t)$ values. As such, we see from Figs. 4 and 5 that the analytic theory curve agrees quite well with the numerical points for the 1000- and 2000-site systems, and that this is especially so at higher values of $t$. The solid numerical points in Fig. 4 designate the points in a similar figure in [1] which indicated $\ln I$ to be linear in $t^{-1}$ over the $t$-range: $0.04 \leq t \leq 0.2$. We see in Fig. 4 that the analytic curve reproduces this linear regime, in agreement with the numerical results. Finally, it was pointed out in [1] — and it is revealed again in Figs. 4 and 5 — that the numerical points for the 2800-site systems lie below the points for the 1000- and 2000-site systems. As discussed in [1], this is apparently due to a shape-dependence of $I(t)$; the 1000- and 2000-site systems were square whereas the 2800-site systems were rectangular, long in the direction of the potential drop. This apparent shape dependence will be the topic of a future publication.
Fig. 4: Comparison between the average $I$ of the overall current, obtained via the numerical simulation, and the analytic theory result $\xi_c$. The boxes indicate an average of 4 1000-site systems, the circles an average of 4 2000-site systems and the triangles an average of 7 2800-site systems. The shaded symbols indicate the points for which $\ln I$ is taken to be linear in $t^{-1}$. The solid curve represents the analytic curve for $-\xi_c(t)$. 
Fig. 5: Same as Fig. 4 but without shaded symbols and with $t^{-1/3}$ as the abscissa.
The analytic theory goes beyond the numerical simulation: it provides us with information as to how the typical hopping distance, and the typical energy change involved in a hop, vary with temperature. $X_{\text{max}}$ provides a good indication of the typical hopping distance. In Fig. 6, we plot $X_{\text{max}}$ vs. $t^{-1/3}$. Previously, we have suggested that the hopping for $t^{-1/3} < 2.4$ was essentially nearest neighbor hopping whereas for $t^{-1/3} > 2.4$ the hopping was over greater and greater distances until, at very low $t$ ($t \leq t_c = 3.98$) the hopping would be variable range. These general ideas are borne out in Fig. 6. Recalling that $X_{ij} \equiv \frac{2r_{ij}}{a}$ and that $\frac{r_{s-s}}{a} = 6$, and defining $r_{\text{max}} = \frac{a}{2} X_{\text{max}}$, we see from Fig. 6 that whereas the hopping distance $r_{\text{max}}$ increases from $1.2r_{s-s}$ at $t^{-1/3} = 1$ to only about $1.4r_{s-s}$ at $t^{-1/3} = 2.4$, $r_{\text{max}}$ further increases to about $2.1r_{s-s}$ at $t^{-1/3} = 3.47$. Also shown in Fig. 6 is the extrapolation of the low $t$ form for $r_{\text{max}}$: 

$$r_{\text{max}} = \frac{a}{2} \xi_c(t),$$

with $\xi_c(t)$ given by equation (VIII.4). As seen in the figure, $r_{\text{max}}$ "overshoots" this asymptote as $t^{-1/3}$ increases from 1 to about 2, instead of merging directly onto the asymptote. As such, $r_{\text{max}}$ remains close to its high $t$ value, and the hopping remains essentially nearest neighbor hopping, down to lower temperatures than it would have had the merging been direct. As $t^{-1/3}$ increases past 3.47, the curve for $\xi_c(t)$ (Fig. 3) will eventually merge onto the low $t$ asymptotic curve for $\xi_c(t)$, and so the curve for $X_{\text{max}}(t)$ will merge onto the low $t$ asymptotic curve for $X_{\text{max}}(t)$.

We can combine our result for $\xi_c(t)$ with equation (1.3) for $R_{ij}$ to get
Fig. 6: Temperature dependence of $X_{\text{max}}$. The solid curve indicates the temperature dependence of the typical hopping distance $\bar{r}(t)$ via

$$X_{\text{max}}(t) = 12\bar{r}(t)/r_{s-s},$$

where $r_{s-s}$ is a measure of the site-to-site distance (see text for details). The dashed line is an extrapolation of the low $t$ asymptotic form of $X_{\text{max}}(t)$: the curve for $X_{\text{max}}$ will merge onto this low $t$ curve when $t$ gets very small.
an idea of the temperature dependence of the typical hopping energy. Taking the hop to be one connecting neighboring sites with energies equally removed from but on opposite sides of $\mu_0$ (i.e. $E_j = -E_i > 0$), setting $R_{ij} = R_0 e^{\xi_c(t)}$ and writing $X_{ij} = X$, we have

$$\frac{1 + e^{-E(t)/2^2}}{e^{E(t)/2^2}} \left(1 - e^{-E(t)/2^2}\right) = X^2 e^{\xi_c(t) - X}, \quad \text{(IX.1)}$$

where $\bar{E}(t) \equiv \frac{\tilde{\varepsilon}_\text{typ}(t)}{t}$ and $\tilde{\varepsilon}_\text{typ}(t)$ is the typical hopping energy (i.e. $\tilde{\varepsilon}_\text{typ}(t) = \tilde{\varepsilon}_j - \tilde{\varepsilon}_i = 2\tilde{\varepsilon}_j$; see equation (I.5)). In Fig. 7, we plot $\tilde{\varepsilon}_\text{typ}$ vs. $t^{-1/3}$ for $X = \frac{2.4r}{a} s - s$ (the high $t$ limit of $X_{\text{max}}$). No curve for $\tilde{\varepsilon}_\text{typ}$ is presented for most of the $t$-range corresponding to nearest neighbor hopping. This is because $\tilde{\varepsilon}_\text{typ}$ depends very sensitively on $\xi_c$ and $X$ in this $t$-range. For example, slight variations in $X$ in equation (IX.1) result in large variations in $\tilde{\varepsilon}_\text{typ}$ in the nearest neighbor hopping regime. A small ($\approx 1\%$) fractional increase (decrease) in $X$ results in an order of magnitude shift downwards (upwards) in the value of $\tilde{\varepsilon}_\text{typ}$ at $t = 1$. Such behavior reveals that there is no typical hopping energy in the nearest neighbor hopping regime: the temperature is high enough that none of the sites may be considered to be unimportant because of its energy. However, slight variations in $X$ in equation (IX.1) result in only slight changes in $\tilde{\varepsilon}_\text{typ}$ for
Fig. 7: Temperature dependence of the typical hopping energy $\tilde{\varepsilon}_{\text{typ}}$ (solid curve) obtained from equation (IX.1) using $X = 14.34$. The dashed curve is an extrapolation of the low $t$ asymptotic form of $\tilde{\varepsilon}_{\text{typ}}$. 
temperatures lower than those corresponding to the nearest neighbor hopping regime. In this lower temperature range, where the hops are over greater distances, the energies of the sites do matter and we may obtain a meaningful curve for \( \tilde{\epsilon}_{\text{typ}} \). The precise curve obtained depends on the precise value of \( X \) used, but the general shape of the curve — decreasing with increasing \( t^{-1/3} \) — always emerges. The low \( t \) asymptotic curve is also shown in Fig. 7; as \( t \) gets very small, the curve for \( \tilde{\epsilon}_{\text{typ}}(t) \) will eventually merge onto the low \( t \) asymptotic curve. We note, in closing, that the same conclusions are reached if we consider two sites with energies on the same side of the chemical potential.
X. Summary

We have used percolation theory to obtain a set of analytic equations which determine the overall resistance of a random resistor network with resistances \( R_{ij} \) specified by equation (1.3). The theory applies to a two-dimensional system with a flat density of states of width \( \Delta\varepsilon \) and is valid for temperatures where \( t \) is of order unity (\( kT \) of order \( \Delta\varepsilon \)). We were able to get agreement between the theory and the results of a numerical simulation performed in [I] (Figs. 4 and 5). We did so by using a very reasonable curve for the temperature dependence of the critical number of bonds (resistors) per site \( B_c(t) \) required for percolation (Fig. 3). For a number density of sites \( \rho_s \) and localization length \( a \) related by \( \rho_s a^2 = \frac{1}{36} \), the analytic theory confirms the following findings of the numerical simulation: that the overall resistance is proportional to \( e^{c_3/kT} \) for \( 0.04 \leq t \leq 0.2 \) with \( c_3 = 0.28\Delta\varepsilon \), that the hopping is essentially just nearest neighbor hopping for \( t^{-1/3} \leq 2.4 \) and that the hops are of longer range for \( t^{-1/3} \geq 2.4 \). The analytic theory provided more information about the high temperature range and intermediate temperature range hopping process. We were able to determine the temperature dependences of the typical hopping distance and the typical hopping energy (Figs. 6 and 7).

In the next paper in this series, we will extend the analytic theory to three dimensions. The final paper of this series will compare the predictions of the three-dimensional theory for various densities of states, other than flat, to experimental results for VRH in LDSs; these experimental results show VRH in a temperature range well in excess of the temperature
range where established theories predict VRH, and have \( T_0 \) values which are much smaller than the theoretical \( T_0 \) values.
In this appendix, we describe a much simplified version of our analytic theory. We begin by reporting that the boundary curve $E_j b$ asymptotically approaches a trapezoidal shape as $x^2 e^{-x} \to \infty$. The trapezoid is given by the four line segments:

\[
\begin{align*}
E_j &= b_{\text{max}} \quad \text{for} \quad 0 < E_i < b_{\text{max}}, \\
E_j &= E_i + b_{\text{max}} \quad \text{for} \quad -b_{\text{max}} < E_i < 0, \\
E_i &= -b_{\text{max}} \quad \text{for} \quad -b_{\text{max}} < E_j < 0, \\
E_j &= E_i \quad \text{for} \quad -b_{\text{max}} < E_i < b_{\text{max}}.
\end{align*}
\]

where $b_{\text{max}}$ is given by equation (IV.4). The approximate version of our theory is based on replacing the boundary curve by this trapezoid.

For a given $\xi_c$, the trapezoid "tops out" when $b_{\text{max}} = \frac{1}{2t}$; this defines a new $X_{\text{top}}$, given by

\[
X_{\text{top}}^2 \xi_c - X_{\text{top}} = 4t \sinh \frac{1}{2t} + 8t \sinh \frac{1}{4t}.
\]  
(A.1)

We have used equation (IV.4) and have put $b_{\text{max}} = \frac{1}{2t}$. Similarly, the trapezoid first fully encloses the triangle when $b_{\text{max}} = \frac{1}{t}$; this defines $X_{\text{full}}$, which is still given by equation (VI.3):
The temperature \( t_{\text{top}} \) is now given by

\[
\frac{X_{\text{full}}}{2} e^{\frac{\xi_c - X_{\text{full}}}{1}} = 2t \sinh \left( \frac{1}{t} + 4t \sinh \left( \frac{1}{2t} \right) \right).
\]  
(A.2)

while the temperature \( t_{\text{full}} \) is still given by equation (VII.7):

\[
\left. b_{\text{max}} \right|_{X=2, \, \xi=\xi_c(t_{\text{top}})} = \frac{1}{2t_{\text{top}}},
\]  
(A.3)

\[
\left. b_{\text{max}} \right|_{X=2, \, \xi=\xi_c(t_{\text{full}})} = \frac{1}{t_{\text{full}}},
\]  
(A.4)

\( X_{\text{mid}} \) and \( t_{\text{mid}} \) do not appear in the simplified version of the theory.

With these simplifications, the following expression for \( p(X;\xi) \) is easily obtained:
\[ p(X;\xi) = \]

1, if \( t > t_{\text{full}} \) and \( 2 < X < X_{\text{full}} \),

\[ t > t_{\text{full}} \) and \( X_{\text{full}} < X < X_{\text{top}} \),

\[ 2 - t_{\text{max}} \), if \( t_{\text{top}} < t < t_{\text{full}} \) and \( 2 < X < X_{\text{top}} \),

\[ t > t_{\text{full}} \) and \( X_{\text{top}} < X < X_{\text{max}} \), \quad (A.5)

\[ 3t_{\text{max}}^2 \), if \( t_{\text{top}} < t < t_{\text{full}} \) and \( X_{\text{top}} < X < X_{\text{max}} \),

\[ t < t_{\text{top}} \) and \( 2 < X < X_{\text{max}} \).

The full set of equations for the approximate theory is as follows:

\[ B_c(t) = B(\xi_c), \quad (A.6) \]

\[ B(\xi_c) = \frac{\pi}{2} \rho a^2 \int_{2}^{X_{\text{max}}} X p(X;\xi_c) dX, \quad (A.7) \]

\[ X_{\text{max}}^2 e^{\xi_c - X_{\text{max}}} = 4, \quad (A.8) \]

with \( p(X;\xi_c) \) given by equations (A.5), \( t_{\text{top}} \) and \( t_{\text{full}} \) by equations (A.3) and (A.4), \( X_{\text{top}} \) and \( X_{\text{full}} \) by equations (A.1) and (A.2), and \( b_{\text{max}} \) by equation (IV.4):
\[
\frac{\sinh b_{\text{max}}}{b_{\text{max}}} + \frac{\sinh(b_{\text{max}}/2)}{b_{\text{max}}/2} = \frac{x^2 e^{\xi-X}}{2}.
\]  

(A.9)

Given an expression for $B_c(t)$, the equations in this appendix may be solved numerically for $\xi_c(t)$ and also, if desired, $t_{\text{top}}$, $t_{\text{full}}$, $X_{\text{max}}(t)$, $X_{\text{top}}(t)$ and $X_{\text{full}}(t)$.

In Fig. 8, we compare the $\xi_c(t)$ obtained via the exact theory to the $\xi_c(t)$ using the approximate version of the theory. Curve "3" of Fig. 3 for $B_c(t)$ was used for both the exact and approximate versions of the theory. Fig. 8 reveals that the two versions of the theory agree to within less than about 5\% for $t^{-1/3} < 3.47$.

We also note that when curve "2" for $B_c(t)$ is used, the approximate version of the theory agrees with the numerical results almost as well as the exact version of the theory does when using curve "3" for $B_c(t)$.

ACKNOWLEDGMENTS

We wish to thank Peter C.W. Holdsworth and Matthew Choptuik for some helpful discussions. This work was supported financially by NSERC Operating Grant A713 and by the Graduate Student Fellowship Programme at UBC.
Fig. 8: Comparison between curves for $\xi_c$ obtained via the exact analytic theory (lower curve) and the approximate analytic theory (upper curve).
REFERENCES

*Present Address: Department of Physics, Dalhousie University, Halifax, N.S., CANADA B3H 3J5

10. Considerations similar to those in this appendix may be found in either K.J. Hayden and P.N. Butcher, Philos. Mag. B 38, 603 (1978), or P.N. Butcher and K.J. Hayden, Philos. Mag. 36, 657 (1977).
Hopping conductivity in lightly doped semiconductors - III:

Three-dimensional, high temperature analytic theory

for flat density of states

Mark R.A. Shegelski* and Robert Barrie

Department of Physics, University of British Columbia,
Vancouver, B.C., CANADA V6T 2A6

Abstract

In previous works, we performed a numerical simulation of and presented an analytic theory for a two-dimensional lightly doped semiconductor. These were first steps towards understanding certain experimental results for lightly doped n-GaAs and n-InP, in which the conductivity \( \sigma \) was found to be

\[
\sigma = \sigma_0 e^{-\left(\frac{T_0}{T}\right)^s}
\]

with \( s \) close to \( \frac{1}{4} \), indicating variable range hopping (VRH). The experimental range of temperatures \( T \), \( 1K \lesssim T \lesssim 7K \), was well above, and the experimental \( T_0 \) values well below, the limits set by established theories of VRH. In this paper, we extend our analytic theory to three dimensions. As before, we choose a flat density of states with width \( \Delta \varepsilon \) and model the semiconductor as a Miller and Abrahams type resistor network; we again use the full form of the resistance and do not take the low temperature asymptotic form because we are interested in temperatures where
the reduced temperature \( t = \frac{kT}{\Delta \varepsilon} \) is of order unity. One of the major results of this paper is that we establish that the small experimental \( T_0 \) values may not be accounted for using a flat density of states in conjunction with the full resistor network model. We also demonstrate that, in the high temperature regime, our theory predicts a conductivity of the form
\[
\sigma = \sigma_0 e^{-\varepsilon_3/kT}
\]
where the activation energy \( \varepsilon_3 = 0.20\Delta \varepsilon \) for three dimensions. For two dimensions, we found \( \varepsilon_3 = 0.28\Delta \varepsilon \). Our values for \( \varepsilon_3 \) differ considerably from those reported by Skal, Shklovskii, and Efros and Hayden and Butcher. We believe that our results differ from those of these other workers because of their use of the low temperature form of the resistances in the high temperature regime. We also investigate the effect of a usually neglected \( Q_{ij} \) factor. We find that inclusion of this factor in the analysis produces a trend of decreasing \( \frac{\varepsilon_3}{\Delta \varepsilon} \) with increasing \( \Delta \varepsilon \), although \( \varepsilon_3 \) itself increases with \( \Delta \varepsilon \); taking the \( Q_{ij} \) factor into consideration does not lead to an explanation of the small \( T_0 \) values. We conjecture that a density of states other than flat will be required for such an explanation.

PACS numbers: 72.20.-i
I. Introduction

Variable range hopping (VRH) in disordered systems, such as amorphous semiconductors or doped semiconductors, is characterized by a conductivity \( \sigma \) of the form

\[
\sigma = \sigma_0 e^{-(T_0/T)^{1/4}},
\]

where \( \sigma_0 \) depends weakly on the temperature \( T \). Mott \(^1\) and others \(^2,3\), by assuming a density of states flat in the vicinity of the Fermi level, have provided theoretical derivations of equation (I.1), finding

\[
T_0 = \frac{C}{k g_p a^3},
\]

where \( C \) is a dimensionless constant approximately equal to 20, \( k \) is Boltzmann's constant, \( g_p \) is the density of states at the Fermi level and \( a \) is the electron localization length.

In a previous paper \(^4\), hereafter referred to as [I], we reported that there exist experimental results \(^5,6\) showing unequivocal VRH in lightly doped semiconductors (LDSs) and that this experimental VRH has at present no suitable theoretical explanation. The experimental VRH was observed in a temperature range well above the temperature range where existent theories have predicted VRH and the experimental \( T_0 \) values were found to be much smaller than the theoretical values.
This paper represents the third of four sequential investigations which comprise a program designed to understand these experimental results. The program is based on the resistor network model of a LDS, which was presented in [I]. In this model, the LDS is replaced by a network of resistors. Every pair of donors in the LDS is mapped into a resistor $R_{ij}$. In [I], we performed a numerical simulation in two dimensions to calculate the overall resistance of the network. To simplify matters, we took the number density $n_D$ of donors to be twice the number density $n_A$ of acceptors, which simplified the chemical potential $\mu$ to $\mu = \mu_0 - kT \ln 2$. We chose a flat density of states with bandwidth $\Delta \varepsilon$ and took $\mu_0 = 0$. The resistor network simplified to

$$R_{ij} = R_0 \frac{Q_{ij} f(E_i, E_j)}{X_{ij}^2} e^{X_{ij}},$$

where $R_0$ was a constant resistance,

$$Q_{ij} = \left[ 1 + \Delta^2 (E_j - E_i)^2 t^2 \right]^4,$$

and

$$f(E_i, E_j) = \frac{(1+e^{-E_j})(1+e^{-E_i})}{(E_j - E_i)(E_j - E_i - 1)}.$$
\[ \lambda \equiv \frac{a \Delta \varepsilon}{2 \pi s}, \quad E_i \equiv \frac{e_i}{t}, \quad \tilde{e}_i \equiv \frac{e_i}{\Delta \varepsilon}, \quad t \equiv \frac{kT}{\Delta \varepsilon}, \quad X_{ij} \equiv \frac{2r_{ij}}{a}, \quad (I.6) \]

where \( \pi = \frac{h}{2\pi} \), \( h \) is Planck's constant, \( s \) is the speed of sound, \( e_i \) is the energy of donor site \( i \), \( r_{ij} \) is the distance separating donor sites \( i \) and \( j \), and by definition \( E_j > E_i \). (Although \( Q_{ij} \) does not depend on the temperature, it proves to be convenient to write \( Q_{ij} \) in the form (I.4); we elaborate on this point later.) Equation (I.3) may be derived using the deformation potential approximation for the electron-phonon interaction.

In a subsequent paper, hereafter referred to as [II], we again calculated the overall resistance of the resistor network (I.3), using an analytic instead of a numerical approach. We formulated the calculation in terms of a percolation problem. The theory was constructed for a two-dimensional system with a flat density of states and with \( n_D = 2n_A \). In both [I] and [II], since the temperature range in the experiments of ref.s [5] and [6] was such that \( kT \) was not necessarily small compared to \( \Delta \varepsilon \), we took \( t \) to be of order unity. This necessitated that we not make the usual simplification of replacing \( R_{ij} \) by its low temperature asymptotic form. In consequence, the analytic theory required the use of a temperature dependent average number of bonds (resistors) per site, \( B_c(t) \), required for percolation. By choosing a reasonable form for \( B_c(t) \), we were able to obtain agreement between the analytic theory and the results of the numerical simulation.
In this paper, we use the knowledge gained from our study of two dimensions to extend our analytic theory to three dimensions. Once again, we consider a flat density of states and we take \( n_D = 2n_A \). One of our main objectives in this paper is to determine whether or not the small \( T_0 \) values observed experimentally can be accounted for by using the full resistor network (1.3) and a flat density of states. To this end, we begin by summarizing in Sec. II the essential features of our theory and outlining the changes required to extend the theory from two to three dimensions; we will show that the overall resistance of the network depends on a critical resistance \( R_c = R_0 e^{-\xi_c(t)} \). We will also see that extending the theory to three dimensions comes down to determining a form for \( B_c(t) \). In Sec. III we indicate how \( B_c(t) \) may be specified. As in [I] and [II], we replace \( Q_{ij} \) by unity and obtain a result for \( \xi_c(t) \). We do not find a temperature range which is wide enough wherein \( T_0 \) is small enough to account for the experimental results; instead, as was the case in our two-dimensional study, we find for three dimensions and a flat density of states that \( \xi_c(t) \) is linear in \( t^{-1} \) for \( t \) of order unity. In Sec. IV, we bring into consideration the \( Q_{ij} \) factor and study the effects of this term on the form of \( \xi_c(t) \). We present a summary of our investigation for three dimensions and a flat density of states in Sec. V.

In order that this paper be self-contained, it has been necessary in this introduction, and it will continue to be necessary in Secs. II and III, to repeat in brief some of the essential concepts discussed in [I] and [II].
Even so, the work presented herein will be understood more fully if this paper is read in conjunction with [I] and [II].
II. Essential Elements of the Theory in Three Dimensions

In this section, we review the basic aspects of our theory, pointing out the modifications required to go from two dimensions to three.

The basic method underlying our theory is the percolation method. In the percolation method, one imagines removing all of the resistors except those which satisfy

\[ R_{ij} < R_0 e^{\xi} \tag{II.1} \]

for some arbitrarily chosen value of \( \xi \). The remaining network of resistors will percolate only if \( \xi \) is equal to or exceeds a critical value \( \xi_c \).

According to percolation theory, the overall conductivity \( \sigma \) of the original network is given by

\[ \sigma = \sigma_0 e^{-\xi c} \tag{II.2} \]

where \( \sigma_0 \) depends only weakly on the temperature \( t \).

Combining equation (I.3) for \( R_{ij} \) with the inequality (II.1) gives

\[ Q_{ij} f(E_i, E_j) X_{ij}^{-2} e^{\frac{X_{ij}}{\xi}} < e^{\xi}. \tag{II.3} \]

We seek the critical value \( \xi_c \) of \( \xi \) for which the network of resistors specified by inequality (II.3) just percolates. Since the \( E_i \) and \( E_j \) in
inequality (II.3) involve the temperature $t$, $\xi_c$ will be a function of $t$. If we know $\xi_c(t)$, we have, via equation (II.2), the dominant temperature dependence of the overall conductivity $\sigma$.

Determining $\xi_c(t)$ begins with the observation that inequality (II.3) specifies, in terms of $\xi$, which pairs of sites will be connected by a resistor. As such, we can determine $\xi_c(t)$ by proceeding as follows. We imagine choosing a site at random. The probability $p(X)$ that there will be another site between $X$ and $X + dX$ away from the chosen site depends on the number of dimensions. For three dimensions,

$$p(X) = \frac{\pi}{2} \rho_s a^3 X^2 dX,$$

where $\rho_s$ is the number of sites per unit volume. Given that there is indeed a site between $X$ and $X + dX$ away from the chosen site, let $p(X;\xi)$ denote the probability that the energies of the two sites are such that the two sites are connected. $p(X;\xi)$ can be determined from inequality (II.3). The result does not depend on the number of dimensions. Once we know $p(X;\xi)$, we can easily calculate the average number $B(\xi)$ of resistors, or "bonds", that are connected to a site. By $B(\xi)$, we mean the ratio of twice the total number of resistors admitted by inequality (II.3) to the total number of sites. It is clear that, for three dimensions,
\[ B(\xi) = \frac{\pi}{2} p_s a^3 \int_{X_{\min}^{(\xi)}}^{X_{\max}^{(\xi)}} x^2 p(x;\xi) dx, \quad (II.5) \]

where \( X_{\min}^{(\xi)} \) and \( X_{\max}^{(\xi)} \) are, respectively, the closest and farthest apart two sites can be in accordance with inequality (II.3). Since the minimum value of \( Q_{ij} f(E, E_j) \) is 4, \( X_{ij} \) is restricted according to

\[ X_{ij}^{-2} \leq e_{ij} < \frac{1}{4} e^\xi. \quad (II.6) \]

Provided \( \xi > 2 \), the inequality (II.6) is satisfied for \( X_{ij} \) such that

\[ X_{\min} < X_{ij} < X_{\max}, \]  

where \( X_{\min} \) and \( X_{\max} \) are the two solutions of \( x^2 e^{\xi-x} = 4 \) \( (X_{\min} < 2 \) and \( X_{\max} > 2) \). Recall that \( X_{ij} = 2 \) corresponds to two sites a distance a apart. Because it is unreasonable to be thinking, in the case of lightly doped semiconductors, in terms of localized electronic states with centers less than about a apart, and since the function \( x^{-2} e^{x} \) is minimized at \( x = 2 \), we take \( X_{\min}^{(\xi)} \equiv 2 \) for all \( \xi > 2 \) and \( 2 < X_{ij} < X_{\max}^{(\xi)} \), where \( X_{\max}^{(\xi)} \) is given by

\[ X_{\max}^{2} e^{\xi - X_{\max}^{\xi}} = 4. \quad (II.7) \]

Equation (II.5) reduces to
\[
B(\xi) = \frac{\pi}{2} \rho S a^3 \int_{\frac{X}{2}}^{X_{\text{max}}} X^2 p(X;\xi) dX. \tag{II.8}
\]

Hereafter, we shall always take \(\xi > 2\) and \(2 < X < X_{\text{max}}\). Once we know \(p(X;\xi)\), we can get \(\xi_c(t)\). The way to do so is to note that if \(B(\xi)\) is too small, there will be no percolation; percolation will occur, at temperature \(t\), only if \(B(\xi) > B_c(t)\), where \(B_c(t)\) is the critical number of bonds per site required for percolation at temperature \(t\). Solving

\[
B(\xi) = B_c(t) \tag{II.9}
\]

will give us \(\xi_c(t)\).

Calculating \(\xi_c(t)\) reduces to determining an expression for \(B_c(t)\). We turn to this task in the next section, where we will also present a result for \(\xi_c(t)\).
III. Form Chosen for $B_c(t)$; Overall Resistance for Flat Density of States in Three Dimensions

To calculate the overall resistance of the resistor network (1.3), we will need an expression for $B_c(t)$ for $t$ of order unity. As such, it will be helpful to start by specifying the high $t$ limit and low $t$ asymptotic form of $B_c(t)$. If $t \gg 1$, the resistor network (1.3) reduces to

$$R_{ij} = R_0 X_{ij}^{-2} e^{i j}, \quad (III.1)$$

where we take $Q_{ij} = 1$ throughout this section. For the resistor network $R_{ij} = R_0 e^{i j}$ in three dimensions, $B_c = 2.77$ (ref. [10]). Since the factor $X_{ij}^{-2}$ will be dominated by the factor $e^{i j}$ in the resistor network (III.1), we take $B_c = 2.77$ for $t \gg 1$. If $t \ll 1$, the resistor network (1.3) may be replaced by the form

$$R_{ij} = R_0 e^{i j + (|\tilde{e}_i| + |\tilde{e}_j| + \tilde{e}_j - \tilde{e}_i)/2t}. \quad (III.2)$$

In three dimensions, $\xi_c$ for this resistor network may be written as

$$\xi_c(t) = \left(\frac{4n_c}{\rho_s a^3 t}\right)^{1/4}, \quad (III.3)$$
where $n_c = 5.3 \pm 0.3$ (ref. [11]). As we saw in [II], for the resistor network (III.2) and a flat density of states we have

$$p(X; \xi_c) = 3t^2(\xi_c - X)^2,$$  \hspace{1cm} (III.4)

provided that $t < t_c$, where $t_c$ is given by

$$\xi_c(t_c) = \frac{1}{2t_c}.$$ \hspace{1cm} (III.5)

Using equations (III.4) and (III.3) in equation (II.8) yields, since $t \ll 1$,

$$B_c(t) = \frac{\pi}{10} n_c(\frac{t}{t_c})^{3/4},$$ \hspace{1cm} (III.6)

where $t_c$ is obtained explicitly upon combining equations (III.3) and (III.5):

$$t_c = \frac{\rho_s a}{4n_c^{1/3}}.$$ \hspace{1cm} (III.7)

Since $\rho_s a^{1/3}$ will be small, for LDSs, $t_c$ will be small. We choose $\rho_s a^{1/3} = 0.0087$, which was the value in the GaAs sample of ref. [5] with $n_A/n_D$
closest to $\frac{1}{2}$. Equation (III.7) then gives $t_c \approx 0.030$.

We next use the knowledge gained from our two-dimensional studies, [I] and [II], to interpolate for $B_c(t)$ between the high $t$ limit $B_c = 2.77$ and the low $t$ ($t < t_c$) asymptotic form (III.6) for $B_c(t)$. Although we anticipate our final result for $\xi_c(t)$ to be reliable, we must acknowledge that, because we are making an extension from two to three dimensions, $\xi_c(t)$ cannot be obtained to arbitrary precision. As such, we use the approximate version of our theory, which was presented in the appendix of [II], and which - as pointed out in [II] - agrees to within a few percent of the exact version of our theory. (We have checked that such agreement also holds in the three-dimensional case; the result is a figure which is very much like Fig. 8 of [II].)

A certain $B_c(t)$ curve was used in conjunction with the approximate version of our theory to give agreement with the numerical results in two dimensions. The essential features of this curve (curve "2" of Fig. 3 of [II]) were as follows.\(^{12}\)

The $B_c(t)$ curve for the high $t$ regime consisted of two line segments. One segment was almost flat, going from the high $t$ limiting value of $B_c$ at $t^{-1/3} = 1$ to about 0.98 of this value at $(t^{-1/3} - 1)/(t_c^{-1/3} - 1) = \frac{1}{2}$ (i.e., approximately halfway from $t^{-1/3} = 1$ to $t^{-1/3} = t_c^{-1/3}$ on the $t^{-1/3}$ axis). The flatness of this segment revealed the nearest neighbor hopping character of this $t$-range, and also showed that the infinite temperature limit had essentially been reached by $t = 1$. The second segment dropped further so
that $B_c$ went to about 0.8 of its $t^{-1/3} = 1$ value at

$$(t^{-1/3} - 1)/(t_c^{-1/3} - 1) = 0.8 \ (i.e., \ approximately \ four \ fifths \ of \ the \ way \ from \ t^{-1/3} = 1 \ to \ t^{-1/3} = t_c^{-1/3} \ on \ the \ t^{-1/3} \ axis). \ This \ final \ point \ lay \ close \ to \ the \ extrapolation \ of \ the \ low \ t \ asymptotic \ curve \ for B_c(t).$$

If we extend these features to three dimensions, for the choice $\rho_s a^3 = 0.0087$, we construct a curve which drops linearly from $B_c = 2.77$ at $t^{-1/4} = 1$ to $B_c = 2.71$ at $t^{-1/4} = 1.666$, and then to $B_c = 2.28$ at $t^{-1/4} = 2.175$. This last point, again, lies very close to the extrapolation of the low $t$ asymptotic curve. Since $\rho_s a^3 = 0.0087$, we have, from equation (III.7), $t_c^{-1/4} = 2.413$.

$Q_{ij} = 1$ in this section; this was also the case in [II]. As such, we can apply here the theory that was derived in [II]. Using the approximate version of the theory, we have determined $\xi_c(t)$ for the $B_c(t)$ curve specified. To determine the form of the temperature dependence of $\xi_c(t)$, we have plotted $-\xi_c(t)$ against $t^{-s}$ for various choices of $s$. (We have plotted $-\xi_c(t)$, as opposed to $\xi_c(t)$, against $t^{-s}$ because we wish to maintain the trend established in [I] and [II].) We have found that, for each value of $s$ chosen in the range $\frac{1}{4} < s < 2$, there is some region of the resultant curve which is linear. Of particular interest to us is the case $s = \frac{1}{4}$. We wish to compare a plot of $-\xi_c(t)$ vs. $t^{-s}$, with $s$ close to $\frac{1}{4}$, to the experimental results of refs. [5] and [6]. The essential features of these results are as follows. $\ln \sigma$ was found to be linear in $T^{-s}$ with $s$ close to $\frac{1}{4}$ and for a
temperature range where $T$ was a few Kelvin. For example, for sample 2 of ref. [5] (where n-GaAs was studied), the linearity was found for $1.4K < T < 6K$ with $s = 0.27$. The experimental $T_0$ value (915K) was such that $T_0^{-1/4} < 3.5$. In order to account for this experimental result, we would need to present a plot of $-\xi_c(t)$ vs. $t^{-s}$, with $s$ close to $\frac{1}{4}$, wherein:

1. $-\xi_c(t)$ would be linear in $t^{-s}$ for $t_{\text{min}} < t < t_{\text{max}}$, and where
2. $\frac{t_{\text{max}}}{t_{\text{min}}} \approx \frac{6}{1.4} \approx 4.3$, and where (3) the magnitude $m$ of the slope of this linear region would satisfy $3.5 < mt_{\text{min}}^{-1/4} < 5.1$ or $mt_{\text{max}}^{-1/4} \approx 5.1$.

The procedure has met with only partial success. We show in Fig. 1 one of the better examples; $-\xi_c(t)$ is plotted against $t^{-0.3}$. The linear segment, indicated by a heavy, solid line, has a slope with magnitude $m = 2.7$. The high and low $t$ limits, $t_h$ and $t_\perp$ respectively, of the linear region are $t_h = 0.18$, $t_\perp = 0.10$. As such, $\frac{t_h}{t_\perp} = 1.8$ or a factor of about 2.4 too small.

The $mt_{\text{min}}^{-1/4}$ values fall into the correct range: $mt_h^{-1/4} = 4.2$, $mt_\perp^{-1/4} = 4.8$.

This all indicates that the slope of the linear region is in accord with the experimental result, but that the extent of the linear region is simply not wide enough to fully describe the experimental curve. The dashed line in Fig. 1 is an example of the linearity required to fully explain the small $T_0$ value found by experiment. Since this linearity is not overly removed from the curve for $-\xi_c(t)$, we will examine, in Sec. IV, the effect of the $Q_{ij}$ factor in order to establish unambiguously whether or not the full resistor
Fig. 1: The $-\xi_c$ curve when plotted against $t^{-0.3}$. The heavy, solid line segment has a slope which agrees with experimental results of ref. [5]. However, in order to fully explain the experimental results, the $-\xi_c$ curve would have to coincide with the line segment over a wider $t$-range: the dashed line segment gives an example of the linearity required to account for the experimental results.
network and a flat density of states can explain the experimental findings.

Meanwhile, however, we will describe the results of our plotting $-\xi_c(t)$ vs. $t^{-s}$ for various $s$ between $\frac{1}{4}$ and 2. Although for any such $s$ there is some linear range, because the most extensive linear region is found for $s \approx 1$, and since $\ln \sigma$ has experimentally often been found to be linear in $T^{-1}$ for doped semiconductors (see, for example, ref.s 13 to 16), we present in Fig. 2 our result for $-\xi_c(t)$ plotted against $t^{-1}$. Since $\xi_c(t)$ was determined numerically, we have used a $\chi^2$ minimization routine to obtain the slope of the curve in Fig. 2. The result is $0.200 \pm 0.002$. This corresponds to a conductivity of the form $\sigma = \sigma_0 e^{-\varepsilon_3/kT}$ with activation energy $\varepsilon_3 = 0.20 \Delta \varepsilon$. The temperature range of the linearity in Fig. 2 has been taken as $5 \leq t^{-1} \leq 22$. In our two dimensional studies, we found the linear regime to be $5 \leq t^{-1} \leq 20-25$, and the corresponding value for $\varepsilon_3$ to be $0.28 \Delta \varepsilon$. We see that $\varepsilon_3$ is less in three dimensions than in two. Our results differ from those of Skal, Shklovskii, and Efros, 17,18 who obtained $\varepsilon_3 = \frac{5}{12} \Delta \varepsilon$ for intermediate compensations in either two or three dimensions, and those of Hayden and Butcher, 19 who claimed $\varepsilon_3 = 0.38 \Delta \varepsilon$ for the two-dimensional case. These other workers' findings, however, were obtained using the low temperature, asymptotic form of the resistor network (1.3).

As shown in [I], the full form of the resistance and the low temperature asymptotic form differ appreciably unless $t$ is quite small, below the $t$-range shown in Fig. 2. We believe that our results differ from theirs
Fig. 2: Plot showing linearity of $-\xi$ in $t^{-1}$. The linear region is taken to be $5 \leq t^{-1} \leq 22$. The slope, obtained from a $\chi^2$ minimization routine, is $-0.200 \pm 0.002$, which corresponds to an activation energy $\varepsilon_3 = 0.20 \Delta \varepsilon$. 
because of their use of the low temperature asymptotic form for $R_{ij}$ in a high temperature regime.

We also wish to point out that, although the $t$-range of the linear regime for three dimensions is essentially the same as it was for two dimensions (see [II]), whereas the linear regime in two dimensions corresponded mostly to the "flat" part of the $B_c(t)$ curve and the nearest neighbor hopping regime, in three dimensions it corresponds mostly to the "dropping" part of the $B_c(t)$ curve.

Since we did not find a wide enough temperature range wherein $-\xi_c(t)$ was linear in $t^{-1/4}$ with a small enough slope to account for the small $T_0$ values observed experimentally as reported in ref.s [5] and [6], we have established that for a flat density of states in three dimensions, and where the $Q_{ij}$ factor in equation (1.3) is replaced by unity, there results no temperature range which is sufficiently extensive and wherein the slope is small enough to account for the small, experimental $T_0$ values. We turn next to the task of bringing into consideration the factor $Q_{ij}$.
IV. Effect of the $Q_{ij}$ Factor

As in [II], we will find it convenient to introduce $b = E_j - E_i$ and work with the three quantities $E_i$, $E_j$, $b$ even though only two are independent. Since $E_i < E_j$ and $-\frac{1}{2t} < E_j < \frac{1}{2t}$, the region in the $E_i, E_j$ plane corresponding to allowed pairs of energies is formed by the lines $E_i = -\frac{1}{2t}$, $E_j = \frac{1}{2t}$, $b = 0$; we refer to this region as "the triangle" (see Fig. 3).

To determine $p(\lambda; \xi)$ when $\lambda \neq 0$ ($Q_{ij} \neq 1$), we begin by rewriting $Q_{ij}$ and $f(E_i, E_j)$ in terms of $b$ and $E_j$:

\[ Q(b) = [1 + \lambda^2 b^2 t^2]^4, \quad \text{(IV.1)} \]

\[ f(b, E_j) = 2\left(\frac{\sinh b}{b} + \frac{\sinh (b/2)}{b/2} \cosh (E_j - \frac{b}{2})\right). \quad \text{(IV.2)} \]

As pointed out in Sec. I, $Q_{ij}$ itself does not actually depend on temperature; $Q_{ij}$ depends only on the energy difference $\varepsilon_j - \varepsilon_i$, viz. $Q_{ij} = [1 + \left(\frac{\varepsilon_j - \varepsilon_i}{2\hbar \epsilon}\right)^2]^4$. However, as we saw in [II], the energy difference of the important pairs of sites does vary with temperature: when the temperature drops below the range corresponding to nearest neighbor hopping, the energy difference for the important pairs decreases with decreasing $t$. As such, the $Q_{ij}$ factors associated with the resistors for such important pairs will show a temperature dependence, and the effect of
Fig. 3: The triangle and various boundary curves. The triangle is specified by the lines $E_j = \frac{1}{2t}$, $E_i = -\frac{1}{2t}$, $E_j = E_i$. The boundary curves, labelled according to the value of $\lambda$ appearing in the factor $Q_{ij}$ (equation (IV.1)), are symmetrical about the line $E_j = -E_i$ (dashed line). The quantity $E_j^{0}$, associated with all boundary curves, and the quantity $b_{\text{max}}$ corresponding to the boundary curve for $\lambda = 0$, are shown and are described in the text.
the $Q_{ij}$ factor on the overall value of $\xi^c$ will vary with temperature. This is reflected in the form chosen to write $Q_{ij}$ in equation (IV.1), and will be revealed further below by the manner in which $Q(b)$ affects the probability that sites with "energies" $E_i$ and $E_j$, separated by a distance $X$, will be connected. Combining (I.3), (II.1), (IV.1) and (IV.2) gives the criterion which $E_i$ and $E_j$ must satisfy if sites $i$ and $j$, separated by a distance $X$, are to be connected:

$$Q(b)f(b,E_j) < \chi^2 e^{\xi-X}.$$ (IV.3)

As such, sites $i$ and $j$ are connected if $E_i$ and $E_j$ are inside "the triangle" and also within the "boundary curve" specified by

$$Q(b)f(b,E_j) = \chi^2 e^{\xi-X},$$ (IV.4)

or

$$\cosh(E_j^b - \frac{b}{2}) = \frac{b/2}{\sinh(b/2)} \left( \frac{\chi^2 e^{\xi-X}}{2Q(b)} - \frac{\sinh b}{b} \right).$$ (IV.5)

We write $E_j^b$ in equation (IV.5) to denote that $E_j^b$, for given $X$ and $\xi$, is a function of $b$. For a flat density of states, the probability $p(X;\xi)$ is simply the fraction of the area of the triangle enclosed by the "boundary
The basic features of the boundary curve for the case $Q(b) \neq 1 (\lambda \neq 0)$, as specified by equation (IV.5), are the same as for the case $Q(b) = 1 (\lambda = 0)$, which were discussed thoroughly in [II], and are as follows. $E_j^b = E_j^0$ and $\frac{\partial E_j^b}{\partial E_i} = -1$ at $b = 0$, where

$$\cosh E_j^0 = \frac{X^2 e^{\xi-X}}{2} - 1. \tag{IV.6}$$

As $b$ increases, $E_j^b$ increases, reaches a maximum value, then decreases steadily with $b$; $b$ can increase to a maximum value $b_{\text{max}}$ where

$$E_j = -E_i = \frac{b_{\text{max}}}{2} \quad \text{and} \quad \frac{\partial E_j^b}{\partial E_i} = +1, \quad b_{\text{max}} \text{ being given by}$$

$$\frac{\sinh b_{\text{max}}}{b_{\text{max}}} + \frac{\sinh(b_{\text{max}}/2)}{b_{\text{max}}/2} = \frac{X^2 e^{\xi-X}}{2Q(b)}. \tag{IV.7}$$

As $b$ increases from 0 to $b_{\text{max}}$, $E_j^b - \frac{b}{2}$ decreases monotonically, meaning that the boundary curve monotonically approaches the line $E_j = -E_i$ (see Fig. 3). The boundary curve is symmetrical about the line $E_j = -E_i$; the value of $E_j^0$ exceeds $\frac{b_{\text{max}}}{2}$ for all values of $\lambda$. All of these features are displayed in Fig. 3.
Although $E_j^0$ is independent of $\lambda$, $b_{\text{max}}$ monotonically decreases with increasing $\lambda$. This is illustrated in Fig. 3 by showing boundary curves corresponding to given values of $\xi$ and $X$ and different values of $\lambda$. The principal characteristic of the curves is that the curves become more and more flattened as $\lambda$ increases, i.e. that $b_{\text{max}}$ decreases as $\lambda$ increases.

Our main goal in this section is to establish how the $Q_{ij}$ factor affects the shape of the $-\xi_c(t)$ vs. $t^{-1/4}$ curve: will inclusion of $Q_{ij}$ result in sufficient straightening of the curve in Fig. 1 to account for the experimental results of ref.s [5] and [6]? To answer this question, we employ the approximate version of our theory, which was described in the appendix of [II], to obtain results for $\xi_c(t)$ for various choices of $\lambda$. The approximation is to replace the boundary curve by its low $t$ asymptotic trapezoidal shape; the trapezoid is given by the line segments $E_j = b_{\text{max}}$ for $0 < E_i < b_{\text{max}}$, $E_j = E_i + b_{\text{max}}$ for $-b_{\text{max}} < E_i < 0$, $E_j = -b_{\text{max}}$ for $-b_{\text{max}} < E_i < 0$ and $E_j = E_i$ for $-b_{\text{max}} < E_i < b_{\text{max}}$. Because $b_{\text{max}}$ depends on $\lambda$, the approximate version of the theory must be slightly modified from the form given in [II] in order to obtain results for the case $\lambda \neq 0$. Notice from equation (IV.7) that $b_{\text{max}}$ decreases as $X$ increases. For a given $t$ and a given $\xi_c(t)$, the trapezoid "tops out", as $X$ decreases, when $b_{\text{max}} = \frac{1}{2t}$. This occurs for $X = X_{\text{top}}$; from equation (IV.7),

$$X_{\text{top}}^2 e^{\xi_c - X_{\text{top}}} = [1 + \frac{k^2}{4}]^4 [4t \sinh \frac{1}{2t} + 8t \sinh \frac{1}{4t}].$$

(IV.8)
Similarly, the trapezoid just fully encloses the triangle when $b_{\text{max}} = \frac{1}{2}$ occurring for $X = X_{\text{full}}$

$$X_{\text{full}} \Sigma e^{X_{\text{full}}} = [1 + \lambda^2] \left[ 2t \sinh \frac{1}{t} + 4t \sinh \frac{1}{2t} \right]. \quad \text{(IV.9)}$$

Note that, if $t$ is small enough, there will be no solution for $X_{\text{full}}$; if $t$ is smaller still, there will be no solution for $X_{\text{top}}$. This is because there are temperatures $t_{\text{top}}$ and $t_{\text{full}}$ such that, for $t < t_{\text{top}}$, the trapezoid for $X = 2$ is wholly inside the triangle, while for $t > t_{\text{full}}$, the trapezoid for $X = 2$ wholly encloses the triangle. These and other details are fully discussed in [II]. Using the theory presented in the appendix of [II], along with equations (IV.7), (IV.8), and (IV.9), we are able to determine $\xi_c(t)$ for various choices of $\lambda$. These various choices of $\lambda \equiv \frac{a\Delta\varepsilon}{2\hbar}$ correspond to various bandwidths $\Delta\varepsilon$. In Fig. 4, we plot the curves for $-\xi_c(t)$ vs. $t^{-1/4}$ for $\lambda = 0.75, 1.5, 2.25, 3$. Using $a = 100\hbar$ and $s = 5 \times 10^3 \text{ m}\cdot\text{s}^{-1}$, these correspond roughly to bandwidths $\Delta\varepsilon$ of 0.5, 1, 1.5 and 2 meV. Also shown in Fig. 4 is the curve obtained when the $Q_{ij}$ factor is replaced by unity.

Fig. 4 reveals that, as $\lambda$ increases, $-\xi_c(t)$ shifts downwards and, since $\xi_c(t)$ varies over a smaller range, the curves become flatter. An attempt to fit the curves to the experimental results of ref. [5] discloses that the curves are not linear over a wide enough $t$-range to explain the experimental
Fig. 4: Effect of the $Q_{ij}$ factor on the resulting curve for $-\xi_c$ vs. $t^{-1/4}$.

The curves are labelled according to the value for $\lambda$ appearing in equation (IV.1) for $Q_{ij}$. $\lambda = 0$ corresponds to replacing $Q_{ij}$ by unity. The other choices of $\lambda$ are related to particular choices for the bandwidth $\Delta t$, as shown in Table I. As was the case in Fig. 1, none of the curves is linear over a wide enough range of temperature to explain the experimental results of ref. [5].

Compare with Fig. 5.
results. We conclude that inclusion of the $Q_{ij}$ factor will not account for the small $T_0$ values observed experimentally.

In Fig. 5, we plot $-\xi_c(t)$ vs. $t^{-1}$ for the various values of $\lambda$. As we saw previously in Fig. 2, for $\lambda = 0$, $-\xi_c(t)$ varies linearly with $t^{-1}$, having a slope of about $-0.20$. This corresponds to a conductivity of the form $\sigma = -\frac{\varepsilon_3}{kT} e_0 e$ where the activation energy $\varepsilon_3 = 0.20\Delta\varepsilon$. As $\lambda$ increases, the magnitude of the slope, which is $\frac{\varepsilon_3}{\Delta\varepsilon}$, decreases. According to Table I, as the bandwidth $\Delta\varepsilon$ increases, the activation energy $\varepsilon_3$ becomes a smaller fraction of the bandwidth. Even so, $\varepsilon_3$ increases with increasing $\Delta\varepsilon$, but more slowly for larger bandwidths, suggesting perhaps a limiting value.

One's immediate inclination might be to think that the decrease in $\frac{\varepsilon_3}{\Delta\varepsilon}$ with $\lambda$ to be due to the linearity appearing in $T$-ranges where $t \equiv \frac{kT}{\Delta\varepsilon}$ was successively smaller. Fig. 5 indicates this to be not so. The decrease in $\frac{\varepsilon_3}{\Delta\varepsilon}$ with increasing $\Delta\varepsilon$ is a consequence of the $Q$ factor.

We must point out that the numbers presented in Table I are only approximate. This is partly because, as shown in Fig. 3, with increasing $\lambda$, the boundary curves become more flattened. The consequence of this is that the trapezoidal approximation described briefly above becomes less reliable as $\lambda$ increases. Moreover, the plots of $-\xi_c(t)$ vs. $t^{-1}$ appearing in Fig. 5 become less linear and more curved as $\lambda$ increases. The kink that appears in the curves corresponding to larger values of $\lambda$ is simply a manifestation of the kink used in the curve for $B_c(t)$ (Sec. III). This kink could be
Fig. 5: As in Fig. 4 but with $-\xi_c$ plotted against $t^{-1}$. The curves are linear for the temperature ranges given in Table I. The figure and table together reveal that, as $\Delta \varepsilon$ increases, the magnitude of the slope, $\frac{\varepsilon_3}{\Delta \varepsilon}$, decreases: the activation energy $\varepsilon_3$ becomes a smaller fraction of $\Delta \varepsilon$ as $\Delta \varepsilon$ increases. See text and Table I for more details.
TABLE I

| \( \lambda \) | \( \Delta \varepsilon \) (in meV) | \( |\text{slope}| (\frac{\varepsilon_3}{\Delta \varepsilon}) \) (in meV) | \( \varepsilon_3 \) (in meV) | T-range of linear regime (in K) |
|---|---|---|---|---|
| 0.75 | 0.5 | 0.19 | 0.10 | 0.3 to 1.2 |
| 1.5 | 1 | 0.17 | 0.17 | 0.6 to 1.7 |
| 2.25 | 1.5 | 0.15 | 0.22 | 1.1 to 2.3 |
| 3 | 2 | 0.12 | 0.24 | 1.8 to 3.9 |

TABLE I: Effect of the \( Q_{ij} \) factor on the slope of the curves in Fig. 5. \( \lambda \) is a parameter entering \( Q_{ij} \) which designates a choice of bandwidth \( \Delta \varepsilon \) of localized states. The tabulated magnitudes of the slopes, \( \frac{\varepsilon_3}{\Delta \varepsilon} \), associated values for the activation energies \( \varepsilon_3 \), and ranges of temperatures \( T \) for the linear regimes are approximate only and are estimated from the curves in Fig. 5.
smoothed away simply by using a smooth curve for $B_c(t)$. The trapezoidal approximation could also be improved upon. For example, one could use the trapezoid described by the four line segments: $E_j = E_j^0$ for $E_j^0 - b_{\text{max}} < E_j < E_j^0$, $E_j = E_j + b_{\text{max}}$ for $-E_j^0 < E_j < E_j^0 - b_{\text{max}}$, $E_j = -E_j^0$ for $-E_j^0 < E_j < -E_j^0 + b_{\text{max}}$, and $E_j = E_j$ for $-E_j^0 < E_j < E_j^0$. Alternatively, one could use the full form of the theory, as presented in [II], and not make any approximation to the boundary curve. We realize that such improvements are expected to produce slight changes in the values presented in Table I. However, we have decided not to undertake a more precise treatment of this problem. Based on the experience we have gained in our work for two dimensions, we believe that use of the full form of the theory will result in $\xi_c(t)$ curves which differ only slightly from those obtained above via the approximate version of the theory. We are certainly convinced that the exact curves for $\xi_c(t)$ will not account for the experimental results. For our purposes, the gains to be reaped from a more careful analysis are outweighed by the human time, computer time, and computer dollars which would need to be invested.

Although it is clear that a more thorough investigation into the $Q$-dependence of the slopes of $-\xi_c(t)$ vs. $t^{-1}$ would prove to be valuable, and would lead to more precise entries in Table I, we are at present content to have established what is the essential significance of the $Q_{ij}$ factor: the basic effects of the $Q_{ij}$ factor are to decrease $-\xi_c(t)$ with increasing bandwidth $\Delta\varepsilon$, to produce a trend of decreasing $\frac{\varepsilon_3}{\Delta\varepsilon}$ with increasing $\Delta\varepsilon$, and to
result in a rise in activation energy $\varepsilon_3$ with $\Delta \varepsilon$, possibly levelling off at large enough $\Delta \varepsilon$. Thus, we do not at present undertake a more precise treatment. Our next task will be to attempt to explain the small $T_0$ values observed experimentally by examining our theory for densities of states other than flat. We will report the results of such an investigation in our next paper.
V. Summary

One of the principal objectives of this paper was to determine whether or not the small $T_0$ values observed experimentally\(^5,6\) in three-dimensional LDSs could be accounted for by using the full resistor network (1.3) and a flat density of states. We began by extending our two-dimensional analytic theory to three dimensions. Although we were able to find, for a plot of $-\xi_c(t)$ vs. $t^{-1/4}$, a $t$-range wherein the curve was linear and with a small enough slope to match the experimental value of $T_0$, the $t$-range was not wide enough to fully account for the experimental results. We found that $-\xi_c(t)$ was better described as being linear in $t^{-1}$, with a slope of about $-0.20$, corresponding to a conductivity of the form $\sigma = \sigma_0 e^{-\varepsilon_3/kT}$ with activation energy $\varepsilon_3 = 0.20\Delta\varepsilon$. We found in [1] the same form for $\sigma$ for two dimensions, but with $\varepsilon_3 = 0.28\Delta\varepsilon$: the effect of going from two dimensions to three dimensions is to reduce the activation energy $\varepsilon_3$ relative to the bandwidth $\Delta\varepsilon$. We have also seen that, in three dimensions, the linear regime corresponds mostly to the "dropping" part of the $B_c(t)$ curve, whereas in two dimensions, it corresponded mostly to the "flat" part of the $B_c(t)$ curve. Our results differ from those of Skal, Shklovskii, and Efros\(^17,18\), who reported $\varepsilon_3 = \frac{5}{12} \Delta\varepsilon$ for both two and three dimensions, and also those of Hayden and Butcher\(^19\), who found $\varepsilon_3 = 0.38\Delta\varepsilon$ for two dimensions. We believe that our results differ from these other workers' because of their use of the low temperature form of the resistance in the high temperature regime.

Since the $-\xi_c(t)$ vs. $t^{-1/4}$ curves were not too far from being in full
agreement with the experimental results, we examined the effect of the usually neglected $Q_{ij}$ factor. Our main goal in this investigation was to establish whether or not inclusion of the $Q_{ij}$ factor would straighten out the $-\xi_c(t)$ vs. $t^{-1/4}$ curves enough to match the experimental curves. We have established this to be not so. As such, we have concluded that the experimental results may not be understood using the full resistor network along with a flat density of states. Beyond this result, we have also demonstrated what the principal effects of the $Q_{ij}$ factor are. The major role played by the $Q_{ij}$ factor is to produce a trend of decreasing $\frac{\xi_3}{\Delta\xi}$ with increasing $\Delta\xi$. The activation energy $\xi_3$ when $Q_{ij}$ is replaced by unity, as quoted above, is $\xi_3 = 0.20\Delta\xi$ for three dimensions. With $Q_{ij}$ taken into consideration, the activation energy can be a much smaller fraction of $\Delta\xi$ than 0.20 (see Table I). Although $\frac{\xi_3}{\Delta\xi}$ decreases with increasing $\Delta\xi$, $\xi_3$ itself rises with $\Delta\xi$, possibly levelling off at large enough $\Delta\xi$. Finally, the $Q_{ij}$ factor also results in a decrease in $-\xi_c$ with increasing $\Delta\xi$.

Although we have presented the nature of the essential effects of the $Q_{ij}$ factor, we realize that a more precise treatment can and should be made. We have not attempted a complete analysis of the role played by $Q_{ij}$. Our approach was based on the deformation potential approximation; a possibility for refinement of our results, beyond those suggestions given in Sec. IV, is to take into consideration the piezoelectric effect (see ref. [15] and other references therein).

In our next paper, we will investigate the effect which different
densities of states have on the shape of the $-\xi_c(t)$ vs. $t^{-1/4}$ curve. We will establish whether or not the full resistor network for any reasonable density of states can explain the low $T_0$ values and high $T$-range reported experimentally in ref.s [5] and [6].

ACKNOWLEDGMENTS

We wish to acknowledge useful discussions with Peter C.W. Holdsworth. This work was supported financially by NSERC Operating Grant A713.
REFERENCES

*Present Address: Department of Physics, Dalhousie University, Halifax, N.S., CANADA B3H 3J5

10. See, for example, J. Kurkijarvi, Phys. Rev. B 9, 770 (1974).
In two dimensions, the $\frac{1}{4}$ in equation (1.1) is replaced by $\frac{1}{3}$; for this reason, $B_c$ is plotted against $t^{-1/3}$ in two dimensions but against $t^{-1/4}$ in three dimensions.


Hopping conductivity in lightly doped semiconductors - IV:
High temperature $T^{-1/4}$ behavior with small $T_0$ values

Mark R.A. Shegelski* and Robert Barrie
Department of Physics, University of British Columbia
Vancouver, B.C., CANADA V6T 2A6

Abstract

In previous works, we performed a numerical simulation of, and presented an analytic theory for, a lightly doped semiconductor with a flat density of states. These were first steps toward the aim of the present work: to understand certain experimental results for lightly doped n-GaAs and n-InP in which the conductivity $\sigma$ was found to be $\sigma = \sigma_0 e^{-(T_0/T)^s}$ with $s$ close to $1/4$. The experimental range of temperatures $T$, $1K \lesssim T \lesssim 7K$, was well above, and the experimental $T_0$ values well below, the limits set by those established theories which predict $\sigma$ to be of this form. In this paper, we extend our analytic theory to study how the form of the density of states affects the resulting conductivity. As before, we model the semiconductor as a Miller and Abrahams type resistor network. We examine densities of states $g(\varepsilon)$ with dependence on energy $\varepsilon$ of the form $g(\varepsilon) \propto |\varepsilon|^{\lambda}$ where $\lambda = +\frac{1}{2}, 0, -\frac{1}{2}$, and the chemical potential is chosen to be zero. We find
that although the densities of states with \( \lambda = + \frac{1}{2}, 0 \) are not compatible with the experimental conductivity, it is highly plausible that a density of states which decreases with \(|\varepsilon|\) for \( \varepsilon \) far enough away from the chemical potential \( (\lambda = - \frac{1}{2}) \) accounts for the experimental conductivity. Crucial in our analysis is the use of the full form of the resistor network; we do not take the low temperature asymptotic form because we are interested in temperatures where \( kT \) is of order of the bandwidth \( \Delta \varepsilon \). We also predict that the \( T_0 \) in the conductivity \( \sigma = \sigma_0 e^{- (T_0/T)^{1/4}} \) will switch from a small value at temperatures where \( kT \) is of order \( \Delta \varepsilon \) to a large value when \( kT \ll \Delta \varepsilon \). We also suggest that whether \( \sigma \) goes like \( T^{-1} \) or \( T^{-1/4} \) when \( kT \) is of order \( \Delta \varepsilon \) may be due simply to the density of states, and that the electron-electron interaction appears to play no role in determining the form of \( \sigma \) when \( kT \) is of order \( \Delta \varepsilon \).

PACS numbers: 72.20.-i
I. Introduction

In a previous paper\textsuperscript{1}, hereafter referred to as [I], we reported that there exist experimental results\textsuperscript{2,3} for lightly doped n-GaAs and n-InP for which the conductivity $\sigma$ is given by

$$\sigma = \sigma_0 e^{-\left(\frac{T_0}{T}\right)^{1/4}},$$

(I.1)

and that these experimental results have not until now had a suitable theoretical explanation. The experimental conductivity (I.1) holds for a range of temperatures $T$ which is well above the temperature range where certain well known theories\textsuperscript{4,5,6} predict $\sigma$ to be of the form (I.1), and the experimental values for the characteristic temperature $T_0$ are much smaller than the theoretical values.

This paper represents the culmination of four investigations which have led to an understanding of these experimental results. The investigations have been based on a resistor network model of a lightly doped semiconductor, which was presented in detail in [I]. In [I], we performed a numerical simulation in two dimensions to calculate the temperature dependence of the overall resistance of the resistor network. In a subsequent paper\textsuperscript{7}, hereafter referred to as [II], we used an analytic approach to calculate the overall resistance of the network. We formulated the calculation in terms of a percolation problem. The analytic theory involved the use of the temperature dependent critical number of resistors, or bonds, per site, $B_c(T)$, required for percolation at temperature $T$. By
choosing a reasonable form for $B_c(T)$, we were able to obtain agreement between the analytic theory and the results of the numerical simulation. In a third paper, [III], we extended our analytic theory to three dimensions. In all three papers, we chose the density of localized donor states to be flat. We established in [III] that the experimental conductivity reported in ref.s [2] and [3] could not be accounted for by using a flat density of states.

In the present paper, we consider densities of states other than flat. The densities of states studied, and the reasons for choosing them, are described in Sec. II. We very briefly outline in Sec. III the key equations involved in our theory. We also describe in detail how our theory — delivered in full in [II] and [III] — must be modified when a flat density of states is replaced by an energy-dependent density of states. Presented in Sec. IV are the results of our investigations for the various densities of states examined. In Sec. V, we discuss how these results reveal an understanding of the experimental conductivity. We also relate the implications of our explanation of the experimental results. A summary and recommendation for further study are offered in Sec. VI.

We assume that the reader is familiar with the work we reported in [I], [II], and [III]. If he or she is not, the essential ideas in, and findings of, our work may be absorbed by skimming Sec.s II through IV, and reading Sec. V and Sec. VI.
II. The Densities of States

In this paper, we consider the same basic problem as that of [II] and [III], and we employ all the same notation. In particular, we continue to regard the density of states $g(\varepsilon)$ to be symmetrical about the zero temperature chemical potential $\mu_0$, we set $\mu_0 = 0$, and for simplicity we take the number density $n_D$ of donors to be twice the number density $n_A$ of acceptors. As such, the resistances $R_{ij}$ are of the same form as in our previous works. However, because we found in [III] that the $Q_{ij}$ factor in $R_{ij}$ did not affect the value of $s$ corresponding to $-\xi_c$ being linear in $t^{-s}$, we will replace $Q_{ij}$ by unity in this paper. As in our previous works, we use the full form of the resistor network, and do not take the low $t$ asymptotic form, because we are interested in temperatures where $kT$ is of the order of the bandwidth $\Delta \varepsilon$ of the density of states. The only new element to be introduced in this paper is the form of the density of states.

We want to examine how different densities of states affect the result obtained for $-\xi_c(t)$ (the critical resistor $R_c$ is given by $R_c = R_0 e^{-\xi_c}$). We are interested in densities of states which differ fundamentally from one another. As such, we consider densities of states of the form

$$g(\varepsilon) = \begin{cases} 
0, & \text{for } |\varepsilon| > \frac{\Delta \varepsilon}{2}, \\
\alpha |\varepsilon|^\lambda, & \text{for } 0 < |\varepsilon| < \frac{\Delta \varepsilon}{2}, \lambda > -1,
\end{cases}$$  \hspace{1cm} (II.1)
where $\alpha$ is fixed by $\int_{-\Delta \varepsilon / 2}^{\Delta \varepsilon / 2} g(\varepsilon) \, d\varepsilon = n_D$. The case $\lambda = 0$ corresponds to a flat density of states; $\lambda = 2$ corresponds to a case like the one envisioned by Shklovskii and Efros where the density of states vanishes at the Fermi level in consequence of the electron-electron interaction. The gap in the density of states is referred to as the Coulomb gap. We will examine the cases $\lambda = + \frac{1}{2}$ and $\lambda = - \frac{1}{2}$. The densities of states for cases $\lambda = + \frac{1}{2}$, $\lambda = 0$, $\lambda = - \frac{1}{2}$ are shown by solid curves in Fig. 1. The first corresponds to a narrower Coulomb gap than the $\lambda = 2$ case; the last models a density of states which decreases with $|\varepsilon|$ as $\varepsilon$ moves away from $\mu_0$, and is characteristic of a case where there is no Coulomb gap.

For the choice $\lambda = - \frac{1}{2}$, it would be more appropriate to take the density of states to be

$$g(\varepsilon) = \begin{cases} g_F, & \text{for } 0 < |\varepsilon| < \frac{\eta}{2}, \\ \alpha |\varepsilon|^{\lambda}, & \text{for } \frac{\eta}{2} < |\varepsilon| < \frac{\Delta \varepsilon}{2}, \\ 0, & \text{for } |\varepsilon| > \frac{\Delta \varepsilon}{2}, \end{cases}$$

(II.2)

where $g_F = \alpha \left|\frac{\eta}{2}\right|^{\lambda}$ (see the dashed curve in Fig. 1). One might also wish to employ equation (II.2) for the case $\lambda = + \frac{1}{2}$. Our attitude in the remainder of this paper will be as follows. We consider the density of states to be of the form (II.2). As such, at low $T (kT \ll \eta)$, we take $g(\varepsilon) = g_F$, since
Fig. 1: Various densities of states. Densities of states of the form (II.1) with $\lambda = +\frac{1}{2}$, $\lambda = 0$ and $\lambda = -\frac{1}{2}$ are indicated by solid curves; the density of states (II.2) for $\lambda = -\frac{1}{2}$ is indicated by a dashed curve. Notice that all the densities of states are symmetric about the zero temperature chemical potential $\mu_0$. The area under each of the curves is equal to the number density $n_D$ of donors. The bandwidth for each case is $\Delta \varepsilon$. 

$\begin{align*}
2\frac{n_D}{\Delta \varepsilon} g(\varepsilon) \\
\frac{n_D}{\Delta \varepsilon} \\
0 \quad \mu_0 - \frac{\Delta \varepsilon}{2} \quad \mu_0 - \frac{\eta}{2} \quad \mu_0 \quad \mu_0 + \frac{\eta}{2} \quad \mu_0 + \frac{\Delta \varepsilon}{2} \quad \varepsilon
\end{align*}$
only those states with small $|\epsilon|$ matter, and $\sigma$ will be of the form (I.1) with $T_0$ large. (This will account for the low temperature conductivity observed by Emel'yanenko et al.\textsuperscript{10}, which was more thoroughly discussed in [1]. We return to this point in Sec. V.) At high $T$ ($t \equiv kT/\Delta\varepsilon$ of order unity), provided $\frac{n}{\Delta\varepsilon}$ is not too close to unity, we replace the density of states (II.2) by the form (II.1). We can do so because the states near the band edges will be as important at high $T$ as those near the middle of the band. We make such an approximation at high $T$ for two reasons. Our main objective is to establish the manner in which fundamentally different types of densities of states — such as the cases $\lambda = \pm \frac{1}{2}$ — will affect the conductivity. In particular, we will determine whether or not, for a reasonable density of states, the conductivity predicted by the resistor network model (see [1]) can be of the form (I.1) at high temperatures ($t$ of order unity) and with $T_0$ much smaller than the value for $T_0$ associated with low temperatures ($t \ll 1$) and a flat density of states $4,5,6$. An answer in the affirmative will constitute at last a possible explanation for the experimental conductivity reported by Benzaquen and Walsh\textsuperscript{2} for lightly doped n-GaAs and by Benzaquen et al.\textsuperscript{3} for lightly doped n-InP. However, it is not our objective at this time to determine the exact density of states required to account for these experimental results. Moreover, replacing the density of states (II.2) by the reduced form (II.1) greatly simplifies the theoretical analysis to be presented in the next section. The form (II.1) for the density of states will be suitable for our purposes.
III. Theory

The overall conductivity of the resistor network (see [I], [II], [III]) is given by

\[ \sigma = \sigma_0 e^{-\xi_c}, \quad (III.1) \]

where \( \xi_c \) is given in terms of the critical resistance \( R_c \) by \( R_c = R_0 e^{-\xi_c} \) (\( R_0 \) is a prefactor common to all resistors in the network), and is fixed by the following equations:

\[ B_c(t) = \frac{\pi}{2} \rho_s a^3 \int_{-\infty}^{\infty} X^2 p(X;\xi_c) dX, \quad (III.2) \]

\[ \frac{e^{\xi_c} - X_{\max}}{X_{\max}^2} = 4, \quad (III.3) \]

where \( B_c(t) \) is the critical number of bonds (resistors) per site required for percolation at temperature \( t \), and \( p(X;\xi_c) \) is the probability that two sites a distance \( X \) apart will have energies such that the sites will be connected.

In this section, we obtain an expression for \( p(X;\xi_c) \) when the density of states is of the form (II.1), and we give explicit expressions for \( \lambda = \pm \frac{1}{2} \) (the case \( \lambda = 0 \) was dealt with thoroughly in [II] and [III]).
symbols used are as defined in [II] and [III].

For the density of states (II.1), the probability that a site chosen at random has energy $\varepsilon_i$ between $\varepsilon$ and $\varepsilon + d\varepsilon$ is

$$\frac{2^\lambda (\lambda + 1) |\varepsilon|^\lambda d\varepsilon}{(\Delta \varepsilon)^{\lambda + 1}}.$$

The probability that, in selecting a pair of sites at random, one has energy between $\varepsilon_i$ and $\varepsilon_i + d\varepsilon_i$, and the other has energy between $\varepsilon_j$ and $\varepsilon_j + d\varepsilon_j$, is

$$\frac{2^{2\lambda + 1} (\lambda + 1)^2 |\varepsilon_i|^\lambda |\varepsilon_j|^\lambda d\varepsilon_i d\varepsilon_j}{(\Delta \varepsilon)^{2\lambda + 2}}.$$

In terms of $E_i = \frac{\varepsilon_i}{kT}$, $E_j = \frac{\varepsilon_j}{kT}$, and $t$, this last quantity above, denoted by $p(E_i, E_j)$, may be written as

$$p(E_i, E_j) = \frac{1}{2} (\lambda + 1)^2 (2t)^{2\lambda + 2} |E_i|^\lambda |E_j|^\lambda dE_i dE_j. \quad (III.4)$$

$p(X; \xi_c)$ is determined by integrating $p(E_i, E_j)$ over the intersection of the triangle and the region interior to the boundary curve (see [II]). In the context of the trapezoidal approximation (described in the appendix of [II]), there are two cases to consider: $0 < b_{\max} < \frac{1}{2t}$ and $\frac{1}{2t} < b_{\max} < \frac{1}{t}$. It is straightforward to obtain the expressions for $p(X; \xi_c)$ for these cases.
p(X; \xi_c) turns out to depend only on \(2t_{b_{\text{max}}} = u\) (where, in turn, \(b_{\text{max}}\) depends on \(X\) and \(\xi_c\) as given by equation (IV.4) of [II]), so we write \(p(X; \xi_c) = p(u)\). The expression for \(p(u)\) for \(\lambda > -1\) is:

\[
p(u) = \frac{1}{2}(1 + (\lambda+1) \frac{\Gamma(\lambda+1)\Gamma(\lambda+2)}{\Gamma(2\lambda+3)}) u^{2\lambda+2}, \text{ for } 0 < u < 1,
\]

\[
\frac{1}{2}(1 + (u-1)^{\lambda+1}) + \frac{1}{2}(\lambda+1) u^{2\lambda+2} \int_{1-u^{-1}}^{1} d\zeta \zeta^{\lambda}(1-\zeta)^{\lambda+1}, \text{ for } 1 < u < 2, \quad (\text{III.5})
\]

1, for \(u > 2\),

when \(\Gamma(x)\) is the Gamma function. For the particular cases \(\lambda = \pm \frac{1}{2}\), we have:

\[
\left(\frac{1}{2} + \frac{3\pi}{64}\right)u^3, \quad \text{for } 0 < u < 1,
\]

\[
p_{+1/2}(u) = \frac{1}{2}(1+(u-1)^{3/2}) + \frac{3}{16}u^3(\frac{\pi}{4} - \arccos u^{-1/2}) \quad (\text{III.6})
\]

\[
+ \frac{3}{16} u(2-u)(u-1)^{1/2}, \text{ for } 1 < u < 2,
\]

1, for \(u > 2\),
and

\[ p_{1/2}(u) = \left( \frac{1}{2} + \frac{\pi}{8} \right) u, \text{ for } 0 < u < 1, \]

\[ p_{1/2}(u) = \frac{1}{2}(1 + (u-1)^{1/2}) + \frac{1}{2}u\left(\frac{\pi}{4} - \arccos u^{-1/2}\right), \text{ for } 1 < u < 2, \quad (III.7) \]

1, for \( u > 2 \).

The reason for using the approximate version of our theory instead of the full form will soon become obvious.

Having found expressions for \( p(X;\xi_c) \), the only ingredient in our theory remaining to be specified is the curve for \( B_c(t) \). In [II], we obtained the \( B_c(t) \) curve for two dimensions and a flat density of states by matching our theoretical result for \( \xi_c(t) \) with the numerical result obtained in [I]. In [III], we used the \( B_c(t) \) curve for three dimensions and a flat density of states. Instrumental in our analyses were considerations of the \( t \to \infty \) limit and the \( t \to 0 \) asymptotic form of \( B_c(t) \). For the three-dimensional case where the density of states is not flat, the \( t \to \infty \) limit of \( B_c(t) \) is still 2.77 (see [III]) but the \( t \to 0 \) asymptotic form is different. In an appendix, we derive the low \( t \) asymptotic form for \( B_c(t) \) for the density of states (II.2) with \( \lambda = -\frac{1}{2} \). The reasons for examining this case will become clear in the next section. The expression we obtain is
\[ B_c(t) = \left( \frac{\Delta \epsilon}{\eta} \right)^{3/4} B_c^0(t), \]  

(III.8)

where \( f \), the fraction of states for which \(|\epsilon| < \frac{\eta}{2}\), is

\[ f = \frac{1}{\sqrt{\Lambda \epsilon^2/\eta} - 1} \]  

(III.9)

and \( B_c^0(t) \) is the expression for the low \( t \) form of \( B_c(t) \) for \( \lambda = 0 \):

\[ B_c^0(t) = \frac{\eta}{10} n^c (t_c/t_c)^{3/4}, \]  

(III.10)

where \( n_c \) was taken in [I] to be \( 5.3 \pm 0.3 \) (ref. [11]) and \( t_c \) was shown in [III] to be

\[ t_c = \frac{\rho_s^{1/3} a}{4 n_c^{1/3}}. \]  

(III.11)

It is important to note that, whereas the \( B_c(t) \) curve for the case \( \lambda = 0 \) is expected to merge onto the low asymptotic form \( B_c^0(t) \) when \( t \lesssim t_c \), for \( \lambda = -\frac{1}{2} \) the merging can be expected only for \( t \lesssim f^{1/3} \frac{\eta}{\Delta \epsilon} t_c \) (see equation (A.5) of the appendix). Combined with equation (III.8), this implies that the merging moves to both smaller \( t \) and to a smaller value of \( B_c \).

\( (t < f^{1/3} \frac{\eta}{\Delta \epsilon} t_c \) implies \( B_c < f \frac{\eta}{10} n_c \sim 1.66f) \)
The consequence of this shift in the merging is as follows. It becomes less clear than it was in [III] how to interpolate for $B_c(t)$ between the high $t$ limit and the low $t$ asymptotic form. In [III], we could obtain the appropriate form for $B_c(t)$ for three dimensions and a flat density of states by appealing to the form for $B_c(t)$ obtained in [II] for two dimensions and a flat density of states. We have less to guide us in choosing $B_c(t)$ for three dimensions and a density of states other than flat. As such, it is clear that it would not be suitable to use the full form of our theory; the approximate version will suffice.

Despite not being able to specify $B_c(t)$ exactly, the circumstances for continued investigation remain favorable, and we adopt the following strategy. We will begin by assuming that $B_c(t)$ does not change appreciably with a change in the density of states, and we will use for our initial investigation the form for $B_c(t)$ presented in [III]. We will inquire as to how close the curves for $\bar{\xi}_c(t)$ vs. $t^{-1/4}$ come to accounting for the experimental conductivities reported in ref.s [2] and [3]. Based on the results obtained, we will ask: for a reasonable density of states, is there a reasonable form for $B_c(t)$ that arranges that the $\bar{\xi}_c(t)$ vs. $t^{-1/4}$ curve does correspond to the experimental data for $\ln\sigma$ vs. $T^{-1/4}$? The answer is given in the next section.
IV. Results

Combining the expressions for $p(X; \xi_c)$ of Sec. III with the approximate form of our theory (the trapezoidal approximation, described in [II] and [III]), we have obtained $\xi_c(t)$ for the density of states (II.1) for $\lambda = \pm \frac{1}{2}$. For the "flat density of states" $B_c(t)$ used in [III], we find that $\xi_c$ is linear in $t^{-1}$ for both values of $\lambda$, as was found in [III] for $\lambda = 0$. In Fig. 2, we plot $-\xi_c$ vs. $t^{-1}$ for the three cases $\lambda = -\frac{1}{2}$, $\lambda = 0$, $\lambda = +\frac{1}{2}$. We see from the figure that the magnitude of the slope of the curve increases as $\lambda$ increases. (For discussions of the linearity of $\xi_c$ in $t^{-1}$, see [II] and [III].)

In Fig. 3, we plot $-\xi_c$ vs. $t^{-1/4}$ for the three values of $\lambda$, with this same $B_c(t)$. Also shown in the figure is a heavy curve, hereafter referred to as "the experimental curve", which is consistent with the experimental conductivity for sample 2 ($\rho_s a^3 = 0.0087$) of ref. [2]. (The reasons for choosing this sample were given in [III].) This curve has the three features required for consistency with experiment (see the discussion in Sec. III of [III]): (1) $-\xi_c$ is linear in $t^{-1/4}$, (2) the range of the linearity, $t_{\text{min}} < t < t_{\text{max}}$, is extensive enough to match the experimental range, viz. $t_{\text{min}} \approx 0.036$, $t_{\text{max}} \approx 0.15$, thereby corresponding to the experimental values $T_{\text{min}} = 1.4K$, $T_{\text{max}} = 6K$, and (3) the magnitude $m$ of the slope is small enough to match the small experimental $T_0$ value (915 K), viz. $m t_{\text{min}}^{-1/4} = (\frac{T_0}{T_{\text{min}}})^{-1/4} \approx 5.1$, $m t_{\text{max}}^{-1/4} = (\frac{T_0}{T_{\text{max}}})^{-1/4} \approx 3.5$ ($m \approx 2.2$). The
Fig. 2: Plots of $-\xi_c$ vs. $t^{-1}$ for the cases $\lambda = \pm \frac{1}{2}$, $\lambda = 0$, $\lambda = -\frac{1}{2}$ with $B_c$ given by the "flat density of states" $B_c(t)$, curve "1" of Fig. 4.
Fig. 3: Plots of $-\xi_c$ vs. $t^{-1/4}$ for the cases $\lambda = +\frac{1}{2}$, $\lambda = 0$, $\lambda = -\frac{1}{2}$ with $B_c$ given by curve "1" of Fig. 4; these plots are given by the curves of normal thickness. The heavy curve corresponds to $\lambda = -\frac{1}{2}$ and $B_c$ as given by curve "2" of Fig. 4. See text for details and discussion.
The experimental curve has been positioned near to the curve for \( -\xi_c(t) \) corresponding to \( \lambda = -\frac{1}{2} \) for an obvious reason: the curve for \( \lambda = -\frac{1}{2} \) is by far the closest of the three to matching the experimental curve. Moreover, notice that the \( \lambda = -\frac{1}{2} \) curve is actually quite close to matching the experimental curve. As such, we ask: what must \( B_c(t) \) be for the case \( \lambda = -\frac{1}{2} \) in order that \( -\xi_c(t) \) coincide with the experimental curve, and is this form of \( B_c(t) \) reasonable?

The curve for \( B_c(t) \) required to give such agreement, hereafter referred to as "the new \( B_c(t) \) curve", is labelled "2" in Fig. 4. The flat density of states \( B_c(t) \) curve, which has been used until now and is hereafter referred to as "the old \( B_c(t) \) curve", is labelled "1" in Fig. 4. The two curves coincide for \( t^{-1/4} < 1.6 \). The new \( B_c(t) \) curve agrees with the \( t \to \infty \) limit, and is also consistent with the low \( t \) asymptotic form described in Sec. III.

To establish the latter claim more clearly, we recall that, for the case under consideration (\( \rho_s a^3 = 0.0087 \)), \( t_c^{-1/4} = 2.4 \). The low \( t \) asymptotic form (III.8) for \( B_c(t) \) corresponding to the density of states (II.2) with \( \lambda = -\frac{1}{2} \), holds, as we saw in Sec. III, for \( t \ll f^{1/3} \frac{\eta}{\Delta \varepsilon} t_c \), thereby implying that \( B_c \ll 1.66f \). The point \( (t^{-1/4}, B_c) = (f^{-1/12} (\frac{\Delta \varepsilon}{\eta})^{1/4} t_c^{-1/4}, 1.66f) \) lies "to the right and down" relative to the rightmost and lowest point of curve "2" in Fig. 4 (the new \( B_c(t) \) curve). For the sake of illustration, the point \( (t^{-1/4}, B_c) \) is (3.0, 0.91) for \( \Delta \varepsilon = 2\eta \), (3.7, 0.55) for \( \Delta \varepsilon = 4\eta \), (7.2, 0.15) for \( \Delta \varepsilon = 36\eta \). As such, we see that the new \( B_c(t) \) curve is a
Fig. 4: $B_c$ plotted against $t^{-1/4}$. The curve labelled "1" corresponds to the form of $B_c$ used in [III]. The curve labelled "2" is the form of $B_c$ required to give $\xi_c$ linear in $t^{-1/4}$ over a wide enough range of $t$ and with a small enough slope to account for the experimental conductivity of ref. [2]; this curve for $\xi_c$ is given in Fig. 3. See text for details and discussion.
reasonable curve. We turn next to a discussion of the implications of our results.
V. Discussion

Although the $B_c(t)$ curve of Fig. 4 which gives agreement with experiment is a reasonable curve, there is no guarantee that it is the true $B_c(t)$ curve. Given a particular form for the density of states, $B_c(t)$ is fixed; the $B_c(t)$ curve for the density of states (II.2) with $\lambda = -\frac{1}{2}$ may or may not coincide closely with the $B_c(t)$ curve of Fig. 4. One of the key points of our analysis is simply that it is quite conceivable that the two curves do coincide closely.

Our contention is not necessarily that the density of states for sample 2 of ref. [2] is given by (II.2) with $\lambda = -\frac{1}{2}$; as we stated in Sec. II, it is not our objective at present to determine the exact density of states required to explain the experimental data. Rather, we have demonstrated that it is highly plausible that the high temperature experimental conductivity (1.1), with its associated small $T_0$ value, may be accounted for, in the context of the resistor network model, by a density of states which has the character of (II.2) with $\lambda = -\frac{1}{2}$. The character referred to is that the density of states decreases with $|\varepsilon|$ for $\varepsilon$ far enough away from $\mu_0$. Even if the true $B_c(t)$ curve for $\lambda = -\frac{1}{2}$ is not close to the curve of Fig. 4, it is very plausible that, for some particular density of states which decreases with $|\varepsilon|$ for $\varepsilon$ far enough away from $\mu_0$, the true $B_c(t)$ curve for that density of states is such that $-\xi_c$ varies as $t^{-1/4}$ with a small slope at high $t$.

For such a density of states, since at low enough temperature the only states that will matter will be those with $\varepsilon$ close to $\mu_0$, the effective
density of states will be flat. (We have built this into the density of states (II.2).) As such, as the temperature is lowered, we expect to see the conductivity go from the form (I.1) with $T_0$ small to the form (I.1) with $T_0$ large. As we argued in [I], this is precisely the behavior that was observed by Emel'yaneenko et al.\textsuperscript{10} in lightly doped n-GaAs. Our interpretation of their data for the temperature range 1.2K to 4.2K agrees with the conductivity reported by Benzaquen and Walsh\textsuperscript{2} for lightly doped n-GaAs in the range 1.4K to 5-7K. Moreover, we predict that, just as Emel'yaneenko et al. found (I.1) to hold for 0.15K to 1K with $T_0$ having a value given by established theories of variable range hopping\textsuperscript{4,5,6}, had the Benzaquen and Walsh data been extended to lower temperatures, the conductivity would have switched to the form (I.1) with a significantly larger value of $T_0$. (For further details, refer to the discussion in [I].)

What we are suggesting is the existence of two $T^{-1/4}$ regimes: one at low temperature, with large $T_0$, which is well understood on the basis of established theories\textsuperscript{4,5,6} of variable range hopping, and one at high temperature, with small $T_0$, which may now be understood on the basis of the ideas presented in this paper.

A puzzle related to the experimental conductivities reported in the literature for lightly doped n-GaAs or n-InP is the following. In some experimental studies (see, for example, ref.s [12] to [15], it was reported that $\ln\sigma$ varied as $T^{-1}$, whereas in others\textsuperscript{2,3} it was reported that $\ln\sigma$ varied as $T^{-1/4}$; in all cases, $T$ was such that $1K \lesssim T \lesssim 10K$. In view of the
results presented in Sec. IV, we are able to offer a suggestion as to why this is so. Inasmuch as a change in the $B_c(t)$ curve used in our theory can lead to a change in the value of $s$ for which $-\xi_c$ is linear in $t^{-s}$ — curve "1" in Fig. 4 gave $s = 1$ for $\lambda = + \frac{1}{2}$, 0, $- \frac{1}{2}$, while curve "2" gave $s = \frac{1}{4}$ for $\lambda = - \frac{1}{2}$ —, different densities of states, with their concomitant different $B_c(t)$ curves, will have different values of $s$ associated with them.

Moreover, just as the two $B_c(t)$ curves in Fig. 4 do not differ nearly as drastically as they might have, it is quite feasible that the density of states corresponding to $s = 1$ is not wildly different from the density of states corresponding to $s = \frac{1}{4}$.

Finally, we point out that, nowhere in our work were we required to invoke a density of states having a dip at the chemical potential in order to account for the experimental conductivity of ref.s [2] and [3]. Indeed, the density of states for $\lambda = + \frac{1}{2}$ in Fig. 1 is unquestionably without a hope of being able to describe the experimental conductivity. It has been suggested that such a dip in the single-particle density of states would arise from electron-electron interaction. Our model of hopping conductivity ignores such interaction.
VI. Summary and Recommendations for Further Study

We have demonstrated that it is highly plausible that the conductivity (I.1), associated with small $T_0$ values ($T_0 \approx 10^3 K$), and observed at high temperatures ($1 K \lesssim T \lesssim 5-7 K$) in lightly doped semiconductors (ref.s [2], [3], and possibly also [10] — see [1]), may be accounted for by using the resistor network model along with a density of states which decreases with $|\varepsilon|$ for $\varepsilon$ far enough away from $\mu_0$. We have stated that at lower temperatures, the conductivity will be of the form (I.1) but with a much larger value of $T_0$ ($T_0 \approx 10^5 K$). We have also suggested that whether the conductivity goes, at high temperatures, as $T^{-1}$ or as $T^{-1/4}$ — or, indeed, as any power in between — may simply be due to the density of states, and we have observed that our results indicate the lack of a need to take into consideration the electron-electron interaction to understand the high temperature hopping conductivity.

We recognize that we have not established unequivocally that the high temperature conductivity (I.1) with small $T_0$ values may be accounted for via the resistor network model. However, we expect that further study in this area will remove any doubt. Our recommendations for further study are as follows. The program we have carried out can be repeated for various densities of states, and the actual form of $B_c(t)$ can be determined by computer simulation. To obtain $B_c(t)$ for a given density of states, one constructs a resistor network as outlined in [I], connects up resistors until percolation occurs, and counts the number of resistors present at
percolation. As such, $B_c(t)$ will be known, whence one can examine how the various densities of states lead to different types of behaviour for the conductivity at high temperatures. The various densities of states may be taken to be of the form (II.2). If so, we suggest that the full form (II.2) be used and not the simplification (II.1). This means the theory will have two parameters, $\lambda$ and $\frac{h}{\Delta e}$, and that one can investigate the transition from $\sigma$ of form (I.1) with small $T_0$ at high temperature to form (I.1) with larger $T_0$ at small temperature. Further improvement can be obtained by using the full form of the theory instead of the approximate version.

It is our belief that such a study will establish without doubt what we have advanced as a strong plausibility argument in this paper. Such a study is being undertaken.
APPENDIX

In this appendix, we derive the low $t$ asymptotic form for $B_c(t)$ corresponding to the density of states (II.2) with $\lambda > -1$. We begin by noting that the low $t$ form of $\xi_c$,

$$\xi_c = \left( \frac{4n_c}{g_F a^3 \eta} \right)^{1/4},$$

may be expressed as

$$\xi_c = \left( \frac{4n_c}{f \rho_s a^3 \eta} \right)^{1/4},$$  \hspace{1cm} (A.1)

where $f$ is the fraction of states for which $|\varepsilon| < \frac{n}{2}$ and $t_\eta = \frac{kT_c}{\eta}$. The equation $kT_c \xi_c(T_c) = \frac{n}{2}$ defines $T_c$; we find

$$t_{\eta, c} = \frac{(f \rho_s)^{1/3} a}{4n_c^{1/3}},$$ \hspace{1cm} (A.2)

where $t_{\eta, c} = \frac{kT_c}{\eta}$.

The probability that a pair of sites selected at random will have one energy between $\varepsilon_i$ and $\varepsilon_i + \delta \varepsilon_i$ and the other between $\varepsilon_j$ and $\varepsilon_j + \delta \varepsilon_j$, where
\[ 0 < |\varepsilon_i| < \frac{\eta}{2} \text{ and } 0 < |\varepsilon_j| < \frac{\eta}{2}, \text{ is } 2f^2 \int dE_i dE_j. \]

In the trapezoidal approximation, \( b_{\text{max}} = \xi_c - X \) at low \( t \), and the area of the trapezoid is \( \frac{3}{2}(\xi_c - X)^2 \). This gives

\[ p(X; \xi_c) = 3f^2 \eta^2 (\xi_c - X)^2 \quad (A.3) \]

at low \( t \) (i.e. \( \xi_c < \frac{1}{2t\eta} \) or \( t_\eta < t_{\eta,c} \)).

Combining equations (III.2) and (A.3) leads to

\[ B_c = \frac{\pi}{10} n_c \left( \frac{t_{\eta}}{t_{\eta,c}} \right)^{3/4} \quad (A.4) \]

in the limit \( t \to 0 \). We may rewrite (A.4) in the form (III.8). At low enough \( t \), the \( B_c(t) \) curve will merge onto the asymptotic form (A.4). The temperature at which this occurs will be of the order of the temperature for which \( t_{\eta} \approx t_{\eta,c} = f^{1/3} t_c \), or

\[ t = f^{1/3} \frac{\eta}{\Delta \varepsilon} t_c. \quad (A.5) \]

\( f \) is given by

\[ f = \frac{\lambda + 1}{\left( \frac{\Delta \varepsilon}{\eta} \right)^{\lambda+1} + \lambda}. \quad (A.6) \]
All of the above is valid for \( \lambda > -1 \). For the case \( \lambda = -\frac{1}{2} \), we have the particular case (III.9) for \( f \).

ACKNOWLEDGMENTS

We are very grateful to Peter C.W. Holdsworth for continued useful discussions. This work was supported financially by NSERC Operating Grant A713.
REFERENCES

*Present Address: Department of Physics, Dalhousie University, Halifax, N.S., CANADA B3H 3J5

1 M.R.A. Shegelski and R. Barrie, submitted to Phys. Rev. B.


7 M.R.A. Shegelski and R. Barrie, submitted to Phys. Rev. B.

8 M.R.A. Shegelski and R. Barrie, submitted to Phys. Rev. B.


CONCLUSION

I. Summary and Highlights

One of the highlights of this dissertation was to establish that certain experimental results, originally thought to be well understood, did not actually have a satisfactory explanation prior to the work of this dissertation. The experimental conductivity $\sigma$ in lightly doped $n$-GaAs was found by Benzaquen and Walsh $^1$ to be

$$\sigma = \sigma_0 e^{-\left(\frac{T_0}{T}\right)^{\frac{1}{4}}}, \quad (I.1)$$

with $s$ close to $\frac{1}{4}$. The same behavior was subsequently reported in lightly doped $n$-InP (ref. [2]). This experimental behavior was initially $^1$ regarded to be evidence of variable range hopping (VRH) as described by established theories $^3,4,5$ of hopping conduction. However, it was revealed in [1] that the experimental range of temperatures $T$ were well above, and the experimental $T_0$ values well below, the limits set by these theories. The principal aim of this dissertation was to explain these results.

In view of the success of the Miller and Abrahams resistor network model $^6$ of hopping conduction in doped and amorphous semiconductors, a lightly doped semiconductor was modelled in this dissertation by this type of resistor network. The specific form of the resistance was given by
equations (4) - (7) of [I]; a derivation of these equations appears in the Appendix. The major assumptions underlying the resistor network model were as follows. The electron-electron Coulomb interaction was treated in only a very simple manner, having been taken into account only through the form of the density of one-electron localized states. Electron-electron correlations were assumed to play no part in the hopping process, so the resistances depended only on quantities pertaining to the two sites at the ends of the resistor, and a rate equation approach was used to describe the site to site hopping. Moreover, the positions of donors and acceptors, and the energy levels associated with the donor sites, were assumed to be uncorrelated. Finally, the electron-phonon interaction was modelled by the deformation potential approximation.

The temperature range in the investigations undertaken in [I] - [IV] was always such that kT was of order of the bandwidth Δε of localized states. For such a temperature range, it is essential to use the full resistor network if meaningful results are to be obtained. This dissertation represents the first work in which the full form of the resistor network was used for the purpose of determining the temperature dependence of σ when kT was of order Δε.

A two-dimensional numerical simulation designed to calculate σ was performed in [I]. The density of states was taken as flat and the number density \( n_d \) of donors was chosen for simplicity to be twice the number density \( n_A \) of acceptors. The results of the simulation could have been
viewed as being in accord with the experimental results of refs. [1] and [2] had the following requirements been met: (i) there would have been a range of temperatures at high t (t ≡ kT/Δε of order unity) wherein σ was given by (I.1) with s having the two-dimensional value for VRH, 1/3; (ii) this behavior would have held over a wide enough temperature range to match the experimental behavior; (iii) the slope of the region with lnσ linear in t⁻¹/³ would have been small enough to match the small experimental T₀ values. Had such behavior been found, it would have strongly suggested that the same behavior would occur in three dimensions. However, such behavior was not found. Instead, it was found that

\[ \sigma = \sigma_0 e^{-\varepsilon_3 / kT} \]  

(I.2)

This finding did not preclude that the above requirements would hold in three dimensions. This issue was eventually settled in [III] and [IV], as will be discussed below.

One of the highlights of the work in [I] was obtaining a trustworthy value of ε₃. For two dimensions, it was found that ε₃ = 0.28Δε. This was a new result which differed markedly from values reported by earlier investigators: Hayden and Butcher⁷ found ε₃ = 0.38Δε while Skal, Shklovskii and Efros⁸ reported ε₃ = \( \frac{5}{12} \Delta \varepsilon \). The result obtained in this dissertation is an improvement on these previous values because this dissertation used the full resistor network. The other investigators employed the low t
asymptotic form of the resistance in the high t regime. That previous
workers found values for $\varepsilon_3$ which differ significantly from the value
obtained in this dissertation established the necessity of using the full
form of the resistance at high t.

In [II], a two-dimensional analytic theory based on the percolation
method was presented. The theory involved the critical number $B_c(t)$ of
bonds (resistors) per site required for percolation. Agreement between the
analytic theory and the numerical simulation was obtained by using a
reasonable form for $B_c(t)$. This form of $B_c(t)$ was obtained in two steps.
First, a conjecture was made for the form of $B_c(t)$. The conjectured form
gave agreement to within a few percent between the analytic and the
numerical conductivities; this form of $B_c(t)$ was then fine tuned by
requiring improved agreement between the analytic and numerical results.
The analytic theory verified the finding (1.2) with $\varepsilon_3 = 0.28\Delta\varepsilon$, and also
gave information about the typical hopping distance and typical hopping
energy. It was found that the hopping was essentially nearest neighbor
hopping for the temperature range where (1.2) held, and that the hopping
distance increased with decreasing temperature below this range. It was
also demonstrated that a typical hopping energy existed only for the
temperature range where the hopping distance increased with decreasing
temperature. When the hops were between nearest neighbors, the energies of
the sites played no role in selecting the pairs of sites which determined
the overall conductivity. As such, although there was an average hopping
energy in the nearest neighbor regime, there was no typical hopping energy.
In [III], the analytic theory was extended to three dimensions. It was established that the experimental conductivity of ref.s [1] and [2] could not be accounted for by using a flat density of states. One of the highlights of [III] was that (1.2), with $\varepsilon_3 = 0.20 \Delta \varepsilon$, was obtained for the conductivity. The result revealed that going from two dimensions to three reduced $\frac{\varepsilon_3}{\Delta \varepsilon}$. This result was different than the value $\varepsilon_3 = \frac{5}{12} \Delta \varepsilon$ due to Skal, Shklovskii and Efros, who claimed that $\varepsilon_3$ would be the same in three dimensions as in two. That the value of $\varepsilon_3$ found in this dissertation differed from the result of these previous workers once again verified the need to use the full form of the resistance at high $t$. Another of the highlights of [III] was to show that the effect of including the $Q_{ij}$ factor in the analytic theory was to produce a trend of decreasing $\frac{\varepsilon_3}{\Delta \varepsilon}$ with increasing $\Delta \varepsilon$. Inclusion of the $Q_{ij}$ factor did not change the value of $s$ for which $\ln \sigma$ was proportional to $T^{-s}$. The conductivity was still of the form (1.2).

In [IV], an inquiry was made as to how fundamentally different forms of the densities of states would influence the temperature dependence of $\sigma$. For simplicity, the densities of states were chosen to be symmetric about the zero temperature chemical potential $\mu_0$, and were taken to be of the form

$$g(\varepsilon) = \begin{cases} 
\alpha |\varepsilon|^{\lambda}, & \text{if } 0 < |\varepsilon| < \frac{\Delta \varepsilon}{2}, \lambda > -1, \\
0, & \text{if } |\varepsilon| > \frac{\Delta \varepsilon}{2}, 
\end{cases}$$  

(1.3)
with \( a \) fixed by the requirement \( \int_{-\Delta \varepsilon}^{\Delta \varepsilon} g(\varepsilon) \, d\varepsilon = n_d \), and \( \mu_0 \) set to zero. The cases \( \lambda = + \frac{1}{2}, 0, - \frac{1}{2} \) were studied, because they described fundamentally different densities of states. It was found to be impossible for the cases \( \lambda = + \frac{1}{2}, 0 \) to account for the high temperature experimental conductivity. However, examination of the case \( \lambda = - \frac{1}{2} \) led to a very important discovery.

It is highly plausible that the resistor network model, used in conjunction with a density of states which decreases with \(|\varepsilon|\) for \( \varepsilon \) far enough away from \( \mu_0 \), will account for the experimental conductivity. This claim was realized as follows. There are two \( \mathcal{B}^c(t) \) functions to consider with regard to the density of states for \( \lambda = - \frac{1}{2} \). One is the \( \mathcal{B}^c(t) \) that is fixed by the density of states, which is referred to as "the true \( \mathcal{B}^c(t) \)". As discussed in [IV], it was not possible to determine this \( \mathcal{B}^c(t) \). The other \( \mathcal{B}^c(t) \), which will be called "the \( \mathcal{B}^c(t) \) curve", is the one required in order that the conductivity of the case \( \lambda = - \frac{1}{2} \) agree with the experimental conductivity. The \( \mathcal{B}^c(t) \) curve was determined in [IV]. Although it was not possible to determine in [IV] how closely the \( \mathcal{B}^c(t) \) curve matched the true \( \mathcal{B}^c(t) \), it was demonstrated that the \( \mathcal{B}^c(t) \) curve was compatible with both the \( t \to \infty \) limit and the \( t \to 0 \) asymptotic form of the true \( \mathcal{B}^c(t) \). The work in this dissertation has not culminated in an unequivocal explanation of the experimental results in ref.s [1] and [2]. Nevertheless, a principal highlight of this work is the establishment of a very plausible theoretical accounting for these results. This point will be discussed further in
Sec. II below. On the basis of the work in [IV], it was possible to present a suggestion as to why some investigators\textsuperscript{1,2} found $\sigma$ to be of the form (I.1) with $s = \frac{1}{4}$ while others\textsuperscript{9-12} found $s = 1$ (in all cases, lightly doped n-GaAs or n-InP was studied, and $1K \lesssim T \lesssim 10K$): different densities of states will have different concomitant $B_c(t)$ which in turn will lead to high $t$ conductivities with different values of $s$ in (I.1).

A key prediction associated with the ideas set forth in the last two paragraphs is as follows. Experimental conductivities of the form (I.1) with $s = \frac{1}{4}$ and $T_0$ small ($T_0 = 10^3K$ in ref. [1]) at high temperatures ($1K \lesssim T \lesssim 7K$ in ref. [1]), will go over to the form (I.1) with $s = \frac{1}{4}$ and $T_0$ large ($= 10^5K$) at low temperatures ($T \sim 1K$). It would be of great value for the authors of ref.s [1] and [2] to extend their measurements to lower temperatures and thereby test this prediction.

Beyond the achievements described above, one of the most important features of this dissertation is that, for the first time, an analytic theory using the full resistor network has been developed. As such, the way has been paved for further theoretical work on high temperature hopping conduction.
II. Suggestions for Further Work

With each successive investigation undertaken in this dissertation, it became less certain that the $B_c(t)$ curve being used was close to the true $B_c(t)$. This was not at all a concern in [II]; in [III], the uncertainty in the $B_c(t)$ curve was still not cause for serious concern. However, in [IV], where the most interesting investigation was carried out, the uncertainty in the form of $B_c(t)$ was enough to merit the implementation of a new approach. Rather than calculating the conductivity resulting from a given $B_c(t)$, an inquiry was made as to whether or not the $B_c(t)$ required to give agreement with the experimental conductivity was a reasonable $B_c(t)$. As such, although this dissertation has presented a highly plausible theoretical explanation for the experimental results presented in ref.s [1] and [2], it cannot be claimed that the work herein constitutes an unequivocal understanding of these experimental results.

In order to remove any doubt in the plausibility argument that was presented in [IV], a further investigation along the same lines is required. In this further investigation, various densities of states will be examined. The $B_c(t)$ concomitant to a given density of states will be determined by computer simulation: a resistor network will be constructed as described in [I], and $B_c(t)$ will be given in terms of the temperature dependence of the number of resistors present at percolation. Using this knowledge of $B_c(t)$ in conjunction with the analytic theory developed in this dissertation, the dependence of the conductivity on the form of the density of states can be studied. The work in this dissertation indicates that it is very likely
that there are densities of states for which $\sigma$ will be of the form (I.1) with either $s = 1$ or $s = \frac{1}{4}$. It is hoped that the investigation will reveal what values $s$ can have in equation (I.1); it may be the case that $s$ is rather unrestricted, or, alternatively, that most densities of states give either $s = 1$ or $s = \frac{1}{4}$. As such, perhaps the investigation will shed light on why experimental studies have furnished either the result $s = \frac{1}{4}$ (ref.s [1] and [2]) or $s = 1$ (ref.s [9]-[12]). If, in the vicinity of the zero temperature chemical potential, the densities of states examined are flat, then the investigation will also provide information on the transition to the low temperature form of the conductivity.

A second project of great interest has to do with simplifying the analytic theory. The analytic theory was developed to allow the best possible determination of the conductivity within the context of the resistor network model. The results obtained were improvements over previous results, as discussed in Sec. I. Unfortunately, the theory in its present form does not provide a clear view of the underlying physical processes. It would be of great value to extend the development of the theory to the point where these underlying processes were revealed. Interesting questions such as the following could then be addressed. In the high temperature regime, where $\sigma$ is given by (I.1) with $s = \frac{1}{4}$, what determines the size of $T_0$? Why is the high temperature $T_0$ smaller than the low temperature $T_0$? What does $T_0$ depend on? If $\sigma$ is of the form (I.1) with $s = \frac{1}{4}$, will the hopping necessarily be variable range?
Another possible investigation is to examine carefully the shape dependence of the numerically determined overall resistance, as was discussed in [I].

A fuller analysis of the role of the $Q_{ij}$ factor could also be carried out, possibly along the lines suggested in [III].

Nowhere in this dissertation was the prefactor $\sigma_0$ worked out. Moreover, no investigations for $n_D \neq 2n_A$ were made. No studies were performed to explore the dependence of $\sigma$ on $n_D$. Further work could involve considerations of these kinds.

The most important result of this dissertation was to establish a plausible explanation for the experimental results of ref.s [1] and [2]. The most important piece of work to be based on this dissertation will be the full investigation of the effect of the density of states on the form of the conductivity. Interesting questions remain to be solved. This dissertation has established a basis from which to proceed to the answers.
REFERENCES

BIBLIOGRAPHY


A. Miller and E. Abrahams, Phys. Rev. 120, 745 (1960).
N.F. Mott, Metal-Insulator Transitions (Taylor and Francis, London, 1974).


APPENDIX

In this appendix, a description is given of how to derive the resistor network used in this dissertation.

The derivation is based on a time-averaged rate equation approach for describing the hopping process. The time-averaged occupation number $n_i$ for donor site $i$ is given by

\[
\frac{\partial n_i}{\partial t} = \sum_j \left( \Gamma_{ij} - \Gamma_{ji} \right),
\]  

(i)

where $\Gamma_{ij}$ is the time-averaged rate at which electrons hop from site $i$ to site $j$, $t$ is time, and the sum is over all donor sites $j$ different from $i$. The hopping rate $\Gamma_{ij}$, given by

\[
\Gamma_{ij} = n_i \gamma_{ij} (1-n_j),
\]  

(ii)

is a product of three probabilities: the probability $n_i$ that site $i$ is occupied, the probability $(1-n_j)$ that site $j$ is empty, and the probability per unit time $\gamma_{ij}$ that an electron on site $i$ will hop to an empty site $j$. The net electric current $I_{ij}$ flowing from site $i$ to site $j$ is given by

\[
I_{ij} = -e(\Gamma_{ij} - \Gamma_{ji}),
\]  

(iii)
where $e$ is the modulus of the electronic charge.

When there is no external electric field (i.e., in equilibrium), there is a detailed balance: $\Gamma_{ij} = \Gamma_{ji}$, $I_{ij} = 0$, and there is no bulk current flow. Application of an electric field $\mathbf{E}$ alters the hopping rates, making it more likely for electrons to hop in the direction $-\mathbf{E}$. The $I_{ij}$ become non-zero and a bulk current results. The network of currents may be replaced by a network of resistors in the manner described below.

When there is an electric field, the occupation numbers change. This is because equilibrium has given way to a steady state situation in which there are electron flows toward, and away from, each donor site. Taking $E$ to be small, $n_i$ may be written in terms of the small unknown $\Delta_i$ as

$$n_i = \left[ \frac{1}{2} e^{(\varepsilon_i + \Delta_i - \mu)/kT} + 1 \right]^{-1}, \quad (iv)$$

where $\varepsilon_i$ and $\mu$ are the localization energy for site $i$ and the chemical potential when $E = 0$, and $T$ is the temperature. (When $E = 0$, $\Delta_i = 0$ and $n_i$ has its equilibrium form.)

Suppose a phonon is absorbed in hopping from site $i$ to $j$. Then

$$\frac{\gamma_{ij}}{\gamma_{ji}} = \frac{n_q}{n_q + 1}, \quad (v)$$

where $n_q$ is the number of phonons of mode $q$. The energy of the mode is
given by

\[ \varpi q = [\varepsilon_j - e\hat{E} \cdot \hat{r}_j] - [\varepsilon_i - e\hat{E} \cdot \hat{r}_i], \quad (vi) \]

where \( \hat{r}_i \) locates donor site \( i \). Assuming the phonons to be in thermal equilibrium, \( n_q \) is given by

\[ n_q = \left( \frac{\varpi q}{kT} \right)^{-1} - 1. \quad (vii) \]

Steady state means the left hand side of equation (i) is zero. As such, the \( \Delta_i \) can in principle be solved for by combining equations (i), (ii), and (iv)-(vii). This formidable task can be circumvented by deriving an expression for \( I_{ij} \) to first order in \( E \). Combining equations (ii)-(vii) gives

\[ I_{ij} = \frac{V_{ij}}{R_{ij}^0}, \quad (viii) \]

where

\[ V_{ij} = e\hat{E} \cdot \hat{r}_{ij} - (\Delta_i - \Delta_j)/e, \quad (ix) \]
and

\[ R_{ij}^0 = \frac{kT}{e^{2E_{ij}^0}} = \frac{kT}{e^{2E_{ij}^0}(1-n_j^0)} , \]  

where the superscripts 0 denote equilibrium values. In steady state, equation (i) is

\[ \sum_{j} I_{ij} = 0. \]  

Equations (viii)-(xi) describe a resistor network with resistances \( R_{ij}^0 \) and potential drops \( V_{ij} \). Kirchhoff's laws are given by equations (ix) and (xi). Thus, the current \( I_{ij} \) that flows in resistor \( R_{ij}^0 \) is the same current that flows in the lightly doped semiconductor (LDS) between donor sites i and j, and the resistor network may be used to calculate the conductivity of the semiconductor. Since \( R_{ij}^0 \) involves only known quantities, the conductivity can be calculated without solving for the \( \Delta_i \).

Note that by writing \( I_{ij} = \tilde{V}_{ij}/\tilde{R}_{ij} \) where \( \tilde{V}_{ij} = \alpha_{ij} V_{ij} \), \( \tilde{R}_{ij} = \alpha_{ij} R_{ij}^0 \), \( \alpha_{ij} = (\tilde{V}_i - \tilde{V}_j)/V_{ij} \), where \( \tilde{V}_i \) and \( \tilde{V}_j \) are arbitrary but unequal, it follows that the resistor network \( \tilde{R}_{ij} \) is an alternative network to model the LDS with. As such, there are infinitely many networks to choose from. The choice \( R_{ij}^0 \) is made because it is the choice involving none of the unknowns \( \Delta_i \). That such freedom of choice is possible implies that no special significance should be attached to the \( \Delta_i \). Some authors have ascribed to
the physical significance of "local chemical potential at site \( i \); the above shows that it is unnecessary to apply a macroscopic concept (chemical potential) at the microscopic level.

Obtaining an expression for \( R_{ij}^0 \) reduces to calculating \( \gamma_{ij}^0 \). \( \gamma_{ij}^0 \) was originally evaluated by Miller and Abrahams \(^2\), and later by Shklovskii and Efros \(^3\). Assuming an isotropic, acoustic phonon branch, no valley degeneracy and an isotropic effective mass for the bottom of the conduction band, and using the deformation potential approximation for the electron-phonon interaction, \( \gamma_{ij}^0 \) is given by

\[
\gamma_{ij}^0 = \frac{E_d^2(\varepsilon_i - \varepsilon_j)}{\pi ds^5} \left[ \frac{2k_0 e^2}{3\pi a} \right] \frac{r_{ij}}{a} \left[ 1 + \frac{(\varepsilon_i - \varepsilon_j)a}{2\bar{m}_s} \right] e^{-2r_{ij}/a}, \tag{xii}
\]

where the symbols are all as defined previously in this dissertation (see, for example, \([1]\)), and where \( n_q \) is given by (vii) with \( t_{q} = \varepsilon_j - \varepsilon_i \). Using equation (xii) in equation (x) gives equations (4)-(7) of \([1]\) for the resistor network.

Apart from the works of this dissertation, the expressions (4)-(7) of \([1]\) for the resistor network do not appear explicitly in the literature. This is partly because most authors use the \( T \to 0 \) asymptotic form for the resistance. Shklovskii and Efros \(^3\) presented equations (x) and (xii) explicitly (their equations (4.2.30), and (4.2.17) combined with (4.7.18), respectively), but when combining them wrote down only the \( T \to 0 \) asymptotic form. Miller and Abrahams, while not taking the \( T \to 0 \) asymptotic form, did
respectively), but when combining them wrote down only the $T \to 0$ asymptotic form. Miller and Abrahams, while not taking the $T \to 0$ asymptotic form, did not explicitly record equation (xii). Instead, they replaced the term $1 + \left(\frac{\epsilon_j - \epsilon_i}{\epsilon_j - \epsilon_i}a^2\right)$ by unity. They also particularized to Ge and Si, materials having valley degeneracies. This resulted in a slightly different prefactor, and $r_{ij}/a$ appearing to the power $\frac{3}{2}$ instead of 2 — see their equations (III-22) - (III-26).

The expression (xii) is suitable for GaAs and InP because these materials have isotropic effective masses.

---

1For a derivation of the equilibrium form of $n_i$, see, for example, N.W. Ashcroft and N.D. Mermin, Solid State Physics (Holt, Rinehart and Winston, Philadelphia, 1976), pp. 581-582.

2A. Miller and E. Abrahams, Phys. Rev. 120, 745 (1960).